A COMBINED SUCCESSIVE AND Σ–Δ A/D CONVERSION SCHEME

Li Ping

Department of Electrical and Electronic Engineering
University of Melbourne, Australia
Tel: (+61 3) 344 4487, Fax: (+61 3) 344 6678
E-mail: lp@ees.mmu.oz.au

ABSTRACT

This paper presents a mismatch-insensitive approach to the multi-bit Σ–Δ modulator design. The multi-bit conversion is realised by a successive approximation and a simple controlled averaging technique is introduced to cancel the first order mismatch error. The cost in circuit complexity increase is very modest.

I. INTRODUCTION

The Σ–Δ modulator adopts an oversampling technique to trade off speed for resolution [1-3]. As a result, the required high oversampling ratio is a major factor limiting the bandwidth. High order (≥3) Σ–Δ modulators have been investigated to reduce the oversampling ratio and the instability problem can be improved by the multi-bit conversion approach [4-7]. A number of techniques have been proposed to eliminate the component matching error in multi-bit D/A conversions but they tend to increase the circuit complexity [4-9].

In this paper a combined successive and Σ–Δ A/D converter is described. With a simple controlled averaging scheme, the first order mismatching error due to the successive approximation is cancelled by the Σ–Δ operation. Stable high order Σ–Δ modulation can be realised to reduce oversampling ratio or to increased resolution.

II. PRELIMINARIES

Oversampling:

Let $f_C$ and $f_{in}$ be the bandwidth and the sampling rate of the input analog signal $x_{in}$, respectively. With oversampling, $f_{in}$ is higher than that required by the Nyquist criterion so that a relatively wider transition band can be tolerated for the analog anti-aliasing filter. Fig. 1. To remove the out-of-band noise, a digital decimation filter is used after the A/D conversion, bringing down the sampling rate at the output to the Nyquist frequency, $f_{out}=2f_C$. The oversampling rate is defined by

$$M = \frac{f_{in}}{2f_C} = \frac{f_{in}}{f_{out}}$$

(1)

Σ–Δ Modulation:

Fig. 2a is an N-th order Σ–Δ structure, with

$$\Phi = \frac{1}{1 - z^{-1}}$$

(2)

In Fig. 2b the ADC and the DAC induced errors are modelled by two additive noise sources, $n_{ad}$ and $n_{da}$.

Fig. 2a: Conceptual diagram for an N-th order Σ–Δ modulator.

(b) The linearised model of (a).

For high order Σ–Δ modulation the quantisation noise, $n_{ad}$, can be reasonably assumed to be white and its spectrum is related to the quantisation bit number, $B$, by,

$$\text{SNR}_{ad}(f_{out}) = 2^B N_0$$

(3)

where $N_0$ is a constant.

The noise $n_{da}$ due to the DAC is zero if all the circuit element values are ideal. In this paper it is assumed that $n_{da}$ is caused by mismatching error only and we will concentrate on its elimination.

From Fig. 2b it can be derived that, in the z-domain,

$$y = \frac{(-\Phi)^N}{1 + z^{-1}(\Phi + \cdots + \Phi N)} x + \frac{1}{1 + z^{-1}(\Phi + \cdots + \Phi N)} n_{ad} + \frac{1}{1 + z^{-1}(\Phi + \cdots + \Phi N)} n_{da}$$

(4)

where $N_0$ is the quantisation noise and $x$ is the input signal. The noise $n_{da}$ due to the DAC is zero if all the circuit element values are ideal. In this paper it is assumed that $n_{da}$ is caused by mismatching error only and we will concentrate on its elimination.

From Fig. 2b it can be derived that, in the z-domain,
where \((1-z^{-1})^N\) is the so-called noise shaping function. Its amplitude is monotonically increasing in \([0, f_{in}]\). The second term in the above expression is the contribution of the ADC noise. At \(f_{s} = f_{c}\) it gives the upper bound of the ADC induced noise at the output (with (1) and (3)),

\[
n_u = (1-e^{-\sqrt{2}\pi f_c f_{in}}) N_{\text{ad}} = (2\sin(\pi/2M))N^2 - B Nq_d
\]

(5)

Notice that the decimation filter removes the noise components beyond \(f_c\), Fig.1.

In Fig.2 the operator \(\Phi\) can be readily realised by a standard switched capacitor integrator [10]. In the following only the final stage is considered in detail.

**III. COMBINED \(\Sigma-\Delta\) AND SUCCESSIVE SCHEME**

In the time domain the output of the last discrete integrator is given by, Fig.2a,

\[
w_N(i) = w_N(i-1) - q[w_N(i-1)] w_N(i-1)
\]

(6)

where \(q[w_N(i-1)]\) is the quantised value of \(w_N(i-1)\). Let,

\[
w_N(i) = w_N(i-1) - q[w_N(i-1)]
\]

(7)

Eqn.(6) can be broken into two recursive operations,

**Sampling:** \(w_N(i) - w_{N-1}(i) \rightarrow w_N(i)\)  
(8a)

**Quantisation and subtraction:** \(w_N(i)-q[w_N(i)] \rightarrow w_N^*(i+1)\)  
(8b)

**Fig. 3 A realisation of the final stage of the diagram in Fig.2, with the switching functions.**

A realisation of (8) is shown in Fig.3. Suppose that at the beginning of the \(i\)-th period, the voltage at \(w_N\) is \(w_N^*(i)\). When \(S_g\) is high the switches labelled \(S_{ga}\) are closed and \(w_{N-1}(i)\) is subtracted from \(w_N^*(i)\) to generate \(w_N(i)\), eqn.(8a). In the rest of this period, \(w_N(i)\) is successively approximated and quantised. Each quantised value is subtracted from \(w_N\). After B cycles, B bits are produced for \(q[w_N(i)]\), which are in the same time subtracted from \(w_N(i)\), establishing the correct initial value, \(w_N^*(i+1)\), for the next period, eqn.(8b).

**The reference generator, the sign control circuit and the DAC induced noise:**

During each cycle the reference generator is able to generate \(V_{Q_k} = V_d/2^k\), \(k = 1,2,..., B\), Fig.3. Each of these reference values is multiplied by a sign factor \(e(k)\) and then subtracted from \(w_N\). The sign factor, \(e(k)\), is determined by the output of the comparator,

\[
e(k) = \begin{cases} 1 & \text{if } k-1\text{th approximation cycle } w_N \geq 0 \\ -1 & \text{if } k-1\text{th approximation cycle } w_N < 0 \end{cases}
\]

(9)

The quantised output \(y\) is

\[
y(n) = B \sum_{k=1}^{B} e(k) V_{Q_k} = V_d \sum_{k=1}^{B} e(k) 2^{-k}
\]

(10)

The quantisation procedure starts half a cycle after the sampling instant, Fig.3, when \(w_N(n)\) is available. The last quantisation bit of \(w_N(n)\) is subtracted from \(w_N\) together with \(w_{N-1}(n+1)\).

A detailed implementation of the reference generator is shown in Fig.4, together with the sign control circuit.

**Fig. 4 (a) The reference generator and the sign control circuit.**

**Fig. 4 (b) The switching functions.**

Suppose that some additional circuitry, which is omitted for simplicity, is used to establish an initial voltage

\[
V_{CA}(1) = V_{CB}(1) = V_{Q_d} = V_d/2
\]

(11)

The following function is to be realised by the reference generator in each cycle,
\[ V_q(k+1) = 0.5 V_q(k) \quad \text{for } k = 2, 3, \ldots \]  
\[ (12) \]

It is important to note that here the superscript \( k \) is the index of cycle, not period. Each approximation cycle is divided into two phases, even and odd. Four switching functions are shown in Fig.4b. If a switch function = even (odd, 1 or 0, respectively) then the corresponding switches are closed only in the even (odd, both or neither, respectively) phase(s). For the sign control circuit,

\[
S_f = S_1 = \text{even} \quad S_g = S_h = \text{odd} \quad \text{if } e(k) = 1 
\]
\[ (13a) \]
\[
S_f = S_1 = \text{even} \quad S_g = S_h = \text{odd} \quad \text{if } e(k) = -1 
\]
\[ (13b) \]

A straightforward realisation of the successively-divided-by-two circuit is to let

\[
S_f = \text{even}, S_g = \text{odd}, S_h = S_h = 1 
\]
\[ (14) \]

In every cycle \( C_b \) is used to share the charge with \( C_h \) during even phase and subsequently it is discharged during odd phase. Thus,

\[
V_q(k+1) = \frac{C_b}{C_a+C_b} V_q(k) 
\]
\[ (15) \]

If \( C_b = C_b \), it is easily seen that eqn.(12) is realised in this ideal case. However, there is inevitably some mismatching between the two capacitors, say, \( C_b = C_a + \Delta C \). Define

\[
\delta = \frac{\Delta C}{2 C_a + \Delta C} 
\]
\[ (16a) \]

Then,

\[
\frac{C_b}{C_a+C_b} = 0.5 + \delta \quad \text{and} \quad \frac{C_a}{C_a+C_b} = 0.5 - \delta 
\]
\[ (16b) \]

Eqn.(15) becomes

\[
V_q(k+1) = 0.5 V_q(k) + \delta V_q(k) 
\]
\[ (17) \]

Compared with (12), the error term \( \delta V_q(k) \) results from the noise \( \eta_{da} \) in Fig.2b.

**Cyclic averaging:**

The error due to odd cycles can be reduced by discharging \( C_a \) and \( C_b \) alternatively. The switching functions are,

\[
S_f = \text{even} \quad S_f = \text{odd} \quad S_g = S_h = 1 \quad \text{for odd } k 
\]
\[ (18a) \]
\[
S_f = \text{even} \quad S_g = \text{odd} \quad S_f = S_h = 1 \quad \text{for even } k 
\]
\[ (18b) \]

From (16), when \( k=3 \),

\[
V_q(3) = \frac{C_b}{C_a+C_b} + \frac{C_a}{C_a+C_b} (V_f/2) = (0.5^2 - \delta^2)(V_f/2) 
\]
\[ (19) \]

It is seen that the first order error has been cancelled, which is also true for other odd \( k \) > 3.

As an example, if the component mismatch results in \( \delta = 0.1\% \), then the level of the second order error is below -120dB, which can be neglected for most applications.

**Controlled averaging:**

For even \( k \), first order error remains a dominant factor with the scheme based on (18). An improved method is that, instead of following a fixed pattern, \( C_a \) or \( C_b \) are discharged adaptively for even cycles.

First suppose that whichever of \( C_a \) or \( C_b \) is discharged in an even cycle, the opposite one will be discharged in the next (odd) cycle.

This guarantees that eqn.(19) and the statement below it are valid. Thus for any odd \( k \),

\[
\eta_q(k+2) = \eta_q(k+1) + \eta_f(k) 
\]
\[ (20) \]

The error term \( \pm \delta e(k+1) + \eta_f(k) \) has a constant magnitude. Its sign depends on whether \( C_a \) or \( C_b \) is discharged. For each sample period the total error due to even cycle quantisation is,

\[
\eta_{da}(i) = \sum_{k=2,4,6...}^{B} \pm \delta e(k+1) + \eta_f(k) 
\]
\[ (21) \]

Since the signs can be controlled, it is always possible to force

\[
\eta_{da}(i) = -\eta_{da}(i-1) 
\]
\[ (22) \]

for every \( i \). This can be done by adaptively selecting the discharging capacitor for even cycles. Consequently the spectrum of \( \eta_{da} \) is non-zero only at \( f_{in}/2 \) which would be filtered out by the decimation filter (Fig.1). Only very simple logic circuit is required for realising (22). It is, for the \( k \)-th cycle in the \( i \)-th period,

(i) when \( k \) even and \( i \) even, discharge \( C_a \) if \( e(k+1) = 1 \), otherwise \( C_b \),
(ii) when \( k \) even and \( i \) odd, discharge \( C_b \) if \( e(k+1) = 1 \), otherwise \( C_a \),
(iii) when \( k \) odd, discharge the capacitor which has not been discharged during the last cycle.

**Bit number, order and stability:**

Given \( B, M \) and \( N \) values, the resolution can be found from (5), which indicates that the resolution increases with order \( N \) for \( M>3 \). However, instability problems occur for high order structures. It has been found by computer simulation (up to \( N=B+10 \)) that the network structure of Fig.2 becomes unstable if \( N>B \) and its performance is satisfactory if \( N>B \) (here by satisfactorily we mean that eqn. (5) can be applied). When \( N>B \) and \( N \leq 3 \), the noise level has been found higher than that given by (5).

Let \( f_r = 1/\tau \) where \( \tau \) is the length of a cycle. \( B \)-cycles are required for a \( B \)-bit conversion scheme so that \( B f_{in} = f_r \). Compared with (1) we have,

\[
BM = f_{in}/\eta_{Out} 
\]
\[ (23) \]

Therefore the BM value is primarily determined by the ratio of the available circuit speed to the required bandwidth of the digitised signal. Setting \( P+BM \) as a constant, it would be interesting to find an optimal trade-off between \( B \) and \( M \). As a rough evaluation, we use \( N=B \) for (5),

\[
\eta_u = |2-P/M|N(2\sin(\pi(2M))/M|N(M)/M|N| 
\]
\[ (24) \]

for \( M \geq 3 \). The minimum of the above equation takes place at

\[
M = \pi/2 \approx 4.2677 
\]
\[ (25) \]

**IV. EXAMPLES**

In Fig.5a a typical simulation spectrum is shown for an ideal 8-bit, 3rd order \( \Sigma - \Delta \) modulator based on Fig.2a-4. The sampling frequency \( f_{in}=6000\text{Hz} \). The input is a \( 5031.25\text{Hz} \) sine wave with amplitude=0.8. If the mismatching is set at 6=0.01, (equivalent to \( \Delta C/C_a=4\% \)), the nonlinear noise sparks are at the level above 60 dB without the averaging technique, Fig.5b. However with the controlled averaging technique the noise floor, caused by the second and higher order mismatching error, is reduced to about 130dB, noticeable only in the low frequency range, Fig.5c. It is clearly seen that the controlled averaging technique can be useful to reduce the non-ideal mismatching noise significantly.
V. CONCLUSION

Generally speaking, successive A/D converters can achieve higher bandwidth than $\Sigma$-$\Delta$ ones at the cost of reduced accuracy. In this paper a compromise has been proposed to attain the advantage of both methods. It has been demonstrated that a controlled average technique can be employed to push the first order mismatching error out of the signal band. Thereby accurate successive conversion and stable higher order $\Sigma$-$\Delta$ modulation can be realised with a combined scheme.

REFERENCES


Fig. 5 (a) Simulated response of an ideal A/D converter.
(b) Simulated response of a non-ideal A/D converter, ($\delta=0.01$), without any averaging.
(c) Simulated response of a non-ideal A/D converter, ($\delta=0.01$), with controlled averaging.