We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,100
Open access books available

116,000
International authors and editors

125M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Noise, Averaging, and Dithering in Data Acquisition Systems

Filippo Attivissimo and Nicola Giaquinto

Dipartimento di Elettrotecnica ed Elettronica
Politecnico di Bari,
Italia

1. Introduction

In any data acquisition system (DAS) many error effects, both of systematic nature (e.g. nonlinearity) and of random nature (e.g. electronic noise) are simultaneously present. While systematic errors are a comparatively stable characteristic of a DAS, random errors may be smaller or larger in different situations, and it is important to understand how they degrade the overall performance of the system. It is even more important to understand that random errors can be actually used to improve the fidelity of the acquisition, i.e. the technique of dithering. This possibility is due to the inherent presence in any DAS of a particular kind of error: the quantization error.

Quantization is a basically simple operation and it is easily understood at an elementary level. However, evaluating its effects on signals, with or without the simultaneous presence of other errors, requires quite complex mathematics, usually not mastered by engineers and even by researchers without a specific interest in the topic. Due to the complexity of the subject (an excellent reference book is [WK08]), misunderstandings and mistakes are common when dealing with noise in DAS. For example, it is true that averaging a particular number of samples is convenient to reduce the noise, but it is easy to disregard the fact that it is useless to increase the number of samples beyond a certain limit (contrary to what happens in analogue measurements). In the same way, even if introducing noise in a DAS may be desirable and effective, and is expressly a feature in commercial DAS (e.g. [Nat97], [Nat07]), few users are aware of how the appropriate level of noise (and other parameters) can be chosen.

The present chapter deals with the topic of performance degradation/improvement in a DAS, deriving by the presence (wanted or unwanted) of noise, and by averaging or filtering the output samples. The aim is making the theory understandable and usable by a wide audience, using ideas and mathematics as simple as possible. Proper reference, when needed, is made to works with rigorous mathematical demonstration of the derived results. The chapter covers only the case of perfectly linear DAS, with no (or negligible) nonlinearity errors. The more general case of nonlinear DAS with noise is a subject for a possible future expanded version of the chapter.

1 corresponding author: http://dee.poliba.it/DEE/Giaquinto.html
2. Effective number of bits

If $x(t)$ is the analogue input of a DAS and $y_n$ are the output samples, the evaluation of the overall acquisition fidelity takes into account, customarily, only transformations involving the shape of $x(t)$. Therefore, the fidelity evaluation excludes:

- linear transformations in the amplitude of the signal (due to fixed gain and offset errors);
- linear transformations in the time of the signal (due to a fixed trigger delay and a fixed error in sampling frequency).

Formally, this means that one has to identify four constants $a, b, c, d$ so that, if $t_n$ are the nominal (ideal) sampling instants, the scaled input samples

$$x_n^s = a + b \cdot x(c + d \cdot t_n)$$

have minimum distance, in the least squares (LS) sense, from the output samples $y_n$

$$\sum_n (y_n - x_n^s)^2 = \min.$$ 

In practical DAS testing, $x(t)$ is often a large sinusoidal signal, i.e. a sinusoid stimulating at least 90% of the full-scale range (FSR) of the acquisition channel (as specified in [IEE94], Clause 3.1.29). Identifying the four constants $a, b, c, d$ means to determine with the LS method the four parameters $C, V, \omega_x, \phi_x$ in the expression

$$x_n^s = C + V \cos(\omega_x t_n + \phi_x).$$

After determining $x_n^s$, a logical fidelity measure is the mean squared error (MSE)

$$\sigma^2 = \frac{1}{2} (y_n - x_n^s)^2 = \sigma_q^2.$$ 

The MSE, however, is an absolute number lacking an immediately clear meaning. It is preferred, therefore, to express the value of MSE in terms of effective number of bits (ENOB), defined by the formula

$$b_e = b - \frac{1}{2} \log_2 \frac{\sigma^2}{\sigma_q^2}$$

where

$$\sigma_q^2 = Q^2 / 12$$

is the MSE of an ideal sampler/quantizer with the same resolution of the DAS. It is obvious that in an actual DAS, which has additional errors besides quantization, it is always true that $\sigma^2 > \sigma_q^2$ and therefore $b_e < b$.

The meaning of the ENOB definition (4) is better understood by considering a conventional input signal with uniform distribution in the whole FSR of the converter, e.g. a triangular signal (or a ramp, a sawtooth, etc.). The FSR has amplitude

$$x_{FSR} = 2^{b_e} \cdot Q.$$
and therefore the full-scale triangular signal has power (without considering a possible dc component)

$$\sigma_T^2 = \frac{x_f^2}{12}$$  \hspace{1cm} (7)

For an ideal quantizer, the resolution $b$ may be expressed in terms of the logarithm of the ratio between the power of the full-scale triangular signal $\sigma_T^2$ and the power of the ideal quantization error $\sigma_q^2$, i.e.

$$\frac{1}{2} \log_2 \frac{\sigma_T^2}{\sigma_q^2} = \frac{1}{2} \log_2 \frac{x_f^2}{Q^2} = \frac{1}{2} \log_2 2^{2b} = b$$  \hspace{1cm} (8)

For an actual DAS, the same ratio, with the actual MSE $\sigma_e^2$ instead of the ideal one $\sigma_q^2$, yields the ENOB:

$$\frac{1}{2} \log_2 \frac{\sigma_e^2}{\sigma_q^2} = \frac{1}{2} \log_2 \frac{x_f^2}{Q^2} = \frac{1}{2} \log_2 2^{2b_e} = b_e$$  \hspace{1cm} (9)

Therefore, expression (4) of the ENOB gives the resolution of an ideal quantizer with the same MSE of the actual DAS (although the result is in general a non-integer number of bits).

It is worth to highlight that, if $10 \log_{10}$ is substituted to $(1/2)\log_2$ in (8), the ideal dynamic range $DR = 6.02 \cdot b$ of the DAS is obtained, and in the same way the quantity $6.02 \cdot b_e$ may be considered a measure of actual dynamic range (although this is not a standardized definition).

The given definition of ENOB, like the MSE $\sigma_e^2$, depends on the actual signal $x(t)$ used to stimulate the input of the DAS. The normal practical choice, which has become a standard, is a sinusoidal signal smaller than the FSR, but larger than 90% of the FSR itself (“large sinusoid”). The main reason for choosing the sinusoid is that the difference between the actually generated signal and its ideal mathematical expression must be a negligible quantity with respect to the error introduced by the DAS itself. This is technologically much more feasible for the sinusoid then for any other waveform. The large sinusoid, on the other hand, has its drawbacks, in practice and in theory.

1. Under a practical point of view, the large sinusoid does not cover exactly the whole range of the DAS, nor it stimulates uniformly the covered range. Therefore, nonlinearity errors near the border of the scale weigh less than errors near the centre, and the errors outside the range of the signal are not accounted for at all [GT97].

2. Under a theoretical point of view, in an ideal quantizer the sinusoid does not produce a MSE exactly equal to $\sigma_q^2 = Q^2 / 12$ [WK08]. Besides, there is a logical inconsistency in evaluating the MSE produced by a sinusoidal signal, and comparing it with the power of a uniformly distributed signal, as the ENOB definition (8) requires.

Because of the aforementioned problems, a perfectly linear ramp or a triangular signal are also used when possible. When the sinusoidal signal is the only feasible choice, a good suggestion (first given and developed in [GT97], and confirmed in [KB05]) is to stimulate the DAS with some overdrive, since in this way the signal laying in the FSR is almost uniformly...
distributed. As a matter of fact, in this way the ENOB evaluation is practically insensitive to small variations in the amplitude and offset of the stimulus sinusoid (contrary to what happens without overdrive), and the evaluation is much more consistent with the results obtained by different tests (e.g. the histogram test of nonlinearity, which uses a sinusoid with overdrive [IEE00]). The issue of practical ENOB testing, however, is not further addressed here.

In this chapter, mainly to avoid theoretical inconsistencies (point 2 above), the stimulus signal \( x(t) \) is always assumed to be uniformly distributed in the FSR of the DAS. Since dynamic effects (like e.g. dynamic nonlinearity, sampling jitter, etc.) are not examined in the chapter and not included in the mathematical analysis and in the simulations, the frequency of the input is inessential. If one wants to obtain practical measurements in good accordance with the theory developed in the chapter, a sinusoidal signal with some overdrive must be used. Using a large sinusoidal signal leads to similar results, but with meaningful differences.

Another convention followed in this chapter is that the quantization step is assumed to be \( Q = 1 \). This is equivalent to express in LSB units all the quantities with the same physical dimension of \( Q \) (voltages), and simplifies many equations and notations. For example, since \( \sigma_q = Q / \sqrt{12} = 1 / \sqrt{12} \), ENOB may be expressed by

\[
b_e = b - \frac{1}{2} \log_2 12\sigma_e^2 \tag{10}
\]

provided that \( Q=1 \), or, equivalently, that \( \sigma_e \) is expressed in LSB units. Under this condition, all the equations in the chapter can be used without modifications.

3. Perfectly linear DAS with noise and no averaging

The case of perfectly linear DAS with noise and no averaging is elementary but is also preliminary to the analysis of more complete and complex cases.

In an actual DAS there are many sources of noise, but the overall effect can be seen (and is quantified by manufacturers) as a single noise generator with power \( \sigma_n^2 \) at the input of the system. If the DAS has negligible nonlinearity, it can be represented by the very simple equivalent model in Fig. 1.

Fig. 1. Equivalent model of linear DAS with noise.

The ideal quantizer adds, of course, a quantization error \( e_q(x') \), which is a function of the input \( x' = x + n \). For a fixed input signal, and in particular for a full-scale triangular signal,
the quantization error has a fixed power. Consequently, the model in Fig. 1 can be
substituted by the fully additive model of Fig. 2 (a typical operation in quantization theory).
Under broad conditions on the quantized signal \( x \), quantization theory assures that
quantization error is: (i) uniformly distributed in \([-Q/2, Q/2]\) and therefore zero-mean
with power equal to \( \sigma_q^2 = Q^2 / 12 \); (ii) white; (iii) uncorrelated with the input. It can be
proven (the more general proof is probably the one given in [SO05]) that \( n \) and \( q\)
are uncorrelated, too, and therefore the overall MSE of the DAS is:

\[
\sigma_n^2 + \sigma_q^2 = \sigma_n^2 + \sigma_q^2. \tag{10}
\]

Taking into account the normalization convention \( Q = 1 \), the term in (10) becomes
\( 12\sigma_n^2 = 1 + 12\sigma_n^2 \), and therefore in this elementary case the ENOB of the DAS is:

\[
b_e = b - \frac{1}{2} \log_2 \left( 1 + 12\sigma_n^2 \right) \tag{11}\]

A simple numerical simulation (performed for \( b \) in the range 8÷16 bits) confirms the
formula (Fig. 3). It is interesting to note the formal similarity of the law of the performance
degradation \( \Delta b = -(1/2)\log_2 (1 + 12\sigma_n^2) \) with that of a first-order low-pass filter, with a cut-off
frequency equal to the pure root mean square (rms) quantization error,
\( \sigma_q = Q / \sqrt{12} \approx 0.289 \) LSB. At the cut-off (\( \sigma_n = \sigma_q \)) the ENOB is half a bit below the nominal
resolution \( b \). After the cut-off, the ENOB decreases with a rate of 1 bit/ octave, or
3.32 bit/ decade, equivalent to a decrease in the dynamic range of 6.02 dB/ octave or
20 dB/ decade.

4. Perfectly linear DAS with noise and averaging: an important case of non-
subtractive dithering

4.1 Oversampling and averaging

When the performance of a measurement system is degraded by noise, the obvious method
to increase accuracy is some form of averaging.
The simple non-weighted averaging is the well-known optimal method to estimate an
unknown constant signal buried in white Gaussian noise (WGN). When the signal is not
constant, averaging is advantageously substituted by other filtering techniques, ranging
from simple low-pass or band-pass filtering to adaptive filtering, etc. The basic principle is,
however, the same: to obtain each output sample by a (weighted) average of many samples
of the input, in order to reduce the acquisition error. This is the principle of oversampling, i.e.
trading bandwidth (and possibly sampling frequency) for accuracy, e.g. in terms of ENOB.
As a side note, it must be highlighted that oversampling is implemented by design in a wide class of analog-to-digital converters (sigma-delta converters, etc.), used in commercial DAS [Nat05]. This chapter does not deal with this “hard” oversampling which involves built-in hardware to improve performance, e.g. in the form of embedded feedback loops. The chapter deals, instead, with the “soft” oversampling implemented by the user in the form of output processing when there is unwanted noise, and an abundance of acquired samples with respect to the signal bandwidth. Soft oversampling does not include the implementation of feedback loops, or similar techniques.

Fig. 3. ENOB of perfectly linear DAS (with resolution in the range 8÷16 bits) affected by input noise (with rms value in the range 0÷10 LSB). Numerical simulations are compared with theoretical equations. The “cut-off” at $\sigma_n = Q / \sqrt{12} = 0.289$ LSB is highlighted.

In the rest of the chapter, attention will be focused on the case of simple (non-weighted) averaging of many output samples of a DAS, when the input is an unknown constant with additive WGN.

Like for the hypothesis of uniformly distributed signal for ENOB evaluation, the choice of the case is primarily justified by theoretical convenience. In this way the problem is mathematically treatable and accurate closed-form equations are derivable. Besides, the analysis and the results provide a good understanding, useful for more general cases, of the interaction between the signal-dependent errors introduced by quantization, and the signal-independent errors introduced by noise.
In practice, the case of WGN is by far the most common, and it is easy to repeat the analysis for other kinds of noise (non-Gaussian and/or non-white). Also the hypothesis of constant signal is verified in many practical cases, e.g. when the sampling frequency is very high with respect to the variations of the input, when a sample-and-hold is used to acquire many samples with “frozen” signal, or when there is a repetitive sampling of many periods of a periodic signal. It is also not too difficult to extend the analysis to specific cases of linear filtering applied to a non-constant signal.

4.2 Dithering

Besides being present as an unwanted disturbing signal, WGN can be purposely added to the input of a DAS in order to improve the final accuracy. This is a particular case of the well-known technique of dithering, which is a main error-correction method among those available for DAS [BDR05]. The basic idea is that, since there is no way to remove or reduce the error introduced by quantization when the input is perfectly constant, random variations in the input are beneficial for error-correction. Indeed, the addition of a random signal to the input randomizes the quantization error which, in turn, can be removed (or reduced) by averaging.

Subtractive dithering in DAS consists in adding a dither signal to the input, and subtracting it from the output before possible further processing [Sch64]. Subtractive dithering inherently requires accurate knowledge of the signal added to the input (or specific hardware to measure it) and is therefore more difficult and expensive to implement.

Non-subtractive dithering, instead, implies averaging/filtering the output without previous subtraction of the dither added at the input. This technique is much easier to implement with respect to subtractive dithering, and has been studied quite deeply in a number of theoretical works (see, e.g. [WLW00], and the bibliography in [WK08]). Even easier is to use simple WGN as a dither signal, since this noise is (almost always) already present at the input of DAS, and may be easily incremented if necessary. This very common kind of dithering may be called “white Gaussian non-subtractive dither” (WGND). Averaging the output of a linear DAS with WGN is therefore also a particular but very common and important case of non-subtractive dithering, the WGND (Fig. 4).

![Diagram](https://via.placeholder.com/150)

Fig. 4. Basic scheme of operation of WGND applied to a perfectly linear quantizer.

For the sake of completeness, it must be clarified that dithering consists in general in purposely altering the signal at the input of a system (the technique is not limited to data acquisition), in order to improve the performance of the system itself. In the field of data acquisition, besides adding an external signal, other kinds of dithering are possible and used, aiming at different performance improvements. For example, an effective anti-aliasing filter can be obtained, without increasing the sampling rate and without introducing
physical filters, by a proper randomization of the sampling instants. Amplitude and time dithering may be combined efficiently [AH98]. Of course random errors in sampling instants can be also undesired effects, and in this case they are studied with specific mathematical models [AD09]. These techniques, dealing with errors in sampling instants and other kinds of alterations of the input signal, are not within the scope of this chapter. The scheme reported in Fig. 4 has been deeply examined in the context of quantization theory using the typical, quite complex mathematical tools of the theory. The analysis reported here is probably the simpler and most direct way to understand the actual benefits given by WGND, and in general by averaging/filtering in presence of noise at the input of the DAS. The analysis is centred on the determination of the attainable ENOB in given conditions.

5. Averaging infinite output samples

The analysis starts considering the average of infinite output samples in the scheme of Fig. 4. Averaging infinite samples transforms the system, which includes random contributions, in a purely deterministic one. The input-output relationship of the system is the convolution of the ideal quantization function quant(x) with the probability density function (pdf) of the dither, i.e. with the zero-mean Gaussian density

$$\phi(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(12)

with $\mu = 0$ and $\sigma = \sigma_n$. The result of the convolution is a nonlinear function which is actually a smoothed quantization, or a dithered quantization $y = \text{quantd}(x)$. The system in Fig. 4 is transformed in that represented in Fig. 5.

![Fig. 5. Representation of a nonlinear system equivalent to a perfectly linear DAS with WGN at the input and averaging of infinite samples at the output.](www.intechopen.com)

The error introduced by the dithered quantization, $e_{\text{qd}}(x) = \text{quantd}(x) - x$, may be directly obtained by the convolution

$$e_{\text{qd}}(x) = e_q(x) * \phi(x) = \frac{1}{\sigma_n\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{2\sigma_n^2}} \cdot e_q(x-\xi) d\xi$$

(13)

where $e_q(x)$ is the ideal quantization error $\text{quant}(x) - x$ (Fig. 6).
For a fixed input signal, and in particular for a full-scale triangular signal, the system in
Fig. 5 may be represented also as an additive error (the dithered quantization error) with a
fixed power $\sigma^2_{qd}$ (Fig. 7). This additive model is perfectly analogous to that used for the
ideal quantization error.

The rms error $\sigma_{qd}$ introduced by dithered quantization may be evaluated by means of a
numerical integration of the square of the smooth curve in Fig. 6, weighted with the
distribution of the input signal. For the case of triangular uniformly distributed signal, there
is no weighting:

$$\sigma^2_{qd} = \frac{1}{Q} \int_{-Q/2}^{Q/2} e^2_{qd}(x) dx .$$

(14)

Fig. 6. Ideal quantization error $e_q(x)$ and dithered quantization error $e_{qd}(x)$ (for the case
$\sigma_n = 0.1$ LSB).

Fig. 7. Additive model of the dithered quantization of Fig. 5.
(If $e_q(x)$ is substituted to $e_{qd}(x)$, the result is trivially $\sigma_q^2 = Q^2 / 12$.) Of course the result of integration (14) with integrand given by (13) depends only on the standard deviation $\sigma_n$ of the input Gaussian noise:

$$\sigma_{qd} = g(\sigma_n).$$

(15)

This function can be easily evaluated numerically. The result is reported in Fig. 8, and the values in a few points are reported in Tab. 1.

The result shows that $\sigma_{qd}$ becomes practically negligible at $\sigma_n \approx 0.5$ LSB: more precisely, at $\sigma_n = 0.5$ LSB the dithered quantization error $\sigma_{qd}$ is about $1.6 \cdot 10^{-3}$ LSB (Fig. 9). This means that $\sigma_n = 0.5$ LSB achieves an almost complete randomization of the quantization error (i.e., $e_{qd}(x) \equiv 0$ for every $x$). A perfectly complete randomization, however, is theoretically achieved only for an infinite $\sigma_n$. The randomized quantization error is removed by averaging a sufficiently high (theoretically, infinite) number of samples.

![Fig. 8. Rms dithered quantization error, $\sigma_{qd}$, as a function of the rms input Gaussian noise, $\sigma_n$.](www.intechopen.com)

For $\sigma_n = 0$, the dithered quantization error becomes pure quantization error with standard deviation $\sigma_q = Q / \sqrt{12}$. A zoom of the curve in the rectangle is represented in Fig. 9.
Noise, Averaging, and Dithering in Data Acquisition Systems

<table>
<thead>
<tr>
<th>$\sigma_n$</th>
<th>$\sigma_{qd}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2887</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1921</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1023</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0381</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0096</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

Tab. 1. Some points of the function $\sigma_{qd} = g(\sigma_n)$ (both in LSB units).

Fig. 9. Zoom of the curve in Fig. 8 in the neighbourhood of $\sigma_n = 0.5$ LSB.

It is worth to recall that according to well-known results of quantization theory [Sch64], [WK08], a perfectly complete randomization of the quantization error is possible with a proper pdf of the input noise, and there are infinite pdfs that lead to such a perfect result. For example, a uniformly distributed noise in $[-Q/2, Q/2]$ yields exactly $e_{qd}(x) = 0$ and therefore $\sigma_{qd} = 0$. However, implementing a uniformly distributed noise with exact amplitude is unpractical; besides, it can be shown that the performance of a uniformly distributed dither, contrary to that of a Gaussian dither, is quite poor if there are nonlinearity errors in the quantization [WK08].
The result depicted in Figs. 8-9 suggests that if one wants to improve the resolution using WGN, the ideal choice is a standard deviation $\sigma_n \cong 0.5$ LSB. This is indeed a typical choice of manufacturers who implement WGN in their DAS [Nat97], [Nat07]. However, even if the hypothesis of perfectly linear DAS is fulfilled, the ideal choice depends actually on the number of averaged samples, as shown in the next Section.

An exact closed-form expression for the function $g()$ represented in Fig. 8 is not available. In [CP94] a series expansion of $e_{qd}(x)$ is derived using typical methods of quantization theory; then, the series is truncated at its first term, squared and integrated. An asymptotic approximation of $g()$, usable for high enough $\sigma_n$, is derived:

$$
g(\sigma_n) \cong \frac{1}{\sqrt{2\pi}} e^{-2\sigma^2\sigma_n^2}. \quad (16)$$

This expression is implicit in [CP94], and written out explicitly in [SO05]; in both papers, it is recommended for $\sigma_n > 0.3$ LSB. For some computations, like those presented in the next Section, an accurate evaluation of $g()$ is needed also for $\sigma_n$ near to zero. This can be achieved with empirical approximate formulae.

The simpler approximation, which makes use of (16), is:

$$
g(\sigma_n) \cong g_1(\sigma_n) = \begin{cases} 
\frac{1}{\sqrt{12\pi}} - \sigma_n & \text{for } \sigma_n \leq 0.11 \text{ LSB} \\
\frac{1}{\sqrt{2\pi}} e^{-2\sigma^2\sigma_n^2} & \text{for } \sigma_n > 0.11 \text{ LSB}
\end{cases} \quad (17)$$

(the threshold 0.11 LSB achieves a nearly optimal approximation of $g()$ for this formula). A more accurate, even if less elegant approximation, is given by the expression (a refinement of that proposed in [AGS08]):

$$
g(\sigma_n) \cong g_2(\sigma_n) = k \frac{\varphi(\sigma_n, \mu_1, \sigma_1)}{\Phi(\sigma_n, \mu_2, \sigma_2)}, \quad (18)$$

where $\varphi(x, \mu, \sigma)$ is the Gaussian pdf (12), and $\Phi(x, \mu, \sigma) = \int_{\mu}^{x} \varphi(x', \mu, \sigma) dx'$ is the Gaussian cumulative distribution function. The five parameters $k, \mu_1, \sigma_1, \mu_2, \sigma_2$, are determined by a nonlinear LS fitting and have the values:

$$
k = 0.0774; \quad \mu_1 = 0.0190; \quad \sigma_1 = 0.1543; \quad \mu_2 = -0.0587; \quad \sigma_2 = 0.1201. \quad (19)$$

Both the approximations are quite good (Fig. 10): in the range $\sigma_n \in [0,1]$ LSB, $g_1(x)$ approximates $g(x)$ with a maximum error of $4 \cdot 10^{-3}$ LSB, while the error introduced by $g_2(x)$ is 20 times lower ($2.1 \cdot 10^{-4}$ LSB). It is to be remarked that the asymptotic expression (16) is usable for $\sigma_n$ as low as 0.11 LSB, and the condition $\sigma_n > 0.3$ is a bit too pessimistic.

6. Averaging a finite number of output samples

In practice, only a finite number of samples may be averaged. In order to evaluate the resulting ENOB it is convenient to derive a system equivalent to the noisy quantizer in
Fig. 1. Since the averaging is at the output, and not at the input, the signal-dependent and the signal-independent error contributions must be in reversed order. For the particular case of averaging infinite samples, the new equivalent system must reduce to that in Fig. 5. Therefore, the system is bound to have the form represented in Fig. 11.

The contribution $e_{qr}$ is a random error with standard deviation $\sigma_{qr}$, which takes into account the effect of the noise as seen at the output. As the error $e_{qd}$ is not removed at all by averaging, so $e_{qr}$ is completely removed by infinite averaging. The analysis of the system requires the introduction of the usual additive model of the deterministic error, obtaining the system in Fig. 12.

Fig. 10. Comparison between the numerically evaluated points of the function $g(\cdot \cdot \cdot)$, the asymptotic expression (16), and the approximations $g_1(\cdot)$ and $g_2(\cdot)$.

Fig. 11. Equivalent representation of the noisy quantization in Fig. 1.
Fig. 12. Equivalent representation of the noisy quantization for a fixed input signal. The input signal determines the parameters $\sigma_{qd}^2$ and $\sigma_{qr}^2$.

For a full-scale uniformly distributed input signal, the standard deviation $\sigma_{qd}$ is given by the function $g(\sigma_n)$ represented in Fig. 8 and approximated by expressions (16) and (18). As regards the determination of the standard deviation $\sigma_{qr}$, even if a formal analysis of the problem is quite complicated, it can be proven [SO05] that $e_{qr}$ is white and uncorrelated with $e_{qd}$. This can be seen as a direct consequence of the equivalence of the system with that in Fig. 2, where the input noise $n$ is white and uncorrelated with the deterministic error $e_q$. From the uncorrelation between $e_{qr}$ and $e_{qd}$ and the equivalence of the systems in Figs. 12 and 2 follows that

$$\sigma_n^2 + \sigma_q^2 = \sigma_{qd}^2 + \sigma_{qr}^2$$

and therefore

$$\sigma_{qr}^2 = \sigma_n^2 + \sigma_q^2 - \sigma_{qd}^2 = \sigma_n^2 + \sigma_q^2 - g^2(\sigma_n).$$

As a particular case, by assuming $\sigma_{qd} \approx 0$ (a condition achieved exactly only for $\sigma_n = +\infty$, and approximately for $\sigma_n \approx 0.5$ LSB) the random output error has variance

$$\sigma_{qr}^2 = \sigma_n^2 + \sigma_q^2 = \sigma_e^2,$$

i.e. the acquisition error is purely random.

Now, by substituting the system of Fig. 12 in the noise + quantization cascade of Fig. 4, it is easy to compute the MSE $\sigma_e^2$, and therefore the ENOB, obtained by averaging $N$ samples (Fig. 13).

Fig. 13. Equivalent representation of the WGN applied to a linear quantizer.

The ENOB of the system, from the uncorrelation between $e_{qd}$ and $e_{qr}$, and the whiteness of $e_{qr}$, is
\[ b_e = b - \frac{1}{2} \log_2 12 \sigma_n^2 = b - \frac{1}{2} \log_2 12 \left( \sigma_{qd}^2 + \sigma_{qg}^2 \frac{\sigma_q^2}{N} \right) \]  

This expression is first derived in [AGS04] (and re-published in [AGS08]), and recovered in the broader work [SO05]. An explicit expression of ENOB in terms of the rms input noise \( \sigma_n \) may be written in a simple approximate form or in exact form. The approximate formula is derived by assuming \( \sigma_{qd} \cong 0 \) and \( \sigma_{qg}^2 \cong \sigma_n^2 + \sigma_q^2 = \sigma_n^2 + 1/12 = \sigma_n^2 \): 

\[ b_e \cong b - \frac{1}{2} \log_2 \left( 1 + 12 \cdot \sigma_n^2 \right) + \frac{1}{2} \log_2 N. \]  

The exact formula is derived by substituting \( \sigma_{qd} = g(\sigma_n) \): 

\[ b_e = b - \frac{1}{2} \log_2 \left( 12 g^2(\sigma_n) + 1 + 12 (\sigma_n^2 - g^2(\sigma_n)) \right) \frac{\sigma_n^2}{N}. \]  

Figs. 14-16 show the result of numerical simulations of an 8-bit quantizer with various levels of input WGN (ranging from 0.05 to 0.5 LSB), and Fig. 17 shows the result of an analogous simulation for a 12-bit quantizer. Simulations results are compared with both expressions (24) and (25). In (25), the approximate function \( g(\cdot) \) has been used (slightly better agreement with simulations is obtained using \( g(\cdot) \); this is especially true for the case in Fig. 17.) Simulations basically demonstrate that (25) is able to predict with great accuracy the ENOB of a perfectly linear DAS with input noise and output averaging. A number of interesting and important facts follow from the validity of (25), and they are well illustrated by the curves in the figures.

1. For a given \( \sigma_n \), the maximum (asymptotic) increase of performance is given by (Fig. 18): 

\[ \Delta b = - \log_2 \sqrt{12} - \log_2 g(\sigma_n) \]  

An accurate evaluation of (26) for low \( \sigma_n \) can be obtained by using the approximation \( g(\cdot) \) given by (18) (values in Tab. 2). The approximation \( g(\cdot) \) given by (17) is also usable, obtaining an extension of the formula given in [CP94]: 

\[ \Delta b \cong \begin{cases} 
- \log_2(1 - \sqrt{12} \sigma_n) & \text{for } \sigma_n \leq 0.11 \text{ LSB} \\
\log_2 \frac{\pi}{\sqrt{6}} + 2 \pi^2 \sigma_n^2 \log_2 e & \text{for } \sigma_n > 0.11 \text{ LSB} 
\end{cases} \]  

The unbounded increase predicted by approximation (24) is untrue.

2. The usability of (24) depends on the actual number \( N \) of averaged samples, and not simply on \( \sigma_n \). Under this viewpoint it is quite inaccurate to say that \( \sigma_n \cong 0.5 \) is the right value to obtain an approximately full randomization of the quantization error. For example, \( \sigma_n = 0.3 \text{ LSB} \) is not too low for the validity of (24), provided that \( N < 32 \). On the other hand, \( \sigma_n = 0.5 \text{ LSB} \) is not sufficient to use (24), if \( N > 2^{15} \).
3. As a consequence, if one wants to add some WGN to increase performance by averaging, the choice is dictated by the number of samples that may be averaged. This is clearly suggested by the intersecting continuous lines in Fig. 18, and better illustrated by Fig. 19, in which ENOB is plotted as a function of $\sigma_n$ for fixed $N$. It is clear, for example, that for $N = 4$ it is convenient $\sigma_n \approx 0.2$ LSB, etc. Quite surprisingly, the very frequent choice $\sigma_n = 0.5$ is optimal only for $N$ of the order of $2^{13}$.

Fig. 14. ENOB of an 8-bit linear DAS with input WGN ($\sigma_n = 0.05$ LSB), as a function of the number $N$ of the averaged samples.
Fig. 15. ENOB of an 8-bit linear DAS with input WGN ($\sigma_n = 0.3$ LSB), as a function of the number $N$ of the averaged samples.
Fig. 16. ENOB of an 8-bit linear DAS with input WGN ($\sigma_n = 0.5$ LSB), as a function of the number $N$ of the averaged samples.
Fig. 17. ENOB of a 12-bit linear DAS with input WGN ($\sigma_n = 0.1$ LSB), as a function of the number $N$ of the averaged samples.
Fig. 18. Variation in the ENOB (with respect to the nominal resolution $b$) as a function of the number $N$ of the averaged samples, for different values of input WGN ($\sigma_n = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$ LSB). The figure compares the approximation given by (24) (approx. 1) with expression (25), in which the approximation (17) of $g(\cdot)$ is used (approx. 2).

<table>
<thead>
<tr>
<th>$\sigma_n$ [LSB]</th>
<th>$\Delta b$ [bit]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.59</td>
</tr>
<tr>
<td>0.2</td>
<td>1.50</td>
</tr>
<tr>
<td>0.3</td>
<td>2.92</td>
</tr>
<tr>
<td>0.4</td>
<td>4.92</td>
</tr>
<tr>
<td>0.5</td>
<td>7.48</td>
</tr>
</tbody>
</table>

Tab. 2. Maximum (asymptotic) increase of ENOB attainable by averaging, for given levels $\sigma_n$ of input WGN.
7. Conclusions

The chapter examines the overall effect, in terms of effective resolution, of input noise and output averaging in linear DAS. The analysis applies to both the cases of unwanted system noise, and of noise purposely added to increase the performance (non-subtractive dithering). After a brief discussion of the ENOB figure of merit, the equations to determine the ENOB in various situations are derived and validated by simulations. The results clarify the nature of the acquisition error in presence of noise — in terms of “dithered quantization error” \( q_{dl} \) and “randomized quantization error” \( q_{re} \) — and can be used, for example, to choose the optimal level of input noise in a non-subtractive dithering scheme. The choice is demonstrated to be non-trivial, even if quite simple with the use of the proper equations. In particular, the very common choice \( \sigma_n = 0.5 \) LSB is demonstrated to be suboptimal in most practical cases.

A very important warning is that the presented analysis is limited to the case of perfectly linear DAS, and is not applicable in the common case of meaningful nonlinearity error affecting the DAS. The case of non-subtractive dithering in nonlinear DAS can be analyzed with means similar to those presented in this chapter. In particular, the optimal levels of...
noise for nonlinear DAS are considerably higher than those derived for linear DAS [AGLS07]. This is, however, the subject of a possible future extended version of the chapter.

8. Acknowledgements

The authors wish to thank prof. Mario Savino for helpful discussions and suggestions.

9. References


The book is intended to be a collection of contributions providing a birdâ€™s eye view of some relevant multidisciplinary applications of data acquisition. While assuming that the reader is familiar with the basics of sampling theory and analog-to-digital conversion, the attention is focused on applied research and industrial applications of data acquisition. Even in the few cases when theoretical issues are investigated, the goal is making the theory comprehensible to a wide, application-oriented, audience.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
