A New High-Level Synthesis Methodology of Cascaded Continuous-Time $\Sigma\Delta$ Modulators

Ramón Tortosa, José M. de la Rosa, Senior Member, IEEE, Francisco V. Fernández, and Angel Rodríguez-Vázquez, Fellow, IEEE

Abstract—This brief presents an efficient method for synthesizing cascaded sigma–delta modulators implemented with continuous-time circuits. It is based on the direct synthesis of the whole cascaded architecture in the continuous-time domain instead of using a discrete-to-continuous time transformation as has been done in previous approaches. In addition to placing the zeroes of the loop filter in an optimum way, the proposed methodology leads to more efficient architectures in terms of circuitry complexity, power consumption and robustness with respect to circuit errors.

Index Terms—Analog–digital conversion, continuous-time sigma–delta ($\Sigma\Delta$) modulation.

I. INTRODUCTION

C O N T I N U O U S - T I M E (CT) sigma–delta modulators ($\Sigma\Delta$Ms) are suited to analog-to-digital (A/D) interfaces in submicrometer CMOS [1]. Although most $\Sigma\Delta$Ms employ discrete-time (DT) circuits, CT techniques are applicable to broadband data communication systems due, among other features, to their intrinsic anti-aliasing filtering and their suitability for fast operation with low power consumption [2], [3].

This brief presents a direct synthesis method for cascaded architectures. On the one hand, this method yields better placing of the noise transfer function zeroes and poles. On the other hand, it reduces the number of analog components. Thus, more robust and efficient architectures are obtained than when using DT-to-CT transformation. The method is illustrated in this brief through the synthesis of a 2-1-1-1 cascaded CT-$\Sigma\Delta$M for 12-bit at 40-MS/s.

II. CASCADED CT $\Sigma\Delta$ MODULATORS

This brief presents a direct synthesis method for cascaded architectures. On the one hand, this method yields better placing of the noise transfer function zeroes and poles. On the other hand, it reduces the number of analog components. Thus, more robust and efficient architectures are obtained than when using DT-to-CT transformation. The method is illustrated in this brief through the synthesis of a 2-1-1-1 cascaded CT-$\Sigma\Delta$M for 12-bit at 40-MS/s.

The authors are with the Institute of Microelectronics of Seville (IMSE), National Microelectronics Center (CNM-CSIC), 41012 Seville, Spain (e-mail: jrosa@imse.cnm.es).

Fig. 1. Conceptual block diagram of a cascaded CT-$\Sigma\Delta$M

Fig. 1 shows the conceptual block diagram of a $m$-stage cascaded CT-$\Sigma\Delta$M. Each stage, consisting of a single-loop CT-$\Sigma\Delta$M (typically either second or first order), modulates a signal that contains the quantization error generated in the previous stage. Once in the digital domain, the stage outputs ($y_k$) are processed and combined by the cancellation logic so that only the quantization error of the last stage remains. Also, this error is shaped by a transfer function whose order equals the sum of the respective orders of all the stages in the cascade.

Cascaded CT-$\Sigma\Delta$Ms are typically synthesized from equivalent DT systems and use the same digital cancellation logic [8] as the DT prototype. This DT-to-CT equivalence can be guaranteed because at each stage the corresponding feedback transfer function from the quantizer output to the quantizer input is of discrete-time nature [2]. Assuming a rectangular impulsive response at the digital-to-analog converter (DAC), this equivalent DT loop filter transfer function is calculated as [10], [11]

$$F(z) = \sum_{k} \text{Re} \left( \frac{F(s)}{s} \cdot \frac{e^{m_k T_c s}}{z - e^{T_{c} s}} \right) - \sum_{k} \text{Re} \left( \frac{F(s)}{s} \cdot \frac{e^{m_k T_c s}}{z - e^{T_{c} s}} \right)$$

(1)
where \( f_s = 1/T_s \) is the sampling frequency; \( m_1 = 1 - t_d/T_s \);
\( m_2 = 1 - (t_d + \tau)/T_s \); \( t_d \) and \( \tau \) are, respectively, the
time delay and pulselength of the DAC waveform; \( p_i \) are the poles of
\( F(s)/s \) and \( \text{Re}(x) \) stands for the residue of \( x \).

To get a functional CT-\( \Sigma \Delta M \) whose cancellation logic coincides
with that of the original DT-\( \Sigma \Delta M \), every state variable
and DAC output must be connected to the integrator input of
the ulterior stages in the cascade [8]. This increases the
number of analog components (transconductors, amplifiers,
and DACs) needed. For instance, Fig. 2(a) shows a cascaded
2-1-1 CT-\( \Sigma \Delta M \) obtained from an existing DT-\( \Sigma \Delta M \) [12].
Note that eight scaling coefficients \( (k_6g_2-a) \) and their corresponding
signal paths are needed to connect the different stages of
the modulator. As a counterpart, if the CT-\( \Sigma \Delta M \) is directly
synthesized by following the technique in this brief, the number
of integrating paths can be reduced—see Fig. 2(b).

III. PROPOSED METHODOLOGY

Let us consider the more general case of the \( m \)-stage cascaded
CT-\( \Sigma \Delta M \) shown in Fig. 1. The overall output \( y_m \) is given by

\[
y_m(z) = \sum_{k=1}^{m} y_k(z)C_{Lk}(z)
\]

(2)

where \( y_k(z) \) and \( C_{Lk}(z) \) denote, respectively, the output and
the cancellation transfer function of the \( k \)th stage.

If the modulator input \( x(t) \) is set to zero, the output of each
stage can be written as

\[
y_k(z) = \frac{E_k(z) + \sum_{i=1}^{k-1} Z\left\{ L^{-1}[H_D(F_{ik})]nT_s \right\} y_k(z)}{1 - Z\left\{ L^{-1}[H_D(F_{kk})]nT_s \right\}}
\]

(3)

where \( Z \) stands for the \( z \)-transform, \( L^{-1} \) is the inverse Laplace
transform, \( H_D = H_DAC(s) \) is the DAC transfer function, and

\[
F_{ij} = F_{ij}(s) = \frac{\text{Input Quantizer } j}{y_k(s)}
\]

(4)

represents the transfer function from \( y_k(s) \) to the input of the \( j \)th
quantizer [see Fig. 2(b)].

Using the notation \( Z(L^{-1}[H_D(F_{km})]nT_s) \equiv Z_{km} \),
the output of each stage is given by

\[
y_k(z) = \frac{E_k(z)}{1 - Z_{kk}} + \sum_{i=1}^{k-1} Z_{ik}y_k(z)
\]

(5)

and the output of the modulator can be written as

\[
y_m = \sum_{k=1}^{m} y_k C_{Lk} = \sum_{k=1}^{m} \left( \frac{E_k(z)}{1 - Z_{kk}} + \frac{1}{1 - Z_{kk}} \sum_{i=1}^{k-1} Z_{ik}y_k(z) \right) C_{Lk}
\]

(6)

The cancellation logic transfer function \( C_{Lk}(z) \) is calculated by
annulling the quantization errors of the first \( m - 1 \) stages. This

\[
C_{Lk}(z) = \frac{-Z_{km}C_{Lm}}{1 - Z_{mm}} = \frac{-Z \left\{ L^{-1}[H_D(F_{km})]nT_s \right\} C_{Lm}(z)}{1 - Z \left\{ L^{-1}[H_D(F_{mm})]nT_s \right\}}
\]

(7)

where the cancellation logic transfer function of the last stage,
\( C_{Lm}(z) \), can be chosen to be the simplest form that preserves
the required noise shaping.

Note that the design equations (2)–(7) do not only take into
account the single-stage loop filter transfer functions \( (F_{ii}) \),
but also the inter-stage loop filter transfer functions \( (F_{ij}, i \neq j) \).
The latter are CT integrating paths appearing only when the
modulator stages are connected to form the cascaded \( \Sigma \Delta M \) and
must be included in the synthesis methodology to obtain a
functional modulator with a minimum number of inter-stage paths.

Therefore, the following systematic procedure can be pro-
posed for the synthesis of cascaded CT \( \Sigma \Delta M \).1

1In this procedure, the modulator order, oversampling ratio, and number of
bits of internal quantizers are assumed to be determined for the given speci-
cfications from well-known expressions [1].
For illustration, the 2-1-1 CT-ΣΔM of Fig. 2(b) was synthesized using (2)–(7) to achieve 16-bit resolution in a 750-kHz bandwidth, with a sampling frequency of 48 MHz (oversampling ratio \( M = 32 \)) [12]. For simplicity, in order to facilitate the comparison of the performance of both modulators in Fig. 2, the coefficients of the first stage \((k_{in1}, k_{g1}, k_{p11}, k_{p2})\) are taken to be equal in both systems and are obtained from a DT-to-CT transformation of the first stage of the DT-ΣΔM in [12]. The rest of the coefficients in Fig. 2(b) are chosen such that the time constant of the integrators is the inverse of the sampling frequency \((T_s = 1/\dot{f}_s)\). In summary

\[
k_{in1} = -k_{f4} = 1/4 \\
k_{p2} = -3/8 \\
k_{g1} = k_{in2} = -k_{f3} = k_{in3} = -k_{f4} = 1.
\] (8)

Thus, the single-loop and inter-stage transfer functions are

\[
F_{11} = \frac{-3T_s}{sT_s + 3/4} \\
F_{22} = F_{33} = \frac{-1}{sT_s} \\
F_{13} = \frac{-3T_s}{sT_s + 3/4} \\
F_{23} = \frac{-1}{(sT_s)^3} 
\] (9)

and the cancellation logic transfer functions can be calculated using (7)–(9). Considering a nonreturn-to-zero (NRZ) DAC, the following are obtained:

\[
CL_1 = \frac{z^{-1}}{48} (7 + 25z^{-1} - 7z^{-2} - 5z^{-3}) \\
CL_2 = \frac{z^{-1}}{2} (1 + z^{-1})(1 - z^{-1})^2 \\
CL_3 = 2(1 - z^{-1})^3
\] (10)

where \(CL_3\) is chosen to have three zeroes at dc, corresponding to the zeroes contributed by the first three integrators.

To compare the robustness of both modulators in Fig. 2, the effect of mismatch on the signal-to-noise ratio (SNR) was simulated using SIMSIDES—a SIMULINK-based time-domain behavioral simulator for ΣΔMs [15]. For this purpose, maximum values of mismatch were estimated for a 0.18-μm CMOS technology. On the one hand, capacitor mismatch \((\sigma_C)\) lower than 1% are standard in modern technologies. On the other hand, MOS transconductance mismatch \((\sigma_{gm})\) is related to drain current mismatch \((\sigma_{Id})\) by the following expression:

\[
2\sigma_{gm}/\sigma_{gm} = \sigma_{Id}/\dot{I}_d.
\] (11)

In the case of the 0.18-μm technology addressed, assuming typical transistor sizes and a maximum spacing between transistors of 1500 μm, a maximum \(\sigma_{Id}/\dot{I}_d \leq 2\%\) is obtained. Then, taking into account that in a linearized transconductor, a minimum of 4 transconductance elements are used and that the tuning circuit introduces an additional error, maximum values of \(\sigma_{gm}/\sigma_{gm} \leq 2.5\%\) are expected.

Both modulators in Fig. 2 were simulated considering a \(\dot{G}_m-C\) implementation. The results are shown in Fig. 3, where the SNR loss is represented as a function of \(\sigma_{gm}\) and \(\sigma_C\). For each point of these surfaces, 150 simulations were carried out using random variations with the corresponding standard deviations. The value of the SNR loss represented in Fig. 3 stands for the difference between the ideal SNR, i.e., with no parameter variation, and the SNR with 90% of the 150 simulations above it. It is shown that the lower analog component count in Fig. 2(b) is reflected in a lower variance of the modulator coefficients, leading to better behavior in terms of sensitivity to mismatch.

The same reasoning can be applied to the requirements in terms of dc gain of the transconductors. Since the number of transconductors connected to the same node is higher in the previous method, the equivalent impedance associated with these nodes tends to be lower and the requirement in terms of the individual transconductor output impedance is higher. Therefore, a higher dc gain is needed. Fig. 4 illustrates this point. Note that the SNR peak starts dropping at a dc gain \(\geq 5\) dB higher in the case of the system designed following the previous method. For high dc gains, the two curves collide because the effect of this error becomes negligible.

IV. SYNTHESIS EXAMPLE

As an application of the proposed methodology, the 2-1-1-1 cascaded CT-ΣΔM shown in Fig. 5 was synthesized to achieve 12-bit resolution within a 20-MHz signal bandwidth. In order to fulfill these specifications without being limited by the clock jitter error, the sampling frequency, \(\dot{f}_s\), and the number of bits of
the internal quantizers (and DACs), $B$, must be properly chosen.
In the case of a fifth-order modulator like the one shown in
Fig. 5, the in-band jitter noise power is minimized for $f_n = 240$ MHz and $B = 4$ [16].

Another critical source of error in CT $\Sigma$ΔMs is the excess loop delay. As shown in [17], this error can be compensated for by adding an extra feedback branch between the output and the input to the quantizer (DAC$_2$ in Fig. 5) and two D-latches. By adding this extra branch with the appropriate gain, the loop impulse response is exactly the same as that of the original. This extra feedback term can be easily included in the calculation of the cancellation logic by modifying appropriately the $F_{ij}$ transfer functions. In practice, this extra feedback branch will only affect the cancellation logic if the input to other stages is tapped at quantizer inputs or if the last stage includes a second DAC. In a practical implementation it might be advantageous to make DAC$_2$ programmable [5].

Considering the factors above, the 2-1-1-1 architecture in Fig. 5 was synthesized using the methodology described in Section III. The first stage is formed by a resonator which has its pole placed at $\omega_p = 2\pi \sqrt{1/9B_w}$, minimizing the quantization noise transfer function in the signal bandwidth, $B_w$. Resistor variations can be tuned out using a combination of a discrete rough tuning of the resistors ($R_{\text{d}}$, $R_{\text{fb}}$ and $R_v$) and a continuous fine tuning of the transconductors $k_{ff}$ and $k_{g1}$. This tuning can be also used to cancel the effect of finite gain-bandwidth product (GB) of the front-end opamp, due to the fact that this error can be modelled as an integrator gain error [5]. All the other transconductors could be tuned in order to keep the time constant $C/g_m$ unchanged over $C$ variations. In this case, the single-loop and inter-stage transfer functions are

$$F_{11} = -\frac{b_{11}s + b_{10}}{s^2 + \omega_p^2}$$

$$F_{ii} = \frac{-1}{T_i s^i}, \quad i = 2, 3, 4$$

$$F_{14} = -\frac{b_{10}}{(T_4 s)^3 (s^2 + \omega_p^2)}$$

$$F_{24} = -\frac{1}{(T_4 s)^2}$$

$$F_{34} = -\frac{b_{10}}{(T_4 s)^2}$$

(12)

where

$$\omega_p = \sqrt{\frac{k_{g1}}{R_v C_1 C_2}}$$

(13)

and coefficients $b_{11}$, $b_{10}$, and $b_{10}$ are given by

$$b_{11} = \frac{k_{ff}}{R_{fb} C_1}; \quad b_{10} = \frac{k_{ff} k_{g2}}{R_{fb} C_1 C_2}; \quad b_{10} = \frac{k_{g1}}{R_{fb} C_1 C_2}$$

(14)

These coefficients are found through an iterative simulation-based process that optimizes the first stage of the modulator in order to maximize the SNR while maintaining stability. The outcome of this process—done entirely in the CT domain—is summarized in Table I. This table includes the values of loop filter coefficients, $k_i$ (implemented as transconductances) as well as the capacitors, $C_i$, and resistances, $R_i$ obtained from the optimization process. The expressions of $CL_4$, can be obtained from (7) and (12), giving

$$CL_4 = -z^{-1}(n_{14} + n_{13}z^{-1} + n_{12}z^{-2})$$

$$+ n_{11}z^{-3} + n_{10}z^{-4})$$

$$CL_2 = \frac{1}{6}z^{-1}(1 + 4z^{-1} + z^{-2})$$

$$\times (1 - 2\cos(T_{\omega p})z^{-1} + z^{-2})$$

$$CL_3 = \frac{1}{2}z^{-1}(1 + z^{-1})(1 - z^{-1})$$

$$\times (1 - 2\cos(T_{\omega p})z^{-1} + z^{-2})$$

$$CL_4 = (1 - z^{-1})^2 (1 - 2\cos(T_{\omega p})z^{-1} + z^{-2})$$

(15)
where coefficients $n_{ij}$ are given by

$$
\begin{align*}
    n_{10} &= n_{14} = -\frac{2H_1}{T_{\Delta p p}} \left[ \sin(T_{\Delta p p}) - T_{\Delta p p} + \frac{1}{3} (T_{\Delta p p})^3 \right] \\
    n_{11} &= n_{13} = -\frac{2H_1}{T_{\Delta p p}} \left[ \left( T_{\Delta p p} \right)^2 \cos(T_{\Delta p p}) - \frac{1}{3} \sin(T_{\Delta p p}) \right] \\
    n_{12} &= \frac{2H_1}{T_{\Delta p p}} \left[ \left( T_{\Delta p p} \right)^3 \frac{1}{3} - 4 \cos(T_{\Delta p p}) + 6 \sin(T_{\Delta p p}) \right]
\end{align*}
$$

(16)

The modulator in Fig. 5 was simulated using SIMSIDES, taking into account the non-ideal effects listed in Table II. Fig. 6 shows an output spectrum for a $-8$ dBV@6-MHz sinewave signal, demonstrating a correct operation within the signal band-width. The effective resolution is 76 dB (12.4 bits) as shown in Fig. 7, where the signal-to-noise-plus-distortion ratio (SNDR) is plotted as a function of the input signal amplitude.

V. Conclusion

This brief presents a new methodology for synthesizing cascaded continuous-time $\Sigma\Delta$Ms. The resulting circuits are less complex and more robust than those obtained with conventional methodologies. Examples supported by realistic behavioral simulations are given to validate the presented approach.

REFERENCES


