A Second-Order Sigma-Delta A/D Converter

Keith Onodera and Andy Abo
EE290Y Project
May 9, 1995

Abstract— **The design of an audio rate, second order sigma-delta analog-to-digital converter is presented. The converter achieves a peak SNDR of 95 dB and consumes 382 μW (figure of merit is 0.19). Careful signal scaling saves power and prevents integrator overloading.**

I. Specification

Our goal was to design a sigma-delta A/D converter that achieves 95 dB SNDR at a Nyquist rate of 44kHz with minimum power consumption. To avoid large signal distortion, the converter integrators were required to settle to 0.003% accuracy at the maximum input step. This introduced a maximum integrator gain variation of 0.003%. The technology available was a 0.6μm CMOS process.

II. Architecture

A first order analysis can be used to determine which converter architecture will consume the least power at a reasonable complexity. The power will be dominated by static class A bias currents in the integrator opamps. The bias currents must be large enough to drive the capacitive loads quickly and accurately. Thus, the sampling capacitor size ultimately limits the minimum power and is determined by the amount of kT/C noise that can be tolerated.

Because the integrator usually directly samples a one-bit DAC, the opamp sees large input signals and settling time is dominated by slewing. Using this assumption, we can compute the approximate bias current required:

$$\left(\frac{I}{C}\right) t_s = \Delta$$
$$I \geq \frac{C\Delta}{t_s}$$

Where I is the static bias current, C is the capacitive load, t_s is the allowed setting time (one-half clock period), and Δ is the quantizer swing (worst case integrator output swing).

Assume $\Delta = 10$V (rail-to-rail differential swing on a 5V supply). We can now roughly compare the power of different architectures.

An ideal first order modulator requires an oversampling ratio of 2048 (must be a power of two for decimation) to achieve *over* 95 dB SNR based on the linearized model (L=1, M=2048):

$$\text{SNR} \approx \frac{3}{2}\frac{2L+1}{\pi^{2L}} M^{2L+1}$$

For a Nyquist rate of 44kHz the converter sampling rate is 90.1MHz giving a settling time t_s of 5.55 ns. If we account for thermal noise, we can compute the minimum C to get *exactly* 95dB SNR (assumes peak SNR occurs at maximum p-p input swing Δ):

$$\text{SNR} = \frac{\frac{\Delta^2}{8}}{\frac{\Delta^2}{8}\frac{2}{3}\frac{\pi^{2L}}{2L+1}\frac{1}{M^{2L+1}} + \frac{4kT}{MC}}$$

This equation requires $C \geq 2.6$fF. The total bias current based on slew rate would then be 4.7μA for the one integrator.

An ideal second order ($L = 2$) modulator requires a smaller oversampling ratio of $M = 256$ to achieve over 95 dB SNR. The corresponding converter sampling rate is 11.3 MHz and t_s is 44.4ns. Accounting for thermal noise $C \geq 4.2$fF to achieve exactly 95 dB SNR. The total bias current would be 1.9μA for two integrators.

Table I summarizes the relative power (assumes fixed voltage supply) of 1st, 2nd, and 3rd order modulators normalized to the second order.

L	M	t_s (ns)	C (fF)	P_{norm}
1	2048	5.55	2.6	2.50
2	256	44.4	4.2	1.00
3	64	178	17.4	1.55

TABLE I

While this simplified analysis is not an accurate predictor of absolute power, it is useful for comparison of relative power. It shows that higher order converters have relaxed settling time, but require larger sampling capacitance because the wideband thermal noise is oversampled less. The resulting slew rate requirement eventually increases for higher order modulators. Modulators higher

than second order also require careful design to ensure stability. Thus, we chose to design a second order modulator for our application.

It is interesting to note, that if the converter is dominated by thermal noise, our slew-limited approximation shows that for a given order L, bias current is independent of oversampling ratio M. Oversampling allows capacitance to be reduced by a factor M, but settling time is simultaneously decreased by M, keeping the slew-rate constant.

III. Practical considerations

The linearized model of sigma-delta converters can only provide an estimate of quantization noise. Because actual quantization noise and distortion is difficult to predict analytically, we used MIDAS to model non-idealities and allocated the remaining noise budget to thermal noise.

A. Signal scaling

A second-order architecture that is easy to implement in the switch-capacitor domain has been proposed by Boser [1]. Figure 1 shows a block diagram of this architecture.

Using ideal integrators, MIDAS simulation shows that this modulator can achieve a peak SNDR of 102.3 dB. If we assume the quantizer output is limited to $\pm$5V differential ($\Delta = 10$V), then simulation also shows that we require a differential swing of $\pm$5.3V at the output of the first integrator. If we assume a folded cascode opamp is used in the integrator, then we can realistically only swing a maximum of $\pm$3.8V differentially ($V_{\rm dsat} \approx 300$mV).

One alternative is to scale Δ down so that the integrator swing is within the limits of the opamp. This approach, however, decreases the amplitude of the input signal at peak SNR. Because the thermal noise does not scale with Δ, the overall SNR will be degraded, forcing an increase in sampling capacitance and hence power.

A better solution is to use signal scaling to lower the relative swing between the quantizer and the integrators. A scaled version of the same architecture is used to meet the opamp swing constraint without lowering Δ. This modulator is shown in figure 2.

The signal and quantization error transfer functions will be approximately unchanged in the signal band if we constrain that $G1 = G2$. This constraint preserves the ratio of the feedback signals to the two integrators. Actually, this introduces a high frequency pole (out of the signal band) in both the signal and quantization error transfer functions (SNR is preserved).

$$Y(z) = \frac{0.5G_1 z^{-2} X(z) + (1 - z^{-1})^2 E(z)}{1 + (G_1 - 2)z^{-1} + (1 - 0.5G_1)z^{-2}}$$

The integrator output swing can now be scaled down by attenuating $G1$ without degrading overall SNR.

MIDAS simulation shows that $G1 = G2 = 0.2$ yields differential swings of $\pm$3.0 and $\pm$1.7V at the outputs of the integrators ($\Delta = 10$V), which is well within the active range of the opamps.

B. Finite gain opamps

The finite gain of the opamp degrades the achievable SNR by making the integrators leaky. The integrator is implemented with switched capacitors as shown in figure 3 (single-ended version). The actual implementation is fully-differential to improve dynamic range and power supply rejection. The z-domain transfer function of this integrator is

$$\frac{V_o(z)}{V_i(z)} = \frac{1}{1 + 1/af}\left(\frac{C_s}{C_f}\right)\frac{z^{-1}}{1 - (1 + 1/af)z^{-1}}$$

$$\begin{aligned} f &= \frac{C_f}{C_f + C_s + C_p} \\ a &= \text{opamp DC gain} \\ C_p &= \text{opamp input parasitic capacitance} \\ C_l &= \text{opamp output parasitic capacitance} \end{aligned}$$

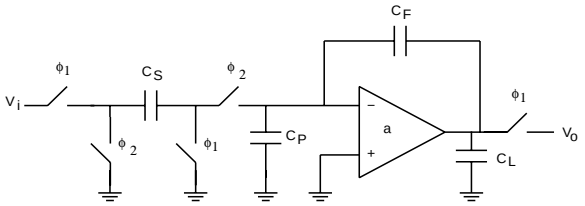

Fig. 3. A DDI integrator

This equation shows that higher loop gain af reduces leakage. The integrator leakage degrades the low-frequency noise suppression and lowers overall SNR. Figure 4 shows the converter peak SNR vs opamp gain for both integrators. Typically a minimum gain equal to the oversampling ratio is required [2], and from figure 4 we see a minimum gain of a few hundred should be sufficient to achieve a overall SNDR of 95 dB.

C. Thermal noise

The thermal noise in the sampling switches of first integrator results in an input referred noise power of approximately $\frac{4kT}{MC_s}$ (differential). This noise is usually the dominant noise source since minimizing capacitance tends to minimize power. The opamp output noise and the kT/C noise of the second integrator is first-order shaped; it is

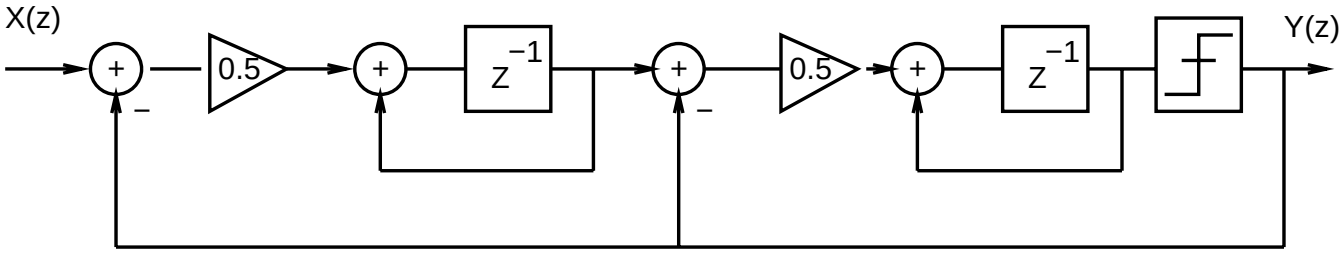

Fig. 1. A sampled-data second-order modulator

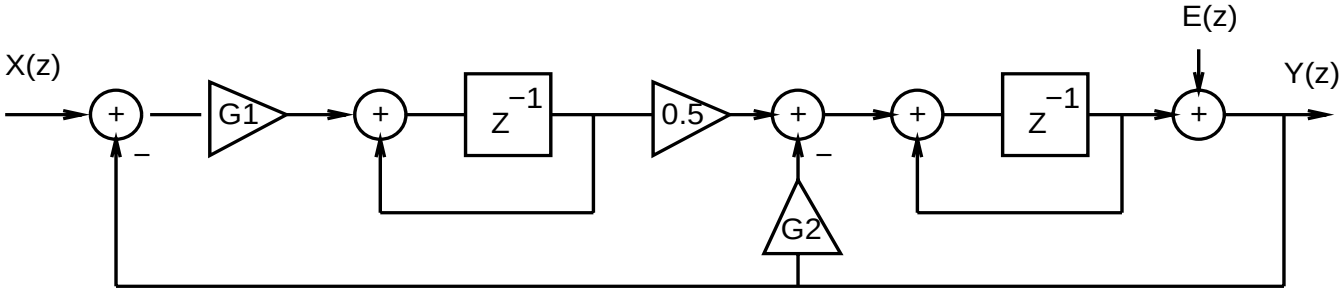

Fig. 2. A scaled second-order modulator

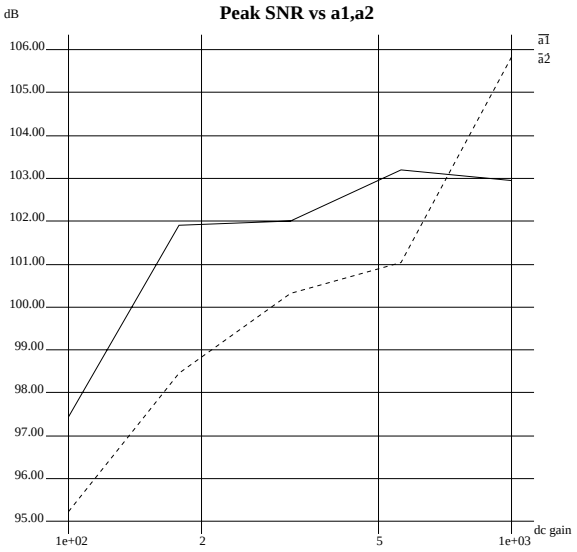

Fig. 4. Converter peak SNR vs opamp gain

shown to be insignificant in section IV:E. The total input-referred noise is the thermal noise plus the quantization noise and distortion.

The quantization and distortion is determined from MIDAS simulation. Using the scaled architecture above with finite opamp gain of 400 and DAC levels of ±5V differential ($\Delta = 10$V), the quantization and distortion power $q_{\rm rms}$ is 27.4μV rms at peak SNDR. The corresponding input power for peak SNDR $V_{\rm rms}$ is 2.38V rms.

Total SNDR including thermal noise is given by:

$$\text{SNDR} = \frac{V_{\rm rms}^2}{4\frac{kT}{MC_s} + q_{\rm rms}^2}$$

For an oversampling ratio of $M = 256$, the minimum C_s to achieve 95dB SNDR overall is 62fF. The corresponding thermal noise is 32.2μV rms. We chose $C_s = 70$fF, which determines $C_f = 350$fF due to the attenuation of 0.2 in the first integrator. For the second integrator, we can use minimum sized capacitors since that thermal noise is first-order shaped by the first integrator. The fully differential, switched-capacitor implementation is shown in figure 5.

IV. Amplifier design

Based on the MIDAS simulation, we require an opamp/integrator that meets the following specifications:

	Integrator 1	Integrator 2
Open loop gain (A_{o1})	> 400	> 400
Output swing (V_o)	±3V	±1.73V
Closed Loop Gain (A_{cl})	0.2	0.4/0.2
Sampling Capacitor	70fF	min (2×50fF)
V_{in}(max)	±1.7V	±1.5V
V_{sum}(max)	±4.2V	±2.75V
t_s	< 44.4 ns	< 44.4 ns

TABLE II

We used a CMOS folded cascode amplifier shown in figure 6.

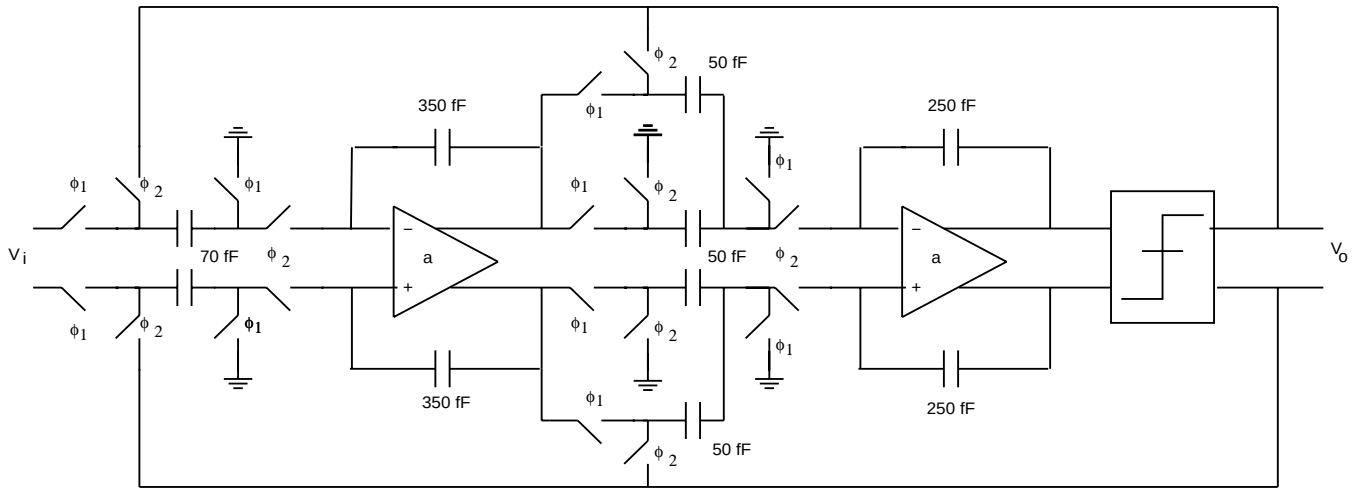

Fig. 5. Complete modulator

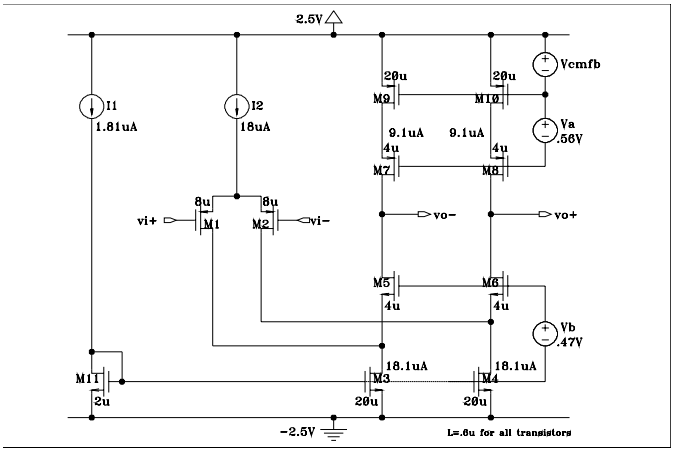

Fig. 6. Folded cascode OTA

A. Setting g_m (pmos inputs)

For an OTA integrator, the devices' g_m determines in part the settling time constant and the DC openloop gain as (see figure 3):

$$\begin{aligned} \tau &= \frac{1}{g_m}(C_s + C_p + \frac{1}{f}C_l) \\ A_{ol} &= g_m R_o \end{aligned}$$

where R_o is the resistance at the output node.

If the OTA does *not* slew then it will take 10.4τ to settle within .003% of the final value (assuming a single pole response). For an oversampling ratio of 256 and Nyquist sampling frequency of 44KHz, half the oversampling period is:

$$\frac{T}{2} = \frac{1}{2 \cdot 256 \cdot 44KHz} = 44.4\text{ns}$$

This requires τ to be:

$$\tau \leq \frac{44.4\text{ns}}{10.4} = 4.3\text{ns}$$

The closed loop gain roughly sets the feedback factor f:

$$\begin{aligned} \frac{1}{f} &= \frac{C_f + C_s + C_p}{C_f} \\ &= 1 + A_{cl} + \frac{C_p}{C_f} \\ &= 1 + .2 + .1 \\ &= 1.3 \end{aligned}$$

where $C_f = 350\text{fF}$, $C_s = 70\text{fF}$, and assuming $C_p = 35\text{fF}$. Assuming $C_l = 50\text{fF}$, requires g_m for Integrator 1 (Int1) to be:

$$\begin{aligned} g_m &\geq \frac{70\text{fF} + 35\text{fF} + 1.3 \cdot 50\text{fF}}{4.3\text{ns}} \\ &\geq 40\mu\text{A/V} \end{aligned}$$

With this minimum value for g_m sets the minimum R_o value as:

$$R_o \geq 400/g_m(\text{min}) = 400/40\mu\text{A/V} = 10\text{M}\Omega$$

For the folded cascode OTA shown in figure 6, I_2 is a key design value. Once the current I_2 is set the W/L of the pmos input devices can be determined. (Note: the second integrator (Int2) can use a minimum capacitor for C_s due to a more relaxed kT/C requirement so Int1 is only considered). Once I_2 is found it determines all the other currents since $I_{d3}, I_{d4} \geq I_2$.

B. Slew rate Limiting

The slew rate is determined by the maximum available current and the total capacitance at the output nodes.

$$\frac{dV}{dt}(max) = \frac{I_{\max}}{C_{\text{out}}} = \frac{I_2}{2Cout}$$

Using the fact that the OTA has a single pole response, you can find a criterion for the minimun current density for the input devices that will ensure no slew rate limiting. However this criterion is too unreasonable in power even for minimum input devices so slew rate limiting cannot be avoided.

To determine how much slew rate limiting there will be, the maximum signal (V_{sum}) at the integrator's summing node needs to be determined.

The maximum positive V_{sum} for Int1 is:

$$V_{\text{sum}}(\max) = V_{\text{in}}(\max+) - (V_{\text{ref}}-) = 3.4V - (-5V) = +8.4V$$

Where the quantizer output swings $\pm 2V_{\text{ref}}$ differentially. Thus, the voltages at v_i+/v_i- of the OTA will be +4.2V/-4.2V [1]. With these large input voltages the OTA will be slewing for most of its settling time. Neglecting C_l and C_p and assuming that the OTA slews 50% of the time, we need a current (single-ended) equal to:

$$\begin{aligned} I_{\text{avail}} &= C_s \frac{dV_{\text{sum}}}{dt} \\ &= \frac{70\text{fF} \cdot 4.2V}{22\text{ns}} \\ &= 3.4\mu A \Rightarrow I_2 = 26.8\mu A \end{aligned}$$

C. Final design values

Since the current levels were so low, initially all devices were minimum (2μm/$.6\mu$m).

The open loop gain (A_{ol}) was marginal so the input devices were made $4X$ to increase the g_m by $2X$.

The nmos and pmos cascodes were made $2X$ since this allowed drain sharing and improved the $g_m r_o$ self gain without increasing the output capacitance (C_l).

The nmos current mirrors (M3 & M4) were made $10X$ to reduce the bias current (I_1). This also improved the output resistance and increased A_{ol}.

The pmos current sources (M9 & M10) were also increased to $10X$ to improve their output resistance and A_{ol}.

[1] Our goal initially was to maximize the input signal power in order to tolerate the highest kT/C noise power and smallest C_s possible.

Although the project allows for ideal switches, actual MOS switches would cause the voltages on the summing nodes to be clamped at a diode above and below the supply rails, i.e. at approximately +/-3.2V (and hence reduce your signal power). This fact was realized after all our simulations were done. We will address the impact of later in the report.

As expected with the worst case settling condition (max V_{sum}), the output is slew rate limited for about 75% of the time. Thus the settling time is dominated by the slew rate performance as opposed to τ.

So the power is solely a function of C_s & $V_{\text{sum}}(\max)$ for a set time of:

$$t = .75 \cdot 44\text{ns} = 33\text{ns}$$

that is:

$$I_{\text{avail}} = \frac{C_s V_{\text{sum}}(\max)}{t}$$

For the above transistor sizes, $C_p = 10$fF and $C_l = 5$fF so neglecting them is okay. The reason for using the summing node rather than the output node when calculating slew rate is that the output node voltage is not as predictable due to its transient behavior.

$I_{\text{avail}}(= I_2/2)$ estimates:

Int1 $\quad 70\text{fF} \cdot 4.2V/33\text{ns} = 8.9\mu A$

Int2 $\quad 140\text{fF} \cdot 2.75V/33\text{ns} = 11.7\mu A$

From simulations the .003% settling could be obtained with $I_{\text{avail}} = 9\mu A$ for both integrators.

D. DC open loop gain

The worst case measurement of open loop gain is when the output is at its largest swing (since this is where R_o is minimum). We measured the OTA's input voltage at this $V_o(\max)$ to calculate the gains.

For both OTA's we measured:

$$A_{ol} > 1500$$

E. Opamp noise

The 2nd stage OTA's noise undergoes the same noise sampling as the quantization noise. Since it will be much smaller than the quantization noise it is neglected.

The first OTA's output noise will be sampled by the 2nd stage so we can derive the kT/C contribution using the circuit of figure 7. By integrating the output spectral noise density we find:

$$\overline{V_o^2} = \frac{R_{eq}}{R}\frac{kT}{C}$$

We first calculate the output current noise of the OTA. The main contributions are from the input devices (M1 & M2) and current mirror devices (M3, M4, M9, M10). They each contribute:

$$\overline{I_o^2} = 4kT\frac{2}{3}g_m\Delta f$$

so the total output noise current is:

$$\overline{I_{\text{tot}}^2} = 4kT\frac{2}{3}\Delta f(g_{m2} + g_{m4} + g_{m10})$$

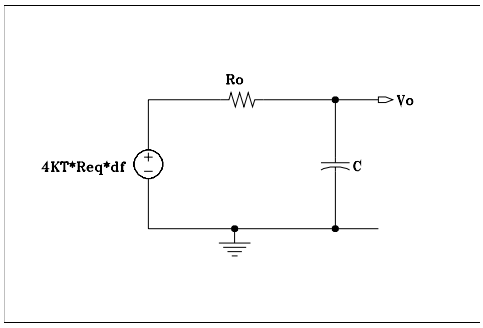

Fig. 7. Circuit for calculating sampled opamp noise

at the V_o+ output. In closed loop the output impedance of the OTA is

$$R_o = \frac{1}{g_{m2} f}$$

here f is the feedback factor so the total mean square output noise voltage is:

$$\overline{V_{\rm tot}^2} = \overline{I_{\rm tot}^2} R_o^2 = 4kTR_{eq}\Delta f$$

where:

$$\begin{aligned} R_{eq} &= \frac{2}{3}\frac{1}{g_{m2}}\frac{1}{f^2}\left(1+\frac{g_{m4}}{g_{m2}}+\frac{g_{m10}}{g_{m2}}\right) \\ &= \frac{2}{3}(g_{m2}+g_{m4}+g_{m10})R_o^2 \\ \Rightarrow \frac{R_{eq}}{R_o} &= \frac{2}{3}\frac{1}{f}(1+\frac{g_{m4}}{g_{m2}}+\frac{g_{m10}}{g_{m2}}) \end{aligned}$$

From this we can now calculate the kT/C contribution of the 2nd stage due to the output noise from this OTA:

$$\overline{V_n^2} = \frac{kT}{C}\frac{R_{eq}}{R_o} = \frac{kT}{C}\frac{2}{3}f\left(1+\frac{g_{m4}}{g_{m2}}+\frac{g_{m10}}{g_{m2}}\right)$$

From Spice:

$$\begin{aligned} \frac{g_{m4}}{g_{m2}} &= 290\mu/85\mu = 3.41 \\ \frac{g_{m10}}{g_{m2}} &= 108\mu/85\mu = 1.27 \end{aligned}$$

and since $C_s = 100$fF, $1/f = (1 + A_{cl} + C_p/C_f) = 1.23$;

$$\begin{aligned} R_{eq} &= 67k \\ \frac{\overline{V_{\rm tot}^2}}{\Delta f} &= 4kT \cdot 67k = (33.4\text{nV})^2 \\ \overline{V_n^2} &= 4.66 \cdot kT/100\text{fF} = (439\mu V)^2 \end{aligned}$$

This analysis ignores any real switch $R_{\rm on}$ which would modify the kT/C as:

$$\frac{kT}{C}\left(\frac{R_{\rm eq}+R_{\rm on}}{R_o+R_{\rm on}}\right)$$

F. Noise shaping

The transfer function of the modulator with the 2nd stage sampled kT/C noise (E_2) included but with all other noise sources neglected is:

$$Y(z) = \frac{z^{-2}X(z) + 5(1-z^{-1})z^{-1}E_2(z)}{12.5(1-1.8z^{-1}+.88z^{-2})}$$

From the transfer function we see that the output has a twice delayed version of the input plus the first difference of a delayed version of the kT/C noise. So the 2nd stage sampled kT/C noise is 1st ordered noised shaped and should have the benefits of being reduced by 9dB per doubling of M.

Even though the complex poles are outside of the signal band their influence would affect both the input and the noise equally and not change the SNR.

Being 1st ordered noised shaped E_2 contributes noise as:

$$\begin{aligned} \overline{N^2} &= \overline{V_n^2}\frac{\pi^2}{3}\frac{1}{M^3} \\ &= (439\mu V)^2 \cdot (443\times 10^{-6})^2 = (194\text{nV})^2 \\ sqrt\overline{N^2} &= 194 \text{ nV rms} \end{aligned}$$

This is negligible compared to the other noise sources.

V. Results

Table III summarizes the contributions of the various noise components at peak SNDR.

Source	(μV) rms
Quantization/Distortion	27.4
Input $\frac{kT}{C}$	30.3
1st opamp	0.2

TABLE III

A MIDAS simulation was performed with Guassian Noise sources used to represent the kT/C noise components. Figure 8 is a SNDR plot from that run. This plot shows a peak SNDR of greater than 105 dB. The power was partitioned evenly between the two integrators as shown in the table IV. Figure 9 shows the hspice verification of a worst case settling time of 42 ns for the first integrator. Figure 10 shows the hspice verification of a worst case settling time of 41 ns for the first integrator.

Figure of Merit	
Peak SNDR	> 95dB
FM	(95dB - 95dB) + .382mW/2mW 0.191

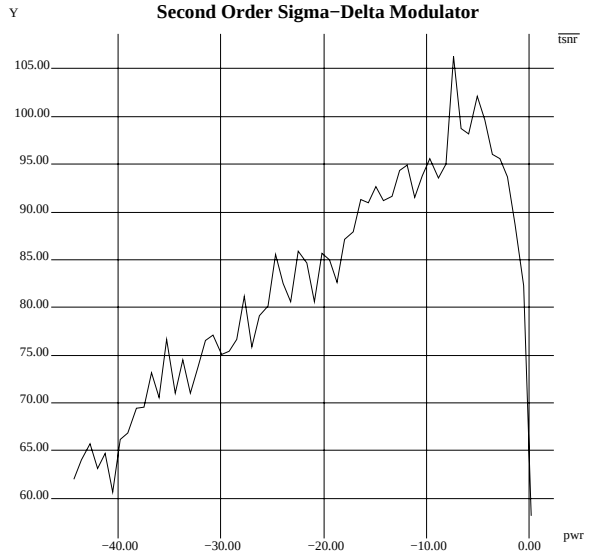

Fig. 8. SNDR (including thermal noise) vs input power

	I_1	I_2	$I_{d9} + I_{d10}$	Amp-total
Integrator 1	2.06μA	18μA	18.1μA	38.2μA
Integrator 2	2.06μA	18μA	18.1μA	38.2μA
total I_{dd}				76.4μA
total Power				382μW

TABLE IV

Appendix

A. Summing node overload

As previously mentioned, the summing node voltage excursions could exceed the supply rails and be clamped by the substrate diodes. Since neither the MIDAS nor Hspice simulations included these diodes, their impact was not realized until all the simulations were completed.

However in an actual implementation this nonlinear clipping would cause significant distortion to the signal and degrade the SNR.

Rather than perform another complete optimization of the modulator including this effect, some quick calculations were performed to estimate the new power consumption.

The problem stems from the quantizer levels being too large (in our case, $V_{\rm dd}$ & $V_{\rm ee}$). The worst case voltage on the summing node can be:

$$\begin{aligned} V_{\rm sum}({\rm max}) &= V_{\rm in}({\rm max}) - (V_{\rm ref}-) \\ &= 1.7V - (-2.5V) = 4.2V({\rm forInt1}) \\ &= 1.5V - (-2.5V/2) = 2.75V({\rm forInt2}) \end{aligned}$$

Since the substrate diodes turn on at .7V, keeping the voltage swings to within .5V of the supplies (i.e. $\pm$3V) should be adequate. We see that integrator 2 does not overload, but Int1 exceeds the limit by 1.2V.

As a quick solution we change Vref proportionately to scale the 4.2V overload to 3.0V. A simulation by MIDAS and some hand calculations gives results shown in table V.

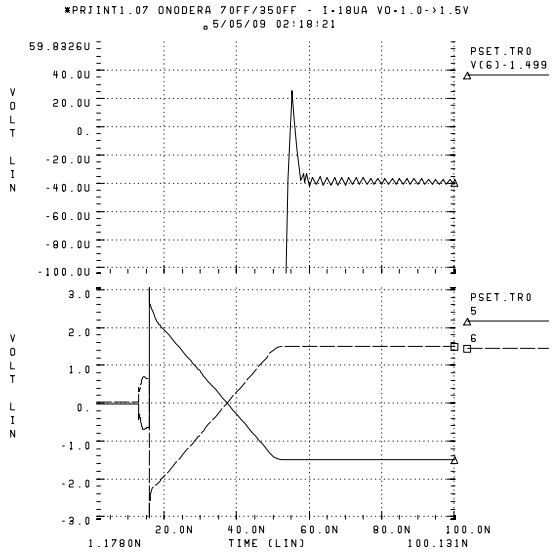

Fig. 9. HSPICE: settling time of first integrator 42 ns

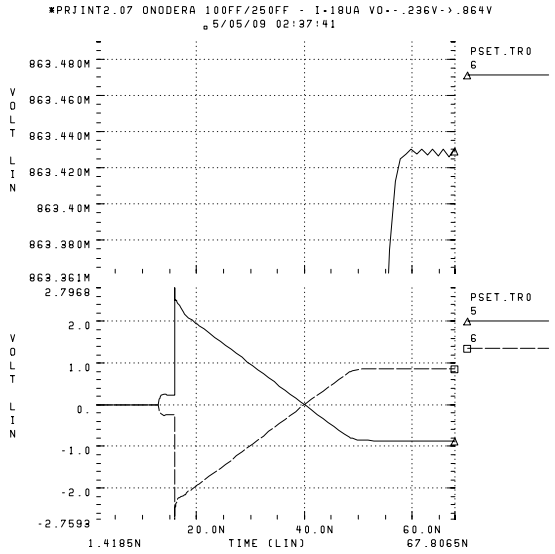

Fig. 10. HSPICE: settling time of second integrator 41 ns

As a quick solution we change Vref proportionately to scale the 4.2V overload to 3.0V. A simulation by MIDAS and some hand calculations gives results shown in table V.

References

[1] B. Boser and B. Wooley, "The Design of Sigma-Delta Modulation Analog-to-Digital Converters", *IEEE J. Solid-State Circuits*, vol. SC-23, pp. 1298–1308, Dec. 1988.

	Unclamped	Clamped
V_{ref} $(\Delta/2)$	2.5V	1.8V
C_{s1}	70fF	121fF
V_{in}	3.4V	2.4V
I(Int1)	38.2μA	66μA
I_{total}	76.4μA	104μA
Power ($V_{\text{supply}} = 5$V)	382μW	521μW

TABLE V

[2] J. Candy and G. Temes, "Oversampling Methods for A/D and D/A Conversion", *in Oversampling Delta-Sigma Data Converters*, 1992.