
©20xx IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
On the Design of Incremental $\Sigma\Delta$ Converters

M. Belloni, C. Della Fiore, F. Maloberti
Department of Electronics
University of Pavia
Pavia, Italy
Email: massimiliano.belloni@unipv.it

M. García Andrade
Department of Electronics
INAOE
Tonantzintla, Puebla, Mexico
Email: mgarcia@inaoep.mx

Abstract—The theoretical basis and design techniques of incremental converters for obtaining optimum performances are presented. The most suitable architectures and signal processing are identified. They allow to limit the loss of resolution caused by active components with finite gain and bandwidth. Matching of elements that enables the use of multi-bit quantizers is also discussed.

I. INTRODUCTION

The sigma-delta technique is suitable for signals that are busy enough for representing the quantization error as noise. For very slow signal, at the limit $dc$, the condition is not verified: the originated limit cycles cause in-band tones whose amplitude and frequency depends on the the input level. Since many critical amplitudes initiate limit cycles even using high order modulators, obtaining high resolution with low oversampling can be problematic.

Because of the poor control of limit cycles sigma-delta based converters with very slow signals at the input use a reset of the initial conditions and renounce to exploit the noise shaping. The use of a first order modulator and just a simple accumulation of the digital output obtains an incremental converter. The architecture requires $2^n$ clock periods for securing $n$ bit of resolution.

The use of an initially reseted high order modulator realizes a high order incremental converter. The higher order partially limits the limit cycles problem and a cascade of integrators (CoI) or more complex digital processing obtains high resolution with a relatively small number of clock periods [1], [2]. However, at some critical amplitude the tones problem is still relevant requiring to use dither for breaking the limit cycles and spreading its power over the Nyquist interval.

The above considerations come from studying the operation of the incremental converters in the frequency domain as it is customary for oversampling schemes. However, since incremental converters use a reset at the beginning of the conversion cycle, the memory of the past history is cancelled making the system time-variant. Therefore, the study in the frequency domain is formally not correct but also is not an effective tool of investigation for predicting performance; instead it is more conveniently studying the operation in the time domain as it is done in this paper.

Since the use of basic blocks with real response and inaccurate matching between elements vanishes the advantage of using high order modulators, the following studies how the limits affect performances and discusses how to keep the modulator within regions of minimum sensitivity.

Even if this paper studies a second order architecture, a third order scheme can possibly obtain higher resolution at the cost of more demanding analog sections. However, as shown in the following, a second order modulator can obtain 20.1, 22.5 and 24.2 bit with 128, 256, and 512 samples respectively by using an affordable circuit implementations.

II. INCREMENTAL $\Sigma\Delta$ MODULATORS

A double ramp converter integrates starting from zero the input signal in the continuous-time or sample-data domain and compares the result with the integration of the reference voltage performed in a smaller time or smaller clock periods. A first order $\Sigma\Delta$ with zero initial condition obtains the same operation but superposes at the same time of the integration an inverse reference. The output of the integrator determines every clock period whether the input and the one of the inverse of the reference. The output of the integrator determines every clock period whether the input is close enough for representing the quantization error as noise.

The operation of the first order modulator of Fig. 1 a) can be modified as shown in Fig. 1 b) to obtain every clock period the output of the integrator. In the sampled data domain we have

$$x + \sum_{1}^{n} y(n) = \frac{P(n)}{n} \quad (1)$$

Since $P(n)$ has limited amplitude the second term decreases in average as $1/n$ making $\sum_{1}^{n} y(n)/n$ closer and closer to $x$. If $P(n)$ is zero the error in the measure of $x$ is zero. Actually, the accuracy of the measure of $x$ increase by making an average of a number of equations (1). For example, accounting for the clock periods from $N - k$ to $N$, it results...
III. **Optimum** **ΣΔ Architecture**

The finite performance of the op-amp, the on-resistance of switches, the clock feed-through and the mismatch between unity elements are key practical limits for ΣΔ modulators.

The finite gain and bandwidth of the op-amp alter the transfer function of the integrator by changing its gain and causing a phase error. Moreover, the slew-rate is responsible of a non-linear incomplete transfer of charge between the input and the integrating capacitor.

The clock feed-through normally faced by using the ’switch before’ technique gives rise to a signal independent charge injection. The limit for dc measures corresponds to an input referred offset.

The mismatch between unity elements is normally corrected by dynamic element matching (DEM) techniques that are effective with busy input because they transform the mismatch into a shaped noise. However, for dc inputs the effectiveness of DEM is not ensured.

Assume that the above described limits give rise, every clock period, to an input referred error $\epsilon_{in}(i)$ which after an $N$ clock periods measure affects the accuracy of $x$ by

$$\delta_x(N) = \frac{\sum_{i=0}^{N} \epsilon(i) H_x(N-i) - \sum_{i=0}^{N} H_y(N-i)}{\sum_{i=0}^{N} H_y(N-i)}$$

If the modulator passes the input through $L$ integrators the coefficient $H_y(k)$ is $k \cdots (k + L - 2)/(L - 1)!$ while the denominator $N(1) \cdots (N + L - 1)/L!$ is the full scale normalization factor.

Observe that the contribution of the error $\epsilon_{in}(i)$ to $\delta_x(N)$ diminishes as $i$ increases and is maximum for the first error of the series. Namely, for a cascade of $L$ integrators the first error (or initial conditions error) is after processing attenuated by $(N + L - 1)/L$ giving rise to the condition

$$\epsilon(1) < \frac{V_{FS}}{2^{N_{bit}}} \cdot \frac{L}{N + L - 1}$$

for securing $N_{bit}$ of resolution.

As known, the offset of the op-amp used in the first integrator gives rise to an equal input referred offset [4]; therefore, it is necessary to use compensation methods like the chopper stabilized technique that modulates the offset with a square wave at half clock period [5]. The method does not completely cancel the limit because, as it results from equation (2), the one polarity is multiplied by coefficients $H_x(k)$ with $k$ even and the other polarity with $k$ odd. Therefore, the input referred offset with $N$ processed outputs is

$$V_{os}(N) = V_{os,1} \frac{\sum_{i=0}^{N/2} H_x(N - 2i) - \sum_{i=1}^{N/2} H_x(N - 2i - 1)}{\sum_{i=1}^{N} H_x(N - i - 1)}$$

The input referred noise caused by the first op-amp and the $kT/C$ noise gives rise to a sequence of uncorrelated samples that are superposed quadratically after the multiplication by $H_x(k)$. The result is that the weighting determined by $H_x(k)$ gives rise to an improvement equivalent to the plain superposition of $N$ samples: $\sqrt{N}$. Therefore, the input referred noise voltage thanks to the equivalent averaging of $N$ samples can be $\sqrt{N}$ times the LSB.

The above study gives rise to the following design hints:

- the output voltage and swing of the first op-amp of the modulator should be as low as possible for remaining...
in the region with higher gain and limiting the slewing requirements;
• it is recommended to obtain a low value of \( P_L(N) \) or possible to measure it for compensating for its effect;
• offset and errors at the initial stage of the conversion cycle must be carefully minimizes.

The suggestions give rise, together with additional considerations that will follow, to the possible optimum architecture shown in Fig. 3. The scheme is a second order modulator with \( n \)-bit quantization and a feed forward branch. The output of the modulator is processed by a digital matching filter and the averaging of a suitable number of results. The accuracy of the signals used for the averaging are improved by the measure of \( P_2 \) with an \( m \)-bit converter (\( m > n \)). The multibit DAC used in the main loop is controlled by a suitable processing that will be discussed shortly. The \( X \) in the first integrator denotes a suitable technique for cancelling the offset also discussed below. The reset of the two integrators is controlled by different signals.

**IV. DESIGN ISSUES**

The voltage swing of the first integrator in Fig. 3 is low thanks to the feed-forward path and the multi-bit quantizer which quantization interval is \( \Delta = \frac{V_{FS}}{2^n_{bit}} \). Since after few clock periods the swing is in the \( \pm \Delta \) interval, the error caused by finite gain and bandwidth are significantly reduced permitting less demanding op-amp specifications. Also a reduced swing benefits the slew-rate specification.

However, at the first clock cycle the signal from the feedback path is still zero and the output of the fist integrator equals the input signal. The real performance of the op-amp determine an error that is much larger than during the successive clock periods.

Unfortunately, as already discussed, an initial conditions error is attenuated by a small factor and limits the accuracy much more than the successive errors. Therefore, it is necessary to have initial conditions with minimum error. The request is ensured by delaying the reset of the first integrator by a clock period so that the feed-forward path establishes the input signal at the input of the second integrator after the first clock period. When the first reset is released the feedback path provides a proper signal for bringing the swing of the first op-amp within the steady limits as shown in Fig. 4.

A second design issue concerns the compensation of the offset. The chopper stabilized method is not fully effective as the result of equation (6) applied to the scheme of Fig. 3 obtains \( V_{OS}(N) = N/2 \cdot 2^{n_{bit}} \). As a matter of facts, the equation (6) can determine a better result by equalizing the contributions in the positive sum with the contributions in the negative sum. A possible solution is to use a single chopping after a suitable number of clock periods such that

\[
\sum_{i=0}^{p} H_x(N - i) = \sum_{i=p+1}^{N} H_x(N - i) \tag{7}
\]

that for the scheme of Fig. 3 obtains the offset accumulation shown in Fig. 5 for \( p = 38 \).

Finally let us consider the use of multibit quantization that reduces the swing of the first op-amp but poses the problem of the mismatch between the unity elements of the internal DAC. Assume to use \( R \) unity capacitances whose value is

\[
C_i = C_u(1 + \epsilon_i), \quad i = 1, \cdots R \tag{8}
\]
and that the average of the capacitances equals, for simplicity, \( C_u \). Therefore

\[
\sum_{i=1}^{R} \epsilon_i = 0
\]  

(9)

If at the \( k \)-th clock period the DAC uses the capacitance \( C_i \) its error \( \epsilon_i \), before the normalization to the full scale \( \sum_{i=0}^{N} H_x(N-i) \) is multiplied by \( H_x(N-k) \). Summing up the effects and accounting for the benefit of the m-bit quantizer, the error caused by the mismatch will be, in LSB

\[
\delta V_{\text{mism}} = 2^m \sum_{i=1}^{R} \epsilon_i \sum_{p_{i}} H_x(N-p_{i})
\]  

(10)

where \( p_{i} \) is the set of time clocks at which the capacitor \( C_i \) is used.

The use of equation (9) shows that the error is zero if the processing used to select the capacitors used by the DAC is such that

\[
\sum_{p_{i}} H_x(N-p_{i}) = K
\]  

(11)

for any \( i \). The condition 11 is ensured by a corresponding processing every clock period.

V. SIMULATION RESULTS

The scheme of Fig. 3 has been extensively simulated at the behavioral level to verify the features discussed above. Fig. 6 shows the equivalent number of bit with \( n = 2 \) and \( m = 5 \). The number of samples of the output stream is 128, 256 and 512 and the average is taken on the last four measures. The result shows that the minimum resolution is 20.1 bit, 22.5 bit and 24.2 bit respectively.

The effect of finite gain of the first op-amp (N=128) is depicted in Fig. 7. The plot compares a conventional structure with a modulator with feed-forward and the version with delayed reset. The result shows that for securing 20 bit the conventional solution requires \( A_0 = 140 dB \) while for the one proposed here a more affordable \( A_0 = 60 dB \) is sufficient. The benefit of averaging of results (equation (2)) is significant for \( k \) in the 4-8 range. For \( k > 8 \) the resolution drops and the benefit of averaging vanishes at \( k \) around 100, as shown in Fig. 8.

VI. CONCLUSIONS

It was demonstrated that it is possible to reduce the number of clock pulses by averaging the last outputs of the digital filter. Feed-forward with a delayed reset is used in order to limit the output amplitude of the first integrator. This relaxes the specification of gain for the first op-amp, which in this case an accuracy of 19 bit is reached with \( A_0 = 60 dB \) and 128 clock periods. A suitable way to reduce offset was also presented.

REFERENCES