Coupling Dithering and Static Linearization in A/D Converters

F. Attivissimo, N. Giaquinto, A. M. L. Lanzolla, M. Savino

Laboratory for Electric and Electronic Measurements
Department of Electrics and Electronics (DEE) – Polytechnic of Bari
Via E. Orabona 4, 70125 Bari
phone +39080 5963318, Fax +39080 5963410, [attivissimo, giaquinto, savino]@misure.poliba.it

Abstract **-** This paper presents a study of the performance attainable by combining two different methods of linearization algorithms for A/D converters: dithering (which removes errors due to quantization) and static look-up table (which removes errors due to INL). The theory and the simulation results show two very important facts, i.e.: (i) the amplitude of the dither signal must be chosen according to the INL of the converter, and should be greater than the usual value of 0.5 LSB rms; (ii) using both the linearization techniques allows one to attain (in absence of other sources of error) an arbitrary high number of effective bits, proportional to the logarithm of the averaged samples.

Keywords: A/D conversion, error compensation, dithering,

1. INTRODUCTION

An analog-to-digital converter (ADC) is an intrinsically nonlinear device which, as such, causes certain degradation in the converted signal. It is often convenient to quantify the fidelity of the ADC output, and therefore the quality of the ADC itself, in terms of "effective number of bits" (ENOB) [1], which is exactly equal to the resolution of the device for an ideal quantizer, and is a lower number in presence of additional errors (besides the mere quantization). The most important source of additional error in actual ADCs is the integral nonlinearity (INL), which is a deviation of the actual staircase characteristic from the ideal quantization, due to an imperfect position of the threshold levels. Removing the static nonlinearity error introduced by quantization and INL is therefore a primary practical issue.

The methods for removing quantization and INL are quite different, because these effects, even if in a sense very similar (they are both a static nonlinearity), are in another sense of a very different kind. The ideal quantization is a "perfectly non-invertible" nonlinear characteristic. It can be corrected, therefore, only by "randomizing" and averaging (or filtering) the associated error: this is usually obtained via the well-known technique called *dithering* [2]. The INL is, instead, a comparatively smooth and large nonlinearity which is little affected by the dithering technique; it can be, instead, effectively reduced by implementing an inverse nonlinear function, for example by means of a look-up table [3].

In commercial instruments both the techniques are used, but rarely at the same time. The simultaneous application of linearization and quantization noise filtering can be found, in a sense, in some multibit sigma-delta converters, where the dithering is substituted (advantageously) by the noise shaping feedback loop, and the linearity errors of the inner multibit ADC are corrected by means of digital linearization [4]. Complete stand-alone instruments, instead, usually offer only one of the two possibilities (dithering and linearization). The possibility of activating a dither circuit is usually available in instruments with comparatively low INL (see for example [5], [6]); the possibility of a full "ADC calibration", which produces a linearization look-up table, is offered in instruments with comparatively high INL (see, for example, [7]). Implementing only one correction technique, dependent of the most prominent error effect, is however a questionable choice, which can prevent one from obtaining a substantial performance increase.

In the following, the simultaneous application of nonsubtractive dither and linearization via look-up table is analyzed. The performance increment is measured in terms of ENOB, and the different results obtained by applying dithering only, look-up table only, and both the methods are highlighted. Some

formulae derived from the theoretical analysis are presented, and compared with simulation results on a "realistic" ADC model. In the whole paper, the ideal quantization step is always assumed to be $Q = 1$: if $Q \neq 1$ all the derived equations concerning the ENOB keep their validity, provided that the involved quantities (rms noise σ_n, INL vector inl_k) are expressed in LSB units.
Experimental results are currently being derived and will be presented in a future, more complete version of the paper.

2. NONSUBTRACTIVE GAUSSIAN DITHER IN AN IDEAL ADC

In commercial instruments dithering is usually implemented, for reason of hardware and software simplicity, in a nonsubtractive scheme, using a Gaussian dither signal which is simply derived by the noise intrinsically present in every electronic devices. "Nonsubtractive scheme" means that the Gaussian noise (dither signal) with standard deviation σ_n is simply added to the ADC input $x(t)$, and the output samples are averaged or filtered (Fig. 1). The manufacturers' choice for σ_n is usually 0.5 LSB [5], [6]. Even if manufacturers do not explain the considerations (theoretical or empirical) leading to this choice, it appears fully justified by a theoretical analysis, presented by the authors in a former study [8], and summarized here in few lines for the reader's convenience.
In the scheme of Fig. 1 the dither signal and the quantization error are mutually independent, and therefore their powers are summed together. As a consequence, the variance σ_e^2 of the error $e = y - x$ is (under the hypothesis of ideal quantization) $\sigma_e^2 = \sigma_n^2 + \sigma_q^2$, being $\sigma_q^2 = 1/12$ the variance of the quantization error. On the other hand, the dither signal has the effect of (partly) randomizing the quantization error, which is instead a completely deterministic effect in perfect absence of noise. Therefore, if the final ADC output $\bar{y}$ is the average of *N* samples taken with the same input x, the variance of the overall error $\bar{e} = \bar{y} - x$ is:

$$\sigma_{\bar{e}}^2 = \sigma_{qd}^2 + \frac{\sigma_{qr}^2}{N} \tag{1}$$

where σ_{qd}^2 is the variance of the residual deterministic part of the quantization error, and $\sigma_{qr}^2 = \sigma_n^2 + \sigma_q^2 - \sigma_{qd}^2$ is the random part of the error. The quantity σ_{qd} depends only upon the dither amplitude σ_n: in [8], the authors have calculated the amount of σ_{qd} for various σ_n, obtaining the result represented in Fig. 2. The graph shows that $\sigma_n \cong 0.5$ LSB is the minimum amount of noise that randomizes almost completely the quantization error, and is therefore an optimal choice. With a simple computation, one derives that the ENOB obtained by averaging N samples in an ideal quantizer with dither $\sigma_n \geq 0.5$ LSB is given by:

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

Fig. 3 compares Eqn. (2) with simulation results, relevant to an ideal 8-bit quantizer with dither $\sigma_n = 0.4$ and $\sigma_n = 0.5$ LSB. The choice $\sigma_n = 0.4$ LSB is clearly suboptimal for large *N*: the residual deterministic part of the quantization error prevents, indeed, the ENOB from following the increase predicted by Eqn. (2) for $N > 2^{11} = 2048$. With $\sigma_n = 0.5$ LSB, instead, the ENOB increase follows closely Eqn. (2) up to $N = 2^{16} = 65536$. In order to follow Eqn. (2) for larger *N*, one needs a larger σ_n, in order to further reduce the residual deterministic quantization error. It is interesting to notice that with $\sigma_n = 0.5$ LSB, Eqn. (2) becomes $b_e = b - 1 + \frac{1}{2}\log_2 N$, and therefore dithering is convenient $(b_e > b)$ if $N > 4$ output samples are averaged for each input samples.
Despite the encouraging result represented by Eqn. (2) and Fig. 3, the choice $\sigma_n = 0.5$ LSB is dubious for an actual ADC, for two separate reasons.
First of all, in an actual ADC the quantization steps are of different sizes, so that many of them have size considerably greater than Q. For this reason $\sigma_n = 0.5$ LSB should not be expected to be a sufficient value for randomizing the quantization error in practice.

A second, and more important reason which prevents formula (2) from being valid for non-ideal ADCs is that, even if one succeeds in randomizing all the quantization error, the distortion of the characteristic caused by the INL makes ENOB far lower than the figure given by (2). If one has the possibility of averaging a large number N of samples, and wants to increase ENOB proportionally to $\log_2 N$, some linearization scheme to reduce the effect of the INL must be adopted.

3. NONLINEARITY AND MIDPOINT LINEARIZATION IN ACTUAL ADC'S

In a former work [9], the authors have calculated the ENOB reduction associated to the presence of INL in a converter, and the ENOB increase produced by an optimal linearization implemented via look-up table. The static nonlinearity of an ADC is described by the INL vector $inl_k = t_k^{id} - Gt_k - O$, where t_k^{id} and t_k are the ideal and the actual positions of the threshold levels, and G, O are the gain and the offset (ideally 1 and 0, respectively). In the equations reported in this section also the differential nonlinearity (DNL) is involved, quantified by the vector $dnl_k = inl_{k+1} - inl_k$.
The variance of the overall conversion error, as demonstrated in [9], is increased exactly of the *mean squared integral nonlinearity* $\overline{inl_k^2}$, so that the ENOB of an ideal ADC affected only by INL errors is $b_e = b - (1/2)\log_2(1 + 12 \cdot \overline{inl_k^2})$. The main problem raised by the presence of the INL is that the dither has very little effect on this "large-scale" nonlinearity, which is an essentially deterministic error even in presence of large input noise. Therefore, the variance of the overall error in a dithered nonlinear ADC is (if the quantization error is fully randomized by the dither):

$$\sigma_{\bar{e}}^2 \cong \frac{\sigma_n^2 + \sigma_q^2}{N} + \overline{inl_k^2} \,. \tag{3}$$

Consequently, the ENOB is given by:

$$b_e \cong b - \frac{1}{2}\log_2\left(\frac{1 + 12 \cdot \sigma_n^2}{N} + 12 \cdot \overline{inl_k^2}\right). \tag{4}$$

This equation can be verified by simulations. We have simulated an 8-bit nonlinear ADC with the INL depicted in Fig. 4 and the DNL depicted in Fig. 5 (this is the actual nonlinearity of a digital instrument, and precisely of a digital oscilloscope), obtaining the results shown in Fig. 6. The results demonstrate that:

- as predicted by (4), neither $\sigma_n = 0.5$ LSB nor $\sigma_n = 1.0$ LSB are able to randomize meaningfully the error associated to INL, so that it is impossible to achieve the big ENOB increase obtained with the ideal (linear) quantizer;
- the asymptotic value of b_e for large N is approximately $b - (1/2)\log_2(12 \cdot \overline{inl_k^2})$; this is slightly smaller than the actual value because the dither has a small "smoothing" effect on the INL, not taken into account by Eqn. (4).

The solution for obtaining an ENOB increase proportional to $\log_2 N$, like in the ideal quantizer case, obviously can consist only in the introduction of a linearization scheme. The study presented in [9] has shown that the linearizing optimal look-up table substitutes the k-th ADC output (y_k) with a new digital output y'_k which is the average of the actual threshold levels t_k, t_{k-1}:

$$y'_k = (t_k + t_{k-1})/2 \,. \tag{5}$$

This method can be named "midpoint linearization". According the results in [9], the midpoint linearization decreases the overall error variance, substituting the term $\overline{inl_k^2}$ with a smaller term, approximately equal to $\overline{dnl_k^2}/4$. The main advantage of this linearization is, however, that the residual error, being associated to the DNL, has a far more "granular", or "small-scale" structure than those associated to the INL. Therefore, it is a very natural idea to couple *dithering* and *midpoint linearization* in order to obtain an ENOB increase proportional to $\log_2 N$, like in the ideal quantizer case.

4. COUPLING DITHERING AND MIDPOINT LINEARIZATION

The examined linearization scheme, which couples nonsubtractive dithering and midpoint linearization, is as follows:

1) add to the input x a Gaussian noise with standard deviation σ_n ;
2) take the ADC output y and apply the midpoint linearization of Eqn. (5);
3) average N output samples in order to obtain the final ADC output $\bar{y}'$.

It should be noted that this scheme (which could be called "linearize+average") is different from another way of coupling dithering and linearization, already analyzed in literature [10], [11]. This alternative consists in *first* averaging and *then* applying the linearization, and could therefore named "average+linearize". The disadvantage of the latter scheme is that the averaged ADC output is *not* coarsely quantized (like the raw output) and therefore cannot be linearized using a simple look-up table. Therefore, an almost continuous linearization function must be implemented in the "average+linearize" method.

The theoretical questions relevant to the "linearize+average" scheme described here are:

- What value of b_e is attained with given N, σ_n and dnl_k ?
- What is the optimal choice for σ_n ?

The answer to the first question is quite simple if one supposes that the dither signal is strong enough to randomize completely the A/D conversion error. As highlighted in the former section, this hypothesis is quite reasonable due to the fact that the, after the midpoint linearization, the residual nonlinearity error has the granular nature typical of the DNL and of the quantization error. Under this hypothesis, the overall conversion error after averaging is:

$$\sigma_{\bar{e}}^2 \cong \frac{\sigma_n^2 + \sigma_q^2 + \overline{dnl_k^2}/4}{N} . \tag{6}$$

Of course this equation derives from (3), by substituting $\overline{inl_k^2}$ with $\overline{dnl_k^2}/4$, and considering the latter term of random nature and therefore affected by the averaging, just like σ_n^2, σ_q^2 . From (6), the ENOB is readily derived and given by:

$$b_e \cong b - \frac{1}{2}\log_2\left(1 + 12\cdot\sigma_n{}^2 + 3\cdot\overline{dnl_k^2}\right) + \frac{1}{2}\log_2 N \tag{7}$$

As regards the second of the two questions above, we have already pointed out that, by theoretical considerations, the full randomization of the error requires $\sigma_n \cong 0.5$ LSB for an ideal quantizer, and a higher value for an actual ADC with irregular quantization steps. Even if the optimal value of σ_n is clearly dependent on the actual nonlinearity pattern of the ADC, it is quite reasonable to consider the *maximum DNL*, and increase the 0.5 LSB value (optimal for an ideal ADC) of a related quantity. The simulations has shown, in particular, that a good rule of the thumb is to choose:

$$\sigma_n \approx 0.5\ \text{LSB} + 2\cdot\max(|dnl_k|) . \tag{8}$$

The most significant simulation results are reported in Figs. 7 and 8. The figures show the ENOB increase obtained with and without midpoint linearization, comparing the latter result with Eqn. (7).

Fig. 7 is relevant to the case $\sigma_n = 0.5$ LSB, and shows first of all that the ADC with midpoint linearization has a much greater ENOB, for large N, than the one without linearization. It should be noted, indeed, that for the same ADC, without dither or noise, the performance increase associated to the linearization is quite modest: for $\sigma_n = 0$ we have $b_e = 7.66$ without midpoint linearization, and $b_e = 7.90$ with midpoint linearization. Therefore the midpoint linearization is truly exploited when coupled with dithering and averaging. Fig. 7 shows, however, that $\sigma_n = 0.5$ LSB is not sufficient for a full randomization of the conversion error, since Eqn. (7) is followed only for $N < 2^9$.

Fig. 8 shows, instead, the results obtained with the rule of the thumb (8), which yields $\sigma_n \approx 1.4$ LSB (since the simulated ADC has $\max(|dnl_k|) = 0.44$ LSB. With this choice, Eqn. (7) is closely followed up to $N = 2^{16}$, just like in the ideal ADC case with $\sigma_n = 0.5$ LSB.

Of course, the larger value of σ_n is disadvantageous if one can average few samples, but is a big advantage with large *N*: for example, $N = 2^{17}$ samples yields $b_e = 13.9$ bits, while only $b_e = 11.9$ bits is reached with $\sigma_n = 0.5$ LSB.

5. CONCLUSIONS

Two are the main results of the study presented in this paper.

The first result is that *coupling midpoint linearization and dithering is highly advisable*, even for ADCs with small INL. For this kind of ADCs the linearization could seem scarcely useful, because it gives a very modest increase of effective bits if employed alone (without dithering). But the linearization greatly improves the effectiveness of the dithering: the effective number of bits of the linearized dithered ADCs is always much greater than the one of the dithered-only ADC, provided a large enough number of samples are averaged.

The second result is that the choice $\sigma_n = 0.5$ LSB for a Gaussian nonsubtractive dither, employed by many manufacturers and justified by the theory of the ideal quantizer, is suboptimal for a non-ideal quantizer affected by INL. Even for an ADC with modest INL and DNL (≈ 0.5 LSB maximum), like to one examined in this paper, the needed amplitude of the dither signal is about 1.5 LSB, if one wants to randomize completely the A/D conversion error. Of course the better choice of σ_n depends heavily on the number *N* of the averaged samples.

6. REFERENCES

[1] IEEE Standard 1057/94, "IEEE Standard on Digitizing Waveform Recorders", IEEE Press, 1994.

[2] R. A. Wannamaker, S. P. Lipshitz, J. Vanderkooy, J. N. Wright, "A theory of nunsubtractive dithering", IEEE Trans. Signal Proc., vol. 48, n. 2, Feb. 2000, pp. 499-516.

[3] E. Balestrieri, P. Daponte, S. Rapuano, "A state of the art on ADC error compensation methods", IEEE Instrumentation and Measurement Technology Conference (IMTC/04), Como, Italy, May 2004, pp. 711-715.

[4] PXI-5922 Data Sheet, National Instruments, 2005.

[5] PCI-1200 User Manual, National Instruments, Jan. 1997.

[6] AT E Series User Manual, National Instruments, May 2002.

[7] 16533/34 Digitizing Oscilloscope Help Volume, Agilent Technologies, Jan. 2003.

[8] F. Attivissimo, N. Giaquinto, M. Savino, "Evaluating measurement uncertainty in A/D converters with and without dither", Proc. 8th International Workshop on ADC Modelling and Testing (IWADC/03), Perugia, Italy, Sept. 2003.

[9] N. Giaquinto, M. Savino, A. Trotta, "Detection, digital correction and global effect of A/D converters nonlinearities", 1st IMEKO Workshop on ADC Modelling, Smolenice, Slovak Rep., May 1996, pp. 122-127.

[10] J. Holub, O. Aumala, "Large scale error reduction in dithered ADC", Proc. 5th International Workshop on ADC Modelling and Testing (IWADC/00), Vienna, Sept. 2000, pp. 42-48.

[11] F. Adamo, F. Attivissimo, N. Giaquinto, A. Trotta, "A/D converters nonlinearity measurement and correction by frequency analysis and dither", IEEE Trans. Instrum. Meas., vol. 52, n. 4, Aug. 2003, pp. 1200-1205.

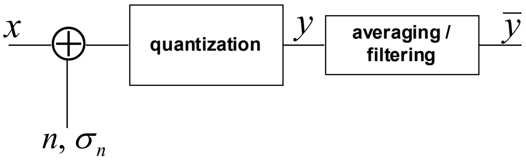

Fig. 1. – Nonsubtractive dithering scheme.

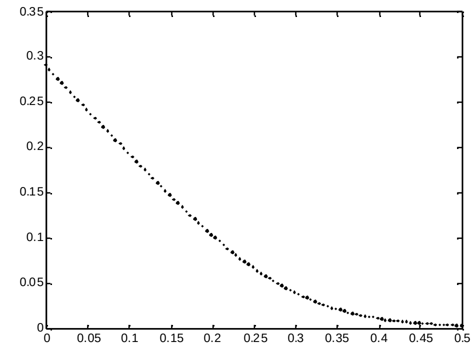

Fig. 2. – Plot of the residual deterministic quantization error σ_{qd} as a function of the input noise σ_n (both are in LSB units).

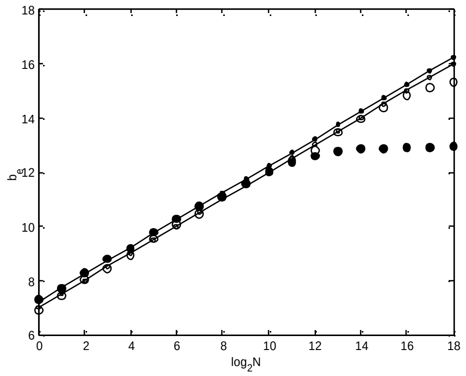

Fig. 3. – ENOB for an ideal quantizer with dither as a function of the number N of averaged samples. Black points refer to $\sigma_n = 0.4$ LSB , white points to $\sigma_n = 0.5$ LSB . The straight lines represent the value predicted by Eqn. (2).

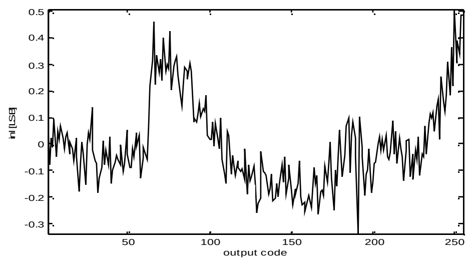

Fig. 4. – INL of the simulated ADC.

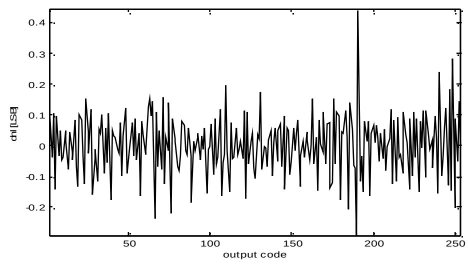

Fig. 5. – DNL of the simulated ADC.

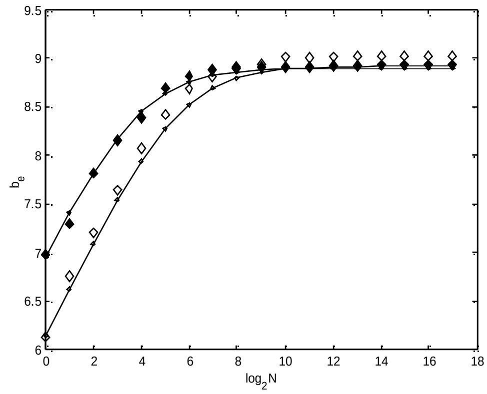

Fig. 6. – ENOB for a nonlinear quantizer with dither as a function of the number N of averaged samples. Black diamonds refer to $\sigma_n = 0.5$ LSB , white points to $\sigma_n = 1.0$ LSB . The straight lines represent the value predicted by Eqn. (4).

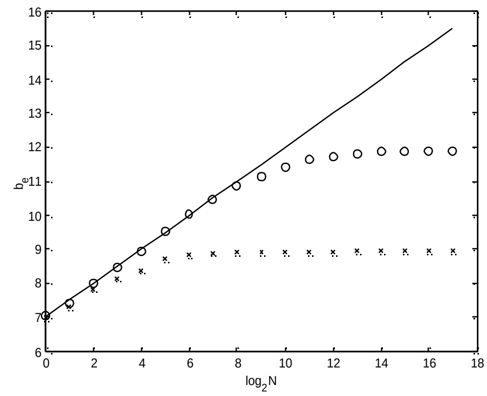

Fig. 7. – Effective bits for a nonlinear quantizer with midpoint linearization and dither ($\sigma_n = 0.5$ LSB), as a function of the number N of averaged samples. Circles report the actual result of the simulation, the straight line is Eqn. (7), while crosses report, for comparison, the ADC performance without midpoint linearization.

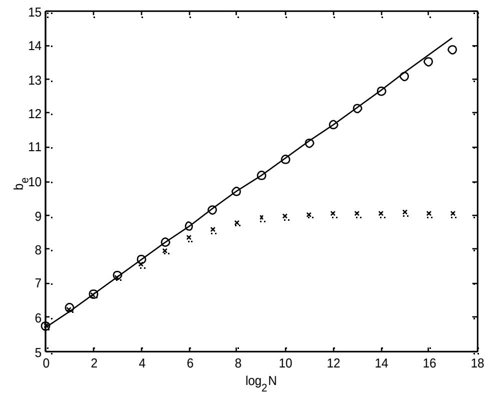

Fig. 8. – Effective bits for a nonlinear quantizer with midpoint linearization and dither ($\sigma_n = 1.4$ LSB), as a function of the number N of averaged samples. Circles report the actual result of the simulation, the straight line is Eqn. (7), while crosses report, for comparison, the ADC performance without midpoint linearization.