Excess Loop Delay in Continuous-Time Delta–Sigma Modulators

James A. Cherry, *Student Member, IEEE*, and W. Martin Snelgrove, *Member, IEEE*

Abstract— **Continuous-time (CT) delta–sigma modulators ($\Delta\Sigma$M's) suffer from a problem not seen in discrete-time (DT) designs, that of *excess loop delay*: nonzero delay between the quantizer clock edge and the time when a change in output bit is seen at the feedback point in the modulator. This paper analytically shows how such delay affects the equivalence between the CT modulator loop filter and its DT counterpart. The effect of this delay on modulator dynamic range is studied through simulation for the standard double-integration (low pass) CT modulator and its equivalent fourth-order $f_s/4$ band pass circuit. For the first time, the results are extended to higher order low-pass and bandpass designs, as well as multibit designs. Methods for alleviating the performance loss caused by excess loop delay are also discussed.**

Index Terms— **Delta–sigma modulation.**

I. INTRODUCTION

DELTA–SIGMA modulators ($\Delta\Sigma$M's) [1], [2] are popular nowadays for analog-to-digital conversion applications. $\Delta\Sigma$M's are usually thought of mathematically in the discrete-time (DT) domain, which means the majority of published designs are built using DT, e.g., switched-capacitor [3] or switched-current [4] circuitry. There is increasing interest in building $\Delta\Sigma$M's using continuous-time (CT) circuitry for the loop filter [5], because it is generally possible to clock CT $\Delta\Sigma$M's at much higher frequencies than DT $\Delta\Sigma$M's. Then, for a given oversampling ratio (OSR), the conversion bandwidth is greatly increased.

In recent years, several published CT $\Delta\Sigma$M circuits have appeared in the literature with clock speeds from over 100 MHz up to a few gigahertz. Table I shows the order and type of some recent designs where "1LP" means first-order low pass, "2BP" means second-order bandpass, etc. The majority of the designs are implementations of the standard double-integration modulator first popularized in [17]. The four bandpass (BP) modulators are for converting analog signals at one quarter of the sampling frequency to digital; ideally, they have the same performance and stability as a low pass (LP) design of half the order. Thus, all the high-speed designs listed are first- or second-order. For each clock rate f_s and OSR, the dynamic range (DR) and maximum signal-to-noise ratio (SNR) achieved are also listed. The performance of an ideal first- or second-order modulator can be found from DT simulation or
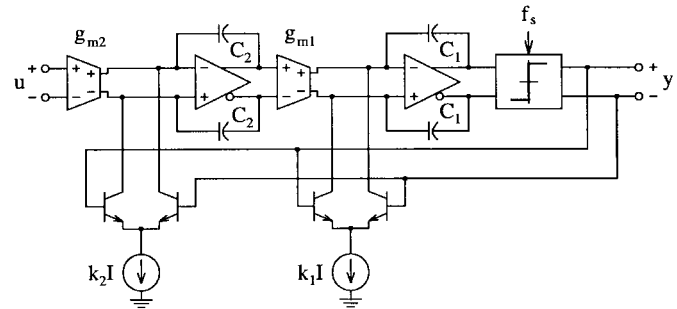
Fig. 1. Example of high-speed double-integration CT $\Delta\Sigma$M.

from [18, Fig. 7]; the last two columns of the table show how the achieved DR and SNR fare compared to ideal. Generally, we see performance falling far short of ideal, particularly for OSR's of 64 or more.

For high-speed CT designs, it is known that performance can be limited by several things. The smallest input signal that can be converted is ideally determined by the magnitude of the noise-shaped in-band quantization noise, but for high-speed designs, thermal noise in the input stage [19], quantizer clock jitter [19], and quantizer metastability [20] can all fill the noise notch with white noise and further limit the minimum convertible input signal. These problems are exacerbated as clock rates and conversion bandwidths become ever higher. Large input signals eventually overwhelm the linearity of the input stage, and any nonlinearity in the input stage appears directly in the modulator output spectrum [10].

There is yet another problem that affects the resolution of very high-speed designs. Consider the high-speed double-integration modulator from [10], depicted in Fig. 1. There are two integrator blocks with voltage inputs and outputs, each consisting of a transconductor for voltage-to-current conversion and an integrator for current-to-voltage conversion. The quantizer is a latched comparator whose output drives differential pair digital-to-analog converters (DAC's); their output currents sum with the transconductor outputs. Thus, the feedback necessary for $\Delta\Sigma$M operation works via Kirchhoff's current law (KCL). Ideally, the DAC currents respond immediately to the quantizer clock edge, but in practice, the transistors in the latch and the DAC have a nonzero switching time. Thus, there exists a delay between the quantizer clock and DAC current pulse, and we call this delay *excess loop delay*, or simply *excess delay* or *loop delay*.

Excess delay has been studied in the literature before; a brief summary of past work is appropriate. Gosslau and Gottwald

TABLE I
HIGH-SPEED CT $\Delta\Sigma$M PUBLISHED PERFORMANCE

Paper	Technology	Type	f_s	OSR	DR (dB)	SNR (dB)	DR loss (dB)	SNR loss (dB)
[6]	1μm GaAs D MESFET	2LP	200MHz	100	58	50	29	30
[7]	2μm CMOS	2LP	150MHz	128	63	57	29	28
[8]	0.5μm GaAs HEMT	2LP	500MHz	100	60	55	27	25
[9]	1.4μm GaAs HBT	2LP	2GHz	20	43	37	9	8
[10]	2μm InP DHBT	2LP	3.2GHz	32	49	50	13	5
[5]	0.8μm Si BiCMOS	2BP	200MHz	500	50	46	32	36
[11]	Si BJT	2LP	1.28GHz	64	–	45	–	25
[12]	2μm InP DHBT	2LP/BP	4GHz	64	44	41	33	29
[13]	0.5μm Si BJT	2BP	3.8GHz	10000	60?	49	60?	71
[14]	AlGaAs/GaAs HBT	4BP	3.2GHz	64	42	41	35	29
[15]	0.4μm InGaP/InGaAs HEMT	2LP	5GHz	50	51	39	21	26
[16]	0.5μm SiGe HBT	4BP	4GHz	500	62	53	59	61

[21], [22] found that excess delay of 25% actually improves the DR of a 1LP CT $\Delta\Sigma$M, compared to no excess delay. Horbach [23] confirmed this and extended the results to higher order LP modulators, showing that excess delay is detrimental to their performance. Chan [6] found that a full sample of feedback delay in his 2LP modulator caused 10 dB of SNR loss. Shoaei [5] found excess delay problematic in 2BP and 4BP modulators. Gao *et al.* [24] propose feedback coefficient tuning, and demonstrate that it alleviates delay problems in a 4BP modulator, while Benabes *et al.* [25] add an extra feedback loop to a 2LP modulator for the same purpose. One of the aims of the present paper is to unify and summarize the past work in the area, but we also contribute new material. First, most authors use the modified $\mathcal{Z}$-transform for studying excess delay, but we explain here why this is inappropriate and demonstrate a preferred method. Second, we consider higher order LP and BP modulators in much more detail than has previously appeared. We also consider multibit modulators, something which seems not to have been done in the past.

The remainder of this paper is organized as follows. Section II illustrates the mathematical equivalence between a CT modulator and a DT counterpart, and shows what excess loop delay does to this equivalence. Section III demonstrates through simulation how the in-band noise (IBN), maximum stable amplitude (MSA), and DR of the double-integration modulator are affected, while Section IV does the same thing for the fourth-order $f_s/4$ bandpass modulator. In Section V, we see what happens when excess loop delay occurs in LP modulators of orders three through five, as well as in a sixth-order $f_s/4$ BP modulator. Section VI briefly studies what happens in a $\Delta\Sigma$M with a multibit quantizer instead of the more traditional single-bit quantizer. Section VII talks about the various methods of compensating for loop delay, including DAC pulse selection, feedback coefficient tuning, and the inclusion of additional feedbacks. Finally, Section VIII draws some conclusions about the work.

II. PRELIMINARIES

A general CT $\Delta\Sigma$M is depicted in Fig. 2. The CT input $\hat{u}(t)$ [possibly prefiltered by $\hat{G}(s)$] is applied to a modulator with a CT *loop filter* $\hat{H}(s)$ whose output we denote $\hat{x}(t)$. The quantizer samples this signal at frequency f_s, or equivalently
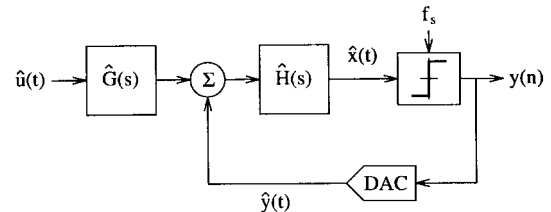

Fig. 2.   General CT $\Delta\Sigma$M block diagram.

with period T_s; this produces a DT output signal $y(n) = \hat{y}(nT_s)$, which is fed back through a DAC.
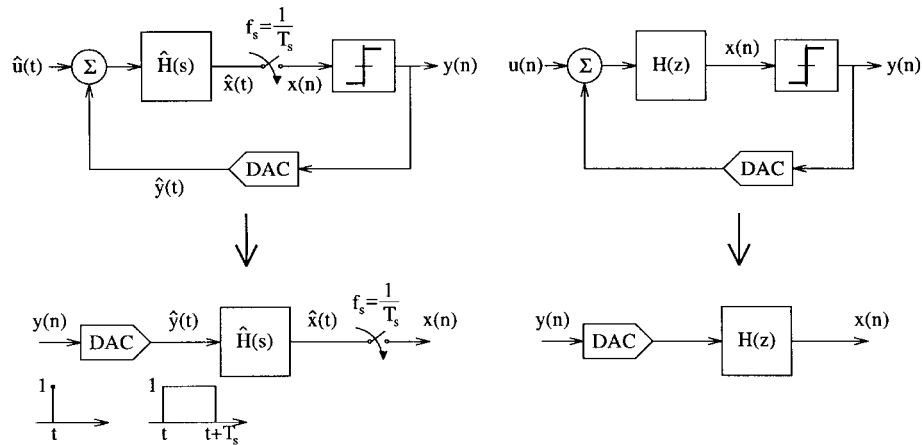
A. CT/DT Modulator Equivalence

It is useful to begin by explaining how to find the equivalent DT loop filter $H(z)$ for a given CT loop filter $\hat{H}(s)$. Why does such an equivalent exist? Because the quantizer in a CT $\Delta\Sigma$M is clocked, which means there is an *implicit* sampling action inside the modulator, and sampled circuits are DT circuits. We can make the sampling explicit by placing the sampler immediately prior to the quantizer, as depicted in the upper left diagram of Fig. 3; this does not change the behavior of the modulator. If we want to know how this is equivalent to a DT modulator, shown in the upper right of Fig. 3, then it is illustrative to zero both inputs and open both loops around the quantizer. This leads to the bottom two diagrams of Fig. 3.

In the CT open-loop diagram, the quantizer output $y(n)$ is a DT quantity, and we may think of the DAC as a "discrete-to-continuous converter." It makes a CT pulse $\hat{y}(t)$ from the output sample $y(n)$. This pulse is filtered by $\hat{H}(s)$ (the CT loop filter) to produce $\hat{x}(t)$ at the quantizer input, which is then sampled to produce the DT quantizer input $x(n)$. The input and output of both the CT and DT open-loop diagrams are thus DT quantities. A CT modulator would produce the same sequence of output bits $y(n)$ as a DT modulator if the inputs to the quantizer in each were identical at the following sampling instants:

$$x(n) = \hat{x}(t)|_{t=nT_s}. \tag{1}$$

This would be satisfied if the impulse responses of the open-loop diagrams in Fig. 3 were equal at sampling times, leading to the condition [26]

$$\mathcal{Z}^{-1}\{H(z)\} = \mathcal{L}^{-1}\{\hat{R}_D(s)\hat{H}(s)\}|_{t=nT_s} \tag{2}$$

Fig. 3.   Open-loop CT $\Delta\Sigma$M and its DT equivalent.

TABLE II
s-DOMAIN EQUIVALENCES FOR z-DOMAIN LOOP FILTER POLES

z-domain pole	s-domain equivalent	Limit for $z_k = 1$
$\frac{1}{z-z_k}$	$\frac{r_0}{s-s_k} \times \frac{1}{z_k^{1-\alpha}-z_k^{1-\beta}}$ $r_0 = s_k$	$\frac{r_0}{s-s_k}$ $r_0 = \frac{1}{\beta-\alpha}$
$\frac{1}{(z-z_k)^2}$	$\frac{r_1 s+r_0}{(s-s_k)^2} \times \frac{1}{z_k(z_k^{1-\alpha}-z_k^{1-\beta})^2}$ $r_1 = q_1 s_k + q_0$ $r_0 = q_1 s_k^2$ $q_1 = z_k^{1-\beta}(1-\beta) - z_k^{1-\alpha}(1-\alpha)$ $q_0 = z_k^{1-\alpha} - z_k^{1-\beta}$	$\frac{r_1 s+r_0}{(s-s_k)^2}$ $r_1 = \frac{1}{2}\frac{\alpha+\beta-2}{\beta-\alpha}$ $r_0 = \frac{1}{\beta-\alpha}$
$\frac{1}{(z-z_k)^3}$	$\frac{r_2 s^2+r_1 s+r_0}{(s-s_k)^3} \times \frac{1}{z_k^2(z_k^{1-\alpha}-z_k^{1-\beta})^3}$ $r_2 = \frac{1}{2}q_2 s_k - q_1$ $r_1 = -q_2 s_k^2 + q_1 s_k + q_0$ $r_0 = \frac{1}{2}q_2 s_k^3$ $q_2 = (1-\beta)(2-\beta)(z_k^{1-\beta})^2$ $\quad + (1-\alpha)(2-\alpha)(z_k^{1-\alpha})^2$ $\quad + [\beta(\beta+3) + \alpha(\alpha+3)$ $\quad - 4(1+\alpha\beta)]z_k^{1-\alpha}z_k^{1-\beta}$ $q_1 = (\frac{3}{2}-\beta)(z_k^{1-\beta})^2 + (\frac{3}{2}-\alpha)(z_k^{1-\alpha})^2$ $\quad + (\alpha+\beta-3)z_k^{1-\alpha}z_k^{1-\beta}$ $q_0 = (z_k^{1-\alpha} - z_k^{1-\beta})^2$	$\frac{r_2 s^2+r_1 s+r_0}{(s-s_k)^3}$ $r_2 = \frac{1}{12}\frac{1}{\beta-\alpha}[\beta(\beta-9)$ $\quad + \alpha(\alpha-9) + 4\alpha\beta + 12]$ $r_1 = \frac{1}{2}\frac{\alpha+\beta-3}{\beta-\alpha}$ $r_0 = \frac{1}{\beta-\alpha}$

or, in the time domain [5]

$$h(n) = [\hat{r}_D(t) * \hat{h}(t)]|_{t=nT_s} = \int_{-\infty}^{\infty} \hat{r}_D(\tau)\hat{h}(t-\tau)\, d\tau|_{t=nT_s} \tag{3}$$

where $\hat{r}_D(t)$ is the impulse response of the DAC. Since we are requiring the CT and DT impulse responses to be the same, the transformation between the two is called the *impulse-invariant transformation* [27].

Without loss of generality, we shall simplify the discussion by assuming a sampling period of $T_s = 1$ for the remainder of this paper.

B. Usefulness of Equivalence

Knowledge of the equivalence allows us to perform CT $\Delta\Sigma$M loop filter design in the DT domain using any design technique we choose, for example, noise-transfer function

(NTF) prototyping [2, Ch. 4]. Once we have chosen $H(z)$, we may find the $\hat{H}(s)$ to implement the CT modulator with identical behavior, given a certain type of DAC pulse. For simplicity, we assume a perfectly rectangular DAC pulse of magnitude 1 that lasts from β to β, i.e.,

$$\hat{r}_{(\alpha,\beta)}(t) = \begin{cases} 1, & \alpha \le t < \beta, 0 \le \alpha < \beta \le 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Table II lists the s-domain equivalents for z-domain $H(z)$ poles of orders one through three. These were found by solving (2) in the symbolic math program Maple [28] where the Laplace transform of (4) is

$$\hat{R}_{(\alpha,\beta)}(s) = \frac{\exp(-\alpha s) - \exp(-\beta s)}{s}. \tag{5}$$

It is found that a z-domain pole of multiplicity l at z_k maps to one at s_k with the same multiplicity, with

$$s_k = \ln z_k. \tag{6}$$

Therefore, to use the table, $H(z)$ is first written as a partial fraction expansion, then we apply the transformations in the table to each term and recombine them to get the equivalent $\hat{H}(s)$. Poles at dc (i.e., $z_k = 1$) end up giving $0^l/0^l$ as the numerator of the s-domain equivalent, which necessitates l applications of l'Hopital's rule; this has been done in the right column of Table II.

Let us illustrate the process. Many designs use DAC's with an output pulse which remains constant over a full period, which we shall term a *nonreturn-to-zero* (NRZ) DAC. For this type of DAC, $(\alpha, \beta) = (0, 1)$ in (4). Moreover, we saw that many of the high-speed designs in Table I were second-order LP designs; these differentiate the quantization noise twice so that $\text{NTF}(z) = (z - 1)^2$ and

$$H(z) = \frac{-2z + 1}{(z - 1)^2}. \tag{7}$$

Writing this in partial fractions yields

$$H(z) = \frac{-2}{z - 1} + \frac{-1}{(z - 1)^2}. \tag{8}$$

Thus $z_k = 1$, which means $s_k = 0$ from (6). Applying the first row of Table II to the first term of (8) and the second row to the second term with $(\alpha, \beta) = (0, 1)$ gives

$$\hat{H}(s) = \frac{-2}{s} + \frac{-1 + 0.5s}{s^2} \tag{9}$$

$$= -\frac{1 + 1.5s}{s^2}. \tag{10}$$

Equation (10) was first derived by Candy [17] as the CT equivalent of the DT double-integration modulator in (7).

A previous paper on CT $\Delta\Sigma$M design by Schreier [29] used state-space representation for modulator equivalence calculations; we choose to use pole–zero representation here, though either method works. We have only been dealing with loop filter equivalence, which affects the noise transfer function in the linearized $\Delta\Sigma$M model; there are some subtleties regarding the signal transfer function [5], [29] which we simplify by assuming a signal transfer function of one in the band of interest. This assumption is approximately valid for most designs.

C. Effect of Excess Loop Delay

As noted in Section I, excess loop delay arises because of nonzero transistor switching time, which makes the edge of the DAC pulse begin *after* the sampling clock edge. We assume that excess loop delay can be expressed by

$$\tau_d = \rho_d T_s \tag{11}$$

which is depicted for an NRZ DAC pulse in Fig. 4.[1] The sampling instant is $t = 0$. The value of τ_d depends on the switching speed of the transistors f_T, the quantizer clock frequency f_s, and the number of transistors in the feedback

[1] A previous paper [25] treats DAC pulses as having delay *plus* a nonzero rise time (which can be either exponential or slewing in behavior). This is more realistic for an actual circuit, but we choose to use rectangular pulses for three reasons: it is mathematically simpler; the general results presented here still hold with nonzero rise time DAC pulses; and the compensation schemes presented in Section VII apply equally to either case.
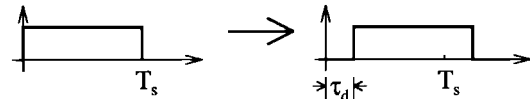

Fig. 4. Illustration of excess loop delay on NRZ DAC pulse.

path n_t, as well as the loading on each transistor. As a crude approximation, we may assume all transistors switch fully after $1/f_T$, in which case

$$\rho_d \approx \frac{n_t f_s}{f_T}. \tag{12}$$

τ_d could end up being a significant fraction of T_s depending on the parameters in (12). For example, in the design in Fig. 1, suppose we desire 12-bit DR in a 50-MHz bandwidth. This will require an OSR of about 50 [18], which means we must clock at $f_s = 50(2 \cdot 50) = 5$ GHz. If the quantizer is an ECL-style latched comparator, its output differential pair must switch; the DAC must also switch, and thus $n_t = 2$. In a $f_T = 30$-GHz process, therefore, (12) predicts

$$\rho_d \approx \frac{2 \cdot 5}{30} = 33\%. \tag{13}$$

Excess loop delay is problematic because it alters α and β, which means it affects the equivalence between $\hat{H}(s)$ and $H(z)$. We can calculate the effect mathematically by using Table III, which lists the z-domain equivalents for s-domain $\hat{H}(s)$ poles of orders one through three. As with Table II, these were calculated with the help of Maple and (2). An s-domain pole of multiplicity l at s_k maps to one at z_k with the same multiplicity, with

$$z_k = \exp s_k. \tag{14}$$

Poles at $s_k = 0$ give numerators of $0^l/0^l$, as before, and the rightmost column gives the formulas that result when l'Hopital's rule is applied l times.

Let us assume that we have designed $\hat{H}(s)$ from (10) assuming NRZ DAC pulses, but that we have excess loop delay τ_d, so that in actuality we have NRZ DAC pulses delayed by τ_d as in Fig. 4. Now, we have $(\alpha, \beta) = (\tau_d, 1 + \tau_d)$. The formulae in Table III only apply for a pulse with $\beta \leq 1$, but we need not worry; it is possible to write a τ_d-delayed NRZ pulse as

$$\hat{r}_{(\tau_d, 1+\tau_d)}(t) = \hat{r}_{(\tau_d, 1)}(t) + \hat{r}_{(0, \tau_d)}(t - 1) \tag{15}$$

that is, as a linear combination of a DAC pulse from τ_d to 1 *and* a one-sample-delayed DAC pulse from 0 to τ_d. Writing (10) in partial fractions gives

$$\hat{H}(s) = \frac{-1.5}{s} + \frac{-1}{s^2}. \tag{16}$$

Applying Table III to each term of (16), for each of the two DAC pulses in (15), yields

$$\frac{-1.5}{s} \rightarrow \frac{-1.5(1 - \tau_d)}{z - 1} + z^{-1} \frac{-1.5\tau_d}{z - 1} \tag{17}$$

$$\frac{-1}{s^2} \rightarrow \frac{(-0.5 + \tau_d - 0.5\tau_d^2)z + 0.5(-1 + \tau_d^2)}{(z - 1)^2}$$

$$+ z^{-1} \frac{\tau_d(-1 + 0.5\tau_d)z - 0.5\tau_d^2}{(z - 1)^2}. \tag{18}$$

TABLE III
z-Domain Equivalences for s-Domain Loop Filter Poles

s-domain pole	z-domain equivalent	Limit for $s_k = 0$
$\frac{1}{s-s_k}$	$\frac{y_0}{z-z_k} \times \frac{1}{s_k}$ $y_0 = z_k^{1-\alpha} - z_k^{1-\beta}$	$\frac{y_0}{z-z_k}$ $y_0 = \beta - \alpha$
$\frac{1}{(s-s_k)^2}$	$\frac{y_1 z + y_0}{(z-z_k)^2} \times \frac{1}{s_k^2}$ $\begin{aligned} y_1 &= z_k^{1-\beta}[1 - s_k(1-\beta)] \\ &\quad - z_k^{1-\alpha}[1 - s_k(1-\alpha)] \\ y_0 &= z_k^{2-\alpha}(1 + s_k\alpha) \\ &\quad - z_k^{2-\beta}(1 + s_k\beta) \end{aligned}$	$\frac{y_1 z + y_0}{(z-z_k)^2}$ $\begin{aligned} y_1 &= \tfrac{1}{2}[\beta(2-\beta) - \alpha(2-\alpha)] \\ y_0 &= \tfrac{1}{2}(\beta^2 - \alpha^2) \end{aligned}$
$\frac{1}{(s-s_k)^3}$	$\frac{y_2 z^2 + y_1 z + y_0}{(z-z_k)^3} \times \frac{1}{s_k^3}$ $\begin{aligned} y_2 &= z_k^{1-\beta}[-1 + s_k(1-\beta) + \tfrac{s_k^2}{2}(1-\beta)^2] \\ &\quad - z_k^{1-\alpha}[-1 + s_k(1-\alpha) + \tfrac{s_k^2}{2}(1-\alpha)^2] \\ y_1 &= z_k^{2-\beta}[2 - s_k(1-2\beta) \\ &\quad + \tfrac{s_k^2}{2}(-1 - 2\beta + 2\beta^2)] \\ &\quad + z_k^{2-\alpha}[2 - s_k(1-2\alpha) \\ &\quad + \tfrac{s_k^2}{2}(-1 - 2\alpha + 2\alpha^2)] \\ y_0 &= z_k^{3-\alpha}(1 + s_k\alpha + \tfrac{s_k^2}{2}\alpha^2) \\ &\quad - z_k^{3-\beta}(1 + s_k\beta + \tfrac{s_k^2}{2}\beta^2) \end{aligned}$	$\frac{y_2 z^2 + y_1 z + y_0}{(z-z_k)^3}$ $\begin{aligned} y_2 &= \tfrac{1}{6}(\beta^3 - \alpha^3) \\ &\quad - \tfrac{1}{2}(\beta^2 - \alpha^2) + \tfrac{1}{2}(\beta - \alpha) \\ y_1 &= \tfrac{1}{3}(\beta^3 - \alpha^3) \\ &\quad - \tfrac{1}{2}(\beta^2 - \alpha^2) - \tfrac{1}{2}(\beta - \alpha) \\ y_0 &= \tfrac{1}{6}(\beta^3 - \alpha^3) \end{aligned}$

Adding (17) and (18) gives (19), shown at the bottom of the next page. If instead of Table III we use the modified $\mathcal{Z}$-transform on (7), the result is [24]

$$H(z, \tau_d) = \mathcal{Z}_m\left[\mathcal{Z}^{-1}\left[\frac{-2z+1}{(z-1)^2}\right]\right]$$
$$= \frac{(-2 + 2\tau_d)z^2 + (1 - 3\tau_d)z + \tau_d}{z(z-1)^2} \qquad (20)$$

which is similar to (19) but not identical. The modified $\mathcal{Z}$-transform assumes the delay happens at the output of $\hat{H}(s)$, but we assume the delay happens prior to the DAC pulse, which turns out to be different mathematically. Our assumption represents what happens in an actual circuit, hence we prefer our method over the modified $\mathcal{Z}$-transform.

We can quickly verify that for $\tau_d = 0$, (19) turns into (7) as it should. However, for $\tau_d \neq 0$, the equivalent $H(z)$ is no longer (7). How well does a modulator with a loop filter given by (19) perform?

III. Double-Integration Modulator

To study the effects of excess loop delay, Matlab [30] code was written to perform the transformations in Tables II and III numerically. The output bit-stream from a modulator was determined by evaluating the difference equation $X(z) = G(z)U(z) + H(z)Y(z)$ in the time domain, with a C program, for given $G(z)$ and $H(z)$. The virtue of using the transformations is it allows us to simulate in the DT domain, a process usually significantly more rapid than simulating using $\hat{H}(s)$ in the CT domain.[2]
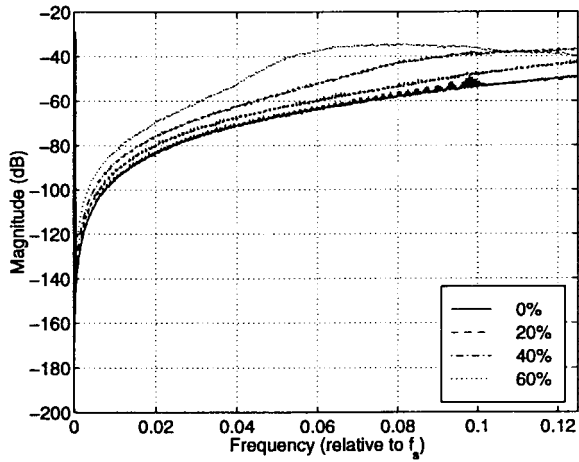
Since first-order modulators with excess delay have been studied already [21], we confine ourselves to modulators of orders two and above. In this section, we commence with the double-integration $\Delta\Sigma$M. We wish to know how its DR is affected by excess delay. DR is defined as the difference between the smallest and largest input levels (in decibels) which give SNR ≥ 0. At low input levels, SNR is limited by IBN, while a large-enough input level eventually compromises the stability of the modulator. There exists a maximum stable input amplitude (MSA); DR may be found from IBN and MSA, as we explain below.
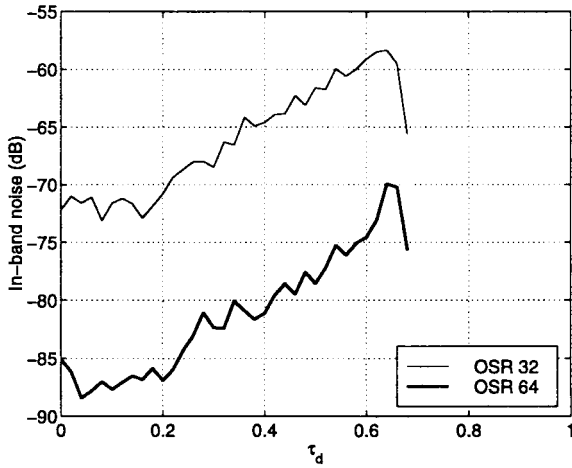
A. In-Band Noise

Fig. 5(a) shows an output spectrum near dc: there were 256 16 384-point Hann-windowed periodograms with random initial conditions averaged, and the input signal was a 0.1-V sinewave. As the delay increases from 0% up to 60%, we see that the noise floor rises slowly. Integrating the IBN for zero input, as a function of τ_d, produces Fig. 5(b). For delays below about 20%, IBN stays roughly constant, but rises as delay increases. If the excess delay exceeds about 65%, the modulator goes unstable. In this paper, instability is defined as the quantizer-input magnitude exceeding ten before the end of a simulation for 1000 successive simulations with random initial conditions. A similar definition was used in [31].

[2] Though we provide no experimental verification of the results throughout this paper, we find simulation of Fig. 1 in Eldo using ideal circuit components and a variable delay in the feedback path gives results that are consistent with those presented in this section. Simulations take much longer with Eldo, however.

$$H(z, \tau_d) = \frac{(-2 + 2.5\tau_d - 0.5\tau_d^2)z^2 + (1 - 4\tau_d + \tau_d^2)z + (1.5\tau_d - 0.5\tau_d^2)}{z(z-1)^2} \qquad (19)$$

Fig. 5. (a) Output spectrum from double-integration CT $\Delta\Sigma$M as a function of loop delay and (b) in-band noise for zero-input against loop delay.

Fig. 6. Maximum stable amplitude for double-integration CT $\Delta\Sigma$M.

Fig. 7. DR for double-integration CT $\Delta\Sigma$M.

The smallest input signal for which SNR = 0 dB is exactly the IBN, adjusted for the gain of the window (0.375 for Hann, or 4.26 dB) and the fact that periodograms measure rms power (3.01 dB). For example, the IBN for $\tau_d = 0$ and OSR = 64 is -85.06 dB, and we find in simulation that an input magnitude of approximately

$$-85.06 + 4.26 + 3.01 = -77.79 \text{ dB} \qquad (21)$$

is needed to get SNR = 0 dB.

B. Maximum Stable Amplitude

To determine the MSA, we once again follow [31]. We apply a ramp input whose amplitude increases slowly from zero to one over 10^5 time steps; when the quantizer-input magnitude exceeds ten, the input level at that instant is the MSA. We could just as well apply a low-frequency sinewave at the input and find the maximum amplitude of such an input for which simulating for many cycles keeps the modulator stable, but we find the method [31] gives approximately the same answer with much less simulation.

Performing this test for 200 runs with random initial conditions and averaging the MSA's so obtained yields the graph in
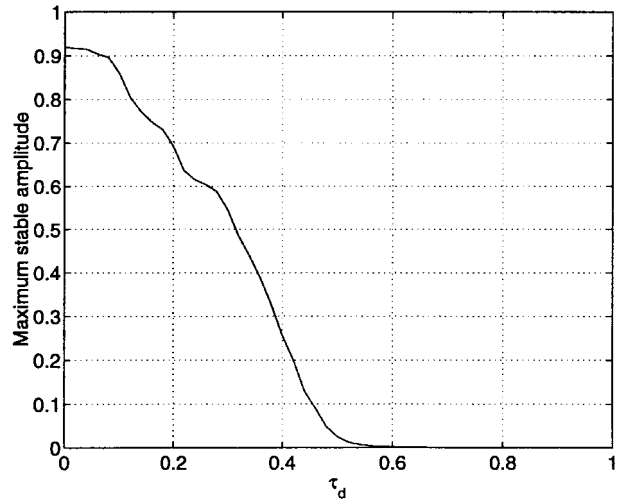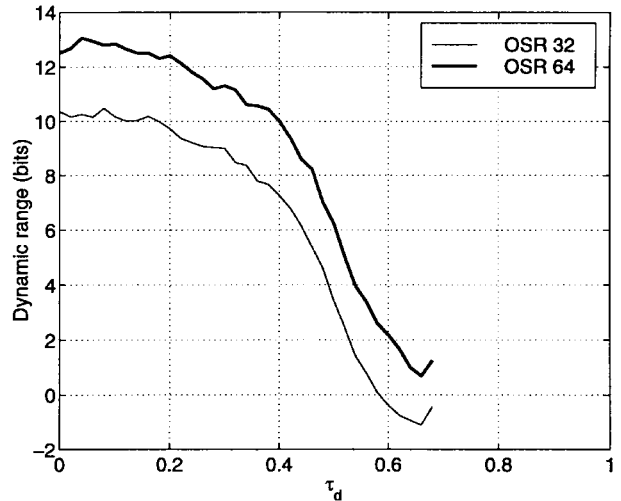
Fig. 6. The modulator is stable for inputs of up to 0.92 for no excess delay, but this falls more or less linearly to near zero at about 50% delay. An unstable modulator has SNR = $-\infty$, so the MSA is precisely the largest input for which SNR > 0. For example, at $\tau_d = 0$, the MSA is

$$20 \log_{10} 0.92 = -0.72 \text{ dB}. \qquad (22)$$

C. Dynamic Range

We can combine the previous two results to plot the modulator DR against delay. DR is exactly the difference between MSA and adjusted IBN; for example, at $\tau_d = 0$, (21) and (22) give
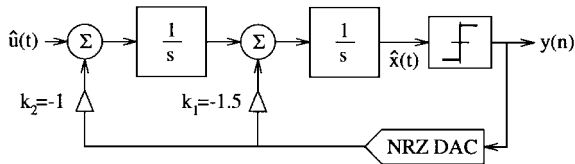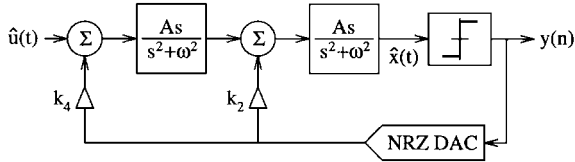
$$\text{DR} = -0.72 - (-77.79) = 77.07 \text{ dB}. \qquad (23)$$

This is converted to bits using [32]

$$\text{DR(bits)} = (\text{DR(dB)} - 1.76)/6.02 \qquad (24)$$

and the result is plotted for $0 \leq \tau_d \leq 1$ in Fig. 7.

This is useful as follows. Our example from earlier (12 bits at 50 MHz) estimated a loop delay of 33% in (13) for

Fig. 8. Block diagram for LP CT $\Delta\Sigma$M from Fig. 1.

Fig. 9. Block diagram for BP CT $\Delta\Sigma$M with integrators replaced by resonators that cannot implement desired equivalent $H(z)$.

OSR $= 50$. We see from the figure that even with OSR $= 64$, it would not be possible to achieve the desired resolution at 33% delay: We could only obtain DR $= 11$ bits. To achieve 12 bits at OSR $= 64$, we must have no more than about 20% excess loop delay. For a 50-MHz bandwidth, an OSR of 64 means clocking at 6.4 GHz, and from (12), we see that the transistors must have $f_T > 32$ GHz or so.

IV. $f_s/4$ FOURTH-ORDER BAND PASS MODULATOR

One type of $\Delta\Sigma$M that has found applications in radio circuits [33], [13] takes a low-pass NTF(z) with a quantization noise notch at dc and performs the substitution $z^{-1} \rightarrow -z^{-2}$. This gives a BP NTF$(z)$ with a noise notch at $f_s/4$, one quarter the sampling frequency [2, Ch. 9] with double the order and identical stability properties to the LP prototype. The substitution can be applied to the loop filter $H(z)$ to yield the same result. Applying this to the double-integration modulator (7) gives
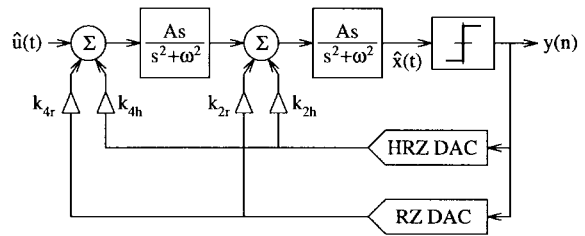
$$H_{\mathrm{LP}}(z) = \frac{-2z^{-1} + z^{-2}}{(1 - z^{-1})^2} \rightarrow H_{\mathrm{BP}}(z) = \frac{2z^{-2} + z^{-4}}{(1 + z^{-2})^2}. \quad (25)$$

This contains two double poles at $z_k = \pm j$; we could find the equivalent $\hat{H}_{\mathrm{BP}}(s)$ by applying the results in Table II to a partial fraction expansion of (22).

Doing this for NRZ DAC pulses yields

$$\hat{H}_{\mathrm{BP}}(s) = \frac{-1.0354s^3 + 1.0652s^2 - 1.3210s + 4.5661}{\left(s^2 + \left(\frac{\pi}{2}\right)^2\right)^2}. \quad (26)$$

How do we build a circuit to implement this? Historically, LP DT modulators have been built as a cascade of integrators $z^{-1}/(1 - z^{-1})$ [34], and building an $f_s/4$ BP DT modulator simply requires replacing the integrator blocks directly with resonator blocks $-z^{-2}/(1 + z^{-2})$. It is likewise possible to build LP CT modulators as a cascade of integrators $1/s$; the block diagram for Fig. 1 is shown in Fig. 8. However, simply replacing integrators with resonators $As/(s^2 + \omega^2)$, $\omega = \pi/2$ as in Fig. 9 does *not* build (26). The numerator of $\hat{H}_{\mathrm{BP}}(s)$ for Fig. 9 does not contain an s^2 or s^0 term, yet each is required in (26). Early designs [26] suffered from this problem.

Fig. 10. Multifeedback BP CT $\Delta\Sigma$M architecture.

One solution is to use resonators with a low-pass term included in the numerator: $(As + B)/(s^2 + \omega^2)$. A second elegant solution first proposed in [35] and [36] is to use resonators $As/(s^2 + \omega^2)$ with two different types of feedback DAC, leading to the so-called *multifeedback architecture* in Fig. 10. There, the DAC's are return-to-zero (RZ), which has $(\alpha, \beta) = (0, 0.5)$ in (4), and half-delayed RZ (HRZ) $(\alpha, \beta) = (0.5, 1)$. Both are easy to fabricate in an ECL-style latched comparator by diode-connecting the final differential pair rather than cross-coupling them [16], as shown in Fig. 11. We could have used any two of NRZ, RZ, and HRZ, or for that matter any other two different pulses, but these three types are easiest to build in a practical circuit. For an NRZ comparator, connect the final differential pair via the dashed lines; for RZ, connect the dotted lines instead.

The numerator of $\hat{H}_{\mathrm{BP}}(s)$ implemented in Fig. 10 can be set by altering the k coefficients. We wish to find how to set the k's so that the equivalent $H_{\mathrm{BP}}(z)$ is that in (25); this is done by converting $\hat{H}_{\mathrm{BP}}(s)$ to the z-domain using Table III for each DAC separately, then linearly combining the results and solving for the k's. The k values that implement (25) when the CT modulator uses RZ and HRZ DAC's can be calculated to be

$$\{k_{r4}, k_{r2}, k_{h4}, k_{h2}\}$$
$$= \{-1.086\,78, -2.133\,88, 0.450\,16, 1.487\,44\}. \quad (27)$$

How does excess delay affect this design? Both leading DAC edges become delayed by τ_d. Exactly the same simulations were carried out for this BP modulator as were done in the previous section (IBN and MSA), only instead of using a ramp input to find the MSA, a sine wave input at $f_s/4$ whose amplitude increases from zero to one over 10^5 time steps is used. Again, this method is rapid, and we find it gives similar results to using a sinewave input with fixed amplitudes and frequencies near $f_s/4$, simulating for many cycles to see if the modulator remains stable, then increasing the amplitude and repeating the simulation.

The resulting DR as a function of τ_d is plotted in Fig. 12; for comparison, the results from Fig. 7 for the double-integration modulator are overlaid with dashed lines. Interestingly, the two designs perform the same until about 30% excess delay, at which point the BP design becomes more severely affected. It goes unstable for about 50% excess delay. These results do not change if a different pair of DAC pulses are selected. Previous examinations of this modulator [5, Sec. 3.1.4], [24] which found 25% delay required for instability made two errors. First, the modified $\mathcal{Z}$-transform was used, which led to an incorrect
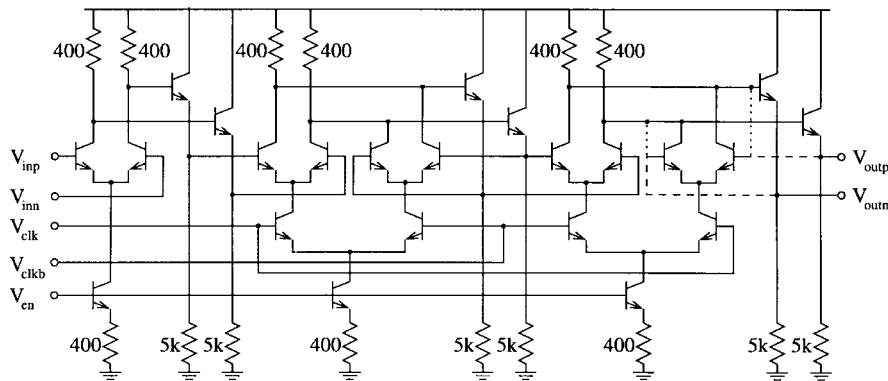
Fig. 11. ECL-style latched comparator with preamplification for enhanced resolution at high speed.

$H_{\mathrm{BP}} z \tau_d$. Second, simulations were carried out using a large fixed-amplitude input tone, which fails to take into account the changing modulator MSA with increasing delay.

V. HIGHER ORDER MODULATORS

We now turn to studying the effects of excess loop delay for low-pass CT $\Delta\Sigma$M's of order higher than two. The architecture we will consider is a generalization of Fig. 8 shown in Fig. 13; it is straightforwardly realizable in VLSI with transconductors, integrators, and DAC differential pairs as in Fig. 1. The loop filter realized by this architecture for $m \geq 2$ is

$$\hat{H}(s) = \frac{\sum_{i=1}^{m} \left(\frac{1}{s}\right)^i \left[k_i - \sum_{j=1}^{i-1} k_{i-j} B_j\right]}{1 - \sum_{i=1}^{m} B_i \left(\frac{1}{s}\right)^i}. \tag{28}$$

Equation (28) shows that the purpose of the B_i's is to allow us to implement NTF(z) zeros at places other than dc (i.e., $z = 1$).

Four types of high-order modulators were designed using NTF(z) prototyping. The NTF's used had:

1) third-order Butterworth poles, all zeros at $z = 1$;
2) third-order Butterworth poles, optimally-spread zeros;
3) fourth-order Butterworth poles, optimally-spread zeros;
4) fifth-order Chebyshev poles, optimally-spread zeros.

The spread-zero modulators had zeros placed according to [18] so that IBN would be minimized for a given OSR. Modulators with out-of-band gains (OOBG's) of 1.3, 1.4, 1.5, and 1.6 were all designed; recall that higher OOBG means lower IBN at the price of MSA [2, Ch. 4] The DR as a function of excess loop delay for NRZ DAC pulses and OSR's of both 32 and 64 are summarized in the graphs in Fig. 14.[3]

The results are most intriguing. The modulators with OOBG =1.3 remain stable, even for one full-sample excess delay, and moreover they only suffer a DR loss of between two and three bits. This contrasts starkly with the results for

[3] The nonmonotonicity in these and certain later DR graphs, for example in the tails of the curves in the upper right graph, is not a real effect: it is an artifact of doing simulations with zero input and no dither. Otherwise, the general trends indicated by the curves are accurate.
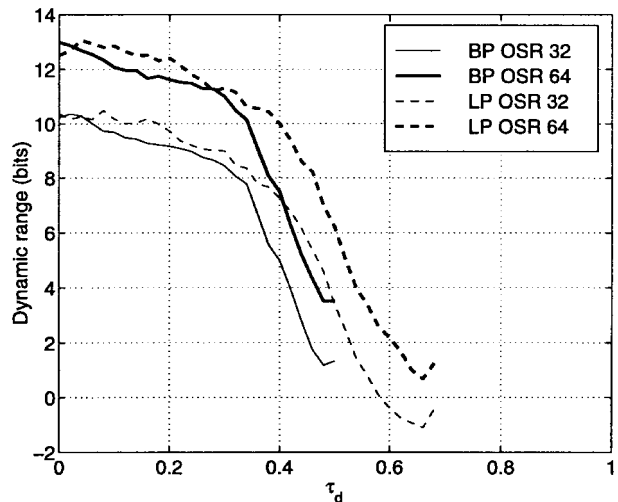

Fig. 12. DR for multifeedback BP CT $\Delta\Sigma$M, with comparison to double-integration results.

the second-order LP and fourth-order BP circuits. Increasing OOBG results in modulators which have generally better resolution at no delay, but which become unstable for less excess delay. This makes perfect sense; higher OOBG means a generally less-stable $H(z)$, and in fact, we see the τ_d needed for instability is roughly inversely proportional to OOBG. This suggests that higher order modulators enjoy an advantage over the lower order ones; the existence of a parameter OOBG which we may select according to our resolution *and* excess delay imperviousness requirements. To be fair, one *can* vary the OOBG in a second-order LP $\Delta\Sigma$M, but it is rarely done in practice.

For interest's sake, a sixth-order $f_s/4$ BP design was also tested by taking the low pass NTF with third-order Butterworth poles and three dc zeros and transforming it to a band pass design using $z^{-1} \rightarrow -z^{-2}$. This can be implemented using the multifeedback architecture in Fig. 10 with a third resonator and an additional feedback coefficient for each DAC. DR is plotted against τ_d in Fig. 15. Comparing these curves to those of the equivalent third-order LP design (the upper-left graph of Fig. 14) illustrates behavior like that in Fig. 12; the BP curves have the same shape as those of the LP curves for low excess delay, but they become unstable sooner as excess delay increases. Significantly, the LP modulator with OOBG =1.3
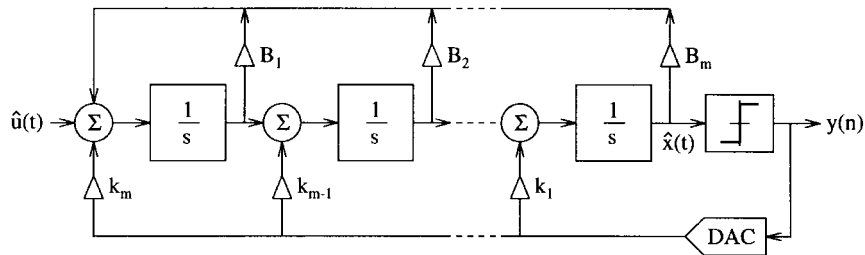
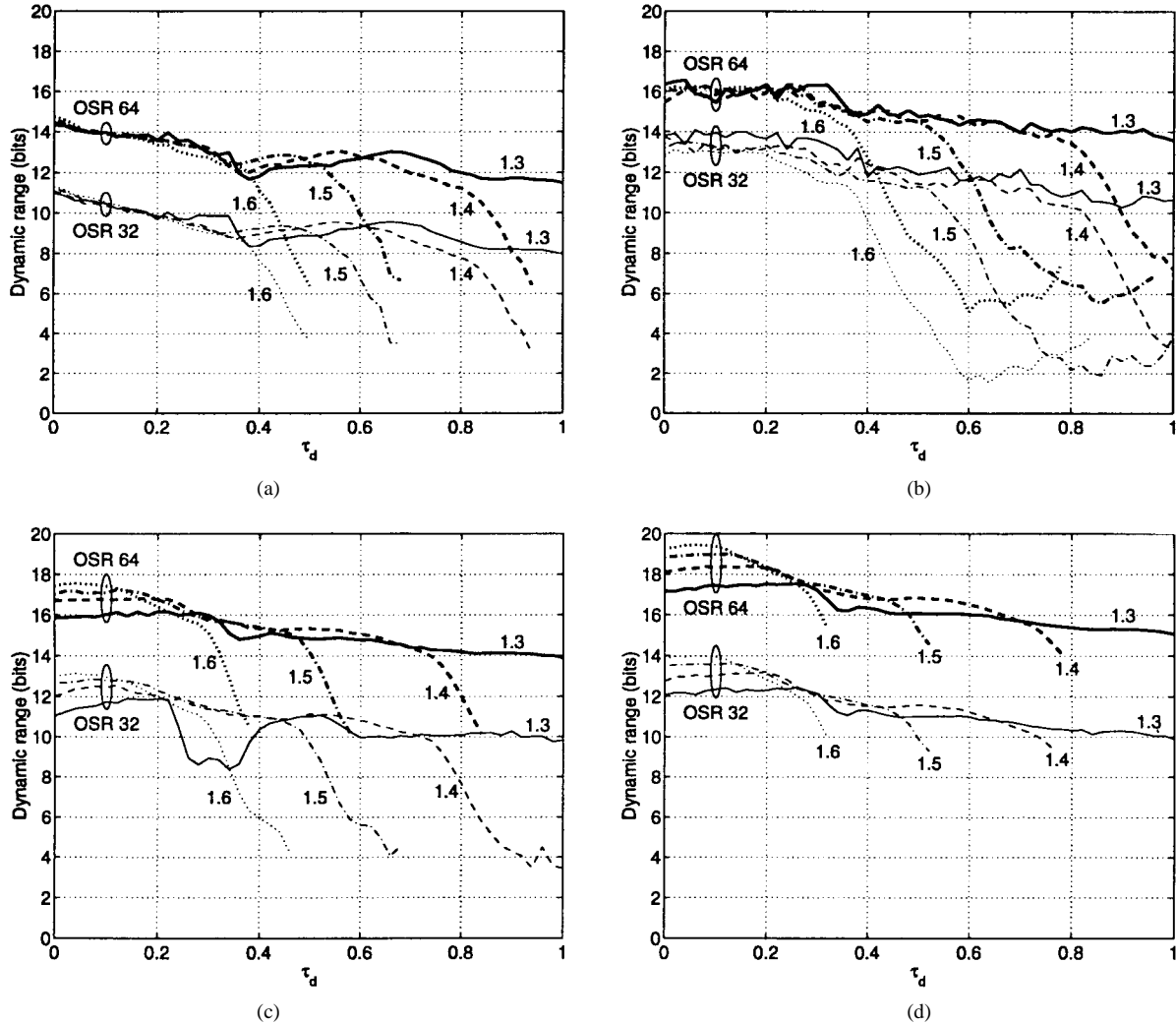Fig. 13.   Block diagram for general high-order LP CT $\Delta\Sigma$M.

Fig. 14.   DR for high-order LP CT $\Delta\Sigma$M's. Numbers on curves are OOBG values. Third-order LP Butterworth with (a) dc zeros and (b) optimal zeros; (c) fourth-order LP Butterworth with optimal zeros; and (d) fifth-order LP Chebyshev with optimal zeros.

was stable for a full sample of excess delay, while the same BP modulator was only stable up until $\tau_d = 0.65$.

In conclusion, LP modulators of order higher than two let us choose OOBG as an anti-delay measure at the cost of resolution. High-order multifeedback BP modulators do likewise, though their immunity to excess delay is not as good as in their LP counterparts.[4] Finally, in fairness, even though the resolution of some of the ideal modulators in

Fig. 14 exceeds 16 bits, it is unlikely that gigahertz-speed modulators would achieve such a high resolution, because other nonidealities such as thermal noise and clock jitter will almost surely limit performance more than quantization noise.

VI. MODULATORS WITH A MULTIBIT QUANTIZER

Thus far, this study has simulated $\Delta\Sigma$M's employing a single-bit quantizer. It is known that multibit quantizers in DT designs improve stability [2, Ch. 8] and sensitivity to clock jitter [37]. If the previous section is any guide, we can hope for an improvement in the immunity of CT designs with a multibit quantizer to excess delay.

[4] The Matlab code written to do the transformations was unfortunately not sophisticated enough to handle BP modulators with noncoincident NTF zeros, though it seems reasonable to assume the results for such modulators would echo those seen in Fig. 15.
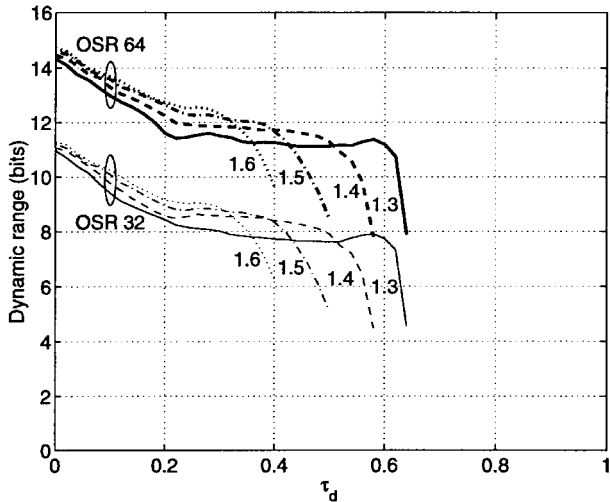
Fig. 15. DR for sixth-order BP CT $\Delta\Sigma$M. Numbers on curves are OOBG values.

There is some improvement, but not a lot. Fig. 16(a) shows the DR against excess delay for the second-order LP modulator for OSR $=64$, while Fig. 16(b) is for the fourth-order BP modulator. The thick lines are from Figs. 7 and 12, the results for a 2-level (1-bit) quantizer, and the other lines are for 3-, 4-, 8-, and 16-level (1.5-, 2-, 3-, and 4-bit) quantizers. Generally, DR improves with quantizer resolution as expected, and furthermore, the τ_d range over which the modulators remain stable improves a little with increasing quantizer resolution. Similar results are seen for the high-order LP modulators as for the second-order LP. We see the fourth-order BP circuit can be stable for τ_d close to 0.7 with a 4-bit quantizer, compared to 0.5 for a 1-bit quantizer. Again, similar results are seen for the sixth-order BP modulator.

The traditional problem in multibit designs is that any level mismatches in the multibit-feedback DAC are directly input-referred, thereby limiting the achievable performance. Techniques such as *dynamic element matching* (DEM) [38]–[40] and digital post-correction [41] have been proposed to alleviate these problems. A difficulty implementing either technique in a high-speed $\Delta\Sigma$M is that they require digital circuitry switching at f_s, which would cause a great deal of switching noise that might couple through the substrate into the forward modulator circuitry and degrade performance. Moreover, DEM would mean switching circuitry in the feedback path, which would add excess delay. To the authors' knowledge, no one has yet attempted to build a high-speed CT $\Delta\Sigma$M with a multibit quantizer.

VII. COMPENSATING FOR EXCESS LOOP DELAY

All is not lost for the second-order LP and fourth-order BP modulators when there is a good deal of excess delay, nor indeed for higher order high-OOBG modulators. We turn now to how to compensate for its effects in single-bit designs, though the results are equally applicable to multibit designs.

A. DAC Pulse Selection

In Section III, we considered the second-order LP $\Delta\Sigma$M with NRZ DAC pulses. A problem with this kind of pulse is
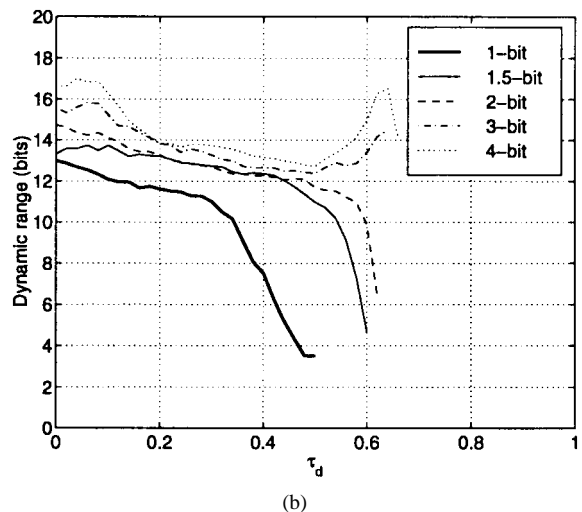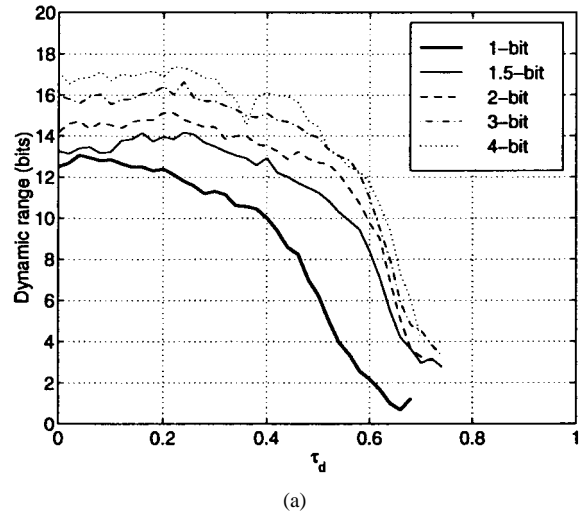

(a)

(b)

Fig. 16. Modulators with multibit quantizers: (a) second-order LP and (b) fourth-order BP.

that any excess loop delay $\tau_d > 0$ causes $\beta > 1$, which means the end of the pulse extends beyond $t = 1$. We saw in (15)–(19) that this increases the order of the resulting equivalent $H(z)$; in (19), $H(z)$ has the two poles at $z = 1$, but it acquires an additional pole at $z = 0$ for $\tau_d > 0$. Thus, the second-order modulator we tried to build actually has a *third* order loop filter.[5] In general, in any CT modulator with enough excess delay to push the falling DAC pulse edge past $t = 1$, the order of the equivalent DT loop filter is one higher than the order of the CT loop filter. Thus, a multifeedback BP modulator using either an NRZ or HRZ pulse increases in order, as do the higher order LP modulators from Section V with NRZ DAC's.

If we were to use DAC pulses with $\beta < 1$, then the pulses would extend past $t = 1$ only if the condition

$$\tau_d > 1 - \beta \qquad (29)$$

held. This suggests the following for the second-order LP modulator in Fig. 8; if we used an RZ DAC instead of an

---

[5]For small τ_d, the NTF has a pole and a zero close to one another which almost cancel, so the design appears approximately second-order in that case.
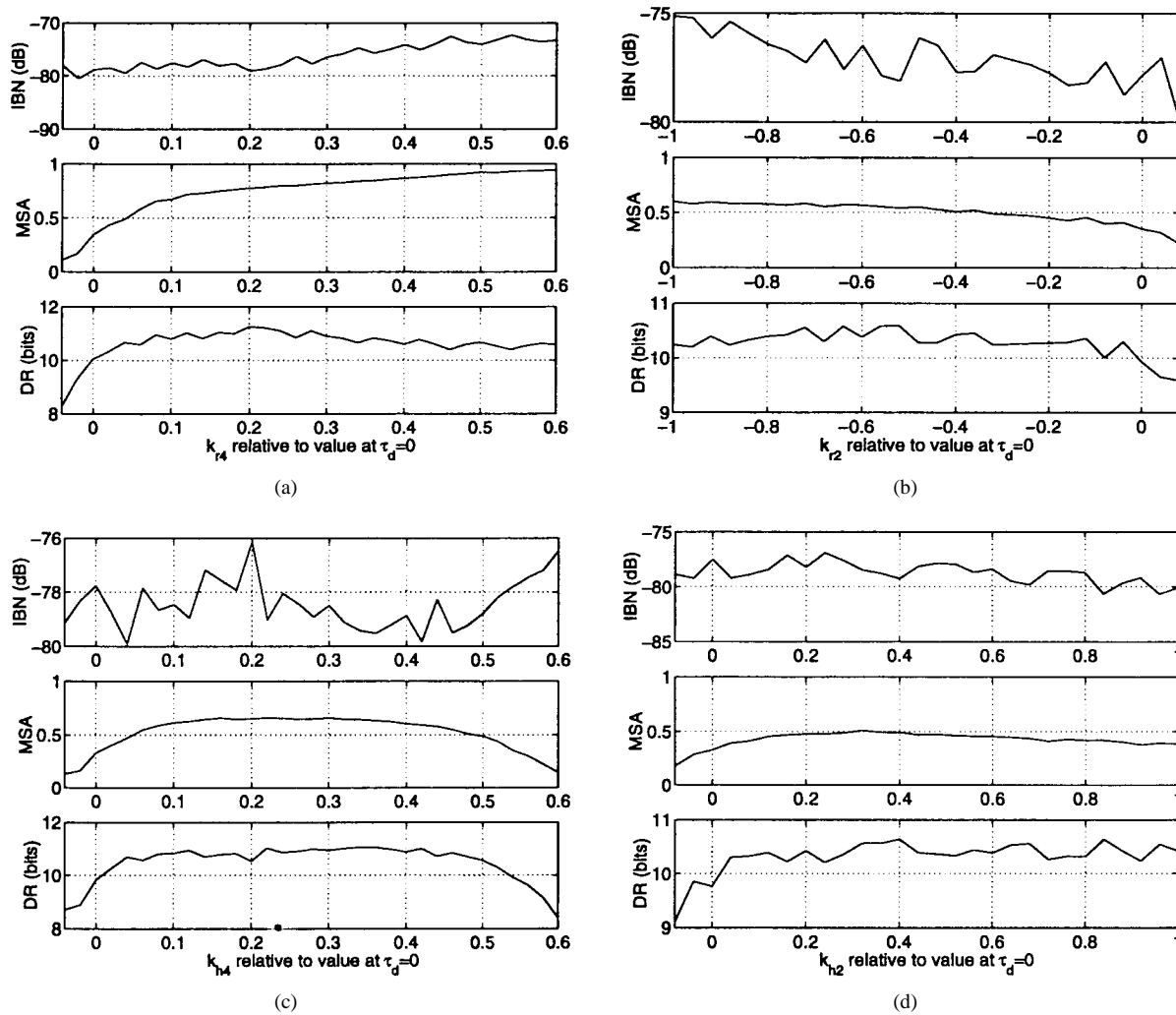
Fig. 17. Fourth-order $f_s/4$ BP CT $\Delta\Sigma$M performance at $\tau_d = 35\%$ delay with feedback coefficient tuning. (a) Tuning k_{r4}. (b) Tuning k_{r2}. (c) Tuning k_{h4}. Tuning k_{h2}.

NRZ DAC, $H(z)$ would remain second-order for $\tau_d \leq 0.5$. If we knew exactly what τ_d was, we could select the feedback coefficients $\{k_2, k_1\}$ to get exactly the equivalent $H(z)$ from (7). Let us demonstrate this: For Fig. 8, the loop filter is

$$\hat{H}(s) = \frac{k_2 + k_1 s}{s^2}. \tag{30}$$

Applying Table III to the partial fraction expansion of this for $(\alpha, \beta) = (\tau_d, \tau_d + 0.5)$ gives

$$H(z, \tau_d) = \frac{[4k_1 + k_2(3 - 4\tau_d)]z + [-4k_1 + k_2(1 + 4\tau_d)]}{8(z - 1)^2}. \tag{31}$$

We wish for this to equal (7); equating powers of z in the numerator and solving yields

$$\{k_2, k_1\} = \left\{-2, -\frac{5}{2} - 2\tau_d\right\}. \tag{32}$$

Thus, for a given $\tau_d \leq 0.5$ and RZ DAC pulses, we can make our $\hat{H}(s)$ match exactly the desired $H(z)$ by *tuning* the parameter k_1. In the particular circuit of Fig. 1, this is accomplished by changing the value of the current source in the rightmost differential pair DAC.

It has long been recognized that it is sensible to use RZ DAC pulses in low pass CT $\Delta\Sigma$M's [6], [9], [43]. Apart from the immunity to excess delay it afford us, an RZ DAC also alleviates intersymbol interference problems caused by nonsymmetric DAC pulse rise and fall times [43]. However, the differential circuit architecture of Fig. 1 also avoids this nonsymmetry [10] even with NRZ pulses.

B. Feedback Coefficient Tuning

As we have noted, if there exists enough excess delay to push the falling edge of a DAC pulse past $t = 1$, the modulator order increases by one. Therefore, there will be $m + 1$ coefficients in the numerator of the equivalent $H(z)$; with only m feedback coefficients k, the system is not fully controllable via these k's alone. Previous examinations of loop delay in $f_s/4$ BP $\Delta\Sigma$M's (notably [5, Sec. 3.1.4] and [24]) have studied the system in Fig. 10 using the modified $\mathcal{Z}$-transform and found the number of parameters in the numerator is m. The multifeedback architecture achieves a numerator coefficient of zero for the z^{-1} term, only because of a perfect cancellation when $\tau_d = 0$. For $\tau_d \neq 0$, the cancellation is ruined, so the coefficient of z^{-1} is nonzero,

yet the modified $\mathcal{Z}$-transform incorrectly finds it to remain zero [24]. There are actually $m+1$ rather than m numerator coefficients for $f_s/4$ modulators with excess delay.

Even though delay causing $\beta > 1$ means the system cannot be controlled perfectly with the k's, *some* degree of control can be exercised. We demonstrate the helpfulness of this on the fourth-order multifeedback modulator in Fig. 10. Suppose there is a fixed excess delay of $\tau_d = 35\%$: Fig. 12 shows that for OSR $=,64$ a DR of 9.9 bits is achieved using the nominal k values in (27). It is found that IBN $= -78$ dB and MSA $= 0.34$ at this τ_d.

Fig. 17 shows how the performance of the modulator is affected when the k's are tuned one at a time away from their nominal values. By adopting a steepest-descent tuning approach where each k is tuned iteratively until the DR is maximized, we find that it is possible to improve the DR from 9.9 bits to 11.3 bits, still at $\tau_d = 0.35$. The IBN and MSA are also improved, IBN to -79 dB and MSA to 0.74. The k values which give this performance are approximately

$$\{k_{r4}, k_{r2}, k_{h4}, k_{h2}\} = \{-0.87, -1.83, 0.48, 1.89\}. \quad (33)$$

The tuned k performance is still not as good as the 13 bits achieved at no excess delay in Fig. 12, but it is an improvement compared to the untuned k performance.

Fig. 18 compares the modulator DR for untuned k parameters from Fig. 12 and tuned k parameters where the steepest-descent algorithm was applied for several different values of excess delay between zero and one. We see that it is possible to find k values which keep the modulator stable for the entire range of τ_d. What is perhaps more surprising is that performance worsens up to 50% excess delay, but then actually starts to improve again until there is a full sample delay, whereupon the performance becomes as good as it was for no delay at all. How can this apparently incongruous result be true?

Recall $H_{\mathrm{BP}}(z)$ in (25). The numerator was $2z^{-2} + z^{-4}$. The z^{-2} means there is a two-sample delay in the feedback; every $\Delta\Sigma M$ must have at least one sample of delay in order to be causal. We found the equivalent $\hat{H}_{\mathrm{BP}}(s)$ in (26); the two-sample delay is implicit in this equation. Note that

$$H_{\mathrm{BP}}(z) = \frac{2z^{-2} + z^{-4}}{(1 + z^{-2})^2} = z^{-1}\frac{2z^{-1} + z^{-3}}{(1 + z^{-2})^2}. \quad (34)$$

This suggests we could place a digital latch that provides one sample of delay (z^{-1}) prior to the DAC's, and then find the equivalent $\hat{H}_{\mathrm{BP}}(s)$ for the $H(z)$ with numerator $2z^{-1} + z^{-3}$. In other words, we have two choices for building a two-sample delay into the CT feedback loop: by matching to an $H(z)$ with two delays in the numerator, *or* by providing a latch which adds one delay and matching to an $H(z)$ with one delay in the numerator. These are denoted, respectively, the *zero* and *one digital delay* schemes in [42]. This choice is peculiar to $f_s/4$ BP modulators; it does not exist for LP modulators or BP modulators with a different center frequency because they invariably have a nonzero z^{-1} term in the numerator, and therefore $H(z)$ would become noncausal if we were to factor out a z^{-1} as we did in (34).
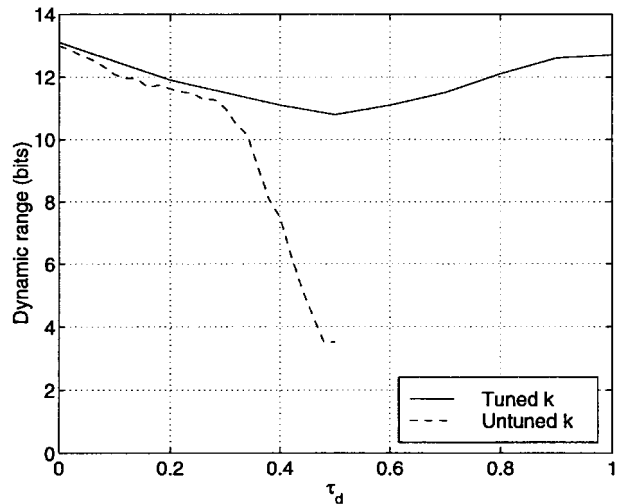

Fig. 18. Multifeedback BP modulator-DR with k tuning.

For each scheme, it is possible to find analytically the feedback k's which implement the desired $H(z)$

$$\{k_{r4}, k_{r2}, k_{h4}, k_{h2}\}$$
$$= \begin{cases} \{-1.086\,78, -2.133\,88, 0.450\,16, 1.487\,44\}, \\ \quad \text{zero digital delay} \\ \{-0.450\,16, -0.633\,88, 1.086\,78, 2.987\,44\}, \\ \quad \text{one digital delay} \end{cases} \quad (35)$$

where the first set of k's is from (27). The reason for the identical DR performance observed at both $\tau_d = 0$ and $\tau_d = 1$ is now clear. For $\tau_d = 1$, the optimal k's are those in the second row of (35), and the steepest-descent algorithm turns out to converge to values close to those. For $0 < \tau_d < 1$, the k's for optimal DR lie in between the zero and one digital delay values—compare, for example, (33) for $\tau_d = 0.35$ to (35)—though unfortunately the relationship between τ_d and the k's which optimize DR is not linear. For example, for $\tau_d = 0.5$, picking k values that lie exactly half way between the values in (35) leads to DR $=9.2$ bits, though the steepest-descent algorithm found k values to make a modulator with DR $=10.8$ bits.

In any case, Fig. 18 is strong encouragement to design the k's to be tunable, possibly even for on-line calibration against process and temperature variations. How to design a tuning algorithm to maximize DR that works on-chip, perhaps even while the modulator is operating, is an interesting topic for future research.

C. Additional Feedback Parameters

If $\beta > 1$ causes the modulator order to increase from m to $m+1$, and we only have m feedback coefficients, then it stands to reason that adding an additional feedback should restore full controllability to the system. This has been suggested in [25]. In the block diagram of Fig. 8, a third NRZ feedback was added whose output goes directly to a summing node after the second integrator (that is, immediately prior to the quantizer). To use this approach in a circuit architecture like Fig. 1, where the quantizer input must be a voltage but summation is done with currents, we would have to add a transconductor followed

by a current-to-voltage converter in between the second op amp and quantizer.

We can avoid adding components in the forward $\Delta\Sigma M$ path by adding one additional feedback with a *different* kind of DAC pulse. This is akin to the multiple feedbacks in the multifeedback BP circuit. By way of example, consider again Fig. 8; let us denote its NRZ feedback parameters $k_{n2} = k_2$ and $k_{n1} = k_1$, and let us suppose there is a third feedback which goes to the same summing node as k_{n1}: An HRZ DAC with coefficient k_{h1}. The z-domain equivalents for the NRZ pulses with excess delay have already been found in (17) and (18); to generalize them to feedbacks k_{n1} and k_{n2} instead of -1.5 and -1 is a trivial change to those equations. For an HRZ pulse delayed by τ_d, the z-domain equivalent is

$$\frac{k_{h1}}{s} \to \frac{k_{h1}(0.5 - \tau_d)}{z - 1} + z^{-1}\frac{k_{h1}\tau_d}{(z - 1)}. \tag{36}$$

Combining this with (17) and (18) yields

$$H(z) = \frac{y_2 z^2 + y_1 z + y_0}{z(z - 1)^2} \tag{37}$$

where $\{y_2, y_1, y_0\}$ are expressions involving $\{k_{n2}, k_{n1}, k_{h1}, \tau_d\}$. We wish for the numerator of this to equal $-2z^2 + z$ from (7), and Maple can be used to solve symbolically for the k values

$$k_{n2} = \frac{\tau_d^2 + 2}{\tau_d^2 - 2}$$
$$k_{n1} = \frac{-\tau_d^4 + 4\tau_d^3 - 12\tau_d^2 + 10\tau_d + 4}{\tau_d^2 - 2}$$
$$k_{h1} = \frac{\tau_d^4 - 4\tau_d^3 + 13\tau_d^2 - 12\tau_d - 2}{\tau_d^2 - 2}. \tag{38}$$

Therefore, given the excess delay τ_d, we can get exactly the $H(z)$ in (7) by tuning the feedbacks to the values given in (38).

We could also use an HRZ pulse fed back to the first summer; this would give us different equations from (38), but it would still be possible to achieve the $H(z)$ in (7). However, we could *not* use an RZ pulse in place of an HRZ pulse. This is because for $\tau_d \leq 0.5$, the RZ pulse would not contribute to y_0 in (37); only k_{n2} and k_{n1} would, and thus to set $y_0 = 0$ (as (7) dictates) would require $k_{n2} = k_{n1} = 0$, which renders the feedback inoperational.

How do we add an additional parameter to the BP multifeedback architecture for delay compensation? Interestingly, adding an NRZ pulse to Fig. 10 turns out not to work. This is because an NRZ pulse is a linear combination of RZ and HRZ pulses, so its feedback parameters are not independent. An independent pulse is needed—for example, a pulse with $(\alpha, \beta) = (0.25, 1)$ in combination with any two of NRZ, RZ, and HRZ—but generating a pulse other than these latter three might be nontrivial at high speed.

VIII. CONCLUSION

Excess loop delay in a CT $\Delta\Sigma M$ is a delay between the sampling clock edge and the change in output bit as seen at the feedback point in the modulator. It arises because of the nonzero switching time of the transistors in the feedback path,

and is significant because it alters the equivalence between the CT and DT representations of the loop filter, $\hat{H}(s)$ and $H(z)$. Its effect on performance is severe if the sampling clock speed is an appreciable fraction (10% or more) of the maximum transistor switching speed; this is becoming more likely nowadays as desired conversion bandwidths increase and delta–sigma modulation with an aggressively high clock rate relative to the transistor switching speed is considered for the converter architecture.

If excess delay is not designed for, then as excess delay increases as a fraction of the clock period, second-order LP and fourth-order $f_s/4$-BP modulators will suffer in terms of in-band noise, maximum stable input amplitude, and DR. Higher order LP designs seem more robust if designed using NTF prototyping because there is a parameter, the out-of-band gain, which can be selected to give some immunity to excess delay. Higher-order BP designs are also more robust than lower order ones, but a multifeedback $f_s/4$-BP design is always found to be less immune to excess delay than the corresponding LP design. The use of a multibit quantizer is somewhat helpful, though incorporating the usually-needed correction circuitry for a feedback DAC with mismatched levels is nontrivial for high-speed designs.

It is more sensible to recognize the presence of excess delay and take it into account in the design process. We have demonstrated that choosing the right DAC pulse shape in combination with tuning of the feedback parameters (either in the design phase or automatically on-line) can greatly mitigate the performance loss due to delay, to the point that excess delay can be rendered effectively a nonproblem in high-speed CT $\Delta\Sigma M$'s.

REFERENCES

[1] J. C. Candy and G. C. Temes, Eds., *Oversampling Delta–Sigma Data Converters: Theory, Design, and Simulation.* New York: IEEE Press, 1992.
[2] S. R. Norsworthy, R. Schreier, and G. C. Temes, Eds., *Delta–Sigma Data Converters: Theory, Design, and Simulation.* New York: IEEE Press, 1997.
[3] B. E. Boser and B. A. Wooley, "The design of sigma–delta modulation analog-to-digital converters," *IEEE J. Solid-State Circuits,* vol. 23, pp. 1298–1308, Dec. 1988.
[4] J. Nedved, J. Vanneuville, D. Gevaert, and J. Sevenhans, "A transistor-only switched current sigma–delta A/D converter for a CMOS speech codec," *IEEE J. Solid-State Circuits,* vol. 30, pp. 819–822, July 1995.
[5] O. Shoaei, "Continuous-time delta–sigma A/D converters for high speed applications," Ph.D. dissertation, Carleton Univ., Ottawa, Canada, 1996.
[6] K. T. Chan and K. W. Martin, "Components for a GaAs delta–sigma modulator oversampled analog-to-digital converter," in *Proc. Int. Symp. Circuits Syst.,* 1992, pp. 1300–1303.
[7] B. Hallgren, "Design of a second order CMOS sigma–delta A/D converter with a 150 MHz clock rate," in *Proc. European Solid-State Circuts Conf.,* 1992, pp. 103–106.
[8] S. Feng, J. Sauerer, and D. Seitzer, "Implementation of GaAs E/D HEMT analog components for oversampling analog/digital conversion," in *GaAs Int. Circuits Symp.,* 1994, pp. 228–231.
[9] K. R. Nary, S. Beccue, R. Nubling, R. Pierson, K.-C. Wang, P. Zampardi, and A. Jayaraman, "Second order $\Delta\Sigma$ modulators using AlGaAs/GaAs HBT's," in *GaAs Int. Circuits Symp.,* 1994, pp. 232–235.

[10] J. F. Jensen, G. Raghavan, A. E. Cosand, and R. H. Walden, "A 3.2-GHz second-order delta–sigma modulator implemented in InP HBT technology," *IEEE J. Solid-State Circuits,* vol. 30, pp. 1119–1127, Oct. 1995.

[11] M. Erbar, M. Rieger, and H. Schemmann, "A 1.28-GHz sigma–delta modulator for video A/D conversion," in *Int. Conf. Consumer Electron.,* pp. 78–79, 1996.

[12] G. Raghavan, J. F. Jensen, R. H. Walden, and W. P. Posey, "A bandpass $\Sigma\Delta$ modulator with 92dB SNR and center frequency continuously programmable from 0 to 70MHz," in *Int. Solid-State Circuits Conf. Tech. Dig.,* pp. 214–215, 1997.

[13] W. Gao and W. M. Snelgrove, "A 950MHz second-order integrated LC bandpass $\Delta\Sigma$ modulator," in *Proc. VLSI Circuits Symp.,* pp. 111–112, 1997.

[14] A. Jayaraman, P. Asbeck, K. Nary, S. Beccue, and K.-C. Wang, "Bandpass delta–sigma modulator with 800 MHz center frequency," in *GaAs Int. Circuits Symp.,* 1997, pp. 95–98.

[15] A. Olmos, T. Miyashita, M. Nihei, E. Charry, and Y. Watanabe, "A 5GHz continuous time sigma–delta modulator implemented in $0.4~\mu$m InGaP/InGaAs HEMT technology," in *Proc. Int. Symp. Circuits Syst.,* vol. 1, pp. 575–578, 1998.

[16] W. Gao, J. A. Cherry, and W. M. Snelgrove, "A 4 GHz fourth-order SiGe HBT band pass $\Delta\Sigma$ modulator," in *Proc. VLSI Circuits Symp.,* pp. 174–175, 1998.

[17] J. C. Candy, "A use of double integration in sigma–delta modulation," *IEEE Trans. Commun.,* vol. 33, pp. 249–258, Mar. 1985.

[18] R. Schreier, "An empirical study of high-order single-bit delta–sigma modulators," *IEEE Trans. Circuits Syst. II,* vol. 43, pp. 461–466, Aug. 1993.

[19] V. F. Dias, G. Palmisano, and F. Maloberti, "Noise in mixed continuous-time switched-capacitor sigma–delta modulators," *Proc. Inst. Electron. Eng. G,* pp. 680–684, Dec. 1992.

[20] J. A. Cherry, W. M. Snelgrove, and P. Schvan, "Signal-dependent timing jitter in continuous-time $\Sigma\Delta$ modulators," *Electron. Lett.,* vol. 33, no. 13, pp. 1118–1119, June 1997.

[21] A. Gosslau and A. Gottwald, "Optimization of a sigma–delta modulator by the use of a slow ADC," in *Proc. Int. Symp. Circuits Syst.,* vol. 3, pp. 2317–2320, 1988.

[22] ——, "Linearization of a sigma–delta modulator by a proper loop delay," in *Proc. Int. Symp. Circuits Syst.,* vol. 1, pp. 364–367, 1990.

[23] U. Horbach, "Design of a 20 bit sigma–delta A/D-converter for audio applications," in *Proc. Int. Symp. Circuits Syst.,* vol. 4, pp. 2789–2792, 1990.

[24] W. Gao, O. Shoaei, and W. M. Snelgrove, "Excess loop delay effects in continuous-time delta–sigma modulators and the compensation solution," in *Proc. Int. Symp. Circuits Syst.,* vol. 1, pp. 65–68, 1997.

[25] P. Benabes, M. Keramat, and R. Kielbasa, "A methodology for designing continuous-time sigma–delta modulators," in *IEEE European Design Test Conf.,* 1997, pp. 46–50.

[26] A. M. Thurston, T. H. Pearce, and M. J. Hawksford, "Bandpass implementation of the sigma–delta A-D conversion technique," in *Int. Conf. on A.-D. and D.-A. Conversion,* pp. 81–86, 1991.

[27] F. M. Gardner, "A transformation for digital simulation of analog filters," *IEEE Trans. Commun.,* vol. 44, pp. 676–680, July 1986.

[28] D. Redfern, *The Maple Handbook.* New York: Springer-Verlag, 1994.

[29] R. Schreier and B. Zhang, "Delta–sigma modulators employing continuous-time circuitry," *IEEE Trans. Circuits Syst.,* vol. 43, pp. 324–332, Apr. 1996.

[30] D. C. Hanselman and B. Littlefield, *Mastering MATLAB 5: A Comprehensive Tutorial and Reference.* Englewood Cliffs, NJ: Prentice-Hall, 1998.

[31] L. Risbo, "$\Sigma\Delta$ modulators—Stability and design optimization," Ph.D. dissertation, Technical Univ. Denmark, 1994.

[32] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.,* pp. 446–472, July 1948.

[33] F. W. Singor and W. M. Snelgrove, "Switched-capacitor delta–sigma A/D modulation at 10.7MHz," *IEEE J. Solid-State Circuits,* vol. 30, pp. 184–192, Mar. 1995.

[34] K. C.-H. Chao, S. Nadeem, W. L. Lee, and C. G. Sodini, "A higher order topology for interpolative modulators for oversampled A/D converters," *IEEE Trans. Circuits Syst.,* vol. 37, pp. 309–318, Mar. 1990.

[35] O. Shoaei and W. M. Snelgrove, "Optimal (bandpass) continuous-time $\Sigma\Delta$ modulator," in *Proc. Int. Symp. Circuits Syst.,* 1994, vol. 5, pp. 489–492.

[36] ——, "A multi-feedback design for LC bandpass delta–sigma modulators," in *Proc. Int. Symp. Circuits Syst.,* vol. 1, pp. 171–174, 1995.

[37] R. Adams, K. Nguyen, and K. Sweetland, "A 113 dB SNR oversampling DAC with segmented noise-shaped scrambling," in *Int. Solid-State Circuits Conf. Tech. Dig.,* 1998, pp. 62–63.

[38] R. Schreier and B. Zhang, "Noise-shaped multibit D/A convertor employing unit elements," *Electron. Lett.,* pp. 1712–1713, Sept. 1995.

[39] T. Kwan, R. Adams, and R. Libert, "A stereo multibit $\Sigma\Delta$ DAC with asynchronous master clock interface," *IEEE J. Solid-State Circuits,* vol. 31, pp. 1881–1887, Dec. 1996.

[40] I. Galton, "Noise-shaping D/A conversion for $\Delta\Sigma$ modulation," in *Proc. Int. Symp. Circuits Syst.,* vol. 1, pp. 441–444, 1996.

[41] M. Sarhang-Nejad and G. C. Temes, "A high-resolution multi-bit $\Sigma\Delta$ ADC with digital correction and relaxed amplifier requirements," *IEEE J. Solid-State Circuits,* vol. 28, pp. 648–660, June 1993.

[42] O. Shoaei and W. M. Snelgrove, "Design and implementation of a tunable 40 MHz–70 MHz Gm-C bandpass $\Delta\Sigma$ modulator," *IEEE Trans. Circuits Syst. II,* vol. 44, pp. 521–530, July 1997.

[43] R. Adams, "Design and implementation of an audio 18-bit analog-to-digital converter using oversampling techniques," *J. Audio Eng. Soc.,* pp. 153–166, Mar./Apr. 1986.

James A. Cherry (S'97) received the B.A.Sc. degree in computer engineering from the University of Waterloo, Ont., Canada, in 1992, graduating at the top of his class, and the M.Eng. and Ph.D. degrees in electronics from Carleton University, Ottawa, Ont., in 1994 and 1998, respectively. His Ph.D. thesis was on the fundamental performance limitations of continuous-time delta–sigma modulators for high-speed data conversion using continuous-time delta–sigma modulators.

From 1996 to 1998, has was a Consultant to the Technology Access Group of Nortel, Ottawa, Canada, in the area of high-speed continuous-time delta–sigma modulators. In 1998, he joined Philsar Electronics Inc., Ottawa, Canada, where he is an analog/mixed-signal circuit designer.

W. Martin Snelgrove (M'81) received the Ph.D. degree in electrical engineering from the University of Toronto, Canada, in 1982.

He taught at the University of Toronto until 1992, then moved to Carleton University, Ottawa, Canada, as a Professor and Chair of OCRI/NSERC Industrial Research in High Speed Integrated Circuits. In 1998, he joined Wireless Systems Technologies, Inc., and Philsar Electronics, Inc., as Director of Research and Development.

Dr. Snelgrove received the 1986 IEEE Circuits and Systems Society Guillemin–Cauer Award for a 1986 paper coauthored with A. S. Sedra. He served as Associate Edtior for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1994 to 1997.