Pipelined Oversampling Analog-to-Digital Converters

by
Susanne A. Paul

B.S., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1988

M.S., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1995

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy

at the
Massachusetts Institute of Technology
September 1999

Signature of Author _____
Department of Electrical Engineering and Computer Science
June 3, 1999

Certified By _____
Hae-Seung Lee
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted By _____
Arthur C. Smith
Chairman, Committee on Graduate Student

Pipelined Oversampling Analog-to-Digital Converters

by

Susanne A. Paul

Submitted to the Department of Electrical Engineering and Computer Science
on June 3, 1999 in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy

Abstract

Oversampling and noise-shaping techniques, such as $\Delta\Sigma$ modulation, are widely used in analog-to-digital conversion to achieve accuracy that exceeds that of integrated-circuit components. Such converters have an inherent tradeoff between accuracy and speed, whereby resolution in amplitude is achieved at the expense of resolution in time. Although much attention has been focused on improving the speed and power of $\Delta\Sigma$ analog-to-digital converters, data rates remain limited to less than a few MHz and are not easily extended.

A pipelined oversampling architecture is described that circumvents the speed-resolution tradeoff of conventional oversampling converters by performing spatial, rather than temporal, oversampling. It combines high-resolution capabilities of $\Delta\Sigma$ techniques with the high speed of pipelined architectures so that both of these attributes are achievable. The architecture also differs from conventional oversampling in that it performs Nyquist-rate sampling. Power is improved as a result of a charge-domain implementation, reduced sensitivity to thermal noise, simplified decimation, and reduced circuit speed, which permits voltage scaling and use of low-power technologies.

Circuit techniques for implementation of a pipelined oversampling converter are also presented. Although CCDs are not essential to the concept, such converters are most practically built using a combination of CCD and CMOS circuits. CCDs make analog pipelines with hundreds of stages feasible by providing fully-depleted operations which are highly accurate, low power, simple, and compact. Other operations are performed using nondepleted circuits. A circuit technique, referred to as dynamic double sampling, is presented, which provides improved linearity and speed over existing techniques and forms a core circuit element for these nondepleted operations.

Two prototype converters have been demonstrated. They were built in standard CMOS processes and show that moderate to high performance is possible from CCD circuits and can be achieved without custom processing. The first prototype uses a 1.2-μm process and operates at an 18-MHz data rate. It achieves 78-dB SFDR, DNL < ±0.15 LSB at 13 bits, 74-dB SNR over a 9-MHz bandwidth, and 324 mW power dissipation. The second prototype uses a 0.6-μm design rule and operates at a 30-MHz data rate. It achieves 70-dB SFDR and 66-dB SNR over a 15-MHz bandwidth.

Thesis Supervisor: Hae-Seung Lee
Title: Professor of Electrical Engineering and Computer Science

Table of Contents

Table of Figures

1 Introduction

1.1 Motivation

1.1.1 Conventional Oversampling

Oversampled ΔΣ modulation is a technique that is commonly used for implementing high resolution analog-to-digital (A/D) converters. It involves converting an analog input into a coarse digital code at a frequency much higher than the Nyquist rate and then digitally decimating this code to produce a higher resolution, lower data rate result [1][5].

Single-loop systems, such as that in Figure 1-1, are the simplest example of ΔΣ converters. (Tables of block diagram notation and common variable notation are included in Appendices A and B.) The modulator consists of a discrete-time integrator, a coarse analog-to-digital converter (ADC), and a coarse digital-to-analog converter (DAC), connected in a feedback configuration. The ADC and DAC are often 1-bit elements. The modulator's front end performs analog operations, while its output is digital. The decimator is strictly digital and consists of a lowpass filter followed by a downsampler.

The modulator can be analyzed by modeling ideal quantization noise and other sources of non-ideal noise, introduced by the feedforward ADC, as an additive error e_r. The modulator output is then described by the difference equation

$$w_r[n] = s_i[n-1] + e_r[n] - e_r[n-1] \quad . \tag{1-1}$$

Variable w_r contains signal information, unchanged except for a delay, and the 1st-order difference of feedforward ADC quantization noise.

A time-domain view of this process is shown in Figure 1-2. An input greater than half of the full-scale value results in a stream of digital bits with more '1's than '0's. Although large error is introduced into each value of w_r by the 1-bit quantization shown here, feedback ensures that the average, or duty cycle, of w_r tracks that of the analog input s_i. Similarly, a signal less than half of the full-scale value results in more '0's than '1's.

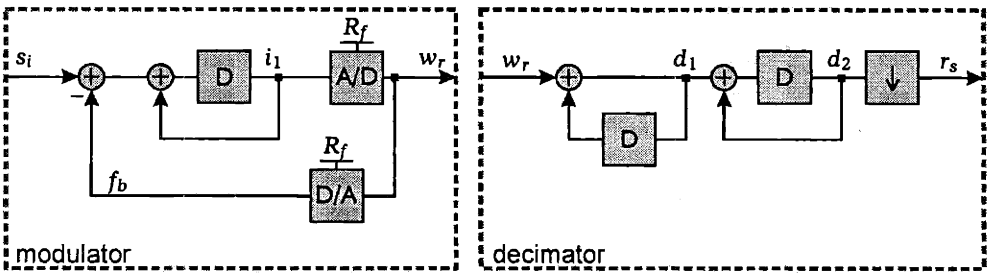


Figure 1-1. First-order ΔΣ architecture contains analog modulator and digital decimator blocks.

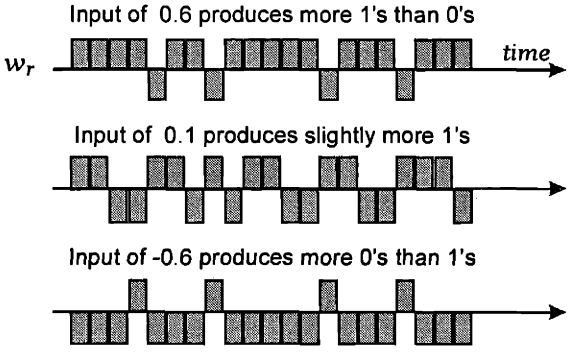


Figure 1-2. Time-domain characteristics of the modulator output w_r. Its duty cycle represents the average analog input.

Differentiation of the quantization noise e_r is one form of noise shaping. A frequency domain representation of this process is illustrated in Figure 1-3. In this simple model, it is assumed that e_r is uniformly distributed in frequency before it experiences noise shaping. At the modulator output, w_r has a z-transform of

$$W_r = S_i z^{-1} + E_r\left(1 - z^{-1}\right) \ . \tag{1-2}$$

The spectral distribution of e_r is shaped as indicated in the figure. Quantization noise energy is reduced at low frequencies but is increased at high frequencies. The response of a simple sinc^2 decimation filter is shown in the figure. Input signals are restricted to frequencies within the decimator passband, typically a small fraction of the total modulator bandwidth. The ratio between modulator sampling frequency and input Nyquist rate is referred to as the oversampling ratio.

The modulator output w_r contains input signal components as well as modulation and circuit noise. The decimation filter attenuates its energy outside of the signal bandwidth so that when w_r is resampled at a lower rate, aliasing does not corrupt the signal. Only noise within the signal bandwidth influences the

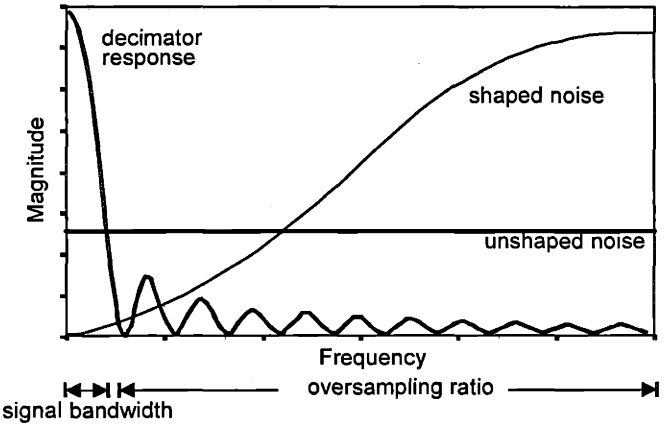


Figure 1-3. Frequency-domain characteristics of $\Delta\Sigma$ modulation. Shaping reduces noise within the signal bandwidth.

converter's result. In the final stages of decimation, downsampling converts the filter output from a series of low resolution, high frequency words to a series of higher-resolution words at the Nyquist rate. Resolution in the converter output words is increased by 9 dB, or 1.5 bits, for every doubling of the oversampling ratio.

The 1st-order architecture described above is the simplest example of an oversampling ΔΣ modulator. It contains only a single feedback loop. However, single-loop modulators require large oversampling ratios to achieve high resolution. The need for a large oversampling ratio is alleviated by use of 2nd or higher-order ΔΣ modulation [9].

A 2nd-order system, such as that in Figure 1-4, contains two nested feedback loops within its modulator. The inner loop is similar to that of a 1st-order modulator, with an integrator, a coarse ADC, and a coarse DAC. The outer loop adds a second analog integrator, cascaded with the first, and an outer feedback path. The decimator is also modified to include a third stage of digital accumulation.

In this architecture, the modulator output has a z-transform of

$$W_r = S_i z^{-2} + E_r\left(1 - z^{-1}\right)^2 , \tag{1-3}$$

which contains delayed signal information plus the 2nd-order difference of quantization and other errors in the feedforward path. In comparison to a first order modulator, this modulator is capable of greater resolution for the same oversampling ratio because its higher-order noise shaping leaves less noise energy within the signal band. Resolution in the converter output word is improved by 15 dB, or 2.5 bits, for every doubling of the oversampling ratio.

1.1.2 Attributes of Conventional Oversampling

ΔΣ techniques offer a number of advantages that make them well suited for high-resolution A/D conversion. First, high resolution is achieved despite inaccuracies of internal circuit components such as the ADC. In contrast, component imperfections generally limit non-oversampling architectures without calibration to about 12-bit resolution. Second, high resolution is achieved from an ADC with only a few levels and these levels do not require great precision. The ADC in a 1-bit system is particularly simple to

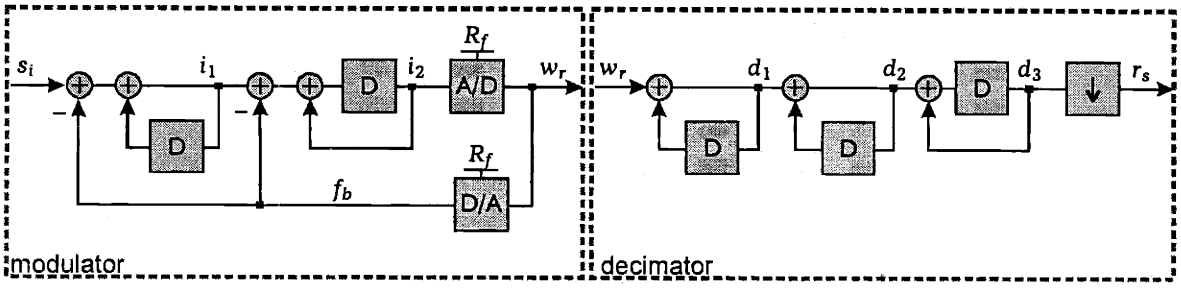


Figure 1-4. Second-order ΔΣ architecture contains two nested feedback loops.

implement because it consists of a single comparator. A third advantage is that high resolution is possible from a DAC with only a few levels. Although full accuracy is required from the placement of these levels, 1-bit systems are capable of excellent linearity because they contain only two DAC levels and variation in the placement of these does not introduce nonlinearity. A final advantage is that anti-alias filtering, at the converter input, is simplified because the decimator serves as a digital anti-alias filter.

The advantages of $\Delta\Sigma$ techniques arise through a series of tradeoffs. First, higher accuracy is achieved at the expense of a drastic increase in the number of operations required. The number of operations performed by non-oversampling architectures, such as successive approximation, increases linearly with converter bits, whereas for $\Delta\Sigma$ converters it increases more rapidly. In a second tradeoff, $\Delta\Sigma$ converters achieve simplicity in their analog circuitry by moving processing complexity into the digital domain. Decimation filters with a flat passband response and a sharp cutoff transition often constitute a majority of the complexity, chip area, and power dissipation of such a device. In a third tradeoff, resolution is improved by a longer decimator impulse response at the expense of increased converter settling time. For this reason, $\Delta\Sigma$ converters require a bandlimited input signal.

Finally, $\Delta\Sigma$ techniques include an inherent tradeoff between speed and resolution, whereby these characteristics are not independently adjustable. An increase in the oversampling ratio improves resolution, but reduces the data rate. Similarly, a decrease in the oversampling ratio improves the data rate at the expense of resolution. Data rates are limited because the modulator must operate over many clock cycles to produce a single result. This speed-resolution tradeoff has been the subject of much attention. Data rates are improved by higher-order modulators [27], but those of greater than second order are susceptible to instability and have increased circuit complexity. Data rates are improved by use of multi-bit quantizers [11][17], but their success has been limited by the accuracy required from multi-bit feedback elements. More recently, multi-bit quantizers with single-bit feedback have been used to reduce oversampling ratios [12] [28], but these devices continue to be limited by a fixed speed-resolution product.

1.1.3 Pipelined Oversampling

Pipelining can be applied to oversampling converters to circumvent their speed-resolution tradeoff and its associated limitations. Conventional and pipelined architectures are compared in Figure 1-5. Conventional oversampling, shown in (a), includes a single modulator operating over many clock cycles to produce a train of digital outputs. Many of these bits are combined within the decimator to generate a single higher-resolution digital result. Each result in a converter with an oversampling ratio of R requires an additional R clock cycles of operation. This type of operation is referred to hereafter as time oversampling.

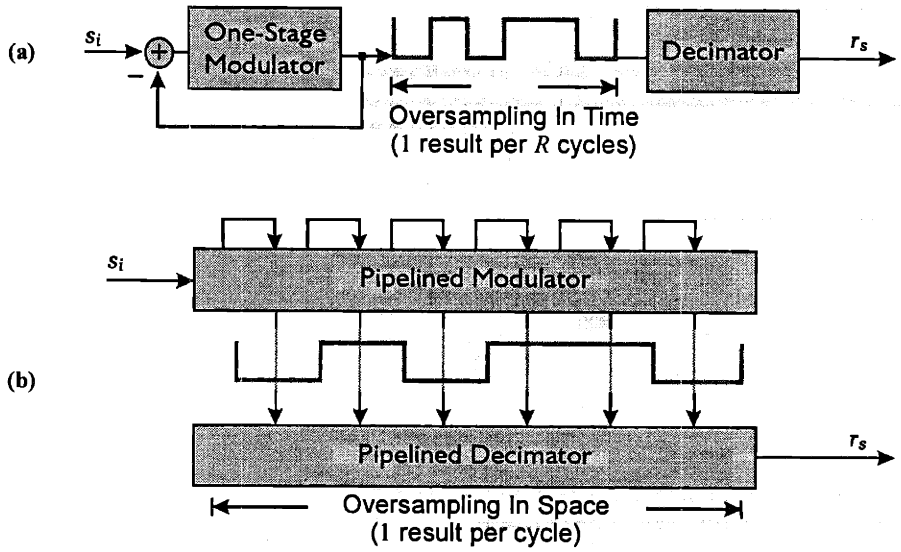


Figure 1-5. Comparison of oversampling techniques. (a) Time oversampling requires many cycles per result. (b) Pipelined approach produces one result per cycle.

The operations performed in (a) can also be implemented in a pipelined fashion. Such a configuration is shown in (b). The analog input is initially sampled by the first pipeline stage. The modulator feedback loop is unraveled into a chain of many stages that operate sequentially on this input sample. The decimator, which is pipelined as well, accepts the modulator outputs at each stage. Signals from all pipeline stages are combined in the decimator to produce a result at the end of the pipeline.

Since oversampling in a pipelined structure occurs sequentially in space rather than sequentially in time, resolution is determined by the pipeline length. The converter data rate is equal to the clock rate at which the pipeline is operated. Since these two parameters are independently adjustable, there is no longer a speed-resolution tradeoff.

Figure 1-6 compares signals in time-oversampling and pipelined devices. The analog input to a time-oversampling converter is shown in (a) as a slowly varying sinewave that is sampled at a frequency much greater than its Nyquist rate. Many input samples are collected during the period for one result. A lowpass filter is applied to samples from a longer time interval, equal to the decimator impulse response length. The input may vary during each of these intervals, although it can not change rapidly.

Lowpass filtering in the decimator for such a device has a number of consequences. First, the decimator serves as a digital anti-alias filter because noise outside of its passband is suppressed. This noise might already be present on the incoming signal or might arise as a result of sampling at the converter input. However, both desirable and undesirable high-frequency components of the input signal are suppressed. Figure 1-7(a) shows an example where lowpass filtering is undesirable. The presampled converter input is shown by the upper waveform. Samples of it, indicated by dots, occur during both plateau and transition

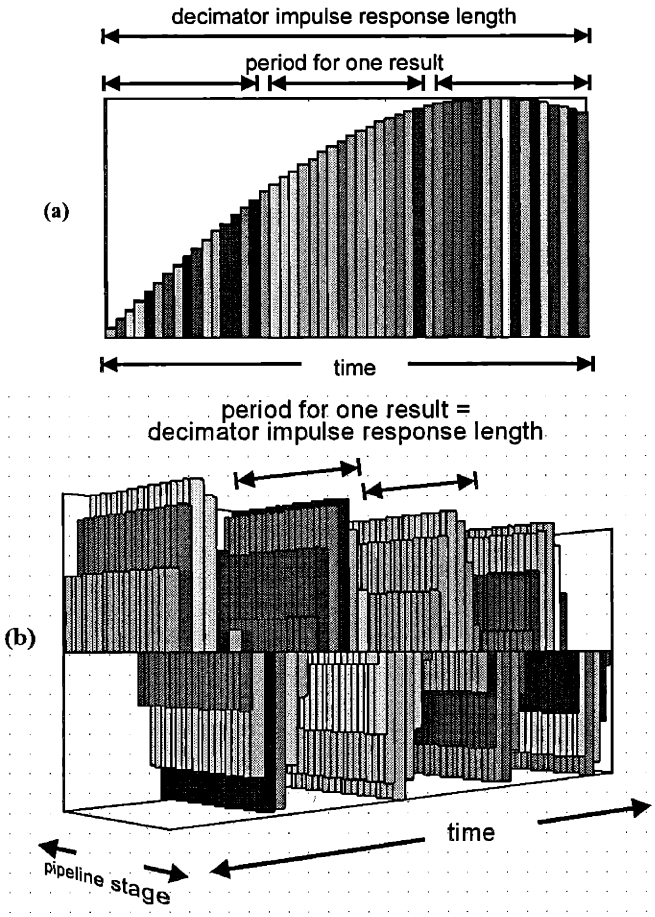


Figure 1-6. Comparison between (a) time-oversampling and (b) pipelined oversampling signals. Pipelined configuration has an additional dimension of space.

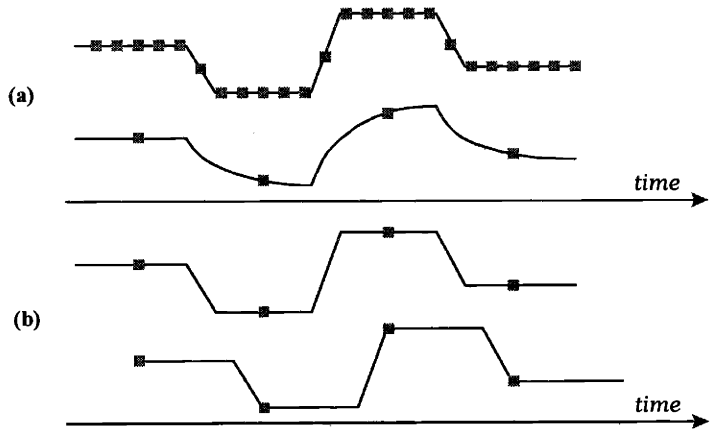


Figure 1-7. Sampling of clocked waveforms. (a) Time oversampling is limited by a long settling time. (b) Pipelined oversampling performs Nyquist sampling.

regions. Since the filter's impulse response length covers multiple plateau regions, there is a mixing between consecutive outputs. The converter output, shown by the lower curve in (a), has a long settling time in response to each step transition. When downsampling is performed, the result does not accurately represent the original input.

Data flow for a pipelined device is shown in Figure 1-6(b). The analog input is shown as a sinewave that is rapidly varying along the time axis. A sequence of samples of this waveform are captured by the first pipeline stage. An additional dimension of space is used to represent position along the pipeline. After each sample is captured, it is passed unchanged along the pipeline and is used by each stage toward computation of a final output. Therefore, from the perspective of the $\Delta\Sigma$ computation, the input is constant.

Computationally a pipelined oversampling converter (POSC) is similar to a time-oversampling converter configured as shown in Figure 1-8. A sampling clock is used to reset integrators in the modulator, reset memory in the decimator, and perform sampling of an analog input. The sampled input is held constant while the modulator and decimator operate on it over many cycles using a higher frequency internal clock. The converter's result is transferred to its output on the next sampling clock edge, at the same time that the next phase of resetting and sampling occurs.

Each output word from a POSC is generated from a single sample of the incoming analog waveform and is processed independently from its neighbors. As a result, a pipelined device performs Nyquist sampling. Figure 1-7(b) illustrates its operation with presampled signals. The converter input, shown by the upper waveform is sampled at the times denoted by dots. One sample is captured during each plateau and each of these is used to compute an output word. There is no mixing between adjacent computations and the converter output, shown by the lower waveform, accurately reproduces its input. The benefits of Nyquist sampling in a POSC do, however, come at a price. As with any Nyquist-sampling converter, an analog anti-alias filter must precede a pipelined converter to prevent high frequency noise from aliasing into the signal passband.

A POSC has a number of unique attributes that differentiate its $\Delta\Sigma$ algorithm from that of a time-

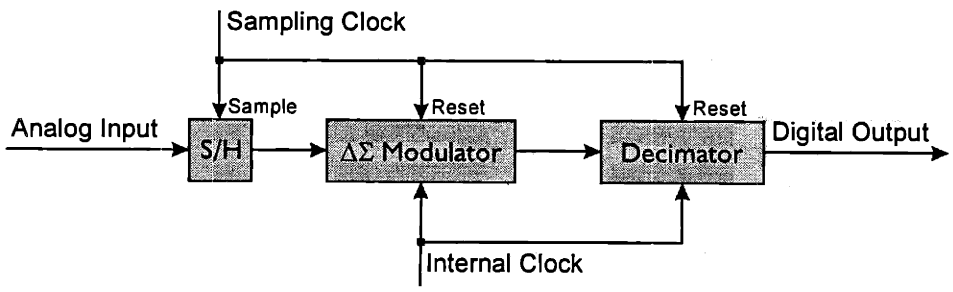


Figure 1-8. Time-oversampling configuration that is computationally similar to a POSC.

oversampling device. These include integrator resetting in the modulator, accumulator resetting in the decimator, an input signal that is effectively constant, an input signal whose form is precisely known, and a decimation filter, whose length is constrained. These aspects are described in detail in the remainder of this chapter.

1.1.4 Analysis of Pipelined Architectures

Theoretical descriptions of ΔΣ A/D conversion are traditionally performed in the frequency domain [27]. However, the analysis in Chapter 2 is performed using a time-domain approach for the following reasons. First, integration of noise energy over the input signal bandwidth is not well defined for a POSC because input signals are effectively constant. Second, a POSC is a transient device and does not naturally have a time-invariant description. Frequency-domain techniques are cumbersome due to resetting at the beginning of the pipeline, truncation at the end of the pipeline, downsampling, and operations that vary as a function of pipeline stage.

The POSC architectures presented in this chapter are described using a common framework. First, their algorithms are presented. These define the computations that occur between each sample at the converter input and its corresponding result at the converter output. For simplicity, algorithms are described using cyclic, rather than pipelined, structures. Variable n, used to denote points in time for a cyclic configuration, is also equivalent to pipeline stage for a pipelined configuration.

Next, pipelined realizations of these algorithms are presented. These describe the structure of an actual device more precisely but are presented without reference to a particular technology or specific circuits. Technology specific design details are described in Chapter 3. Variable n, used to denote pipeline stage, corresponds to the 'space' axis in Figure 1-6(b). Variable m, used to denote sample points along the analog input, corresponds to the 'time' axis in Figure 1-6(b). Operations in pipeline stage n occur n cycles after their corresponding signal was sampled by the first pipeline stage.

2 Pipelined Oversampling Architectures

2.1 First-Order Pipelined Oversampling

2.1.1 Modulation Algorithm

POSC architectures in this section are based on the 1st-order block diagrams of section 1.1.1, but with two modifications. First, integrator i_1 is replaced with a pair of integrators, u_1 and v_1. This is done for illustrative purposes but in general would not be included as part of an actual implementation. Second, each accumulator loop in the decimator is modified to include a delay in its feedforward path. This configuration is more appropriate for a pipelined implementation where each operation occupies one pipeline stage. The resulting block diagram is shown in Figure 2-1.

An incoming analog signal is sampled at the converter input, where n=0. S_g is a constant equal to the value of this sample and the modulator input sequence s_i is equal to

$$s_i[n] = S_g\, u[n] \ , \tag{2-1}$$

where $u[n]$ is a unit step function. Integrator u_1 integrates signal information, via s_i, and integrator v_1 integrates feedback information, via f_b. The difference between their values is similar to integrated quantity i_1 of Figure 1-1. However, as shown, this configuration is single-ended, not differential, and u_1 and v_1 are not a differential pair.

All integrators are preset to zero at time n=0 and their values at cycle n are given by

$$\begin{aligned} u_1[n] &= \begin{cases} 0 & \text{for } n \le 0 \\ u_1[n-1] + s_i[n-1] & \text{for } n \ge 1 \end{cases} \\ v_1[n] &= \begin{cases} 0 & \text{for } n \le 0 \\ v_1[n-1] + f_b[n-1] & \text{for } n \ge 1 \end{cases} \end{aligned} \tag{2-2}$$

The value of u_1 increases linearly with time because S_g is unconditionally added to it each cycle. Since feedback ensures that u_1 and v_1 coarsely track each other, v_1 increases with the same slope.

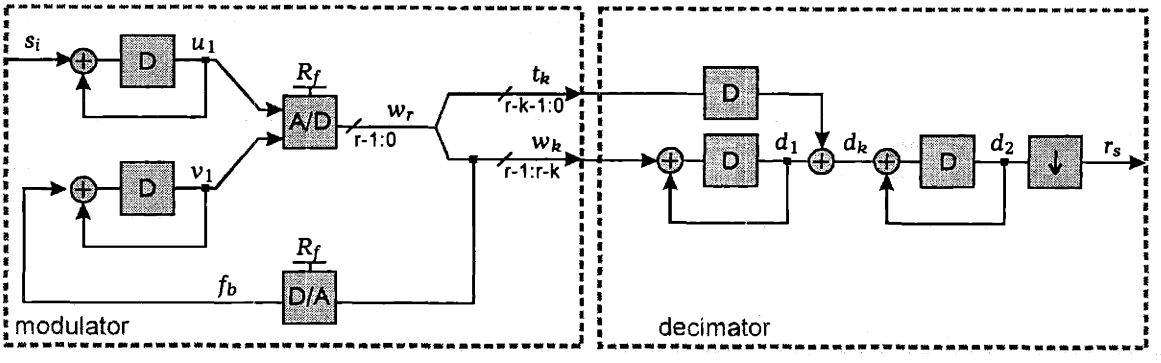


Figure 2-1. First-order POSC algorithm in cyclic form.

A coarse ADC with a full-scale value of R_f generates an r-bit digital representation

$$w_r[n] = \frac{1}{R_f}(u_1[n] - v_1[n]) + e_r[n] \tag{2-3}$$

of the difference between u_1 and v_1. The expression for w_r contains an error term e_r that is due to r-bit quantization as well as circuit inaccuracies in the coarse ADC. As defined in (2-3), e_r is normalized to the full scale.

The k most significant bits of w_r are used to form the first modulator output and the feedback value

$$w_k[n] = \frac{1}{R_f}(u_1[n] - v_1[n]) + e_r[n] - t_k[n] \ . \tag{2-4}$$

This signal contains an additional error term t_k that is due to truncating an r-bit word to k bits. In contrast to error e_r, error t_k is a known quantity and can be compensated for in downstream digital processing. It is provided as an input to the decimator for this purpose. Errors e_r and t_k are both zero for $n<1$ since no comparisons occur during this time.

Node w_k is supplied to a feedback DAC with a full-scale value of R_f that generates a corresponding analog quantity

$$f_b[n] = u_1[n] - v_1[n] + R_f\left(e_r[n] - t_v[n]\right) \ . \tag{2-5}$$

Since the average of f_b approximates the value of S_g, the average of w_r is always greater than or equal to S_g.

Solving equations (2-2), (2-4), and (2-5) recursively yields a time-domain and frequency-domain modulator output of

$$\begin{aligned} w_k[n] &= \frac{S_g}{R_f}u[n-1] + (e_r[n] - e_r[n-1]) - (t_k[n] - t_k[n-1]) \\ W_k(z) &= \frac{S_g}{R_f}\frac{z^{-1}}{\left(1-z^{-1}\right)} + (E_r - T_k)\left(1-z^{-1}\right) \end{aligned} \ . \tag{2-6}$$

Signal w_k contains a delayed version of the input sample plus the 1st-order difference of errors e_r and t_k.

An alternative understanding of $\Delta\Sigma$ operation is evident when it is compared to algorithmic A/D conversion. Both architectures contain a feedback loop and are topologically similar. Both algorithmic and oversampling architectures keep a running estimate (v_1) formed from a known quantity (R_f) that best approximates a known amplification of the input signal (u_1). Both architectures start with small signal amplification when their running estimate is less refined and gradually build to a larger amplification, as this estimate becomes more accurate. They judiciously increase amplification at the same rate that quantization noise is reduced, so that their feedforward ADC input remains small and can be quantized with only a few bits.

The primary difference between successive approximation and ΔΣ oversampling lies in the rate at which their estimates improve and in the rate at which they amplify signal information. Because amplification in an algorithmic device increases exponentially with n, no combination of bits after stage n can reverse an earlier decision, unless error correction techniques are used. Amplification in oversampling increases much more slowly with n so that decisions can easily be reversed in later stages.

2.1.2 Decimation Filtering

The purpose of the decimator is to amplify signal information and attenuate quantization noise in the digital modulator output. One example of such a decimator is shown in Figure 2-1. Its first stage accumulates w_k according to

$$d_1[n] = \begin{cases} 0 & \text{for } n \le 1 \\ d_1[n-1] + w_k[n-1] & \text{for } n \ge 2 \end{cases} . \tag{2-7}$$

Solving equations (2-7) and (2-6) yields the response

$$d_1[n] = \frac{S_g}{R_f}(n-1)u[n-1] + (e_r[n-1] - t_k[n-1]) , \tag{2-8}$$

containing signal information that increases linearly with n at a rate of S_g/R_f, and unamplified error that has a constant mean-square value.

Truncation error is digitally corrected after the first accumulator by adding t_k to d_1. The result,

$$d_k[n] = \frac{S_g}{R_f}(n-1)u[n-1] + e_r[n-1] , \tag{2-9}$$

represents the number of times R_f was added to integrator v_1 over the previous n-1 cycles. Equation (2-9) could be used directly to reconstruct the input sample S_g. However, better resolution is achieved from a second stage of accumulation. The accumulated result

$$d_2[n] = \begin{cases} 0 & \text{for } n \le 2 \\ d_2[n-1] + d_k[n-1] & \text{for } n \ge 3 \end{cases} \tag{2-10}$$

has a value at time n of

$$d_2[n] = \frac{S_g}{2R_f}(n-2)(n-1)u[n-2] + \sum_{i=1}^{n-2} e_r[i] \tag{2-11}$$

and a corresponding z transform of

$$D_2(z) = \frac{S_g}{R_f}\frac{z^{-3}}{\left(1-z^{-1}\right)^3} + (E_r - T_k)\frac{z^{-2}}{\left(1-z^{-1}\right)} . \tag{2-12}$$

Signal information in d_2 is amplified as n^2 while the quantization noise term increases more slowly as $\sqrt{n}$. The value of d_2 is independent of both truncation error t_k and of the number of feedback bits k. It does, however, depend on the number of ADC bits r because this determines the mean-square value of e_r. This

configuration is referred to hereafter as an error-averaging decimator because the final term in (2-11) contains an average of e_r.

The input sample is easily reconstructed from d_2. Because the sequence s_i is constant, equation (2-11) can be solved directly for S_g in terms of d_2 and n. The quantization noise term is assumed to be negligible. For a P-stage pipeline, the final value of d_2 is completed at time $n=P+2$ and reconstruction is given by

$$S_g \approx d_2[P+2]\left(\frac{2R_f}{P^2+P}\right) . \tag{2-13}$$

Equation (2-13) has a particularly simple implementation if the pipeline length P is chosen so that the term $\left(P^2+P\right)$ is a power of two. In this case the result is identically equal to d_2 with a change in significance of its bits. Alternatively, the term $\left(P^2+P\right)$, which only impacts converter gain, could be treated as part of the full-scale value.

2.1.3 Downsampling

In the analysis above, $\Delta\Sigma$ operations are performed continuously for all $n>0$. Finite pipeline length is taken into account in (2-13) by downsampling the value of d_2 at time $n=P+2$. This correctly models a configuration with identical modulator and decimator pipeline lengths. But in general, modulator and decimator lengths could differ and downsampling does not necessarily occur at the end of the modulator pipeline. The following discussion illustrates the concept of downsampling in a POSC and describes optimal placement of this operation.

Consider a general configuration with modulator length P and decimator length Z, where $Z \geq P$. Figure 2-2 illustrates its operation for the example error sequence

$$e_r[n] = \delta[n-1] + \delta[n-P] . \tag{2-14}$$

where $\delta[n]$ is a unit impulse function. A P-stage modulator performs P comparisons, so w_k is nonzero for $1 \leq n \leq P$. Equation (2-6) is modified to account for this truncation as

$$w_k[n] = \left(\frac{S_g}{R_f} + \left(e_r[n]-e_r[n-1]\right) - \left(t_k[n]-t_k[n-1]\right)\right)\left(u[n-1]-u[n-P-1]\right) . \tag{2-15}$$

Signal and error components of this sequence are plotted in (a) and (b). Although e_r is differentiated in w_k for $n<P$, its value at time P is not differentiated as a result of truncating feedback operations.

Equation (2-8) for d_1 is generalized to

$$\begin{aligned} d_1[n] = {} & \frac{S_g}{R_f}\left((n-1)u[n-1]-(n-P-1)u[n-P-1]\right) \\ & +\left(e_r[n-1]-t_k[n-1]\right)\left(u[n-2]-u[n-P-2]\right)+\left(e_r[P]-t_k[P]\right)u[n-P-2] \end{aligned} . \tag{2-16}$$

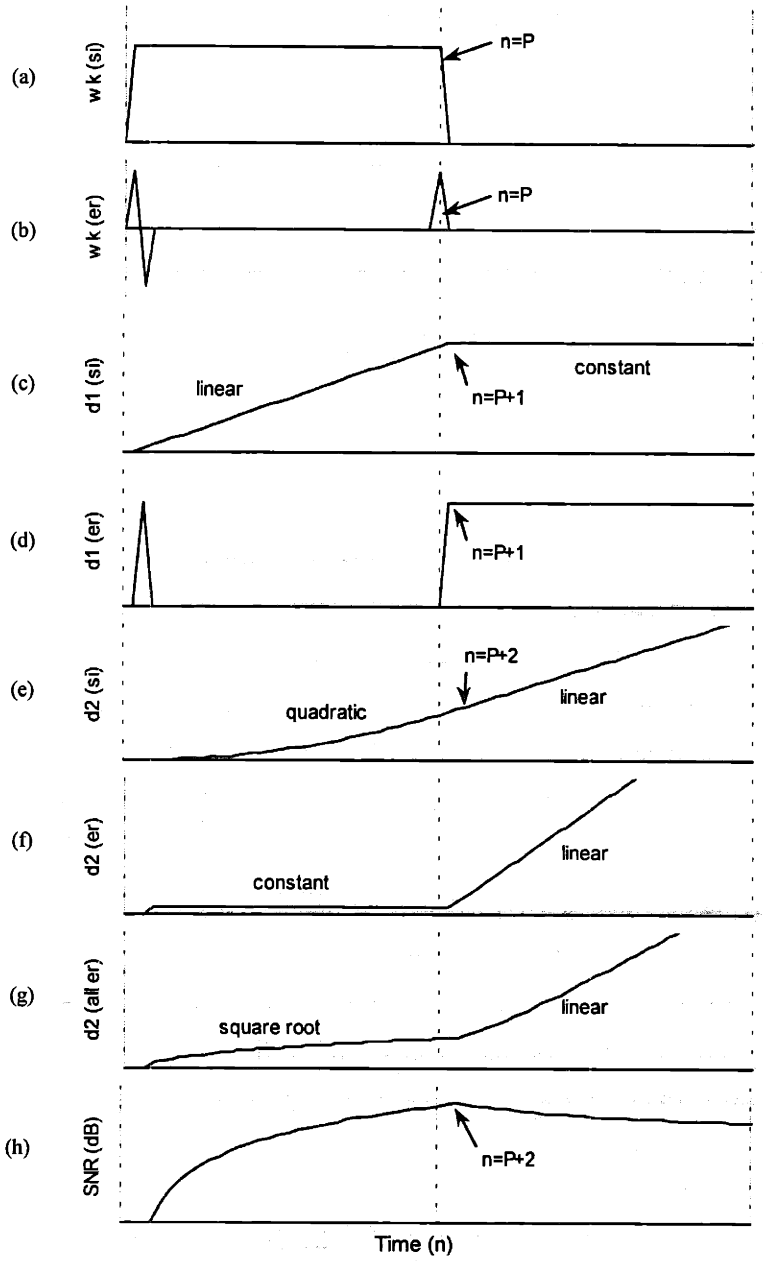


Figure 2-2. Example signal flow for a generalized 1st-order configuration with unequal modulator and decimator lengths. (a), (c), (e) Signal terms of w_k, d_1, and d_2. (b), (d), (f) Error terms of w_k, d_1, and d_2. (g) Averaged error term for general e_r sequence. (h) SNR of filter output.

Its signal component in (c) is a ramp for $n \leq P+1$ and constant for $n > P+1$. Its error component in (d) contains the original e_r sequence for $n \leq P+1$ and a step function, associated with $e_r[P]$ for $n > P+1$. Equation (2-10) for d_2, is generalized to

$$d_2[n] = \frac{S_g}{2R_f}((n-2)(n-1)u[n-2] - (n-P-2)(n-P-1)u[n-P-2]) + \sum_{i=1}^{n-2} e_r[i] + e_r[P](n-P-2)u[n-P-2] \quad (2\text{-}17)$$

Its signal component in (e) increases quadratically for $n \leq P+2$ and linearly for $n > P+2$. Its error component is shown in (f). Because feedback operations are truncated at $n=P$, d_2's error term increases rapidly beyond the end of the pipeline.

The plot in (g) shows the error in d_2 for a case in which e_r is a sequence of white noise. In this case, the sum of e_r has a root-mean-square (rms) value that increases as $\sqrt{n}$ before the end of the pipeline, and linearly thereafter. For best performance, downsampling should occur when the signal-to-noise ratio (SNR) in d_2 is maximized. (SNR is derived in section 2.1.4 and only the results are included here). SNR, equal to the ratio of signal terms in (e) to error terms in (g), is plotted on a logarithmic scale in (h). Maximum SNR occurs at $n=P+2$, after P stages in both the modulator and decimator pipelines. Outputs from the filter are only used at this point and the decimator does not need to compute any other values along the curve in (g).

The example above illustrates a more general fact that for optimal performance, decimator and modulator pipeline lengths should be the same. A longer decimator does not improve, and even degrades, converter resolution. A shorter decimator ignores later modulator outputs and also degrades resolution. POSC architectures in the remainder of this chapter are assumed to have identical modulator and decimator pipeline lengths and downsampling at the end of their decimator pipeline.

2.1.4 Converter Resolution

The analysis above can be used to arrive at an expression for POSC resolution as a function of pipeline length P and number of ADC bits r. To do this, an additional dimension of time must be considered. The variable n, used above, corresponds to time steps along the $\Delta\Sigma$ computation for a single input sample. It does not represent time along a sequence of input samples. This is denoted by variable m. For example, $s_i[n=0,m]$ is the converter's incoming analog signal at time m and $r_s[n=P+2,m]$ is its associated output word, generated from pipeline stage $n=P+2$, in clock cycle $m+P+2$.

The relevant measure of SNR for a POSC is that for its downsampled output sequence r_s over time m. At each point m, SNR is defined as the ratio of rms signal and error terms. Modulator overloading effects are described in section 2.4.9. The largest input range that a 1st-order modulator can accommodate without

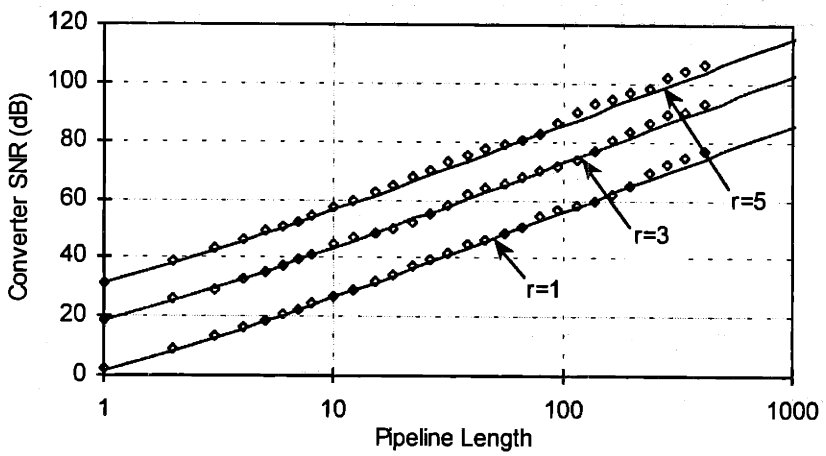


Figure 2-3. First-order POSC resolution versus pipeline length. Lines show theoretical SNR. Diamonds show simulated SNR. SNR improves by 1.5 bits for every doubling of the pipeline length.

A well-known aspect of 1st-order time-oversampling $\Delta\Sigma$ architectures is that each doubling of the oversampling ratio provides an additional 1.5 bits of resolution [1]. For a pipelined architecture a similar result is achieved by evaluating

$$\frac{\mathrm{SNR}[2P]}{\mathrm{SNR}[P]} = \frac{2^{3/2}P^{3/2} + 2^{1/2}P^{1/2}}{P^{3/2} + P^{1/2}} \approx 2^{3/2} \qquad \text{for } P >> 1 \ . \tag{2-22}$$

In a POSC, every doubling of the pipeline length provides an additional 1.5 bits of resolution.

2.1.5 Frequency-Domain Description

Although time-domain analysis provides the most natural description of POSC architectures, a frequency-domain approach offers insight as well. Such an approach differs from that of time-oversampling converters. Since a POSC is a transient device, effects such as integrator resetting, finite pipeline length, and downsampling of the response at the end of the pipeline must be accounted for. In the time-domain analysis above, integrators and accumulators are modeled with infinite impulse response lengths and finite pipeline length is accounted for by downsampling. In a frequency-domain description, integrators and accumulators are modeled with finite impulse response lengths to ensure that their responses converge and their Fourier transforms exist. The description that is chosen accurately models signals before the end of the pipeline but contributes a nonphysical response after the end of the pipeline. It is an acceptable model because details occurring after the end of the pipeline take place after downsampling and do not affect the downsampled result.

Variable n is used here to denote time along the pipeline. The dual of variable n is frequency ω. However, variables n and ω have an unusual significance since they relate to a single input sample. They do not describe time or frequency of signals at the input or output of the converter.

overloading is $R_f\left(1-2^{-k}\right)$. A sinusoidal input is assumed with this peak-to-peak span and an rms value about its mean of $R_f\left(1-2^{-k}\right)/2\sqrt{2}$.

The mean-square value of e_r is given by

$$\overline{e_r^2[n]} = \frac{2^{-2r}}{12} + \overline{q_n^2[n]} \ . \tag{2-18}$$

Its first term represents ideal quantization noise from the coarse r-bit ADC and its second term is due to non-ideal circuit inaccuracies of elements within the feedforward modulator path. Error in r_s is given by the summation in (2-11) of e_r for $1 \le n \le P$. Knowledge of the spectral distribution of these errors is required to evaluate the mean-square value of their sum. Conventional $\Delta\Sigma$ analysis is based on an assumption that errors are uncorrelated and white. This assumption simplifies analysis and provides an accurate description for most configurations, even when errors are not strictly white [8]. However, certain configurations, such as those with 1st-order modulation, single-bit feedforward ADCs, and slowly varying input signals, are more likely to have reduced performance as a result of correlated errors. Such effects are considered in section 2.4.

Under the assumption of uncorrelated noise, the error term in (2-11), evaluated at $n=P+2$, has an rms value of

$$\sqrt{\overline{r_s^2[m]}} = \sqrt{P}\sqrt{\frac{2^{-2r}}{12} + \overline{q_n^2[n]}} \qquad \text{(error term only)} \ . \tag{2-19}$$

It increases as the square root of the pipeline length. Random circuit noise q_n is significant when r is large. However, in most cases r is small and quantization noise dominates the above sum. This case is considered below.

The signal term in (2-11), evaluated at $n=P+2$, has an rms value of

$$\sqrt{\overline{r_s^2[m]}} = \frac{\left(1-2^{-k}\right)}{4\sqrt{2}}\left(P^2+P\right) \qquad \text{(signal term only)} \ , \tag{2-20}$$

which increases quadratically with pipeline length. SNR, as determined by the ratio of (2-20) to (2-19) for $q_n=0$, is given by

$$\mathrm{SNR}[P] = \frac{2^r\left(1-2^{-k}\right)\sqrt{6}}{4}\left(P^{3/2}+P^{1/2}\right). \tag{2-21}$$

Each additional bit in the r-bit feedforward ADC improves resolution by 1 bit. The number of DAC feedback bits k impacts resolution in that it determines the useful input signal range. The result in (2-21) is plotted as a function of pipeline length in Figure 2-3 for $r=1$, 3, and 5 and $k=r$. Theoretical results are shown as solid lines. Simulated results are shown as diamonds.

Consider the configuration of section 2.1.3, with equal modulator and decimator lengths of P. Downsampling occurs immediately after the end of the decimator pipeline to assure convergence of the Fourier transforms. The modulator output in (2-6) is changed to the finite-impulse-response description

$$w_k[n] = \frac{S_g}{R_f}(u[n-1]-u[n-P-1]) + (e_r[n]-e_r[n-1]) - (t_k[n]-t_k[n-1]) \ . \tag{2-23}$$

The Fourier transform of w_k equals

$$W_k(\omega) = \frac{S_g}{R_f}\frac{e^{-(P+1)j\omega/2}\sin(P\omega/2)}{\sin(\omega/2)} + (E_r(\omega)-T_k(\omega))e^{-j\omega/2}e^{j\pi/2}2\sin(\omega/2) \ . \tag{2-24}$$

Its magnitude is plotted in Figure 2-4 for an example of P=16. The noise component is multiplied by $\sin(\omega/2)$ due to differentiation. The signal component has an envelope that decreases with an inverse relationship as $1/\sin(\omega/2)$.

The spectrum in Figure 2-4 differs from that of a time-oversampling converter. In time oversampling, a narrow passband is designated for signal information and incoming signals must be bandlimited. As a result of integrator resetting, s_i is effectively a step function and has a spectrum that is not bandlimited. Such an input would result in poor performance from a time-oversampling device. However, it is ideally suited for a pipelined device. Another unique aspect of Figure 2-4 is that the spectral distribution of s_i is signal independent to within the scaling factor S_g. This can be seen from equation (2-24). The purpose of filtering is to extract the value of this scale factor. However, filtering does not need to preserve the spectral distribution of s_i. Matched filtering techniques, described later, make use of this unusual characteristic.

Accumulators in the decimator are modeled using the finite-impulse-response description

$$H(\omega) = \frac{\sin(P\omega/2)}{\sin(\omega/2)} \ . \tag{2-25}$$

Equations (2-8) and (2-11), are modified to

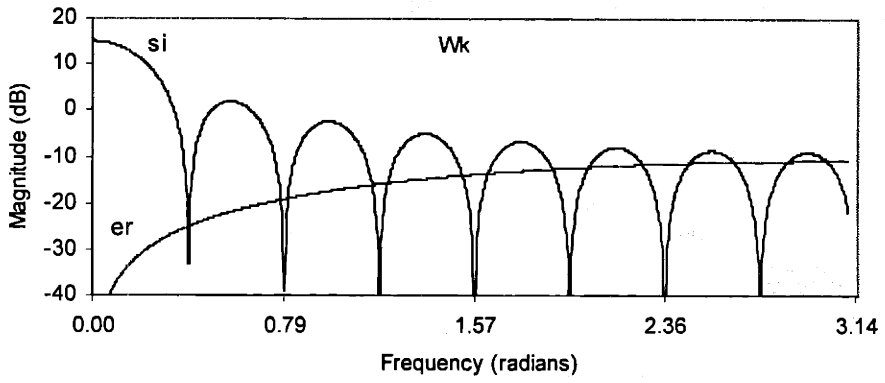


Figure 2-4. Spectrum of w_k. Signal and noise components have inverse amplification.

$$\begin{aligned} d_1[n] &= d_1[n-1] + w_k[n-1] - w_k[n-P-1] \\ d_2[n] &= d_2[n-1] + d_k[n-1] - d_k[n-P-1] \end{aligned} \quad . \tag{2-26}$$

The Fourier transforms for d_1 and d_2 are

$$D_1(\omega) = \frac{S_g}{R_f}\left(\frac{e^{-(P+1)j\omega}\sin^2(P\omega/2)}{\sin^2(\omega/2)}\right) + \left(E_r(\omega) - T_k(\omega)\right)e^{-j\omega}e^{-Pj\omega/2}2e^{j\pi/2}\sin(P\omega/2) \tag{2-27}$$

and

$$D_2(\omega) = \frac{S_g}{R_f}\left(\frac{e^{-3(P+1)j\omega/2}\sin^3(P\omega/2)}{\sin^3(\omega/2)}\right) + E_r(\omega)\frac{e^{-(P+3/2)j\omega}2e^{j\pi/2}\sin^2(P\omega/2)}{\sin(\omega/2)} \quad . \tag{2-28}$$

These results are plotted in Figure 2-5. Downsampling is performed to select the value of d_2 at time n=P+2. In the time-domain this operation is described by

$$r_s[m] = d_2[n]\,\delta[n-P-2] \quad . \tag{2-29}$$

In the frequency domain, the spectrum of d_2 is convolved with that of a delayed impulse. The result

$$r_s[m] = \int_0^{2\pi} d\omega \left[\frac{s_i[m]}{R_f}\frac{e^{-(P-1)j\omega/2}\sin^3(P\omega/2)}{\sin^3(\omega/2)} + E_r(\omega)\frac{e^{j\omega/2}2e^{j\pi/2}\sin^2(P\omega/2)}{\sin(\omega/2)}\right] \quad . \tag{2-30}$$

is equivalent to integrating the product $\left(D_2\, e^{-j\omega(P+2)}\right)$ over all frequencies. Equation (2-30) provides a direct relationship between the converter's input and output as a function of time m, without any

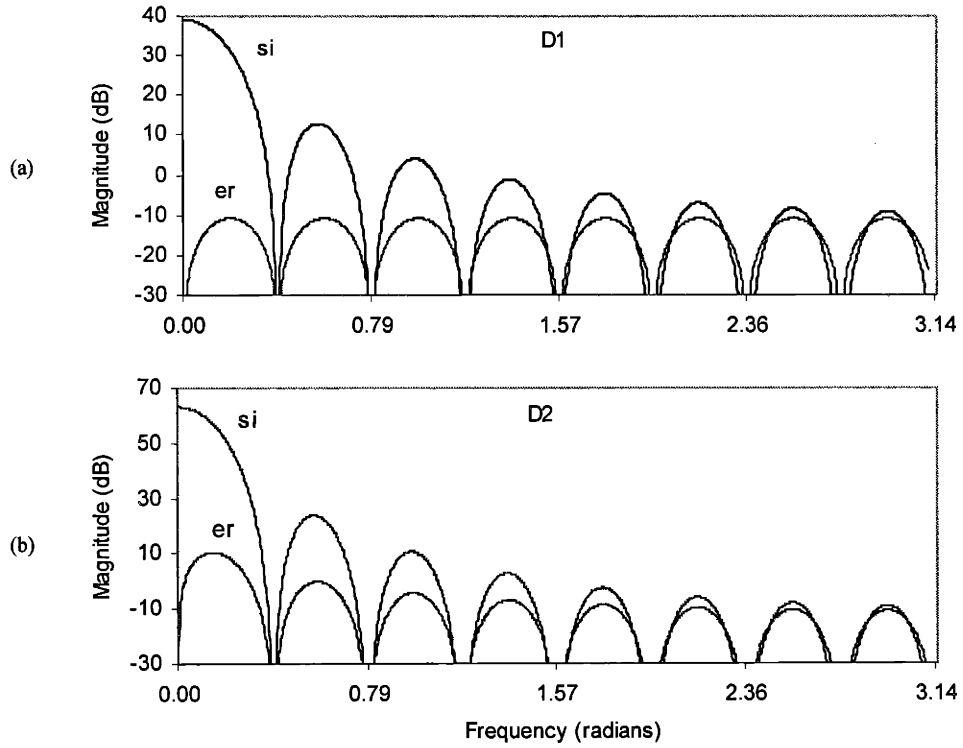


Figure 2-5. Spectrum of signal and noise terms in d_1 (a) and d_2 (b). Each frequency is amplified in proportion to its SNR.

dependency on ω or n.

The ratio between signal and noise terms in w_k at the decimator input is proportional to $1/\sin^2(\omega/2)$. The two stages of accumulation in the decimator multiply w_k by this same factor. The result after downsampling is an integral, over all ω, of w_k multiplied at each frequency by its SNR at that point.

2.1.6 Matched-Filter Decimation

A matched-filter decimator provides better performance than an error-averaging decimator. Matched-filtering techniques can be applied to systems, such as a POSC, meeting two criteria. First, the form of the modulator input must be known so that the decimator can attenuate frequencies having little relative signal energy and amplify those where relative signal energy is high. Second, distortion of the modulator input must be acceptable since these techniques do not provide a flat passband response within the decimator.

A POSC meets the above criteria for matched filtering. The first criterion, when applied to a pipeline device, states that the form of s_i must be known as a function of pipeline stage n. No knowledge is necessary about s_i as a function of time m. The time response and frequency spectrum of this signal depend only on the input sample and scaling factor S_g. Otherwise it is precisely known.

The second criterion is met by a POSC because the decimator produces only a single value for each pipeline sequence. The downsampled value r_s depends only on the total signal energy, not on its time or frequency distribution. Although d_2 is a highly distorted version of the decimator input w_k, distortion is introduced as a function of n and does not impact r_s as a function of time m.

A matched-filter configuration for a 1st-order POSC is shown in Figure 2-6. The structure in (a) replaces the entire decimator block in Figure 2-1. For simplicity, equal numbers of ADC and DAC bits are

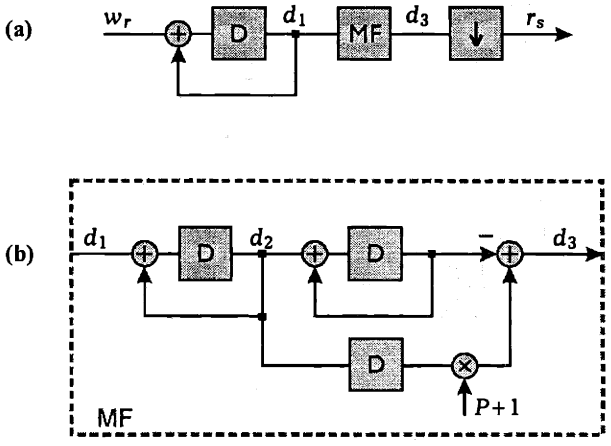


Figure 2-6. Improved decimator for a 1st-order POSC has one stage of accumulation followed by a matched filter. (a) Decimator configuration. (b) Contents of MF block.

assumed so that truncation error is zero. However, the techniques described here apply identically to a configuration with nonzero truncation error, provided this error is digitally cancelled before the matched filter.

Operation of a matched-filter decimator is described with reference to Figure 2-7. The modulator output in (a) is identical to that of (2-6). It contains delayed signal information and differentiated noise. After one stage of integration, d_1 has a signal component in (b) that increases linearly with a slope of S_g/R_f and has a noise component that is white. According to matched-filter theory, the greatest SNR that is achievable from a linear time-invariant filter occurs with a filter whose impulse response is a time-inverted version of the anticipated input. Or expressed differently, the filter weights as a function of time should match those of the signal. The MF block in Figure 2-6(b) represents such a filter. It has an impulse response shown in (c) that is given by

$$h[n] = (P + 2 - n)u[n - 2] \ . \tag{2-31}$$

Convolution of sequences (b) and (c) yields the filter output in (d). Downsampling occurs at the point of maximum SNR in the filter output. The plot in (e) confirms that with a matched filter, as for decimators described earlier, this point occurs at n=P+3, where modulator and decimator pipeline lengths are the

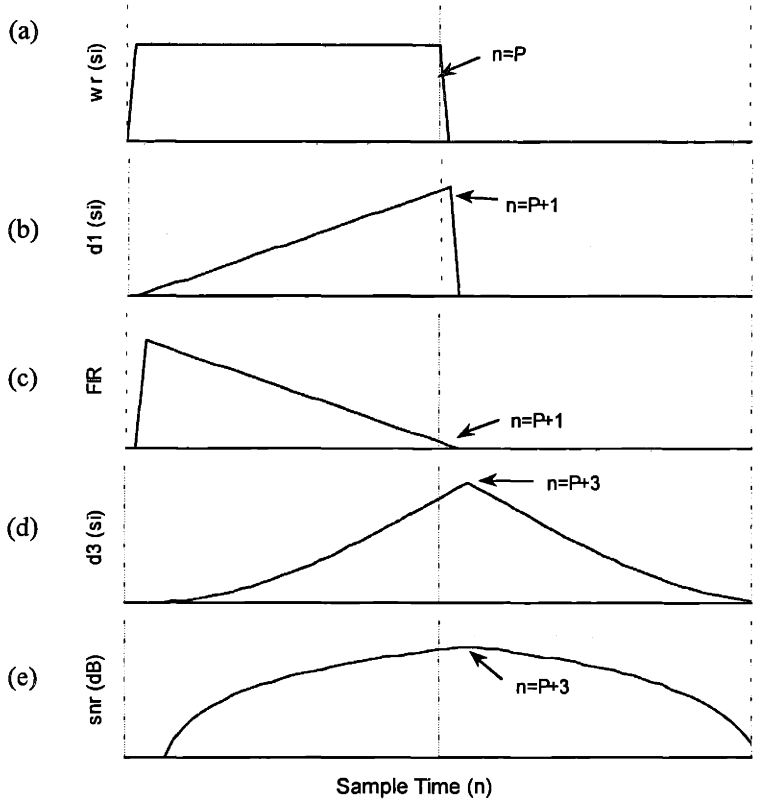


Figure 2-7. Matched filter operation for 1st-order modulation. (a), (b), (d) Signal components of w_r, d_1, and d_3. (c) Matched filter impulse response. (e) SNR in the filter output.

same. This value of d_3 is used to form the converter result and it is the only value along the curve in (d) that is computed.

Figure 2-8 illustrates the reason behind a matched filter's optimal performance. Since the sequence d_1 has noise amplitude that is uniform over n, its SNR increases linearly with n. Later values of w_k have a greater SNR than earlier ones. The matched filter weights also increase linearly with n. At the point of downsampling, sequences are aligned as shown in the figure, with maximum overlap. They are multiplied at each point and summed over all n. This filter provides optimal SNR because each point in the d_1 sequence is weighted by its SNR.

The resolution of a matched-filter decimator is derived in a similar way to that in section 2.1.4. The downsampled filter response at time $n=P+3$is given by

$$r_s[m]=\sum_{n=2}^{P+1}\left[(n-1)^2\frac{S_g}{R_f}+(n-1)e_r[n-1]\right] . \tag{2-32}$$

For a sinusoidal input with a peak-to-peak range of $R_f\left(1-2^{-k}\right)$, signal and quantization noise terms in (2-32) are given by

$$r_s[m]=\frac{\left(1-2^{-k}\right)}{12\sqrt{2}}(P)(P+1)(2P+1) \quad \text{(signal term only)} \tag{2-33}$$

and

$$r_s[m]=\frac{2^{-r}}{\sqrt{12}}\sqrt{\frac{1}{6}(P)(P+1)(2P+1)} \quad \text{(error term only)} . \tag{2-34}$$

The signal term is nearly equal to the square of the quantization noise term. Their ratio,

$$\text{SNR}[P]=\frac{2^r\left(1-2^{-k}\right)}{2}\cdot\sqrt{(P)(P+1)(2P+1)}\approx\frac{2^r\left(1-2^{-k}\right)}{\sqrt{2}}P^{3/2} \quad \text{for } P>>1 , \tag{2-35}$$

gives SNR as a function of pipeline length P, number of quantizer bits r, and number of feedback bits k. Each doubling of the pipeline length improves SNR by the factor

$$\frac{\text{SNR}[2P]}{\text{SNR}[P]}\approx 2^{3/2} \quad \text{for } P>>1 , \tag{2-36}$$

and provides an additional 1.5 bits of resolution. A comparison of equations (2-35) and (2-36) for a

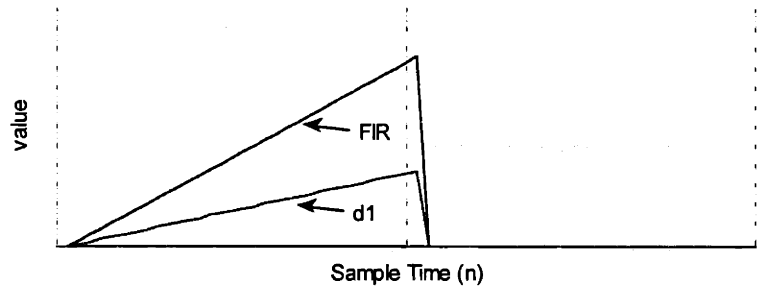


Figure 2-8. Matched filter output is computed by weighting each point by its SNR.

matched-filter decimator with (2-21) and (2-22) for an error-averaging filter shows that both approaches provide an additional 1.5 bits for every doubling of the pipeline length. However the matched filter provides an additional factor of $\sqrt{4/3} \approx 1.3$ dB improvement in resolution. Section 2.4.1 compares the performance of these decimation techniques to that of time-oversampling decimation.

2.1.7 Pipelined Implementation

The term 'pipeline stage' is used above to describe both the hardware comprising one feedback operation and the hardware exercised during one clock period. This terminology must be modified here because the pipelined structures used to implement a POSC often utilize multiple clock cycles to perform one feedback operation. To differentiate between these, the former is referred to as a 'conversion block' and the latter is referred to as a 'stage'. Using this convention, a conversion block might contain multiple stages.

A pipelined implementation of the algorithm in Figure 2-1 is shown in Figure 2-9. The pipeline in (a) contains P conversion blocks. Two analog channels, s_i and i_1, flow through the pipeline. Two digital channels, d_1 and d_2, carry decimator signals. Before the first conversion block, the states of i_1, d_1, and d_2 are all set to zero. The analog converter input is sampled by the first block and is passed unchanged along the pipeline. The final converter result is generated by normalizing d_2 from the last block using equation (2-13). Multiple input samples are processed in parallel along the pipeline so that a new digital word is completed at each clock cycle.

The contents of a single conversion block are shown schematically in (b). Subscripts of n are used to denote outputs from the nth block. Each element labeled D performs delay and latching operations. The ADC element is a coarse r-bit A/D converter, with a full-scale reference of R_f, that is implemented using a non-oversampling architecture such as one-bit-per-stage successive approximation. The DAC element represents a coarse k-bit D/A converter, with a full-scale reference R_f, that is also implemented using a non-oversampling architecture.

Each block accepts feedback signal w_k from its predecessor and generates a corresponding analog quantity that is subtracted from the delayed input sample s_i. The result, after subtraction, is added to integrator i_1. An r-bit quantized representation of this integrated value is divided into upper and lower bits. The k upper bits form the feedback signal w_k. These are added to the first decimator channel and are passed forward to the next block. The only digital signal from block n that is needed by block $n+1$ is w_k. Signals d_1 and d_2 are not needed until the end of the pipeline and their computations can, therefore, be reordered or spread across multiple conversion blocks. An example of this is shown in the figure. The delay shown in Figure 2-1 for t_k is eliminated and t_k is added directly to d_2. At the completion of these operations all results are passed forward to a subsequent conversion block.

The configuration in Figure 2-9 is not well suited for high-speed operation because each conversion block contains only one delay. All of its operations, including D/A conversion, addition, subtraction, and A/D conversion, must be completed within a single clock period. Throughput is improved if the conversion

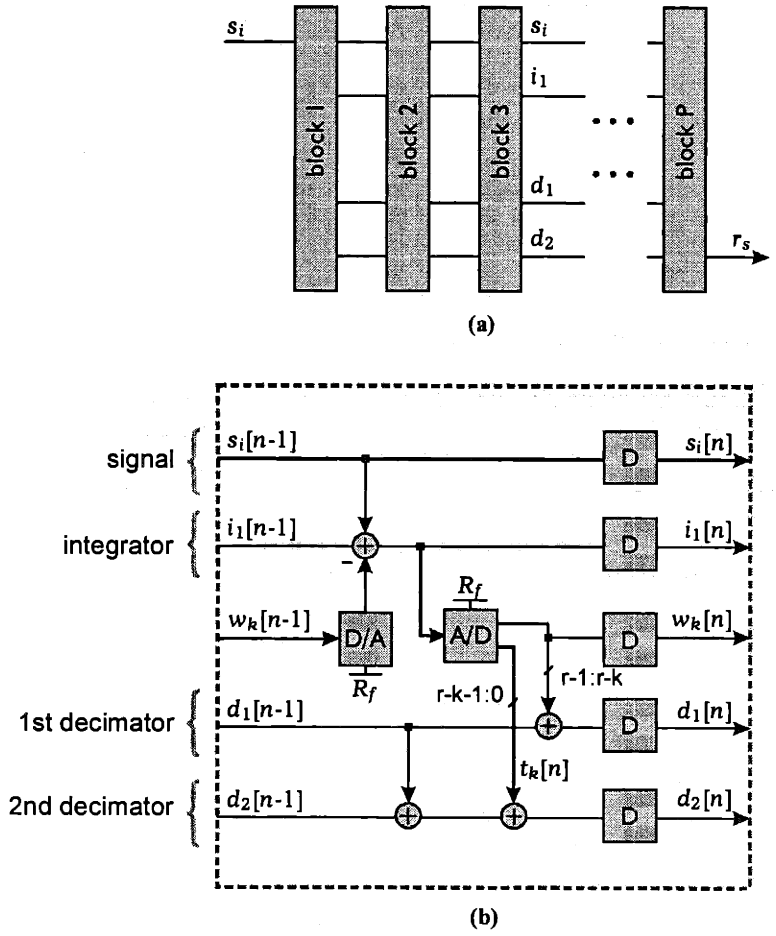


Figure 2-9. Pipelined implementation of a 1st-order POSC. (a) Pipeline contains P identical conversion blocks. (b) Contents of the nth block. Each block implements one feedback operation.

block configuration in Figure 2-9 is divided into multiple pipeline substages as shown in Figure 2-10. In this example, operations are spread over four substages so that each feedback operation is allowed multiple clock cycles to complete. Fewer operations are performed per stage, within a single clock cycle, and a higher frequency clock can be used. Latency is increased in a subpipelined approach, but throughput remains one result per cycle.

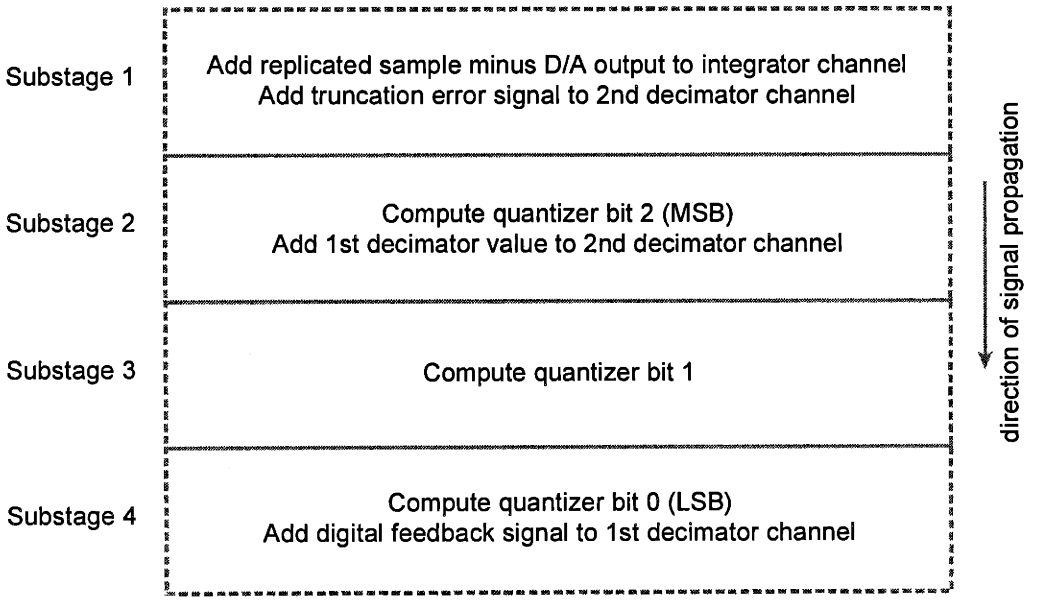


Figure 2-10. Example multi-stage conversion block. Speed is improved by spreading each feedback operation over many substages.

2.2 Second-Order Pipelined Oversampling

2.2.1 Modulation Algorithm

The 1st-order POSC architecture, described in section 2.1, can be modified as shown in Figure 2-11, to accommodate 2nd-order modulation. The configuration shown here differs from the time-oversampling configuration of Figure 1-4 in three ways. First, for illustrative purposes, integrators i_1 and i_2 are replaced with pairs of integrators, u_1/v_1 and u_2/v_2. Second, delays are included in the feedforward paths of all integrators and accumulators so that each operation occupies one pipeline stage. Third, a gain of two is added into the inner feedback loop to compensate for the additional delay from f_b to v_1. Algorithms are illustrated here using a cyclic configuration. A pipelined realization of these algorithms is described in section 2.2.6.

At the converter input, an incoming analog signal is sampled. S_g is a constant used to denote the value of this sample. It forms the input to the $\Delta\Sigma$ algorithm

$$s_i[n] = S_g\, u[n] \quad . \tag{2-37}$$

Integrators u_1 and u_2 accumulate signal information via s_i and integrators v_1 and v_2 accumulate feedback information via f_b. The difference between their values is similar to integrated quantities i_1 and i_2 of Figure 1-4.

The first and second integrators are preset to zero at time n=0 and their values at cycle n are given by

$$\begin{aligned} u_1[n] &= \begin{cases} 0 & \text{for } n \le 0 \\ u_1[n-1] + s_i[n-1] & \text{for } n \ge 1 \end{cases} \\ v_1[n] &= \begin{cases} 0 & \text{for } n \le 0 \\ v_1[n-1] + f_b[n-1] & \text{for } n \ge 1 \end{cases} \end{aligned} \tag{2-38}$$

and

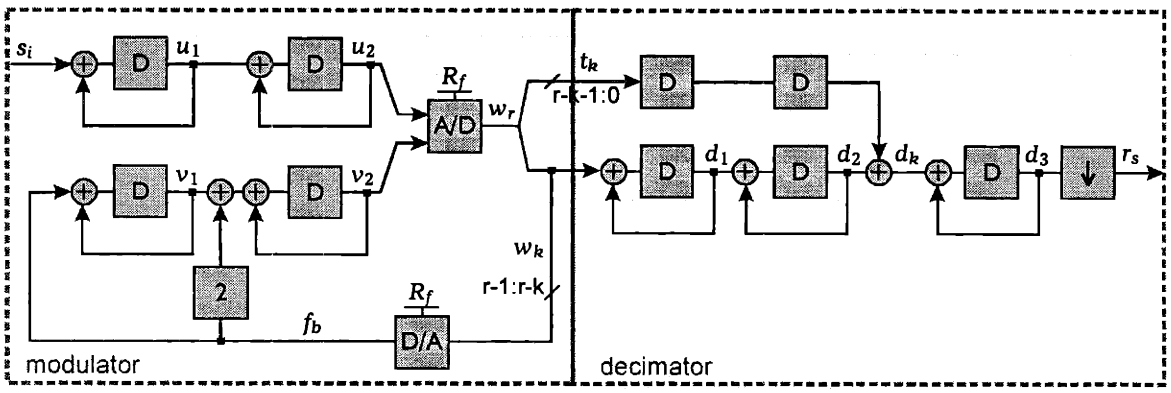


Figure 2-11. Second-order POSC algorithm in cyclic form.

$$u_2[n]=\begin{cases}0 & \text{for } n\le 1\\ u_2[n-1]+u_1[n-1] & \text{for } n\ge 2\end{cases}$$
$$v_2[n]=\begin{cases}0 & \text{for } n\le 1\\ v_2[n-1]+v_1[n-1]+2f_b[n-1] & \text{for } n\ge 2\end{cases} \quad . \tag{2-39}$$

The values of u_1 and u_2 increase linearly and quadratically with time. Feedback ensures that v_1 tracks u_1 and v_2 tracks u_2.

A coarse ADC with a full-scale value of R_f generates an r-bit digital representation

$$w_r[n]=\frac{1}{R_f}(u_2[n]-v_2[n])+e_r[n] \tag{2-40}$$

of the difference between u_2 and v_2. The error term e_r in the ADC output contains r-bit quantization noise as well as circuit inaccuracies.

The modulator's output,

$$w_k[n]=\frac{1}{R_f}(u_2[n]-v_2[n])+e_r[n]-t_k[n] \ , \tag{2-41}$$

consists of the k most significant bits of w_r. It also forms the feedback value. The additional term t_k is due to truncating an r-bit word to k bits and can be digitally corrected for downstream in the decimator. Errors e_r and t_k are zero for $n<1$ since no comparisons occur during this time.

Node w_k is supplied to a feedback DAC with a full-scale value of R_f. The DAC generates a corresponding analog quantity

$$f_b[n]=u_2[n]-v_2[n]+(e_r[n]-t_k[n])R_f \ . \tag{2-42}$$

An iterative solution of equations (2-38), (2-39), (2-41), and (2-42) yields a modulator output of

$$w_k[n]=\frac{S_g}{R_f}u[n-2]+(e_r[n]-2e_r[n-1]+e_r[n-2])+(t_k[n]-2t_k[n-1]+t_k[n-2]) \tag{2-43}$$

and a z-transform of

$$W_k(z)=\frac{S_g}{R_f}\frac{\left(z^{-2}-z^{-P-2}\right)}{\left(1-z^{-1}\right)}+(E_r-T_k)\left(1-z^{-1}\right)^2 \ . \tag{2-44}$$

The modulator output contains a delayed version of the input sample, plus the 2nd-order difference of errors e_r and t_k.

2.2.2 Decimation Filtering

The decimator processes the modulator output w_k to amplify signal information and attenuate quantization noise. Its first and second accumulators are given by

$$d_1[n]=\begin{cases}0 & \text{for } n\le 2\\ d_1[n-1]+w_k[n-1] & \text{for } n\ge 3\end{cases}$$
$$d_2[n]=\begin{cases}0 & \text{for } n\le 3\\ d_2[n-1]+d_1[n-1] & \text{for } n\ge 4\end{cases} \quad (2\text{-}45)$$

Solving equations (2-41) and (2-45) yields responses in the decimator of

$$d_1[n]=\frac{S_g}{R_f}(n-2)u[n-2]+(e_r[n-1]-e_r[n-2])-(t_k[n-1]-t_k[n-2]) \quad (2\text{-}46)$$

and

$$d_2[n]=\frac{S_g}{2R_f}(n-3)(n-2)u[n-3]+e_r[n-2]-t_k[n-2]\ . \quad (2\text{-}47)$$

Signals d_1 and d_2 track u_1 and u_2. They increase linearly and quadratically with n at a rate of S_g/R_f. Error terms e_r and t_k, which were differentiated at the decimator input, are restored to their original form, with a constant mean-square value, by point d_2.

Truncation error is digitally removed from the output of the second digital accumulator,

$$d_k[n]=\frac{S_g}{2R_f}(n-3)(n-2)u[n-3]+e_r[n-2]\ . \quad (2\text{-}48)$$

The final accumulator has a value at time n of

$$d_3[n]=\begin{cases}0 & \text{for } n\le 4\\ d_3[n-1]+d_k[n-1] & \text{for } n\ge 5\end{cases} \quad (2\text{-}49)$$

and time and frequency responses of

$$d_3[n]=\frac{S_g}{6R_f}(n-4)(n-3)(n-2)u[n-4]+\sum_{i=2}^{n-3}e_r[i] \quad (2\text{-}50)$$

and

$$D_3(z)=\frac{S_g}{R_f}\frac{\left(z^{-5}-z^{-P-5}\right)}{\left(1-z^{-1}\right)^4}+E_r\frac{z^{-3}}{\left(1-z^{-1}\right)}\ . \quad (2\text{-}51)$$

The value of d_3 is independent of both truncation error t_k and the number of feedback bits k. It does, however, depend on the number of ADC bits r, since this determines the quantization step size and the mean-square value of e_r. Like the decimator in section 2.1.2, this configuration is referred to as an error-averaging decimator because the last term in (2-50) contains an average of e_r. This is a general aspect of configurations in which the decimator order is one greater than that of the modulator.

The converter's input sample can be easily reconstructed from d_3 because the sequence s_i is known to be constant along the pipeline. Equation (2-50), evaluated at $n=P+4$, can be solved directly for S_g in terms of d_3 and P, under the assumption that its quantization noise term is negligible. Reconstruction is given by

$$S_g \approx d_3[P+4]\left(\frac{6R_f}{P^3+3P^2+2P}\right) . \tag{2-52}$$

Equation (2-52) has a particularly simple implementation if the factor multiplying d_3 is chosen as a power of two. In this case the result is equal to d_3 except for a change in significance of its bits. Alternatively, this scale factor, which only impacts overall converter gain, could be treated as part of the full-scale reference.

2.2.3 Downsampling

Operation of a 2nd-order POSC is illustrated in Figure 2-12. A general configuration is considered with a modulator length of P and a decimator length of Z, where $Z \geq P$. The signal w_k is only nonzero for $2 \leq n \leq P+1$ because the modulator contains P stages and performs P comparisons. For illustrative purposes an example is chosen with

$$e_r[n] = \delta[n-2] + \delta[n-P-1] , \tag{2-53}$$

where $\delta[n]$ is a unit impulse function. Signal and error components of w_k are plotted in (a) and (b). While e_r is differentiated in this sequence at all points for $n<P+1$, it is not differentiated at time $n=P+1$ because of feedback truncation.

The signal components of d_1, d_2, and d_3 are shown in (c), (e), and (g), respectively. Each stage of accumulation increases signal amplification. The error components of d_1, d_2, and d_3 are shown in (d), (f), and (h), respectively. The original e_r sequence appears in d_2 for $n \leq P+3$. As a result of truncation, the undifferentiated point $e_r[P+1]$ causes the error in d_2 to increase rapidly beyond $n=P+3$.

A specific e_r was chosen for the illustration above. In general, e_r is a sequence of white noise. In this case, the rms error in d_3 is as shown in (i). It increases as $\sqrt{n}$ for $n \leq P+4$ and quadratically with n thereafter. For best performance, downsampling occurs when SNR in d_3 is maximized. SNR is plotted on a logarithmic scale in (j). Its maximum occurs at time $n=P+4$, when both modulator and decimator pipelines have operated for P cycles. Downsampling occurs at this point and no other values of d_1, d_2, or d_3 need to be computed.

The example above illustrates a result that is analogous to that of section 2.1.3 for a 1st-order architecture. Optimal resolution is achieved when modulator and decimator pipeline lengths are the same and when downsampling occurs at the end of the decimator pipeline. Adding or removing stages from the decimator pipeline degrades SNR.

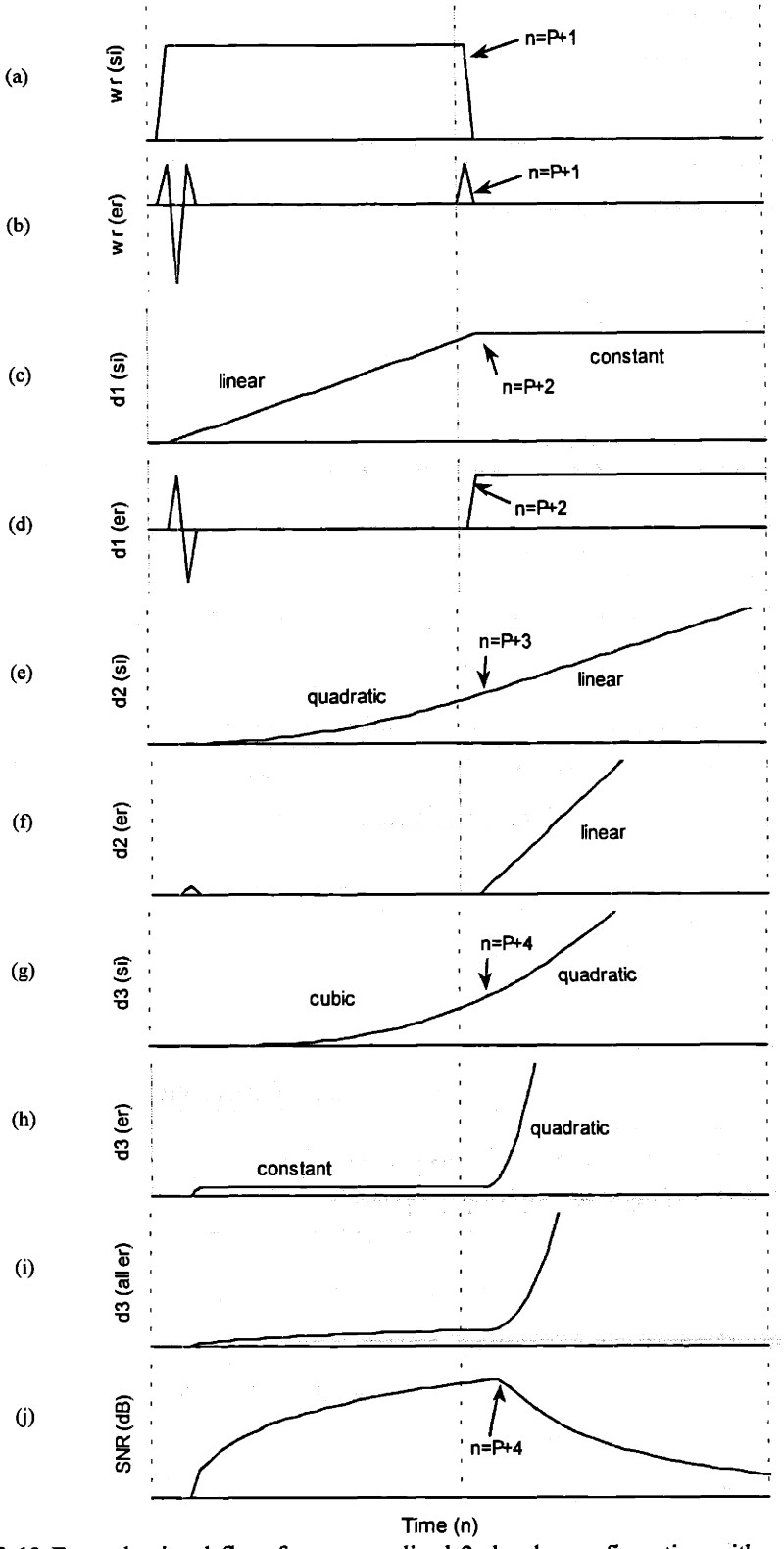


Figure 2-12. Example signal flow for a generalized 2nd-order configuration with unequal modulator and decimator lengths. (a), (c), (e), (g) Signal terms of w_k, d_1, d_2, and d_3. (b), (d), (f), (h) Error terms of w_k, d_1, d_2, and d_3. (i) Averaged error term for general e_r sequence. (j) SNR in the filter output.

2.2.4 Converter Resolution

POSC resolution can be expressed as a function of pipeline length P and number of ADC bits r. The approach followed here is similar to that described in section 2.1.4. An additional dimension of time is included, with sample points denoted by variable m. SNR for the output sequence is defined as the ratio of rms signal to error terms in (2-50). Modulator overloading effects are described in section 2.4.9. The largest input range that a 2nd-order modulator can accommodate without overloading is $R_f\left(1-2^{-k}\right)$. A sinusoidal input is assumed with this peak-to-peak span and an rms value about its mean of $R_f\left(1-2^{-k}\right)/2\sqrt{2}$.

The mean square value of e_r is given by

$$\overline{e_r^2[n]} = \frac{2^{-2r}}{12} + \overline{q_n^2[n]} \ . \tag{2-54}$$

Its first term represents ideal quantization noise in the coarse r-bit ADC and its second term is due to nonideal circuit inaccuracies of elements within the feedforward modulator path. The analysis below considers the case when r is small, the ADC has only a few bits, and quantization noise dominates e_r.

The quantization noise term for r_s is given by the summation in equation (2-50), of all values of e_r along the pipeline. To evaluate the rms value of this sum, noise is assumed to be white. This assumption provides an accurate representation of a 2nd-order modulator, where any tones that might occur in the outer modulator feedback loop are randomized by noise shaping in the inner feedback loop.

The quantization noise term in (2-50), evaluated at n=P+4, has an rms value of

$$\sqrt{\overline{r_s^2[m]}} = \sqrt{P} \cdot \sqrt{\frac{2^{-2r}}{12}} \qquad \text{(error term only)} \ . \tag{2-55}$$

The corresponding signal term is given by

$$\sqrt{\overline{r_s^2[m]}} = \frac{\left(1-2^{-k}\right)}{12\sqrt{2}} \cdot (P)(P+1)(P+2) \qquad \text{(signal term only)} \ . \tag{2-56}$$

SNR, as determined by the ratio of (2-56) to (2-55), is then equal to

$$SNR[P] = \frac{2^r\left(1-2^{-k}\right)(P)(P+1)(P+2)}{2\sqrt{6}\sqrt{P}} \approx \frac{2^r\left(1-2^{-k}\right)}{2\sqrt{6}} P^{5/2} \quad \text{for } P >> 1 \ . \tag{2-57}$$

The signal increases as P^3 due to integration, the quantization noise increases as $\sqrt{P}$ due to averaging, and the resulting resolution improves as $P^{5/2}$. Each additional bit in the r-bit feedforward ADC improves resolution by 1 bit. The number of DAC bits k limits the useful input signal range and, therefore impacts resolution as well. These results are plotted in Figure 2-13 for r=1, 3, and 5 and k=r. Solid lines show theoretical SNR, as given by (2-57). Simulated results are shown as diamonds.

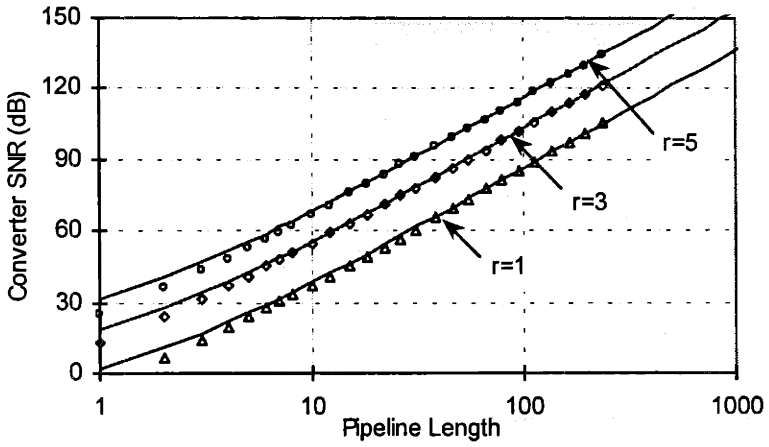


Figure 2-13. Second-order POSC resolution versus pipeline length. Lines show theoretical SNR. Diamonds show simulated SNR. SNR improves by 2.5 bits for every doubling of the pipeline length.

It is a well-known aspect of 2nd-order time-oversampling $\Delta\Sigma$ architectures that each doubling of the oversampling ratio provides an additional 2.5 bits of resolution. A similar result is achieved for a POSC by evaluating

$$\frac{\text{SNR}[2P]}{\text{SNR}[P]} = \frac{4(2P+1)}{\sqrt{2}(P+2)} \approx 2^{5/2} \text{ for } P >> 1 \ . \tag{2-58}$$

In a 2nd-order POSC, every doubling of the pipeline length provides an additional 2.5 bits of resolution.

2.2.5 Matched-Filter Decimation

The technique of matched-filter decimation was described in section 2.1.6. A matched-filter provides the best performance of any linear time-invariant filter for decimation of signals from a pipelined modulator. Application of this technique to a 2nd-order POSC is shown in Figure 2-14(a). This configuration replaces the decimator block of Figure 2-11. For simplicity, it is assumed that the number of ADC and DAC bits in the modulator are equal so that truncation error is zero. However, the techniques described here apply identically to a configuration with nonzero truncation error, provided this error is digitally cancelled before the matched filter.

Signal flow within a matched-filter decimator is illustrated in Figure 2-15. The modulator output w_r, described by equation (2-43), contains signal information and the 2nd-order difference of noise. After two stages of accumulation, d_2 has a signal component that is quadratically increasing and a noise component that is once again white. The M_F block in Figure 2-14(b) represents the matched filter for this system. It should have an impulse response that is a time-inverted version of d_2 in (b). Then the shape of the filter weights as a function of time matches that of its input and each point in d_2 is amplified in proportion to its SNR. The impulse response of the matched filter is shown in (c). Convolution of sequences (b) and (c)

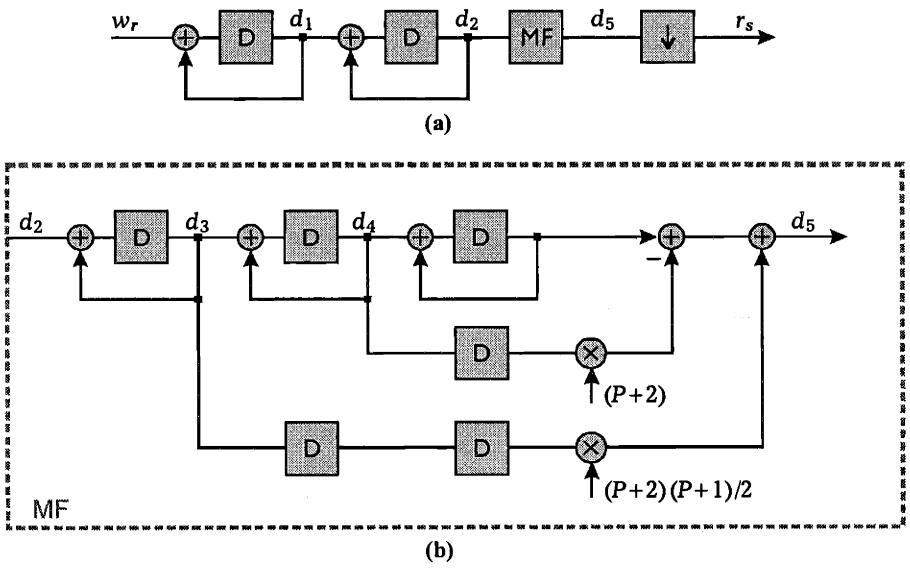


Figure 2-14. Improved decimator for a 2nd-order POSC has two stages of accumulation followed by a matched filter. (a) Decimator configuration. (b) Contents of MF block.

yields the filter output d_5, shown in (d). Maximum SNR occurs at the point of peak filter output at $n=P+6$, where the two sequences overlap completely. At this point, the filter output is equal to

$$d_5[P]=\sum_{n=4}^{P+3}\left[\frac{1}{4}(n-3)^2(n-2)^2\frac{S_g}{R_f}+\frac{1}{2}(n-3)(n-2)e_r[n-2]\right] . \qquad (2\text{-}59)$$

This value is captured by the downsampler and used to form the converter output word. No points other than this peak value are computed for the filter output.

The SNR achieved from this configuration equals

$$\mathrm{SNR}[P]=\frac{2^r\left(1-2^{-k}\right)}{4}\sqrt{\frac{1}{6}\left(7P^5+38P^4+53P^3+46P^2\right)}\approx\frac{2^r\left(1-2^{-k}\right)}{4}\sqrt{\frac{7}{6}}P^{5/2} \qquad \text{for } P>>1 . \qquad (2\text{-}60)$$

The improvement in SNR for each doubling of the pipeline length is given by

$$\frac{\mathrm{SNR}[2P]}{\mathrm{SNR}[P]}\approx 2^{5/2} \quad \text{for } P>>1 . \qquad (2\text{-}61)$$

A comparison of equations (2-60) and (2-61) for a matched-filter decimator with (2-57) and (2-58) for an error-averaging filter shows that both approaches provide an additional 2.5 bits for every doubling of the pipeline length. However, the matched filter provides an additional factor of $\sqrt{7/4}\approx 2.4$ dB of resolution. Section 2.4.1 compares the performance of these decimation techniques to that of time-oversampling decimation.

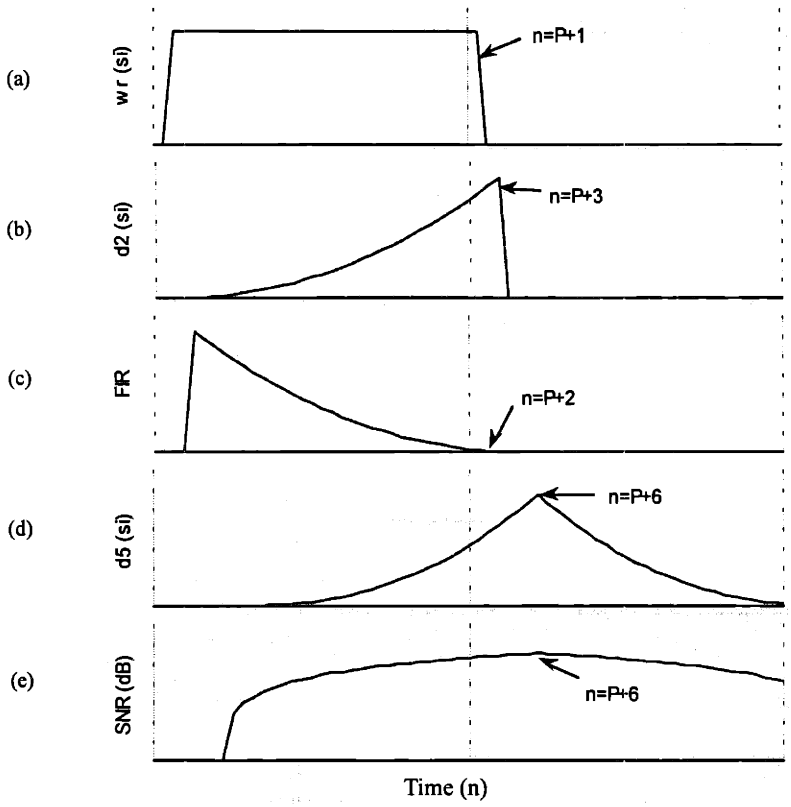


Figure 2-15. Matched filter operation for 2nd-order modulation. (a), (b), (d) Signal components of w_r, d_2, and d_5. (c) Matched filter impulse response. (e) SNR in the filter output.

2.2.6 Pipelined Implementation

A pipelined implementation of the algorithm in Figure 2-11 is shown in Figure 2-16. The pipeline in (a) contains P identical conversion blocks. Three analog channels, s_i, i_1, and i_2, flow through the pipeline. Three digital channels, d_1, d_2, and d_3, carry decimator signals. Inputs to all integrators and accumulators are zero. The analog converter input is sampled before the first pipeline stage. The result is passed along the pipeline and used in each stage as part of $\Delta\Sigma$ computations, but it is never altered. The converter output word is generated by the last stage. Multiple input samples are processed in parallel along the pipeline so that one digital word is completed each clock cycle.

The contents of a single pipeline stage are shown schematically in (b). The integrated quantity i_1, from the previous stage, is added to integrator i_2. The digital feedback signal w_k is used by a DAC, with a full-scale value of R_f, to generate an analog quantity f_b. This value is subtracted twice from i_2 and once from i_1. At the same time, the delayed input sample s_i is added to i_1. An r-bit ADC with a full-scale value of R_f quantizes the newly modified value of i_2. The k upper bits of the ADC result form w_k, the output from this stage. Decimator signals are updated as shown in the figure. The only digital signal from block n that is needed by block $n+1$ is w_k. The computations for d_1, d_2, and d_3 can be spread across multiple stages since these signals are not needed until the end of the pipeline. As shown, t_k is added directly to d_3,

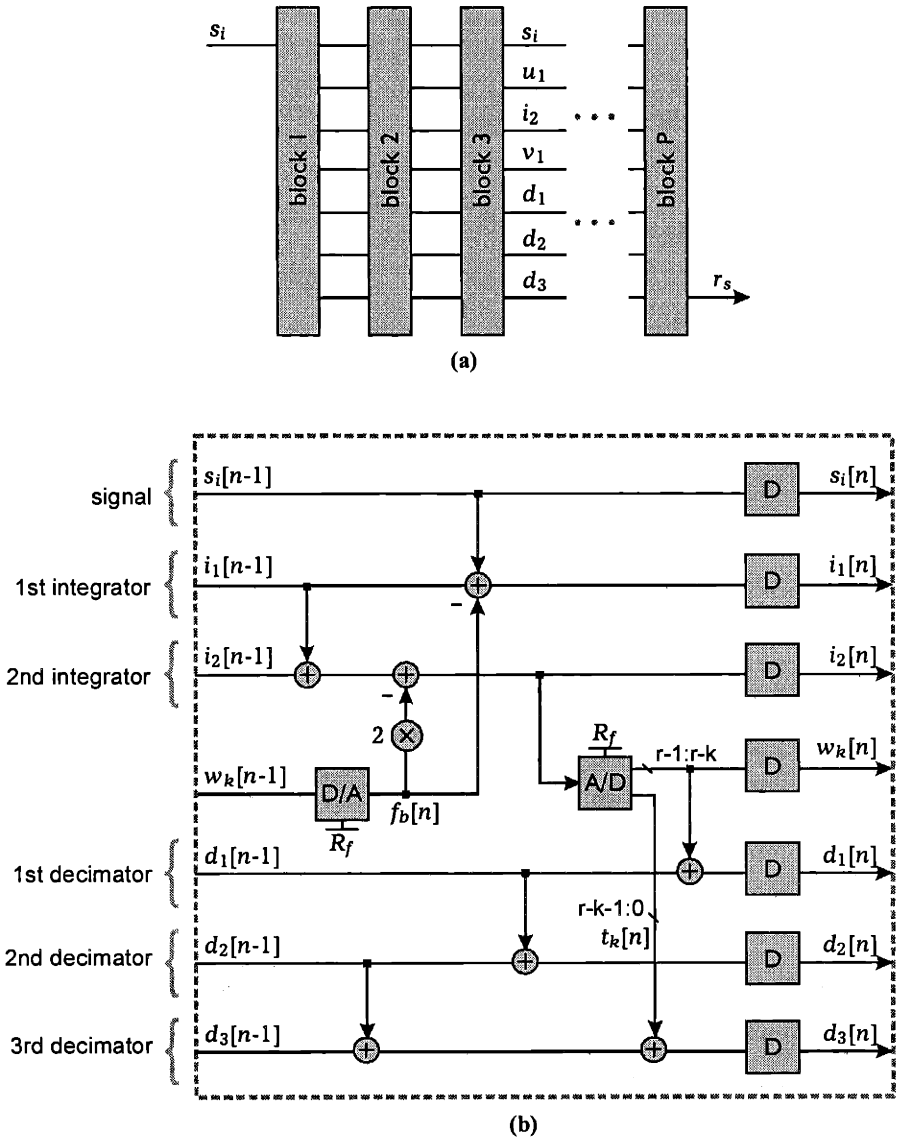


Figure 2-16. Pipelined 2nd-order POSC implementation. (a) Pipeline contains P identical conversion blocks. (b) Contents of block n. Each block implements one feedback operation.

without the delays present in Figure 2-11. The final result is not impacted by these delays. Finally, all signals are passed on to the following stage and the process is repeated. The converter result is generated by normalizing d_3 from the last stage using (2-52).

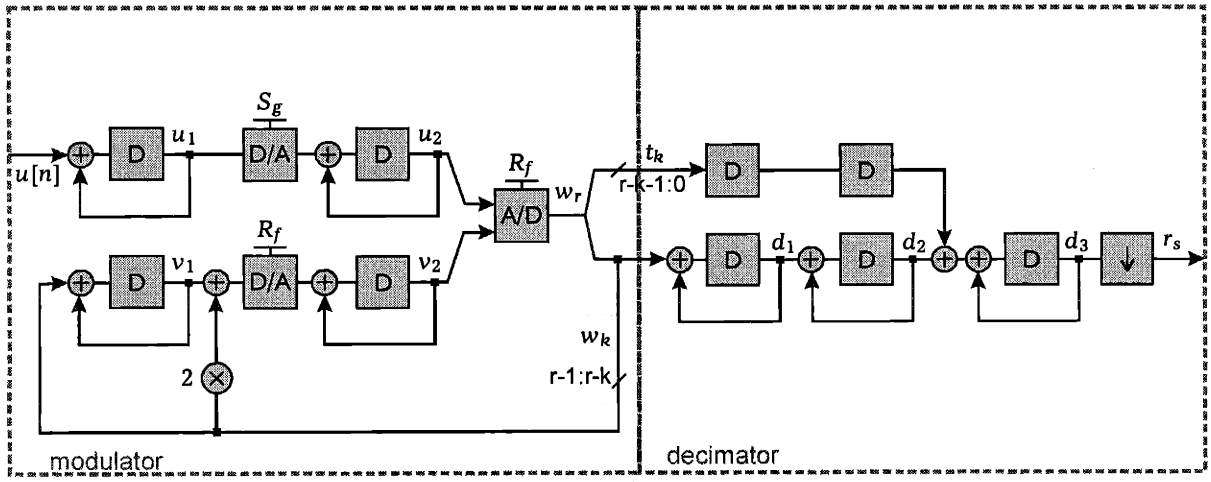
2.3 Digital-Integration Architectures

2.3.1 Path to Digital Integration

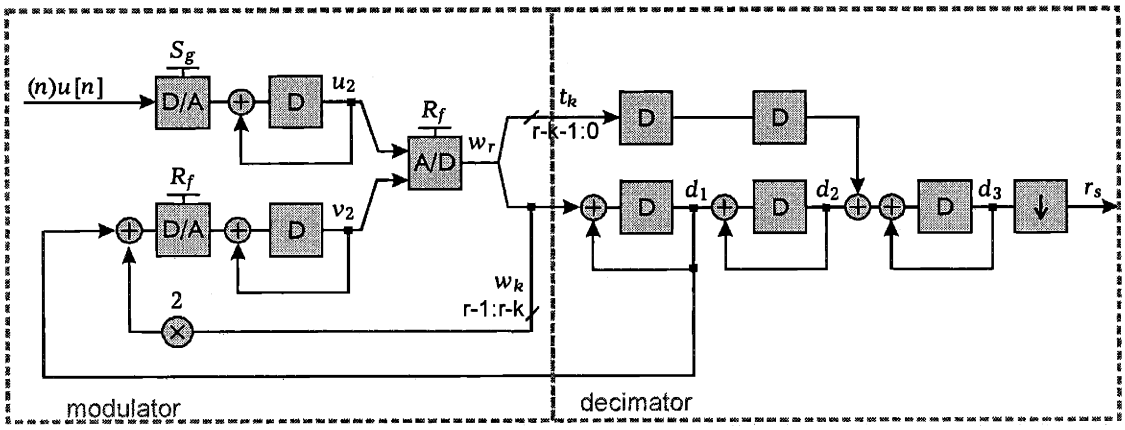
The 2nd-order architecture described in section 2.2 includes two stages of cascaded analog integration. Alternatively, one or both of these integration operations may be performed digitally. Figure 2-17 illustrates the transition from a 2nd-order analog-integration architecture, such as that in Figure 2-11, to a 2nd-order architecture with a combination of analog and digital integration. The block diagram in (a) differs from that of Figure 2-11 in that the DAC is repositioned. It was originally located before v_1 but here is positioned between v_1 and v_2. To balance signal and reference paths, a second DAC is also placed within the upper path, between u_1 and u_2. In this configuration, u_1 and v_1 are digital values. Analog quantities S_g and R_f enter the converter through the DAC full-scale references.

The transition from (a) to (b) is based on the fact that integrators for u_1 and v_1 are redundant. They can be removed to reduce hardware and power, without altering the $\Delta\Sigma$ algorithm. Node v_1 may be replaced by d_1 since the values of these two signals are identical. Node u_1 may be replaced with the fixed expression $nu[n]$, since this signal is a function of pipeline stage but is independent of the converter's input sample. The diagram in (c) differs from that in (b) in that u_2 and v_2 are combined into a single integrator i_2.

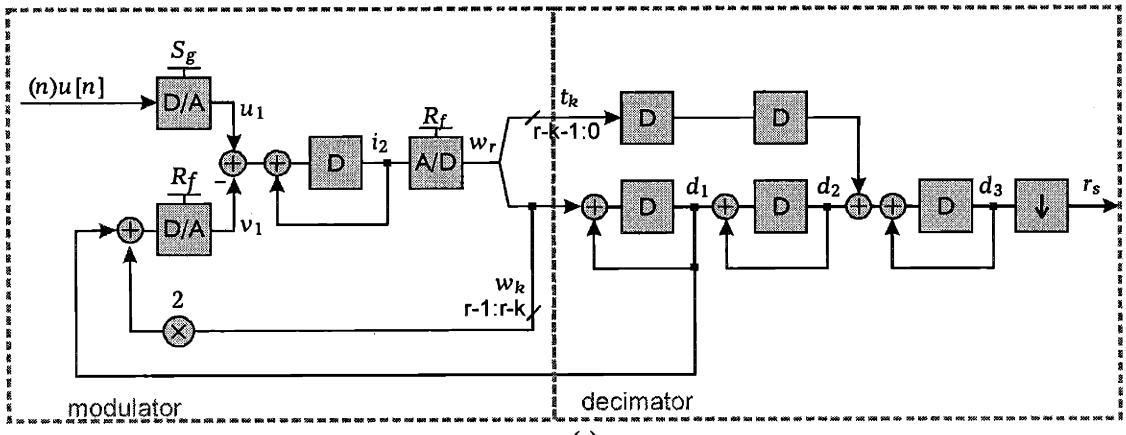
The configuration in (c) is referred to as a digital-integration architecture. Other varieties of digital-integration architectures are possible as well. For example, Figure 2-18 shows a configuration with two stages of digital integration and no analog integration. The only analog operations in this converter are sampling, subtraction, D/A conversion, and coarse A/D conversion. The advantages and disadvantages of this architecture, in comparison to analog integration, are described in Chapter 4.



(a)



(b)



(c)

Figure 2-17. Transition from analog-integration to digital-integration architectures. (a) DAC is repositioned. (b) Nodes u_1 and v_1 are eliminated. (c) Nodes u_2 and v_2 are combined.

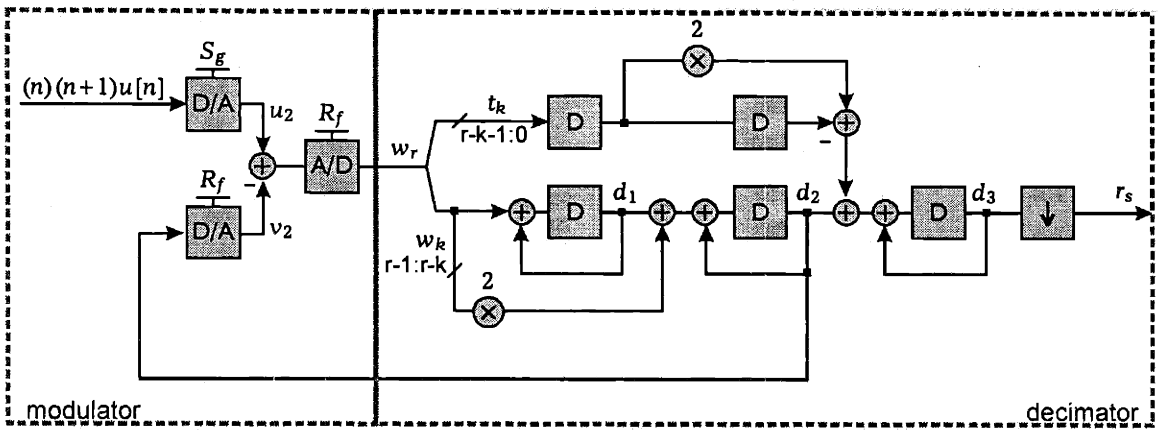


Figure 2-18. Alternative digital-integration architecture. Both stages of integration occur digitally.

2.3.2 Modulation Algorithm

Digital-integration and analog-integration architectures are algorithmically equivalent. The differences between these approaches lie in their implementations. The upper DAC has a digital input that is a function of pipeline stage and an analog input equal to the converter's input sample. It generates a signal

$$u_1[n] = S_g\, n\, u[n] \tag{2-62}$$

equal to the product of these inputs. The result is similar to u_1 in Figure 2-11. The lower DAC has a digital input that is signal dependent and an analog input that is constant. It generates a signal

$$v_1[n] = R_f\left(d_1[n] + 2w_k[n]\right) \ , \tag{2-63}$$

whose value is similar to v_1 in Figure 2-11. It represents the modulator's prediction of the value of u_1.

The difference between u_1 and v_1 is combined in the analog integrator according to

$$i_2[n] = \begin{cases} 0 & \text{for } n \le 1 \\ i_2[n-1] + u_1[n-1] - v_1[n-1] & \text{for } n \ge 2 \end{cases} . \tag{2-64}$$

Feedback in the modulator keeps the value in this integrator small so that it may be quantized using only a few bits. An r-bit ADC with a full-scale value of R_f generates a digital representation of i_2. The result,

$$w_r[n] = \frac{1}{R_f} i_2[n] + e_r[n] \ , \tag{2-65}$$

contains quantization noise e_r. The k most significant bits of this signal form the feedback quantity

$$w_k[n] = \frac{1}{R_f} i_2[n] + e_r[n] - t_k[n] \ . \tag{2-66}$$

An expression for the modulator output is arrived at by iteratively solving equations (2-62), (2-63), (2-64), and (2-66). The resulting expression,

$$w_k[n] = \frac{S_g}{R_f}\, u[n-2] + \left(e_r[n] - 2e_r[n-1] + e_r[n-2]\right) - \left(t_k[n] - 2t_k[n-1] + t_k[n-2]\right) \ , \tag{2-67}$$

is the same as that previously derived for an analog integration architecture in (2-43).

Decimator operation is identical to that described for analog integration in section 2.2.2 and is not repeated here.

2.3.3 Pipelined Implementation

A pipelined implementation of the digital-integration architecture in Figure 2-17(c) is shown in Figure 2-19. The pipeline in (a) differs from that of analog integration in that it has only two analog channels, s_i and i_2, that flow through the pipeline. The contents of a single pipeline stage are shown schematically in (b). An incoming digital feedback signal f_b is used by a DAC with a full-scale value of R_f to generate an analog quantity that is subtracted from i_2. At the same time, the delayed input sample s_i is provided as a full-scale reference to a DAC with a digital input of n, and the result is added to i_2. An r-bit ADC with a

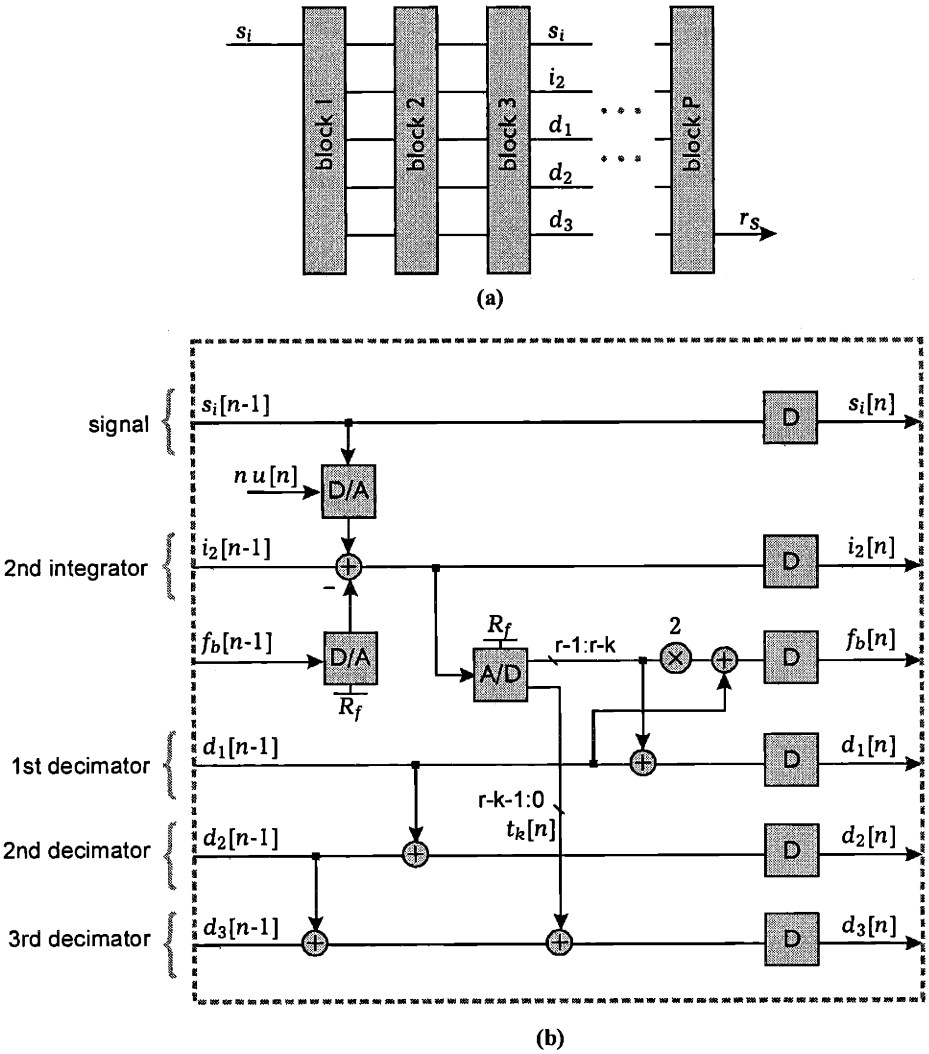


Figure 2-19. Block diagram of a 2nd-order digital-integration POSC. (a) Pipeline contains P identical conversion blocks. (b) Contents of block n. Each block implements one feedback operation.

full-scale value of R_f quantizes the newly modified value of i_2. The k upper bits of the ADC result are combined with the first decimator value d_1 to generate the feedback output from this stage. Decimator signals are updated in the same way as those for an analog-integration architecture. Finally, all signals are passed on to the following stage and the process is repeated. The converter's result is generated by normalizing d_3 from the last stage using (2-52).

2.4 Design Considerations

2.4.1 Resolution Comparison

Resolution of a time-oversampling converter depends on its oversampling ratio. Resolution of a pipelined oversampling converter depends on its pipeline length. To be meaningful, a comparison of the resolution provided by each of these architectures must be based on a single parameter that is applicable to both. However, oversampling ratio is not clearly defined in a pipelined device and pipeline length has no direct analogue in a time-oversampling converter.

A more meaningful parameter for comparison is a parameter referred to here as the 'sample aperture'. It is defined as the number of samples of the modulator output w_k that are combined to form a single downsampled word. A time-oversampling converter has a sample aperture that is a function of its oversampling ratio. A POSC has a sample aperture that is a function of its pipeline length. The relationships between sample aperture, oversampling ratio, and pipeline length are described below.

Two example filter impulse responses for time-oversampling decimation are shown in Figure 2-20. The higher frequency modulator internal clock and the lower frequency decimator downsampling clock are shown in (a) as ticks and arrows, respectively. The decimation filter operates continuously to produce outputs. But with an oversampling ratio of R, its result is only used every R internal clock cycles, at the times denoted by arrows. A 2nd-order sinc filter is shown in (b). The sequence of modulator outputs is multiplied by the triangular weights shown and summed to generate the filter result. Because the impulse response of each filter is greater than its oversampling ratio, adjacent impulse responses overlap, and each input contributes to two adjacent results. In (b) the sample aperture is equal to the filter impulse response length of $2R$-1. A second example in (c) consists of a 3rd-order sinc filter and has a sample aperture of $3R$-2.

Time-Oversampling Resolution

The following method is used to calculate the theoretical maximum SNR for a time-oversampling converter [20]. A fixed signal bandwidth is selected and differentiated noise at the modulator output is integrated over it. The integrated noise is compared to the rms signal for a full-scale sinusoidal input. Their ratio represents SNR of the modulator alone. Equivalently, it represents the maximum SNR achievable from the modulator when combined with an ideal lowpass decimation filter.

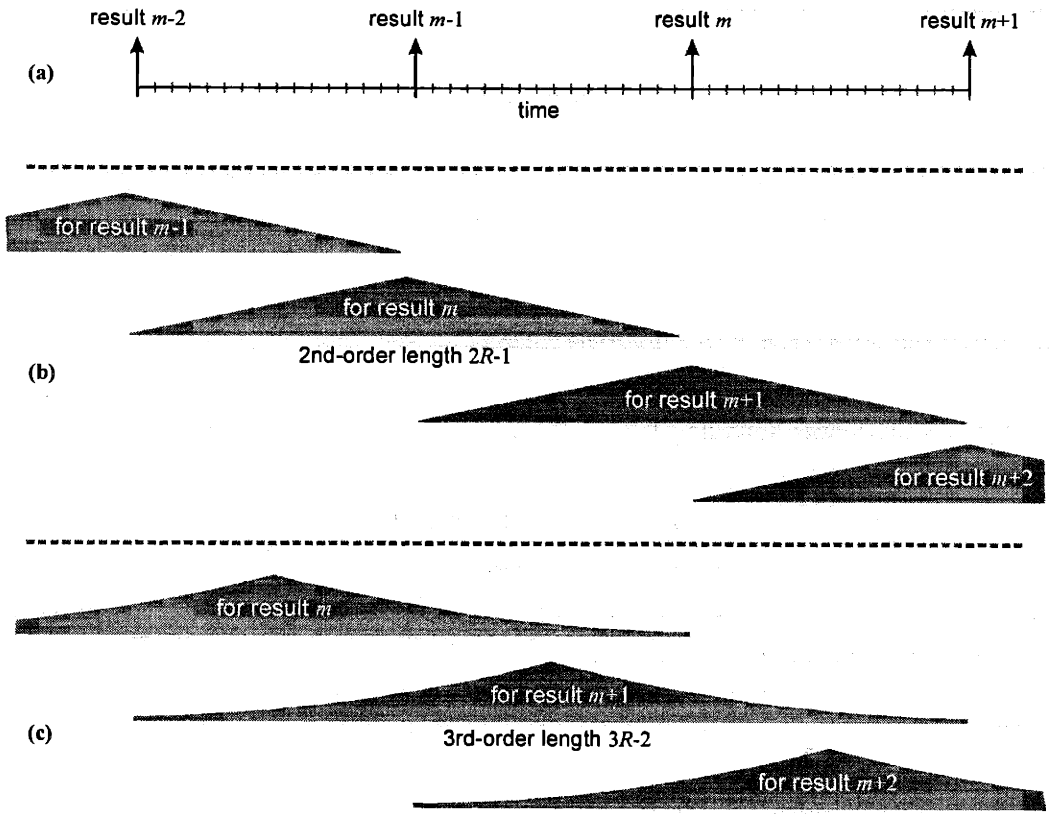


Figure 2-20. Sample aperture for a time-oversampling converter is greater than its oversampling ratio. (a) Modulator clock (ticks) and decimator clock (arrows). (b) Sinc^2 window length is $2R$-1. (c) Sinc^3 window length is $3R$-2.

Using this method, the greatest SNR that is possible for a 1st-order time-oversampling $\Delta\Sigma$ converter with an oversampling ratio of R is

$$\text{SNR} = \frac{\sqrt{4.5}\,2^r\left(1-2^{-k}\right)}{\pi} R^{3/2} \tag{2-68}$$

and that for 2nd-order is

$$\text{SNR} = \frac{\sqrt{7.5}\,2^r\left(1-2^{-k}\right)}{\pi^2} R^{5/2} \quad . \tag{2-69}$$

Since this value is based on an ideal decimation filter, it has an associated decimator impulse response that is infinite.

In practice, any realizable decimator produces worse SNR than that given by (2-68) and (2-69) because some higher-frequency noise is passed by the filter and this noise aliases into the signal bandwidth upon downsampling. The performance of a particular decimator is evaluated by comparing its noise to that from an ideal lowpass filter. It has been shown [19] that for a fixed decimator impulse response length of

R, the maximum SNR possible from a 1st-order time-oversampling converter occurs for the filter with parabolic weights of

$$w[i]=\frac{i(R-1-i)}{R} \quad . \tag{2-70}$$

Triangular sinc^{L+1} filters perform nearly as well as those with parabolic weights and have a much simpler implementation. Both of these approaches provide performance that is within a few dB of that from an ideal filter [20].

The relationship that is assumed here between resolution and sample aperture for an L_{th}-order time-oversampling converter is that for parabolic or triangular filters because they provide the greatest resolution for a given sample aperture. It is derived as follows. First, the sample aperture W is expressed in terms of the oversampling ratio as

$$\begin{aligned} W &= 2R-1 \qquad \text{(1st - order)} \\ W &= 3R-2 \qquad \text{(2nd - order)} \end{aligned} \quad . \tag{2-71}$$

Then (2-68), (2-69), and (2-71) are combined to arrive at an upper bound on SNR as a function of W. For 1st-order modulation the result is given by

$$\text{SNR}=\frac{\sqrt{4.5}\,2^r\left(1-2^{-k}\right)}{\pi}\left(\frac{W+1}{2}\right)^{3/2}\approx\frac{\sqrt{4.5}\,2^r\left(1-2^{-k}\right)}{\pi}\left(\frac{W}{2}\right)^{3/2} \qquad \text{for } W>>1 \ . \tag{2-72}$$

For 2nd-order modulation it equals

$$\text{SNR}=\frac{\sqrt{7.5}\,2^r\left(1-2^{-k}\right)}{\pi^2}\left(\frac{W+2}{3}\right)^{5/2}\approx\frac{\sqrt{7.5}\,2^r\left(1-2^{-k}\right)}{\pi^2}\left(\frac{W}{3}\right)^{5/2} \qquad \text{for } W>>1 \ . \tag{2-73}$$

Filter responses in a time-oversampling decimator can overlap because integrators operate continuously and no resetting occurs.

Pipelined Oversampling Resolution

Figure 2-21 illustrates how decimator filter responses for a POSC differ from those for time oversampling. In a POSC, the modulator generates P outputs over a pipeline length of P and its integrators are effectively reset between each result. As a consequence of resetting, the sequence $w_k[n]$ for $1\le n\le P$ is only relevant to one of the converter input samples and can only be used for one of its output words. In a POSC, the best SNR is achieved if filter impulse responses are constrained to a length of P, regardless of the modulator or decimator order.

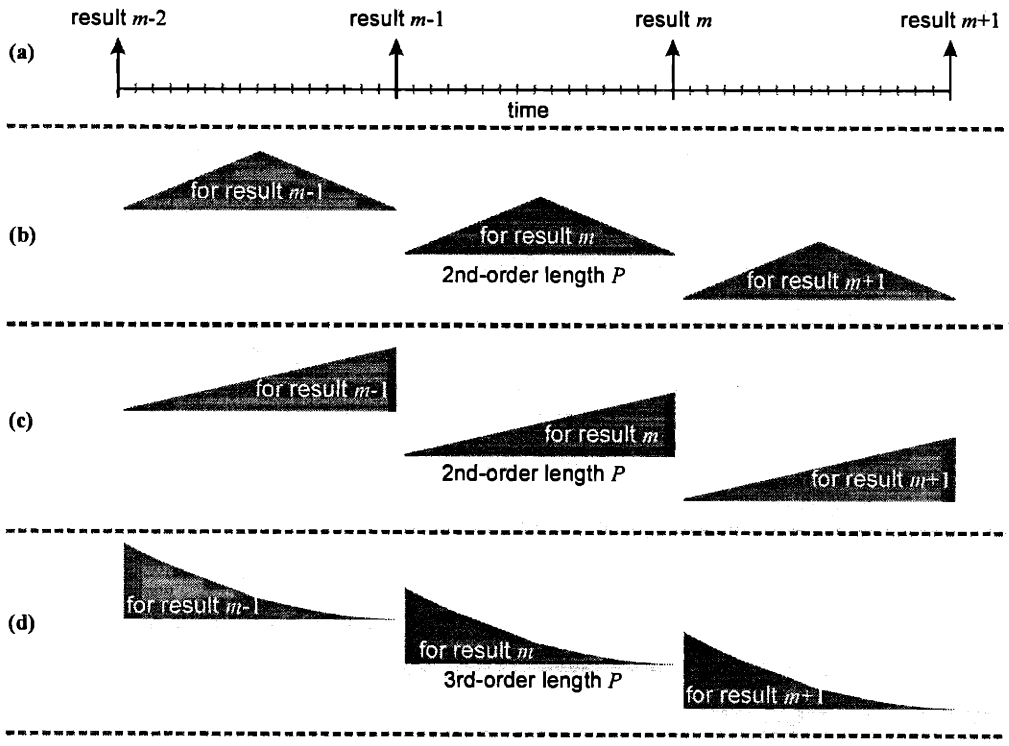


Figure 2-21. Sample aperture for a POSC equals its pipeline length. (a) Time m (arrows) and pipeline stage n (ticks). (b) Compressed sinc^2 filter. (c) Error averaging filter. (d) Matched filter.

The time axis in Figure 2-21(a) represents a combination of sample points (m), denoted by arrows, and pipeline stages (n), denoted by ticks. The 2nd-order sinc filter in (b) has the same form as the filter in Figure 2-20(b), but it is compressed to a length of P. Although a sinc^2 filter, such as the example in (b), provides near-ideal resolution in a time-oversampling device, greater performance is possible for a pipelined device. Truncated, rather than compressed, responses such as those in (c) and (d), are better choices for a POSC. The filter in (c) corresponds to the error-averaging filter of section 2.1.2 and that in (d) represents the matched filter of section 2.1.6. The relation $W=P$ for a pipelined device can be used to express the formulas for SNR in terms of sample aperture W.

Table 2-2 shows expressions for the SNR of 1st and 2nd-order time-oversampling and POSC architectures, based on the assumption $W>>1$. Table 2-1 shows the improvement that pipelined oversampling provides compared to time oversampling for these architectures. Entries are computed as a ratio of POSC SNR to time-oversampling SNR, for the same values of W, L, and r. The resulting expressions are independent of W, L, and r since all entries in Table 2-2 have the same dependence on these variables.

	1st-Order Modulation	2nd-Order Modulation
POSC with Error-Averaging Filter	$\frac{\sqrt{6}\,\pi}{3} = 8.2\,\text{dB}$	$\frac{\pi^2\,3\sqrt{3}}{2\sqrt{5}} = 21.2\,\text{dB}$
POSC with Matched Filter	$\frac{\pi 2\sqrt{2}}{3} = 9.4\,\text{dB}$	$\frac{\pi^2\,3\sqrt{21}}{4\sqrt{5}} = 23.6\,\text{dB}$

Table 2-1. Ratio of POSC SNR to time-oversampling SNR. Pipelined architectures achieve greater SNR than time-oversampling devices for the same sample aperture.

The primary factor contributing to improved resolution for a POSC is that the form of its input is precisely known. This fact enables decimation techniques with better performance. For both pipelined and time-oversampling architectures, the sum of weights in the decimator determines signal amplification at the converter output. The derivative of these weights determines quantization noise power at the converter output. Optimizing decimator performance amounts to maximizing the sum of the decimator weights while minimizing the sum of the magnitude of their derivative.

This concept is illustrated in Figure 2-22. Best performance from a time-oversampling decimator is achieved with a symmetric response. One example of such a filter is shown in the figure. It has weights that taper to zero at its beginning and end. Improved performance is achieved for a POSC with an asymmetric response such as that shown in the figure. It begins at zero and ramps to a maximum at its final point. These two filters produce similar noise power at their outputs because their derivatives have identical magnitudes everywhere, except at one point. However, the asymmetric filter, that is applicable to a POSC provides greater signal amplification because the area under its curve is larger.

	1st-Order Modulation	2nd-Order Modulation
Time-Oversampling with Sinc Filter	$\frac{\sqrt{4.5}\,2^r\left(1-2^{-k}\right)}{\pi}\left(\frac{W}{2}\right)^{3/2}$	$\frac{\sqrt{7.5}\,2^r\left(1-2^{-k}\right)}{\pi^2}\left(\frac{W}{3}\right)^{5/2}$
POSC with Error-Averaging Filter	$\frac{\sqrt{6}\,2^r\left(1-2^{-k}\right)}{4}W^{3/2}$	$\frac{2^r\left(1-2^{-k}\right)}{2\sqrt{6}}W^{5/2}$
POSC with Matched Filter	$\frac{2^r\left(1-2^{-k}\right)}{\sqrt{2}}W^{3/2}$	$\frac{2^r\left(1-2^{-k}\right)}{4}\sqrt{\frac{7}{6}}W^{5/2}$

Table 2-2. Comparison between the SNR of time-oversampling and POSC architectures as a function of sample aperture.

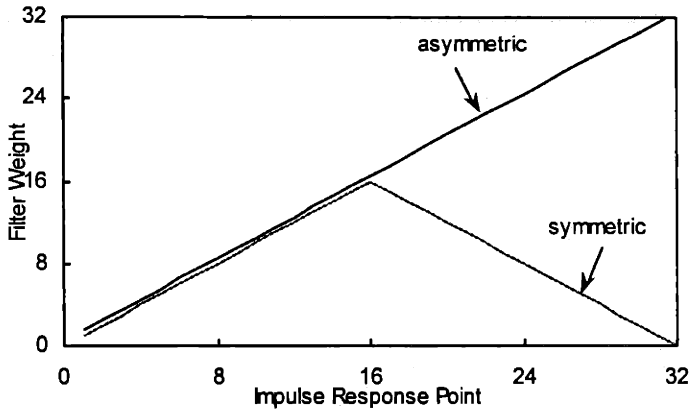


Figure 2-22. Decimation filter weights for symmetric time-oversampling and asymmetric pipelined oversampling responses. Signal amplification equals the area under each curve.

2.4.2 Modulator Order

Although only 1st and 2nd-order modulation techniques are described above, higher-order modulation can be used for a pipelined architecture as well. The criteria used to select a modulator order differ for pipelined versus time-oversampling devices. In a time-oversampling converter, an increase in modulator order decreases its required oversampling ratio and this translates into increased converter bandwidth. The primary disadvantage of higher-order modulation is that modulators of order greater than two are susceptible to instability. Careful design and accurate characterization of circuit components are necessary to assure stability for these devices. Higher-order modulation also has increased analog circuit complexity. Despite these factors, modulators of order greater than two are often used for time-oversampling converters when the advantages of increased converter bandwidth outweigh the associated disadvantages.

Hardware versus Modulator Order

Higher-order modulators do not improve the bandwidth of a POSC because this architecture is not subject to a speed-resolution tradeoff. However, an increased modulator order decreases the pipeline length that is required to achieve a given resolution, and this impacts other device aspects such as hardware, power, and area. A decreased pipeline length does not necessarily translate into reduced hardware. The advantage of fewer conversion blocks is partially offset because the number of stages per conversion block and the number of circuit elements per stage increase with modulator order.

Figure 2-23 plots an estimate of the total number of hardware units required for 1st, 2nd, 3rd, and 4th-order architectures as a function of desired resolution. A 1-bit ADC, a 1-bit DAC, and an error averaging decimator are assumed for all curves. The measure of hardware units is based on a configuration such as that in Figure 2-10 and is defined as

$$\text{Hardware} = (\#\ \text{of stages}) \cdot \begin{bmatrix} L \cdot (\text{Integrators}) + 1 \cdot (\text{Comparators}) + (\text{L}+1)^2 \cdot (\text{Delays}) \\ +L \cdot (\text{DAC / Subtractions}) + (L+1) \cdot (\text{Digital Adders}) \end{bmatrix} . \qquad (2\text{-}74)$$

Each of the elements listed above is assigned a weight of one. One integrator and one DAC are added per conversion block with each increase in L. The number of delay elements increases quadratically with L because an increase in L adds an analog delay channel to each stage and one additional stage to each existing channel.

A jump from 1st to 2nd-order modulation results in a substantial hardware savings. However, the magnitude of this improvement decreases with increasing order. Architectures with $L>3$ provide only a minimal hardware reduction over those for $L=3$. Any hardware reduction that is achieved must be weighed against the disadvantages associated with these higher-order architectures.

Latency versus Modulator Order

A second impact that a decreased pipeline length has is on converter latency. This in turn impacts SNR because noise is introduced by circuits each time they transfer their signals forward. And noise is introduced for each clock cycle that a signal is stored in a sample-and-hold circuit. Therefore, the total number of transfers and the pipeline latency are a consideration in selecting a modulator order. Practical device implementations, such as those of sections 2.1.7 and 2.2.6, have multiple stages per conversion block and require multiple clock cycles to perform one feedback operation. This number increases with modulator order because each step of integration occupies one stage and one clock cycle. The number of

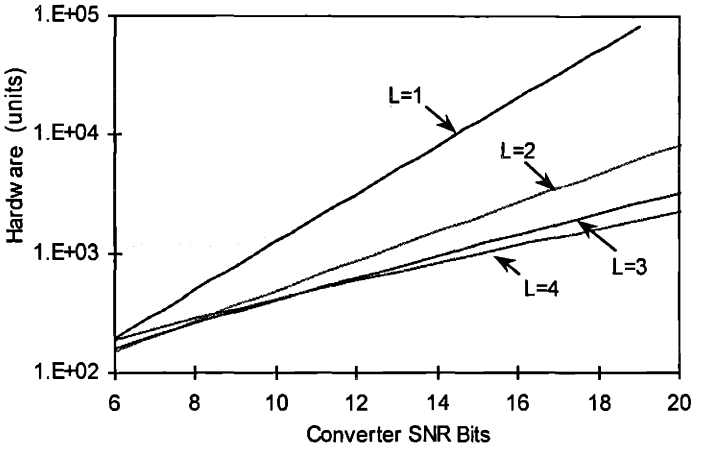


Figure 2-23. Hardware versus resolution for modulators of order 1, 2, 3, and 4.

transfers and the pipeline latency for different modulator orders is plotted in Figure 2-24 as a function of resolution. An increase from 1st to 2nd-order modulation produces a substantially reduced pipeline length. However, the magnitude of this improvement decreases with increasing order and beyond L=3 there is little additional benefit.

A 1st-order architecture is not a suitable choice for a POSC because its resolution is reduced from the theoretical values described above by pattern noise or tones. Pattern noise, which appears in both time-oversampling and pipelined oversampling converters, occurs when the modulator output settles into oscillations whose frequency lies within the passband of the decimator [3]. In this case, quantization noise is no longer white, and the assumptions of sections 2.1.2 and 2.2.2 no longer apply. When errors are correlated, the quantization noise component of the decimator output is larger and SNR is reduced. Susceptibility to pattern noise is aggravated for devices, such as a POSC, with slowly varying or constant inputs and for devices with a single-bit quantizer. However, its effect is significantly reduced for architectures of order two or higher because noise shaping within the inner feedback loops randomize potential patterns in the outer loop.

2.4.3 Decimator Order

Primary objectives in selecting a decimator design are to minimize converter hardware, reduce stress on the analog modulator, and reduce the accuracy required from the modulator circuit elements. Decimator design options are presented here for the 2nd-order modulator of Figure 2-11. Either d_1, d_2, d_3, or the matched filter output, d_5, can be used to reconstruct the original input sample. Table 2-3 lists the reconstruction formulas for, and the SNR of each of these configurations. Truncation error is assumed to be zero, but the results apply identically to a configuration with nonzero truncation error since this error

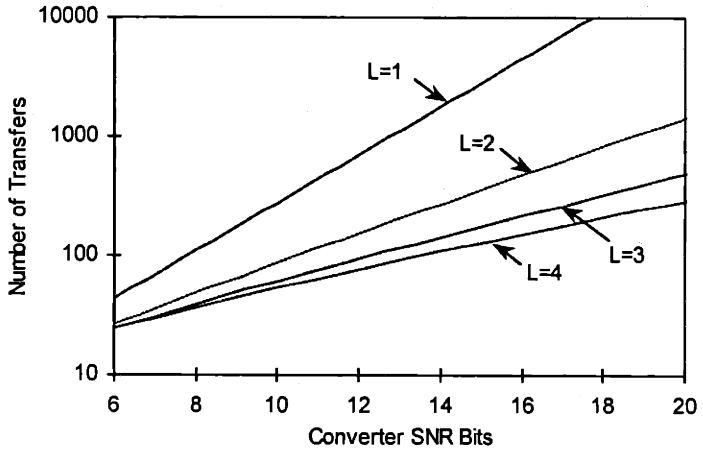


Figure 2-24. Number of transfers and pipeline latency versus resolution for modulators of order 1, 2, 3, and 4.

Signal	Reconstruction	SNR
$d1[n]$	$d1[n+2]\frac{Rf}{P}$	$\frac{2^r\left(1-2^{-k}\right)P\sqrt{3}}{2}$
$d2[n]$	$d2[n+3]\frac{2\cdot Rf}{(P)(P+1)}$	$\frac{2^r\left(1-2^{-k}\right)\sqrt{6}P(P+1)}{4}$
$d3[n]$	$d3[P+4]\left(\frac{6Rf}{P^3+3P^2+2P}\right)$	$\frac{2^r\left(1-2^{-k}\right)\left(P^3+3P^2+2P\right)}{2\sqrt{6P}}$
$d5[n]$	$d5[P+6]\frac{(36)(4)Rf}{\left(7P^5+38P^4+53P^3+46P^2\right)}$	$\frac{2^r\left(1-2^{-k}\right)}{4}\sqrt{\frac{1}{6}\left(7P^5+38P^4+53P^3+46P^2\right)}$

Table 2-3. Signal reconstruction for different order decimation filters.

can always be cancelled digitally before decimation.

Resolution versus Decimator Order

Figure 2-25 shows the resolution for each of these decimation filters as a function of pipeline length P. All curves are based on a 2nd-order architecture with r=k=1. Decimator order is denoted by D. There is a large difference in performance between 1st and 2nd-order decimation. Node d_1 is only capable of producing P distinct outputs and a configuration that uses this signal for reconstruction requires nearly 2^N pipeline stages for N-bit resolution. Node d_2 provides significantly greater resolution than d_1 because it has a signal term that is P times larger and a quantization noise term that is smaller by a factor of $\sqrt{2}$.

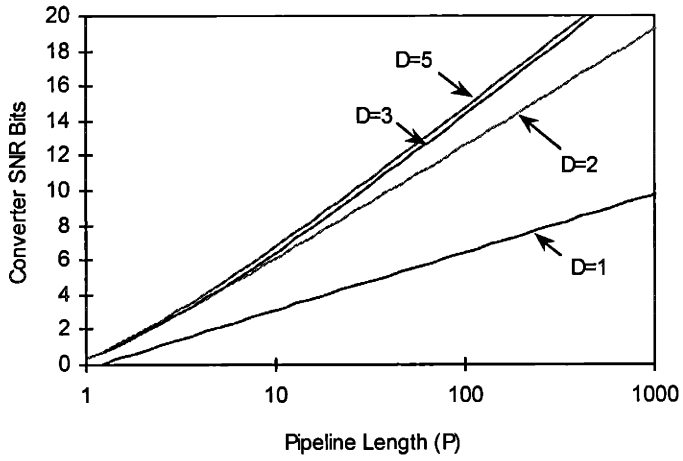


Figure 2-25. SNR for a 2nd-order modulator as a function of pipeline length and decimation order. Higher-order decimation filters provide better SNR up to D = 5.

Neither d_1 nor d_2 benefit from quantization noise averaging.

An error-averaging decimator uses d_3 for signal reconstruction. Node d_3 provides resolution that is a factor of $\sqrt{P}$ larger than that from d_2 because it averages quantization noise and, therefore, has a reduced error term. The greatest SNR is achieved from a matched filter. Both an error-averaging and a matched filter perform noise averaging, but a matched filter does so more effectively. While an error-averaging filter computes an unweighted sum of d_2 values, a matched filter weights each point by its time of occurrence so that earlier values, with less SNR, are weighted less and later values, with greater SNR, are amplified more. Increasing the decimator order beyond that of a matched filter does not improve converter resolution because signal and quantization noise terms are amplified identically.

Hardware versus Decimator Order

Increasing the decimator order reduces the pipeline length needed for a given resolution. But it also increases the number of digital adders in each conversion block. These competing effects are merged in Figure 2-26 as a plot of total converter hardware versus desired resolution. A 2nd-order modulator and 1-bit quantizer are assumed for the plot. Hardware is computed as

$$\text{Hardware} = (\#\ \text{of stages}) \cdot \begin{bmatrix} 2 \cdot (\text{Integrators}) + 1 \cdot (\text{Comparators}) + 3^2 \cdot (\text{Delays}) \\ +2 \cdot (\text{DAC / Subtractions}) + D \cdot (\text{Digital Adders}) \end{bmatrix} . \tag{2-75}$$

Hardware is reduced with increasing decimator order but the magnitude of this improvement decreases with D up to D=5. Increasing D beyond D=5 has no merit because hardware increases but resolution does not.

2.4.4 DAC Matching for Analog Integration

$\Delta\Sigma$ architectures are tolerant of inaccuracies that occur in their feedforward paths. However, their

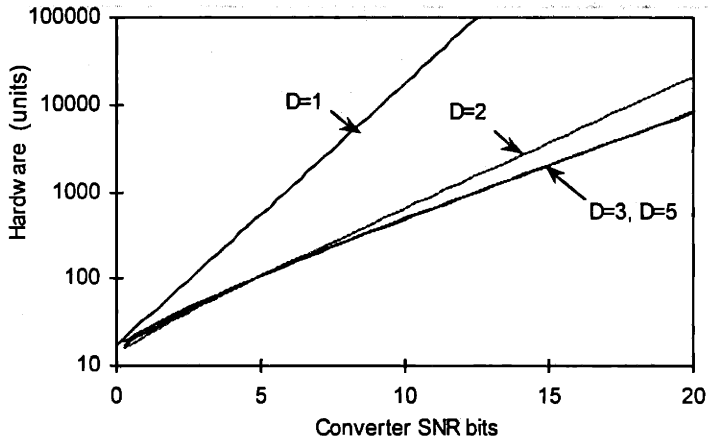


Figure 2-26. Converter hardware as a function of resolution and decimator order. Hardware decreases with increasing decimator order up to $D = 5$.

feedback signals do not experience noise shaping and inaccuracies in them are indistinguishable from an incoming signal. Since the converter interpolates higher-resolution values between its coarse feedback DAC levels, any misplacement in these levels introduces nonlinearity. The feedback path must be as accurate as the overall converter [29]. This requirement applies to both pipelined and time-oversampling converters.

A time-oversampling converter has only one modulator. The same DAC elements are used repeatedly in this modulator, so their inaccuracies add directly and averaging occurs over only a few elements. Fortunately, time-oversampling converters can circumvent the need for high accuracy DAC elements by using single-bit feedback [30]. In this approach, only two feedback levels are utilized and each of these is generated by the same element during every cycle. The placement of their levels impacts converter gain but does not introduce nonlinearity or noise. The price of these benefits is a nearly 6 dB reduction in input signal range and SNR compared to an architecture with a multi-bit DAC.

Time-oversampling converters with multi-bit feedback do not experience this inherent linearity advantage or the benefits of averaging. Nonetheless, the effects of DAC mismatches can be reduced. In one such approach, referred to as dynamic element matching, logic is used to select between a set of DAC elements in a way that randomizes feedback errors so that their effect is reduced by averaging [32]. In another approach, referred to as mismatch shaping, DAC elements are selected with a repetition rate whose frequency lies above the passband of the decimator [33].

A pipelined converter differs from a time-oversampling converter in that single-bit feedback does not have an inherent linearity advantage over multi-bit feedback. Oversampling in a POSC occurs along a series of pipeline stages and is performed using different feedback elements, regardless of the number of feedback bits. Nonlinearity and noise are introduced even when a single-bit DAC is used because its level placement is determined by different elements in each stage. Multi-bit feedback is an attractive option for a POSC because it brings little disadvantage in terms of DAC matching requirements and it also provides other advantages. In comparison to modulators with one-bit feedback, modulators with multi-bit feedback have a larger input signal range that is free from overloading effects, more predictable behavior, and are less susceptible to pattern noise [27].

A pipelined device naturally achieves the benefits of averaging without the use of dynamic element matching techniques because its DAC contains independent circuits from many stages. As a result, the need for highly accurate feedback circuits is moderately alleviated. However, after averaging, DAC matching must have the full converter resolution for either single-bit or multi-bit feedback.

Matching with Respect to DAC Full Scale

The impact of DAC mismatches on an analog-integration POSC is analyzed by considering the 2nd-order configuration in Figure 2-11 with an additive error c_v in the feedback path. The modulator feedback signal becomes

$$f_b[n] = u_2[n] - v_2[n] + (e_r[n] - t_k[n] + c_v[n])R_f \tag{2-76}$$

and its output is given by

$$w_k[n] = \frac{S_g}{R_f}u[n-2] + (-2c_v[n-1] + c_v[n-2]) + (e_r[n] - 2e_r[n-1] + e_r[n-2]) + (t_k[n] - 2t_k[n-1] + t_k[n-2]) \tag{2-77}$$

While c_v is differentiated in the feedback path, it is not differentiated at the modulator output.

An error-averaging decimator generates the result

$$d_3[n] = \frac{S_g}{6R_f}(n-4)(n-3)(n-2)u[n-4] + \sum_{i=2}^{n-3} e_r[i] + \sum_{i=2}^{n-4}(-c_v[n-2-i][i(i+1)/2+i]) \ , \tag{2-78}$$

which contains a weighted sum of c_v values. Earlier DAC errors are weighted more heavily than later ones.

A relevant measure of the accuracy required from DAC elements is given by the point for which the rms values of ideal quantization noise and DAC-related noise are equal. Equating the mean-square values of e_r and c_v terms in (2-78) yields the expression

$$\sqrt{\overline{c_v^2}} = \sqrt{\frac{2^{-2r}P}{3\sum_{i=1}^{P}(i^2+3i)^2}} \ . \tag{2-79}$$

As defined above, c_v is a quantity normalized to R_f, and accuracy requirements refer to the total error produced by each DAC with respect to its full-scale reference. Equation (2-57), which relates SNR to P, is then used to transform this into an expression for DAC mismatch tolerance that is independent of P. The results are plotted in Figure 2-27 for r=k=1 and 7. Solid lines represent the theoretical expressions above. Points show simulated values for these same results. The impact of mismatches is reduced by $\sqrt{P}$ as a result of averaging among elements from each pipeline stage. As r increases, tighter matching is required because fewer pipeline stages, and therefore fewer elements, participate in averaging.

Matching with Respect to Nominal Size

The results above can also be expressed as a tolerance of each DAC element with respect to its nominal size. In this case, equal numbers of ADC and DAC bits are assumed, so that k=r and each k-bit DAC is assumed to consist of 2^k identically sized elements. The total number of DAC elements is increased by a factor of 2^k and each element has a nominal size of 2^{-k}. Tolerance is then given by

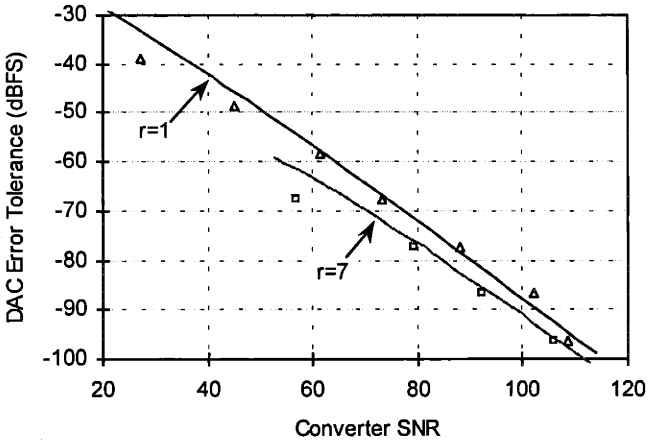


Figure 2-27. Tolerance to DAC errors for a 2nd-order analog-integration architecture. Tolerance for each DAC is relative to its full-scale value. Weighted averaging reduces the impact of DAC errors.

$$\frac{\sqrt{\overline{c_v^2}}}{2^{-k}} = \sqrt{\frac{2^{-2r+k}P}{3\sum_{i=1}^{P}\left(i^2+3i\right)^2}} \quad . \tag{2-80}$$

Results are plotted in Figure 2-28 for r=k=1 and 7. As r and k increase, the size of DAC elements is reduced and the relative accuracy required from each decreases.

The accuracy requirements above assume that the same DAC circuits are used in each stage so that all DACs contribute equal rms noise. However, different c_v values are amplified differently in the converter result. With an error-averaging decimator, earlier errors contribute more heavily than later ones. With a matched-filter decimator, later results contribute more heavily. As an example, consider the error-averaging decimator with weights plotted as a function of n in Figure 2-29. Hardware, power, and area could potentially be reduced by including different DAC circuits in different stages to equalize their error

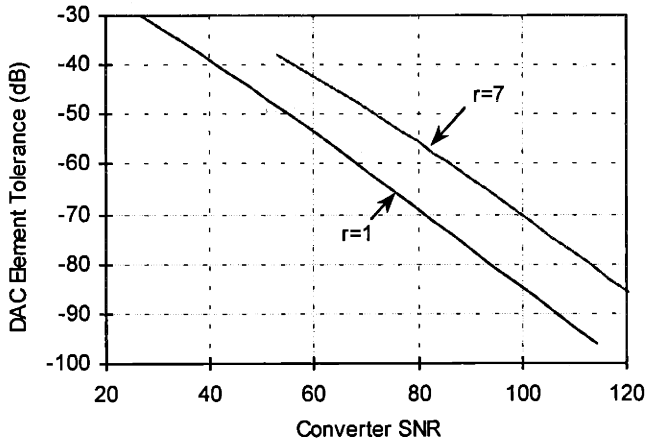


Figure 2-28. Tolerance to DAC element mismatches for a 2nd-order analog-integration architecture. Tolerance for each element is expressed relative to its nominal size.

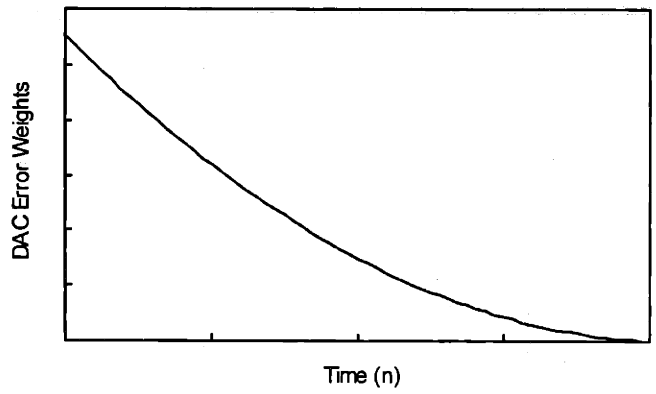


Figure 2-29. Amplification of DAC errors as a function of their time of occurrence, for a 2nd-order error-averaging decimator. An inverse weighting can be used to equalize error contributions.

contributions. In this case, each stage's accuracy should be inversely proportional to its amplification in the decimator.

2.4.5 DAC Matching for Digital Integration

DAC matching requirements are different for a digital-integration architecture such as that of Figure 2-17(c). For digital integration there are upper and lower DACs within the inner feedback loop. The outer loop of feedback occurs digitally and, therefore, contains no errors. In contrast to analog integration, a digital-integration architecture has DAC references that vary as a function of pipeline stage with values of $\left(n R_f\right)$. Mismatches in the upper and lower DACs, expressed as absolute quantities, are denoted by c_u and c_v, respectively. Mismatches, expressed as fractions of the full-scale reference are denoted by $c_u[n]/n$ and $c_v[n]/n$.

The modulator output is given by

$$w_k[n]=\frac{S_g}{R_f}\left(u[n-2]+c_u[n-1]-c_u[n-2]\right)+\left(-c_v[n-1]+c_v[n-2]\right)+ \\ \left(e_r[n]-2e_r[n-1]+e_r[n-2]\right)+\left(t_k[n]-2t_k[n-1]+t_k[n-2]\right) \tag{2-81}$$

The output from an error-averaging decimator equals

$$d_3[n]=\frac{S_g}{R_f}\left(\frac{1}{6}(n-4)(n-3)(n-2)u[n-4]+\sum_{i=1}^{n-4} i(n-3-i)\left(\frac{c_u[n-3-i]}{n-3-i}\right)\right) \\ +\sum_{i=2}^{n-3} e_r[i]-\sum_{i=1}^{n-4} i(n-3-i)\left(\frac{c_v[n-3-i]}{n-3-i}\right) \tag{2-82}$$

DAC mismatches in a digital integration architecture experience 1st-order noise shaping and, as a result, have weights in (2-82) that increase linearly, rather than quadratically, with n. Matching requirements are determined by equating the mean-square values of the error terms in (2-82) to the mean-square value of quantization noise.

Mismatches in the upper path have no impact on converter resolution provided they are not large enough to overload the modulator. This is due to the fact that, since the digital input $nu[n]$ to the upper DAC is signal independent, precisely the same circuit elements are used for each sample that is processed. The c_u error term in (2-82) alters converter gain but the reconstruction formula in (2-52) can be modified to

$$S_g \approx d_3[P+4]\left(\frac{6R_f}{P^3+3P^2+2P}+\sum_{n=1}^{P}(i)c_u[n-3-i]\right) \tag{2-83}$$

to compensate for this. Since c_u does not enter any of the noise terms in (2-82), it does not introduce noise or nonlinearity. Mismatches will not introduce overloading effects when they are small compared to a k-bit quantization step size. In some cases this level of insensitivity can be used to advantage by deliberately introducing mismatches or different gains in the different stages.

The lower DAC in Figure 2-17(c) has more stringent accuracy requirements than the upper one because the combination of its elements that is used in each stage depends on the signal level. The value of c_v enters the error term in (2-82) and, consequently, introduces noise and nonlinearity. The weights applied to $c_v[n]/n$ in each stage are shown in Figure 2-30. The resulting expression,

$$\sum_{i=1}^{P} i^2(P+1-i)^2\frac{{c_v}^2[P+1-i]}{(P+1-i)^2}=\frac{2^{-2r}P}{12}\ , \tag{2-84}$$

is used to solve for the rms value of $c_v[n]/n$. The result,

$$\sqrt{\overline{\left(\frac{c_v[n]}{n}\right)^2}}=\sqrt{\frac{2^{-2r}P}{12\sum_{i=1}^{P} i^2(P+1-i)^2}}\ , \tag{2-85}$$

is plotted in Figure 2-31 as a function of desired resolution. Curves are shown for r=k=1 and 7. This chart shows, in dB, the variation that is tolerable from each DAC output with respect to its full-scale value.

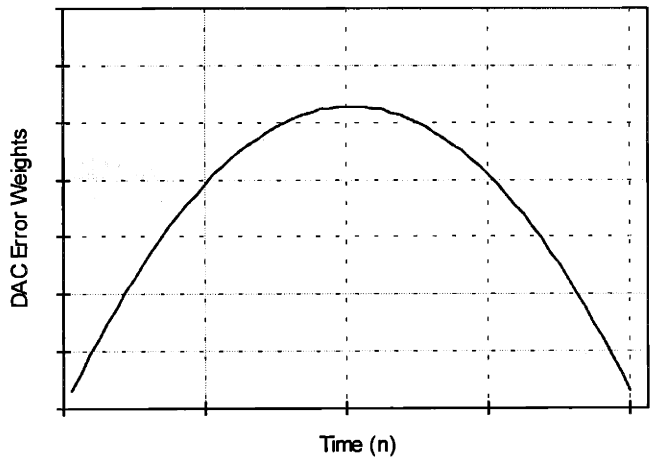


Figure 2-30. Amplification of DAC errors $c_v[n]$ / n at each pipeline stage.

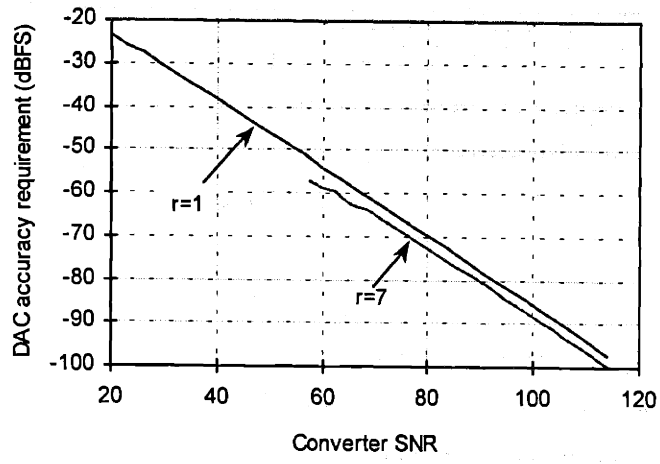


Figure 2-31. Tolerance to DAC errors, expressed with respect to their full-scale values, for a digital-integration architecture. Weighted averaging reduces the impact of DAC errors.

A comparison between Figure 2-27 and Figure 2-31 shows the similarity in DAC requirements of analog and digital-integration architectures. Mismatches in an analog-integration architecture are not subject to noise shaping and enter the converter result with the quadratic dependence in (2-78). Mismatches in a digital-integration architecture experience 1st-order noise shaping and their absolute weights in (2-82) have a linear dependence on n. On an absolute scale, the impact of mismatches in a digital-integration device are significantly attenuated by noise shaping. However, the accuracy required from each DAC with respect to its full-scale value is similar for both architectures.

2.4.6 Feedforward ADC and Feedback DAC Bits

Time-oversampling modulators often use multi-bit quantizers to improve their speed [28]. In this approach, each additional quantizer bit improves overall converter resolution, for a given oversampling ratio, by one bit. Expressed differently, each bit reduces the required oversampling ratio, thereby improving converter speed. Adding quantizer bits is less effective at reducing the converter oversampling ratio than increasing the modulator order, but it is an easy and power-efficient way to do so without incurring the difficulties of higher-order modulation.

A POSC does not experience improved speed as a result of multi-bit quantization because its speed and oversampling ratio are decoupled. However, adding quantizer bits impacts other device aspects such as hardware, power, and area. An increase in r decreases the pipeline length that is required for a given resolution. Fewer conversion blocks are required, but each conversion block has an increased number of stages and more circuit elements per stage. These competing effects partially offset one another. Any reduction in hardware is further offset when k is increased along with r.

Multi-bit feedback improves SNR by increasing the useful range of input signal amplitudes. Since this improvement is proportional to $\left(1-2^{-k}\right)$, it decreases rapidly with increasing k. Small values of k are sufficient to achieve a large fraction of the benefits to be had. The value k=2 is a good choice because it achieves a 2-X improvement in SNR over a single-bit configuration with only a moderate increase in DAC circuit complexity.

The greatest advantage of multi-bit quantization occurs when it is used in a configuration, such as that in [12], where the ADC and DAC are not constrained to have equal numbers of bits. In this approach, the number of ADC bits is increased while the number of DAC bits remains small. For a fixed k, the number of ADC bits should be increased until the random circuit noise term in (2-50) equals the quantization noise term and resolution ceases to improve with r.

Resolution versus ADC bits

Figure 2-32 shows hardware versus resolution for a 2nd-order POSC with 2-bit feedback. The measure of hardware is based on a configuration such as that in Figure 2-10. It is computed as

$$\text{Hardware} = (\#\ \text{of stages}) \cdot \begin{bmatrix} 2 \cdot (\text{Integrators}) + r \cdot (\text{Comparators}) + (2 + \text{r}) \cdot 3 \cdot (\text{Delays}) \\ 2 \cdot 2 \cdot (\text{DAC / Subtractions}) + 3 \cdot (\text{Digital Adders}) \end{bmatrix} . \qquad (2\text{-}86)$$

Each hardware element is assigned a weight of one. There is one comparator for each ADC bit. The number of delays increases with r because the number of stages per conversion block increases with this value. Curves are shown for r=2, 5, and 8.

Latency versus ADC bits

An increase in ADC bits can also be used to improve the sensitivity of a POSC to transfer noise. This is

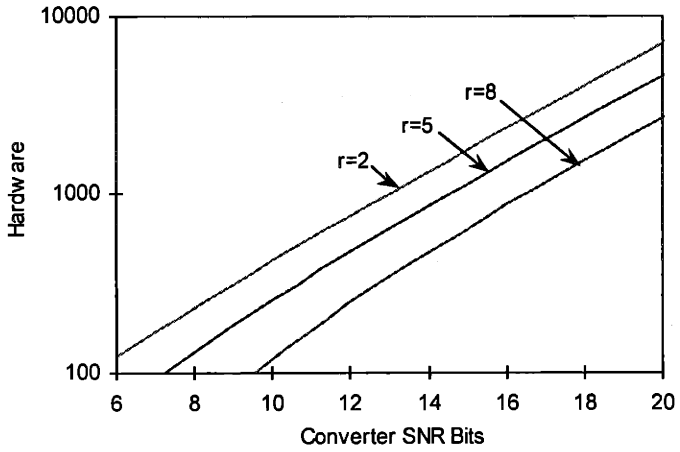


Figure 2-32. Hardware requirements versus resolution for a 2nd-order POSC with 2-bit feedback, shown for r=2, 5, and 8. Multi-bit quantization reduces total converter hardware.

accomplished when the pipeline length is decreased because noise is introduced by circuits during each transfer and each sample-and-hold operation. Each additional ADC bit reduces the number of conversion blocks but introduces an additional stage of delay into each block. The total number of transfers, which equals the pipeline latency, for r=2, 5, and 8 is plotted in Figure 2-33 as a function of converter resolution.

2.4.7 Transfer Inefficiency

Time-oversampling ΔΣ architectures are known to be relatively tolerant of leakage in their integrators [31]. Integrators lie in the feedback path for the loop between quantization noise signals and the output. Their dc poles produce dc zeros for the quantization noise transfer function and finite gain causes increased low frequency noise within the passband of the decimator. The impact of this noise on resolution has been shown to be small, provided the dc integrator gain is at least as large as the converter oversampling ratio [27]. The effect of nonideal integrator characteristics may be problematic for modulators of order greater than two, where gain impacts the position of the modulator poles and instability can result if actual and anticipated gains are not well matched.

Like a time-oversampling converter, a pipelined converter is tolerant of integrator leakage. In a pipelined device, integrator leakage is more appropriately referred to as transfer inefficiency because it is introduced when signals are transferred from one pipeline stage to the next with an inadequate slew rate and less than unity gain. Tolerance to transfer inefficiency is an important attribute for a POSC because a large number of transfers are performed and their inaccuracies compound along the pipeline.

Consider the 1st-order implementation of Figure 2-9 with nonideal transfer efficiency in both the signal and integrator channels. Operations in the signal channel are assumed to have an arbitrary transfer

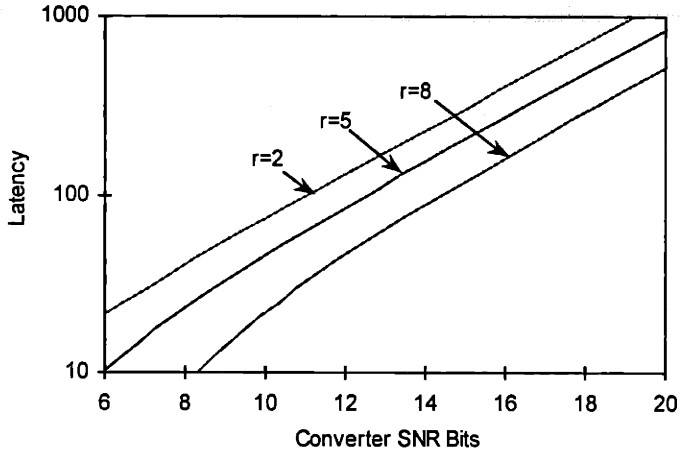


Figure 2-33. Pipeline latency versus resolution for a 2nd-order POSC with 2-bit feedback, shown for r=2, 5, and 8. Multi-bit quantization reduces the number of transfer operations.

function $H(z)$. Values of the delayed input sample are given by

$$S_i = S_g\, H(z) \quad . \tag{2-87}$$

An integrator gain of α results in the transfer function

$$I_1 = \left(S_i - W_k\, R_f\right)\frac{z^{-1}}{1-\alpha z^{-1}} \tag{2-88}$$

for the integrator channel. The corresponding modulator output is

$$W_k = \frac{S_g}{R_f}H(z)\frac{1}{\left(1+(1-\alpha)z^{-1}\right)} + (E_r - T_k)\frac{\left(1-\alpha z^{-1}\right)}{\left(1+(1-\alpha)z^{-1}\right)} \quad . \tag{2-89}$$

and the output from an error averaging decimator is given by

$$D_2 = \frac{S_g}{R_f}H(z)\frac{1}{\left(1+(1-\alpha)z^{-1}\right)\left(1-z^{-1}\right)^2} + E_r\frac{\left(1-\alpha z^{-1}\right)}{\left(1+(1-\alpha)z^{-1}\right)\left(1-z^{-1}\right)^2} \quad . \tag{2-90}$$

Signal Channel Gain Requirements

$H(z)$ is present in the signal term but not the error term. Provided it is signal independent, deterministic and close enough to unity to prevent modulator overloading, $H(z)$ alters converter gain but does not produce noise or distortion. This is true regardless of its form since only the total signal energy, not its spectral distribution, determines the final result. The modified converter gain can be compensated for if the reconstruction formula, given by (2-13), is adjusted to account for $H(z)$. Transfer inefficiency can even improve performance by eliminating the potential for pattern noise.

Integrator Channel Gain Requirements

Transfer inefficiency in the integrator channel is more severe. It impacts the quantization noise term in (2-89) and generates an increase in low frequency quantization noise. Simulations of this effect are shown in Figure 2-34 for a 1st-order architecture in (a) and a 2nd-order architecture in (b) with r=3. Curves indicate points for which SNR is degraded from its nominal level by 1, 2, and 3 dB, respectively. For example, a 2nd-order converter designed for 80-dB SNR experiences a 2-dB reduction in SNR for transfer gain of 0.98.

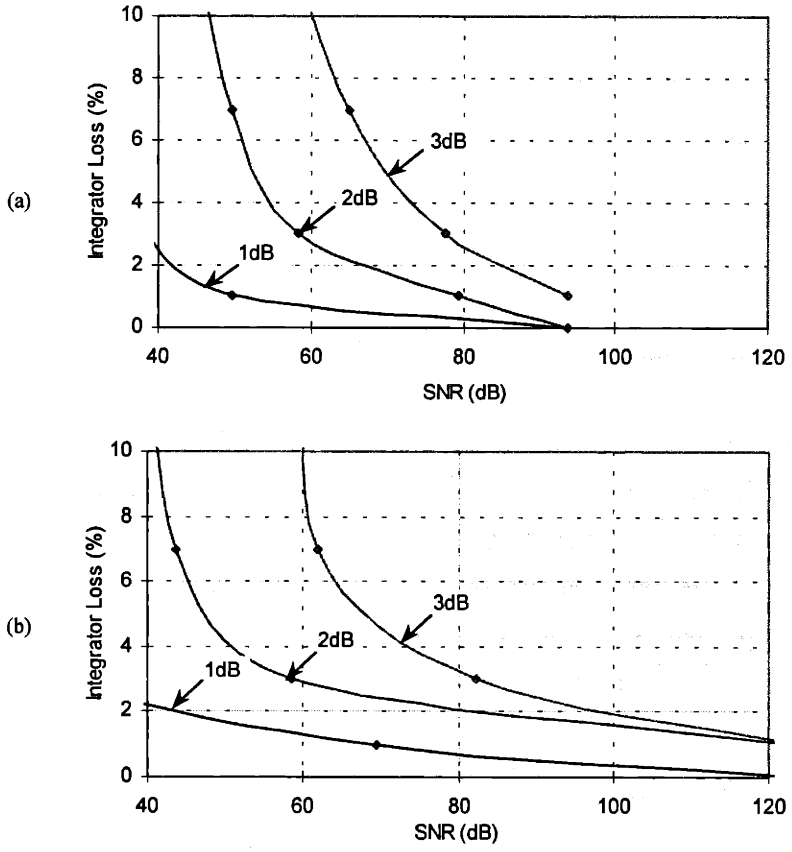


Figure 2-34. Simulated POSC integrator leakage requirements. Curves indicate tolerable loss for a 1, 2, and 3-dB reduction from theoretical SNR. (a) 1st-order. (b) 2nd-order.

2.4.8 Transfer Noise

Noise that is introduced during transfer operations is a critical concern for a pipelined architecture. Its impact is described with reference to the pipelined 1st-order implementation of Figure 2-9. Two sources of transfer noise are considered. That in the signal channel is denoted by c_s and that in the integrator channel is denoted by c_i. Both c_s and c_i are defined as quantities normalized to the full-scale reference R_f. The converter's input sample, captured by the first pipeline stage, is assumed to precisely represent the signal to be quantized. From that point on, noise compounds with each successive transfer. The value of s_i at time n,

$$s_i[n] = S_g\, u[n] + R_f \sum_{i=0}^{n} c_s[i] \ , \tag{2-91}$$

contains a sum of noise terms from all previous stages and has an rms noise that increases as the square root of n. Integrator i_1, given by

$$i_1[n]=\begin{cases}0 & \text{for } n\le 0\\ i_1[n-1]+S_g\,u[n-1]+R_f\sum_{i=0}^{n-1}c_s[i]+c_i[n-1] & \text{for } n\ge 1\end{cases}, \tag{2-92}$$

generates a sum of c_i and an amplified sum of c_s. The modulator output is given by

$$w_k[n]=\frac{S_g}{R_f}u[n-1]+\sum_{i=0}^{n-1}c_s[i]+c_i[n-1]+(e_r[n]-e_r[n-1])-(t_k[n]-t_k[n-1]) \tag{2-93}$$

and the output d_2 from the error-averaging decimator equals

$$d_2[n]=\frac{S_g}{R_f}\frac{1}{2}(n-2)(n-1)u[n-2]+\sum_{i=0}^{n-3}(n-2-i)c_i[i]+\sum_{i=0}^{n-3}\frac{1}{2}(n-2-i)(n-1-i)c_s[i]+\sum_{i=1}^{n-2}e_r[i]\ . \tag{2-94}$$

Figure 2-35 shows simulated SNR as a function of pipeline stage for different values of c_i in (a) and c_s in (b). SNR follows the theoretical curve closely until transfer noise begins to dominate over quantization noise. Thereafter the curve in (a) increases more slowly at a rate determined by averaging and that in (b) decreases because errors are amplified more rapidly than the signal. A resolution requirement for signal and integrator channel transfers is determined by setting the rms value of transfer noise terms at the converter output equal to that of quantization noise.

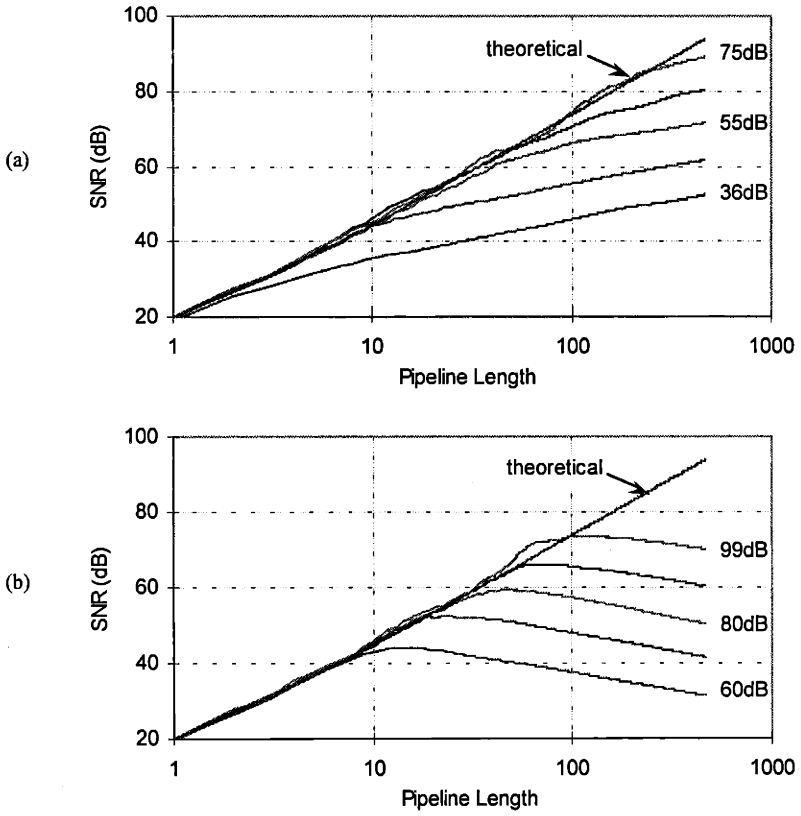


Figure 2-35. Simulated SNR as a function of pipeline stage in the presence of (a) signal channel transfer noise and (b) integrator channel transfer noise.

First-order transfer requirements

First, c_i is considered. Equating noise terms yields the expression

$$\overline{c_i^2}\frac{1}{6}(P)(P+1)(2P+1)=\frac{2^{-2r}P}{12} \quad . \tag{2-95}$$

Equation (2-95) is solved for $\sqrt{\overline{ci^2}}$ as a function of pipeline length P. Then (2-21), which relates SNR to P, is used to express $\sqrt{\overline{ci^2}}$ in terms of SNR under the approximation $P>>1$. Noise due to integrator transfers does not limit converter resolution provided

$$\sqrt{\overline{c_i^2}} < \frac{6^{1/3}2^{-r/3}\left(1-2^{-k}\right)^{2/3}}{4^{7/6}}\text{SNR}^{-2/3} \quad . \tag{2-96}$$

Equating noise terms for c_s yields the expression

$$\overline{c_s^2}\frac{1}{(36)(4)}\left(7P^5+38P^4+53P^3+46P^2\right)=\frac{2^{-2r}P}{12} \quad . \tag{2-97}$$

This is solved for $\sqrt{\overline{cs^2}}$ and (2-21) is used to express the result in terms of SNR for the case when $P>>1$. Transfer noise in the signal channel does not limit resolution when

$$\sqrt{\overline{c_s^2}} < \frac{3^{1/2}2^{r/3}6^{2/3}\left(1-2^{-k}\right)^{4/3}}{7^{1/2}4^{5/6}}\text{SNR}^{-4/3} \quad . \tag{2-98}$$

Figure 2-36 shows these results for r=k=3. Solid lines represent the theoretical expressions in (2-96) and (2-98). Circles and triangles show results from simulation. The SNR needed from integrator channel transfers increases slower than the overall converter SNR as a result of noise averaging. The SNR needed from signal channel transfers increases faster than the overall converter SNR because an additional stage of integration exists in the transfer function for c_s.

Second-order transfer requirements

A similar process of equating noise terms is used to determine transfer noise requirements for the 2nd-

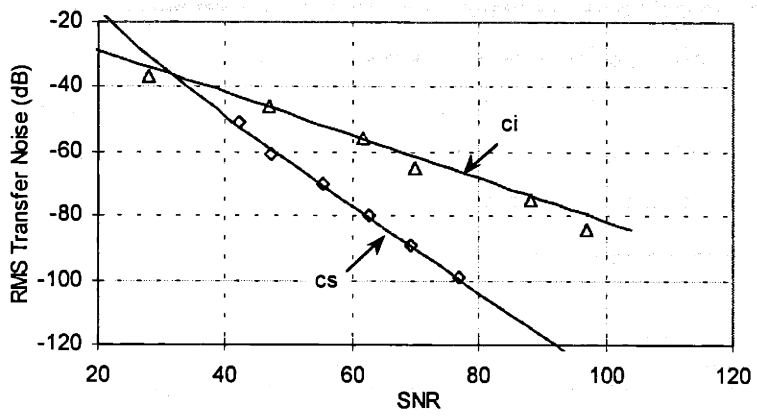


Figure 2-36. Theoretical and simulated SNR of a 1st-order POSC as a function of pipeline stage in the presence of signal and integrator channel transfer noise.

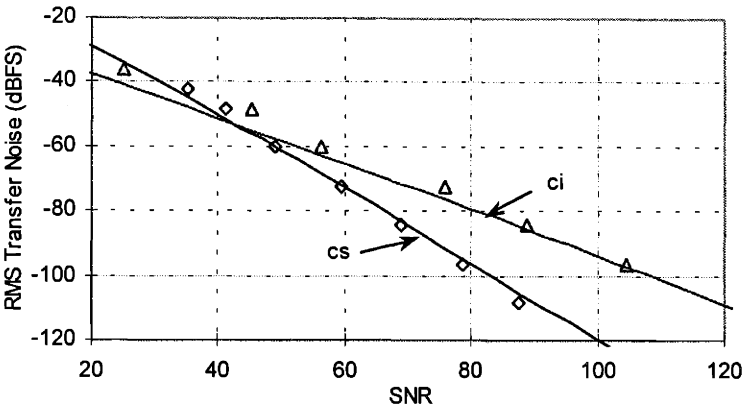


Figure 2-37. Theoretical and simulated SNR of a 2nd-order POSC as a function of pipeline stage in the presence of signal and integrator channel transfer noise.

order POSC implementation in Figure 2-16. Only noise in the first integrator is considered because the effect of noise in the second integrator is reduced by the gain of the first integrator. The output from a 2nd-order modulator is given by

$$w_k[n]=\frac{S_g}{R_f}u[n-1]+\sum_{i=0}^{n-2}c_s[i]+c_i[n-1]+(e_r[n]-2e_r[n-1]+e_r[n-2])-(t_k[n]-2t_k[n-1]+t_k[n-2]) \quad (2\text{-}99)$$

The corresponding output, d_3, from an error-averaging decimator is

$$\begin{aligned} d_3[n]=&\frac{S_g}{6R_f}(n-4)(n-3)(n-2)u[n-4]+\sum_{i=2}^{n-3}e_r[i] \\ &+\sum_{i=0}^{n-5}\frac{1}{2}(n-4-i)(n-3-i)c_i[i]+\sum_{i=0}^{n-5}\frac{1}{6}(n-5-i)(n-4-i)(n-3-i)c_s[i] \end{aligned} \quad (2\text{-}100)$$

The SNR requirement for c_i, given by

$$\sqrt{\overline{c_i^2}}<\frac{3^{1/10}\,2^{-r/5-1/5}\left(1-2^{-k}\right)^{4/5}}{7^{1/2}}\mathrm{SNR}^{-4/5}, \quad (2\text{-}101)$$

increases slower than linearly with converter SNR because of averaging. Figure 2-37 shows theoretical and simulated requirements for transfer noise in the signal and integrator channels. The results for a 2nd-order architecture differ slightly from those for 1st-order. For a given resolution, a 2nd-order device has fewer pipeline stages, and therefore fewer transfers. Requirements for c_i are more stringent because integrator noise experiences less averaging. On the other hand, c_s requirements are more relaxed because signal channel errors have less opportunity to compound.

2.4.9 Modulator Overloading

The 1st and 2nd-order configurations shown in Figure 2-1 and Figure 2-11 contain separate integrators for signal and reference quantities. No subtraction is included before integration and subtraction occurs implicitly at the input to the feedforward ADC. Such a configuration is not practical because a common-

mode accumulation builds with each feedback operation. An actual device should include subtraction before integration.

For example, consider the device in Figure 2-11. The values of u_1 and v_1 increase linearly with n with a slope equal to the input sample S_g. The values of u_2 and v_2 increase quadratically with n. The worst-case common-mode value occurs for $S_g = R_f\left(1-2^{-k}\right)$ and n=P at which point

$$u_2 = v_2 = \frac{R_f\left(1-2^{-k}\right)}{2}(P)(P-1) \ . \tag{2-102}$$

As a function of SNR, this common-mode equals

$$\max(u_2) = \max(v_2) \approx R_f \cdot \frac{1}{2}\left(\mathrm{SNR}\frac{2\sqrt{6}}{2^r\left(1-2^{-k}\right)}\right)^{4/5} \ . \tag{2-103}$$

Two difficulties arise for such a configuration. First, comparators in the coarse ADC must resolve differentials on the order of R_f upon a much larger common-mode signal that grows as $\mathrm{SNR}^{4/5}$. Second, integrators must be capable of representing signals as large as $u_{2\max}$ and $v_{2\max}$.

To eliminate the buildup of a common-mode offset, signal and reference quantities in an analog-integration device should be subtracted before integration occurs. Examples of this approach for 1st and 2nd-order modulation were shown above in Figure 2-9 and Figure 2-16. Separate u_1 and v_1 integrators are replaced by a single integrator, i_1, which operates on their difference. Similarly, u_2 and v_2 integrators are replaced by i_2. The maximum integrator range required for these configurations is described below.

The value of i_1 for a 1st-order architecture equals

$$i_1[n] = S_g\, u[n-1] - R_f\left(e_r[n-1] - t_k[n-1]\right) \ . \tag{2-104}$$

The random noise component of e_r is assumed to be small compared to quantization noise. The deterministic error component of i_1 has a value of

$$\max\left(\left|e_r[n] - t_k[n]\right|\right) = 2^{-k-1} \tag{2-105}$$

that is a function of the number of feedback bits k but is independent of r.

The modulator can experience overloading in two ways. The first, referred to as quantizer clipping, occurs when integrator values lie outside of the feedforward ADC's quantization range. The second, referred to as integrator clipping, occurs when the integrator attempts to produce a value outside of its linear region of operation. Either of these situations arise when input samples lie outside of the range

$$2^{-k-1} R_f < S_g < \left(1-2^{-k-1}\right)R_f \ . \tag{2-106}$$

When quantizer clipping occurs, quantization noise is larger than that given by (2-18) and SNR is reduced. $\Delta\Sigma$ architectures are relatively forgiving of such errors. Integrator clipping is more serious. It

introduces noise that is indistinguishable from incoming signals and causes a rapid degradation in converter performance.

For an ADC with a full-scale reference of R_f, the range of input samples over which the converter is free from overloading effects is $\left(1-2^{-k}\right)R_f$. This approximation is used for both 1st and 2nd-order architectures [20]. In practice, integrators are often given a larger operating range with the first integrator typically clipped at $2R_f$ and the second integrator clipped at $4R_f$.

3 Charge-Domain Circuit Techniques

3.1 Overview of CCD Elements

Charge-coupled device (CCD) elements have a unique set of capabilities for analog signal processing. Their operations are classified into two categories, fully-depleted, and nondepleted, which differ fundamentally in nature. Those in the first category are performed with extremely high accuracy. But only a few processing functions reside in this group. In comparison to depleted functions, nondepleted functions are performed with less accuracy. But a larger, more varied set of them is available. Both depleted and nondepleted CCD circuits are described below.

3.1.1 Bucket Diagrams

The operation of CCD elements is often represented by a simple hydraulic model and depicted by bucket diagrams [39]. Hydraulic models are a very valuable tool for analyzing circuit elements and are used throughout this chapter. An illustration of the bucket analogy, as it applies to n-type surface channel devices, is shown in Figure 3-1. Lower levels in the figure represent increasing potential and decreasing electron energy. MOS structures in deep depletion have a surface potential that increases nearly linearly with gate voltage. This is represented by the depth of a potential well, or bucket, formed beneath the overlying gate. An increase in potential on the gate increases the surface potential and results in a deeper well that is attractive to electrons. The wells corresponding to three possible gate voltages are shown in (a).

As inversion charge is added beneath the gate, the surface potential decreases nearly linearly with the amount of charge. This is analogous to pouring liquid into the well, as shown in (b). The surface potential in the presence of inversion charge is represented by the height of the surface of the liquid. It corresponds to the most negative potential of electrons stored underneath the gate. Charge can be added until the surface potential decreases to twice the Fermi level, ϕ_F, at which point the surface underneath the gate

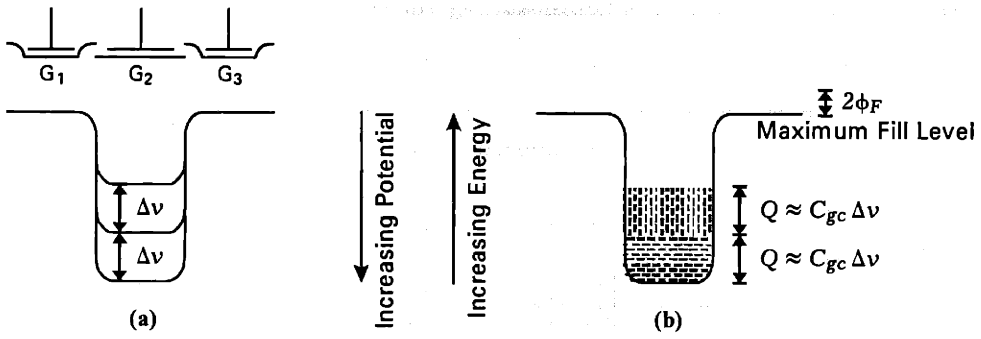


Figure 3-1. CCD bucket diagram representation. (a) Surface potential varies nearly linearly with gate voltage. (b) Surface potential varies nearly linearly with inversion charge.

enters weak inversion. A difference in gate voltage of Δv produces a difference in channel potential of approximately Δv and is equivalent to adding or removing a charge of $Q \approx C_{gc}\,\Delta v$ electrons from a well with gate-to-channel capacitance of C_{gc}.

Bucket diagram representations are a useful qualitative tool for both buried and surface channel CCDs. They provide an accurate quantitative description of surface channel CCDs, where distance between the charge centroid and the gate does not vary appreciably with channel charge. In buried channel devices, they are less accurate because surface potential varies more nonlinearly with gate voltage and channel charge.

3.1.2 Fully-Depleted Operations

Nondepleted MOS Transfers

Depleted CCD transfers are illustrated by comparing them to nondepleted transfers in an MOS transistor. A bucket diagram representation of an NMOS switching operation is shown in Figure 3-2. The drain is initially at a 5-V potential and the source is held at a fixed bias of 1V. In (b), the potential on gate G_1 is raised, charge floods the channel region, and the drain is rapidly equilibrated with the source. In (c), the potential on G_1 is lowered, the channel region turns off, and the source and drain are isolated.

This circuit is subject to a number of noise sources. First, thermal noise, which is due to random fluctuations in the quantity of charge present on the drain, is captured on the drain when G_1 is shut off. Second, clock feedthrough, caused by coupling of the negative transition on the gate through parasitic gate-to-drain capacitance, alters the potential on the drain. Third, charge injection is introduced when the

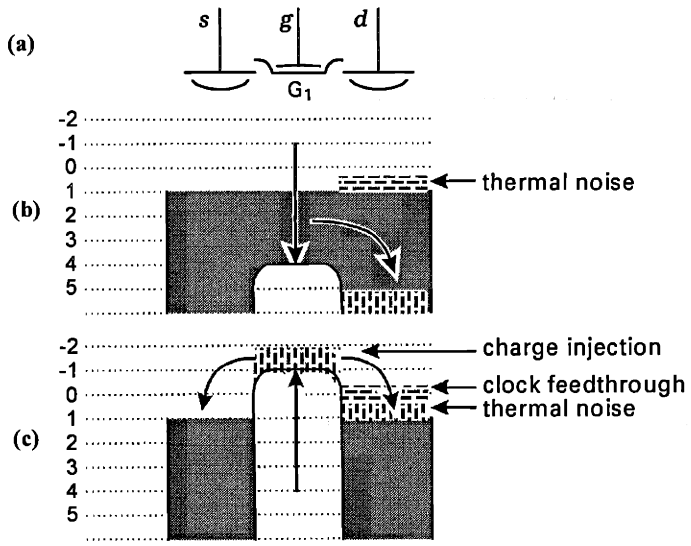


Figure 3-2. Nondepleted MOS transfer. (a) Schematic diagram. (b) The source and drain equilibrate when the gate is high. (c) Various noise sources impact the drain when the gate is turned off.

The process of transferring a signal forward through this register is illustrated in Figure 3-3(b) through (e). During phase 1, a falling transition on ϕ_{s3} causes the charges underneath G_2 and G_6 to be transferred forward. In phase 2, a falling transition on ϕ_{b4} confines charges in their storage wells. During phase 3, the potential of ϕ_{s3} is raised and charges are transferred forward into the wells of G_2 and G_6 upon the falling transition of ϕ_{s1}. Finally, in phase 4, the potential on ϕ_{b2} is lowered and charges are confined to their new storage wells.

In this circuit, no source or sink of carriers exists between the different CCD elements. Instead, the channel region within each well provides electrons to, and accepts electrons from, its neighboring wells. As a charge packet is transferred forward, all its electrons are removed and the supply of carriers in the source well is depleted. The efficiency with which such transfers are performed is typically 0.9999 or better [64].

The fully-depleted nature of CCD transfers brings about a number of unique advantages. First, thermal noise is eliminated because no carriers remain in the originating well at the end of a transfer. Second, no charge injection occurs because the barrier wells are empty when they are turned off. The built in offset between barrier and storage gates ensures that, even though storage wells contain charge when they are turned off, all charge transfers in the same direction. Finally, CCDs have a unique attribute of an analog noise margin, similar to the noise margin enjoyed by digital logic. Any voltage-mode interference that couples to the CCD channel region, such as clock feedthrough or substrate noise, changes the potential of electrons in the channel. But as long as this coupling remains less than a threshold value, at which electrons spill outside the well, it does not change the number of electrons contained inside the wells. Gate-to-drain capacitance, which is considered detrimental in MOS circuits, is desirable in CCD circuits, where its source of fringing fields aids in charge transfer.

Depleted Charge Addition, Integration, and Division

The circuit in Figure 3-3 implements both a charge transfer and a sample-and-hold. Other CCD functions that share this same fully-depleted characteristic are shown schematically in Figure 3-4. The circuit in (a) performs charge addition by combining the charges from two incoming wells, G_1 and G_2, into a single well by shifting them through the same set of gates. The output from this circuit is given by

$$Q_3 = Q_1 + Q_2 \ . \tag{3-1}$$

The charge integration circuit in (b) utilizes a similar process of addition, but does so repeatedly over multiple stages to accumulate packets from each stage. Its output equals

$$Q_4 = Q_1 + Q_2 + Q_3 \ . \tag{3-2}$$

The circuit in (c) performs charge division by shifting a single well of charge from G_2 into two receiving wells, G_3 and G_4 [36][37]. The charge division ratio of this circuit is proportional to the relative

gate is turned off and charge in the channel flows to the source and drain in unpredictable proportions. Each of these noise components is a significant concern for high-resolution CMOS circuits.

Depleted CCD Transfers

The origin of the noise sources mentioned above lies in the fact that the MOS transfer in Figure 3-2 is not fully depleted. More charge is present than is ultimately transferred. In contrast, CCD transfers are fully depleted and are not subject to thermal noise, charge injection, or clock feedthrough. An example of a CCD shift register is shown in Figure 3-3. CCD charge packets reside in potential wells created by MOS capacitors that are driven into deep depletion. Structurally, a CCD is identical to an MOS transistor. The only difference between MOS and CCD devices lies in the way that they are interconnected and used. The source or drain of an MOS transistor is a heavily doped diffusion that provides an infinite source of carriers to the device. The source or drain of a CCD is formed by placing adjacent wells sufficiently close that their underlying potential wells touch. Such a configuration is shown schematically in (a).

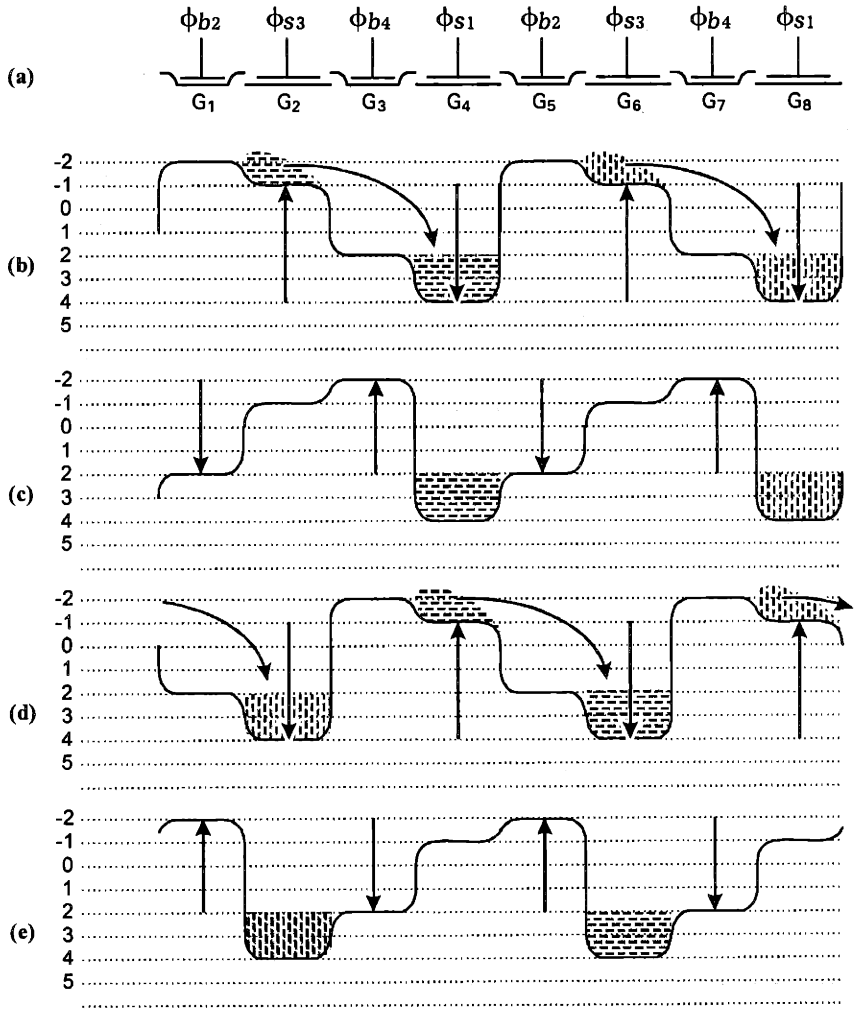


Figure 3-3. Fully-depleted CCD shift register. The supply of carriers is exhausted as charges are transferred forward. (a) Schematic diagram. (b)-(e) Four phases of operation.

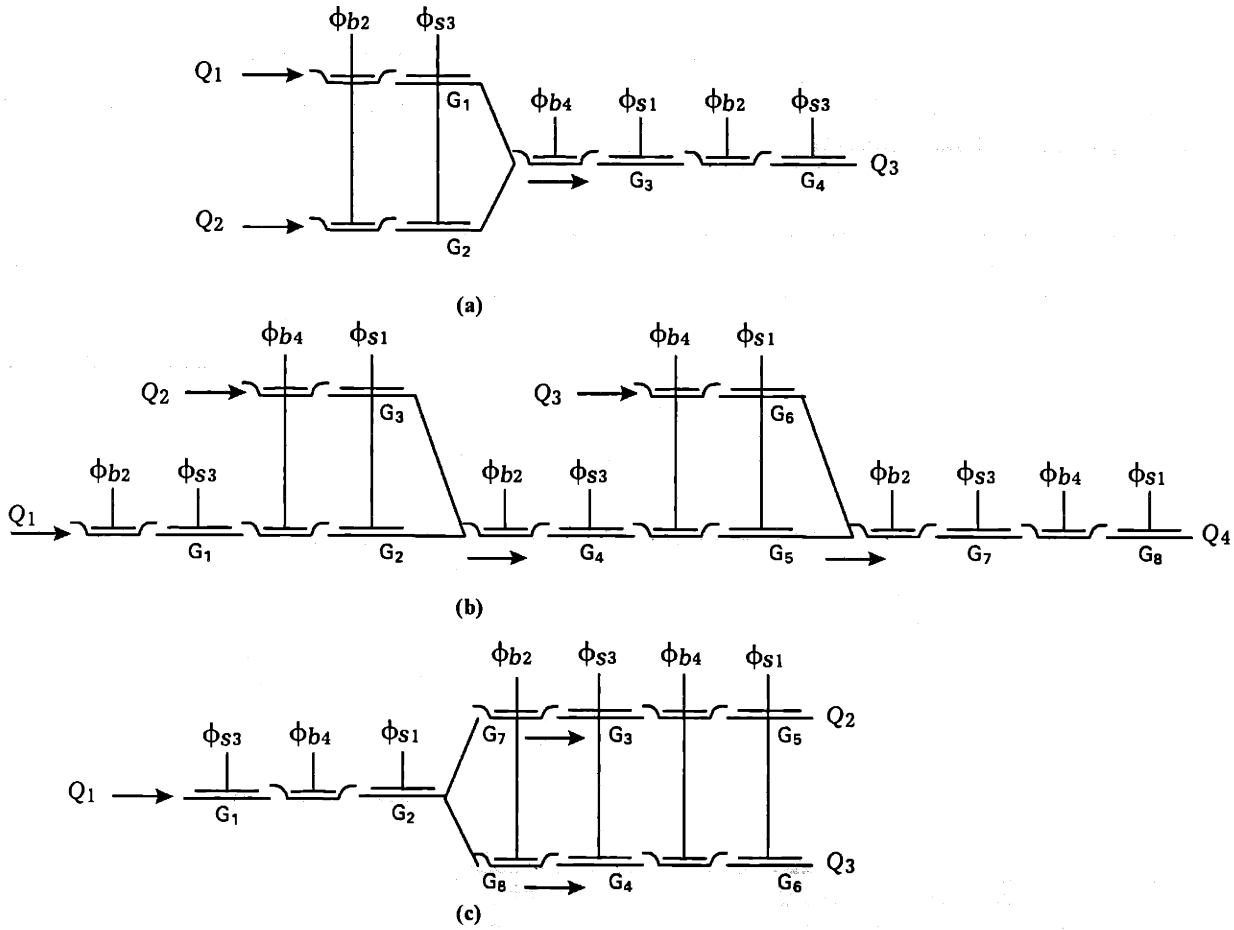


Figure 3-4. Other CCD functions that are fully depleted and highly accurate. (a) Addition. (b) Integration. (c) Fixed ratio division.

transconductances of the barriers, G_7 and G_8, through which the charge is split. The two outputs from this circuit have values of

$$Q_2 = Q_1\left(\frac{W_7}{L_7}\right)\left(\frac{L_7L_8}{W_7L_8 + W_8L_7}\right)$$
$$Q_3 = Q_1\left(\frac{W_8}{L_8}\right)\left(\frac{L_7L_8}{W_7L_8 + W_8L_7}\right) \tag{3-3}$$

and their relative ratio is

$$\frac{Q_2}{Q_3} = \left(\frac{W_7L_8}{L_7W_8}\right) . \tag{3-4}$$

When the transconductances and width to length ratios of the barrier gates G_7 and G_8 are equal, two identical packets are generated, each with half the original size.

3.1.3 Nondepleted Transfers

Complex signal processing devices require other functionality in addition to that of the circuits in Figure 3-3 and Figure 3-4. Examples of these functions include charge generation, wire transfer, charge sensing, charge comparison, subtraction, and D/A conversion. These functions require a source of carriers and are, therefore, nondepleted.

An example of a nondepleted operation, is a wire transfer. It is used when charge packets must be transferred between nonadjacent or distant wells, via a metal line. The conventional circuit for performing wire transfers is shown in Figure 3-5. It contains a combination of a CCD register, formed by gates G_1, G_5, G_6, and G_7, and a single-stage bucket brigade, consisting of G_2, D_1, D_2, G_3, and G_4 [49][50]. The bucket brigade stage is inserted in the midst of the CCD register and is functionally equivalent to a single

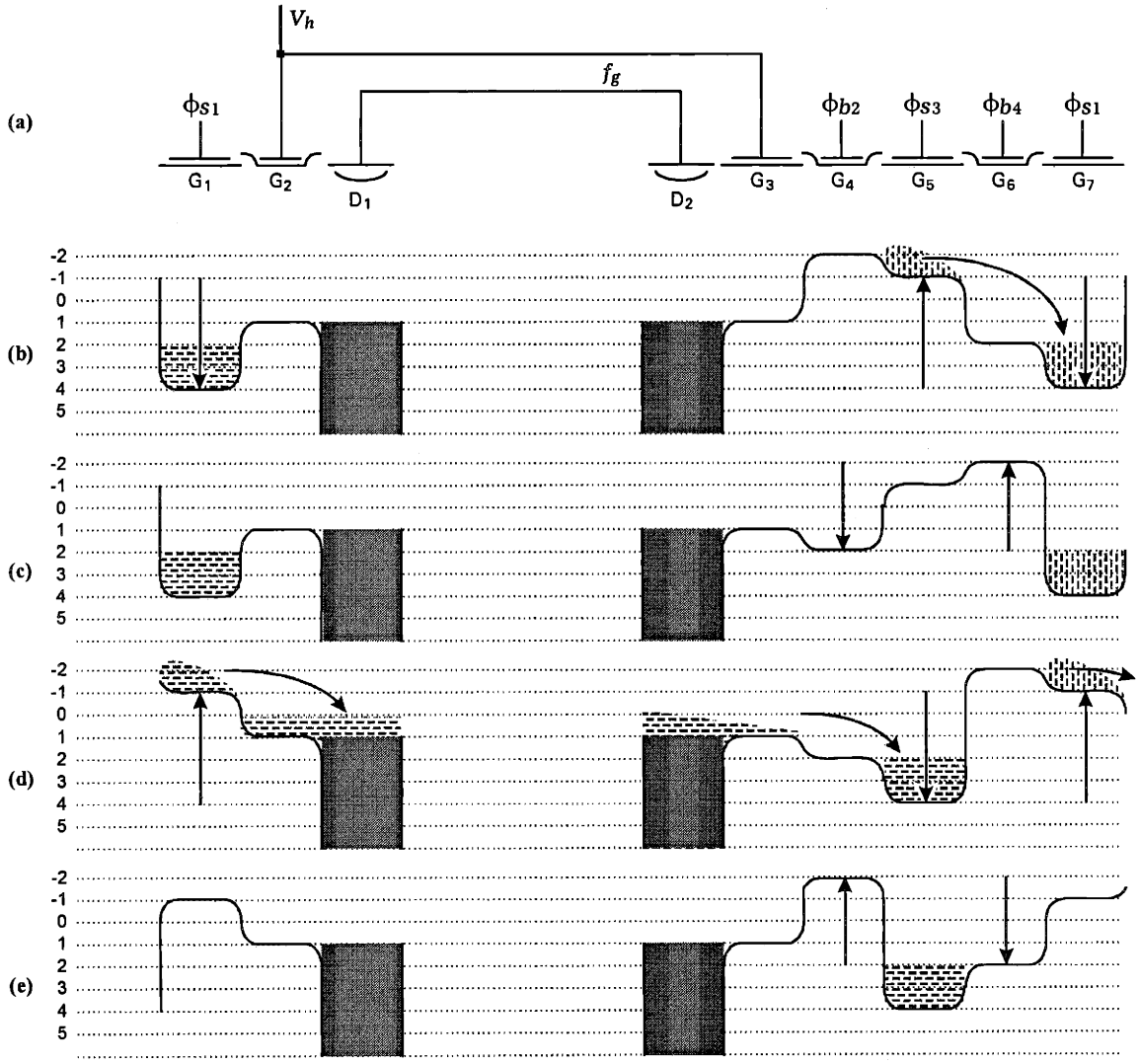


Figure 3-5. Nondepleted wire transfer operation. (a) Schematic diagram. (b)-(e) Four phases of operation. Linearity and speed are poor.

barrier gate. Gates G2 and G3 are tied to the fixed bias, V_h.

The operation of this circuit during four clock phases is illustrated in (b) through (e). During phase 1, incoming charge is received by the well underneath G1. At the same time, the packet underneath G5, which was generated during the previous cycle, is transferred forward to G7. During phase 2, the potential of the G4 barrier is raised, thereby opening the transfer path. Node f_g remains floating during this time at a potential that was set during the previous clock cycle.

Charges are transferred across the wire during phase 3. The potential of G5 is raised in anticipation of it receiving charge. Meanwhile, the potential of G1 is lowered and charge flows onto f_g, which falls temporarily in response. The current from the receiving well to f_g is determined by the voltage $(V_h - v_{th} - f_g)$, where v_{th} denotes the threshold of G3. As current flows, the voltage on f_g rises, and G3's gate-to-source voltage and current are reduced. Eventually, the potential on f_g approaches $(V_h - v_{th})$, equal to the channel potential underneath G3, and current flow ceases. In the final phase, the output packet underneath G5 is isolated by lowering the voltage on G4.

The speed of the wire transfer circuit in Figure 3-5 is severely limited. Charge transfer through G3 occurs rapidly at first but slows considerably as the operation progresses. Theoretically the time constant associated with this operation grows infinitely large and the transfer of charge through f_g is never fully completed. At high speed, a residual portion of the signal is left behind in each cycle. This is referred to as signal lag. Lag is a source of nonlinearity and its magnitude increases exponentially as cycle times are reduced. These speed and linearity limitations of the conventional wire transfer circuit are addressed by the circuit technique presented later in this chapter.

3.2 General CCD/CMOS Circuit Approach

Although previous CCD signal processing devices have been demonstrated, the technology has not been widely accepted for analog signal processing. The limited success of CCDs in this regard is due to a number of factors. CCD devices were traditionally built in custom processes. The high cost and poor availability of these processes hindered development efforts. Most CCD processes are based on NMOS technology and do not include complementary MOS devices. Both analog or digital supporting circuits for the CCDs are awkward to implement in NMOS and are limited in capability. As a result, digital control circuitry is not easily incorporated on chip, and complex off-chip control hardware is required. The few CCD/CMOS processes that have been developed have met with limited success because of difficulties of custom fabrication and the fact that their technology lagged behind that of state-of-the-art CMOS by many generations. [34]

The performance of previous CCD signal-processing devices has been limited. Their power was determined by their operation at high voltages, often 10V or greater. Their speed was limited because of their need for off-chip control hardware and their dependence on slow clock fall times for good transfer efficiency. And their accuracy was limited because of the lack of established high-resolution CCD circuit techniques.

The approach taken here differs in a number of regards from that of other CCD devices. First, the availability of CMOS elements is considered crucial for such a device. The capabilities of CCD and CMOS devices complement one another. CCDs make pipelines with hundreds of stages feasible by contributing sample-and-hold, delay, transfer, and integration operations that are highly accurate, low power, simple and compact. CMOS also plays a vital role by providing digital logic, CCD support circuitry, and other analog functions such as comparison. A merged CCD/CMOS approach makes possible devices that would be difficult using either technology alone.

A second factor that is critical to this approach is compatibility with standard CMOS processing. Traditionally, CCD processing has included one or more buried channel implants, a lightly doped substrate, and, in some cases, an oxide-nitride gate insulator stack. These features are intended to improve CCD charge transfer efficiency, equalize thresholds between the two gate levels, and improve device yield. Although these provisions were necessary in the past, they are less critical today because material quality, implant uniformity, and fabrication yields have improved dramatically.

The most basic requirements for fabricating CCD circuits are identical to those for MOS devices and are met by standard CMOS processing. Many mixed-signal CMOS processes can be used to form overlapping CCD structures because they include two levels of polysilicon for capacitors. CCD gates are formed by placing second-level polysilicon over gate oxide as shown in Figure 3-6. In this configuration, second-level gates lie above both gate oxide and some amount of inter-poly oxide. The resulting second-

poly oxide has a thickness that depends on the fabrication method used for inter-poly oxide. But in most cases, second-level gates have a smaller gate capacitance and a higher threshold voltage than first-level gates.

Overlapping gate structures are ideal for CCDs, but they are not essential. Their advantage is that the gate separation, indicated in Figure 3-6, can be made small in comparison to the polysilicon thickness. If a gap of this size can be achieved by etching slots in a single level of polysilicon, then overlapping gates are not needed. Current lithographic capabilities are beginning to make single-poly CCDs possible.

The approach taken here assumes 5-V or 3.3-V operation. These voltages are desired for compatibility with state-of-the-art CMOS processing. They also allow competitive power performance. CCD clock voltage swings are determined by factors such as charge handling capacity, charge transfer efficiency, and desired operating speed. Lower voltages reduce the charge packet size that can be stored per unit area. They also decrease fringing fields in the CCD channel, thereby reducing the speed and accuracy of charge transfers. These detrimental effects are offset by the shorter CCD gate lengths made possible by state-of-the-art CMOS fabrication.

Finally, the approach taken here uses CCD clocking techniques that are capable of higher speed than that of traditional CCD devices. They are less sensitive to clock edge rates and do not demand slow fall times. Speed is also improved for nondepleted CCD operations using the circuit techniques described later in this chapter.

3.2.1 Voltage and Timing Conventions for this Chapter

A variety of electrode structures and clocking techniques, including two-phase, three-phase, two-and-a-half phase, and four-phase, are well known for CCD registers [35] [39]. These differ in their number of electrodes per stage, their number of independent clock waveforms, and in the waveforms applied to these clocks. The choice of an appropriate clocking scheme depends on many factors. Three-phase and four-phase clocking schemes, which are commonly used for CCD imagers, are not acceptable. First, they rely heavily on slow fall times for good transfer efficiency. Second, their slow edges are not well suited for merged CCD/CMOS circuits. Third, they are incompatible with the need for charge sensing elements because they store charge underneath multiple gates simultaneously. One-and-a-half phase clocking,

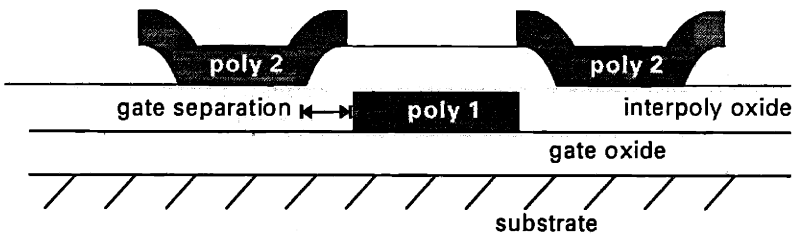


Figure 3-6. Overlapping gate structure in standard CMOS uses parasitic second-poly gates.

which has been previously used for CCD signal processing devices, is incompatible with low-voltage operation because it includes unclocked gates and, therefore, requires twice the operating voltage of other techniques. Finally, two-phase clocking meets all of the functionality needs of CCD/CMOS circuits, but is not suitable because it requires specialized processing.

A hybrid of these conventional clocking techniques is chosen for the circuits in this chapter. The clock period, denoted by P, is divided into four phases with an equal time of $H=P/4$ allocated to each phase. CCD and CMOS clocks are electrically independent, but have similar logical timing. The waveforms of all signals are shown in Figure 3-7. The energy level diagrams associated with this timing were described above with reference to Figure 3-3. In merged CCD/CMOS circuits, both CCD and CMOS elements require multiple clock phases but they have different clocking requirements. CMOS clocks, ϕ_1 through ϕ_4, are each asserted for one quarter of a cycle, during their respective phases. Signals $\overline{\phi_1}$ through $\overline{\phi_4}$ are negated versions of ϕ_1 through ϕ_4. Additional CMOS clocks, ϕ_{12} through ϕ_{41}, are asserted for one half of a cycle, during their respective phases. Signals $\overline{\phi_{12}}$ through $\overline{\phi_{41}}$ are negated versions of them. CCD storage clocks ϕ_{s1} and ϕ_{s3} are asserted for one half of a cycle, during phases 1 through 2 and 3 through 4, respectively. CCD barrier clocks ϕ_{b2} and ϕ_{b4} are asserted for one half of a cycle, during phases 2 through 3 and 4 through 1, respectively. CCD clock timing is chosen so that barrier gates transition low one phase

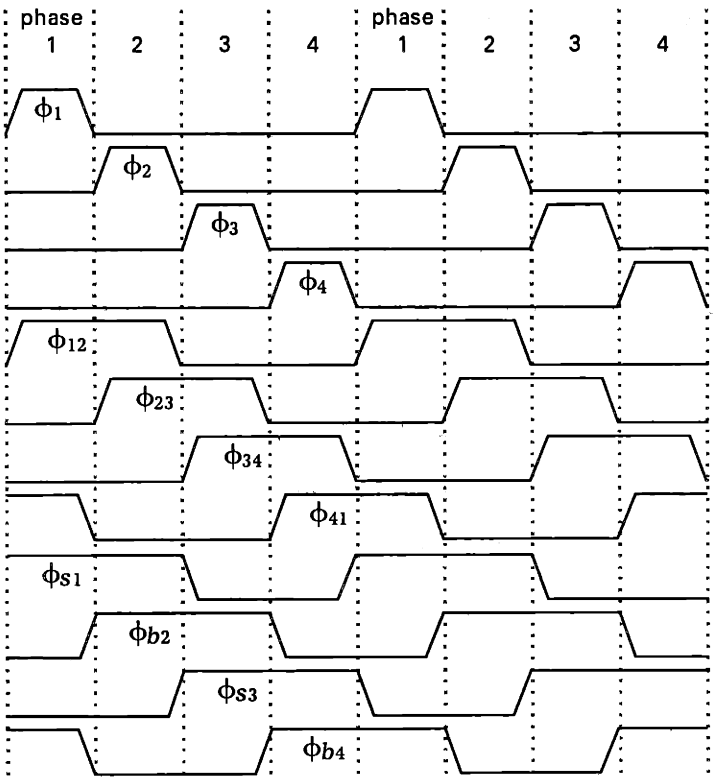


Figure 3-7. Four-phase CMOS and CCD control signal timing. CMOS clocks are nonoverlapping with 5-V swings. CCD barrier and storage clocks are overlapping with 5 and 4-V swings.

before their associated storage gates. This confines carriers in the storage well, prevents charge from spilling backward, and reduces sensitivity to clock skew and fall times.

These clocks are used to synchronize CCD and CMOS elements and, therefore, must be generated and distributed with minimal skew. All CMOS clocks are nonoverlapping. For example, the signal ϕ_2 does not overlap with clocks ϕ_3, ϕ_{34}, ϕ_1, or ϕ_{41}. Such a relationship is necessary to assure that operations during phase n are completed before those during phase $n+1$ begin. All CCD control signals are overlapping. For example, ϕ_{s1} overlaps with ϕ_{s3} at both its beginning and end. Overlapping clocks are needed for CCDs so that charge can be transferred forward from well n to well $n+1$ before well n is closed. All of these clock relationships are enforced using logical operations in a CMOS control block.

Threshold voltages for the first and second gate levels are assumed to be 1V and 2V, respectively. CCD storage wells are formed from the first gate level, while barrier gates are formed from the second gate level. Storage well clocks swing from 0 to 5V. Their empty-well channel potentials, therefore, have low and high levels of -1V and 4V. Barrier clocks swing from 0 to 4V and their empty-well channel potentials range from -2V to 2V. The application of these levels to register clocking can be seen in Figure 3-3. At their high levels, the 2-V offset in surface potential between barrier and storage gates assures that carriers are only stored under the storage gate. This is necessary for compatibility with charge sensing elements. At their low levels, the 1-V offset in surface potential between barrier and storage wells assures that all charge is transferred forward, rather than backward, upon a falling storage gate transition.

3.3 Conventional CCD Circuit Techniques

3.3.1 Charge Generation

Charge generation is necessary whenever CCD signal processing circuitry must interface with voltage-domain signals. A variety of techniques are well known for generating charge packets from voltages [36]. In all of these, an input diffusion is used to provide a source of electrons and the amount of charge that is collected from this source is controlled by either the source potential or the voltages on an initial set of register gates.

The surface channel fill-and-spill technique, shown in Figure 3-8, is the most commonly used method of charge generation [61]. In this approach, the input signal v_s is provided to gate G_2 and a fixed reference level V_r is placed on G_1. During the fill phase in (b), node f_g is pulsed low and charge is injected into the channel regions underneath G_1 and G_2. During the spill phase in (c), f_g is returned to a high potential so that excess charge from the output well spills back. The spill transition in this circuit is similar in nature to that in the wire transfer configuration described earlier. It occurs rapidly at first, but then slows considerably as it nears completion.

The channel potentials of G_1 and G_2 depend nearly linearly on v_s and V_r when they are empty. At the end of the spill process, when the output well contains charge, these channel potentials are equal and the size of the resultant charge packet is approximately

$$Q_2 = C_{gc}\left((v_s - v_{th2}) - (V_r - v_{th1})\right) , \tag{3-5}$$

where C_{gc} represents gate-to-channel capacitance and v_{th1} and v_{th2} represent the thresholds of G_1 and G_2, respectively. Matching between different fill-and-spill circuits is poor because they are sensitive to threshold variations, implant nonuniformity, and variations in CCD well capacitance.

The circuit shown in Figure 3-8 requires a presampled input because it does not perform accurate sampling and it has a rectifying characteristic. Otherwise, if a time varying input is applied, the result depends on both the signal's lowest voltage and its derivative.

At speeds above a few MHz, there is also a long time constant associated with the spill transition. For typical device parameters of current technology, this factor becomes significant at speeds above a few MHz.

Nonlinearity is introduced by the fill-and-spill technique because capacitance of the CCD receiving well is signal dependent. The nonlinearity of various charge injection techniques has been theoretically and experimentally studied [58]. Experimental results range from about 3% for the diode-cutoff method and 3% for a buried channel fill-and-spill circuit to about 1% for a surface channel fill-and-spill circuit.

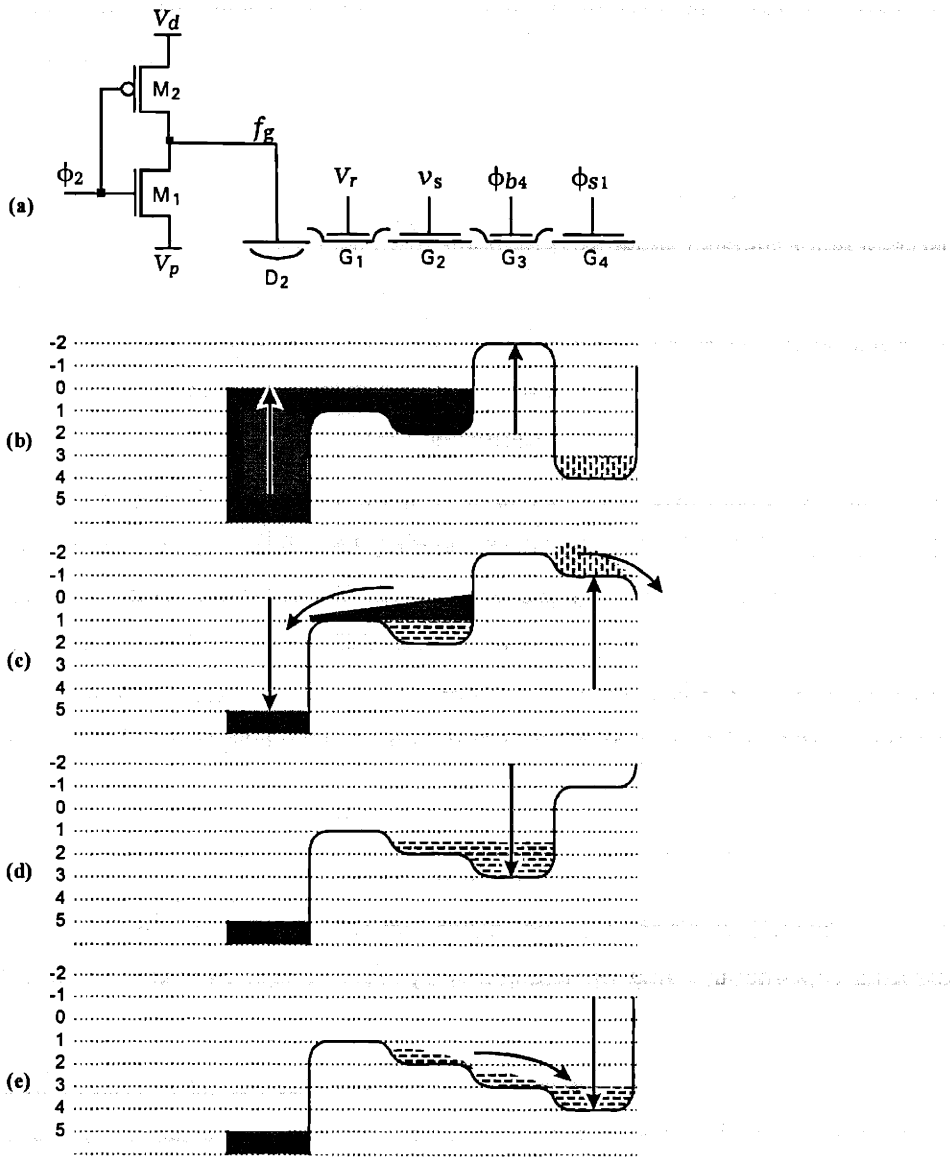


Figure 3-8. Conventional fill-and-spill charge generator. (a) Circuit schematic. (b) Phase 2. (c) Phase 3. (d) Phase 4. (e) Phase 1.

The linearity, speed, and matching limitations of conventional charge generation circuits and the need for accurate sampling are addressed by the alternative charge generation techniques presented in sections 3.4 and 3.5.

3.3.2 Charge Sensing

Charge-domain signals can be used by non-CCD elements in two ways. The first method, referred to as destructive sensing, involves transferring the charge packet onto a receiving element. In this case, the charge packet is consumed by the operation. The second method, referred to as nondestructive sensing, involves creating a separate representation of the charge packet, without altering the original. Multiple nondestructive operations can be performed on the same charge packet.

A conventional nondestructive floating gate amplifier circuit is illustrated in Figure 3-9 [62][45]. It includes a sensing gate G4 within a CCD channel. The sensing gate and the barrier that precedes it are not clocked. In preparation for receiving charge, node v_g is preset to the intermediate bias level V_h during the time that the G4 channel is empty. The precharge is then turned off and v_g is left floating. During the next phase, a charge packet is transferred underneath G4 by lowering the voltages on G1 and G2. As charge collects underneath G4, it couples through the gate-to-channel capacitance and causes the potential on v_g to fall. The final voltage on v_g is

$$v_g = V_h - Q_s \frac{C_{gc}}{C_{gc}C_{p2} + C_{cs}\left(C_{gc} + C_{p2}\right)} , \tag{3-6}$$

where C_{gc} and C_{cs} represent the gate-to-channel and channel-to-substrate capacitances of G4, and C_{p2} represents parasitic and load capacitance on v_g. The clocked v_g waveform is then buffered to form the voltage-domain output v_b.

This approach to charge sensing is not compatible with the low voltages associated with aggressive CMOS processing. The reason for this is that unclocked CCD gates, G3 and G4, reside in the charge transfer path. The remaining gates require twice the voltage swing of those in a register where all gates are clocked.

A second limitation of this circuit is a limited signal range. Charge carrying capacity of the floating gate well is proportional to the potential difference between G3 and G4. If this capacity is exceeded, charge collects underneath G3 and is not sensed by the floating gate. During precharge, this offset is $\left(V_h - V_h'\right)$. As voltage v_g falls during sensing, this offset and its corresponding storage capacity are reduced. The output of this circuit is limited to a range that is typically less than 20% of the power supply voltage. Since signals are sensed as voltage quantities, this limitation applies regardless of charge packet or CCD well sizes.

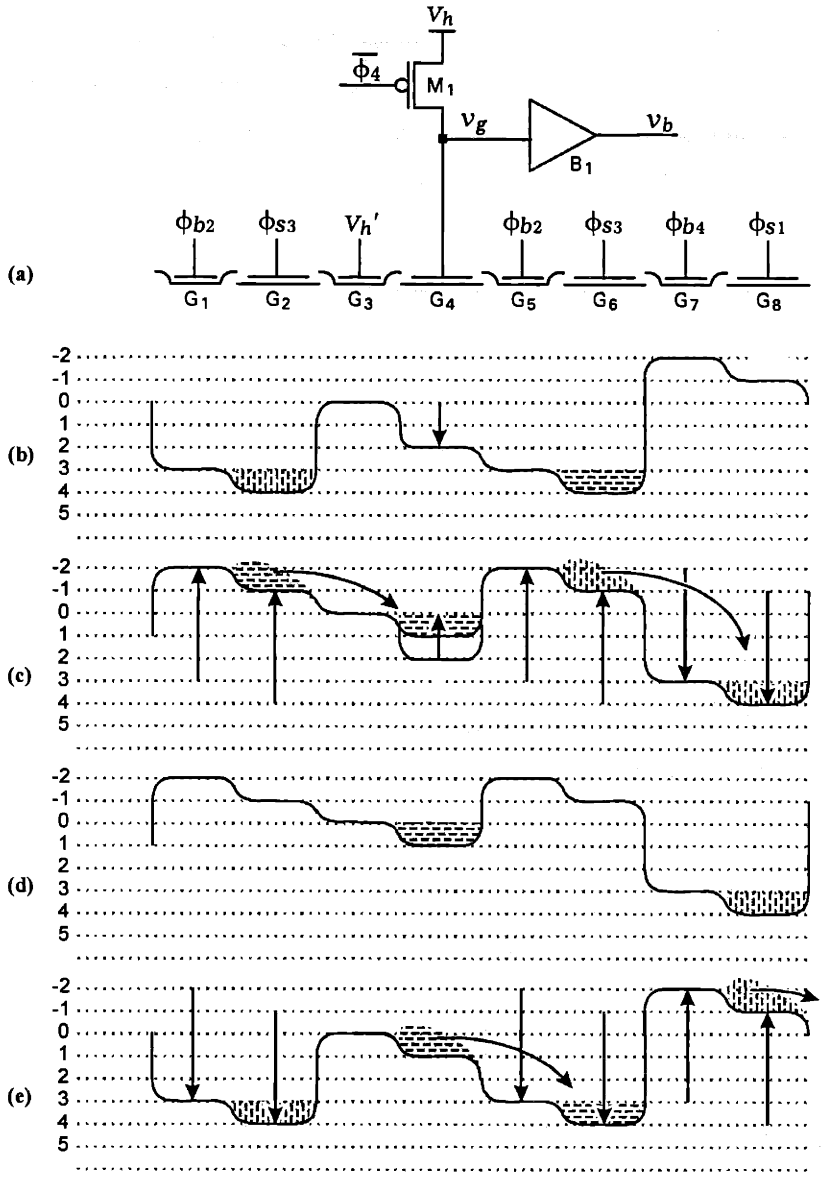


Figure 3-9. Conventional floating gate amplifier. (a) Circuit schematic. (b) Phase 1. (c) Phase 2. (d) Phase 3. (e) Phase 4.

The linearity of a conventional floating gate amplifier is derived in section 3.6.2 and is typically a few percent. Nonlinearity is introduced by voltage dependence of the various capacitances in (3-6). These include channel-to-bulk depletion capacitance and parasitic junction capacitance, both of which are voltage dependent. The alternative charge sensing techniques described later in this chapter provide improved linearity and avoid the disadvantages mentioned above.

3.4 Static Double-Sampling Charge Generator

3.4.1 Circuit Operation

The technique of double sampling provides a number of performance improvements over the conventional fill-and-spill and diode-cutoff methods described in section 3.3. In its simplest form, referred to as static double sampling (SDS), improvements occur in linearity, device-to-device matching, and sensitivity to low frequency noise. Such a configuration is described in this section. However, double sampling is most effectively applied in a dynamic double sampling (DDS) configuration that is presented in section 3.5. Additional improvements of the DDS technique include increased speed and significantly improved linearity.

The simplest example of an SDS charge generator is shown schematically in Figure 3-10(a). An initial voltage-mode sample-and-hold stage is formed from capacitors C_1 and C_2 and transistors M_3, M_4, M_2, and M_5. Remaining portions of the circuit operate in the charge domain. Capacitor C_2 is connected through a series of precharge transistors to diffusion D_1, and through a series of sensing transistors to an output well G_5. Two gates, both labeled G_{24}, are shown on either side of diffusion D_2. They are shown in the figure as two separate gates for illustrative purposes but in an actual circuit are implemented as a single device. G_{24} is connected to a fixed bias, V_h. The channel region underneath it is referred to as a cascode barrier because of its similarity to cascode devices common to CMOS amplifier design.

The following discussion of circuit operation references three figures. Energy level diagrams for the charge-domain portion are provided in Figure 3-10(b) through (e). Voltage waveforms for the sample-and-hold portion are included in Figure 3-11(a) through (d) for the example of a linear input ramp. The voltage and timing conventions, outlined in section 3.2.1, are used for all clocks in this circuit. An idealized representation of the circuit during each of the four phases is shown in Figure 3-12(a) through (d). Capacitors C_{p1} and C_{p2}, included in this model, represent parasitic junction and routing capacitances on nodes v_f and f_g, respectively. C_{p2} includes an additional component equal to the capacitance of the cascode channel underneath G_{24}. Saturation and subthreshold modes of G_{24} are represented by a single model, consisting of resistor R_{24} in series with a voltage source of value $(V_h - v_{th})$. Although it does not accurately represent the current-voltage characteristics of G_{24}, this simple model provides a useful description of circuit operation. It is quantitatively correct when the clock frequency is slow compared to device time constants involved because, as clock periods become infinitely long, details of current flow versus time do not impact the result. At higher frequencies, it still provides a qualitatively accurate representation.

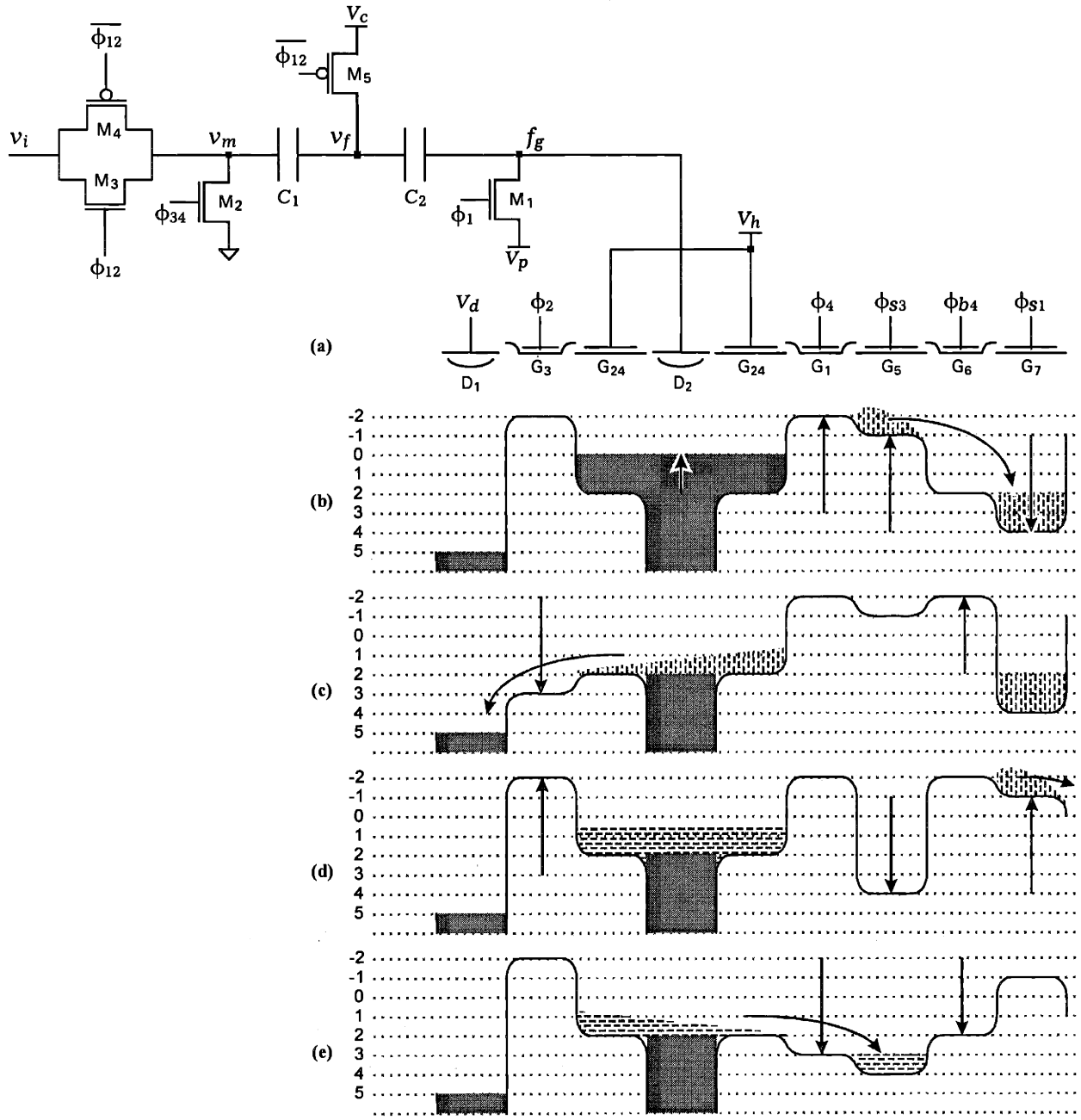


Figure 3-10. Static double-sampling charge generator and associated energy level diagrams. (a) Schematic. (b) Fill phase. (c) Spill Phase. (d) Collection phase. (e) Sensing phase.

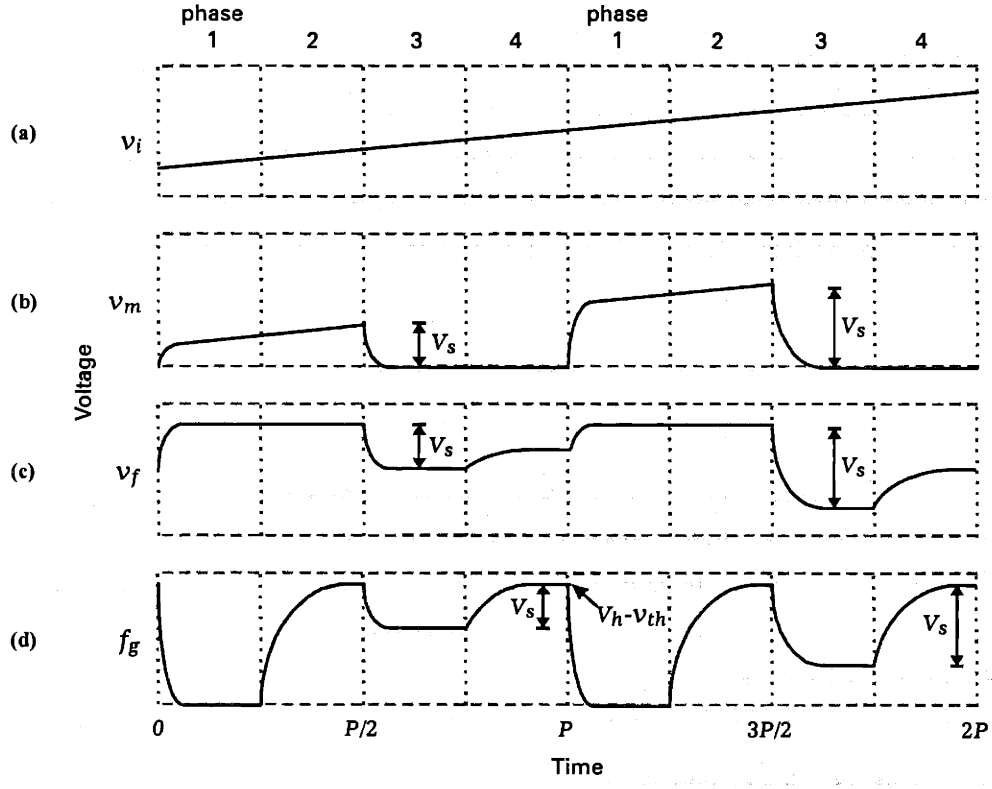


Figure 3-11. Example voltage waveforms for an SDS charge generator. (a)-(c) Signals v_i, v_m, and v_f. (d) Signal f_g is restored to its precharged level at the end of phase 4.

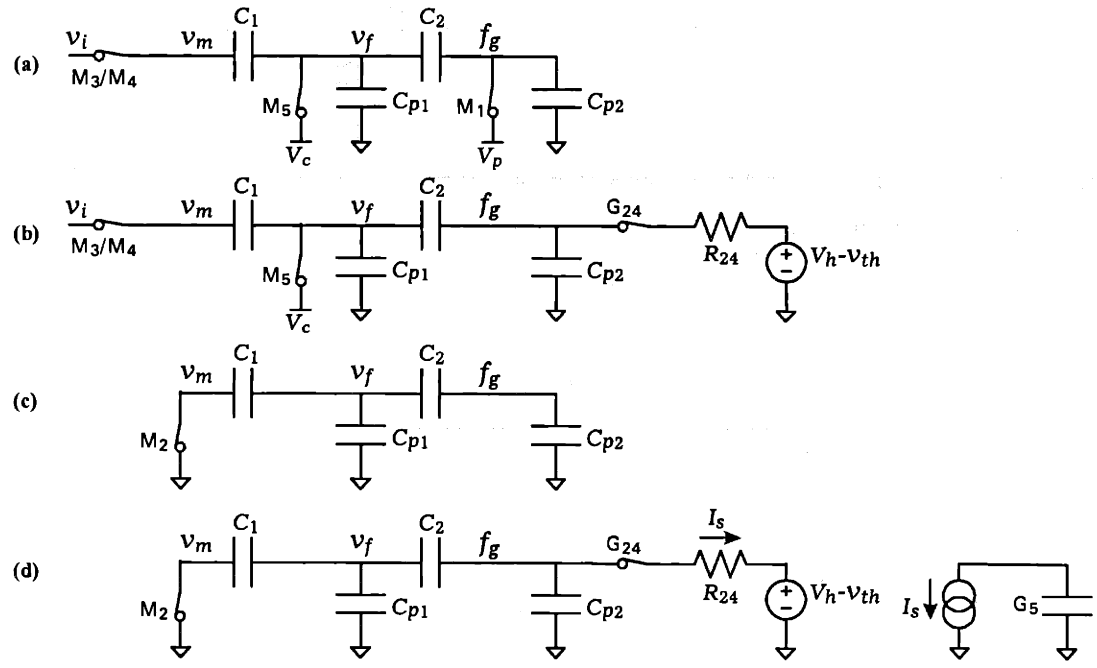


Figure 3-12. Simplified linear model of an SDS charge generator. (a) Fill phase. (b) Spill phase. (c) Collection phase. (d) Sensing phase.

Fill Phase

Phase 1 is referred to as the fill phase. Node v_m tracks the ramp on the analog input signal v_i, and v_f is clamped to V_c. At the same time, f_g is pulled low, to V_p, and the region underneath G_{24} is flooded with charge. This is necessary to replenish electrons consumed during the previous generation cycle. The charge packet underneath G_5, which was generated during the previous cycle, is transferred forward by pulling G_1 and G_5 to ground. Charge can not flow backward because of the built-in threshold offset between barrier and storage gates.

Spill Phase

Phase 2 is referred to as the spill phase. M_1 is turned off and f_g is left floating. The potential on G_3 is raised to enable the precharge path and immediately thereafter, electrons flow from f_g to the drain, V_d. Because the channel potential under G_3 is more positive than that under G_{24}, current flow is determined by $\left(V_h - f_g\right)$, the gate-to-source voltage of G_{24}. As charge flows from f_g to V_d, the voltage on f_g rises. Initially $\left(V_h - f_g\right)$ is large, G_{24} is in saturation, and charge flows rapidly. But as f_g rises, $\left(V_h - f_g\right)$ is reduced and current is reduced accordingly. Eventually, $\left(V_h - f_g\right)$ approaches v_{th}, the threshold voltage of G_{24}, and the device enters subthreshold. From this point on, the voltage on f_g continues to rise very slowly toward the channel potential underneath G_{24}.

The value of f_g during this phase is given by

$$f_g(t) = V_p\, e^{-(t-H)/\tau_2} + (V_h - v_{th})\left(1 - e^{-(t-H)/\tau_2}\right) \qquad H \le t \le 2H \ , \tag{3-7}$$

where the time constant τ_2 is

$$\tau_2 = R_{24}\left(C_{p2} + C_2\right) \ . \tag{3-8}$$

Nodes v_m and v_f are not impacted by the transition on f_g because they remain actively driven. At the end of phase 2 these signals have values of

$$v_m(2H) = v_i(2H) = V_s \tag{3-9}$$

and

$$v_f(2H) = V_c \ . \tag{3-10}$$

Variable V_s denotes the value of the analog input v_i at the end of the sampling operation.

Collection Phase

The collection phase is phase 3. The clamping device G5 is turned off and v_f is left floating. The precharge path is closed by lowering G3 and G5 is raised in anticipation of receiving charge. Meanwhile, v_m is pulled to ground. The falling transition on v_m,

$$v_m(3H) - v_m(2H) = -V_s \ , \tag{3-11}$$

causes the intermediate node v_f to fall toward a final voltage of

$$v_f(3H) = V_c - V_s K_1 \ . \tag{3-12}$$

The unitless constant K_1 is defined as

$$K_1 = \frac{C_1}{\left(C_{p1} + \frac{C_2 C_{p2}}{C_2 + C_{p2}} + C_1 \right)} \ . \tag{3-13}$$

Node f_g, which was precharged high during phase 2, falls and incoming charge fills the channel region underneath G24. The final value of f_g is given by

$$f_g(3H) = V_p\, e^{-H/\tau_2} + (V_h - v_{th})\left(1 - e^{-H/\tau_2}\right) - V_s K_1 \frac{C_2}{\left(C_2 + C_{p2}\right)} \ . \tag{3-14}$$

Sensing Phase

Phase 4 is referred to as the sensing phase. The voltage on G1 is raised to enable the sensing path, and immediately thereafter, electrons flow from f_g to the receiving well underneath G5. The resulting transition on f_g is similar to that during the spill phase. As charge flows to G5, the voltage on f_g rises toward the channel potential of G24. Eventually G24 enters subthreshold and current flow gradually ceases.

Device G24 is represented by the same model that was used for the spill phase. It consists of a resistor in series with a voltage source of value $(V_h - v_{th})$. The value of f_g as a function of time is given by

$$f_g(t) = \left(V_p\, e^{-H/\tau_2} - V_s K_1 \frac{C_2}{\left(C_2 + C_{p2}\right)} \right) e^{-(t-3H)/\tau_4} + (V_h - v_{th})\left(1 - e^{-H/\tau_2} e^{-(t-3H)/\tau_4}\right) \ , \tag{3-15}$$

where the time constant τ_4 equals

$$\tau_4 = R_{24}\left(C_{p2} + \frac{C_2\left(C_1 + C_{p1}\right)}{C_1 + C_2 + C_{p1}} \right) \ . \tag{3-16}$$

By the end of the sensing transition, f_g has been restored to its original precharged level.

Since v_f is floating during this phase, it rises according to

$$v_f(t) = V_c - V_s K_1 + \left(f_g(t) - f_g(3H)\right) \frac{C_2}{C_1 + C_{p1} + C_2} \qquad 3H \le t \le 4H \ . \tag{3-17}$$

In contrast to f_g, the value of v_f at the end of the sensing phase differs from its precharged value.

The charge packet that is collected underneath G5 is given by the integral of the current I_s over the entire sensing operation. Its value,

$$Q_o = \left(f_g(4H) - f_g(3H)\right)\left(C_{p2} + \frac{(C_1 + C_{p1})C_2}{C_1 + C_2 + C_{p1}}\right), \tag{3-18}$$

is proportional to the change in f_g's voltage between the end of phases 3 and 4. Combining (3-18), (3-15), and (3-14) yields the result

$$Q_o = V_s\left(\frac{C_1C_2}{C_1 + C_2 + C_{p1}}\right)\left(1 - e^{-H/\tau_2}\right) + \left(V_h - v_{th} - V_p\right)\left(C_{p2} + \frac{(C_1 + C_{p1})C_2}{C_1 + C_{p1} + C_2}\right)e^{-H/\tau_4}\left(1 - e^{-H/\tau_2}\right). \tag{3-19}$$

This result is easily understood in an idealized case when $H >> \tau_2$, $H >> \tau_4$, and C_{p1}=0. The assumption that $H >> \tau_2$ and $H >> \tau_4$ states that the operating frequency is slow compared to the speed of component devices. A sufficient number of time constants are allowed for precharge and sensing operations that their transitions are approximately complete. The value of f_g at the end of the spill phase is then identical to that at the end of the sensing phase and f_g behaves as a virtual ground. Since no charge is consumed to change the potential on f_g, all displacement charge is collected in the receiving well, and parasitic capacitance, C_{p2}, has no impact on the result. The assumption that C_{p1}=0 states that gain is not reduced, despite the fact that the potential on v_f at the end of sensing differs from that at the end of precharge.

When these assumptions apply, the resulting charge packet,

$$Q_o = V_s\left(\frac{C_1C_2}{C_1 + C_2}\right), \tag{3-20}$$

depends only on the sampled input signal and on C_1 and C_2. The threshold of G24 determines the precharge value in (3-7) and the sensing value in (3-15). But because of double sampling, both precharge and sensing operations are performed with respect to this same threshold, the circuit is autozeroed, and v_{th} does not impact the result in (3-20). Such a circuit is insensitive to absolute threshold values and remains well matched across multiple devices even when large variations exist in their thresholds.

None of the bias voltages V_d, V_c, V_h, or V_p impact the result in (3-20). Bias V_p should be low enough that charge can fill the region underneath G24 during the fill operation. Any level, such as ground, that is less than V_h-v_{th} is acceptable. Bias V_d should be high enough that it can remove electrons from f_g during the spill operation. It is typically held at the most positive supply. The level of V_c must be high enough that $v_f(3H)$, given by (3-12), remains greater than zero for all possible values of V_s. Typically, any level greater than half of the supply voltage can be used. Similarly, the level of V_h must be high enough that $f_g(3H)$, given by (3-14), remains greater than zero at all times.

The following sections examine the result in (3-19) under nonideal conditions. The linearity of this circuit at high speeds is described in section 3.4.2, without the assumptions $H >> \tau_2$ and $H >> \tau_4$. Section 3.4.3

examines linearity when parasitic capacitance, C_{p1}, is greater than zero. Finally, section 3.4.4 explores the impact of various noise sources on circuit performance.

3.4.2 Transient Linearity

At high speed, the second term in (3-19) represents a small offset, proportional to the difference between fill and spill values of f_g. It is independent of the signal being processed and does not introduce distortion. The signal term in (3-19) is multiplied by a factor of $\left(1-e^{-H/\tau_2}\right)$. According to this expression, as frequency is increased, circuit gain is reduced because the spill and sensing values of f_g differ by an amount that is signal dependent. But the resulting response remains linear. This linear signal term is a consequence of the linear models of Figure 3-12, used for the analysis above. In an actual circuit, distortion is introduced because the resistance of G24 is nonlinear and varies with the voltage f_g.

The distortion characteristics of the circuit in Figure 3-10 are examined below using more accurate models for G24. At the beginning of the sensing process this device is in saturation. It has a current-voltage relationship given by

$$I_s(t)=\frac{KW}{2L}\left(V_h-f_g(t)-v_{th}\right)^2 \quad , \tag{3-21}$$

where I_s is the current flowing onto f_g, K is its transconductance parameter measured in A/V^2, and W/L represents its width-to-length ratio.

The effect of drain voltage on I_s is ignored in (3-21). This is appropriate during the spill phase because the drain of the G24/G3 combination is at a fixed bias, V_d. It is also a good approximation during the sensing phase because G24 and G1 form a cascode combination whose output impedance is the product of that for each individual device. At the end of the sensing operation, when currents are small, this combined impedance is large. Cascoding can also be understood in terms of the bucket diagrams of Figure 3-10. The channel potential underneath G1, at the G24 edge, is the effective drain of G24. Although charges collect in the receiving well and the surface potential underneath G5 changes, no charges are stored underneath G1 and its surface potential does not change appreciably.

Near the end of the sensing process, when f_g has risen to within $2kT/q$, or 52 mV, of its final value, G24 is in weak inversion. The constant k is Boltzmann's constant, T is the temperature in degrees Kelvin, and q is the charge of an electron. Its current-voltage relationship during this time is given by

$$I_s(t)=\frac{2K(kT/q)^2}{e^2}\frac{W}{L}e^{\left(V_h-f_g(t)-v_{th}\right)/(nkT/q)} \quad . \tag{3-22}$$

During both saturation and subthreshold periods, I_s determines the derivative of f_g. This relationship is described by

$$I_s(t)=\left(\frac{C_2\left(C_1+C_{p1}\right)}{C_1+C_{p1}+C_2}+C_{p2}\right)\frac{df_g(t)}{dt} \quad . \tag{3-23}$$

Combining (3-21) and (3-23) yields the nonlinear differential equation

$$\frac{df_g(t)}{dt} = \frac{KW}{2L}\left(\frac{C_2(C_1+C_{p1})}{C_1+C_{p1}+C_2}+C_{p2}\right)^{-1}\left(V_h - f_g(t) - v_{th}\right)^2 \tag{3-24}$$

that governs f_g during saturation. The relationship for f_g during the time that G24 is in subthreshold,

$$\frac{df_g(t)}{dt} = \frac{2K(kT/q)^2}{e^2}\frac{W}{L}\left(\frac{C_2(C_1+C_{p1})}{C_1+C_{p1}+C_2}+C_{p2}\right)^{-1} e^{\left(V_h - f_g(t) - v_{th}\right)/(nkT/q)}, \tag{3-25}$$

is a combination of (3-22) and (3-23). Equations (3-24) and (3-25) are solved numerically to arrive at the results presented below. The example set of parameters

$$\begin{aligned}
&\left(\frac{C_2(C_1+C_{p1})}{C_1+C_{p1}+C_2}+C_{p2}\right)^{-1} = 0.5 \text{ pF}\\
&K = 100\ \mu\text{A}/\text{V}^2\\
&\frac{W}{L} = 40\\
&V_h - v_{th} = 1 \text{ V}\\
&n = 1.4
\end{aligned} \tag{3-26}$$

is used for all simulation results.

The sensing operation begins at time t=3H. The value of f_g at this time, is referred to as f_{g0}. It is a function of the input sample V_s as given by (3-14). A larger V_s causes more charge to be transferred and produces a smaller value of f_{g0}. Figure 3-13 shows a plot of the transition on f_g during the first two nanoseconds of phase 4. Curves are shown for f_{g0}=0, 0.25, 0.5, and 0.75. All curves asymptote toward their final value of $V_h - v_{th} = 1$ as $t \to \infty$.

Figure 3-14 plots the relationship between f_g and f_{g0} as a function of the time allowed for sensing. Curves are shown for values of phase time H ranging from 2.5 to 15 nanoseconds. The signal dependence of

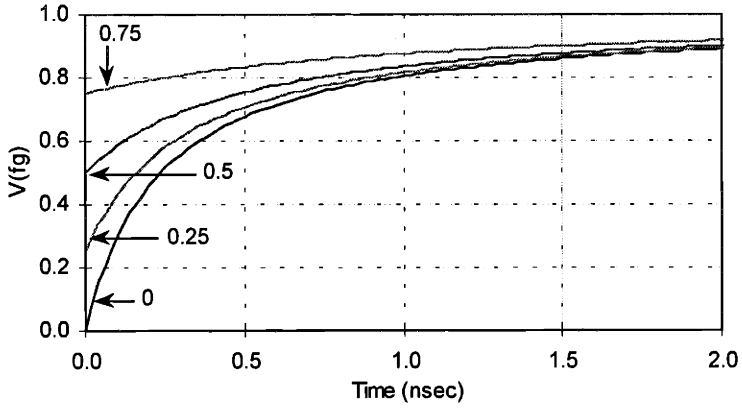


Figure 3-13. Transition on f_g during the sensing phase for initial values of 0, 0.25, 0.5, and 0.75V. All curves rise rapidly and then asymptote toward 1V.

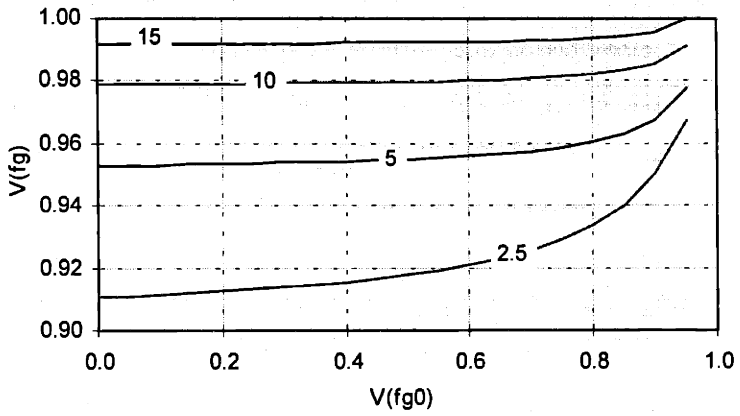


Figure 3-14. Final voltage on f_g at the end of phase 4, as a function of its starting value, for phase times of 2.5, 5, 10, and 15 nsec. Distortion increases rapidly as f_{g0} approaches 1V.

$f_g(H)$ decreases as H increases. Distortion is caused by any deviation of these curves from linear and is particularly pronounced as f_{g0} approaches 1.0, where the charge packets being generated are small. Distortion can be reduced by operating the circuit over a smaller range of f_{g0}, such as that from 0 to 0.8. In this approach, a 25% background level of charge is added into each resulting packet.

A moderate background level of charge is usually not a problem for CCD circuits. First, background charge is beneficial in that it keeps traps in the CCD channel filled and improves the efficiency of charge transfers. Second, CCD circuits are almost always implemented differentially and, in this case, any additional offset increases the CCD well areas that are needed to hold charge packets and increases the common mode signal but does not otherwise change circuit operation.

The magnitude of distortion in the generated result is found by evaluating the nonlinear portion of the curves in Figure 3-14. Since low-order harmonics are of greatest concern, the analysis here considers the second and third. The curves are fit to a third-order polynomial,

$$fg(4H) = d\,{f_{g0}}^3 + c\,{f_{g0}}^2 + b\,f_{g0} + a \ , \tag{3-27}$$

using a least-squares fit. Equations (3-18) and (3-27) are then combined to the produce an expression,

$$Q_o = \left(d\,{f_{g0}}^3 + c\,{f_{g0}}^2 + (b-1)\,f_{g0} + a\right)\left(C_{p2} + \frac{(C_1 + C_{p1})C_2}{C_1 + C_2 + C_{p1}}\right) , \tag{3-28}$$

for Q_o in terms of the polynomial coefficients. The first-order term (b-1) represents a linear reduction in circuit gain. At an input frequency of f_0, the ratio between the second harmonic and the fundamental is

$$\frac{Q_o(2f_0)}{Q_o(f_0)} = \left(\frac{-2c}{4(b-1)+3d}\right) f_{g0} \ . \tag{3-29}$$

The ratio between the third harmonic and the fundamental is

$$\frac{Q_o(3f_0)}{Q_o(f_0)} = \left(\frac{-d}{4(b-1)+3d}\right) {f_{g0}}^2 \ . \tag{3-30}$$

Numerical techniques are used to solve for the coefficients in (3-28) and the results of (3-29) and (3-30) are plotted as a function of clock period, $4H$, in Figure 3-15. The largest harmonic is the second. Most CCD devices include differential signal flow and use pairs of differential charge generators. In such devices, even-order harmonics are cancelled and linearity is limited by the third harmonic.

Distortion can be improved by increasing G_{24}'s transconductance K, its size ratio W/L, or the clock period $4H$. But the long time constants associated with subthreshold operation in this circuit seriously limit its speed and linearity. Distortion increases rapidly as frequency is increased. Improving the linearity and speed performance of this circuit is a primary objective of the dynamic double-sampling technique that is presented in section 3.5. Improvements are accomplished by eliminating the subthreshold portion of the spill and sensing transitions.

3.4.3 Capacitance-Dependent Linearity

A second source of distortion in the charge generator of Figure 3-10 is voltage dependence in its capacitors. Parasitics on the sensing node f_g only impact the offset term in (3-19) and do not add distortion. Other capacitors, present in the signal term of (3-19), can cause nonlinearity. These include C_1, C_2, and C_{p1}. The impact of their voltage dependence is described below.

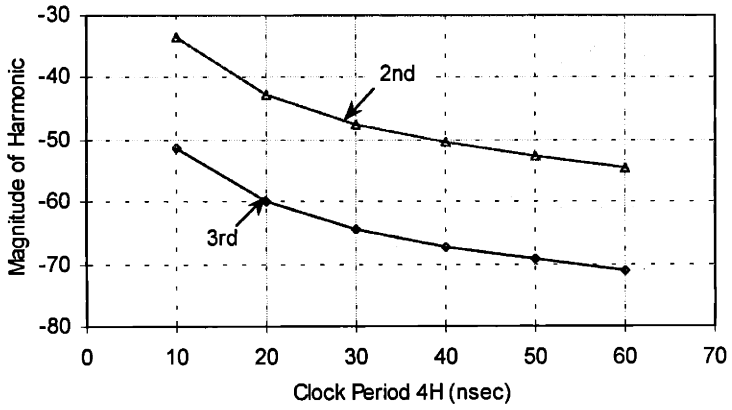


Figure 3-15. Second and third harmonic transient distortion as a function of clock period. Distortion increases exponentially as clock periods are reduced.

For the analysis, each capacitor is defined to have a nominal value, equal to its capacitance at the end of sensing for an input sample V_s of zero. Using this definition, the nominal voltage across both C_1 and C_{p1} is V_c, while that across C_2 is $(V_c - V_h + v_{th})$. When the input sample V_s is greater than zero, the change in voltage across these capacitors is given by the difference between (3-15) and (3-17) and their nominal values. All three capacitors experience an identical change in voltage equal to

$$\left(V_c - v_f(4H)\right) = V_s K_3 \ , \tag{3-31}$$

where the unitless constant K_3 is defined as

$$K_3 = \frac{C_1}{\left(C_{p1} + \dfrac{C_2 C_{p2}}{C_2 + C_{p2}} + C_1\right)}\left(1 - \frac{{C_2}^2}{\left(C_1 + C_{p1} + C_2\right)\left(C_2 + C_{p2}\right)}\right) . \tag{3-32}$$

Input Capacitors

The first case that is considered is one where C_{p1} is constant but C_1 and C_2 have a nonzero voltage dependence. The output packet depends on the values of C_1' and C_2' at the end of phase 4 but does not depend on them at any other times. These values are a function of the input signal, V_s and are represented by the polynomials

$$\begin{aligned} C_1'(V_s) &= b_1 {K_3}^2 {V_s}^2 + a_1 K_3 V_s + C_1 \\ C_2'(V_s) &= b_2 {K_3}^2 {V_s}^2 + a_2 K_3 V_s + C_2 \end{aligned} . \tag{3-33}$$

The constants C_1 and C_2 represent the nominal values of these elements. Coefficients a_1 and a_2 are measured in F/V and b_1 and b_2 are measured in F/V^2.

At low speed, when $H >> \tau_2$ and $H >> \tau_4$, equation (3-19) becomes

$$Q_o = V_s\left(\frac{C_1'(V_s) C_2'(V_s)}{C_1'(V_s) + C_2'(V_s) + C_{p1}}\right) . \tag{3-34}$$

The expressions in (3-33) are substituted for C_1 and C_2 into (3-34). Assuming that the voltage-dependent portion of these capacitors is small compared to their nominal values, Q_o is approximated by

$$Q_o \approx V_s \frac{C_1 C_2}{\left(C_1 + C_2 + C_{p1}\right)}\left(\begin{aligned} &1 + V_s K\left(\frac{a_2\left(C_1 + C_{p1}\right)}{C_2\left(C_1 + C_2 + C_{p1}\right)} + \frac{a_1\left(C_2 + C_{p1}\right)}{C_1\left(C_1 + C_2 + C_{p1}\right)}\right) \\ &+ {V_s}^2 K^2\left(\frac{b_2\left(C_1 + C_{p1}\right)}{C_2\left(C_1 + C_2 + C_{p1}\right)} + \frac{b_1\left(C_2 + C_{p1}\right)}{C_1\left(C_1 + C_2 + C_{p1}\right)}\right) \end{aligned}\right) . \tag{3-35}$$

As an example, consider a case where C_1 and C_2 have the same nominal value, denoted by C_{12}, the same voltage coefficients, denoted by a_{12} and b_{12}, and $C_{p1} << C_{12}$. At an input frequency of f_0, the ratio between the second harmonic and the fundamental is

$$\frac{Q_o(2f_0)}{Q_o(f_0)} \approx \frac{-(a_{12}/C_{12})}{4} V_s \quad . \tag{3-36}$$

The ratio between the third harmonic and the fundamental is

$$\frac{Q_o(3f_0)}{Q_o(f_0)} \approx \frac{-(b_{12}/C_{12})}{8} V_s^2 \quad . \tag{3-37}$$

The distortion given by (3-36) and (3-37) depends only on the ratios a_{12}/C_{12} and b_{12}/C_{12}, not on absolute capacitance values. Results are plotted in Figure 3-16 as a function of these ratios for V_s=1V. Like conventional charge generation techniques, such as fill-and-spill or diode cutoff, the charge packet produced by this circuit depends on input capacitor voltage dependence. However, the SDS technique presented here has an advantage in this regard. Its input capacitance is formed from elements C_1 and C_2 and these can be implemented using materials, such as polysilicon or metal, that have a very low voltage dependence.

Parasitic Capacitors

As a second example, consider the case where C_1 and C_2 are ideal, but C_{p1} has a non-zero voltage dependence. A primary component of C_{p1} is the drain-bulk junction of M_5, which is strongly voltage dependent. At low speed, where $H >> \tau_2$ and $H >> \tau_4$, equation (3-19) becomes

$$Q_o = V_s \left(\frac{C_1 C_2}{C_1 + C_2 + C_{p1}'(V_s)} \right) \quad . \tag{3-38}$$

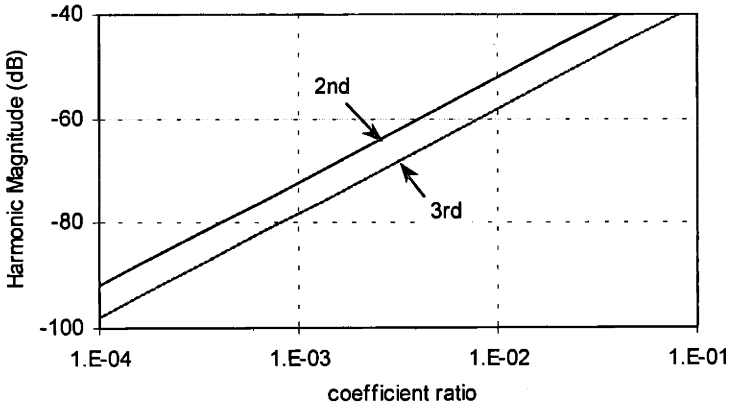


Figure 3-16. Second and third harmonic distortion caused by input capacitor voltage dependence. An advantage of SDS is that low voltage dependence is possible from these capacitors.

The generated result depends on the value of C_{p1}' at the end of phase 4. This value is modeled using the polynomial

$$C_{p1}'(V_s) = b_3 K_3^2 V_s^2 + a_3 K_3 V_s + C_{p1} \ , \tag{3-39}$$

where C_{p1} represents its nominal value and the coefficients a_3 and b_3 are measured in F/V and F/V^2, respectively.

Equation (3-39) is substituted for C_{p1} into (3-38). In most cases, the coefficients a_3 and b_3 are a significant percentage of the nominal value of C_{p1} but a small percentage of the values of C_1 and C_2, and the approximations $a_3 << C_1, C_2$ and $b_3 << C_1, C_2$ are valid. Under this assumption, the resulting expression is simplified to

$$Q_o \approx V_s \frac{C_1 C_2}{(C_1 + C_2 + C_{p1})} \left(1 - V_s K_3 \frac{a_3}{C_1 + C_2 + C_{p1}} - V_s^2 K_3^2 \frac{b_3}{C_1 + C_2 + C_{p1}} \right) . \tag{3-40}$$

As an example, consider a case where both C_1 and C_2 have the same nominal value C_{12} and where $C_{p1} << C_1, C_2$. At an input frequency of f_0, the ratio between the second harmonic and the fundamental is

$$\frac{Q_o(2f_0)}{Q_o(f_0)} \approx \frac{(a_3/C_{12})}{8} V_s \tag{3-41}$$

and the ratio between the third harmonic and the fundamental is

$$\frac{Q_o(3f_0)}{Q_o(f_0)} \approx \frac{(b_3/C_{12})}{32} V_s^2 \ . \tag{3-42}$$

Equations (3-41) and (3-42) depend only on the ratios a_3/C_{12} and b_3/C_{12}, not on the absolute value of C_{p1} or its percentage voltage dependence. The resulting distortion, plotted in Figure 3-17 as a function of these ratios for V_s=1V, is dominated by the second harmonic. When circuits are implemented differentially, even-order harmonics are cancelled and the third harmonic is dominant.

In an n-well process, the voltage dependence of C_{p1} can be reduced by setting V_c to a level less than that of the bulk of transistor M_5, thereby increasing the reverse bias across the drain-bulk junction. However, a

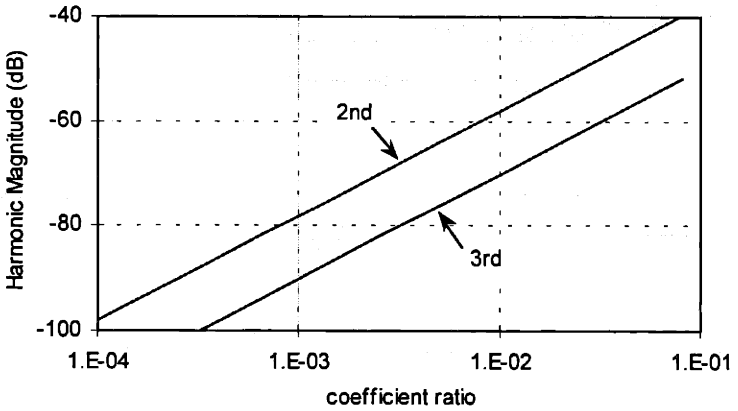


Figure 3-17. Second and third harmonic distortion caused by parasitic capacitor voltage dependence.

more effective technique for reducing this source of distortion is presented in section 3.5 as part of the DDS charge generator.

3.4.4 Noise

A number of deterministic and random noise sources impact the signal-to-noise performance of an SDS charge generator. Thermal noise is introduced by the random fluctuation of carriers due to thermal energy. Flicker noise causes device thresholds to vary over time. Supply noise is present on circuit biases. Clock jitter, or timing variations, occur in controlling clocks. Clock feedthrough is caused by parasitic coupling from transistor gates. And charge injection occurs when channel charge divides unpredictably. First, the general noise transfer characteristics are derived for this circuit. Then the impact of each of these noise sources is described.

General noise transfer characteristics for this circuit are derived using the models in Figure 3-12 and the assumptions that $H >> \tau_2$, τ_4 and C_{p1}, $C_{p2} << C_1$, C_2. During the first half cycle, v_m is connected to the analog input through a switch. Immediately after this switch is opened, the value that is captured on v_m is

$$v_m(2H) = V_s + n_m(2H) \ . \tag{3-43}$$

The noise quantity n_m represents a combination of thermal noise, clock feedthrough, and charge injection from the switch. During the second half cycle, v_m is clamped to ground through a second switch. Its value immediately before the end of the sensing operation,

$$v_m(4H) = n_m(4H) \ , \tag{3-44}$$

includes ground noise and thermal noise but no clock feedthrough or charge injection. The input charge produced by this clamp-and-sample operation is proportional to

$$\Delta v_m = V_s + n_m(2H) - n_m(4H) \ . \tag{3-45}$$

Node v_f is clamped during the first half cycle and is floating during the remainder of the cycle. Shortly after the clamp is opened, it has a value of

$$v_f(2H) = V_c + n_f(2H) \ , \tag{3-46}$$

where n_f represents a combination of supply noise on the bias V_c as well as thermal noise, clock feedthrough, and charge injection from the clamping switch.

At the end of phase 1, when the clamp on f_g is removed, thermal noise, clock feedthrough, and charge injection are introduced. The clamping bias also adds supply noise. If the sum of these noises is represented by n_g, then the value of f_g at the end of phase 1 is given by

$$f_g(H) = V_p + n_g(H) \ . \tag{3-47}$$

At the end of the spill and sensing phases, f_g has a precharge value of

$$f_g(2H) = (V_h - v_{th}) + n_g(2H) \tag{3-48}$$

and the sensing value of

$$f_g(4H) = (V_h - v_{th}) + n_g(4H) \ . \tag{3-49}$$

Noise n_g represents supply noise on V_h, flicker noise in the cascode gate, and thermal noise from G_{24}. No charge injection or clock feedthrough occurs because no carriers reside underneath G_3 or G_1 when they are turned off, and no diffusions are present on the source or drain of these devices.

The output charge is proportional to the difference between the floating gate voltage immediately after the end of collection and that in (3-49), immediately before the end of sensing. It equals

$$Q_o = V_s \frac{C_1 C_2}{C_1 + C_2} + \left(n_g(4H) - n_g(2H) + n_m(2H) - n_m(4H)\right)\frac{C_1 C_2}{C_1 + C_2} \ . \tag{3-50}$$

Because of double sampling, the output packet depends only on the difference between noise sources at the end of phases 2 and 4. The frequency response of this circuit is

$$\frac{|Q_o(\omega)|}{|N_g(\omega)|} = \frac{|Q_o(\omega)|}{|N_m(\omega)|} = \left(\frac{C_1 C_2}{C_1 + C_2}\right)\sin(\omega H) \ . \tag{3-51}$$

Noise at low frequencies is common to both precharge and sensing phases and is attenuated, while noise near the Nyquist frequency is passed directly to the output. Because (3-50) does not depend on $n_m(2H)$ or $n_g(H)$, the circuit is tolerant of noise on the biases V_c and V_p and thermal noise in the switches M_1 and M_5. The CCD gates after G_1 do not add any additional noise because they operate in a fully-depleted manner.

Thermal Noise

Circuits that generate charge packets are fundamentally not depleted because they require a source of electrons. Like other nondepleted operations, charge generation is subject to thermal noise. The result in (3-50) depends on thermal noise components at the end of phases 2 and 4. The mean-square value of n_m at time $2H$ is

$$\overline{n_m(2H)^2} = kT\frac{1}{C_1} \tag{3-52}$$

and that at time $4H$ is

$$\overline{n_m(4H)^2} = kT\left(\frac{1}{C_1} + \frac{1}{C_2}\right) \ . \tag{3-53}$$

Variable k refers to Boltzmann's constant and T is the temperature in degrees Kelvin. The thermal noise component of n_g at these times equals

$$\begin{aligned}\overline{n_g(2H)^2} &= kT\frac{1}{C_2} \\ \overline{n_g(4H)^2} &= kT\left(\frac{1}{C_1}+\frac{1}{C_2}\right)\end{aligned} \quad . \tag{3-54}$$

Under the assumption that these four noise components are independent, the combined rms noise at the output of this circuit is

$$\sqrt{\overline{Q_o^2}} = \sqrt{3kT\frac{C_1C_2}{(C_1+C_2)}} \quad . \tag{3-55}$$

Charge-domain circuits have the general property that thermal noise decreases as capacitance is reduced and in most cases this translates into an increased SNR. However, a charge generator differs in this regard. Like other charge-domain circuits, its thermal noise, in (3-55), decreases as capacitance is reduced, but its signal level also decreases and does so more rapidly. Equations (3-54) and (3-20) are combined to arrive at the expression

$$\text{SNR} = V_s\sqrt{\frac{C_1C_2}{3kT(C_1+C_2)}} \tag{3-56}$$

for an input sinusoid with an rms amplitude of V_s. In contrast to other charge domain circuits, a charge generator has an SNR that is proportional to the square root of capacitance. The result in (3-56) is plotted as a function of capacitance $C_1C_2/(C_1+C_2)$ for V_s=0.707V in Figure 3-18.

Supply Noise and Flicker Noise

Conventional charge generation techniques, such as fill-and-spill or diode cutoff, pass supply and flicker noise directly to their outputs. In contrast, these noise sources are attenuated by a double-sampling charge

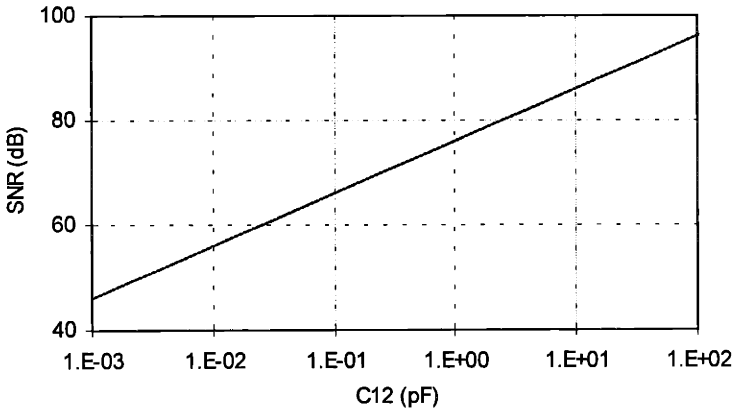


Figure 3-18. Signal-to-thermal-noise ratio of an SDS charge generator for sinusoidal inputs as a function of capacitance C_{12}=C_1=C_2. Increasing capacitance increases noise but improves SNR.

generator. Supply noise on the bias V_h and flicker noise in device G_{24} both influence the cascode channel potential and change the level that precharge and sensing transitions settle to. These noise terms are correlated between the spill and sensing phases because they occur in the same gate and with respect to the same supply. This is an advantage in a double-sampling circuit because noise quantities in (3-50) are differentiated and low frequency noise is attenuated.

The SDS noise response described by (3-51), is shown in Figure 3-19. The spectral density of flicker noise, which decreases inversely with frequency, is also shown in the figure. Most of its energy lies at low frequency, where the circuit response is attenuating. At typical operating frequencies for this circuit, flicker noise is insignificant. For example, above 1 MHz attenuation is greater than 97 dB. Power supply noise is also multiplied by the transfer function in (3-51) and attenuated at low frequencies.

Clock Jitter and Charge Injection

Timing variations in the sampling clocks, ϕ_{34} and $\overline{\phi_{34}}$, of the voltage-mode sample-and-hold block are a source of jitter. They introduce noise into the sampled input value that increases as the input frequency increases. This sample-and-hold block, which is built using conventional voltage-domain CMOS circuit techniques, is the dominant source of jitter for an SDS charge generator. These errors have been widely studied [65] and their analysis is not repeated here.

Additional clock jitter and charge injection errors, introduced by the charge-domain portion of this circuit, are described below. In the charge domain, noise due to clock jitter and charge injection increases with the current I_s that is flowing at the time that clocks are shut off. A shift in the falling edge of the ϕ_2 or ϕ_4 clocks by a time Δt causes a variation in output charge of approximately,

$$\Delta Q_o = I_s \Delta t \quad . \tag{3-57}$$

Charge injection also depends on I_s because the amount of charge in the channel at the time of turn-off is

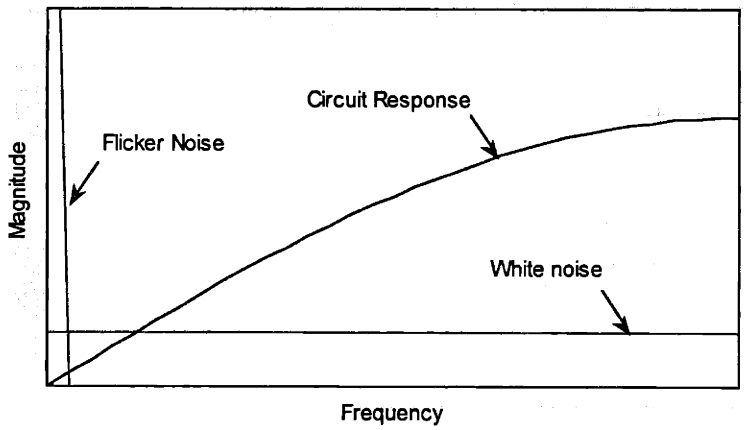


Figure 3-19. Flicker noise and SDS spectral response. Flicker and other low frequency noises are attenuated.

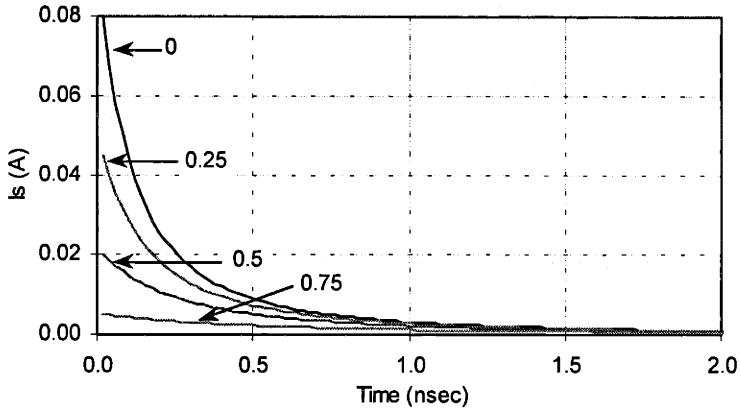


Figure 3-20. SDS sensing current versus time into phase 4. Current is reduced, but nonzero, at the time that clocks are turned off.

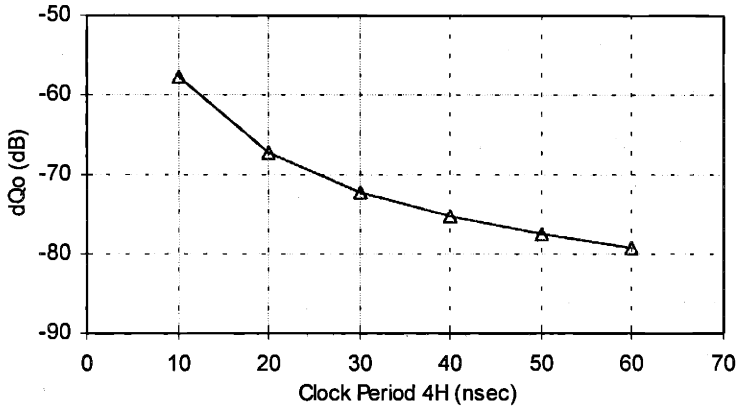


Figure 3-21. Variation in output charge as a function of clock period for a fixed 20-psec displacement of the sensing clock edge.

determined by the current flow.

Fully depleted circuits are not impacted by clock jitter or charge injection because their current flow is terminated when the supply of carriers is exhausted, not by a clock edge. To a good approximation, current is zero and no carriers are present in the channel at the time that clocks are deasserted.

This advantage does not apply to the nondepleted portion of an SDS charge generator. Simulated current through G_{24} during the spill and sensing phases is plotted as a function of time in Figure 3-20 for initial values of f_{g0} of 0, 0.25, 0.5, and 0.75V with the device parameters listed in (3-26). Current decreases rapidly within the first few nanoseconds. By the time that clocks are turned off, current and charge present in the cascode channel are both reduced. Figure 3-21 plots simulated values of $20\log(\Delta Q_o)$ as a function of clock period $4H$ for a fixed jitter of 20 picoseconds. Both jitter noise and distortion depend in a similar way on the current, $I_s(4H)$. This is evident from a comparison of the curves in Figure 3-21 with those in Figure 3-15.

3.5 Dynamic Double-Sampling Charge Generator

3.5.1 Circuit Operation

The SDS charge generator, described in section 3.4, provides a number of improvements over conventional charge generation techniques. First, it achieves better linearity because its result does not depend on the voltage-dependent capacitance of an MOS well. Instead, linearity is proportional to capacitors that can be implemented from materials with low voltage dependence. Second, it is less sensitive to flicker noise and low frequency power supply noise because double sampling differentiates noise quantities. Finally, it does not depend critically on absolute device thresholds because both precharge and sensing operations are performed with respect to the same threshold.

Despite these advantages, an SDS charge generator is not well suited for high performance signal processing applications. It does not have sufficient speed or linearity as a result of the long time constants associated with subthreshold transitions at the end of its spill and sensing phases. For example, if a linearity of 60 dB is desired for a device using the process parameters in (3-26), then device speed is limited to less than 10 MHz. Subthreshold time constants are also the cause of charge injection and jitter noise and they limit the useful distortion-free signal range to about 80% of the full-scale. Finally, the maximum achievable linearity is limited by voltage dependence in the parasitic capacitance C_{p1}.

These limitations are addressed by the dynamic double-sampling (DDS) technique. Speed and linearity are improved by turning off spill and sensing transitions before they enter subthreshold and by sensing charge over a time-varying, rather than a static, cascode barrier. At the same time, the useful operating range is increased to nearly the full scale and the impact of spill and sensing clock jitter is nearly eliminated. Parasitic capacitor voltage dependence is decreased as well. Because of these improvements, a DDS charge generator is capable of meeting the requirements of high performance signal processing applications.

One example of a DDS charge generator is shown schematically in Figure 3-22(a). The primary voltage-mode portion of this circuit is the sample-and-hold stage formed from capacitors C_1 and C_2 and transistors M_3, M_4, M_2, and M_5. In addition there is a feedback amplifier A_1 that computes the difference between f_g and the fixed reference voltage, V_r. A signal-dependent voltage source, which drives the bulk of M_5, is also included. Other circuit elements operate in the charge domain. A series of precharge transistors connect capacitor C_2 to the diffusion D_1 and a series of sensing transistors connect it to an output well, G_5. Two gates, both labeled G_{24}, are shown on either side of the input diffusion. As described earlier, these elements are shown as separate gates in the figure but are implemented as a single gate. A secondary static sensing path, formed from G_8, G_9, and G_{10}, is located between the input diffusion and the output well. Its cascode gate G_8 is tied to the fixed bias V_h'.

The following description of circuit operation references three figures. First, energy level diagrams for the charge-domain portion are provided in Figure 3-22(b) through (e). Second, voltage waveforms for the sample-and-hold portion are included in Figure 3-23(a) through (d) for the example of a linear input ramp. The voltage and timing conventions, outlined in section 3.2.1, are used for all clocks in this circuit. Third, an idealized representation of the circuit during each of the four phases is shown in Figure 3-24(a) through (d). Capacitor C_{p1}, included in this model, represents parasitic junction and routing capacitances on node v_f. Capacitor C_{p2} models capacitance of the cascode channel, underneath G24 and G8, as well as parasitic junction and routing capacitances on f_g. Saturation and subthreshold modes of G24 are modeled by resistor R24 in series with a voltage source. The value of this voltage source is $\left(A\left(V_r - f_g\right) - v_{th}\right)$, where A represents the amplifier gain. In a static charge generator, this type of model provides a qualitatively correct representation when the clock frequency is slow compared to device time constants. In the dynamic circuit described here, details of current flow versus time have an impact on the result, regardless of operating frequency. However, this simple model still provides an illustrative and qualitatively accurate description.

Fill Phase

Phase 1 is referred to as the fill phase. Signal v_m tracks the ramp on the analog input v_i and the intermediate node v_f is clamped to the bias V_c. When the sensing node f_g is pulled low, to V_p, the region underneath G24 is flooded with charge, and electrons consumed during the previous generation cycle are replenished. The charge packet underneath G5, which was generated during the previous cycle, is transferred forward upon the transition of G1 and G5 to ground. Charge can not flow backward during this transfer because of the built-in threshold offset between barrier and storage gates.

Spill Phase

Phase 2 is the spill phase. M1 is turned off and f_g is left floating. Initially the voltage on f_g is much lower than that on V_r and the amplifier output f_b is saturated at its high level, V_h. After the potential on G3 is raised and the precharge path is enabled, electrons flow from f_g to the drain, V_d. Current flow is determined by the value of $\left(f_b - f_g\right)$, the gate-to-source voltage of G24. Initially this voltage is large, G24 is in saturation, and the voltage on f_g rises rapidly. But as f_g rises, currents are reduced and the process slows considerably. Once f_g approaches V_r, the amplifier output falls and G24 transitions quickly through subthreshold and is shut off.

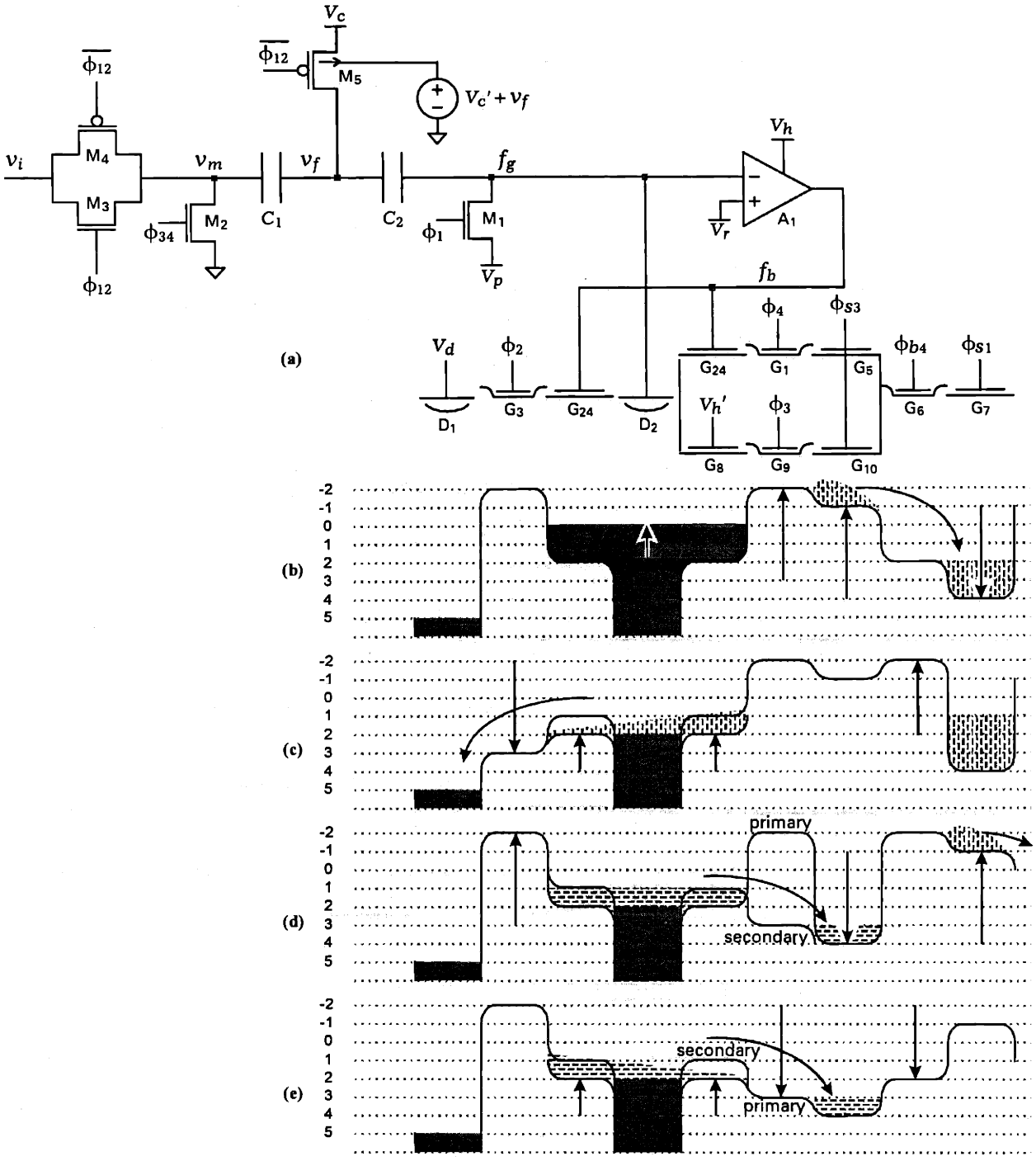


Figure 3-22. Dynamic double-sampling charge generator and associated energy level diagrams. (a) Circuit schematic. (b) Fill phase. (c) Spill Phase. (d) Collection phase. (e) Sensing phase.

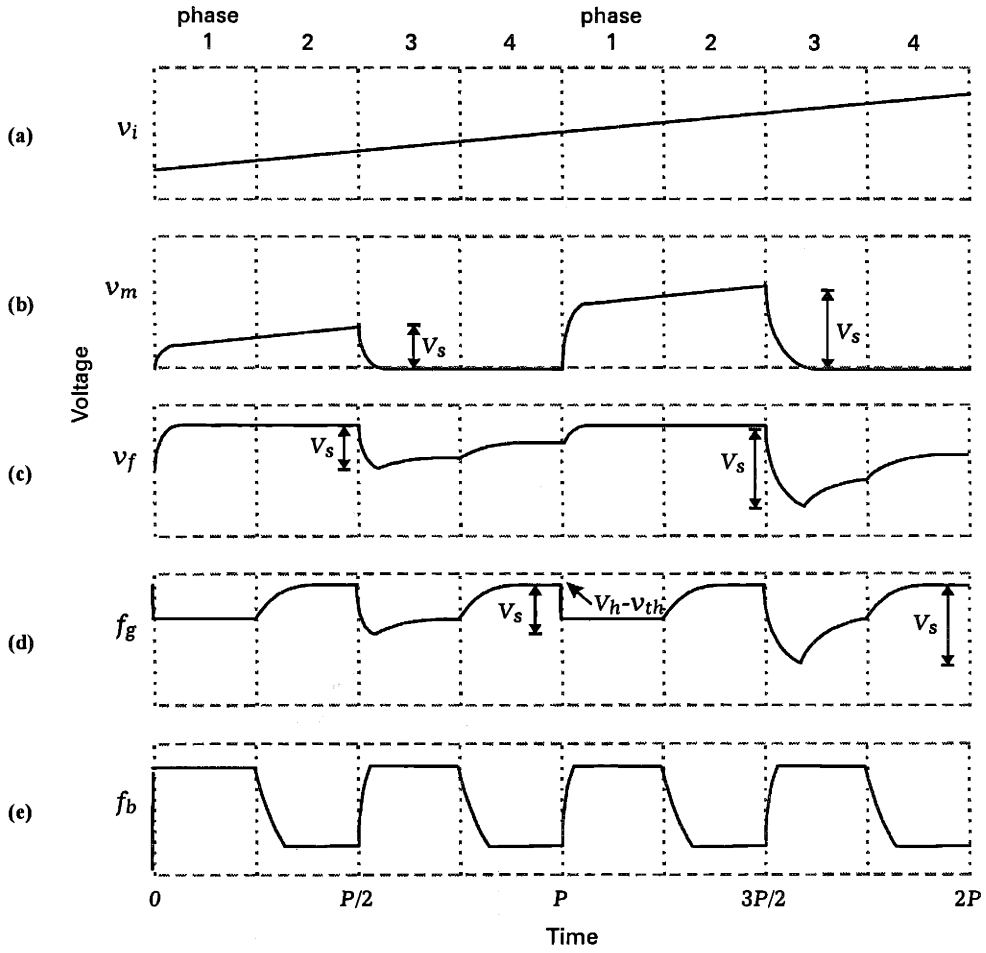


Figure 3-23. Example voltage waveforms for a DDS charge generator. (a)-(c) Signals v_i, v_m, and v_f. (d)-(e) Signals f_g and f_b follow identical paths during phases 2 and 4.

For simplicity in the analysis below, the amplifier is modeled with a linear transfer characteristic, and its saturation at high and low levels is ignored. The value of f_g as a function of time during the spill process is

$$f_g(t) = V_p\, e^{-(t-H)/\tau_2} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-(t-H)/\tau_2}\right) \qquad H \le t \le 2H \ . \tag{3-58}$$

A comparison between the time constant

$$\tau_2 = \frac{R_{24}\left(C_{p2} + C_2\right)}{(A+1)} \tag{3-59}$$

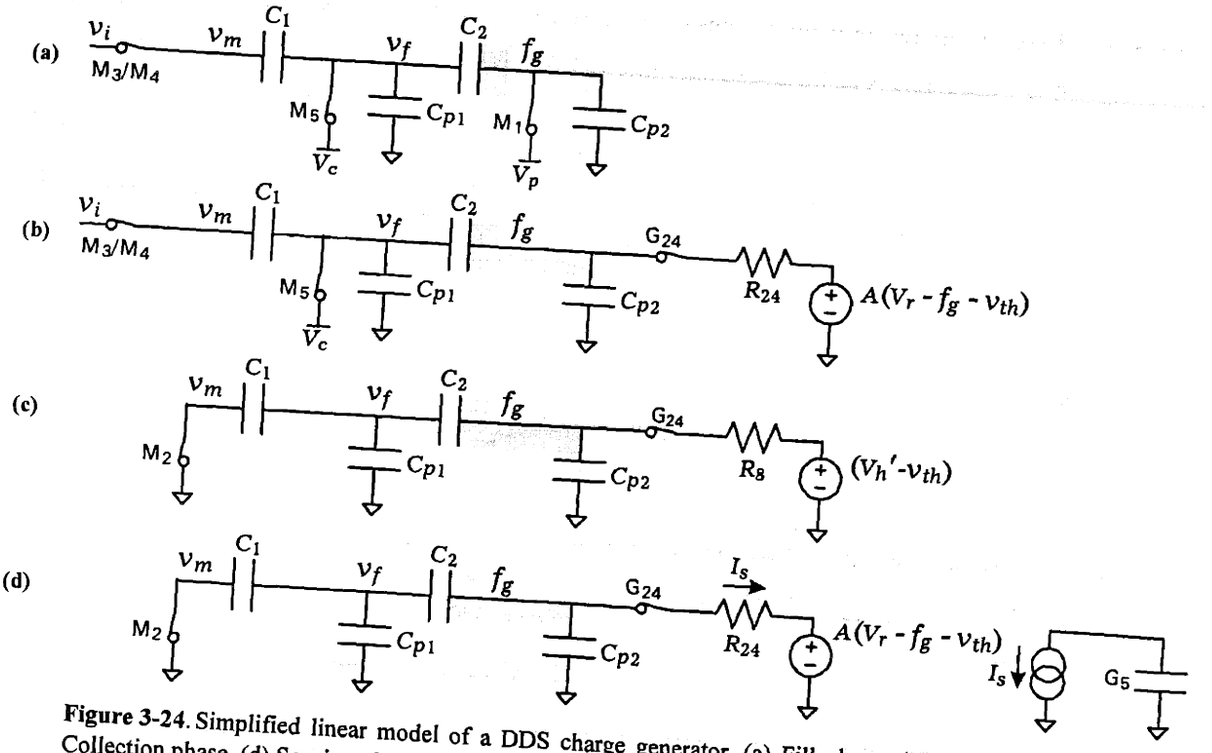


Figure 3-24. Simplified linear model of a DDS charge generator. (a) Fill phase. (b) Spill phase. (c) Collection phase. (d) Sensing phase.

for this operation with that in (3-8) for static sensing shows that the time constant for this circuit is a factor of $(A+1)$ times faster and the incomplete portion of the transition in (3-58) is attenuated by a factor of e^{A+1}. For example, with an amplifier gain of 10, this attenuation is 96 dB. To a very good approximation, the precharge path is completely shut off for the remainder of the cycle with f_g floating at a level of

$$f_g(2H) = \left(\frac{AV_r - v_{th}}{A+1}\right) . \tag{3-60}$$

The final voltage on f_b is

$$f_b(2H) = \left(\frac{A}{A+1}\right)(V_r + v_{th}) . \tag{3-61}$$

Nodes v_m and v_f are actively driven during this phase and are not impacted by the transition on f_g. At the end of phase 2 they have values

$$v_m(2H) = v_i(2H) = V_s \tag{3-62}$$

and

$$v_f(2H) = V_c . \tag{3-63}$$

Variable V_s is used to refer to the analog input v_i at the end of the sampling operation, at time $2H$.

Collection Phase

Phase 3 is referred to as the collection phase. G5 is raised in anticipation of receiving charge and the precharge path is closed by lowering G3. Meanwhile v_m, which was previously tied to the analog input, is clamped to ground. The falling transition on v_m,

$$v_m(2H+\Delta) - v_m(2H) = -V_s \ , \tag{3-64}$$

is assumed to complete during a short time, Δ. It couples through C_1 and C_2 onto v_f and f_g. Node v_f falls toward a final voltage of

$$v_f(2H+\Delta) = V_c - V_s K_1 \ , \tag{3-65}$$

where the unitless constant K_1 is defined as

$$K_1 = \frac{C_1}{\left(C_{p1} + \frac{C_2 C_{p2}}{C_2 + C_{p2}} + C_1\right)} \ . \tag{3-66}$$

The voltage on f_g also falls from its precharged level to

$$f_g(2H+\Delta) = f_g(2H) - V_s K_1 \frac{C_2}{\left(C_2 + C_{p2}\right)} \ . \tag{3-67}$$

The channel capacitance of G24 is included in C_{p2} because charge fills this region during both the fill and collection phases.

For optimum performance, the amplifier should remain saturated at its high level V_h during the entire collection phase. To assure this condition, the input signal range is limited to

$$V_s > \left(\frac{v_{th}}{A+1} - \frac{V_r}{A(A+1)}\right)\left(\frac{C_2 + C_{p2}}{C_2 K_1}\right) \ . \tag{3-68}$$

Meanwhile, the potential on G9 is high and the secondary sensing path through G8 and G9 is open. Electrons flow through this path to the output well by a process similar to that of sensing in a static double-sampling circuit. The voltage on f_g rises toward the channel potential underneath G8. Eventually G8 enters subthreshold and current gradually ceases. The rising transition on f_g during this time is described by

$$f_g(t) = \left(f_g(2H) - V_s K_1 \frac{C_2}{\left(C_2 + C_{p2}\right)}\right) e^{-(t-2H)/\tau_3} + \left(V_h' - v_{th}'\right)\left(1 - e^{-(t-2H)/\tau_3}\right) \qquad 2H \le t \le 3H \ , \tag{3-69}$$

where the time constant τ_3 is

$$\tau_3 = R_8\left(C_{p2} + \frac{C_2\left(C_1 + C_{p1}\right)}{C_1 + C_2 + C_{p1}}\right) \ . \tag{3-70}$$

Sensing Phase

The final phase, phase 4, is referred to as the sensing phase. The sensing path is enabled by raising the voltage on G1. Immediately thereafter, any charge remaining on f_g from the collection phase is transferred

to the output well. The resulting transition on f_g is similar to that during the spill phase. The amplifier output f_b is originally clamped at its high level, V_h. Current flow, determined by $(f_b - f_g)$, slows as f_g approaches V_r. Once f_b begins to fall, G_{24} transitions rapidly through subthreshold and is shut off.

The value of f_g as a function of time is given by

$$f_g(t) = f_g(3H)e^{-(t-3H)/\tau_4} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-(t-3H)/\tau_4}\right) \quad 3H \le t \le 4H \ . \tag{3-71}$$

Like the spill transition, the time constant, τ_4, for the sensing transition,

$$\tau_4 = \frac{R_{24}}{(A+1)}\left(C_{p2} + \frac{C_2(C_1 + C_{p1})}{C_1 + C_2 + C_{p1}}\right) , \tag{3-72}$$

is a factor of $(A+1)$ times faster than that in (3-16) for a static double-sampling circuit.

Since v_f is floating during phase 4, it rises in response to the level

$$v_f(t) = v_f(3H) + \left(f_g(t) - f_g(3H)\right)\frac{C_2}{C_1 + C_{p1} + C_2} \qquad 3H \le t \le 4H \ . \tag{3-73}$$

In contrast to f_g, the value of v_f at the end of the sensing phase differs from its precharged value.

After the sensing path is closed, f_g remains floating for the remainder of the cycle at a level

$$f_g(4H) = \left(\frac{AV_r - v_{th}}{A+1}\right) . \tag{3-74}$$

The feedback signal f_b remains at the voltage

$$f_b(4H) = \left(\frac{A}{A+1}\right)(V_r + v_{th}) \ . \tag{3-75}$$

Current I_s is integrated in both the primary and secondary output wells over the sensing and collection phases. The sum of these two packets represents the final output,

$$Q_o = \left(f_g(4H) - f_g(2H + \Delta)\right)\left(C_{p2} + \frac{(C_1 + C_{p1})C_2}{C_1 + C_2 + C_{p1}}\right) . \tag{3-76}$$

Combining (3-76), (3-71), and (3-67), and retaining only the most significant terms yields the result

$$Q_o = V_s \frac{C_1 C_2}{C_1 + C_2 + C_{p1}} + \left(\left(\frac{AV_r - v_{th}}{A+1}\right)\left(e^{-H/\tau_4} - e^{-H/\tau_2}\right) - V_p e^{-H/\tau_2} + \left(V_h' - v_{th}'\right)e^{-H/\tau_4}\right)\left(C_{p2} + \frac{(C_1 + C_{p1})C_2}{C_1 + C_2 + C_{p1}}\right) . \tag{3-77}$$

When the assumptions $H >> \tau_2$, $H >> \tau_4$, and C_{p1}=0 apply, the result in (3-77) is simplified to

$$Q_o = V_s\left(\frac{C_1 C_2}{C_1 + C_2}\right) . \tag{3-78}$$

This depends only on the sampled input signal and on C_1 and C_2. Although v_{th}, A, and V_r determine the precharge and sensing values of this circuit in (3-58) and (3-71), they do not impact the result in (3-78)

because both precharge and sensing are performed with respect to these same values. Matching between different circuits does not depend critically on any of these parameters.

In the SDS charge generator, described earlier, spill and sensing transitions are slow and never entirely completed. They limit speed and produce a speed-dependent distortion. In contrast, a DDS charge generator does not suffer from the same speed or linearity limitations. Spill and sensing transitions in a DDS circuit are, to a good approximation, performed to completion and the approximations $H >> \tau_2$, $H >> \tau_4$ are nearly always valid. The values of f_g at the end of the spill and sensing phases are then equal, speed-dependent distortion can be eliminated, and parasitic capacitance, C_{p2}, has no impact on the output.

None of the bias voltages V_d, V_c, V_h, V_r, or V_p influence the result in (3-78). Bias V_p should be low enough that charge can fill the region underneath G24 during the fill operation. Any level, such as ground, that is less than V_h-v_{th} is acceptable. Bias V_d should be high enough that it can remove electrons from f_g during the spill operation. It is typically held at the most positive supply. The level of V_c must be high enough that $v_f(3H)$ remains greater than zero for all possible values of V_s. Similarly, the levels of V_h and V_r must be high enough that $f_g(3H)$, given by (3-69), remains greater than zero at all times.

The following sections examine the result in (3-77) under nonideal conditions. The transient linearity of this circuit is considered in section 3.5.2 using more accurate nonlinear device models. Section 3.5.3 examines linearity in the presence of voltage dependent capacitors. Finally, section 3.5.4 explores the impact of various noise sources on circuit performance.

3.5.2 Transient Linearity

The distortion characteristics of the circuit in Figure 3-22 are examined below using more accurate models for G24 and A1. The amplifier is represented by the transfer characteristic

$$f_b(t) = \begin{cases} A\left(V_r - f_g(t-t_d) - a_1 f_g{}^2(t-t_d)\right) + V_h/2 & 0 \le f_b \le V_h \\ V_h & f_b > V_h \\ 0 & f_b < 0 \end{cases} . \qquad (3\text{-}79)$$

This model includes a steady-state bias point of $V_h/2$, output saturation at 0 and V_h, a time delay of t_d, and second harmonic distortion determined by the coefficient a_1.

At the beginning of the sensing process, while G24 is in saturation, it is represented by a current-voltage characteristic of

$$I_s(t) = \frac{KW}{2L}\left(f_b(t) - f_g(t) - v_{th}\right)^2 , \qquad (3\text{-}80)$$

where I_s is the current flowing from the receiving well toward f_g, K is the transconductance parameter of G24, measured in A/V^2, and W/L represents its width-to-length ratio. The effect of drain voltage on I_s is ignored because, as was described in section 3.4.1, G24 and G1 form a cascode combination whose output

impedance is large. This approximation is especially accurate for a DDS circuit, because most of its signal charge is stored in the secondary output well, and the signal dependent potential of the drain of G24 is reduced by a factor of e^{-H/τ_3}.

Near the end of the sensing process, when f_g has risen to within $2kT/q$, or 52 mV, of its final value, G24 is in weak inversion with a current-voltage relationship of

$$I_s(t) = \frac{2K(kT/q)^2}{e^2}\frac{W}{L}e^{\left(A(V_r - f_g(t)) - f_g(t) - v_{th}\right)/(nkT/q)} \quad . \tag{3-81}$$

In an SDS circuit, the cascode gate G24 is held at a fixed bias and does not couple to f_g. In a DDS circuit, this gate is controlled by the amplifier output and coupling, which occurs between it and f_g, alters signal charge. To model this effect, parasitic gate-to-source capacitance, denoted by C_{p3}, is included for G24. Its current equals

$$I_3(t) = C_{p3}\frac{d\left(f_b(t) - f_g(t)\right)}{dt} \quad . \tag{3-82}$$

During both saturation and subthreshold periods, the relationship between I_s, I_3, and f_g is

$$I_s(t) + I_3(t) = \left(\frac{C_2\left(C_1 + C_{p1}\right)}{C_1 + C_{p1} + C_2} + C_{p2}\right)\frac{df_g(t)}{dt} \quad . \tag{3-83}$$

Numerical simulations are used to solve the system given by equations (3-79), (3-80), (3-81), (3-82) and (3-83). Simulations are based on the example set of parameters,

$$\begin{aligned}
&\left(\frac{C_2\left(C_1 + C_{p1}\right)}{C_1 + C_{p1} + C_2} + C_{p2}\right)^{-1} = 0.5 \text{ pF} \\
&C_{p3} = 10 \text{ fF} \\
&K = 100\ \mu\text{A} / \text{V}^2 \\
&\frac{W}{L} = 40 \\
&V_r = 1 \text{ V} \\
&\text{A} = 12 \\
&t_d = 0.125 \text{ nsec} \\
&\text{a}_1 = 0.25 \\
&n = 1.4
\end{aligned} \quad . \tag{3-84}$$

Two sets of simulation results are discussed below. The first set illustrates the dynamic turn-off of the spill and sensing transitions. When the secondary sensing path consisting of G8, G9, and G10 is eliminated from the circuit of Figure 3-22. This path is included in the second set of simulations presented later in this section.

Without the Secondary Sensing Path

Without the secondary sensing path, charge dumped onto f_g during collection remains there until sensing. The value of f_g before sensing begins is referred to as f_{g0}. Larger values of V_s produce smaller values of f_{g0}. The ensuing transitions on f_g, f_b, and I_s are plotted as a function of time in Figure 3-25. Curves are shown for V_h values of 2 and 1.3V and f_{g0} values of 0, 0.25, 0.5, and 0.75V. These curves show the drastic speed improvement of the sensing turnoff that occurs once f_b begins to fall.

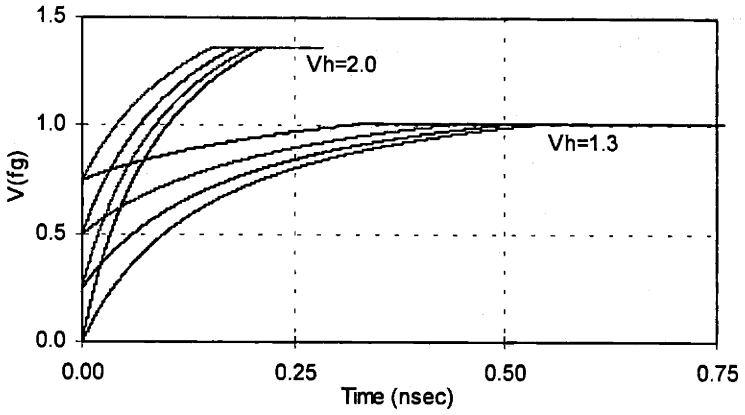
The model described by equations (3-79), through (3-83) is used to determine circuit linearity using a technique similar to that of section 3.4.2. First, the signal dependence of $f_g(4H)$ is determined. Then the resulting curves are fit to a third order polynomial and the second and third-order coefficients are evaluated. These coefficients are translated into a measure of distortion using (3-29) and (3-30) and the results are plotted as a function of clock period, $4H$, in Figure 3-26 for V_h values of 2 and 1.3V. Longer cycle times are chosen for the simulation with V_h=1.3 because its transitions are slower. During the time that f_b is clamped at V_h, the distortion decreases slowly as clock periods are increased. Once f_b transitions low, the distortion decreases rapidly and then remains constant. At their lower level, curves are limited by numerical simulation accuracy. A comparison between these results and those in Figure 3-15, for an SDS charge generator, shows that the DDS technique provides a dramatic improvement in linearity.

A critical aspect of the operation of this circuit is that precise details of each transition in Figure 3-25 are unimportant, provided transitions are signal independent and consistent from one cycle to the next. This is a consequence of the fact that, to a very good approximation, its spill and sensing transitions are completed. The different curves in Figure 3-25 follow identical paths, except for an initial signal-dependent delay. The turn-off transition and the final value of f_g are independent of f_{g0}. This applies even when transitions and settling values differ between the spill and sensing phases.

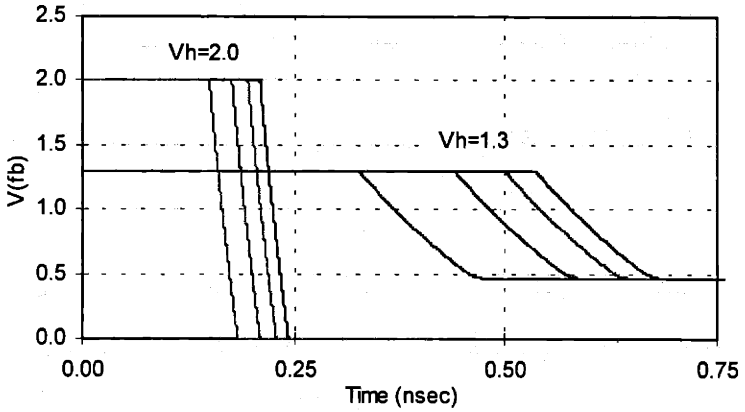
Because transition details are unimportant, this circuit is insensitive to the specific values of its bias voltages, amplifier gain, amplifier nonlinearity, amplifier delays, transistor current-voltage characteristics, transistor threshold voltages, parasitic capacitances, and clock frequency. A wide range of bias conditions and device parameters work equally as well for this circuit and large amplifier gains are not needed.

The simulation results in Figure 3-25 include a delay of 0.125 nsec. If this delay were zero, then all f_g curves would settle to the same final value, given by (3-74), regardless of the value of V_h. This value is approximately equal to V_r. When delays are nonzero, the settling value, $f_g(4H)$, depends on the value of V_h. An increase in V_h produces a greater settling voltage and a more rapid sensing transition. A decrease in this level produces a settling voltage that is closer to V_r and a slower sensing transition.

(a)



(b)



(c)

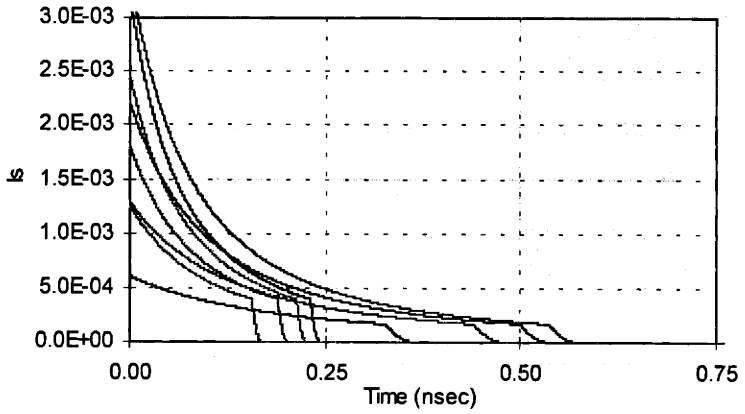


Figure 3-25. Transitions of (a) f_g, (b) f_b, and (c) I_s during the sensing phase for f_{g0} = 0, 0.25, 0.5, and 0.75V and V_h = 2 and 1.3V. As V_h approaches V_r settling times increase.

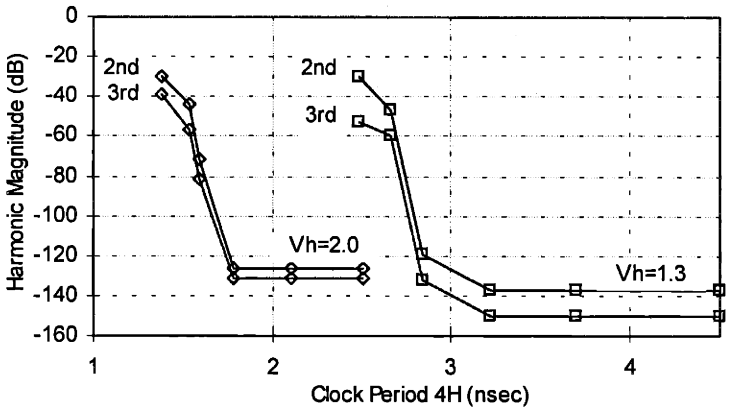


Figure 3-26. Second and third harmonic distortion as a function of clock period.

The points on each curve in Figure 3-26 are independent of clock frequency, provided that enough time is allowed for the spill and sensing transitions to complete. As a result, distortion in this circuit is not improved by increasing the transconductance K, the size ratio W/L, or the clock period $4H$. Once the sensing paths are closed, f_g remains fixed for the remainder of the phase and the resulting charge packet is not altered.

In contrast to a static charge generator, this circuit operates linearly down to charge packets given by the bound in (3-68), which is typically less than 5% of the full scale. Little background charge is required. This bound is equivalent to a requirement that f_b is saturated at V_h at the beginning of the spill and sensing phases.

With the Secondary Sensing Path

This secondary sensing path does not add much complexity to the circuit, and at the same time, it brings advantages. Example transitions on f_g, f_b, and I_s during phases 3 and 4 are plotted as a function of time in Figure 3-27. The same set of parameters, given by (3-84), is used for the simulation. Curves are shown for V_h equal to 2 and 1.3V and f_{g0} equal to 0, 0.25, 0.5, and 0.75V. The bias V_h' is set to 0.9V and the collection phase is assumed to last for 1 nsec. Charge dumped onto f_g during the collection phase is skimmed off into the secondary receiving well by a process identical to that for sensing in a static charge generator.

One purpose of the secondary sensing path is to reduce nonlinearities caused by higher-order effects not accounted for in the simulation model described above. This process reduces nonlinearity in two ways. First, the signal level on f_g is attenuated by a factor of e^{-H/τ_3} before the start of the sensing phase. Any nonlinearity caused by dependence of the sensing turn off transition on $f_g(3H)$ is reduced by this same factor. Second, most of the signal charge is stored in the secondary output well and signal dependence in the primary output well is reduced by a factor of e^{-H/τ_3}. Any nonlinearity caused by dependence of I_s

41) and (3-42) for the parasitic capacitor C_{p1}. Input capacitors have the strongest effect on linearity. However, they can be made from materials with a low voltage dependence. A larger relative voltage dependence is acceptable from the parasitic capacitor but this element contains components, such as drain-bulk junctions, that are strongly voltage dependent.

Two approaches are possible to reducing nonlinearity caused by parasitic capacitors. Either their voltage dependence or their voltage swing can be decreased. The first approach can be accomplished by setting the voltage V_c to a level lower than the bulk of M_5, and thereby operating M_5 with a larger bulk-to-drain reverse bias. The second approach can be accomplished using the technique shown in Figure 3-22. A voltage-dependent voltage source is used to drive the bulk of M_5 with a signal that is an offset version of that on its drain. The voltage drop across the M_5 drain-bulk junction therefore remains constant between time $t=2H$ and $t=4H$. The voltage-dependent voltage source can be implemented from a buffer or by using the output from a second clamp-and-sample circuit with the same input.

3.5.4 Noise

A DDS charge generator is subject to similar noise sources as the SDS charge generator described above. These include thermal noise, flicker noise, supply noise, clock jitter, clock feedthrough, and charge injection. First, the general noise transfer characteristics are derived for this circuit. Then the impact of each of these noise sources is described.

General noise characteristics of this circuit are derived using the models in Figure 3-24 and the assumptions $H >> \tau_2, \tau_4$ and $C_{p1}, C_{p2} << C_1, C_2$. Noise, produced by the clamp-and-sample portion of this circuit, is identical to that previously derived in section 3.4.4. It is given by

$$\Delta v_m = V_s + n_m(2H) - n_m(4H) \ , \tag{3-85}$$

where $n_m(2H)$ represents thermal noise, clock feedthrough, and charge injection from the sampling switch and $n_m(4H)$ contains ground noise and thermal noise but no clock feedthrough or charge injection.

Immediately after the end of the spill phase, f_g has a precharge value of

$$f_g(2H) = \left(\frac{AV_r - v_{th}}{A+1}\right) + n_g(2H) + n_r(2H)\left(\frac{A}{A+1}\right) + n_b(2H)\left(\frac{1}{A+1}\right) \ . \tag{3-86}$$

Shortly before the end of the sensing phase it has a value of

$$f_g(4H) = \left(\frac{AV_r - v_{th}}{A+1}\right) + n_g(4H) + n_r(4H)\left(\frac{A}{A+1}\right) + n_b(4H)\left(\frac{1}{A+1}\right) \ . \tag{3-87}$$

In these expressions, n_g represents thermal noise in the spill and sensing paths through G_{24}. Signal n_r represents supply noise in the reference V_r and thermal noise in the input stage of the amplifier. Noise introduced at the amplifier output is denoted n_b. It is attenuated by the amplifier gain and is, therefore, ignored in the discussion below. No charge injection or clock feedthrough occurs in this process.

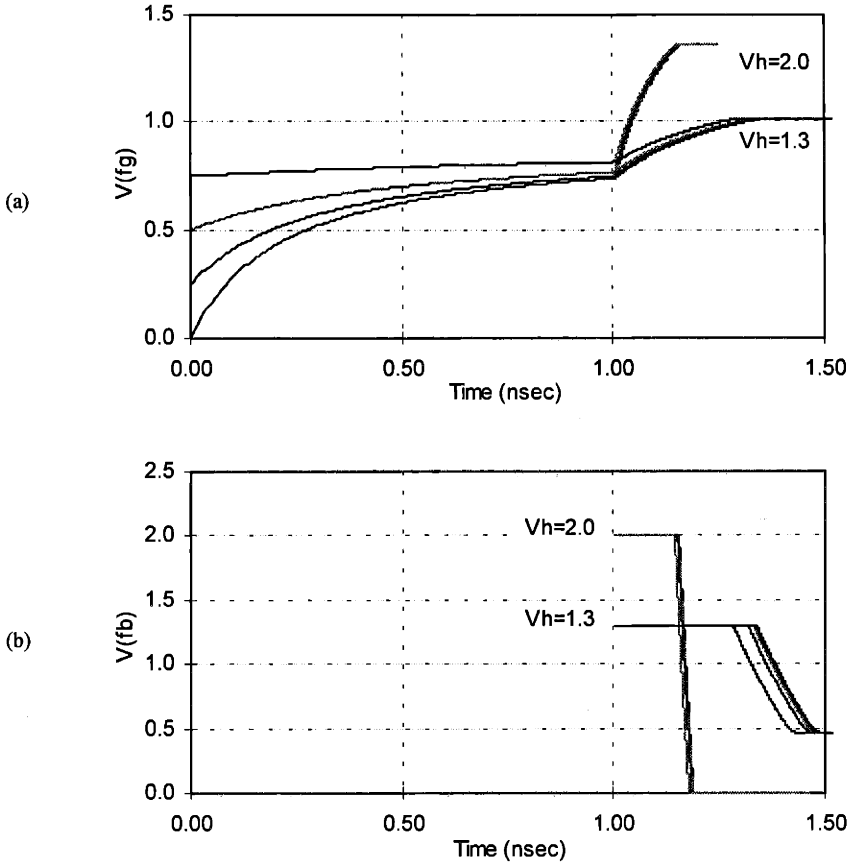


Figure 3-27. Transitions of (a) f_g and (b) f_b during the sensing phase for initial f_g values of 0, 0.25, 0.5, and 0.75. Two sets of curves correspond to V_h values of 2.0 and 1.3V. As V_h approaches V_r settling times increase.

and the sensing turn-off transition on the drain potential of G_{24} is reduced by this same factor. For example, even if the phase time spans only two time constants, a 17-dB improvement is achieved in these two categories of nonlinearity.

A final advantage of the secondary sensing path is that during the collection phase, charge is transferred immediately to the secondary output well and is not stored on the input node, f_g. As a result, thermal noise can be reduced by reducing the capacitance, C_{p2}. This benefit applies to most circuits that use the DDS technique, but does not apply to a charge generator, where input capacitance is fixed by the size of the charge packet to be generated.

3.5.3 Capacitance-Dependent Linearity

A DDS charge generator experiences distortion as a result of voltage dependence in its capacitors. The relationship between voltage dependence and harmonic distortion is identical to that for the SDS charge generator in Figure 3-10 and is given by (3-36) and (3-37) for the input capacitors, C_1 and C_2, and by (3-

The output that is generated by the DDS core depends on the input charge in (3-85) and the difference between the precharge and sensing voltages in (3-86) and (3-87). It equals

$$Q_o = V_s \frac{C_1 C_2}{C_1 + C_2} + \left(n_g(4H) - n_g(2H) + \left(n_r(4H) - n_r(2H)\right)\left(\frac{A}{A+1}\right) + n_m(2H) - n_m(4H) \right) \frac{C_1 C_2}{C_1 + C_2} \quad . \tag{3-88}$$

Like an SDS charge generator, the output packet in a DDS circuit depends only on the difference between noise sources at the end of phases 2 and 4. Noise at low frequencies is common to both precharge and sensing phases and is attenuated, while noise near the Nyquist frequency is passed directly to the output. Because (3-88) does not depend on V_c or V_p, the circuit is tolerant of supply noise on these biases and thermal noise in their associated switches. The result in (3-88) also does not depend on details of the secondary sensing transition, and neither noise on V_h' nor thermal noise, captured at the end of phase 3, are important. The CCD gates after G1 do not add any additional noise because they operate in a fully-depleted manner.

Thermal Noise

The thermal noise response of a DDS charge generator is similar to that of an SDS charge generator. The n_m terms in (3-88), introduced within the clamp-and-sample circuit, have mean-square values of

$$\begin{aligned} \overline{n_m(2H)^2} &= kT\frac{1}{C_1} \\ \overline{n_m(4H)^2} &= kT\left(\frac{1}{C_1} + \frac{1}{C_2}\right) \end{aligned} \quad . \tag{3-89}$$

The n_g terms in (3-88), introduced within the DDS circuit, have mean-square values of

$$\begin{aligned} \overline{n_g(2H)^2} &= kT\frac{1}{C_2} \\ \overline{n_g(4H)^2} &= kT\left(\frac{1}{C_1} + \frac{1}{C_2}\right) \end{aligned} \quad . \tag{3-90}$$

The additional noise component n_r is determined by details within the amplifier. It is ignored in the description below because it does not relate to CCD capacitances. Under the assumption of independence, the combined mean-square noise is

$$\overline{Q_o{}^2} = 3kT\left(\frac{C_1 C_2}{C_1 + C_2}\right) \quad . \tag{3-91}$$

Thermal noise in the output signal decreases with decreasing capacitance. But signal levels decrease with decreasing capacitance as well. Equations (3-78) and (3-91) are combined to describe SNR,

$$\text{SNR} = V_s \sqrt{\frac{C_1 C_2}{3kT(C_1 + C_2)}} \quad , \tag{3-92}$$

for an input sinusoid with an rms amplitude of V_s. SNR improves for this circuit as input capacitances are increased because the signal increases more rapidly than the noise. Noise in the amplifier front end should be kept small. Noise in the amplifier output is less critical because it is attenuated by the amplifier gain.

Supply Noise and Flicker Noise

Supply noise on V_h and V_r and flicker noise in G24 influence the cascode channel potential and cause the sensing and precharge values of f_g to differ. These noise terms are correlated between the spill and sensing phases because they occur in the same gate. Like an SDS charge generator, this circuit is not sensitive to flicker noise or low-frequency power-supply noise because noise quantities in (3-88) are differentiated and only changes in noise values between the spill and sensing phases impact the result.

Clock Jitter and Charge Injection

The dominant source of jitter in this circuit is the voltage-mode sample-and-hold block at its input. This type of sampling noise impacts SDS and DDS charge generators identically and is described in section 3.4.4.

For best linearity and sampling accuracy, the precharge device M5 should be turned off before the input sampling switch is disabled. In this case, charge injection that occurs in the input sampling switch when it is disabled, does not effect the circuit's result. This feature is a result of charge-domain operation and can be understood in the following way. Nodes v_f and f_g are precharged during the time that the input switch is enabled. Charge injection, which occurs after the precharge path is closed, causes the potential on v_m to change temporarily. This in turn causes the potentials on v_f and f_g to change temporarily but does not cause any charge loss from either node. The final result depends only on the difference between the precharge value of v_m, while the sampling switch is on, and the sensing value of v_m, while it is clamped to ground. It does not depend on any intermediate transitions on these nodes.

A second benefit that occurs when M5 is turned off before the input sampling switch is disabled is improved sampling linearity. In this configuration, sampling occurs, not when the sampling switch is disabled, but rather when M5 is disabled. The precharge potential on v_f is independent of the incoming signal and this results in a signal independent sampling instant.

Neither jitter nor charge injection are significant contributors of noise in the charge-domain portion of a DDS charge generator. These noise sources are proportional to the current that is flowing at the time that clocks are shut off. An example of current during the spill and sensing transitions of a DDS circuit are shown in Figure 3-25(c). Since currents are to a good approximation zero at the end of the spill and sensing phases, timing variations in the ϕ_2 and ϕ_4 clocks do not change the result. When the fall time of the feedback signal f_b is large compared to the transit time of carriers across gate G24, carriers in the channel of this gate are gradually reduced to zero and no charge injection error occurs.

3.6 Other Applications of Double-Sampling

The DDS charge generator in Figure 3-22(a) is composed of two blocks; a sample-and-hold front end, and a dynamic double-sampling core. The purpose of the sample-and-hold block is to produce a displacement charge proportional to the analog input voltage, v_i. The purpose of the DDS core is to integrate this displacement charge in a CCD output well. The DDS technique can also be applied to a number of other circuit functions, in addition to charge generation. In this approach, the sample-and-hold block is removed and an alternative source of electrons is provided to the DDS input. Examples of such circuits, including wire transfer, charge sensing, charge subtraction, and D/A conversion are presented below.

3.6.1 Double-Sampling Wire Transfer

The originating and receiving wells in a fully depleted CCD transfer must be located adjacent to one another so that their potential wells touch. Because of this requirement, only a limited set of signal processing topologies can use fully depleted circuits. Wire transfer circuits provide an additional degree of flexibility by allowing charge packets to be transferred between nonadjacent or distant wells, via a metal line. The objective in this approach is to use fully depleted operations whenever possible, and to interconnect groups of these circuits by a wire transfer.

The conventional wire transfer circuit, described in section 3.1.3, has the undesirable attribute that its speed and linearity are severely limited by subthreshold time constants and signal lag. The SDS technique could potentially be applied to a wire transfer circuit to eliminate signal lag, but it would provide only a minor improvement in linearity and speed. Additional improvements in linearity and speed are achieved with the DDS technique.

To incorporate dynamic double sampling into a wire transfer circuit, the output from a CCD register is connected to the input of a DDS core as shown in Figure 3-28(a). A CCD register, formed by gates G_{11}, G_{12}, and G_{13}, serves as the source of electrons. The DDS unit reconstructs incoming charge packets by integrating their electrons in a receiving well.

The operation of this circuit is similar to that of the DDS charge generator of section 3.5.1. It is illustrated using energy level diagrams in Figure 3-28(b)-(e) During the fill phase, a new supply of charges is provided to the circuit when f_g is pulled low to V_p. Gate G_{12} is off at this time to prevent charge from flooding backward into the source register. During the spill phase, f_g is reset high to a level of

$$f_g(2H) = V_p\, e^{-H/\tau_2} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-H/\tau_2}\right) . \tag{3-93}$$

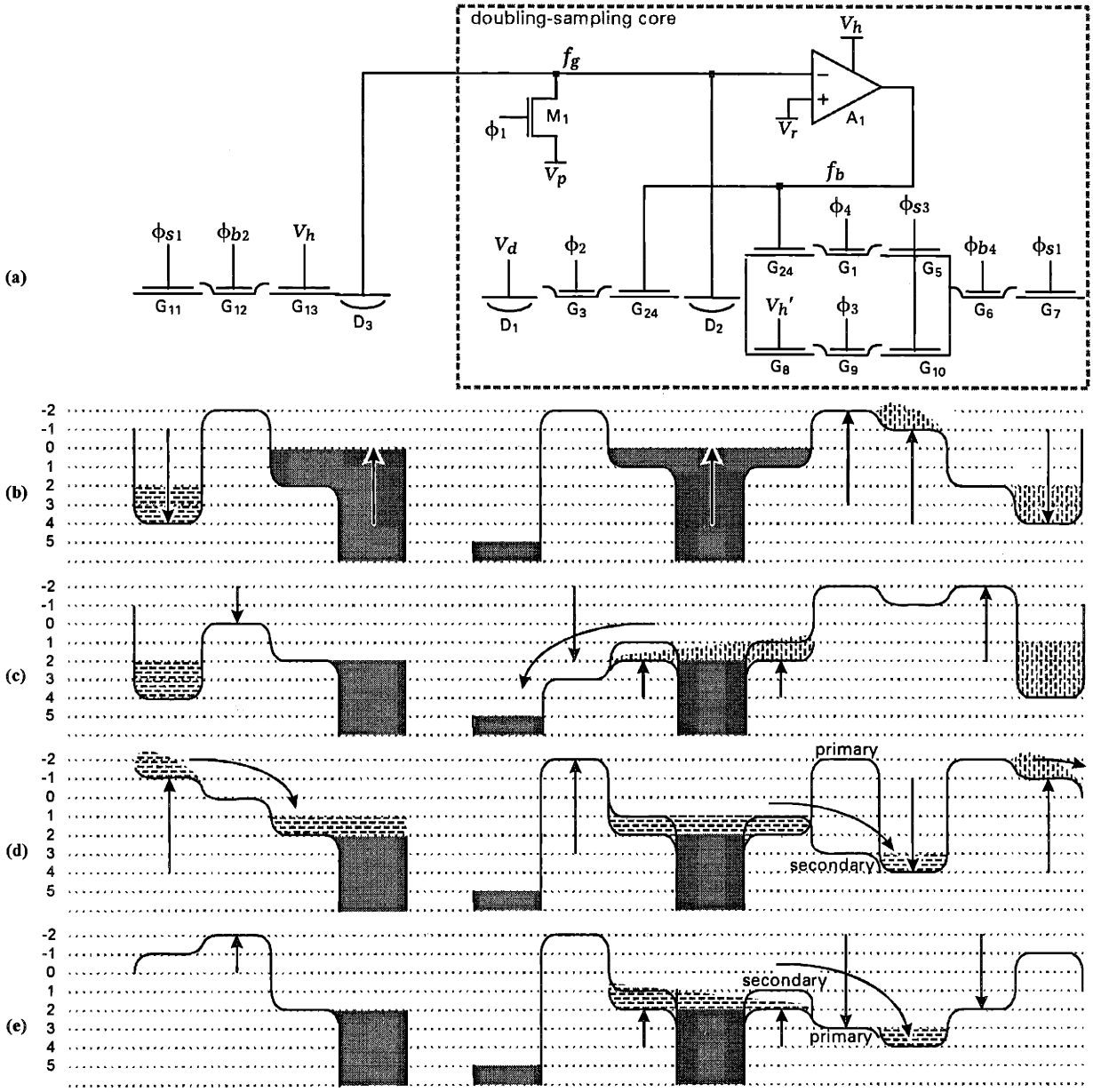


Figure 3-28. DDS wire transfer circuit. (a) Schematic. (b) Fill phase. (c) Spill phase. (d) Collection phase. (e) Sensing phase.

The time constant for this transition is

$$\tau_2 = \frac{R_{24}C_{p2}}{(A+1)} \quad , \tag{3-94}$$

where C_{p2} represents parasitic junction and routing capacitances on f_g as well as channel capacitance of G_{24}, G_8, and G_{13}.

During the collection phase, a quantity of charge Q_s is transferred onto f_g by lowering the potential of G_{11}. The potential on f_g falls in response, and charge flows through the secondary sensing path, into the secondary receiving well G_{10}. The voltage on f_g is restored to a high level of

$$f_g(3H) = \left(f_g(2H) + \frac{Q_s}{C_{p2}} \right) e^{-H/\tau_3} + \left(V_h' - v_{th}' \right) \left(1 - e^{-H/\tau_3} \right) \quad . \tag{3-95}$$

During the sensing phase, charge, any remaining on the wire transfer node, flows through G_{24} to the primary output well, G_5. The final value of f_g is

$$f_g(4H) = f_g(3H) e^{-H/\tau_4} + \left(\frac{AV_r - v_{th}}{A+1} \right) \left(1 - e^{-H/\tau_4} \right) \quad . \tag{3-96}$$

The time constant for the sensing transition, τ_4, equals that for the spill transition, τ_2.

Combining (3-93), (3-95), and (3-96), and eliminating higher-order exponential terms, yields the resultant charge packet

$$Q_o = Q_s - \left(V_h' - v_{th}' + V_p \right) e^{-H/\tau_4} C_{p2} \quad . \tag{3-97}$$

In most cases, the approximation $H >> \tau_4$ is valid for this circuit and the first term, which represents a constant offset, is negligible. The result is then identically equal to the input Q_s, regardless of any parasitic capacitances.

Linearity

The transient linearity of a DDS wire transfer circuit is similar to that described in section 3.5.2 for a DDS charge generator. Precise details of the spill and sensing transitions are unimportant, provided they are signal independent and consistent from one cycle to the next. As a result, this circuit is not sensitive to specific bias voltages, amplifier gain, amplifier nonlinearity, amplifier delays, transistor current-voltage characteristics, transistor threshold voltages, parasitic capacitances, and clock frequency. Figure 3-29 shows a comparison between the linearity and speed of SDS versus DDS wire transfer techniques. The DDS approach has significantly less distortion and operates down to clock periods of a few nanoseconds.

A DDS wire transfer circuit is not subject to capacitance-dependent distortion. Its only capacitor is the parasitic capacitance, C_{p2} and this element does not enter the signal term in (3-97).

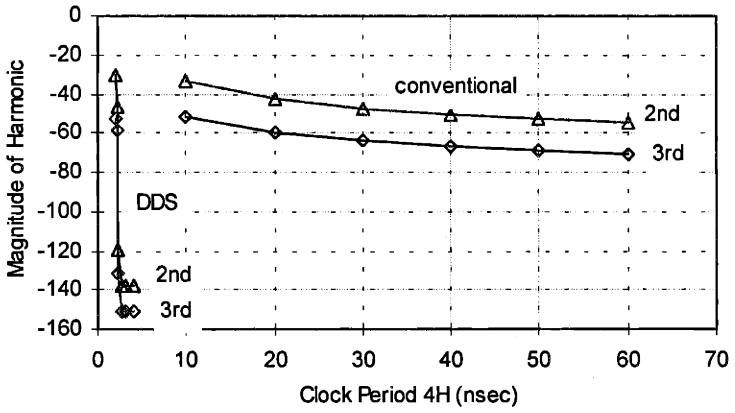


Figure 3-29. Transient linearity and speed comparison between conventional and DDS wire transfer circuits. Both of these attributes are greatly improved by the DDS technique.

Thermal Noise

In contrast to a charge generator, a wire transfer circuit is not strongly impacted by thermal noise. Its only source of thermal noise is the transfer of charge from f_g through G_{24}. If this is represented by variable n_g, then the result in (3-96) becomes

$$Q_o = Q_s + \left(n_g(4H) - n_g(2H)\right) C_{p2} \quad . \tag{3-98}$$

The mean square value of n_g equals

$$\overline{n_g(2H)^2} = \overline{n_g(4H)^2} = kT\frac{1}{C_{p2}} \quad . \tag{3-99}$$

The resulting charge packet has an rms variation of

$$\sqrt{\overline{Q_o^2}} = \sqrt{2kT\,C_{p2}} \quad . \tag{3-100}$$

As is generally the case for charge-domain signals, thermal noise decreases with decreasing capacitance. Yet, the signal level is independent of this capacitance because the secondary sensing path transfers charge directly through G_8, into the receiving well. Charge is never stored on the wire transfer node. The resulting SNR for an rms input charge of Q_s is

$$\text{SNR} = Q_s\sqrt{\frac{1}{2kT\,C_{p2}}} \quad . \tag{3-101}$$

SNR, power, and speed are all improved by a reduction in capacitance on the wire transfer node and this capacitance can generally be kept small.

Clock Feedthrough

The final gate G_{13} in the input register is tied to a fixed bias, V_h. This is done to eliminate clock feedthrough that would otherwise occur from the last barrier gate G_{12} to the wire transfer node. Since no diffusion is present on the drain of G_{12}, coupling from ϕ_{b2} to f_g is nearly eliminated. The dc gate G_{13}

does not alter circuit operation. However, the value of C_{p2} is increased by the channel capacitance of G_{13} because charge resides underneath this gate.

3.6.2 Double-Sampling Charge Sensing

Nondestructive charge sensing is used when a charge packet must be used multiple times. It replicates the original packet in either charge or voltage form, but does not alter the original in the process. The conventional floating gate sensing technique of section 3.3.2 has a number of disadvantages. First, it has poor linearity because its output depends on MOS channel-to-substrate capacitance and parasitic capacitances, both of which are voltage dependent. Second, its charge-referred resolution is low because it has an output signal range that is limited to a small fraction of the power supply voltage. Third, it requires twice the clock voltage swing of other CCD shift registers. Finally, matching between multiple floating gate amplifiers is poor because circuit gain depends on CCD well capacitances, which are difficult to control precisely.

The dynamic double-sampling technique can be applied to charge sensing to improve performance in some of these regards. The resulting circuit, shown in Figure 3-30, is formed by connecting floating gate G_{13} from within a CCD register, to the input of a DDS block. Input charge is provided by coupling signal electrons from the floating gate channel to the DDS sensing node. These carriers are integrated in a CCD well by the DDS block to reconstruct the original charge packet.

The operation of this circuit occurs as follows. During the fill phase, f_g is pulled low to V_p. This serves two purposes. First, it provides a new supply of electrons to the DDS circuit. Second, it presets the floating gate low and forces electrons, underneath it from a previous cycle, to be transferred forward.

During the spill phase, f_g is preset high to a level of

$$f_g(2H) = V_p\, e^{-H/\tau_2} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-H/\tau_2}\right) . \tag{3-102}$$

The time constant for this transition is

$$\tau_2 = \frac{R_{24}}{(A+1)}\left(C_{p2} + \frac{C_{gc}C_{cs}}{\left(C_{gc} + C_{cs}\right)}\right) , \tag{3-103}$$

where C_{p2} represents parasitic junction and routing capacitances on f_g as well as channel capacitance of G_{24} and G_8. The second term in this expression corresponds to effective capacitance of the CCD well when it is empty. (Effective capacitance is described in section 3.8.) The high potential on the floating gate at this time leaves it well situated to receive charge during the following phase.

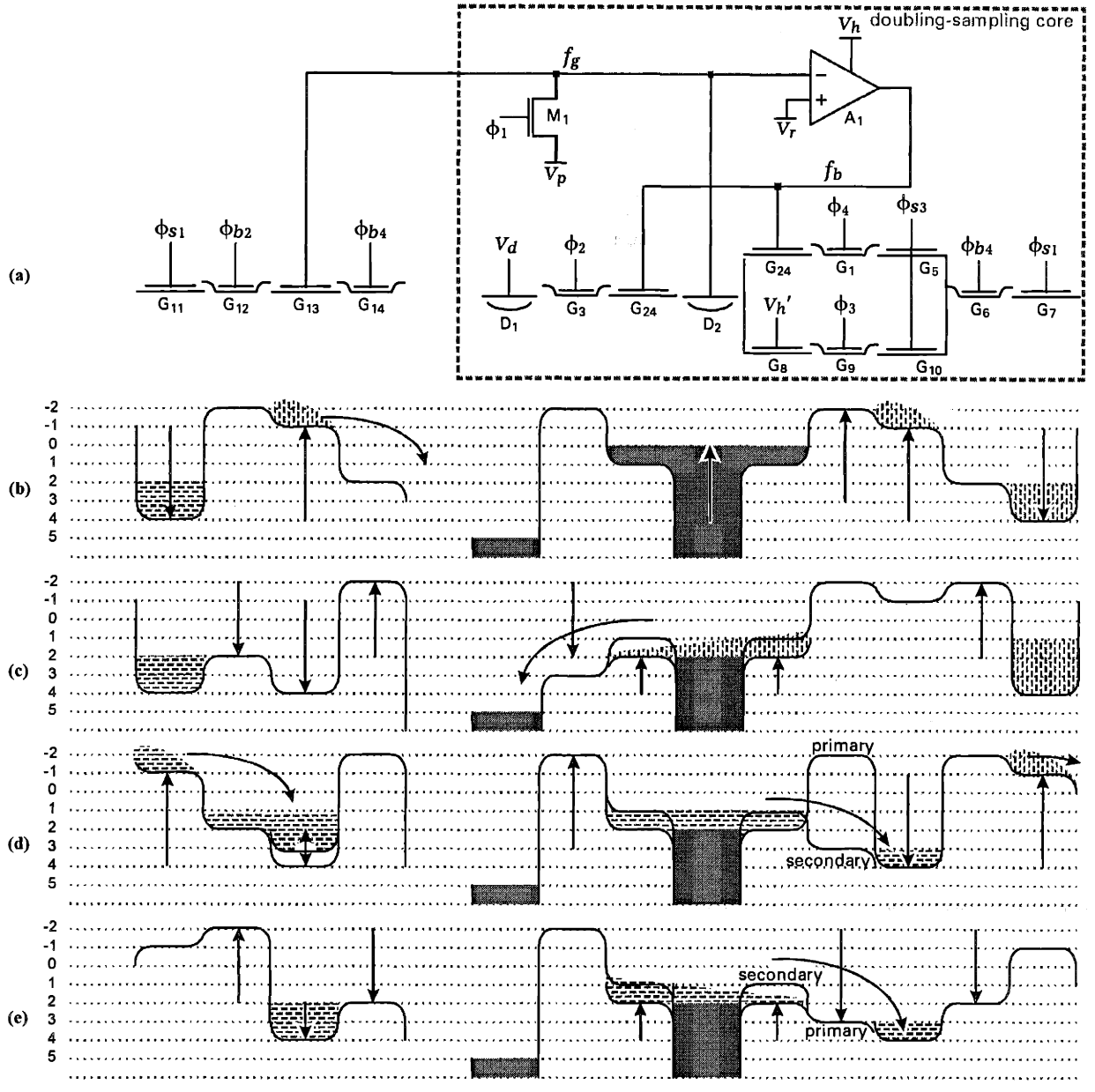


Figure 3-30. DDS charge sensing circuit with charge output. (a) Schematic. (b) Fill phase. (c) Spill phase. (d) Collection phase. (e) Sensing phase.

At the beginning of the collection phase, the floating gate well contains no charge. Its storage capacity, which is proportional to the difference between the channel potentials of G13 and G12, is approximately equal to

$$|Q|_{\max} = (V_r - v_{th1} - 4 + v_{th2})(C_{gc} + C_{cs}) \ . \tag{3-104}$$

Variables v_{th1} and v_{th2} refer to threshold voltages of the first and second level gates and 4V represents the high level of the ϕ_{b2} clock. A quantity of charge Q_s is transferred underneath the floating gate by lowering the potential of G11. The floating gate potential falls temporarily in response and its storage capacity is reduced to approximately

$$|Q|_{\max} = \left(V_r + Q_s \left(\frac{C_{gc}}{C_{gc}C_{p2} + C_{cs}(C_{gc} + C_{p2})} \right) \left(\frac{C_{gc}}{C_{gc} + C_{cs}} \right) - v_{th1} - 4 + v_{th2} \right) (C_{gc} + C_{cs}) \ . \tag{3-105}$$

When Q_s is large, this reduction in storage capacity may cause some electrons to be temporarily stored underneath the G12 barrier, where they are no longer sensed by the floating gate. But as charge flows through the secondary sensing path, the voltage on f_g rises, storage capacity is restored, and electrons are once again confined to the floating gate. At the end of collection, the level on the floating gate is

$$f_g(3H) = \left(f_g(2H) + Q_s \frac{C_{gc}}{C_{gc}C_{p2} + C_{cs}(C_{gc} + C_{p2})} \right) e^{-H/\tau_3} + (V_h' - v_{th}')(1 - e^{-H/\tau_3}) \ . \tag{3-106}$$

During the sensing phase, any charge that remains on f_g is passed to the primary output well, and the potential on f_g rises to its final value of

$$f_g(4H) = f_g(3H)e^{-H/\tau_4} + \left(\frac{AV_r - v_{th}}{A+1} \right) (1 - e^{-H/\tau_4}) \ . \tag{3-107}$$

The time constant for this transition is

$$\tau_4 = \frac{R_{24}}{(A+1)} \left(C_{p2} + (C_{gc} + C_{cs}) FK_{eff} \right) \ , \tag{3-108}$$

As described in section 3.8, the factor FK_{eff} is used to translate storage capacity into effective capacitance seen by the gate.

Combining (3-102), (3-106), and (3-107), and eliminating higher-order exponential terms, yields the resultant charge packet

$$Q_o = Q_s \frac{C_{gc}}{C_{gc} + C_{cs}} - (V_h' - v_{th}' + V_p) e^{-H/\tau_4} C_{p2} \ . \tag{3-109}$$

In most cases, the approximation $H >> \tau_4$ is valid and this expression can be simplified to

$$Q_o = Q_s \frac{C_{gc}}{C_{gc} + C_{cs}} \ . \tag{3-110}$$

The output packet depends on capacitances of the CCD well, but does not depend on parasitics on the floating gate node.

Charge-to-Voltage Configuration

The nondestructive sensing circuit in Figure 3-30 produces an output charge that is a replica of its input packet. A modified version of this circuit, shown in Figure 3-31, can be used to produce an output voltage instead. In this configuration, the final well G7 that was otherwise used to store the result, is replaced by a capacitor, C_3. In this circuit, charge packets described by (3-110) are generated and stored in the wells of G5 and G10. During phases 3 and 4, the output node v_o is precharged to V_c. During phase 1, the result is transferred onto v_o by lowering the potential of G5 and G10. The potential on v_o after all charge has transferred is given by

$$v_o = V_c + \frac{Q_s}{C_3}\frac{C_{gc}}{C_{gc}+C_{cs}} \quad . \tag{3-111}$$

One advantage of DDS charge sensing is that it requires about half the clock voltage of conventional charge sensing. Floating gates in a conventional floating gate amplifier lie at a constant potential, midway between the minimum and maximum clock voltages and CCD clocks must swing about this level by the full range of gates in a simple register. DDS charge sensing has the same voltage requirements as a simple CCD register because its floating gates are clocked to the same levels as other CCD gates.

Two other advantages of DDS charge sensing are that it provides greater storage capacity in its floating gate wells and permits a larger output voltage swing. As charge is sensed in a conventional floating gate amplifier, the floating gate voltage falls, and charge storage capacity is reduced. If storage capacity is exceeded, charge flows out from under the floating gate and the sensed output is not a correct indicator of signal charge. The source of this problem is that charge-to-voltage translation occurs on the floating gate. A DDS charge sensing circuit has a larger signal range. Signals are converted to voltages on the output v_o, rather than within the CCD register. Storage capacity is also improved because the floating gate is a virtual ground and output swing does not reduce storage capacity.

Linearity

The result of a DDS sensing circuit, given by (3-111), does not depend on C_{p2} because the floating gate potential is identical at the beginning and end of sensing. It does, however, depend on C_{gc} and C_{cs} and voltage dependence of these capacitors causes nonlinearity. These capacitors are represented by the polynomials

$$\begin{aligned} C_{cs}'(Q_s) &= b_6\left(\frac{1}{C_{cs}+C_{gc}}\right)^2 Q_s^2 + a_6\left(\frac{1}{C_{cs}+C_{gc}}\right)Q_s + C_{cs} \\ C_{gc}'(Q_s) &= b_7\left(\frac{1}{C_{cs}+C_{gc}}\right)^2 Q_s^2 + a_7\left(\frac{1}{C_{cs}+C_{gc}}\right)Q_s + C_{gc} \end{aligned} \tag{3-112}$$

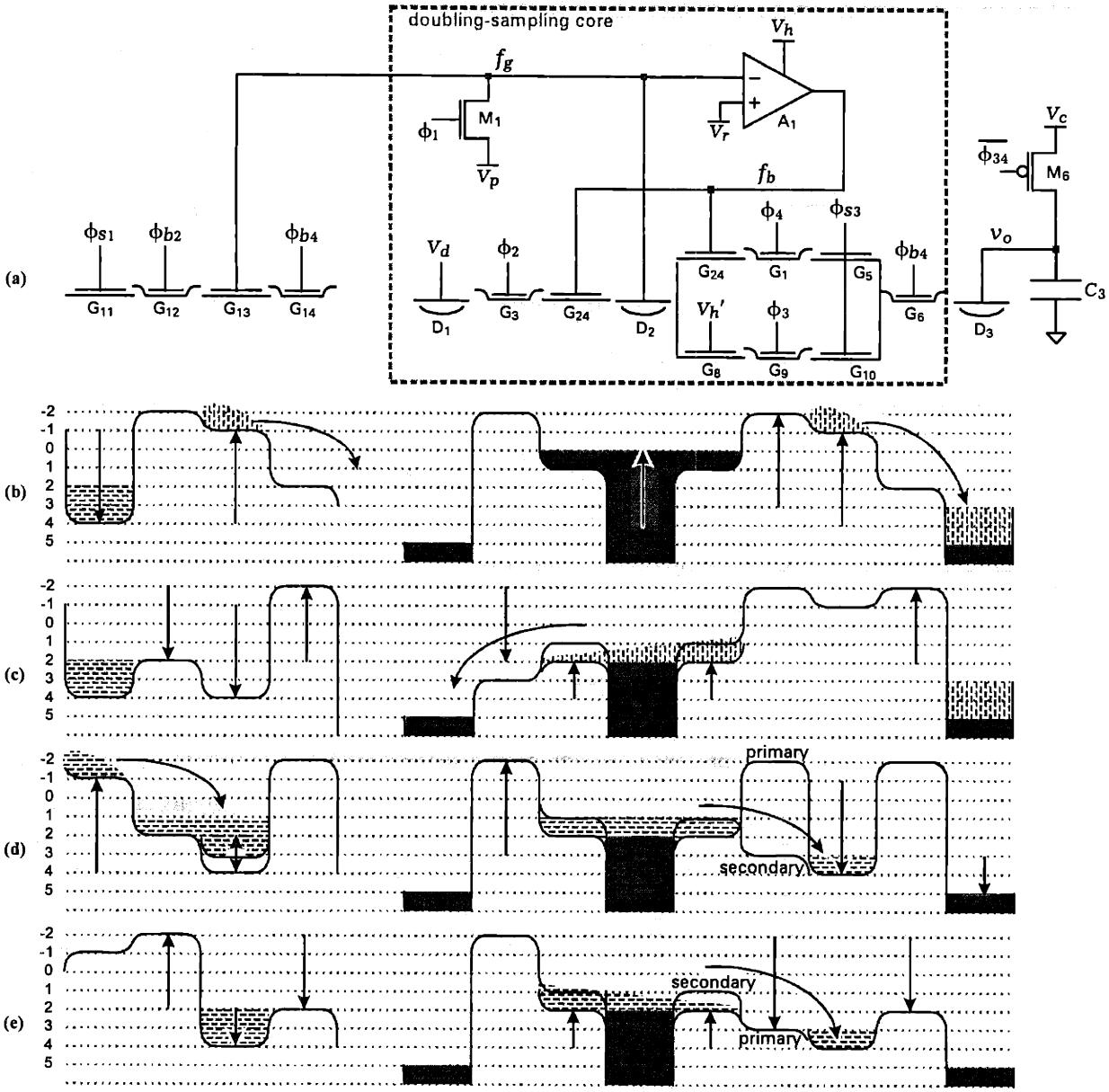


Figure 3-31. DDS charge sensing circuit with voltage output. (a) Schematic. (b) Fill phase. (c) Spill phase. (d) Collection phase. (e) Sensing phase.

The values C_{gc} and C_{cs} represent nominal capacitance. When voltage dependence is small compared to nominal values, the output is given by

$$v_o \approx V_c + \frac{Q_s C_{gc}}{C_3\left(C_{gc}+C_{cs}\right)}\left(1 - Q_s\left(\frac{a_6 C_{gc} - a_7 C_{cs}}{\left(C_{gc}+C_{cs}\right)^2 C_{gc}}\right) - {Q_s}^2\left(\frac{b_6 C_{gc} - b_7 C_{cs}}{\left(C_{gc}+C_{cs}\right)^3 C_{gc}}\right)\right) . \tag{3-113}$$

The ratio between 2nd harmonic and fundamental is

$$\frac{v_o(2f_0)}{v_o(f_0)} \approx \left(\frac{a_6 C_{gc} - a_7 C_{cs}}{2\left(C_{gc}+C_{cs}\right)^2 C_{gc}}\right) Q_s . \tag{3-114}$$

That between 3rd harmonic and fundamental is

$$\frac{v_o(3f_0)}{v_o(f_0)} \approx \left(\frac{b_6 C_{gc} - b_7 C_{cs}}{4\left(C_{gc}+C_{cs}\right)^3 C_{gc}}\right) {Q_s}^2 . \tag{3-115}$$

As an example, consider a case with a C_{gc} to C_{cs} ratio of 10, a CCD storage voltage of 1V, and equal voltage coefficients a_6 through b_7. The results of (3-114) and (3-115) are plotted as a function of voltage coefficient a_6/C_{gc} in Figure 3-32.

Thermal Noise

Two sources of thermal noise are present in a DDS charge sensing circuit. The first is due to the transfer of charge from f_g through G24 during the spill and sensing phases. Thermal noise in this process alters the charge packet Q_o that is integrated in the DDS receiving well. If this noise component is represented as a charge quantity n_g, then (3-110) becomes

$$Q_o = Q_s \frac{C_{gc}}{C_{gc}+C_{cs}} + \left(n_g(4H) - n_g(2H)\right) . \tag{3-116}$$

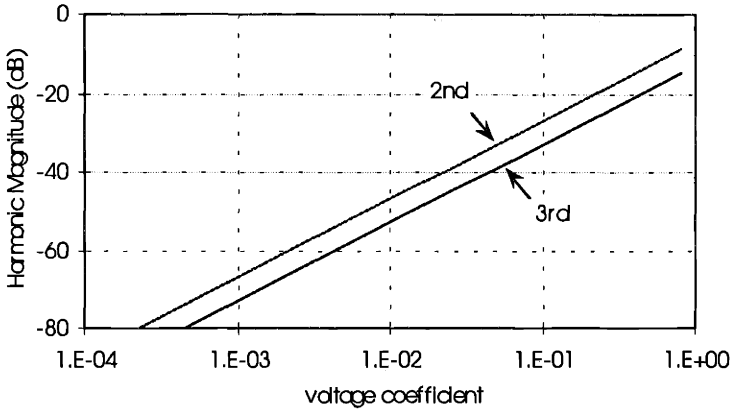


Figure 3-32. Linearity of DDS charge sensing as a function of voltage coefficient.

The mean square value of n_g during precharge is

$$\overline{n_g(2H)^2} = kT\left(\frac{C_{gc}C_{cs}}{(C_{cs}+C_{gc})} + C_{p2}\right) \tag{3-117}$$

and that during sensing is

$$\overline{n_g(4H)^2} = kT\left((C_{cs}+C_{gc})FK_{eff} + C_{p2}\right) . \tag{3-118}$$

Thermal noise during precharge is less than that during sensing because the well contains no charge and has a lower effective capacitance. Total noise in the replicated packet is

$$\overline{Q_o^2} = kT\left(\frac{C_{gc}C_{cs}}{(C_{cs}+C_{gc})} + (C_{cs}+C_{gc})FK_{eff} + 2C_{p2}\right) \tag{3-119}$$

and the resulting SNR is

$$\text{SNR} = \frac{Q_s}{\sqrt{kT\left(\frac{C_{gc}C_{cs}}{(C_{cs}+C_{gc})} + (C_{cs}+C_{gc})FK_{eff} + 2C_{p2}\right)}} . \tag{3-120}$$

SNR improves when parasitic capacitances, C_{cs} and C_{p2}, on the floating gate node are kept small. However, the dominant capacitance is usually $n_g(4\text{H})$, which depends primarily on the total signal charge.

In a charge replicator, (3-119) describes the total circuit noise. However, when a charge replicator is used in a charge-to-voltage configuration, additional noise is introduced when the output capacitor C_3 is precharged. This noise component is represented by a voltage quantity n_o with a mean square value of

$$\overline{n_o(4H)^2} = kT\frac{1}{C_3} . \tag{3-121}$$

Combined noise in the result of (3-111) then becomes

$$\overline{v_o^2} = kT\frac{\left(\frac{C_{gc}C_{cs}}{(C_{cs}+C_{gc})} + (C_{cs}+C_{gc})FK_{eff} + 2C_{p2}\right) + C_3}{C_3^2} \tag{3-122}$$

and the resulting SNR is

$$\text{SNR} = v_o\sqrt{\frac{C_3^2}{kT\left(\frac{C_{gc}C_{cs}}{(C_{cs}+C_{gc})} + (C_{cs}+C_{gc})FK_{eff} + 2C_{p2}\right) + kTC_3}} . \tag{3-123}$$

SNR is improved by reducing parasitic capacitances on the floating gate node and increasing capacitance on the output sensing node.

3.6.3 Dynamic Double-Sampling D/A Conversion and Subtraction

An operation that is commonly required in A/D conversion is subtraction of the output from a multiplying D/A from an incoming analog signal. DDS circuit techniques can be used to perform this function. In a

DDS approach, an input quantity of positive charge, generated through capacitor elements, is combined on a wire with an incoming packet of negative charge, from a CCD register. The D/A full scale can be either a time-varying voltage or a constant reference voltage. Since diffusions are necessary on the wire, the operation is nondepleted and subject to thermal noise, coupling, and clock feedthrough. However, DDS subtraction has the advantage that its positive charge is formed through polysilicon capacitors. These capacitors provide good element-to-element matching and low voltage dependence, which is important if the D/A full-scale is a time varying signal. DDS subtraction also has the advantage that its incoming negative signal is not subject to nonlinearities due to capacitor voltage dependence.

A circuit for performing this function is shown in Figure 3-33. A CCD register, formed by gates G11, G12, and G13, serves as the source of electrons. Its output is connected to the input of a DDS core. N-bit D/A conversion is performed by an array of $(2^N\text{-}1)$ identically sized units, each one consisting of a clamp-and-sample circuit in series with a capacitor. The D/A elements provide positive charge to the input of the DDS core. The resulting combination of negative and positive charge is integrated in receiving well G5. The DDS core of this circuit differs from that used for charge generation in that the secondary sensing path is eliminated.

The operation of this circuit is similar to that of the DDS wire transfer of section 3.5.1. During the fill phase, f_g is pulled low to V_p to provide a new supply of electrons to the circuit. Gate G12 is off at this time to prevent charge from flooding backward into the source register. Meanwhile the array of signals $v_f(i)$, within the D/A, is forced low by the precharge through M3.

The D/A precharge remains on during the spill phase. Node f_g is reset to a level of

$$f_g(2H) = V_p\, e^{-H/\tau_2} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-H/\tau_2}\right) . \tag{3-124}$$

The time constant for this transition is

$$\tau_2 = \frac{R_{24}\left(\left(2^N - 1\right)C_1 + C_{p2}\right)}{(A+1)} . \tag{3-125}$$

Capacitor C_{p2} represents parasitic junction and routing capacitances on f_g as well as channel capacitance of G24 and G13.

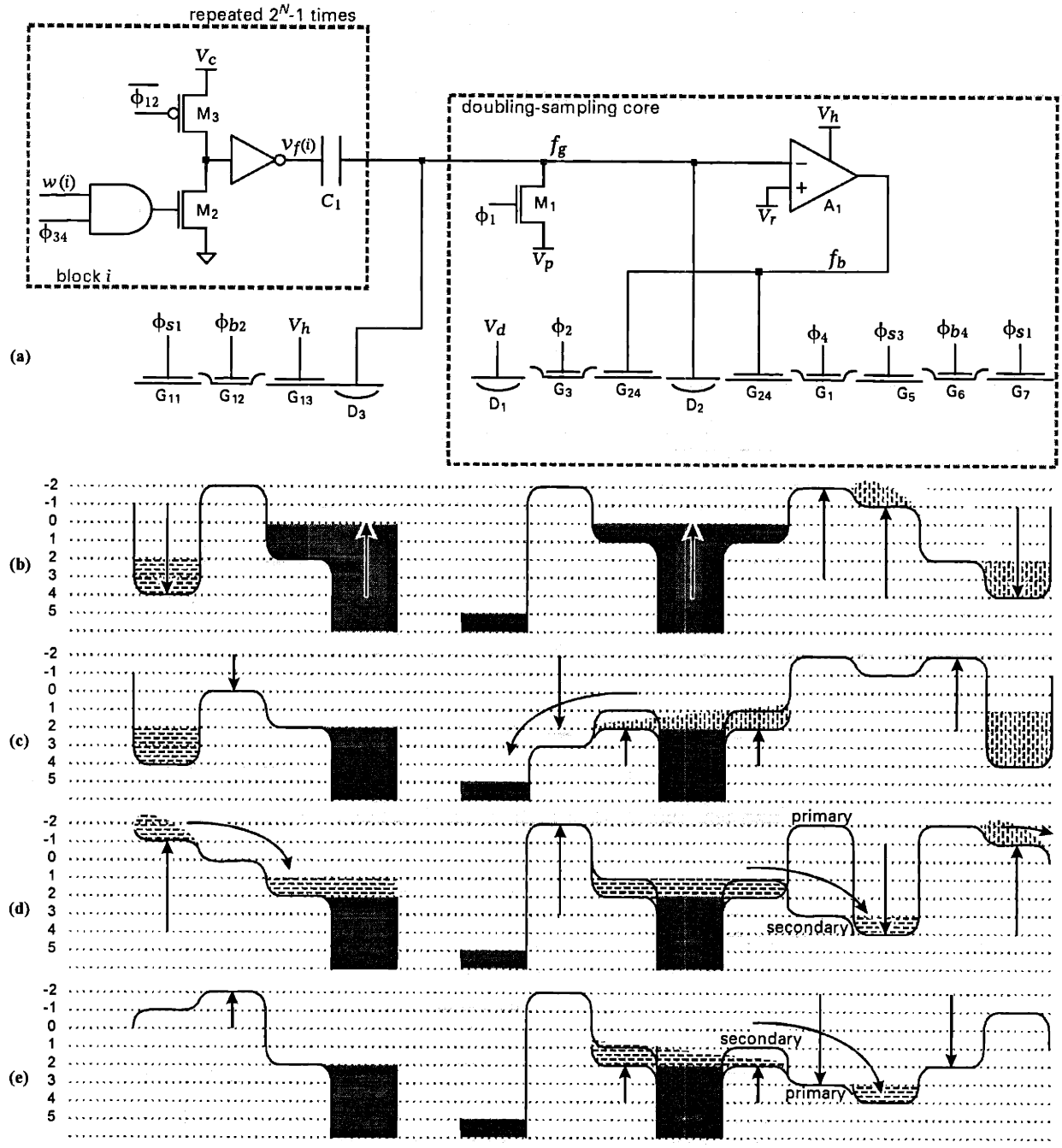


Figure 3-33. DDS D/A and subtraction circuit. (a) Schematic. (b) Fill phase. (c) Spill phase. (d) Collection phase. (e) Sensing phase.

During the collection phase, a quantity of charge Q_s is transferred onto f_g by lowering the potential of G11. At the same time, the digital inputs $w(i)$, which are in a thermometer code format, are used to selectively assert some of the signals $v_f(i)$. When $w(i)$ is high, a positive transition of V_c is introduced across $C_1(i)$. When $w(i)$ is low, the capacitor voltage remains unchanged. Electrons, from the source register, are cancelled by positive displacement charge, and f_g is restored to a higher voltage. The resulting potential on f_g from the combination of these competing effects is

$$f_g(3H) = f_g(2H) + \frac{Q_s + \sum_{i=0}^{2^N-1} w(i) C_1 V_c}{\left(\left(2^N - 1\right) C_1 + C_{p2}\right)} \quad . \tag{3-126}$$

During the sensing phase, the net charge on f_g flows through G24 to the output well, and the potential on f_g rises to its final value of

$$f_g(4H) = f_g(3H) e^{-H/\tau_4} + \left(\frac{AV_r - v_{th}}{A+1}\right)\left(1 - e^{-H/\tau_4}\right) \quad . \tag{3-127}$$

The time constant for the sensing transition, τ_4, equals that for the spill transition, τ_2.

Combining (3-124), (3-126), and (3-127), and eliminating higher-order exponential terms, yields the resultant charge packet

$$Q_o = \left(Q_s + \sum_{i=0}^{2^N-1} w(i) C_1 V_c\right)\left(1 - e^{-H/\tau_4}\right) - \left(\frac{AV_r - v_{th}}{A+1} - V_p\right) e^{-H/\tau_4}\left(\left(2^N - 1\right) C_1 + C_{p2}\right) \quad . \tag{3-128}$$

The approximation $H >> \tau_4$ is nearly always valid for this circuit and the first term, which represents a constant offset, is negligible. The result is then identically equal to the difference between the input charge Q_s and the D/A displacement charge, regardless of parasitics on the wire transfer node, f_g.

Thermal Noise

The noise response of a DDS D/A conversion and subtraction circuit is given by

$$Q_o = \left(Q_s + \sum_{i=0}^{2^N-1} w(i) C_1 V_c\right) + \left(n_g(4H) - n_g(2H)\right)\left(\left(2^N - 1\right) C_1 + C_{p2}\right) \quad , \tag{3-129}$$

where n_g is the noise quantity defined in section 3.5.4. The result depends only on the difference in noise values at the end of phases 2 and 4. The n_g terms in (3-129) have mean-square values of

$$\overline{n_g(2H)^2} = \overline{n_g(4H)^2} = kT \frac{1}{\left(\left(2^N - 1\right) C_1 + C_{p2}\right)} \quad . \tag{3-130}$$

Unlike a charge generator, capacitance and thermal noise in this circuit are the same during phases 2 and 4. Under the assumption that these components are independent, the combined mean-square noise is

$$\overline{Q_o{}^2} = 2kT\left(\left(2^N - 1\right) C_1 + C_{p2}\right) \quad . \tag{3-131}$$

Thermal noise in the output signal decreases with decreasing capacitance. But signal levels decrease with decreasing capacitance as well. When $C_{p2} << (2^N - 1)C_1$, the resulting SNR for a sinusoidal input with rms amplitude V_S is approximately

$$\text{SNR} = V_c \sqrt{\frac{(2^N - 1)C_1}{2kT}} \quad . \tag{3-132}$$

3.6.4 Double-Sampling Charge Comparator

Comparison is a central operation needed for A/D conversion. In most cases, signals are compared nondestructively so that the original charge packets are not altered. One method of performing this function uses a floating gate amplifier, such as that described in section 3.3.2, to produce an output voltage [62]. These voltages are then compared by a CMOS comparator. Two factors limit the resolution that is achievable from this technique. First, floating gate amplifiers have a limited range of signal swing. For a given comparator input offset error, resolution is fixed by the ratio between signal swing range and input offset error. This limit is independent of sensing capacitances or charge packet sizes. A second factor that limits resolution is that floating gate amplifiers have poor capacitance and gain matching and their mismatches are amplified by the common-mode charge signal.

The accuracy of charge comparisons can be improved by applying a pair of DDS sensing circuits, such as those in section 3.6.2, to implement comparison. In this approach, two differential input packets are each replicated, using a DDS replicator circuit. The resulting packets are stored in CCD receiving wells and are translated to voltages at the input to a CMOS comparator during the following phase. This circuit can also incorporate differential amplification and common-mode rejection in the charge domain using the feedback techniques described in [66] and [67].

The resolution of this technique is greater than that for a configuration that performs charge-to-voltage translation on the floating gates. First, it supports a larger signal range so that the impact of comparator or amplifier input-offset voltages is reduced. Second, matching in the charge-to-voltage translation is improved because gains are determined by polysilicon capacitors and are only weakly dependent on CCD well capacitances.

3.7 Mirrored Charge Comparator

Applications, such as Δ-Σ modulation, do not require high comparator accuracy. In such cases, the high linearity and mismatch tolerance of a DDS comparator is unnecessary and a simpler technique referred to as mirrored charge comparison can be used. An example of a mirrored charge comparator is shown in Figure 3-34. Elements in the positive and negative halves of this circuit are denoted by subscripts '+' and '-'. A pair of input registers is formed from gates G_1 through G_4. On each side, charge is sensed on the floating gate labeled G_3 and mirrored through the charge mirror formed from gates M_1 and M_2. The output of the charge mirror is connected to the input of a simple CMOS latching comparator formed from transistors M_5 through M_{10}.

The operation of this circuit proceeds as follows. During phase 1, each floating gate is pulled low through M_{13}. This forces charge, underneath it from a previous cycle, forward. Since the timing of this transition must be aligned with that of other CCD gates, it is accomplished by connected the floating gates to the CCD clock ϕ_{S3}, which is low at this time. Also during phase 1, the comparator outputs v_o are pulled to ground.

During phase 2, the pull-up path through M_{11} is enabled. In the analysis below, M_{1+} and M_{1-} are modeled as resistors of value R_1 in series with voltage sources of value v_{th1}, where v_{th1} represents the threshold voltage of these gates. Initially, the voltage on f_g is much lower than that on V_d and M_1 is in saturation. The potential on f_g rises until it approaches V_d+v_{th1}, at which point M_1 is in subthreshold and slowly turns off. The resulting transition is similar to the spill transition in an SDS charge generator, except that sensing occurs through the drain of a PMOS rather than the source of an NMOS.

At the end of phase 2, the floating gate lies at a precharged level of

$$f_{g\pm}(2H) = (V_d + v_{th1})\left(1 - e^{-H/\tau_2}\right) \quad . \tag{3-133}$$

This high voltage allows them to receive charge during the following phase. The time constant for this transition is given by

$$\tau_2 = R_1\left(C_{p2} + \frac{C_{gc}C_{cs}}{\left(C_{gc} + C_{cs}\right)}\right) \quad , \tag{3-134}$$

where C_{p2} represents parasitic capacitance on f_g and gate capacitance of M_1 and M_2, and the second capacitance term is the effective floating gate capacitance when wells are empty.

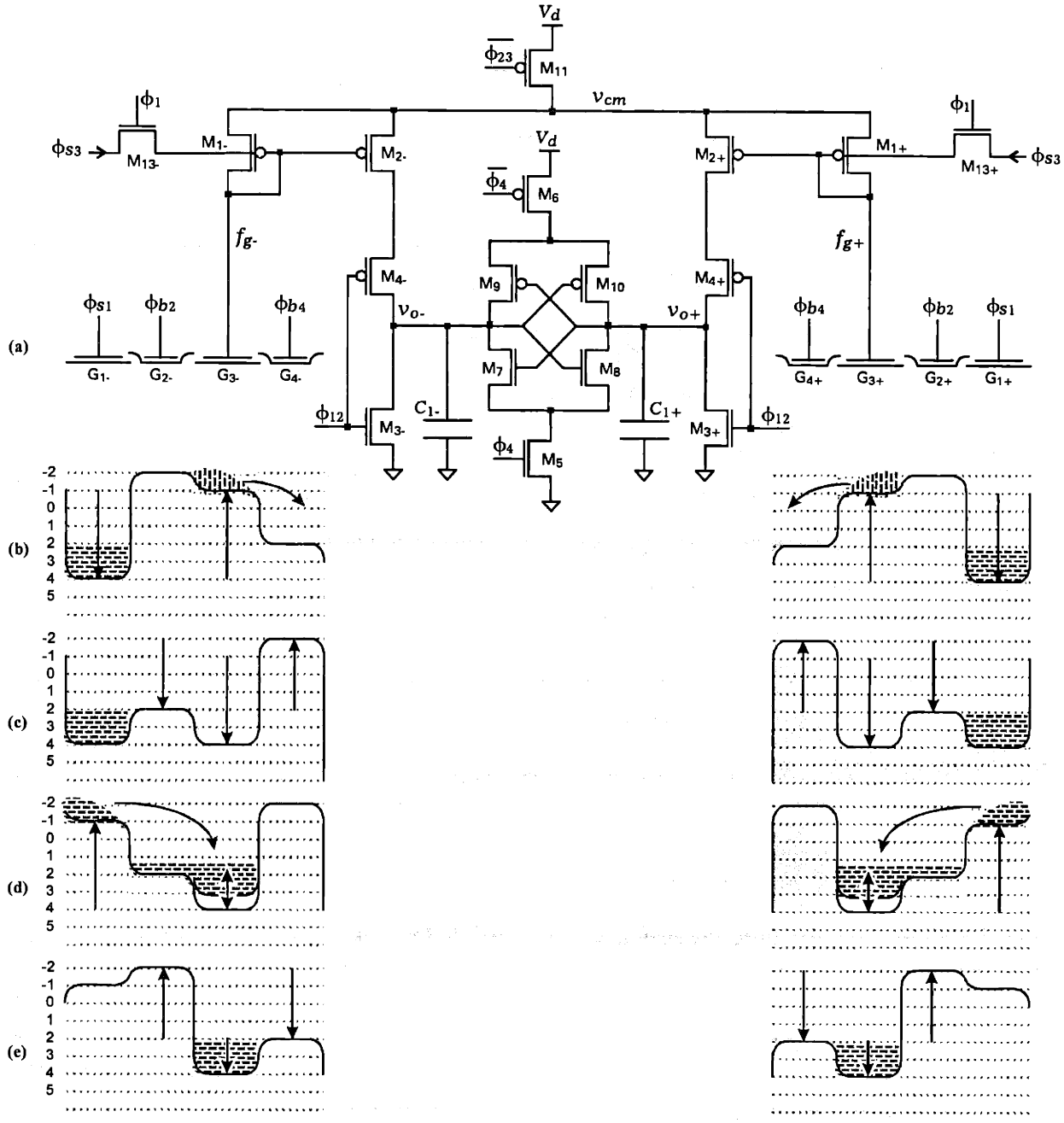


Figure 3-34.Mirrored nondestructive sensing circuit. (a) Schematic. (b) Fill phase. (c) Spill phase. (d) Collection phase. (e) Sensing phase.

During phase 3, charge is transferred underneath the floating gate upon a low transition of the ϕ_{S1} clock. The floating gate voltage falls in response. If this transition is assumed to complete within a short time Δ then the signal on f_g is

$$f_{g\pm}(2H+\Delta) = (V_d + v_{th1})\left(1 - e^{-H/\tau_2}\right) + Q_{s\pm}\frac{C_{gc}}{C_{gc}C_{p2} + C_{cs}\left(C_{gc} + C_{p2}\right)} \quad . \tag{3-135}$$

As the floating gate falls, M1 is turned on again and has a current of

$$I_{1\pm}(t) = \frac{1}{R_1}\left(V_d + v_{th1} - f_{g\pm}(t)\right) \quad . \tag{3-136}$$

The current in (3-136) is mirrored through M2, with a gain that is proportional to the ratio of M1 and M2 gate geometries. Mirrored current at the input to the comparator is

$$I_{2\pm}(t) = \frac{1}{R_2}\left(V_d + v_{th2} - f_{g\pm}(t)\right) \quad , \tag{3-137}$$

where v_{th2} and R_2 represent the threshold voltage and resistance of M2.

The floating gate potential rises in response to this current, with a transition,

$$f_{g\pm}(t) = f_{g\pm}(2H+\Delta)e^{-(t-2H)/\tau_3} + (V_d - v_{th1})\left(1 - e^{-(t-2H)/\tau_3}\right) \quad , \tag{3-138}$$

that is similar to that during precharge. As f_g approaches a voltage of V_d+v_{th1}, M1 approaches subthreshold and turns off. The time constant for sensing is given by

$$\tau_3 = R_1\left(C_{p2} + \left(C_{gc} + C_{cs}\right)FK_{eff}\right) \quad , \tag{3-139}$$

where C_{p2} represents parasitic capacitance on f_g and gate capacitance of M1 and M2, $\left(C_{gc} + C_{cs}\right)$ is the floating gate storage capacity, and FK_{eff} is the effective capacitance factor described in section 3.8.

The comparator input v_o rises in response to I_2. Ignoring higher order exponential terms, the final value on this node is described by

$$v_{o\pm}(3H) = Q_{s\pm}\frac{1}{C_1}\left(\frac{R_1}{R_2}\right)\left(\frac{C_{gc}}{C_{gc} + C_{cs}}\right)\left(1 - e^{-H/\tau_2}\right) + (v_{th1} - v_{th2})\frac{H}{R_2 C_1} \quad . \tag{3-140}$$

As a final step, the flip-flop, consisting of M7 through M10, is enabled during phase 4 and the comparator's digital result is latched.

The first term in (3-140) is proportional to the input charge and a scale factor that depends on the amount of time allowed for sensing. At high speed, signal gain is reduced. However, a moderate gain reduction is usually tolerable in this circuit because the end result is comparison. At lower speed, when the approximation $H >> \tau_2$ is valid and thresholds v_{th1} and v_{th2} are equal, inputs to the comparator immediately before its result is latched are

$$v_{o\pm}(3H) = Q_{s\pm}\frac{1}{C_1}\left(\frac{R_1}{R_2}\right)\left(\frac{C_{gc}}{C_{gc} + C_{cs}}\right) \quad . \tag{3-141}$$

Resolution

The resolution of a mirrored charge comparator is determined by three factors. First, threshold mismatches between the current mirror gates M_{1+} and M_{2+} or between M_{1-} and M_{2-} cause charge, reflected at the comparator input, to differ from charge removed from the floating gate. Second, transistor imbalance between the positive and negative sides of the flip-flop creates an input-offset error. Third, capacitive mismatches on the comparator input nodes result in voltages at the flip-flop input that do not accurately reflect the differential charge signal. These error sources are described below.

Threshold mismatches between M_{1+} and M_{1-} are not critical because the floating gates are autozeroed with respect to their values. However, threshold mismatches between M_{1+} and M_{2+} or between M_{1-} and M_{2-} are critical. The error voltage that is introduced at the comparator input by such a mismatch is given by the second term in (3-140). This value can be translated into a charge-referred error of

$$\Delta Q_{o\pm}(3H) = (v_{th1} - v_{th2})\frac{H}{R_2} \ , \tag{3-142}$$

where the time for each clock phase is denoted by H. This error can also be expressed as

$$\Delta Q_{o\pm}(3H) = (v_{th1} - v_{th2})Z C_2 \ , \tag{3-143}$$

where Z is defined as the number of time constants per phase and is given by

$$Z = H/R_2C_1 \ . \tag{3-144}$$

The impact of threshold mismatches depends on the number of time constants per clock phase. At higher frequencies, a smaller charge-referred error is accumulated at the comparator input. To minimize the impact of threshold mismatch error and cancel their frequency dependence, a current source of value

$$I = C_1\frac{V_d}{H} \tag{3-145}$$

can be placed in series with M_{11} to limit the sensing current.

Comparator errors are also introduced as a result of capacitive imbalances on v_{o+} and v_{o-}. In this case, identical input charges produce different output voltages. The size of this error with respect to nominal output voltage equals

$$\frac{v_{o+}(3H) - v_{o-}(3H)}{v_o(3H)} = \left(\frac{C_{1-} - C_{1+}}{C_1}\right) \ , \tag{3-146}$$

where capacitances on v_{o+} and v_{o-} are denoted C_{1+} and C_{1-} and nominal capacitance is denoted C_1.

3.8 CCD Power Dissipation

Nearly all CCD power dissipation occurs in the gate driver when it is switched. Two factors contribute to power dissipation. The first is MOS deep depletion capacitance between a CCD gate and the substrate. The second is capacitance between a CCD gate and charge in its channel. To estimate power dissipation, these two capacitive components are combined into a measure of effective capacitance, C_{eff}, to ground seen by the CCD driver. Relationships between well capacitances and effective capacitance are derived below.

A simple model of a CCD well is shown in Figure 3-35. Its gate-to-channel and channel-to-substrate capacitances are represented by C_{gc} and C_{cs}. Charge Q_s that enters the CCD channel is represented by a current source. The gate and channel potentials are denoted by V_g and v_c. Effective capacitance is defined as

$$C_{eff} = I_g \frac{dt}{dV_g} = C_{gc} \frac{dV_g - dv_c}{dV_g} . \tag{3-147}$$

The relationship between gate voltage, channel charge, and channel potential is

$$dv_c = dV_g \frac{C_{gc}}{C_{gc} + C_{cs}} + Q_s \frac{1}{C_{gc} + C_{cs}} . \tag{3-148}$$

The first term represents depletion capacitance and is independent of signal charge. The second term depends on signal charge and well storage capacity and is a measure of how full the well is filled.

Equation (3-148) is substituted into (3-147) to arrive at the relationship

$$C_{eff} = \frac{C_{gc}C_{cs} - C_{gc}\left(Q_s/V_g\right)}{C_{gc} + C_{cs}} \tag{3-149}$$

for effective capacitance as a function of channel charge. Storage capacity of a CCD well equals

$$C_s = \left(C_{gc} + C_{cs}\right) . \tag{3-150}$$

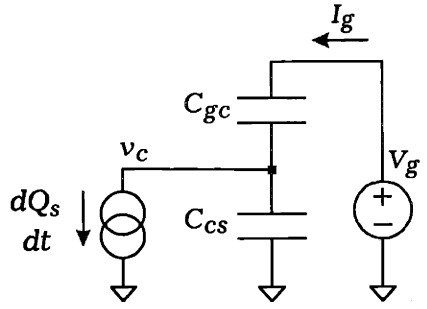


Figure 3-35. Model of a CCD well used to derive effective gate capacitance.

Storage capacity is set by the number of signal electrons per well and the voltage range, V_w, across which these electrons are stored. The largest charge packet that can be stored in a channel is

$$Q_s = -C_s V_w \,, \tag{3-151}$$

Equations (3-151) and (3-149) are then combined to produce an expression for effective capacitance,

$$C_{eff} = \frac{C_{gc}C_{cs}\left(1 + V_w/V_g\right) + C_{gc}{}^2\left(V_w/V_g\right)}{C_{gc} + C_{cs}} \,, \tag{3-152}$$

that depends only on well capacitance and the percentage of the well that is filled with charge. As channel charge approaches zero, so does V_w, and the effective capacitance approaches that of an MOS gate in deep depletion. As charge is added and V_w occupies nearly the full well, C_{eff} approaches the gate-to-channel capacitance.

The fraction to which a CCD well is filled with charge is given by $\left(V_w/V_g\right)$. Since this quantity is a function of signal level, it is convenient to define an average value, denoted by F. In most signal processing CCD devices, circuits are implemented differentially, with charge storage ranging from 0 to a maximum of V_w. In this case, the average value to which wells are filled,

$$F = \left(\frac{\overline{V_w}}{2V_g}\right) \quad \text{for } V_w > 0 \,, \tag{3-153}$$

is determined by the zero-level reference, $V_w/2$. Because of constraints described in section 5.4.1, F is typically chosen between 0.2 and 0.5.

Variable R is defined as

$$R = \left(\frac{C_{gc}}{C_{cs}}\right) \,, \tag{3-154}$$

which is a measure of the relative gate-to-channel and channel-to-substrate well capacitances. For a given process technology, R is fixed and has a typical range between 5 and 15.

Using equations (3-153) and (3-154), storage capacity is translated into a measure of effective capacitance by the relation

$$\frac{C_{eff}}{C_s} = FK_{eff} \,, \tag{3-155}$$

where

$$K_{eff} = \frac{R}{(R+1)}\left(1 + \frac{1}{F(R+1)}\right) \,. \tag{3-156}$$

A plot of the effective capacitance factor K_{eff} is shown in Figure 3-36 as a function of capacitance ratio R for $F=0.5$. In almost all processes, R is much larger than 1, $Keff$ has a value near unity, and $Keff$ depends only weakly on process parameters.

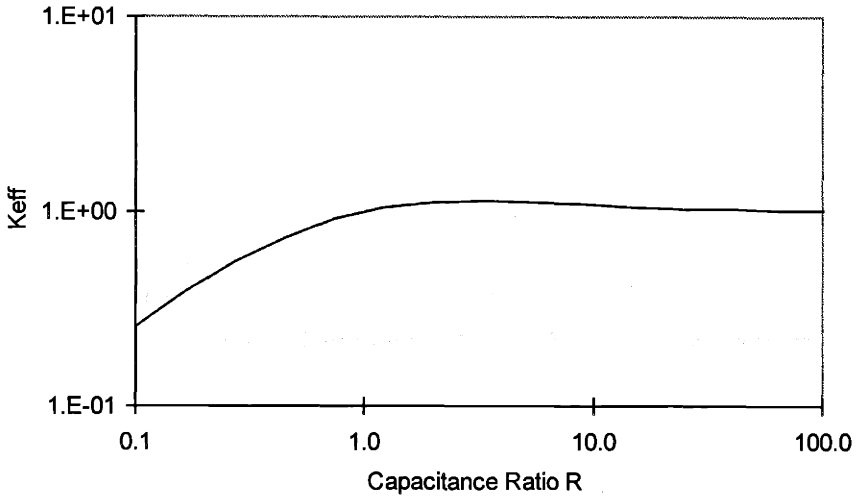


Figure 3-36. Effective capacitance factor versus capacitance ratio R. K_{eff} is typically near 1 and largely process independent.

This relationship between capacitance and signal charge is used to arrive at an expression for average power. Sinusoidal inputs are assumed. CCD power is dissipated predominantly in the gate driver and increases linearly with frequency. Expressed per unit frequency per well, the average power is

$$\frac{P}{f} = C_s F K_{eff} V_g^2 . \tag{3-157}$$

Equation (3-155) shows that capacitance, seen by the CCD gate driver, depends on the storage capacity of the well and how full the well is. Expressed differently, effective capacitance depends primarily on the total quantity of charge stored in the well. Since power is proportional to effective capacitance, it is determined by charge packet sizes as well.

The area required per CCD well follows from storage capacity and gate-to-channel capacitance per unit area.

$$A = \frac{C_s}{C_a}\left(\frac{R}{R+1}\right) . \tag{3-158}$$

The factor (C_{gc}/A) in this expression represents the gate-to-channel capacitance per unit area and for surface channel devices is nearly equal to the gate oxide capacitance.

4 Pipelined Oversampling Converter Implementations

4.1 Example Converter Designs

The architectural considerations, described in Chapter 2, and the circuit considerations, described in Chapter 3, are merged into a description of four example POSC designs, referred to as designs 1 through 4. The differences between these implementations are outlined in Table 4-1. The first two are based on an analog integration architecture, while the last two use a digital integration approach. Two performance objectives are considered for each of these architectures, 72-dB and 84-dB SNR. In both of these cases, quantization and thermal noise components are set equal, to a level 3-dB below the overall converter specification.

First, section 4.1.1 outlines a choice of architectural design parameters for the designs, including pipeline length, modulator order, decimator order, and number of ADC and DAC bits. The discussion applies identically to both analog and digital integration approaches. Next, section 4.2.1 presents implementation details specific to analog integration designs 1 and 2, including pipeline structure, circuit components, and device sizes. Section 4.3 describes implementation details specific to digital integration designs 3 and 4. The POSC prototypes, which are similar to design 3, are presented in section 4.4. Finally, measured performance for these devices is described in section 4.5.

4.1.1 Architectural Parameters

Targeted signal to thermal noise ratio is the starting point for selecting architectural parameters. Choices are constrained by the circuit building blocks that are available and by the accuracy that can practically be realized from them. Other factors such as converter hardware, power, pipeline length, and area are also considered in the selection.

	Design 1	Design 2	Design 3	Design 4
Architecture	Analog Integration	Analog Integration	Digital Integration	Digital Integration
SNR	72 dB	84 dB	72 dB	84 dB
SNR (quantization)	75 dB	87 dB	75 dB	87 dB
SNR (thermal)	75 dB	87 dB	75 dB	87 dB
SFDR	72 dB	84 dB	72 dB	84 dB

Table 4-1. Architecture and design objectives of four example POSC designs. Quantization and thermal noise components are set equal.

Modulator Order

Device requirements for 1st, 2nd, and 3rd-order modulators with the SNR targets described above are listed in Table 4-2. These are determined from the SNR relationships described in Chapter 2. A 1-bit ADC and 1-bit DAC are assumed for all entries. The number of conversion blocks determines the degree of Δ-Σ noise shaping. Converter latency, or number of pipeline delays, is proportional to the number of conversion blocks and the number of substages per conversion block. An increase in latency is undesirable because it increases the accuracy required from signal transfers along this pipeline length. In most cases it also translates into increased hardware and power. As modulator order is increased, the number of conversion blocks decreases but the number of substages per conversion block and the number of circuit elements per substage increases. These factors determine hardware and power.

A 2nd-order modulator is chosen for the present designs. In comparison to 1st-order, 2nd-order is clearly preferable. Not only does it provide a substantial reduction in pipeline length, hardware, power, and area, but it is also less susceptible to pattern noise. In comparison to 3rd-order, a 2nd-order configuration does bring a moderate increase in hardware, power, and area and a longer pipeline length. But despite this, 2nd-order is preferred because it is more easily implemented with available circuit building blocks and is not subject to stability concerns.

Number of ADC Bits

Some of the resolution benefits lost by use of 2nd rather than 3rd-order modulation are more easily regained by use of a multi-bit quantizer. An increase in ADC bits decreases the number of conversion blocks but increases the number of substages and the amount of hardware per block. The combination of these competing effects is summarized in Table 4-3 for 3, 5, and 7-bit ADCs. Two-bit feedback is assumed for all entries.

In general, adding quantizer bits does reduce hardware, power, area, and pipeline length for a given performance target. However, the cost of this is a reduced degree of noise shaping, which increases the SNR and linearity requirements of elements within the feedback loop. Two factors determine the number of ADC bits that should be used. The first is the degree of noise shaping reduction that is tolerable. In

	1st-order	2nd-order	3rd-order
Number of conversion blocks	448 / 1131	59 / 104	26 / 40
Number of pipeline delays	897 / 2262	178 / 313	105 / 159
Number of hardware units	4036 / 10179	1011 / 1775	706 / 1073

Table 4-2. Architecture specifications versus modulator order for 75-dB / 87-dB SNR. Second-order modulation is chosen for the example designs.

most cases this is determined by the characteristics of elements within the outer feedback loop, where errors experience only 1st-order noise shaping. A second factor is the minimum resolution and maximum dynamic range of the quantizer. Minimum resolution is determined by the accuracy of comparators and binary scaled feedback elements within the quantizer. Maximum dynamic range is limited by common-mode accumulation and limited supply voltage. If ADC bits are added beyond this point, hardware increases but resolution does not.

Designs 1 and 3 use 5-bit quantizers based on their minimum resolvable charge packets. Design 2 is limited to a 3-bit ADC because of its need to suppress circuit nonlinearity. Design 4 does not have this same requirement and, based on its minimum resolvable charge packet, includes a 7-bit ADC.

Number of DAC Bits

Multi-bit feedback is an attractive option for a POSC because feedback matching requirements are the same for single-bit and multi-bit systems. Although hardware per conversion block increases slightly with the number of DAC bits, multi-bit feedback improves the range of inputs that can be processed without saturation, decreases susceptibility to pattern noise, and results in more linear and predictable converter behavior. In comparison to single-bit feedback, 2-bit feedback brings significant improvements in these regards. Only minor additional improvements are achieved from each additional bit. The example designs use 2-bit feedback.

Decimator Order

As decimation filter order is increased, the number of conversion blocks that are needed for a given performance is decreased. Since the number of substages per conversion block is independent of decimator order, pipeline latency is reduced as well. However, higher-order filters do increase the hardware and complexity within each conversion block. The combination of these competing effects is summarized in Table 4-4. A 2nd-order modulator, 5-bit ADC, and 2-bit DAC are assumed for all entries.

	3-bit ADC	5-bit ADC	7-bit ADC
Number of conversion blocks	28 / 49	16 / 28	8 / 16
Number of pipeline delays	135 / 243	109 / 189	56 / 109
Number of hardware units	728 / 1311	545 / 944	280 / 545

Table 4-3. Architecture specifications versus quantizer bits for 75-dB / 87-dB SNR. The example designs use 5, 3, 5, and 7-bit ADCs.

	1st-order	2nd-order	3rd-order	5th-order
Number of conversion blocks	6713 / 26097	97 / 194	59 / 104	53 / 93
Number of hardware units	100697 / 391461	1546 / 3109	1011 / 1775	1009 / 1770

Table 4-4. Architecture specifications versus decimator order for 75-dB / 87-dB SNR. The example designs use 3rd-order error averaging decimation.

The example designs include 3rd-order error averaging decimation. In comparison to decimators of lower order, this choice provides a significant reduction in pipeline latency, hardware, power, and area. In comparison to a 5th-order matched filter decimator, 3rd-order has a slightly larger number of conversion blocks and pipeline latency. However, the hardware requirements of these two approaches are almost identical. A 3rd-order approach is, therefore, favored because of its reduced complexity.

4.1.2 Fabrication Technology

The pipelined oversampling concept described in this chapter is not specific to any technology or to specific circuit techniques. However, its architecture presents challenging circuit requirements that limit the range of technologies that can practically be used.

In a time-oversampling converter, pipeline lengths are limited to a few stages. Since integrator signals remain within the same hardware elements, signal transfer accuracy is not a serious concern and only moderate integrator-to-integrator matching is required. Device area is manageable because the hardware of a single modulator is small. And modulator and decimator circuits are located on separate chips to prevent digital decimation circuits from corrupting the integrity of sensitive analog signals.

In contrast, hundreds of pipeline stages are necessary for a POSC. Signal transfers along this pipeline must be performed with extremely low noise and accuracy greater than that of the overall converter. Integrators within the pipeline must also be matched with nearly full converter accuracy. Area and power for this many pipeline stages is a concern. The modulator and decimator must be located on the same chip because a large number of signals communicate between them. As a result, high-precision analog operations must coexist alongside noisy digital logic.

Because of these requirements, a POSC is not easily implemented using traditional CMOS circuits. Long CMOS pipelines with high accuracy are difficult to achieve because each stage requires sample-and-hold amplifiers, buffers, switches, and holding capacitors. Noise, introduced by switches and other circuit elements in each pipeline stage, compounds along the pipeline as signals transfer forward. Mismatches between the gains of different CMOS integrator circuits are also a concern. In addition, the power and area of a CMOS approach is prohibitive for hundreds of stages. Finally, high precision is not readily

achieved from CMOS circuits in the presence of neighboring digital logic because signals are susceptible to coupling from sources such as the substrate.

Tailoring Architectures to CCD/CMOS

A POSC is most practically accomplished using a combination of CCD and CMOS circuits. The POSC designs described here are based on this approach. A combination of these two circuit techniques enables performance that would be difficult from either one alone. CCDs make pipelines with hundreds of stages feasible by contributing fully depleted operations, such as charge transfer, addition, integration, and conditional transfer, that are highly accurate, low power, simple, and compact. Because these operations are not subject to thermal noise or coupling from clocks or the substrate, high signal integrity is possible throughout hundreds of transfers, amidst noisy digital circuitry. Fully depleted circuits have a gain and linearity that is determined by charge conservation and is not impacted by the device parameters of any circuit elements. In addition, they require only capacitive switching current and can be performed with low power and high speed. A POSC implementation should be tailored to make use of fully depleted CCD circuits whenever possible.

Nondepleted CCD circuits, including charge generation, wire transfer, D/A conversion, and D/A subtraction, are also available. They do not have the same fundamental benefits of depleted transfers. They are subject to thermal noise and coupling, and their linearity is determined by the characteristics of circuit elements. Nonetheless, high linearity and low noise are possible from these circuits and they can be used in circuit paths requiring full converter linearity and SNR.

Other nondepleted CCD circuits, such as charge comparison, division, replication, and sensing, are also available. Like the nondepleted operations described above, they are subject to thermal noise and coupling. However, this set of operations has other inaccuracies as well. The resolution of charge comparison and the accuracy of charge division are limited to about 60 dB by device mismatches. The linearity of charge replication and sensing is also limited to about 60 dB by floating gate elements. These circuits can not be used in paths requiring full converter accuracy. Their use is restricted to the inner feedback loop or the feedforward path of the outer loop, where noise is suppressed by noise shaping and nonlinearities are suppressed by the 1st integrator gain.

CMOS also plays a vital role in the implementation of any CCD-based device. Most CCD circuits are not self-sufficient. They depend on CMOS for analog supporting functions and for digital clocks and controls. It is essential that CCD and CMOS control signals are synchronized because of the tight interaction between their circuits. Since CMOS provides this function on chip, it permits high-speed operation. In addition, CMOS offers functions, such as comparison and subtraction, that are not easily accomplished

using CCDs alone. Finally, CMOS makes large amounts of on-chip digital computation possible for functions such as decimation.

A CCD/CMOS converter is best implemented in a differential configuration for a number of reasons. First, fully depleted charge transfers can only represent signals of one sign. Zero levels are set by an offset of charge and positive and negative values are represented by deviations about this offset. Differential signal flow eliminates the need for accurate control of this or other charge reference levels. Second, charge subtraction, which is a nondepleted operation and subject to thermal noise, can be accomplished instead with very high accuracy as complementary addition, using a fully depleted transfer. Third, even-order harmonics, introduced by circuits such as charge generation and charge sensing, are cancelled when these circuits are implemented differentially. Finally, common-mode noise, such as substrate coupling, that affects nondepleted CCD circuits is suppressed in a differential configuration.

4.1.3 Comparison Between Analog and Digital Integration

The choice between analog and digital integration is based primarily on four circuit and layout issues. First, these architectures differ in their number of feedback DAC bits. Analog integration has only one or two bits, whereas digital integration requires many more. Both architectures have similar matching requirements for their DAC elements. However, digital integration has a practical disadvantage in that the difficulty of implementing DACs with highly accurate level placement increases with the number of bits.

Second, these architectures differ in their ability to make use of truncated feedback. Analog integration requires charge replication circuits with high linearity between its first and second stages of integration. The linearity requirement for these circuits can be alleviated by reducing the number of feedforward ADC bits. However, this increases pipeline length, power, and area. Digital integration does not need charge replication and can make full use of truncated feedback to reduce power and area.

Third, these architectures differ in their process requirements. Analog integration contains two integrators and a replicator within each pipeline stage. The stage pitch is strictly constrained by CCD gate lengths, which are fixed by desired operating frequency. Support circuitry, whose area is determined by CMOS design rules, is more readily fit within this pitch as process geometries are reduced. Accurate charge replication also requires more than two levels of metal, so that sensitive nodes can be shielded from crossover coupling. A digital integration layout is better suited to larger geometries and fewer levels of metal. However, analog integration becomes practical at geometries of 0.35-μm or below and three levels of metal.

A final consideration in the choice between analog and digital integration is the build-up of common-mode charge. Although some common-mode rejection is incorporated into feedback operations, complete common-mode rejection is not possible because charge packet signals have only one sign. In digital

integration, no zero-reference offset is needed in the first integrator because it is digital and the common-mode is attenuated at the input to the second integrator. In analog integration, the first integrator's common-mode signal can not be suppressed digitally by the feedback path and builds in the second integrator. Unless this signal can be suppressed by other means, CCD storage capacity and power is increased, and comparators must resolve small differentials on a large common-mode.

4.2 Analog Integration Designs

4.2.1 Subpipelined Implementation

This section presents implementation details of designs 1 and 2 which use analog integration. Figure 4-1 shows a simplified block diagram of this architecture. (Tables of block diagram notation and common variable notation are included in Appendices A and B.) In this simplified representation, each conversion block contains only one delay and all of its operations must be completed within a single clock period.

Throughput for the example designs is improved by dividing each conversion block into multiple pipeline substages. In general, $n+1$ pipeline substages are needed for a device with an n-bit ADC. The distribution of the operations in Figure 4-1 across a subpipeline of six stages is shown in Figure 4-2. Signals enter the first substage from a preceding conversion block, pass vertically through the pipeline, and leave substage 6 for a subsequent block. In this approach, each feedback operation is allowed multiple clock cycles to complete. Since fewer operations are performed per cycle, a higher frequency clock can be used. Latency is increased, but throughput remains one result per cycle. The example shown is that of design 1. Design 2 is similar, but contains only 4 substages.

Contents of the six substages are shown in Figure 4-3, Figure 4-4, and Figure 4-6. Successive substages are separated by delay elements. Signal names include indices for conversion block and substage number and subscripts of p and m for differential polarity. For example, the name s_{ip}[4,2]represents the positive

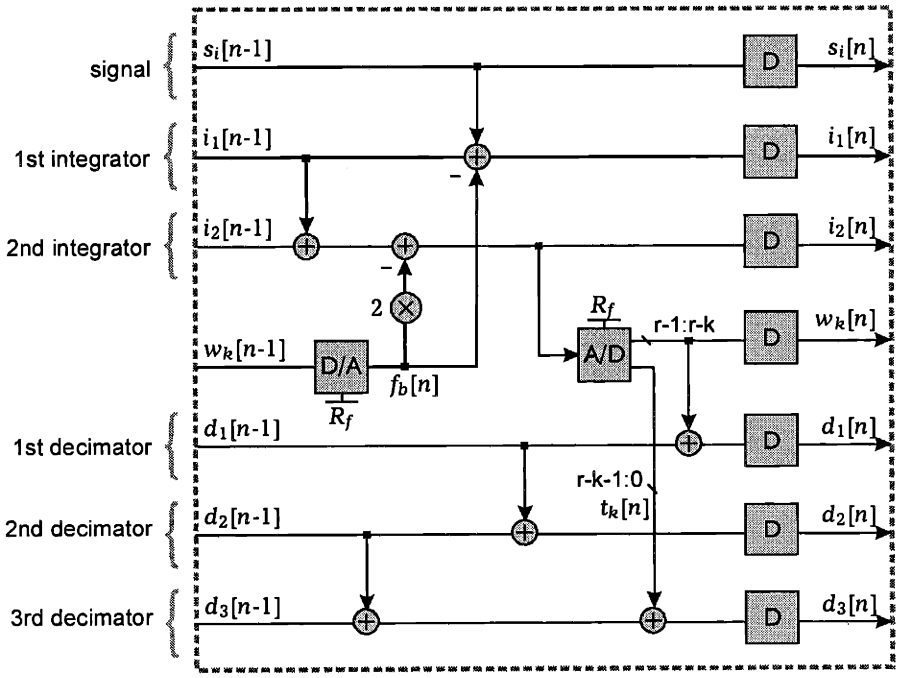


Figure 4-1. Second-order analog integration architecture. As shown, each conversion block includes a single delay.

input sample within substage 2 of conversion block 4. Bus widths are indicated in the figure for all signals with a width greater than one.

Substage 1

The 1st integrator value, $(i_{1p} - i_{1m})$, is modified in substage 1. Negative charge from the delayed input sample $(s_{ip} - s_{im})$ is combined on a wire with positive charge from the feedback quantity $(f_{bp} - f_{bm})R_f$. The result is added to the 1st integrator channel by means of a DDS subtraction circuit.

Positive charge at the input to this circuit is generated from a D/A conversion block with an array of identically sized capacitors. Capacitors are controlled by the complementary 2-bit signals f_{bp} and f_{bm}, which were generated by a previous conversion block. After a $\sqrt{4P}$ reduction due to averaging between the DACs in all stages, the matching between these elements must be as accurate as the overall converter.

Negative charge, representing the delayed input sample, is also needed at the input to the subtraction circuit. One approach to generating this signal is to capture a single charge packet at the beginning of the

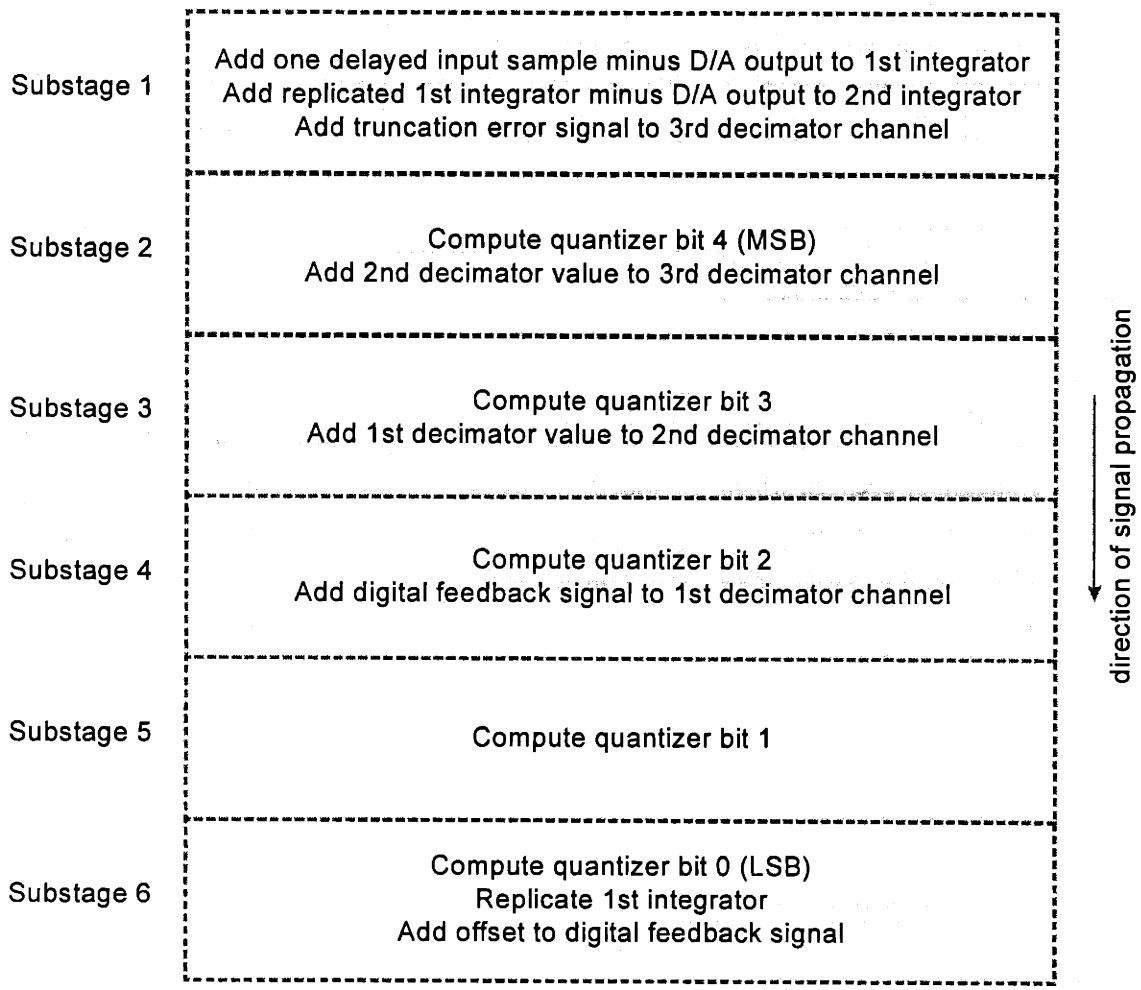


Figure 4-2. Subpipelined analog integration conversion block. Operations are distributed across six pipeline substages.

pipeline and use a DDS charge replicator to reproduce it within each conversion block. Replicated packets can then be used for subtraction. This approach is not acceptable because full converter linearity, which would be required from replicated packets, is not possible from available charge replication circuits. This approach also has the disadvantage that full converter SNR is required from the single delayed input sample.

A preferred approach that does not involve charge sensing elements is shown in the figure. A total of P separate charge packets, with nominally identical values, are captured at the beginning of the pipeline. The entire array of them is transferred along the pipeline, and one packet is used in each conversion block. Since the SNR required from each charge packet in this approach is reduced by a factor of $\sqrt{P}$, each charge packet can be a factor of P smaller and the total amount of charge involved is the same in either approach. Matching among the charge generators in this array is not a concern. The reason for this is that a POSC is tolerant of gain and offset variations in its input signal paths. Since every incoming signal exercises the same set of elements, such variations do not add nonlinearity.

The 2nd integrator is also modified in substage 1. Negative charge, generated by replicating $(i_{1p} - i_{1m})$, is combined on a wire with positive charge $(f_{bp} - f_{bm})2R_f$, from the feedback D/A, and the result is added to the 2nd integrator channel by means of a DDS subtraction circuit. Because the linearity required from this operation is reduced by the gain of the outer feedback loop, charge replication circuits are acceptable in this path.

A final operation within substage 1 is addition of the 3-bit truncation signal t_k, from a previous conversion block, to the 3rd decimator channel.

Substage 2

Substage 2 generates the most significant bit of the quantizer outputs w_r and w_k. This is accomplished by sensing the 2nd integrator packets, which were modified during substage 1, using a pair of floating gate elements and comparing them using a circuit such as mirrored charge comparison. Only moderate accuracy is needed from this operation because comparator inaccuracies and circuit noise are suppressed by 2nd-order noise shaping. For designs 1 and 2, comparator inaccuracies are attenuated by 49 dB and 73 dB.

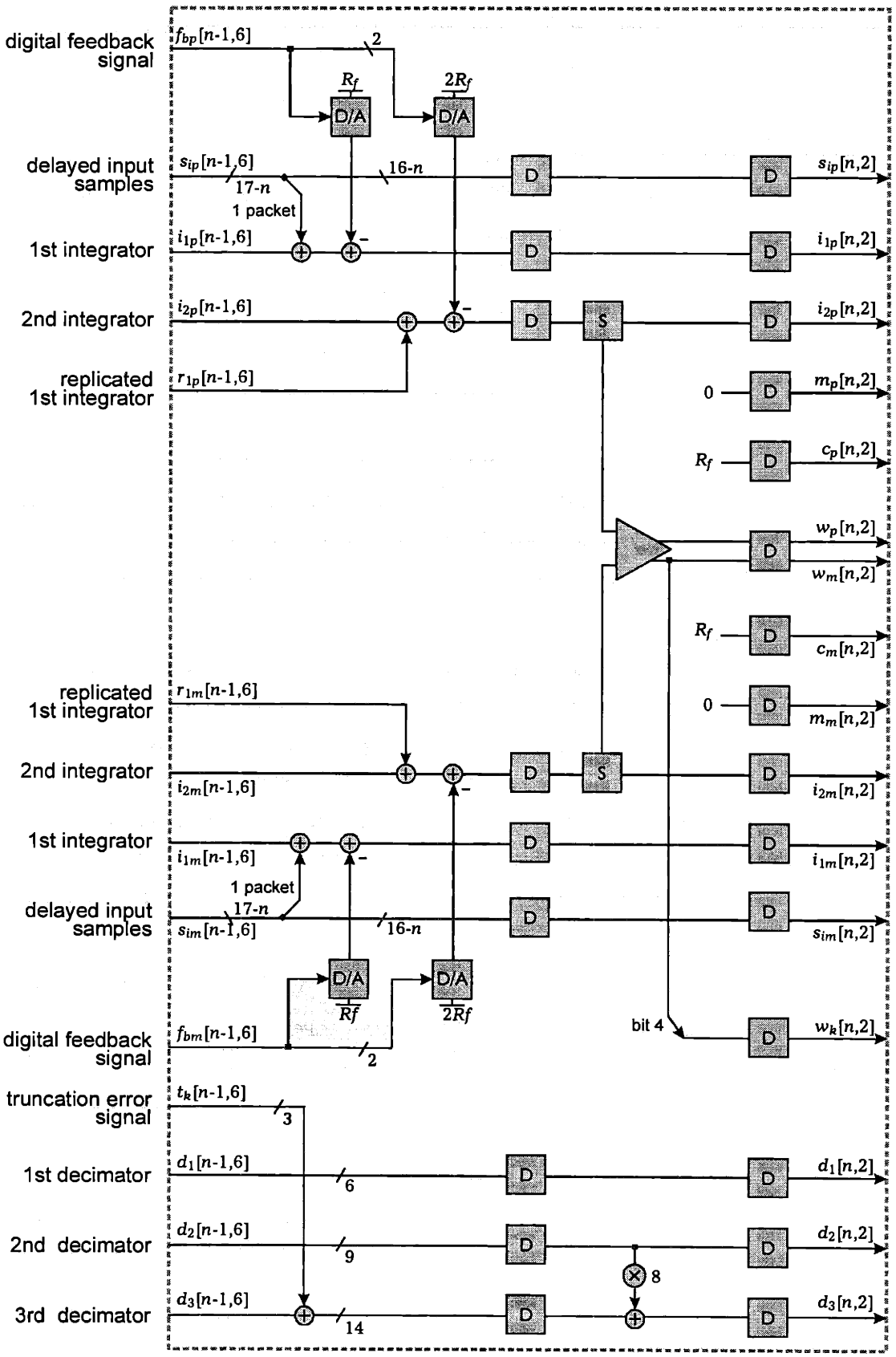


Figure 4-3. Substages 1 and 2 of a subpipelined analog integration design. D/A operations are performed in substage 1. The quantizer MSB is computed in substage 2.

At the end of substage 2, two new sets of CCD channels are formed. They are used to compute the remaining 4 ADC bits. The scaling channel, with signals c_p and c_m, begins with a full-scale packet. The modification channel, with signals m_p and m_m, begins with zero charge.

A final operation within substage 2 is the addition of d_2 to the 3rd decimator channel. Its value is shifted left by 3 bit positions before the add. This is necessary because the least significant bits of d_1 and d_2 occupy position 3 in the modulator output, whereas that for d_3 occupies position 0.

Substage 3

Substage 3 generates bit 3 of the ADC output. First, full-scale reference packets in the scaling channel are divided in half using charge splitting techniques. Second, the comparator result from substage 2 is used to control conditional transfer elements, labeled CT. Using these elements, the $R_f/2$ scaling packet is added to the modification channel on either the positive or negative side, whichever has a smaller value. No addition occurs to side with a larger value. Next, floating gate sensing elements nondestructively sense the sum of charge packets $\left(i_{2p} - i_{2m}\right)$ and $\left(m_p - m_m\right)$. This is accomplished by covering wells in the 2nd integrator and modification channels with a single floating gate. Finally, the results are compared to produce bit 3 of w_k.

The modification channel after substage 4 contains a value which is the result of a 2-bit DAC formed by the binary scaled reference signals c_p and c_m and the conditional transfer elements. It is identical to the signal generated by feedback DACs in substage 1 of the next conversion block. In principle, the modification signal could be used, in place of substage 1 DACs, to modify the 2nd integrator. However, two disadvantages would occur with such an approach. First, high accuracy, required in the feedback path, is not possible from binary-scaled references. Second, common mode charge would accumulate in the integrator channels since the modification signal achieves subtraction by performing complementary addition.

The approach that is shown in the figure is preferable. Binary-scaled references are used, because of their simplicity, to generate a temporary feedback signal that is used as part of the ADC computation. But this temporary feedback charge is kept separate from the integrator so that it can be discarded once the ADC computation is complete. It is never permanently added to the integrator, does not appear in the Δ-Σ feedback path, and experiences 2nd-order noise shaping. The actual feedback signal is generated using DDS D/A conversion circuits in substage 1. In this approach, higher DAC accuracy is possible because signals are formed from an array of identically sized elements. Thermal noise is introduced only once in the feedback path since only one DAC operation is performed. An additional advantage of this approach is that common-mode negative and positive charges are canceled so that signals do not build up along the pipeline.

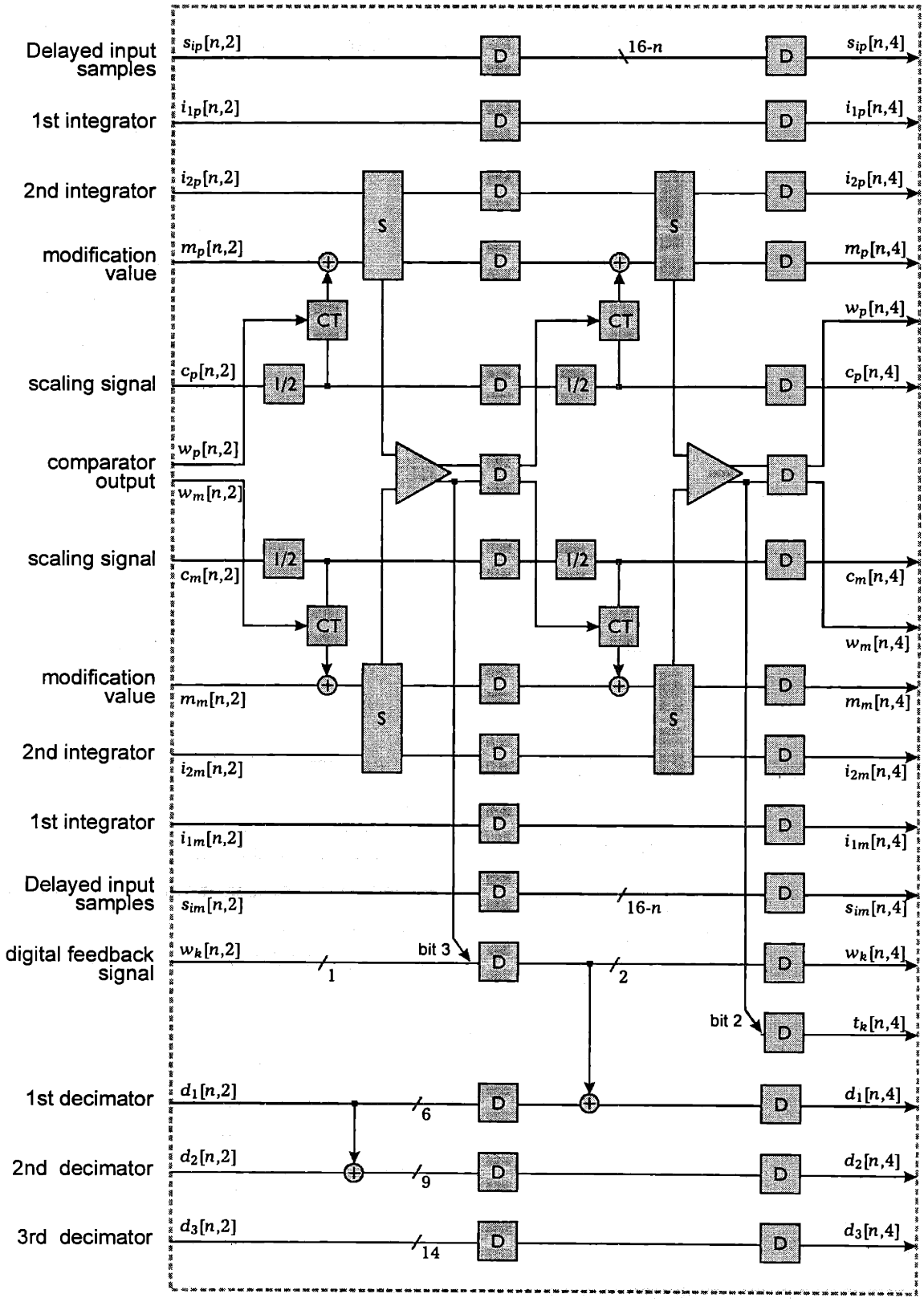


Figure 4-4. Substages 3 and 4 of a subpipelined analog integration design. Bits 3 and 2 of the modulator output are computed.

A final operation in substage 3 is the addition of d_1 to the 2nd decimator.

Substage 4

Substage 4 generates bit 2 of the modulator output in a manner similar to that described above. This signal forms the most significant bit of the truncation quantity t_k.

The 2-bit signal w_k, which was completed during substage 3, is added to the 1st decimator channel in substage 4. The number of bits in this signal sets the required widths of all decimator busses. The magnitudes of d_1, d_2, and d_3 are given by (2-46), (2-47), and (2-50). They increase nearly linearly, quadratically, and cubically from zero in the first pipeline stage to a maximum in the last pipeline stage. Their minimum and maximum values occur when $S_g = \pm R_f$. The number of bits needed to represent d_1 in its nth stage, stage n+2, is

$$\text{width}(d_1[n]) = \log_2(3n+1) \ . \tag{4-1}$$

The number of bits needed to represent d_2 in stage n+3 is

$$\text{width}(d_2[n]) = \log_2\left(\frac{3n(n+1)}{2}+1\right). \tag{4-2}$$

The least significant bit of these signals occupies position 3 in the decimator result. The least significant bit of d_3 occupies position 0 in the decimator's result. The number of bits needed to represent d_3 in the nth d_3 stage, stage n+4, is

$$\text{width}(d_3[n]) = \log_2(4n(n+1)(n+2)+1) \ . \tag{4-3}$$

Bus widths for design 1 are shown in Figure 4-4. Signals d_1, d_2, and d_3 have widths of 6, 9, and 14 bits. After d_3 is normalized, using (2-52), the result is a 13-bit quantity.

The simplest method of sizing decimator busses is to use identical bus widths, given by (4-1), (4-2), and (4-3) at n=P, for all conversion blocks. Alternatively, power can be reduced if widths are varied with conversion block number so that only necessary adders and latches are included. The hardware and power reduction that is achieved from a variable bus width approach is plotted in Figure 4-5 as a function of pipeline length.

Digital additions in the decimator are shown in Figure 4-4 occupying one clock cycle. However, when bus widths are large and adder speed is a concern, additions can be distributed across multiple substages and multiple clock cycles. This is possible because neither d_1, d_2, nor d_3 is needed until the end of the pipeline.

Substage 5

Substage 5 computes bit 1 of the modulator output in a manner similar to that described above.

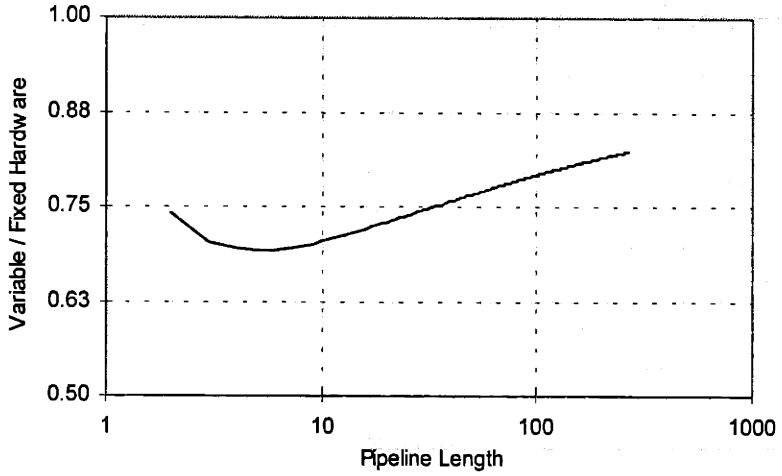


Figure 4-5. Ratio of variable-bus-width to fixed-bus-width decimator hardware and power. Variable approach eliminates unnecessary hardware.

Substage 6

The final substage generates the least significant bit, bit 0, of the modulator output. Its result is merged with comparator outputs from substages 5 and 4 to form the 3-bit truncation signal t_k that is added to d_3 in the next conversion block. After the final comparison is complete, scaling and modification channels are terminated and their signals are discarded.

An additional operation occurring in substage 6 is the replication of signals i_{1p} and i_{1m} to generate two new charge packets, r_{1p} and r_{1m}. These signals are passed forward, along with f_{bp} and f_{mb}, to the next conversion block for use in D/A conversion and subtraction circuits.

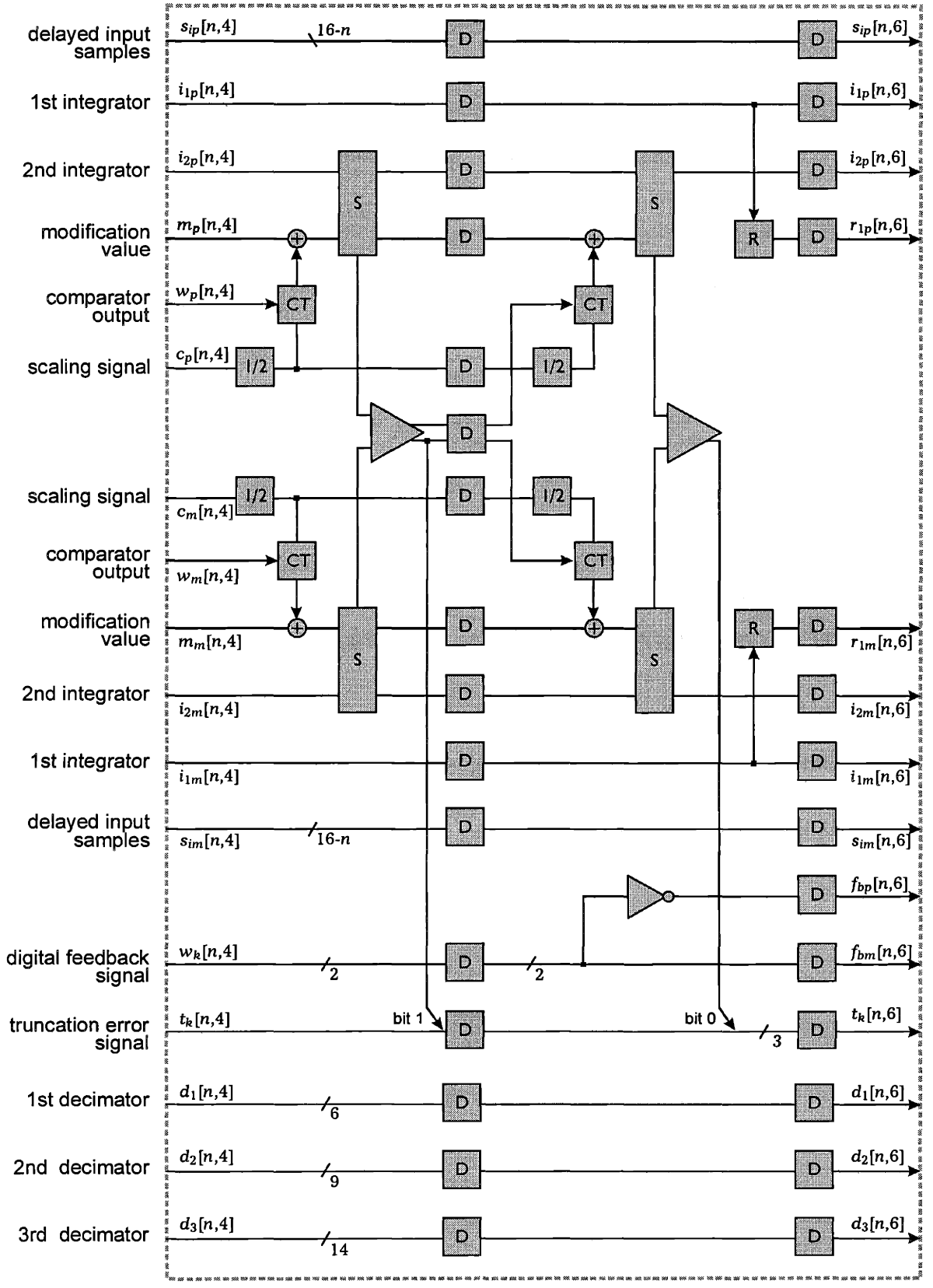


Figure 4-6. Substages 5 and 6 of a subpipelined analog integration design. Bits 1 and 0 of the modulator output are computed and the 1st integrator channel is replicated.

4.2.2 Device Sizing

Architectural parameters, determined in section 4.1.1 for designs 1 and 2, are summarized in Figure 4-7. For design 1, the number of conversion blocks was chosen to achieve its targeted signal to thermal noise ratio, assuming a 5-bit ADC. In contrast, the number of conversion blocks for design 2 was chosen to achieve targeted linearity. Nonlinearity, introduced by i_1 charge replicators, is attenuated by a factor of $(P+2)/3$ due to gain in the outer feedback loop. If replicator linearity is assumed to be 60 dB, then at least 47 conversion blocks are needed for design 2. Given this pipeline length, the number of ADC bits for design 2 was selected to achieve its targeted signal to thermal noise ratio. Because of replicator nonlinearity, design 2 can not make full use of multi-bit quantization and truncated feedback.

Once the parameters in Figure 4-7 have been determined, CCD and capacitor device sizes follow directly from targeted signal to thermal noise ratio. The approach to sizing these circuits is outlined below and the resulting values for designs 1 and 2 are listed in Figure 4-8.

Signal Channel Sizes

First, capacitance values are derived for charge generators at the converter input. In a P-stage pipeline, P separate differential pairs of charge packets are generated, and one is used in each conversion block. Charge generator thermal noise is represented as an additive error, n_{si}, in each of the input samples. A second source of thermal noise is the differencing of delayed input samples and D/A levels at the input to the 1st integrator channel. Noise n_{i1}, introduced by this DDS subtraction circuit, is indistinguishable from noise in the delayed input samples. When both of these sources are included, signal channel packets are given by

$$\left(s_{ip}[n]-s_{im}[n]\right)=2S_g+\sqrt{2\,\overline{n_{si}^2[n]}+2\,\overline{n_{i1}^2[n]}}\ . \tag{4-4}$$

All variables in (4-4) represent charge quantities.

	Architectural Parameters	Design 1	Design 2
1	Desired SNR of the overall converter (dB)	72	84
2	Desired converter signal to quantization noise ratio (dB)	75	87
3	Desired converter signal to thermal noise ratio (dB)	75	87
4	Number of pipeline conversion blocks	16	51
5	Number of substages per conversion block	6	4
6	Number of feedforward ADC bits	5	3
7	Number of feedback bits	2	2

Figure 4-7. Summary of architectural parameters for analog integration designs 1 and 2 with targeted SNR of 72 dB and 84 dB.

The decimator weight that is assigned to each s_i packet is a function of the conversion block in which the packet is used. The final decimator output, found from (2-50), is given by

$$d_3[P+4] = \sum_{i=1}^{P} \left(s_{ip}[n] - s_{im}[n]\right) \frac{(P+2-i)(P+1-i)}{2R_f} . \tag{4-5}$$

The sample S_g has an amplification of

$$\frac{d_3[P+4]}{S_g/R_f} = \frac{1}{3}\left(P^3 + 3P^2 + 2P\right) . \tag{4-6}$$

Charge generator noise, which adds out of phase, appears with an amplification of

$$\frac{d_3[P+4]}{\sqrt{\overline{n_{si}^2} + \overline{n_{si}^2}}/R_f} = \sqrt{\frac{1}{2}\sum_{i=1}^{P} (P+2-i)^2 (P+1-i)^2} . \tag{4-7}$$

Combining (4-6) and (4-7) yields a signal to thermal noise ratio at the converter output of

$$\text{SNR} = \frac{\sqrt{\overline{S_g}}\left(P^3 + 3P^2 + 2P\right)}{3\sqrt{\left(\overline{n_{si}^2} + \overline{n_{i1}^2}\right)\frac{1}{2}\sum_{i=1}^{P} (P+2-i)^2 (P+1-i)^2}} . \tag{4-8}$$

When n_{si} and n_{i1} are the dominant sources of thermal noise, capacitor and CCD well sizes throughout the remainder of the converter follow directly from (4-8). The validity of this assumption is demonstrated later in this section. Figure 4-9 lists device parameters that are used in the analysis below.

The thermal noise of each charge generator is found from (3-55). It increases with the size of its input capacitors C_1 and C_2 and has a mean-square value in Coulombs of

$$\overline{n_{si}^2} = 3kT\left(\frac{C_1 C_2}{C_1 + C_2}\right) . \tag{4-9}$$

Storage capacity, equal to the sum of gate-to-channel and channel-to-substrate capacitances, that is required from signal channel CCD wells is

$$C_{si} = 0.8\left(\frac{C_1 C_2}{C_1 + C_2}\right) . \tag{4-10}$$

It depends on the rms input and peak CCD storage voltages, given by Figure 4-9(8) and Figure 4-9(9).

The thermal noise of each DDS D/A conversion and subtraction circuit is

$$\overline{n_{i1}^2} = \frac{4}{5}kT\left(\frac{C_1 C_2}{C_1 + C_2}\right) . \tag{4-11}$$

	Derived Values for the Signal Channel	Design 1	Design 2
15	Amplification of charge generator signals at the converter output	1632	46852
16	Amplification of charge generator and DAC noise at the converter output	376	6164
17	Series input capacitance of each charge generator (F)	5.3E-14	2.7E-13
18	Thermal noise of each charge generator (electrons)	160	364
19	Thermal noise of DAC/subtraction circuits in the 1st integrator (electrons)	83	188
20	SNR of each charge generator (dB)	63	70
21	Maximum number of electrons in each signal channel packet	6.6E+05	3.4E+06
22	Attenuation of charge replicator nonlinearity (dB)	16	25
23	Capacity of each signal channel storage well (F)	4.2E-14	2.2E-13
	Derived Values for the First Integrator		
24	Ratio of 1st integrator to signal channel storage well capacitance	1.5	2.5
25	Maximum number of electrons in each first integrator packet	9.9E+05	8.5E+06
26	Thermal noise of charge replicators in the 1st integrator (electrons)	79	231
27	Thermal noise of DAC/subtraction circuits in the 2nd integrator (electrons)	117	266
28	Amplification of items 26 and 27 at the converter output	55	302
29	Ratio, at the converter output, of signal to items 26 and 27 (dB)	94	115
	Derived Values for the 2nd Integrator		
30	Ratio of 2nd integrator to signal channel storage well capacitance	2	3
31	Maximum number of electrons in each 2nd integrator packet	1.3E+06	1.0E+07
32	Ratio of modification to signal channel storage well capacitance	1	1
33	Ratio of scaling to signal channel storage well capacitance	1	1
34	Thermal noise of comparator floating gate sensing elements (electrons)	111	292
35	Amplification of item 34 at the converter output	6	10
36	Ratio, at the converter output, of signal to item 34 (dB)	116	146
	Derived Values for the Comparator		
37	Quantization step size of coarse ADC (electrons)	2.1E+04	4.3E+05
38	Comparator input capacitance (F)	6.3E-14	4.4E-13
39	Charge referred error due to mirror threshold mismatches (electrons)	3.6E+03	2.5E+04
40	Attenuation of 2nd-order shaped quantization noise (dB)	49	73
41	Ratio, at the converter output, of signal to item 39 (dB)	86	107

Figure 4-8. Computed capacitance, noise, and SNR values for designs 1 and 2.

It increases with DAC element capacitance, the size of which is in turn determined by charge generator capacitances and the DAC supply voltage. Combining (4-8), (4-9), and (4-11) yields a relationship of

$$\left(\frac{C_1C_2}{C_1+C_2}\right)=\mathrm{SNR}^2\left(\frac{171kT\sum_{i=1}^{P}(P+2-i)^2(P+1-i)^2}{5\left(P^3+3P^2+2P\right)^2}\right) \tag{4-12}$$

between input capacitance and converter signal to thermal noise ratio. The results for 75 dB and 87 dB SNR for designs 1 and 2 are listed in Figure 4-8(17). The SNR required from each individual charge generator, listed in Figure 4-8(20), is less than that of the overall converter.

First Integrator Sizes

The size of other CCD wells is a function of the value in (4-12). To avoid the possibility of integrator saturation, the storage capacity of wells in the 1st integrator are made larger than those in the signal channel by two DAC LSBs, with one on either edge of the full scale range. The storage capacity of these wells is

$$C_{i1}=1.2\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-13}$$

These values and their associated full-scale number of electrons are listed in Figure 4-8(24) and (25).

Thermal noise is introduced by replicators in the 1st integrator channel and by D/A conversion and subtraction circuits at the input to the 2nd integrator channel. These noise components are indistinguishable from one another and are treated as a single noise source n_{i2}. The amplification of n_{i2} at the converter output,

Device Parameters

8	RMS signal voltage at each charge generator input (V)	0.71
9	Peak storage range in non-floating CCD wells (V)	2.50
10	Ratio of storage well to barrier well area	4
11	Capacitance ratio Cgc/Ccs for CCD storage wells	12
12	Analog supply voltage (V)	5
13	Voltage swing for barrier gate clocks (V)	2
14	Voltage swing for storage gate clocks (V)	5
15	Charge mirror threshold mismatches (V)	3.0E-03

Figure 4-9. Device characteristics and operating parameters used to determine CCD and capacitor sizes.

$$\frac{d_3[P+4]}{\sqrt{\overline{n_{i2}{}^2}}} = \sqrt{2\sum_{i=1}^{P}(P+1-i)^2}\ , \tag{4-14}$$

has values listed in Figure 4-8(28).

Charge replicator thermal noise, given by (3-119), depends on the effective capacitance of their floating gates to ground. The relationship between this effective capacitance and storage capacity is derived in section 4.2.3. It depends on the ratio between gate-to-channel and channel-to-substrate capacitance and on storage voltage in the CCDs. These parameters are listed in Figure 4-9(11) and (9). For the capacitance values in (4-13), this noise is

$$\overline{n_{i2}{}^2} = 0.73kT\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-15}$$

Thermal noise in the D/A conversion and subtraction circuit,

$$\overline{n_{i2}{}^2} = 1.6kT\left(\frac{C_1C_2}{C_1+C_2}\right), \tag{4-16}$$

is larger than that for the 1st integrator channel because it contains a larger full-scale reference.

When only these noise sources are considered, SNR is

$$\text{SNR} = \frac{\left(P^3+3P^2+2P\right)}{\sqrt{84kT\sum_{i=1}^{P}(P+1-i)^2}}\sqrt{\left(\frac{C_1C_2}{C_1+C_2}\right)}\ . \tag{4-17}$$

Values for (4-15), (4-16), and (4-17) are listed in Figure 4-8(26), (27), and (29). The noise due to these sources is small compared to that in the signal channel.

Second Integrator Device Sizes

Wells in the scaling and modification channels are sized the same as those in the signal channel. Wells in the 2nd integrator are made larger than those in the signal channel by four DAC LSBs to avoid the possibility of integrator saturation. Their storage capacity,

$$C_{i2} = 1.6\left(\frac{C_1C_2}{C_1+C_2}\right), \tag{4-18}$$

and their associated full-scale number of electrons are listed in Figure 4-8(30) and (31).

Errors are introduced when signals $\left(i_{2p}+m_p\right)$ and $\left(i_{2m}+m_m\right)$ are sensed and compared. The amplification of this noise at the converter output is

$$\frac{d_3[P+4]}{\sqrt{\overline{n_m{}^2}}} = \sqrt{2P}\ . \tag{4-19}$$

A first source is thermal noise from floating gates at the comparator input. It has a mean square value of

$$\overline{n_m{}^2} = 1.5kT\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-20}$$

A second source is comparator errors. These are introduced by component mismatches, such as threshold variations between the charge mirror gates in a mirrored comparison circuit. The mean square charge referred value of this error, given by (3-143), is

$$\overline{n_m{}^2} = \overline{\Delta v_{th}{}^2}\left(1.2Z\left(\frac{C_1C_2}{C_1+C_2}\right)\right)^2, \tag{4-21}$$

where Z is the ratio between the length of each clock phase and the sensing time constants. Values for these noise sources and the resulting SNR for designs 1 and 2 are listed in Figure 4-8(34), (38), and (41). The level of this noise is also small compared to that in the signal channel.

4.2.3 Power Dissipation

The relationship between power dissipation and channel charge, given by (3-157), is used to determine the average CCD power dissipation for designs 1 and 2. For wells in the signal channel alone

$$\overline{P} = (\text{total \# of } si \text{ delays})(2 \text{ storage wells / delay})\left(C_{eff}\left(\overline{V_w}\right)V_{gs}{}^2 + C_{eff}\left(\overline{V_w}=0\right)V_{gb}{}^2\right)f\ . \tag{4-22}$$

The first term is due to storage wells with a gate voltage swing of V_{gs} and an effective capacitance evaluated at an average storage voltage of $\overline{V_w}$. In a differential structure, $\overline{V_w}$ normally equals the common-mode offset. The second term is due to barrier wells with a gate voltage swing of V_{gb}. Since barrier gates do not hold charge, their effective capacitance is that for deep depletion, where V_w=0.

A similar computation is used to arrive at power estimates for wells in the remaining CCD channels. The number of wells in each channel and the total CCD driver power are listed in Figure 4-10. When the power of CMOS control logic and supporting circuits is included as well, total device power will always be larger than these values.

Total power for design 2 is much larger than that for design 1 as a result of two factors. First, larger charge packets are needed in design 2 to reduce thermal noise of the input charge generators. Second, a large number of conversion blocks are needed for design 2 because of the need to suppress nonlinearity of charge replicators in the 1st integrator channel.

	Derived Power	Design 1	Design 2
42	Number of signal channel storage wells	2880	20400
43	Number of first integrator storage wells	360	800
44	Number of 2nd integrator storage wells	384	816
45	Number of modification storage wells	320	612
46	Number of scaling storage wells	128	408
47	Delayed signal channel (si) power (mW/MHz)	0.926	33.949
48	First integrator channel (i1) power (mW/MHz)	0.174	3.328
49	Second integrator channel (i2) power W/MHz	0.247	4.074
50	Modification channel (m) power (mW/MHz)	0.103	1.018
51	Scaling channel (s) power (mW/MHz)	0.041	0.679
52	Total CCD well capacitance seen by driver (F)	7.38E-11	2.13E-09
53	Total CCD driver power (mW/MHz)	1.491	43.049

Figure 4-10. Power dissipation for analog integration designs 1 and 2. Design 2 has larger power as a result of thermal noise and linearity constraints.

4.3 Digital Integration Designs

One disadvantage of analog integration is that signals in the 1st integrator must be replicated before they are added to the 2nd integrator. The need to suppress charge replicator nonlinearity sets a lower limit on the number of conversion blocks and on power dissipation and prevents analog integration devices from making full use of truncated feedback or of its associated power reduction.

In comparison to analog integration, digital integration performs similar Δ-Σ computations but also has the advantage that it does not require charge replicators within its outer feedback loop. A block diagram of a 2nd-order digital integration architecture is shown in Figure 4-11 and the distribution of its operations across a six-stage subpipeline is shown in Figure 4-12. In this approach, the signal input to the 2nd integrator is the result of a DAC whose full-scale reference is the input sample and whose digital weights are fixed in each stage. The contents of each substage in design 3 are described below with reference to Figure 4-13, Figure 4-14, and Figure 4-15. Design 4 is similar to design 3, except that it includes 7, rather than 5, bits in its quantizer, and has 8 substages per conversion block.

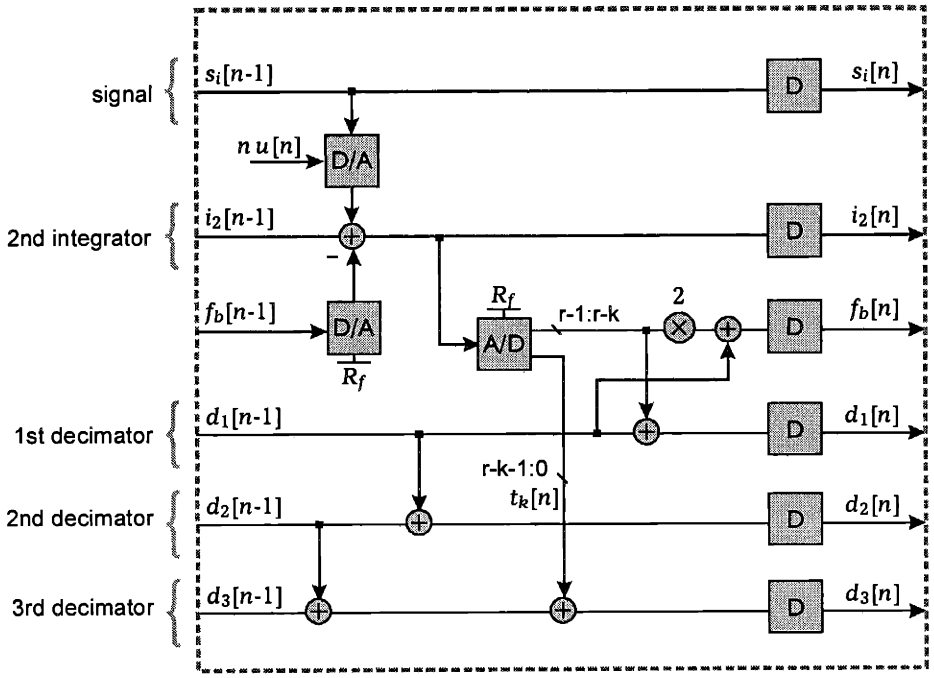


Figure 4-11. Second-order digital integration architecture. As shown, each conversion block includes a single delay.

4.3.1 Subpipelined Implementation

Substage 1

Substage 1 is used to modify the 2nd integrator, $(i_{2p} - i_{2m})$. Negative charge from the signal DAC is combined on a wire with positive charge $(f_{bp} - f_{bm})R_f$ from the feedback DAC, and the result is added to the 2nd integrator channel by means of a DDS subtraction circuit. Positive charge at the input to this circuit is generated through an array of identically sized capacitors in a D/A converter. These elements are controlled by complementary signals f_{bp} and f_{bm}. The D/A full-scale range is PR_f and its least significant bit has a weight of $R_f/4$, since 2-bit feedback is used. Consequently, a total of $4P$ distinct DAC outputs are required. Design 3, shown in the figure, has 6 bits in its DAC.

Noise and mismatches in the reference DACs experience 1st-order noise shaping and are attenuated by a

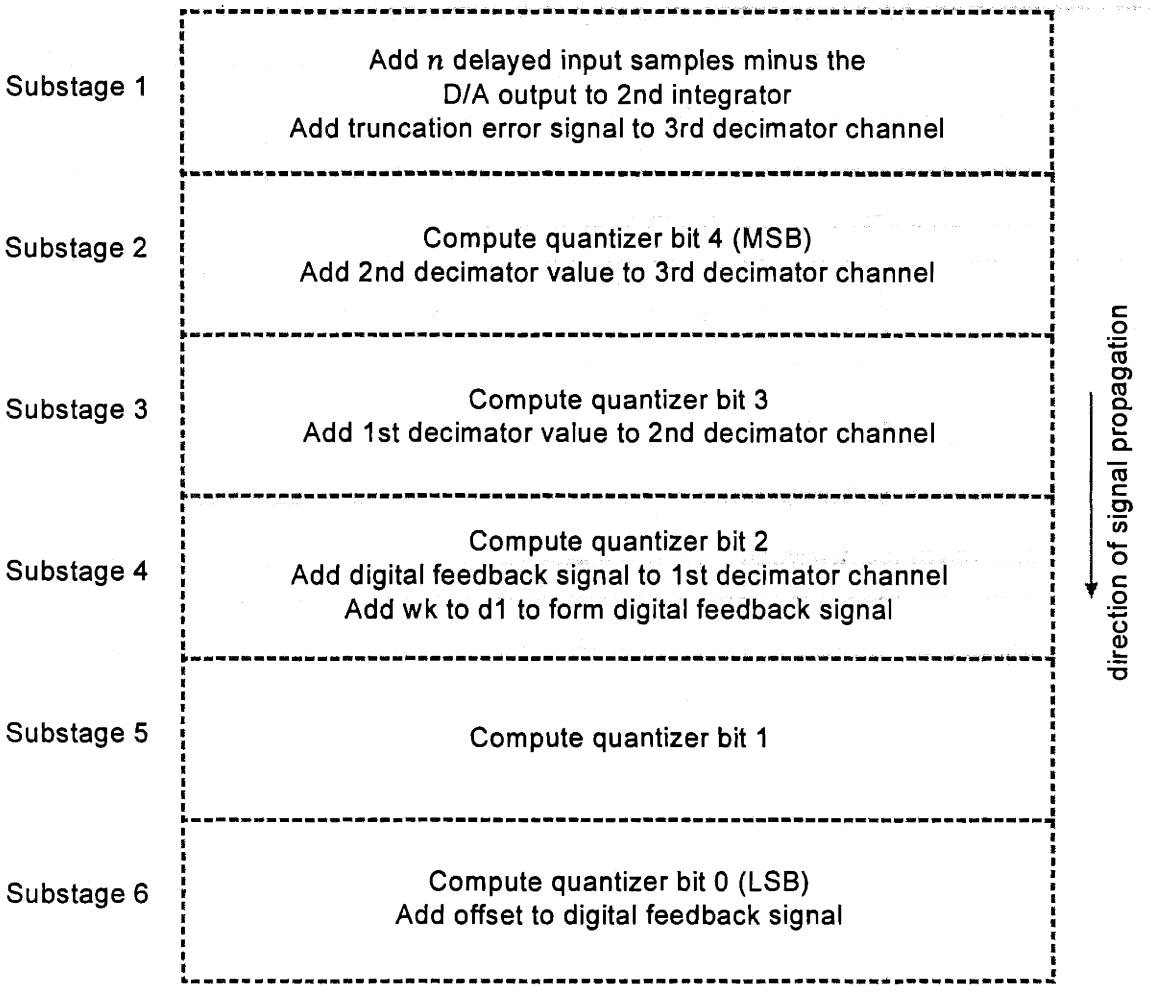


Figure 4-12. Subpipelined digital integration conversion block. Operations are distributed across six pipeline substages and are allowed six clock cycles to complete.

factor of $P^{3/2}/\sqrt{3}$. However, the full scale reference of these DACs is greater than that of the converter by a factor of P. The combination of these effects is that, after averaging between the DACs in all stages, the accuracy of each DAC with respect to its full scale reference must be similar to that of the overall converter.

Negative charge at the input to the 2nd integrator is provided by a signal DAC. An array of $P(P+1)/2$ identically sized charge packets is produced at the beginning of the pipeline. These packets are transferred along the pipeline, and a total of n of them are used at the input to the nth conversion block so that by the final block, all packets have been used. Matching among the charge generators in this array is not a concern because every incoming signal exercises the same set of elements and gain and offset variations within the signal DAC only cause gain and offset variations in the converter output.

Although a factor of $(P+1)/2$ more packets are needed in digital versus analog integration, the total quantity of generated charge is identical in both approaches. In digital integration, noise is attenuated by an additional factor of $\sqrt{(P+1)/2}$, due to averaging, and each charge packet can be made a factor of $(P+1)/2$ smaller. Digital integration also has the advantage that thermal noise is introduced only once before the 2nd integrator.

The primary difference between digital and analog integration lies in the point at which signal and reference values are differenced. In both approaches, the input to the 2nd integrator at stage n contains n signal packets minus an accumulated feedback quantity. In analog integration the feedback quantity is subtracted before the 1st integrator. In digital integration, signal and feedback quantities remain separate until they are subtracted at the input to the 2nd integrator.

Substages 2 through 6

Substages 2 through 6 for digital integration are similar to those described earlier for analog integration, except that no replicators are needed in substage 6. The 2-bit signal w_k is completed after substage 3. A corresponding feedback signal $(d_1[n]+2w_k[n])$, which is needed by the next block, is calculated in the decimator of substage 4. After substage 4, computations within the next conversion block can begin. Decimator bus widths are the same for digital and analog integration.

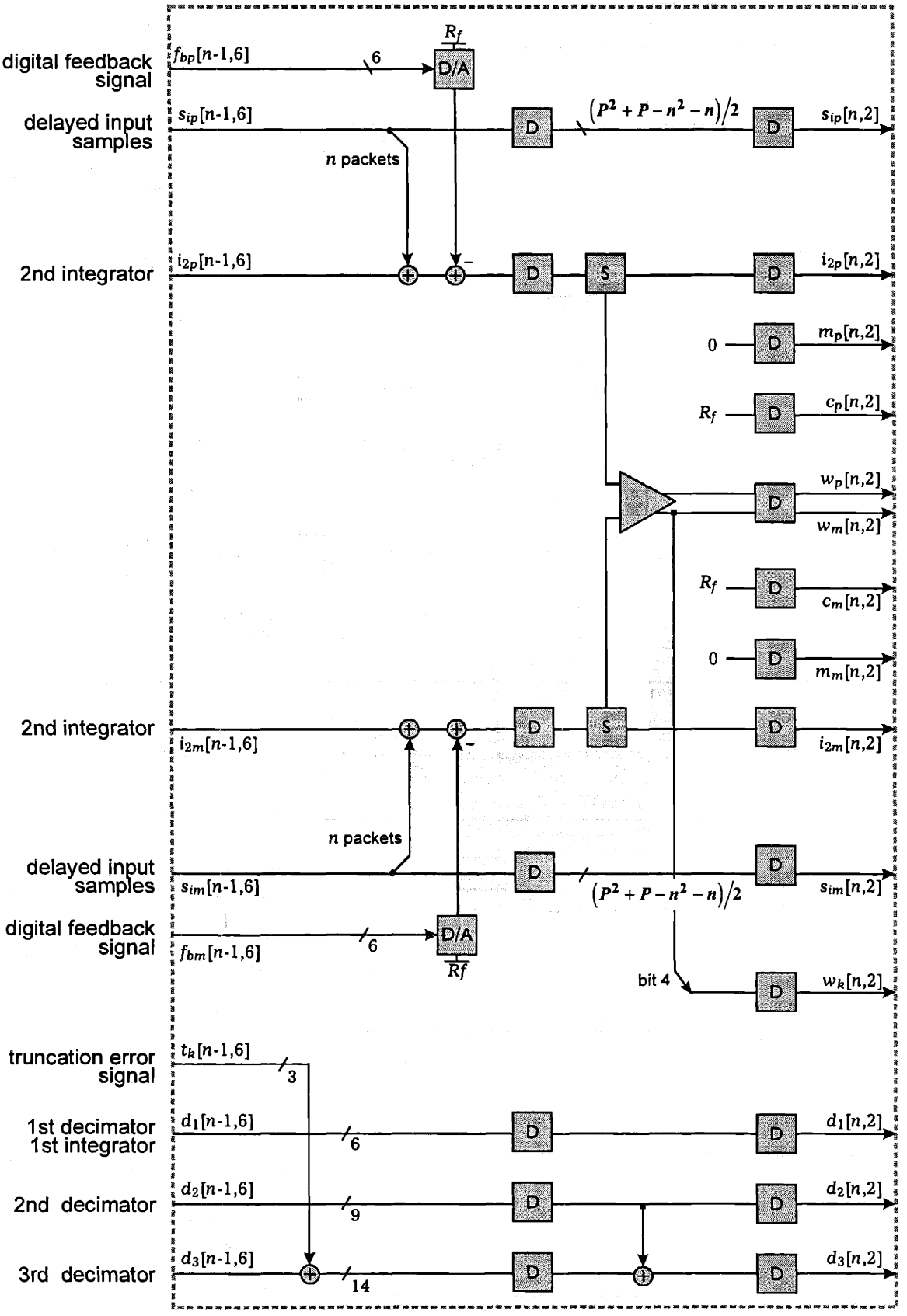


Figure 4-13. Substages 1 and 2 of a subpipelined digital integration design. The 2nd integrator is modified in substage 1. Bit 4 of the modulator output is computed in substage 2.

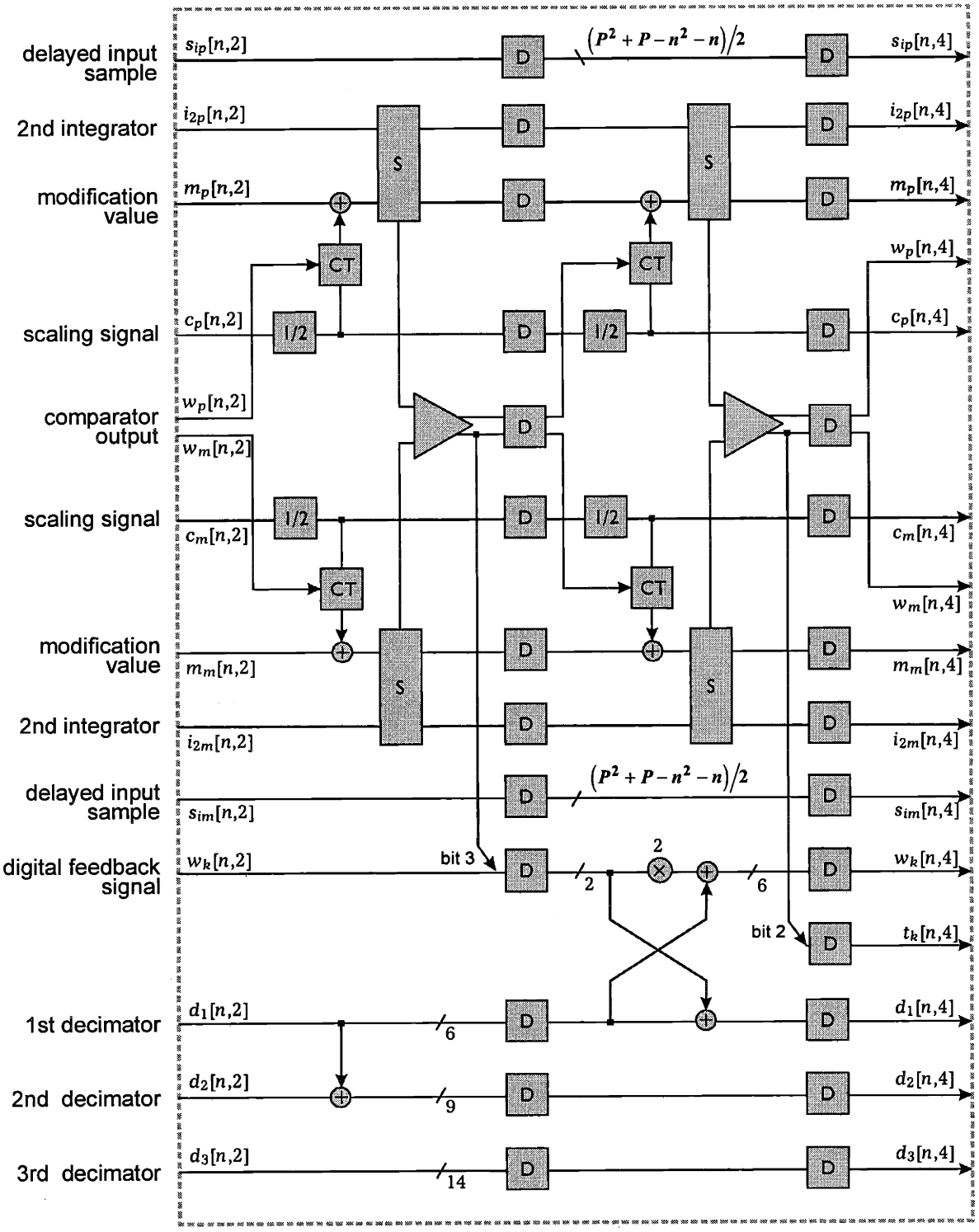


Figure 4-14. Substages 3 and 4 of a subpipelined digital integration design. Bits 3 and 2 of the modulator output are computed. The digital feedback signal is generated in substage 4.

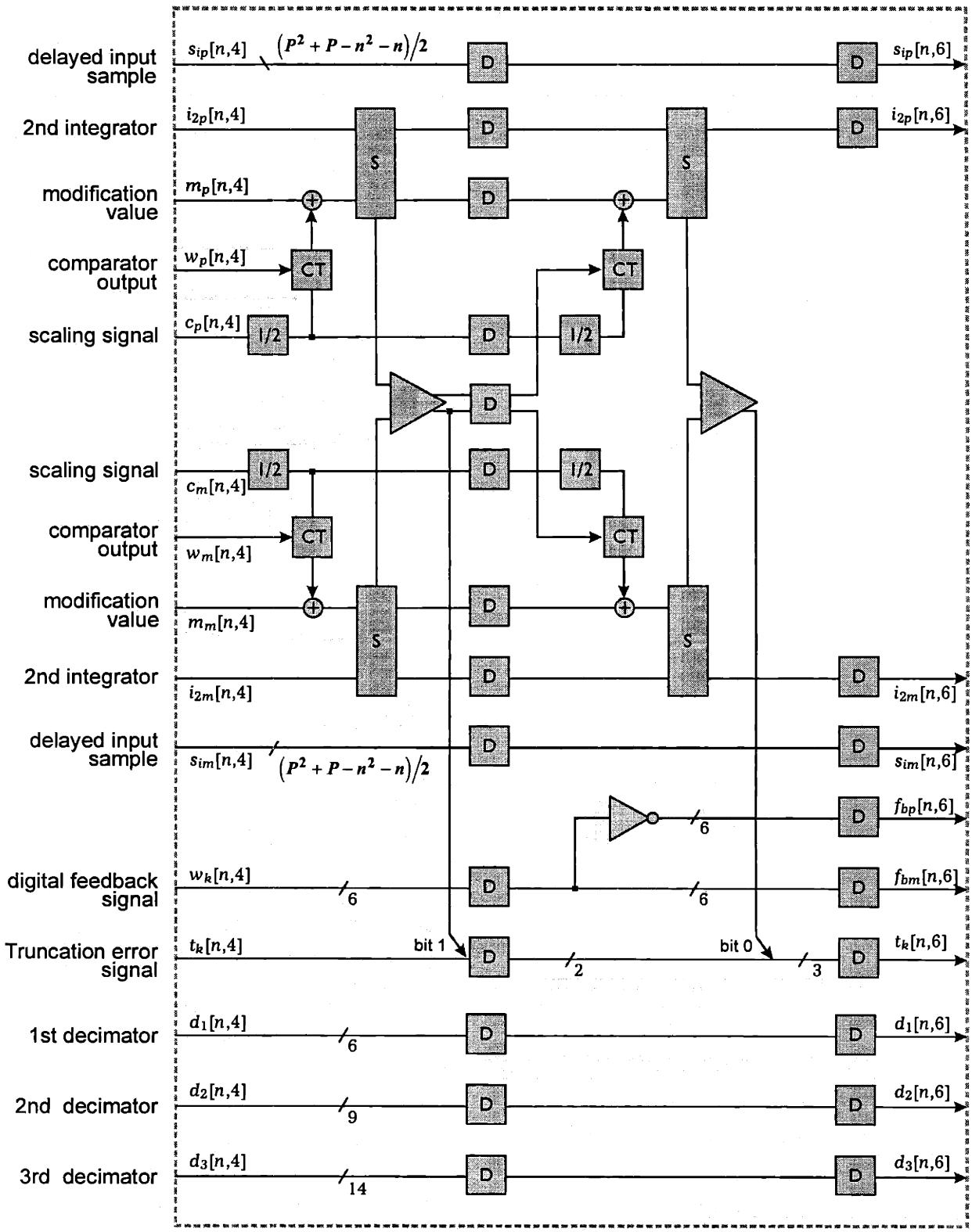


Figure 4-15. Substages 5 and 6 of a subpipelined digital integration design. Bits 1 and 0 of the modulator output are computed.

4.3.2 Device Sizing

Architectural parameters, outlined above for designs 3 and 4, are summarized in Figure 4-16. The number of conversion blocks for design 3 was chosen to achieve the targeted signal to thermal noise ratio, assuming a 5-bit ADC. Design 4 has larger charge packets and can therefore support a greater quantizer range. Its number of conversion blocks was chosen to achieve the desired signal to thermal noise ratio, assuming a 7-bit ADC.

CCD and capacitor device sizes are evaluated below for designs 3 and 4. Calculated values are included in Figure 4-17. Device parameters and circuit operating conditions used for the analysis are the same as those used for analog integration and are listed in Figure 4-9.

Signal Channel Sizes

In a digital integration architecture, a total of $P(P+1)/2$ separate charge packets are generated and P of them are used in each conversion block. Each charge packet is modeled as the sum of an input sample S_g and an additive thermal noise term n_{si}. Equation (2-50) is used to arrive at the decimator response

$$d_3[P+4] = \sum_{i=1}^{P}\sum_{j=1}^{i} 2\left(S_g + n_{si}[i,j]\right)\frac{(P+1-i)}{R_f} \; . \tag{4-23}$$

The weight that each packet experiences depends on the conversion block in which it is used. The input signal S_g has an amplification of

$$\frac{d_3[P+4]}{S_g/R_f} = \frac{1}{3}\left(P^3 + 3P^2 + 2P\right) \; . \tag{4-24}$$

The noise, which adds out of phase, has an amplification of

$$\frac{d_3[P+4]}{\sqrt{\overline{n_{si}^2}}/R_f} = \sqrt{2\sum_{i=1}^{P} i(P+1-i)^2} \; . \tag{4-25}$$

	Architectural Parameters	Design 3	Design 4
1	Desired SNR of the overall converter (dB)	72	84
2	Desired converter signal to quantization noise ratio (dB)	75	87
3	Desired converter signal to thermal noise ratio (dB)	75	87
4	Number of pipeline conversion blocks	16	16
5	Number of substages per conversion block	6	8
6	Number of feedforward ADC bits	5	7
7	Number of feedback bits	2	2

Figure 4-16. Summary of architectural parameters for digital integration designs 3 and 4 with targeted SNR of 72 dB and 84 dB.

Signals are amplified by the same factor in analog and digital integration. However, the digital integration amplification of noise in (4-25) is less than that in (4-7) for analog integration.

An additional source of thermal noise, denoted n_{i2}, is the DDS D/A conversion and subtraction circuits at the input to the 2nd integrator. Since this noise occurs after the 1st integrator, it experiences a reduced amplification of

$$\frac{d_3[P+4]}{\sqrt{\overline{n_{i2}^2}}/R_f} = \sqrt{2\sum_{i=1}^{P}(P+1-i)^2}\ . \tag{4-26}$$

The resulting SNR, found by combining (4-24), (4-25), and (4-26), is given by

$$\mathrm{SNR} = \frac{\sqrt{\overline{S_g}}\left(P^3+3P^2+2P\right)}{3\sqrt{2\sum_{i=1}^{P}\left(\overline{n_{si}^2}\cdot i+\overline{n_{i2}^2}\right)(P+1-i)^2}}\ . \tag{4-27}$$

The thermal noise of each charge generator has a mean-square value in Coulombs of

$$\overline{n_{si}^2} = 3kT\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-28}$$

The thermal noise of each D/A conversion and subtraction circuit, containing a total of $4P$ DAC elements, is

$$\overline{n_{i2}^2} = \frac{4}{5}kTP\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-29}$$

Equation (4-27) is solved for C_1 and C_2 and the resulting capacitance values for designs 3 and 4 are listed in Figure 4-17(18). The corresponding values of n_{si} and n_{i2} are listed in Figure 4-17(19) and (20).

Second Integrator Device Sizes

In most cases, signal channel noise is the dominant noise component and CCD well sizes throughout the remainder of the converter follow directly from (4-27). The storage capacity of signal channel wells,

$$C_{si} = 0.8\left(\frac{C_1C_2}{C_1+C_2}\right), \tag{4-30}$$

depends on rms input voltage and CCD storage voltage. The 2nd integrator wells are sized four DAC LSBs larger than those in the signal channel to avoid the possibility of integrator saturation. Their storage capacity is given by

$$C_{i2} = 1.6\left(\frac{C_1C_2}{C_1+C_2}\right). \tag{4-31}$$

	Derived Values for the Signal Channel	Design 3	Design 4
15	Amplification of charge generator signals at the converter output	1632	1632
16	Amplification of charge generator noise at the converter output	118	118
17	Amplification of DAC noise at the converter output	219	219
18	Series input capacitance of each charge generator (F)	7.9E-15	1.2E-13
19	Thermal noise of each charge generator (electrons)	62	246
20	Thermal noise of DAC/subtraction circuits in the 1st integrator (electrons)	128	508
21	SNR of each charge generator (dB)	55	67
22	Maximum number of electrons in each signal channel packet	9.8E+04	1.6E+06
23	Capacitance of each signal channel storage well (F)	6.3E-15	1.0E-13
	Derived Values for the 2nd Integrator		
24	Ratio of 2nd integrator to signal channel storage well capacitance	2	2
25	Maximum number of electrons in each 2nd integrator packet	2.0E+05	3.1E+06
26	Ratio of modification to signal channel storage well capacitance	1	1
27	Ratio of scaling to signal channel storage well capacitance	1	1
28	Thermal noise of comparator floating gate sensing elements (electrons)	43	171
29	Amplification of item 28 at the converter output	4	4
30	Ratio, at the converter output, of signal to item 28 (dB)	110	122
	Derived Values for the Comparator		
31	Quantization step size of coarse ADC (electrons)	3.1E+03	1.2E+04
32	Comparator input capacitance (F)	7.9E-15	1.2E-13
33	Charge referred error due to mirror threshold mismatches (electrons)	4.4E+02	7.0E+03
34	Attenuation of 2nd-order shaped quantization noise (dB)	52	52
35	Ratio, at the converter output, of signal to item 33 (dB)	90	90

Figure 4-17. Computed capacitance, noise, and SNR values for designs 3 and 4.

These values and their corresponding full-scale number of electrons are listed in Figure 4-17(24) and (25). Scaling and modification channel wells have the same sizes as wells in the signal channel.

In addition to signal channel noise, errors are introduced when signals $\left(i_{2p}+m_p\right)$ and $(i_{2m}+m_m)$ are sensed and compared. These components, including floating gate thermal noise and comparator mismatches, are computed using the same expressions in (4-20) and (4-21) for analog integration. Their values for designs 3 and 4 are listed in Figure 4-17(28) and (33), their decimator amplifications are in (29), and SNR due to only these components is included in (30) and (35). When wells are sized to meet thermal noise requirements of the input charge generators, thermal and comparator noise in the ADC do not limit converter SNR.

4.3.3 Power Dissipation

CCD power for designs 3 and 4 is computed in a similar manner to that described above for designs 1 and 2. Thermal noise limits set charge packet storage capacity throughout the converter. Storage capacity can be translated into power by means of an effective CCD capacitance, given by (3-152). Power in each of the channels is then computed from effective capacitance according to (4-22). The number of wells and corresponding power for each of the signal, integrator, scaling, and modification channels are listed in Figure 4-18.

	Derived Power	Design 3	Design 4
36	Number of signal channel storage wells	32640	43520
37	Number of 2nd integrator storage wells	384	512
38	Number of modification storage wells	320	448
39	Number of scaling storage wells	128	128
40	Delayed sample channel (si) power (mW/MHz)	1.563	33.037
41	Second integrator channel (i2) power W/MHz	0.037	0.777
42	Modification channel (m) power (mW/MHz)	0.015	0.340
43	Scaling channel (s) power (mW/MHz)	0.006	0.097
44	Total CCD well capacitance seen by driver (F)	8.03E-11	1.70E-09
45	Total CCD driver power (mW/MHz)	1.622	34.251

Figure 4-18. Power dissipation for digital integration designs 3 and 4. Design 4 is higher power because of thermal noise constraints.

4.4 POSC Prototypes

Two prototype devices, referred to as the POSC1 and POSC2 were built to demonstrate the pipelined oversampling concept. They are based on a digital integration architecture and are similar to design 3 in both their implementation and their targeted performance. However, the prototypes differ from design 3 in two ways. These modifications are described below.

4.4.1 Dual Pipeline Configuration

The subpipelined digital integration approach described above has a latency of $P(k+1)$. The $k+1$ term arises because the number of substages per conversion block is one greater than the number of ADC bits. The 2-bit signal w_k is completed during substage 3, added to d_1 during substage 4, and provided to DACs in the next conversion block. None of the operations occurring after substage 4 in block n are needed for operation in block $n+1$. Because of this, dual pipelines can be used to reduce latency.

Figure 4-19 shows such a configuration that is used for the prototypes. The pipeline is divided into two halves, containing even and odd blocks. Each block contains 8 substages. Signals flow simultaneously through both pipelines. Even and odd blocks are offset from one another by 4 clock cycles so that operations in substage 1 of block $n+1$ coincide in time with those in substage 5 of block n.

One advantage of this approach is that latency is reduced nearly in half, to $(P+1)(k+1)/2$. A second

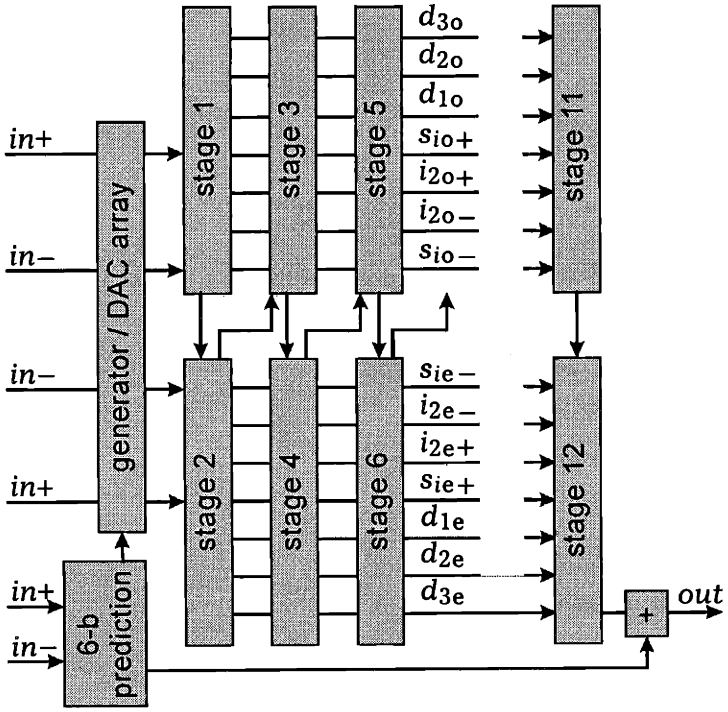


Figure 4-19. Prototype dual pipeline configuration. Signals flow simultaneously through even and odd pipelines.

advantage is that mismatches that are inverted between the even and odd blocks appear as high frequency error, as signals alternate back and forth between the pipelines, and are attenuated by the decimator. A third advantage of a dual pipeline is a 2-X reduction in power. There are a total of

$$2(k+1)\sum_{i=1}^{P}[i(i-1)+1] \approx \frac{2P^3}{3}(k+1) \qquad P>>1 \tag{4-32}$$

delay elements in the signal channels of a single pipeline configuration. The total number in a dual pipeline structure is

$$(k+1)\sum_{i=1}^{P}\left[\frac{(i-1)}{2}(2i-1)+1\right] \approx \frac{2P^3}{3}(k+1) \qquad P>>1\,. \tag{4-33}$$

Both approaches have approximately the same number of delay elements. But since a dual pipeline configuration generates twice as many packets at its input, each packet can be made half as large, and total capacitance and power are reduced in half.

4.4.2 Two-Stage DAC

The second prototype modification, that of a two-stage DAC, is not a recommended approach, but was done because of constraints imposed by a 1.2-μm process. Without this modification, a 6-bit range is required from DACs in a single pipeline configuration and a 7-bit DAC range is needed for dual pipelines. As implemented in the prototype, a two-stage DAC allows 4 bits to be removed from the DAC in each conversion block.

One source of these process constraints is geometries that are required for CCD registers. The speed of charge transfers is dominated by diffusion time constants near the end of the transfer operation, when only a small fraction of the original charge remains behind. This time constant increases quadratically with CCD gate length and sets an upper limit on the allowable pitch of each pipeline stage and the amount of CMOS circuitry that can be included in it for a given process geometry. Barrier and storage gate lengths of 2.4 and 4.6 μm were chosen for the prototype based on its desired speed. The corresponding pitch of each register stage is 14 μm. The need for 7 digital inputs and 5 digital outputs for each conversion block is difficult to accommodate in this pitch.

The number of digital inputs to each conversion block is reduced in the prototypes by moving some of the DAC operations forward to the beginning of the converter. Because of feedback, the input $\left(f_{bp} - f_{bm}\right)$ to the DACs in each conversion block has a slope that is approximately equal to the analog sample. Its value is

$$\left(d_1[n-1]+2w_k[n]\right) \approx nS_g/R_f\,. \tag{4-34}$$

If an estimate E_s of S_g/R_f is known at the beginning of the pipeline, then the same operation can be performed before the pipeline by subtracting a quantity approximately equal to S_g from each of the n

packets used for block n. The result, shown in Figure 4-19, is used for the prototypes. A 6-bit prediction is computed at the beginning of the pipeline, using a quantizer similar to the feedforward ADC in each block. This prediction is used to control a thermometer code array of 256 DAC elements across all channels that subtract from the input samples before they are passed forward to the remainder of the pipeline. The resulting signal channel packets are

$$s_i[n] = S_g - E_s R_f \ . \tag{4-35}$$

And four bits are removed from each DAC in the pipeline.

The resulting configuration is similar to a two-stage converter with error correction. High accuracy is not needed from the prediction because the range of DACs in the pipeline is 2-bits wider than a prediction LSB. Any error in the original prediction can be reversed in the decimator. The configuration differs from two-stage converters in that signals are transferred between the two stages with unity gain. As a result, minimum resolution of the oversampling blocks is not changed by the front end. The unity interstage gain is accurately controlled because it is determined by charge conservation. DAC accuracy requirements are increased in this approach because fewer DAC elements are used throughout the converter and this reduces the degree of averaging that is experienced.

4.5 Prototype Measurements

4.5.1 Performance Targets

The POSC prototypes were developed as part of a program to integrate a scientific CCD imager with an oversampling A/D converter. A parallel effort in this program involved development of the imager and a CCD/CMOS process for this system. The ultimate system objective is to achieve a fully integrated, high performance imaging system with reduced power and weight. Performance specifications for the A/D converter were chosen based on the needs of this system.

The imager readout path in typical imaging systems is illustrated in Figure 4-20(a). A charge packet, originating from a CCD register, is first dumped onto node v_1, containing a capacitor that has been preset. The voltage on v_1 falls in proportion to the packet size, as shown in (b). The signal is then buffered and amplified. Since the reset operation introduces both thermal noise and clock feedthrough, correlated-double-sampling circuitry is used to produce the v_2 waveform in (c), which contains only the difference between the signal's preset level and sampled level. Finally, the resulting voltage is digitized by an A/D converter.

The charge packets in such a system are discrete-time quantities. Each packet represents the value of a different image pixel and should be processed in the A/D converter independent of its neighboring pixels. For reasons described in section 1.1.3, conventional oversampling converters do not perform Nyquist

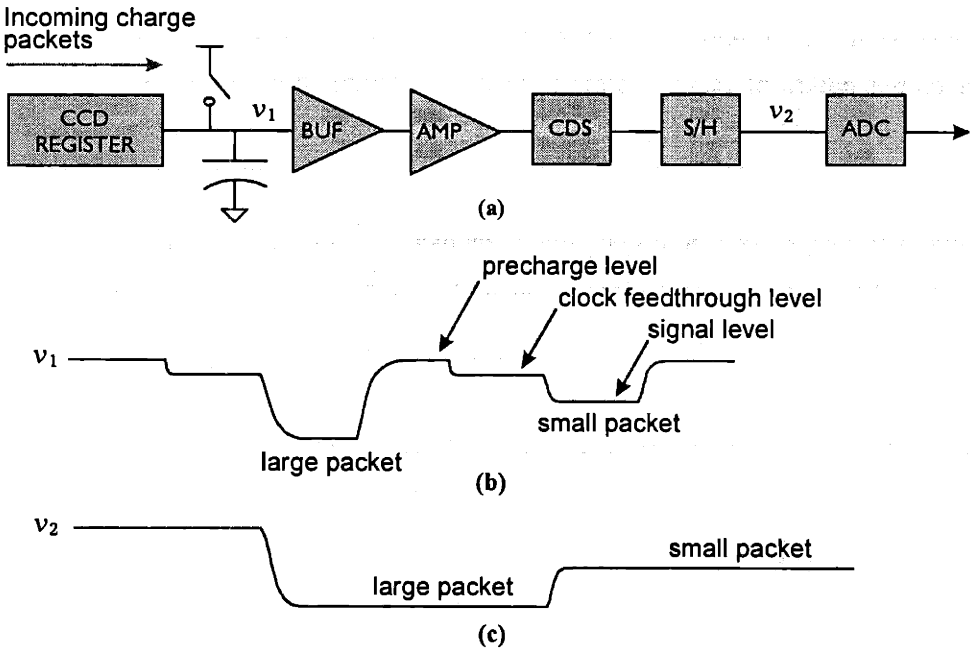


Figure 4-20. (a) Typical imager readout path. (b) Signal after charge-to-voltage conversion. (c) Signal after correlated-double sampling.

sampling and are not well suited to this type of data. However, pipelined oversampling treats each sample independently and can be used in such a system.

The A/D converter requirements for this application differ from those of general purpose applications. The end result, an image to be displayed, should not be significantly degraded by the quantization process. A/D converter requirements needed to achieve this are 72 dB dynamic range and DNL < ±0.5 LSB @ 12 bits. INL specifications are not critical because low-order distortion of a few percent is introduced by the rest of the charge readout path. A data rate of 10 MSPS is needed to support the desired number of pixels and frame rate and less than 10 mW/MSPS power dissipation is required. Finally, the converter must be capable of this performance when integrated with an imager, where large amounts of substrate noise are present.

4.5.2 Experimental Test Setup

Testing was done using an automated A/D testbed, with synchronized signal and clock sources, at MIT Lincoln Laboratory. A block diagram of the printed circuit board, used to support the converter, and the surrounding test setup is shown in Figure 4-21. The input waveform is supplied by a sinewave generator and passed through a signal conditioning block, which is used to remove source harmonics. The result is converted to differential through a transformer and passed through an anti-alias filter before entering the converter. A 4X clock input, supplied by a second sinewave generator, is offset, using a passive level translator, to CMOS levels. All clock and control signals needed by the chip are generated internally from this signal. A set of 5-V and 3.3-V power supplies are used for analog and digital circuits. Separate voltage regulators are also used to power the CCD driver circuitry. This permits independent measurement of the power of these circuit groups. For debugging purposes, on chip source follower pullups are placed at the CCD register outputs. Using them, charge generator operation and CCD transfers can be verified.

A computer controls the setup of all test equipment by means of the GPIB. It also controls the buffer and data acquisition unit, which is used to collect sequences of events and transfer them to memory. When the buffer is full, the collection process is stopped. The data may then be retrieved by the computer from memory at a 300kHz rate in blocks of arbitrary size up to 256K words.

Outputs from the POSC's 6-bit front-end and 12-bit back-end are passed through a wire wrap board which performs the necessary overlap and add. This is done in hardware, rather than in the computer, because the data buffer interface supports only 16 bits. The chip supports 3 modes of operation, in which words are overlapped by 1, 2, and 3 bits.

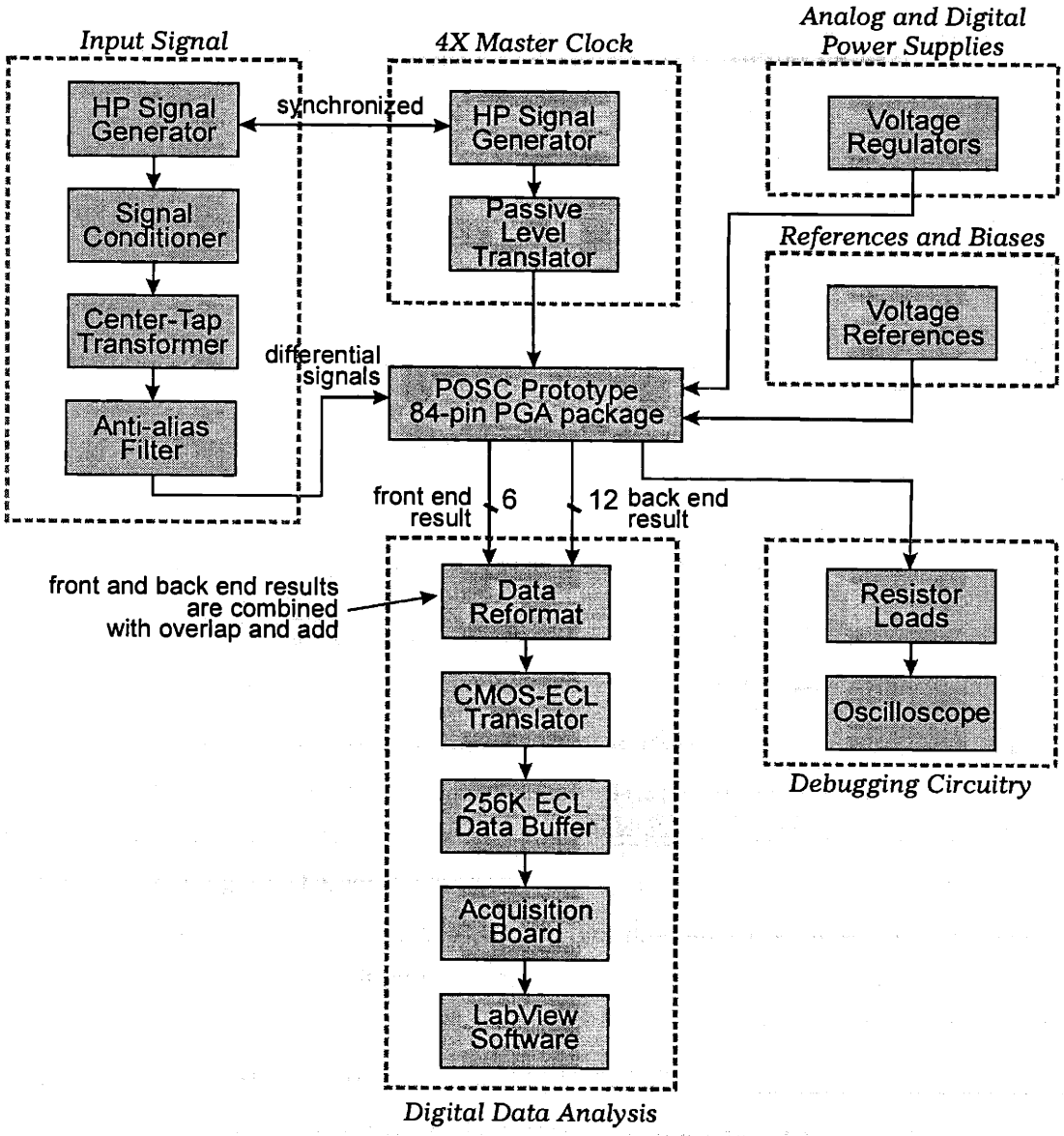


Figure 4-21. Block diagram of the POSC support board and A/D converter testbed.

4.5.3 Measured Performance

The POSC1 and POSC2 designs differ from each other only in their fabrication. Since a goal of these prototypes was to demonstrate that high-performance CCD devices are achievable using only standard CMOS processing, both of them were fabricated in commerical foundaries. The POSC1 was built in a 1.2-μm double-poly double-metal process from Orbit Semiconductor. The POSC2 was built in a 0.35-μm double-poly double-metal TSMC process from MOSIS. A comparison between these processes is listed in Table 4-5. The POSC2 was a straight 2X shrink of the POSC1, with no signficant design changes, and

	Orbit	TSMC
Minimum Gate Length	1.2 μm	0.35 μm
Minimum Design Rule Used	1.2 μm	0.6 μm
Gate Oxide Thickness	225 A	Both 150 and 90 A
Maximum Gate Voltage	5 V	Both 5 and 3.3 V
Interpoly Oxide	700 A	
Metal Levels	2	2
Poly1 Threshold	0.9 V	0.5 V
Estimated Poly2 Threshold	1.5 V	2.0 V
Planarized	No	Yes

Table 4-5. Processes used for the POSC1 and POSC2 prototypes.

therefore, uses a 0.6-μm minumum process geometry. To minimize impact on the design during the shrink, the thicker 5-V gate oxide and 5-V threshold implant option was used throughout the chip.

Parasitic 2nd-poly is used to form overlapping CCD structures, despite the fact that neither process supports such use. The topology of these structures, which are also covered by metal-1 and metal-2 crossings, is particularly a concern in the 1.2-μm process, which is not planarized. Since neither process allows poly1-poly2 overlaps, the potential for poly stringers in the CCD regions also exists. A third nonstandard aspect of the designs is in their poly design rules. Poly1-poly2 overlaps of 0.6 μm and 0.3 μm are used for the designs, whereas process design rules are 1.0 μm and 0.8 μm. Despite these concerns, the fully functional yield of both devices was better than 90%.

The prototypes contain 6200 CCD and 8800 CMOS devices in their modulators and 22,600 CMOS devices in their decimators. The modulator, decimator, and all necessary support logic are integrated onto a single chip. Major functional blocks are indicated on the die micrograph in Figure 4-22. The POSC1 device is a large chip, with a dimension of 6.9 mm per side. However, the POSC2 is about 3 times smaller, at 4 mm per side. Although circuit geometries in the POSC2 are shrunk by precisely 2X, its pad frame is not reduced by this same factor.

Even and odd halves of the modulator pipeline are indicated in the figure. The decimator is configured as a single pipeline of 12 stages because of the large number of signals that transfer between its adjacent stages. Separate control logic blocks, operating at 5V and 3.3V are used for the modulator and decimator circuits. A total of 106 sequential delays, each containing 4 CCD gates, are present in the longest CCD channels. Despite the large number of CCD gates, less than 5% of the total chip area is occupied by CCD elements.

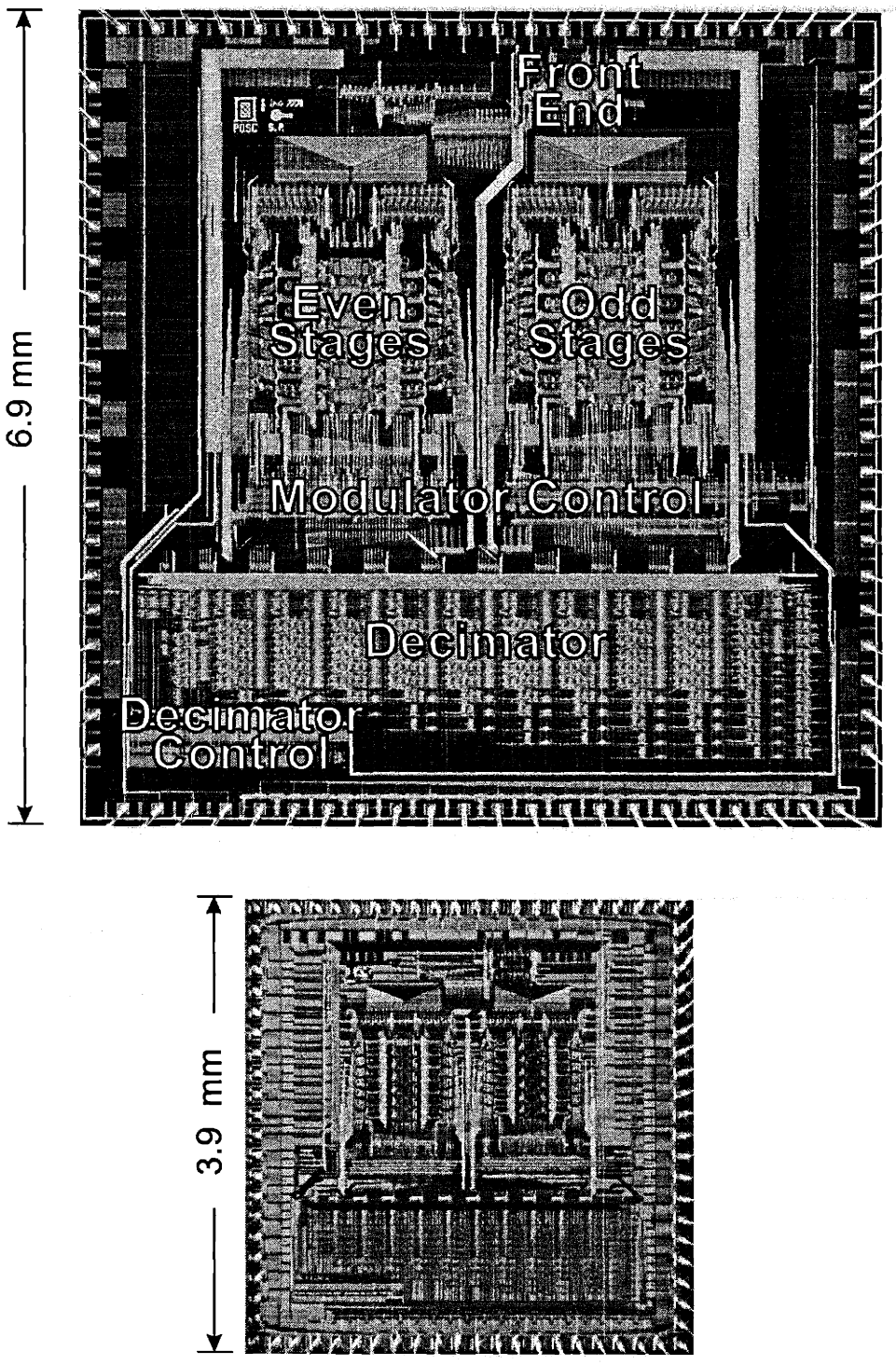


Figure 4-22. Micrographs of the POSC1 and POSC2 prototypes. The POSC2 is about 3 times smaller.

Measured prototype performance is summarized in Table 4-6. POSC1 measurements were done at an 18-MHz sampling rate with an input sinusoid of approximately 8 MHz. A 16,384-point spectral response plot is shown in Figure 4-23. The worst harmonic is the 3rd at 78 dB below the fundamental. The rms variation in the front-end DAC element values is calculated to be about 0.25%. Nonlinearity due to pipeline DACs is reduced because some of it is translated into wide-band noise by inherent dynamic element matching. The presence of higher-order harmonics, with slowly decreasing magnitudes, near the others, was anticipated for this design because of its sensitivity to DAC mismatches. Among the devices tested, SFDR ranges from 78 dB to 73 dB due to incomplete cancellation of even-order harmonics.

Figure 4-24 shows POSC1 SNR, with a peak of 74 dB, and SNDR, with a peak of 71 dB, as a function of input power. Ratios are computed over a 9 MHz bandwidth. Since the converter produces a total of 16 bits, theoretical quantization noise is not a significant contributor. Sampling noise is reduced in the design because 32 separate samples are captured at the converter input. Other significant noise components may include wide-band dynamic element matching noise that is passed by the decimator, coupling in some nondepleted circuits, and surface traps in the CCDs. Thermal noise in the input charge generators is not a limiting factor because the total generator capacitance, spread across 16 circuits, on each differential input is 4 pF.

	POSC 1	POSC 2
Data Rate	18 MSPS	30 MSPS
Peak SFDR	78 dB	70 dB
Peak SNR / SNDR	74 / 71 dB	66 / 63 dB
DNL / INL @ 13 bits	±0.15 LSB / ±1.0	±0.25 LSB / ±2.5
Analog / Digital Supply Voltage	5 V / 3.3 V	5 V / 3.3 V
Input Range	2 V p-p	1.5 V p-p
Power	324 mW @ 18 MHz	230 mW @ 30 MHz
Process	Commercial CMOS	Commercial CMOS
Design Rule	1.2 μm	0.6 μm
Yield	22 of 24	26 of 27

Table 4-6. Summary of measured performance for the POSC1 and POSC2 prototypes.

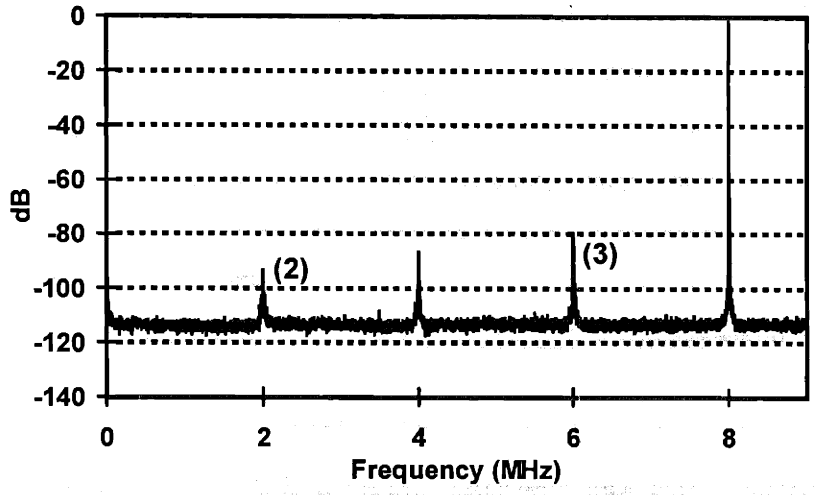


Figure 4-23. Measured spectral response at an 18-MHz sampling rate with an input near 8 MHz. The worst harmonic is the 3rd.

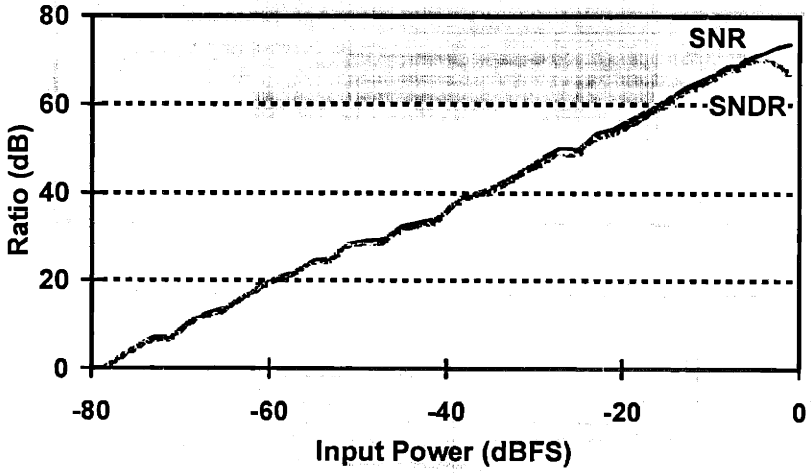


Figure 4-24. Measured SNR and signal-to-noise plus distortion over a 9-MHz bandwidth.

Integral and differential nonlinearity plots in Figure 4-25 were generated using histogram techniques on an 8-MHz input.

POSC2 measurements, listed in Table 4-6, were done at a 30-MHz sampling rate on an input of approximately 13.3 MHz. A few factors are thought to contribute to the reduced performance of this device. First, capacitance values and CCD well capacity are reduced by the shrink, resulting in smaller signal charge packets throughout the device. However, charge-referred resolution of the comparators, which is determined by mismatches, is not improved. Second, unequal scaling of CCD and polysilicon capacitances between the two processes result in incorrectly sized wells. Despite the reduction of capacitance values in this device, thermal noise in the input charge generators is not a limiting factor, because the total capacitance, spread across 16 circuits, on each differential input is 1.7 pF.

The prototype devices operate with four clock phases. CMOS clocks must be nonoverlapping, CCD clocks must be overlapping, and each of these must be synchronized with the others. Generating and distributing these signals is a limiting factor in prototype speed. Since the designs do not contain any static circuits, settling issues are not a concern. Slew rate limitations in the nondepleted CCD circuits

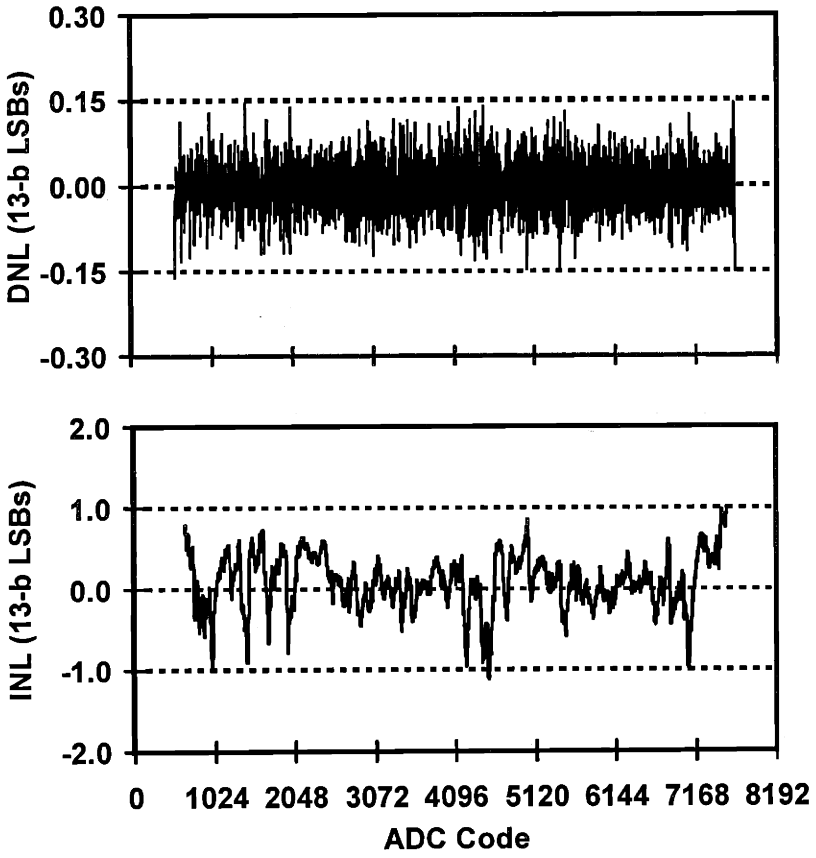


Figure 4-25. Measured integral and differential nonlinearity.

become evident when the POSC2 is operated up to a speed of 50 MHz. In both devices at their maximum speed, speed-dependent CCD charge transfer inefficiency is thought not to be significant because performance is unchanged at lower operating frequencies.

At a given operating voltage, POSC power scales linearly with sampling rate since its circuits are strictly dynamic. At full speed the POSC1 operates from 5V, 4V, and 3.3V for analog CMOS, CCD, and digital CMOS, respectively. It consumes 324 mW at 18 MHz, but at a reduced speed of 10 MHz, voltage levels can be reduced to 4V, 3.3V, and 3.3V and power is reduced to 122 mW. The POSC2 operates from 5V, 5V, and 3.3V for analog CMOS, CCD, and digital CMOS and consumes 230 mW at 30 MHz. Its CCD clock voltages are higher because of an increase in 2nd-poly thresholds.

Table 4-7 shows a breakdown of the power for each device by circuit group. Between the POSC1 and POSC2, total gate capacitance is reduced by 2.6X because oxide capacitance per area is increased by 1.5 but areas are reduced by 4. Total poly-poly capacitance is reduced by 2.3X because capacitance per area is increased by 1.75 but areas are reduced by 4. Analog CMOS power is reduced by 3X, which is slightly more than expected, probably because parasitic loading was reduced by more than 3. CCD power is reduced by less than 2X in the POSC2 because of its higher clock voltages. Digital CMOS power is reduced by slightly less than expected at 2.3X.

	Analog CMOS	CCD	Digital CMOS
POSC1	65%	20%	15%
Total = 18 mW/MSPS	11.7 mW/MSPS	3.6 mW/MSPS	2.7 mW/MSPS
POSC2	50%	35%	15%
Total = 7.7 mW/MSPS	3.9 mW/MSPS	2.6 mW/MSPS	1.2 mW/MSPS

Table 4-7. Breakdown of power by circuit group in the POSC1 and POSC2 prototypes. Entries list percent of total power and absolute power per frequency.

5 CCD Performance Limits

5.1 Dominant Noise Sources in CCD Devices

CCD signal processing devices typically contain a path such as that shown in Figure 5-1. Device inputs are rarely in charge packet form. Instead, signals entering the chip are in the form of photons, voltages, or currents. An initial input stage is used to translate these signals into charge packets for use by the remainder of the device. Once signals are in the charge domain, they are processed over many stages using a combination of fully-depleted and nondepleted CCD-based circuits. There are advantages to preserving the charge-domain nature of these signals and intermediate charge-to-voltage translation stages are generally avoided. However, the output from a CCD device must ultimately be translated back into a format, such as a voltage or current, that can be driven off chip.

The charge-domain portion of the path in Figure 5-1 is typically not a significant determinant of overall SNR. Its fully-depleted CCD operations do not contribute thermal noise or shot noise. Its nondepleted CCD operations do contribute thermal noise, but the magnitude of this noise is easily kept small. For a parasitic node capacitance of C, thermal noise in these circuits is proportional to $C^{1/2}$, the signal is independent of C, and SNR is proportional to $C^{-1/2}$. SNR, power, and speed are simultaneously improved by a reduction in C.

Unfortunately, this low-noise, low-power capability of CCDs is not fully realized for a configuration, such as that in Figure 5-1, because of constraints introduced by interface circuits at the device input and output. Two possible input circuits are shown in the figure. In the first, charge packets are generated by the absorption and collection of incoming photons. In the second, charge packets are generated from an incoming voltage by means of a charge generator. Both of these techniques have a noise level that is proportional to the square root of the charge packet being generated and an SNR that is proportional to

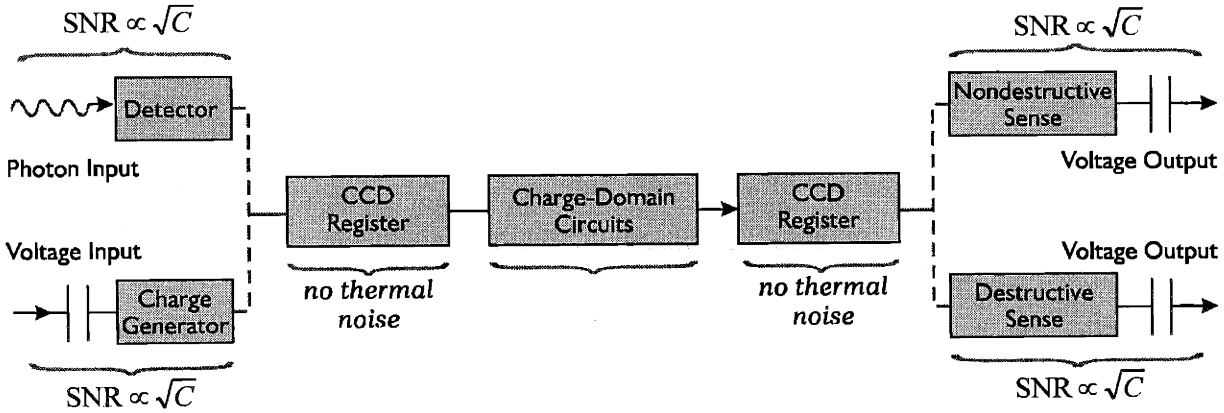


Figure 5-1. The charge-domain signal path in a CCD device is preceded and followed by circuits that interface to voltage, current, or photon signals.

$C^{1/2}$. Two possible outputs, destructive and nondestructive sensing circuits, are shown in the figure. Both of these translate charge to a voltage across a capacitor. When this capacitor includes an MOS gate, the output is available in either voltage or current form. Thermal noise in both of these circuits is proportional to the square root of the charge packet being sensed and, for a sensing capacitance of C, SNR increases as $C^{1/2}$.

Overall device noise is the sum of contributions from these interface circuits, and is largely independent of the number of charge-domain operations performed between input and output. For a given SNR, interface circuit capacitances are fixed. This in turn set charge packet sizes throughout the remainder of the device. CCD storage capacity, area, and power follow directly from charge packet sizes and are, therefore, directly related to SNR.

5.2 CCD Scaling

Minimum Gate Geometries

The layout of a CCD well is determined by its desired storage capacity and by process geometry limits. Figure 5-2 shows an example CCD register layout. Barrier gates are formed from poly2 and storage gates are formed from poly1. Barrier gates do not hold charge and are made as short as is allowed by the minimum poly1 spacing. In a typical process, barrier length, L_b, is approximately equal to λ, where λ denotes the minimum process geometry. Storage gates do hold charge and their length is chosen based on desired speed and transfer efficiency. At low speed or low resolution, storage gate lengths can be made large to achieve the necessary storage capacity. At high speed or high resolution, the minimum storage gate length, L_S, is used. L_S equals the minimum poly2 spacing plus twice the minimum poly1-poly2 overlap and is typically 4λ.

When charge packets are large, the width of CCD gates is chosen based on their length and their storage capacity requirements. For sufficiently small charge packets, storage capacity requirements can not be met. The minimum allowable width of active area definition sets a lower limit on gate width, W_S, and storage capacity, C_S. W_S is typically 4λ. Because of these constraints, the smallest storage capacity that can be achieved is

$$C_S = L_S W_S C_a \left(\frac{R+1}{R} \right) , \tag{5-1}$$

where C_a is the gate-to-channel capacitance per unit area and R is as defined in (3-154). A well with this capacity holds a charge packet of size

$$Q_S = L_S W_S C_a V_g \tag{5-2}$$

when it is full.

Charge packets smaller than that in (5-2) could potentially be accommodated by oversized wells with an

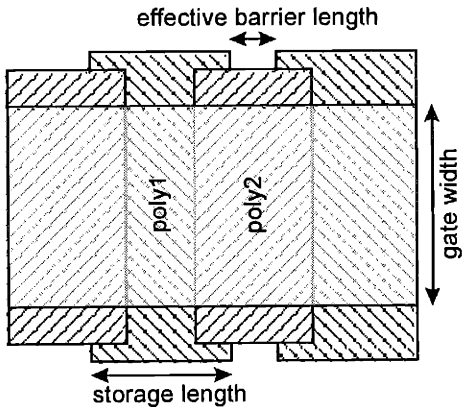


Figure 5-2. Example CCD register layout. Gate geometries are set by desired storage capacity and by process geometries.

area of $L_s W_s$. Since CCD power depends primarily on the amount of charge being transferred, rather than on storage capacity, only a minor increase in power would result. However, two factors make this approach undesirable. First, charge transfer efficiency and speed are significantly degraded when wells are oversized and mostly empty. This effect is described in section 5.4.1. Second, process imperfections provide a practical lower limit on capacitance values that can be reliably formed and accurately matched.

For these reasons, the following analysis assumes that the smallest charge packet size that is available from a process is that given by (5-2). This translates into a lower bound on the power dissipation of a CCD well. Combining (5-1), (3-156) and (3-157) yields a minimum power dissipation of

$$\frac{P}{f} = \frac{1}{2} L_s W_s C_a \left(\frac{R+2}{R+1} \right) V_g^2 \quad . \tag{5-3}$$

When process geometries are scaled by a factor of λ, the terms L_s, W_s, and V_g are decreased by λ, C_a is increased by λ, and the resulting power is decreased by a factor of λ^3.

Constant-Charge-Density Scaling

Charge packet sizes and charge density determine the SNR and speed of CCD circuits. These parameters should be preserved when circuits are scaled. Figure 5-3 illustrates such an approach, referred to as constant-charge-density scaling. In constant-field process scaling, geometries are multiplied by a factor of λ, where a λ of less than 1 corresponds to a geometry reduction. Between Figure 5-3(a) and (b) the gate

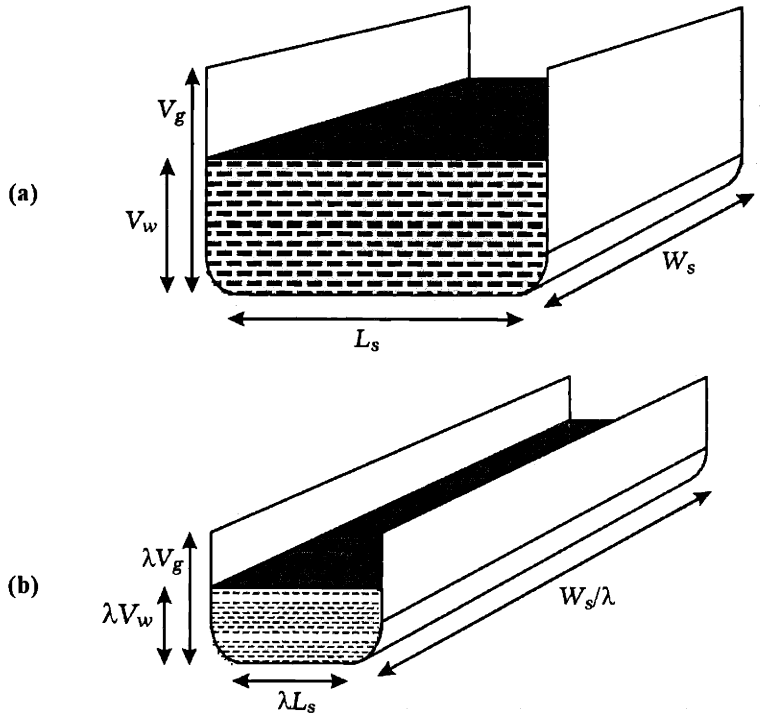


Figure 5-3. Scaling of a CCD well. Charge density and total charge packet size are preserved during scaling.

length is decreased from L_s to $L_s\lambda$, the gate oxide capacitance is increased from C_{ox} to C_{ox}/λ, and voltages are reduced by a factor of λ. To preserve the total charge packet size, the gate width of W_s in (a) is increased to W_s/λ in (b). Both wells then have the same charge density, the same area, and the same total charge packet size.

5.3 Power and Area

The power dissipation of CCD circuits was derived in section 3.8. Power is determined primarily by the product of storage capacity and how full the wells are. In other words, power is nearly proportional to the total stored charge and is only weakly dependent on process parameters. Barrier gates, which do not hold charge, have an effective capacitance equal to their empty well capacitance and do not contribute much power. Storage gates, which do hold charge, have a much larger effective capacitance and dominate overall CCD power dissipation.

As described above, storage capacity of the wells in a CCD device is fixed by the SNR relationship of input and output circuits that interface to the charge, voltage, or current domains. Given this storage capacity, power and area follow directly from (3-157) and (3-158). Thus, for a given set of interface circuits, there is a firm relationship between SNR and CCD power dissipation. These relationships are described below for four categories of interface circuits.

5.3.1 Input Circuits

Charge Generation

One technique for generating charge packets at the input to a CCD device is voltage-to-charge conversion. An example of such a charge generator is the DDS circuit described in section 3.5. A common characteristic of charge generation circuits is that they translate voltage to charge through a capacitance C, which consists of either a poly-poly capacitor, a CCD well, or an MOS channel. In all cases, charge-referred thermal noise is proportional to $C^{1/2}$, the signal level is proportional to C, and SNR is proportional to $C^{1/2}$. In a general case, where m independent sources of thermal noise are introduced per cycle, the relationship between input capacitance and SNR is

$$C = \left(\frac{\text{SNR}}{V_s}\right)^2 8mkT \ . \tag{5-4}$$

SNR in this expression is defined for a sinusoidal input with a peak-to-peak voltage swing of V_s. The value of m for a fill-and-spill charge generator is 1, whereas that for a DDS charge generator is either 2 or 3.

CCD wells receiving this signal must have sufficient storage capacity to hold the generated charge. If charge is stored across a voltage of V_w, then a storage capacity of

$$C_s = \mathrm{SNR}^2\left(\frac{8mkT}{V_s V_w}\right) \tag{5-5}$$

is required from each well. The effective capacitance of these wells translates into an average power per well of

$$\frac{P}{f} = \mathrm{SNR}^2(4mkT)\left(\frac{V_g}{V_s}\right)K_{eff}\ . \tag{5-6}$$

The corresponding area per well is

$$A = \mathrm{SNR}^2\,\frac{8mkT}{V_s V_w C_a}\left(\frac{R}{R+1}\right). \tag{5-7}$$

When a process scale of λ is applied, power remains largely unchanged, whereas area is increased by a factor of $1/\lambda$.

Photon Collection

Shot noise is introduced when charge packets are generated by the absorption of photons. Since the absorption of each photon is an independent process, shot noise equals the square root of the number of signal electrons. For a sinusoidal input, with a range from 0 to N_s electrons, SNR is given by

$$\mathrm{SNR} = \frac{N_s}{2\sqrt{2}\sqrt{\frac{1}{2\pi}\int_0^{2\pi}\frac{N_s}{2}(\sin x + 1)dx}} = \frac{\sqrt{N_s}}{2}\ . \tag{5-8}$$

A full scale signal of

$$N_s = 4\mathrm{SNR}^2 \tag{5-9}$$

electrons is required in each well and a storage capacity of

$$C_s = \frac{4\mathrm{SNR}^2 q}{V_w} \tag{5-10}$$

is needed to store this signal. At a given SNR, lower limits on power dissipation and well area are

$$\frac{P}{f} = 2\mathrm{SNR}^2\, V_g\, q\, K_{eff} \tag{5-11}$$

and

$$A = \frac{4\mathrm{SNR}^2 q}{V_w C_a}\left(\frac{R}{R+1}\right). \tag{5-12}$$

When a process scale of λ is applied, power is reduced by a factor of λ, while area remains unchanged.

Thermal Generation

Thermal generation causes additional carriers to be added to signal packets during the time that they are held in CCD wells. The resulting offset of charge is not a problem when CCD circuits are implemented

differentially because it is common-mode. However, thermal generation introduces shot noise, which impacts SNR and necessitates larger signals. Storage capacity must also be increased to hold the offset charge.

The total number of electrons that are thermally generated within a CCD well is expressed as

$$N_g = \frac{J_g}{2fq}\left(\frac{C_s}{C_a}\right)\left(\frac{R}{R+1}\right), \tag{5-13}$$

where J_g is the thermal generation current density and the factor of 2 accounts for the fact that charge spends only half of a clock cycle within the well. Variable f represents an equivalent operating frequency equal to the inverse of the total time it takes for a signal to traverse the device.

SNR for a sinusoidal input with a peak-to-peak range of N_s electrons equals

$$\text{SNR} = \frac{N_s}{2\sqrt{\frac{J_g}{fq}\left(\frac{C_s}{C_a}\right)\left(\frac{R}{R+1}\right)}}. \tag{5-14}$$

Wells must have sufficient storage capacity to hold the full-scale signal N_s plus the additional thermally generated offset. This storage capacity determines the area of the well, which in turn determines the total generated charge. The relationship governing storage capacity is

$$\sqrt{C_s}\left(1-\frac{J_g}{2f C_a V_w}\frac{R}{R+1}\right) = 2SNR\sqrt{\frac{J_g q}{f C_a V_w{}^2}\frac{R}{R+1}}. \tag{5-15}$$

Using the approximation

$$\frac{J_g}{2f C_a V_w}\frac{R}{R+1} << 1, \tag{5-16}$$

which states that the offset contributes negligibly to required storage capacity, this expression is simplified to

$$C_s = 4\text{SNR}^2\frac{qJ_g}{f C_a V_w{}^2}\frac{R}{(R+1)}. \tag{5-17}$$

Average power corresponding to this storage capacity equals

$$P = 2\text{SNR}^2\frac{qJ_g}{C_a}\left(\frac{V_g}{V_w}\right)\left(\frac{R}{R+1}\right)K_{eff}. \tag{5-18}$$

In contrast to other noise sources described in this section, the power limit introduced by thermal generation is measured in W, rather than W/Hz. Total power is independent of frequency.

The corresponding well area is

$$A = 4\text{SNR}^2\frac{qJ_g}{fC_a{}^2}\frac{1}{V_w{}^2}\left(\frac{R}{R+1}\right)^2. \tag{5-19}$$

When a process scale of λ is applied, power is decreased by a factor of λ and area is increased by a factor of $1/\lambda$.

Nondestructive Charge Sensing

Nondestructive charge sensing operations include both a precharge and a sensing phase. A floating CCD gate is first precharged, when its underlying well is empty of charge. Then a charge packet is transferred underneath the floating gate and the signal is sensed. The thermal noise of this operation depends on the precise circuit that is used. However, SNR can generally be expressed as

$$\mathrm{SNR} = \frac{Q_s}{2\sqrt{2}\sqrt{kT\left(C_{eff}(\mathrm{empty}) + mC_{eff}(\mathrm{full})\right)}}, \tag{5-20}$$

for a sinusoidal signal with a peak-to-peak range of Q_s. The mean-square thermal noise

$$kTC_{eff}(\mathrm{empty}) = kT\frac{Q_s}{V_w}\frac{R}{(R+1)^2} \tag{5-21}$$

is introduced during precharge and is always present. It is proportional to the effective capacitance of an empty CCD well. The mean-square noise

$$mkTC_{eff}(\mathrm{full}) = mkT\frac{Q_s}{2V_g}K_{eff} \tag{5-22}$$

is introduced during sensing and is proportional to the effective capacitance of a full well. The specific sensing circuit that is used determines the factor m and the magnitude of this noise. In the simplest such circuit, where the sensed signal is received by an MOS gate, m=0. In circuits, such as a DDS replicator, where the sensed signal is received through an MOS source or drain, m=1. Since the value of C_{eff} when the well is full is much larger than that when the well is empty, power and area limits are significantly reduced in circuit configurations with m=0.

Combining (5-20), (5-21), and (5-22) yields the expression

$$C_s = \mathrm{SNR}^2 8kT\left(\frac{R}{V_w{}^2(R+1)^2} + m\frac{K_{eff}}{2V_gV_w}\right), \tag{5-23}$$

which relates well storage capacity to SNR. The corresponding power per unit frequency is

$$\frac{P}{f} = \mathrm{SNR}^2 4kT\left(\frac{V_g}{V_w}\frac{R}{(R+1)^2}K_{eff} + \frac{m}{2}K_{eff}{}^2\right) \tag{5-24}$$

and the area per well equals

$$A = \mathrm{SNR}^2 8kT\left(\frac{R}{V_w{}^2(R+1)^2} + m\frac{K_{eff}}{2V_gV_w}\right)\frac{1}{C_a}\frac{R}{(R+1)}. \tag{5-25}$$

When a process scale of λ is applied, power remains unchanged and area increases as $1/\lambda$.

Destructive Charge Sensing

Like nondestructive sensing, destructive charge sensing includes both a precharge and a sensing phase. However, destructive sensing differs in that its capacitance is the same during both operations. The

thermal noise of a destructive charge sensing circuit depends on the precise circuit that is used. However, SNR can generally be expressed as

$$\mathrm{SNR} = \frac{Q_s}{2\sqrt{2}\sqrt{kTC(1+m)}} . \tag{5-26}$$

A sinusoidal input with a peak-to-peak range of Q_s is assumed and C represents capacitance of the sensing node. In the simplest such circuit, where the sensed signal is received by an MOS gate, no thermal noise is introduced during the sensing phase, and m=0. In circuits such as DDS sensing, where the sensed signal is received by an MOS source or drain, thermal noise is present during sensing, and m=1.

Small sensing capacitances are desirable because they increase the gain of charge-to-voltage translation. However, when the sensed signal must remain within an available voltage range, V_s, the minimum value of C_1 is limited to

$$C_1 = \frac{Q_s}{V_s} . \tag{5-27}$$

Combining (5-26) and (5-27) yields the expression

$$C_s = \mathrm{SNR}^2 8kT \frac{1}{V_s V_w}(1+m) , \tag{5-28}$$

which relates well storage capacity to SNR. The corresponding power per unit frequency is

$$\frac{P}{f} = \mathrm{SNR}^2 4kT\left(\frac{V_g}{V_s}\right)(1+m)K_{eff} \tag{5-29}$$

and the area per well equals

$$A = \mathrm{SNR}^2 8kT \frac{1}{V_s V_w}(1+m)\frac{1}{C_a}\frac{R}{(R+1)} . \tag{5-30}$$

When a process scale of λ is applied, power remains unchanged and area increases as $1/\lambda$.

5.3.2 Power and Area Limits

Figure 5-4 plots the power-SNR relationship for each of the input and output circuits described in section 5.3.1. The curve for each circuit represents a lower limit on power per unit frequency per CCD well when this circuit is present. Since power and area are proportional, the curves also show area versus SNR. Table 5-1 lists parameters used for the plot.

All curves increase as SNR^2. For the same SNR, the photon collection power requirement is 48 times larger than that for charge generation. Even at an effective frequency as low as 100 Hz, the thermal generation limit is negligible compared to the others. For the values of m in Table 5-1, charge generation noise is a more significant determinant of power than sensing. However, this is not always the case. For other combinations of m, sensing and generation limits may be comparable.

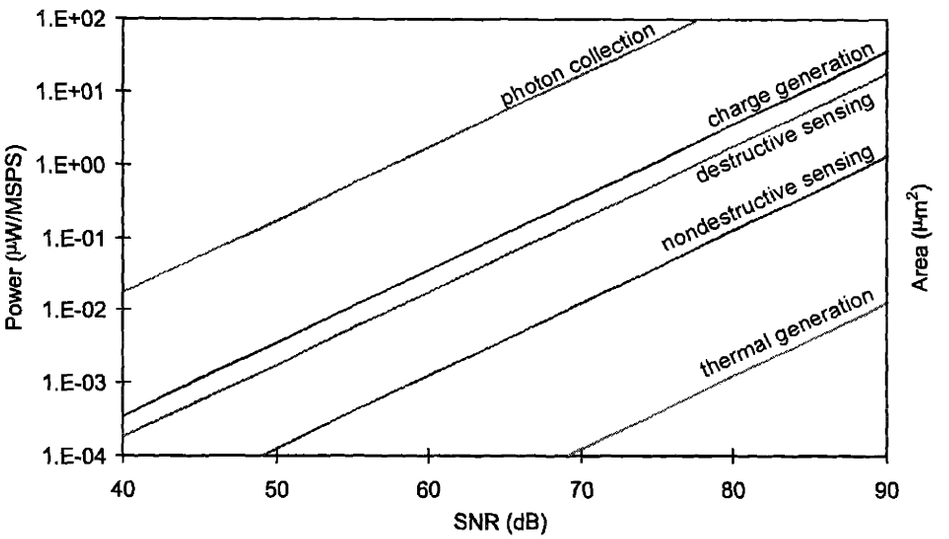


Figure 5-4. CCD power and area versus SNR for various input and output circuits. All curves increase as SNR^2.

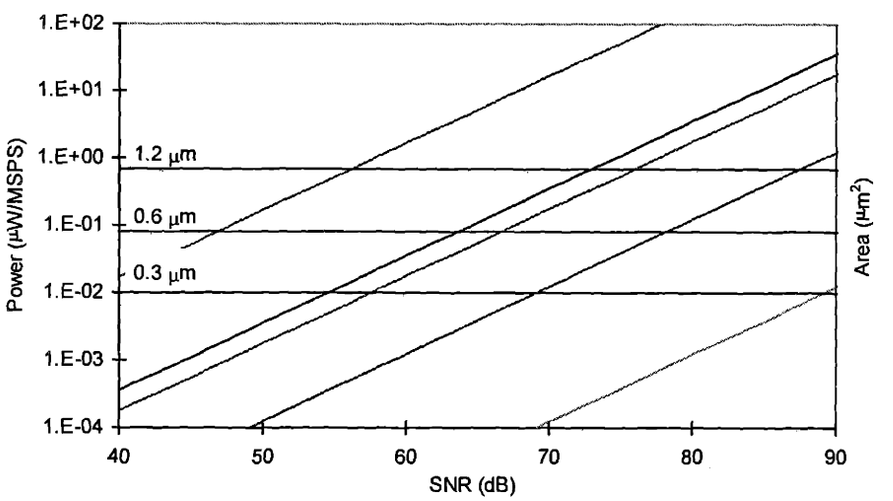


Figure 5-5. Process geometry limits. Power and area are limited by process geometries at low SNR

All Curves	$T = 27\,^{\circ}\mathrm{C}$ $R = 12$ $V_s = V_g = V_w$ $C_a = 1.5\,\mathrm{fF}/\mu\mathrm{m}^2$
Charge Generation	$m = 2$
Photon Collection	$V_g = 5\,\mathrm{V}$
Nondestructive Sensing	$m = 0$
Destructive Sensing	$m = 0$
Thermal Generation	$f = 100\,\mathrm{Hz}$ $J_g = 0.5\,\mathrm{nA}/\mathrm{cm}^2$

Table 5-1. Parameters used for power and area curves.

The relationships in Figure 5-4 are largely independent of process parameters and do not take into account minimum process geometries. In practice, process geometries limit CCD power to a minimum value, given by (5-3), that is independent of SNR. The process limits for 1.2-μm, 0.6-μm, and 0.3-μm geometries are superimposed on the graph in Figure 5-5. At high SNR, power is limited by circuit noise and decreases quadratically as SNR is reduced. Beyond a certain point, power is limited by process geometries and a further reduction in SNR provides no additional power improvement.

The curves at 60-dB SNR in Figure 5-5 are plotted as a function of process geometry, λ, in Figure 5-6. Constant-charge-density scaling is assumed. All circuit curves, except that for photon collection, are independent of geometry. The photon collection limit lies well above the others over the entire range. At large process geometries, power is limited by the process curve and decreases as λ^3 as geometries are reduced. Beyond a certain point, power is limited by circuit noise and further scaling provides no additional improvement. In the past, CCD power has decreased with geometry. However, beyond current process generations, power will be limited by circuit noise and will no longer improve with scaling.

A similar plot is shown in Figure 5-7 for area. Area is plotted as a function of process geometry, λ, for an SNR of 60-dB. All circuit curves, except that for photon collection, are proportional to $1/\lambda$. The photon collection curve is independent of geometry and lies well above the others over the entire range. At large process geometries, area is limited by the process curve and decreases as λ^2 as geometries are reduced. As processes are scaled beyond a certain point, area becomes limited by circuit noise and increases as $1/\lambda$. In the past, CCD area has decreased with geometry. However, as processes are scaled beyond current generations, area will increase.

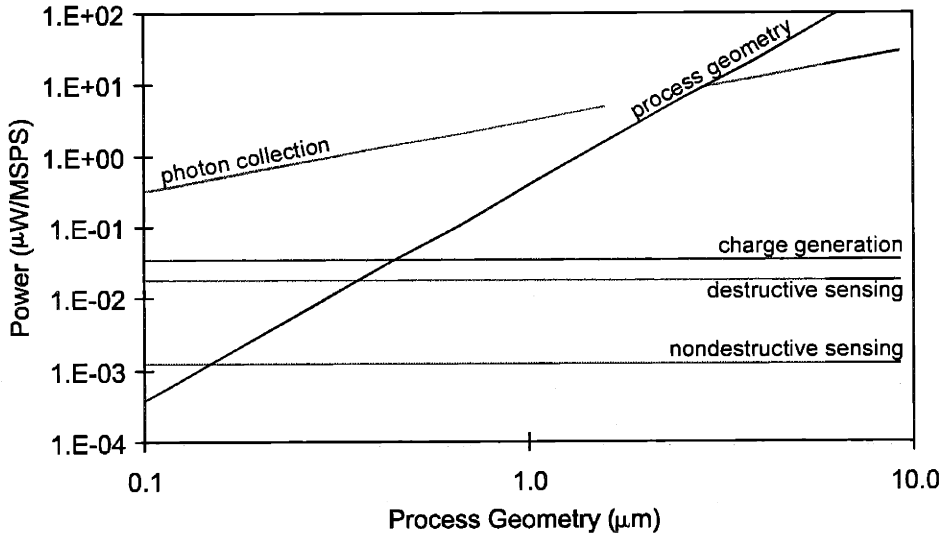


Figure 5-6. Circuit and process power limits versus process geometry for 60-dB SNR and constant-charge-density scaling.

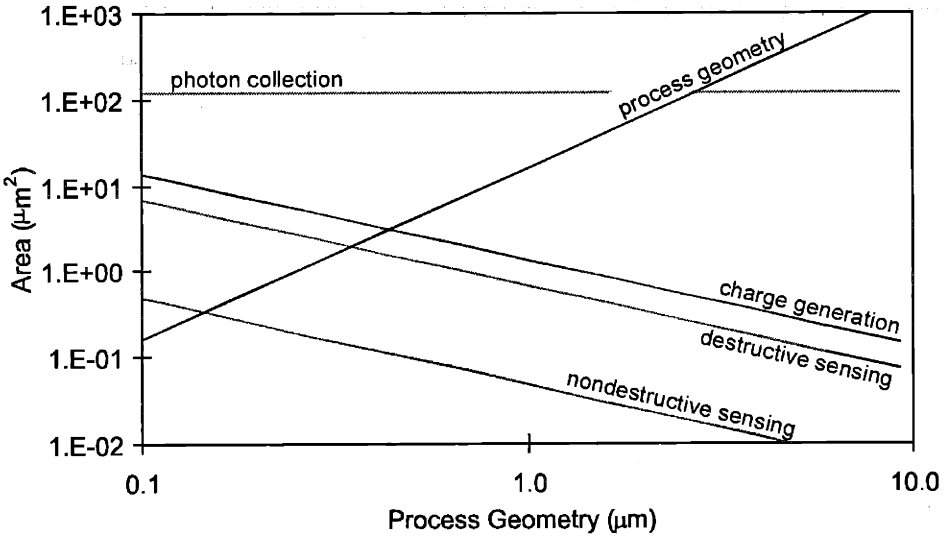


Figure 5-7. Circuit and process area limits versus process geometry for 60-dB SNR and constant-charge-density scaling.

5.4 Speed Limits

5.4.1 Charge Transfer Mechanisms

The transfer of charge out of a CCD well is governed by the relationship [39]

$$\frac{d\rho}{dt} = \frac{\mu}{C_{ox}}\frac{d}{dx}\left(\rho\frac{d\rho}{dx}\right) + \frac{kT}{q}\mu\frac{d^2\rho}{dx^2} + \mu\frac{d}{dx}\left(\rho E_f\right) \ , \tag{5-31}$$

where ρ represents electron density within the well, x is distance along the direction of charge transfer, and μ represents electron mobility. Three mechanisms govern the transfer; drift, diffusion, and fringing field assisted drift. These are illustrated in Figure 5-8. They correspond to the first, second, and third terms in (5-31), respectively. The speed of each of these charge transfer mechanisms and their relationship to process geometries is considered below. Numerical simulations of the expression in (5-31) are used to arrive at approximate speed-accuracy relationships. Three CCD gate lengths are considered. They are 4.8-μm, 2.4-μm, and 1.2-μm, and are assumed to correspond to processes with minimum geometries of 1.2-μm, 0.6-μm, and 0.3-μm. Constant charge density scaling is assumed, so that gate-oxide thickness and operating voltages are reduced along with geometry. Typical oxide thickness values and a constant electron mobility of 500 cm/V-sec are used for all simulations.

Drift

The drift term in Figure 5-8(a) is caused by the mutual repulsion of electrons in the well. Since charge density is constrained to be zero adjacent to the receiving well, current density increases with ρ/C_{ox}, which equals the well storage voltage. Drift is the dominant component of charge transfer initially, when a

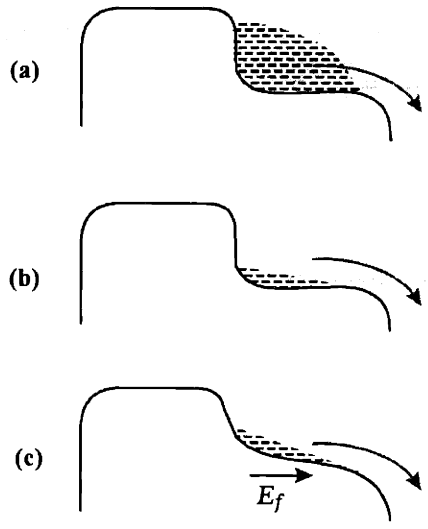


Figure 5-8. Three components of charge transfer. (a) Drift dominates initially. (b) Diffusion dominates at the end of transfer. (c) Fringing field assisted transfer improves speed at the end.

large number of carriers are present in the source well. However, it slows considerably as the transition progresses. Figure 5-9 illustrates this process for the example of a 4.8-μm long well. At time t=0, charge is uniformly distributed across the well. A large fraction of the original charge is removed rapidly by drift.

Figure 5-10(a) shows the density of untransferred charge, as a function of time, for transfers with only a drift component. Charge density is normalized to its initial value. The result in (a) corresponds to the case of a well that is full before the transition begins. The entire gate voltage is used to store charge and F=0.5. More than 95% of the charge is transferred by drift within the first few nanoseconds, but the transfer slows considerably thereafter. As the gate length is reduced from L_s to $L_s\lambda$, the drift component in (5-31) becomes

$$\frac{d\rho'}{dt} = \frac{\mu}{(C_{ox}/\lambda)} \frac{d}{d\lambda x}\left(\rho \frac{d\rho}{d\lambda x}\right) = \frac{1}{\lambda} \frac{d\rho}{dt} \quad . \tag{5-32}$$

and transfer speed increases with 1/λ.

The result in Figure 5-10(b) corresponds to the case of a well that is oversized and only filled to 1/100th of its full capacity before the transition begins. In this case, F=0.005 and currents are reduced. For a given transfer time, the fraction of untransferred charge is about 100 times larger in (b) than in (a). When speed and accuracy are of concern, wells should be filled to near capacity.

Diffusion

The diffusion term in Figure 5-8(b) is caused by a gradient in carrier density along the direction of transfer. Initially, diffusion currents are much smaller than drift currents. However, diffusion becomes a dominant transfer mechanism near the end of transfer, when only a small fraction of the original charge remains and drift currents are small. Figure 5-11 shows untransferred charge, as a function of time, for a transfer that includes only diffusion. The transfer occurs with a uniform time constant, given by $L^2q/kT\mu$. When gate lengths are scaled from L_s to $L_s\lambda$, the diffusion component in (5-31) becomes

$$\frac{d\rho'}{dt} = \frac{kT}{q}\mu \frac{d^2\rho}{d(\lambda x)^2} = \frac{1}{\lambda^2} \frac{d\rho}{dt} \quad . \tag{5-33}$$

Since the diffusion time constant decreases quadratically with λ and charge transfer time is usually dominated by diffusion, speed improves rapidly as geometries are reduced. Other than its dependence on gate length and mobility, diffusion is independent of process parameters, operating voltages, or how full the well is.

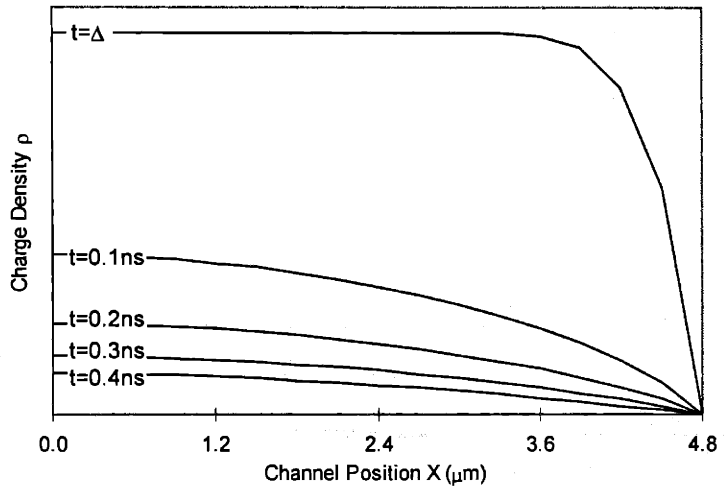


Figure 5-9. Distribution of charge remaining in the source well during the beginning of a CCD transfer. A large fraction of the original charge is removed rapidly by drift.

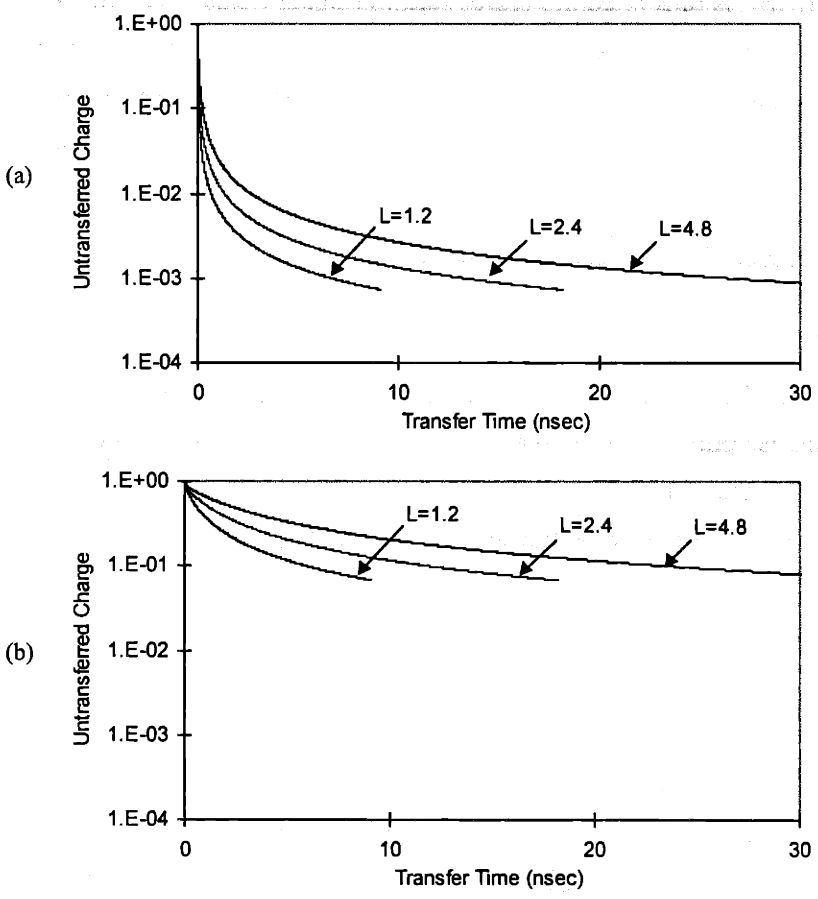


Figure 5-10. Simulated transition for a drift component only. (a) A full well of charge with F=0.5. (b) A nearly empty well with F=.005. Time constants decrease nearly linearly with L and $1/F$.

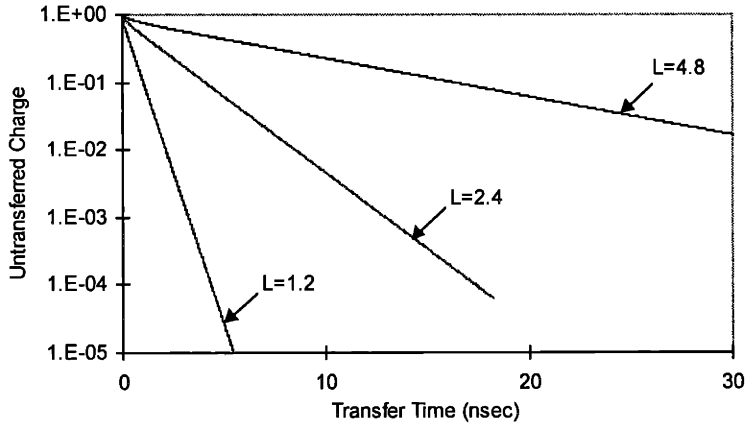


Figure 5-11. Simulated transition for diffusion current only. The diffusion time constant decreases quadratically as gate lengths are reduced.

Fringing Fields

Fringing field assisted drift is illustrated in Figure 5-8(c). Fringing fields at both the source and drain of a CCD well introduce a potential gradient along the direction of charge transfer and electrons are carried by the resulting field. Fringing fields are typically small and, like diffusion, their currents are not significant until the end of a transfer. But they do provide a substantial improvement in speed over that from diffusion alone.

The magnitude of a fringing field E_f can be described by [39]

$$E_f = \left(\frac{t_{ox}}{L_s}\right)\left(\frac{V_g}{L_s}\right) f\left(\frac{x_d}{L_s}\right) , \tag{5-34}$$

where $f(x)$ is a function that depends on two-dimensional effects and x_d is depletion depth in the well. This expression contains ratios of terms that, to first order, remain unchanged with process scaling. Although the magnitude of E_f is largely process independent, speed is still improved when geometries are reduced. At a fixed value of E_f, a reduction in gate lengths from L_s to $L_s\lambda$, results in a fringing field component,

$$\frac{d\rho'}{dt} = \mu \frac{d}{d\lambda x}\left(\rho E_f\right) = \frac{1}{\lambda}\frac{d\rho}{dt} , \tag{5-35}$$

which improves linearly with λ.

Two fringing field models, illustrated in Figure 5-12, are used for the simulations described below. In the first, fringing fields are treated as uniform across the length of the gate. This is an appropriate model for buried channel CCDs. In the second, fringing fields are represented with a maximum value of $2E_f$, at the edges of the well, and a ramp that decreases linearly over a distance of λ. This model is appropriate for surface channel CCDs. In all cases, CCD gate lengths of 4λ are used.

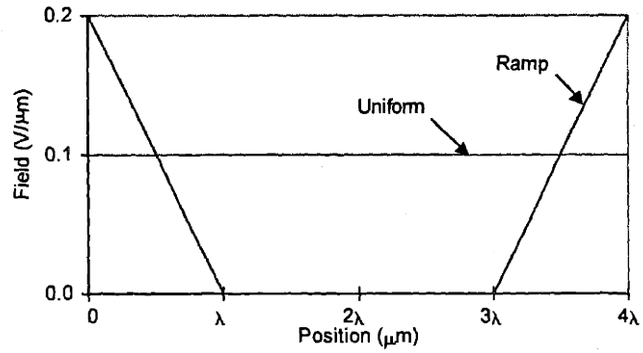


Figure 5-12. Two simplified models, used for simulation, of fringing fields within a CCD well.

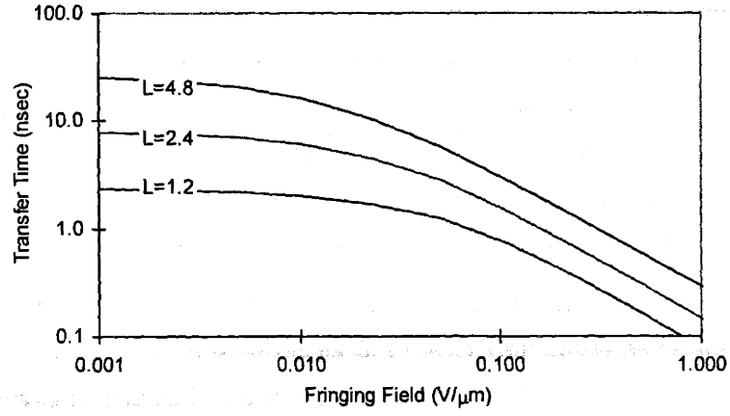


Figure 5-13. Time to reach a transfer efficiency of 10^{-4} as a function of fringing field amplitude for the case of a uniform fringing field.

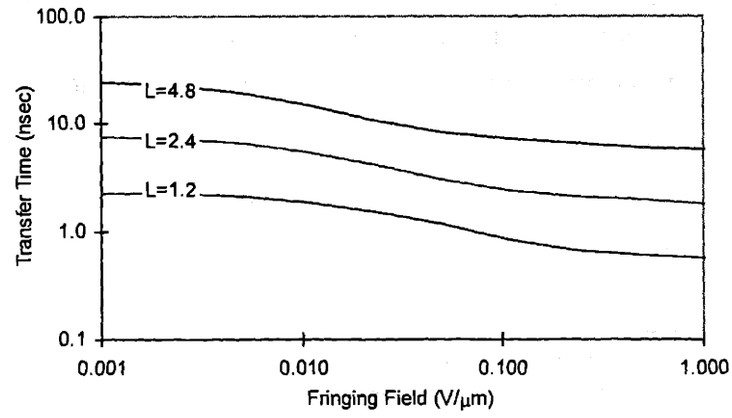


Figure 5-14. Time to reach a transfer efficiency of 10^{-4} as a function of fringing field amplitude for the case of a ramped fringing field.

Simulated results for the uniform field are shown in Figure 5-13 and those for the ramp field are shown in Figure 5-14. All three transfer mechanisms are included in the simulations. Plots show the transfer time that is needed to achieve a residual charge $< 10^{-4}$ as a function of fringing field amplitude Ef. In the case of a uniform field of magnitude greater than 50 mV/μm, transfer time is inversely proportional to Ef. In the case of a ramped field, charge is removed rapidly from the field regions but continues to flow by diffusion elsewhere. The effective CCD gate length is shortened to 2λ, the distance over which the field is zero, and transfer time decreases quadratically with this length reduction, by a factor of 4.

5.4.2 Transfer Efficiency versus Speed

Figure 5-15 shows simulated transfer results with a combination of drift, diffusion, and fringing field components for 4.8-μm, 2.4-μm, and 1.2-μm gates and λ values of 1.2-μm, 0.6-μm, and 0.3-μm. A surface channel fringing field model is used, consisting of the ramp in Figure 5-12 with a peak field amplitude of 0.2 V/μm and a fringing distance of λ. Untransferred charge for a single transfer operation is plotted versus frequency. Since charge has only half a clock cycle to transfer, frequency equals $1/2T$, where T is the required transfer time. Since untransferred charge compounds over each of transfer operation, each transfer must generally be more accurate than that desired from the overall path.

An additional constraint on the speed of a CCD device is the speed of its CMOS supporting circuitry. Four clock phases are used to control the CCDs. CCD clocks must be overlapping, CMOS clocks must be nonoverlapping, and each of these clocks must be synchronized to the others. For a given process there is a maximum frequency at which these clocks can be generated and distributed. This limit is superimposed on the chart in Figure 5-16 for processes of 1.2-μm, 0.6-μm, and 0.3-μm. At low transfer efficiency, device speed is limited by CMOS circuits, while at high transfer efficiency it is limited by CCD circuits. Charge transfer efficiency is a significant factor at large process geometries. However, as geometries are scaled, the speed of CCDs improves more rapidly than that of CMOS. In the past, charge transfer

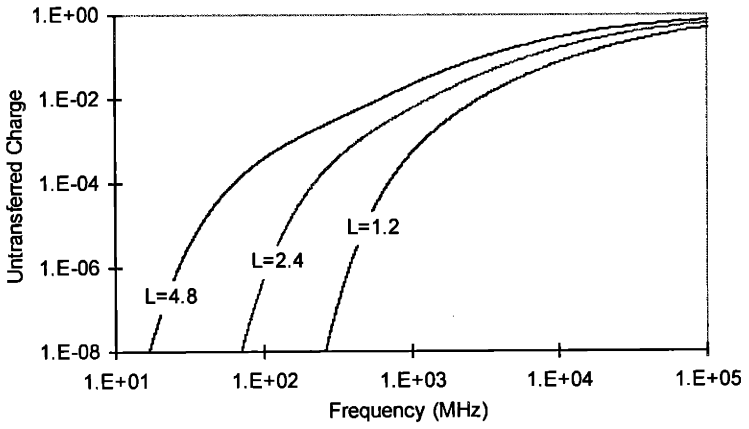


Figure 5-15. Untransferred charge versus operating frequency for all three transfer mechanisms.

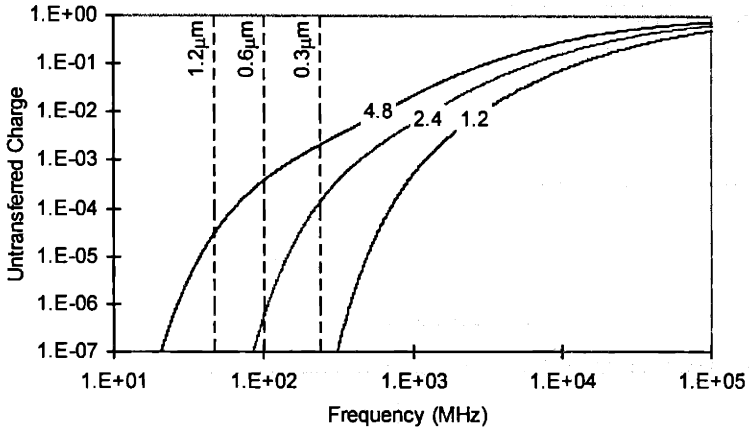


Figure 5-16. Comparison between the speed limitations of CCD transfers and CMOS support operations. CMOS limits speed for small process geometries or low transfer efficiencies.

efficiency has been a significant determinant of clock frequency. However, beyond current process geometries, frequency will be determined by CMOS supporting circuitry and will follow CMOS scaling curves.

6 Summary and Conclusions

6.1 Architecture

Design considerations for a pipelined oversampling converter are summarized below. A 1st-order device is highly susceptible to pattern noise because its inputs are effectively constant. It consumes the most hardware, power, and area of any modulator and also requires the highest accuracy from its analog transfers. In comparison to 2nd order, a 3rd-order device provides a moderate reduction in hardware, power, area, latency, and is less sensitive to signal-channel transfer noise. Modulators of order greater than 3 provide little additional improvement in these regards. In comparison to 2nd order, higher-order modulators bring no speed improvement, the danger of instability, and increased circuit complexity.

In most cases, a 2nd-order architecture is the best option. It has a simple and robust implementation without instability. The benefits, otherwise provided by 3rd-order modulation, are more easily accomplished using a 2nd-order device with a multi-bit quantizer. The greatest advantages are achieved from multi-bit quantization when only a subset of ADC bits is included in feedback. In this approach, each additional ADC bit reduces hardware, power, area, latency, and sensitivity to signal-channel transfer noise and brings little disadvantage. The number of ADC bits should be increased until quantization and random circuit noises are comparable. Beyond this point, hardware increases but resolution does not.

Single-bit feedback in a POSC does not have an inherent linearity advantage over multi-bit feedback because different DAC circuit elements are used in every stage. As a result, multi-bit feedback is an attractive option. It improves resolution by increasing the range of input amplitudes that the converter can accommodate. It also decreases susceptibility to pattern noise. In most cases, a DAC with three or fewer bits is the best option. It achieves most of these benefits with only a small increase in hardware and complexity.

Decimation of signals from a modulator of order L is best accomplished using a decimator of order D such that $L+1 \le D \le 2L+1$. Values of D below this range result in increased converter hardware because they provide inferior resolution per pipeline stage. Values of D above this range require more hardware but do not provide additional resolution. The error-averaging decimator of order $L+1$ and the matched-filter decimator of order $2L+1$ are both well suited options and have similar hardware requirements. A matched filter brings some increase in digital circuit complexity over error-averaging, but it also provides a slight reduction in pipeline latency and in sensitivity to signal-channel transfer noise.

To prevent the buildup of a large common-mode signal along the pipeline, an analog-integration device should compute the difference between signal and reference quantities before integration. A digital-integration device inherently accumulates a large common-mode signal in its digital integrators and the

range of this offset must be accommodated in its DACs. However, DAC errors in a digital integration device experience 1st-order noise shaping so that, with respect to their full-scale references, the SNR requirements of DACs and integrators are identical in analog and digital approaches. The choice between analog-integration and digital-integration architectures is based on implementation and circuit details.

Transfer inefficiency and gain mismatches in the signal and signal-integrator channels have no impact on converter performance, provided they are deterministic and do not cause the modulator to overload, because every sample experiences the same transfer function. Transfer inefficiency in the reference channel does introduce noise, but its accuracy requirements are easily met.

The resolution that is achievable from a POSC is primarily limited by DAC mismatches and random or signal-dependent transfer noise. DAC matching requirements are similar to those of a time-oversampling converter with multi-bit feedback that utilizes dynamic element matching techniques. Feedback noise and mismatches are averaged among independent circuit elements from each pipeline stage, but after averaging, feedback operations must have the full converter accuracy.

Random or signal-dependent transfer noise is most critical in the signal channel where it compounds over many stages and is amplified in the decimator. Greater accuracy is required from these operations than that of the overall converter. The impact of this noise is reduced by techniques, such as multi-bit quantization, which reduce pipeline lengths. It can also be reduced by proper selection of a decimator architecture because the decimator's transfer function determines the degree to which this noise is amplified.

6.2 CCD Circuit Techniques

Pipelined oversampling converters are most practically accomplished using a combination of CCD and CMOS circuits, fabricated in a generic CMOS process. Although CCDs are not essential to the converter concept, a combination of these circuit techniques enables performance that would be difficult from either one alone. CMOS plays a vital role in such devices by providing digital logic and CCD support circuitry. CCDs provide fully-depleted circuits, such as charge transfer, addition, integration, and conditional transfer, which are highly accurate, low power, simple, and compact. Because they are not subject to thermal noise, charge injection, or coupling from clocks or the substrate, high signal integrity is possible throughout hundreds of transfers, amidst noisy digital circuitry. Since their gain and linearity are determined by charge conservation, circuit transfer characteristics are insensitive to device parameters and highly accurate circuit-to-circuit matching is possible. Finally, fully-depleted circuits are strictly dynamic, with only capacitive switching current, and can, therefore, be performed with low power and high speed. These features make analog pipelines with hundreds of stages feasible in a CCD device.

Structurally, CCDs are similar to NMOS transistors. Their difference lies in their methods of interconnection. In a CCD circuit, adjacent gates are brought sufficiently close that their channel regions overlap and no diffusion is present between them. Although CCD devices are traditionally built using specialized fabrication, their most basic requirements are met by standard CMOS processes that include double-poly for capacitors. Overlapping structures are formed by use of parasitic 2nd-poly active gates. CCDs, built using this approach, have the advantage of being compatible with CMOS voltage levels.

Since only a limited set of fully-depleted operations is available from CCDs, other operations are performed using nondepleted circuits. These are accomplished using a technique, referred to as dynamic double sampling (DDS), whose purpose is to integrate incoming charge, on a nondepleted node, in a depleted CCD receiving well. Previous techniques for performing this function have limited speed and linearity and an undesirable sensitivity to absolute device parameters. A DDS circuit is capable of higher speed because its time constants are reduced by the gain of an amplifier. Linearity is also improved because of greater speed and an autozeroed operation. Nondepleted circuits do not have the same inherent accuracy advantages as fully-depleted ones and are subject to thermal noise and coupling.

Functions such as charge generation, wire transfer, charge sensing, D/A subtraction, and charge replication are all based on the DDS circuit technique. They differ in their sources of input signal charge. Wire transfer is used to move charge packets between nonadjacent CCD wells via a wire. Charge generation is used to convert an incoming voltage into a charge packet. D/A subtraction combines negative charge from a CCD register and positive charge from a D/A circuit on a wire. Charge replication is used to produce a replica of a charge packet without altering the original and is also used for charge-to-voltage conversion at the input to a CMOS comparator.

6.3 Converter Implementations

The design of a pipelined oversampling converter must take into account both the architectural considerations, described in Chapter 2, and the circuit considerations, described in Chapter 3. Architectural considerations determine design parameters such as the number of pipeline stages, the number of ADC and DAC bits, and the modulator and decimator order. However, circuit considerations determine which combinations of these design parameters can practically be implemented.

Examples of both analog and digital integration designs are described. Second-order architectures are chosen because of their inherent stability and immunity to pattern noise. Multi-bit ADCs are used to reduce the number of pipeline stages and multi-bit DACs are chosen because they reduce overloading effects and allow a greater input range. Device throughput is increased by dividing each pipeline stage into multiple substages. Dual pipeline configurations are used to reduce device latency and the number of sequential signal transfers. The size of CCD wells and their corresponding power dissipation are determined by both architectural parameters and desired SNR.

The choice between analog and digital integration is based primarily on circuit, layout, and process issues. Analog integration has the advantage of fewer bits in its feedback DACs. However, it requires circuits for highly linear charge replication and charge-domain common-mode suppression. It also requires smaller process geometries for practical implementation. In state-of-the-art processes, analog integration is possible and from an architectural perspective, is probably preferable. However, these circuit challenges must be met first.

Two prototype pipelined oversampling A/D converters have been demonstrated. They were built from CCD/CMOS circuits in standard CMOS processes and have demonstrated that moderate to high performance is achievable from CCD circuits without custom processing. The first prototype uses a 1.2-μm process and achieves 78-dB SFDR and 74-dB SNR within a 9-MHz bandwidth using Nyquist sampling at an 18-MHz data rate. The second prototype uses a 0.35-μm process and achieves 70-dB SFDR and 66-dB SNR within a 15-MHz bandwidth at a 30-MHz data rate.

6.4 Performance Limits

At the input and output of a CCD device, charge-domain signals must be translated into a form, such as a voltage or current, which is suitable for off-chip communication. Power and area are determined primarily by the noise relationships of these interface circuits, rather than by noise within the charge-domain circuits themselves. At a given SNR, thermal or shot noise in these input and output circuits set the minimum charge packet sizes that can be used. Since CCD power dissipation is nearly proportional to charge packet sizes, there is a direct relationship between SNR and power. For each of the interface circuits described, power and area increase quadratically with SNR. In most cases, at a given SNR, power remains unchanged and area increases during process scaling.

Another factor that is not fundamental, but impacts power and area in practice, is minimum process geometries. The minimum size of a CCD well is constrained by process geometries and speed considerations dictate that these wells must be filled to near capacity. The power and area of minimum sized wells decrease as λ^3 when process geometries are reduced by a factor of λ.

At higher SNR or smaller process geometries, noise limits are dominant. At lower SNR or larger geometries, process limits are. In the past, process limitations have been a dominant factor and power and area have improved rapidly with scaling. For future CMOS generations, power and area will be limited by noise constraints and will not improve with further process scaling.

CCD transfer speed is determined by well gate lengths and stored charge density. To preserve speed, charge density is held constant as CCDs are scaled. Diffusion time constants, which usually dominate transfer times when wells are full, decrease quadratically as gate lengths are reduced. Short channel

effects cause fringing fields which improve transfer speed as well. A second determinant of the speed of a CCD device is its CMOS supporting circuitry.

When large process geometries are used or high transfer accuracy is required, CCDs limit overall device speed. At smaller geometries or lower accuracy, CMOS speed is limiting. In the past, charge transfer accuracy has been a dominant factor. However, the speed of CCDs improves more rapidly than that of CMOS as devices are scaled. In the future, frequencies will be limited by CMOS and will improve along a CMOS scaling curve.

6.5 Conclusions

A new pipelined oversampling A/D architecture has been described that has the potential for significantly increased data rates and reduced power over conventional oversampling techniques. Initial prototypes of the architecture have been demonstrated with 74-dB SNR and 324 mW at 18 MSPS and 66-dB SNR and 230 mW at 30 MSPS. The architecture does not incorporate the inherent linearity advantage of conventional oversampling techniques, but is nonetheless thought to be applicable to resolutions up to about 16 bits. The extension of this technique to higher speeds will require additional improvements in input sampling techniques.

The potential of CCDs for performing analog functions with moderate-to-high resolution, using generic two-poly CMOS processing, has been demonstrated. Techniques for accomplishing these functions and for integrating and synchronizing CCD and CMOS circuit elements have been developed. The power and speed performance limits of these circuits have been studied. The results indicate that, in comparison to static CMOS circuit techniques, CCDs may be capable of significant improvements in power and speed. Because they are compatible with many analog CMOS processes, CCDs are thought to be a useful circuit technique for general purpose analog applications as well as A/D conversion.

Appendix A. Variable Definitions

τ_3 = Time constant for the secondary sensing transition in phase 3

τ_4 = Time constant for the sensing transition in phase 4

τ_2 = Time constant for the spill transition in phase 2

δ = Unit impulse function.

α = Gain with which integrator signals are transferred forward from one stage to the next. Nominally equal to 1.

ϕ_{b2}, ϕ_{b4} = CCD barrier clock signals asserted high over two phases

ϕ_{s1}, ϕ_{s3} = CCD storage well clock signals asserted high over two phases

ϕ_1, ϕ_2, ϕ_3, ϕ_4 = CMOS clock signals for phases 1 to 4, asserted high

ϕ_{12}, ϕ_{23}, ϕ_{34}, ϕ_{41} = CMOS clock signals asserted high over two phases

$\overline{\phi_{12}}$, $\overline{\phi_{23}}$, $\overline{\phi_{34}}$, $\overline{\phi_{41}}$ = CMOS clock signals asserted low over two phases

$\overline{\phi_1}$, $\overline{\phi_2}$, $\overline{\phi_3}$, $\overline{\phi_4}$ = CMOS clock signals for phases 1 to 4, asserted low

A = Gain of the amplifier in a DDS core circuit

a, b, c, d = Best-fit polynomial coefficients for transient-related distortion

a_1, a_2, a_{12} = First-order coefficients of input capacitor voltage dependence

a_3 = First-order coefficient of parasitic capacitor voltage dependence

b_1, b_2, b_{12} = Second-order coefficients of input capacitor voltage dependence

b_3 = Second-order coefficient of parasitic capacitor voltage dependence

C = Capacitance of each input, output, and CCD circuit in a CCD device

C_1, C_2 = Input capacitors for the double-sampling circuits

C_{12} = Input capacitors for the double-sampling circuits when $C_1 = C_2$

C_a = Gate-to-channel capacitance of a CCD well per unit area

C_{cs} = CCD channel-to-substrate capacitance

C_{eff} = Effective capacitance of a CCD well as seen by its driver.

C_{gc} = CCD gate-to-channel capacitance

c_i = Transfer noise introduced in the integrator channel, normalized to the full-scale reference.

C_{ox} = CCD gate oxide capacitance.

C_{p1} = Parasitic capacitance on voltage-domain nodes

C_{p2} = Parasitic capacitance on an SDS or DDS input node

C_s = Storage capacity of a CCD well.

c_s = Transfer noise introduced in the signal channel, normalized to the full-scale reference.

c_u = Mismatches in the signal DAC of a digital integration architecture.

c_v = Mismatches in the feedback DAC of either digital or analog integration architectures.

D = Decimator order. Determined by the number of series accumulators in the decimator.

d_1 = Value of the 1st digital accumulator in a pipelined decimator.

d_2 = Value of the 2nd digital accumulator in a pipelined decimator. Filter output for a 1st-order architecture with an error-averaging decimator.

d_3 = Value of the 3rd digital accumulator in a pipelined decimator. Filter output for a 2nd-order architecture with an error-averaging decimator. Filter output for a 1st-order architecture with a matched-filter decimator.

d_5 = Value of the 5th digital accumulator in a pipelined decimator. Filter output for a 2nd-order architecture with a matched-filter decimator.

E_f = Fringing field within a CCD well.

e_r = Quantization noise modeled as an additive error and normalized to the full-scale reference. Introduced in the coarse feedforward ADC.

fe = Effective frequency, equal to the inverse of latency, used to compute thermal generation charge.

F = Fill ratio of a CCD well. Equal to the average storage voltage divided by the gate voltage.

f_0 = Fundamental frequency of an analog input signal

f_b = Amplifier feedback signal in a DDS core circuit

f_b = Analog modulator feedback signal. Generated by its feedback DAC.

f_g = Input node to an SDS or DDS core circuit.

f_{g0} = Minimum value of f_g during the collection phase

H = Period of one clock phase, equal to one quarter of the clock period

i_1 = Value of the first modulator analog integrator when signal and reference quantities are combined in a single integrator.

i_2 = Value of the second modulator analog integrator when signal and reference quantities are combined in a single integrator.

I_s = Signal current that is integrated over a cascode barrier

J_g = Thermal generation current density.

k = Boltzmann's constant.

k = Number of bits in the modulator feedback path and used by the feedback DAC. Less than or equal to r.

K = Transistor transconductance parameter measured in A/V^2

K_1, K_2, K_3 = Unitless constants given by capacitive ratios

L = Modulator order. Determined by the number of zeros in its transfer function from quantization errors to the output.

L_b = Length of a CCD barrier gate.

L_s = Length of a CCD storage gate.

m = Factor representing the number of thermal noise terms introduced by a circuit.

m = Time multiplied by the sampling frequency at the converter input. An integer equal to the sample point number.

n = Pipeline conversion block number. Also represents clock cycle in cyclic representations.

n_f = Noise on node v_f

n_g = Noise on node f_g

n_m = Noise on node v_m

N_s = Full scale number of signal electrons.

P = Period of one clock cycle, equal to four times the period of one phase

P = Pipeline length as determined by the number of its conversion blocks. Less than or equal to the number of pipeline stages.

q = Electron charge equal to 1.6×10^{-19}

q_n = Random circuit noise introduced in the modulator feedforward path.

Q_o = Output charge produced by an SDS or DDS circuit

Q_s = Full-scale charge signal.

Q_s = Input charge packet to wire transfer, floating gate, and comparison circuits

R = Ratio of gate-to-channel and channel-to-substrate capacitances.

r = Number of bits generated by the coarse feedforward ADC. Equivalent to the number of bits passed from the modulator to the decimator.

R = Oversampling ratio of a time-oversampling $\Delta\Sigma$ converter.

R_f = Analog full-scale reference for a pipelined $\Delta\Sigma$ converter. Also denotes the full-scale reference for ADC and DAC elements within an analog integration modulator.

r_s = Final converter output generated by downsampling the decimation filter result at a single point.

S_g = Value of the converter input sample. A constant for each $\Delta\Sigma$ computation along the pipeline.

s_i = Analog signal sample passed along the pipelined signal channel.

T = temperature in degrees Kelvin

t = time

t_d = amplifier delay in a DDS core circuit

t_k = Truncation error introduced when only a subset of the modulator output bits are included in its feedback path.

u = Unit step function.

u_1 = Value of the first modulator signal integrator when separate signal and reference integrators are used.

u_2 = Value of the second modulator signal integrator when separate signal and reference integrators are used.

v_1 = Value of the first modulator reference integrator when separate signal and reference integrators are used.

v_2 = Value of the second modulator reference integrator when separate signal and reference integrators are used.

V_c = High-level circuit bias used for clamping

V_d = High-level circuit bias used to drain electrons

v_f = Voltage-domain node of a charge generator or D/A converter input capacitor

V_g = CCD gate voltage swing.

v_g = Voltage-domain representation of a floating gate signal

V_h = Mid-level bias used to set precharge and sensing levels of SDS or DDS inputs

v_i = Analog input voltage to charge generators

V_i = Rms value of input waveform on v_i

v_m = Sampled analog input voltage in a charge generator

V_p = Low-level bias used to inject electrons at the input to an SDS or DDS core

V_r = Reference bias used to set the precharge and sensing levels of a DDS core

V_s = Analog operating voltage and signal swing range.

V_s = Sampled value within a clamp-and-sample circuit

v_{th} = Threshold voltage of cascode gate

v_{th1}, v_{th2} = Threshold voltage of CCD gates configuration

V_w = CCD storage voltage. Equal to signal charge divided by storage capacity.

W/L = transistor width-to-length ratio

W = Sample aperture length. Equal to the number of modulator outputs that are combined in the decimator to form a single result.

w_k = Digital modulator feedback signal. Contains the k most significant bits of its output.

w_r = Digital modulator output signal.

W_s = Width of a CCD storage well.

x = Distance within a CCD well along the direction of transfer.

x_d = Depletion depth beneath a CCD gate.

Z = Impulse response length of the decimation filter.

Appendix B. Block Diagram Notation

= sample-and-hold, delay, and latching operations

= downsampling operation at a single point in time

= coarse non-oversampling ADC with a full-scale reference of *Rf*

= coarse non-oversampling DAC with a full-scale reference of *Rf*

= coarse non-oversampling DAC with a full-scale reference of *Sg*

= multiplication by 2

= analog or digital addition

= conditional transfer or switching operation

= division by 2

= sample-and-hold operation

= digital matched filter

= comparator

= comparator or amplifier

= replicator

= bits 4 through 2 of a digital or analog bus

References

[1] J. C. Candy, G. C. Temes, "Oversampling Methods for A/D and D/A Conversion," *Oversampling Delta-Sigma Data Converters, Theory, Design and Simulation*, The Institute Of Electrical And Electronics Engineers, Inc., New York, pp. 1-25, 1991.

[2] W. R. Bennett, "Spectra of quantized signals," *Bell Sys. Tech. J.*, vol. 27, pp. 446-472, July 1987.

[3] J. C. Candy and O. J. Benjamin, "The structure of quantization noise from sigma-delta modulation," *IEEE Trans. Commun.*, vol. COM-29, pp. 1316-1323, Sept. 1981.

[4] W. Chou and R. M. Gray, "Dithering and its effects on sigma-delta and multi-stage sigma-delta modulation," *IEEE Proc. ISCAS'90*, pp.368-371, May 1990.

[5] R. M. Gray, "Oversampled sigma-delta modulation," *IEEE Trans. Commun.*, vol. COM-35, pp. 481-489, May 1987.

[6] R. M. Gray, "Spectral analysis of quantization noise in a single-loop sigma-delta modulator with dc input," *IEEE Trans. Commun.*, vol. COM-37, pp. 588-599, June 1989.

[7] Ning He, F. Kuhlmann, and A. Buzo, "Double-loop sigma-delta modulation with dc input," *IEEE Trans. Commun.*, vol. COM-38, pp. 487-495, April 1990.

[8] R. M. Gray, "Quantization noise spectra," *IEEE Trans. Information Theory*, vol. IT-36, pp.1220-1244, Nov. 1990.

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

[10] T. Cataltepe, G. C. Temes, and L. E. Larson, "Digitally corrected multi-bit sigma-delta data converters," *IEEE Proc. ISCAS'89*, pp. 647-650, May 1989.

[11] L. E. Larson, T. Cataltepe, and G. C. Temes, "Multibit oversampled sigma-delta A/D converter with digital error correction," *Elect. Lett.*, vol. 24, pp. 1051-1052, Aug. 1988.

[12] T. C. Leslie and B. Singh, "An improved sigma-delta modulator architecture," *IEEE Proc. ISCAS'90*, pp. 372-375, May 1990.

[13] P. J. Hurst and R. A. Levinson, "Delta-sigma A/Ds with reduced sensitivity to op amp noise and gain," *IEEE Proc. ISCAS'89*, pp. 254-257, May 1989.

[14] D. Ribner, R. Baertsch, et al., "A 16b third-order sigma-delta modulator with reduced sensitivity to non-idealities," *ISSCC Dig. Tech. Pap.*, pp. 66-67, Feb. 1991.

[15] D. B. Ribner, "A comparison of modulator networks for high-order oversampling sigma-delta analog-to-digital converters," *IEEE Trans. Circuits Sys.*, vol. CAS-38, pp. 145-159, Feb. 1991.

[16] Y. Shoji and T. Suzuki, "Improved double integration delta-sigma modulations for A/D and D/A conversion," *IEEE Proc. ISCAS'87*, pp.451-454, May 1987.

[17] R. H. Walden, T. Cataltepe, and G. C. Temes, "Architectures for high-order multibit sigma-delta modulators," *IEEE Proc. ISCAS'90*, pp. 895-898, May 1990.

[18] Y. Sakina, "Multi-bit sigma-delta A/D converters with nonlinearity correction using dynamic barrel shifting," University of California, Berkeley Report, Dept. of EECS, 1990.

[19] J. C. Candy, Y. C. Ching, and D. S. Alexander, "Using triangularly weighted interpolation to get 13-bit PCM from a sigma-delta modulator," *IEEE Trans. Commun.*, vol. COM-24, pp. 1268-1275, Nov. 1976.

[20] J. C. Candy, "Decimation for sigma delta modulation," *IEEE Trans. Commun.*, vol. COM-34, pp. 72-76, Jan. 1986.

[21] R. E. Crochiere and L. R. Rabiner, "Interpolation and decimation of digital signals – a tutorial review," *Proc. IEEE*, vol. 69, pp. 300-331, Mar. 1981.

[22] S. Chu and C. S. Burrus, "Multirate filter designs using comb filters," *IEEE Trans. Circuit Sys.*, vol. CAS-31, pp. 913-924, Nov. 1984.

[23] E. Dijkstra, O. Nys, C. Piguet, and M. Degrauwe, "On the use of modulo arithmetic comb filters in sigma delta modulators," *IEEE Proc. ICASSP'88*, pp. 2001-2004, April 1988.

[24] H. Meleis and P. LeFur, "Novel architecture design for VLSI implementation of a FIR decimation filter," *IEEE Proc. ICASSP'85*, pp. 1380-1383, Mar. 1985.

[25] T. Saramaki, T. Karema, T. Rotoniemi, H. Tenhunen, "Multiplier-free decimator algorithms for superresolution oversampled converters," *Proc. ISCAS'90,* pp. 3275-3278, May 1990.

[26] T. Saramaki and H. J. Tenhunen, "Efficient VLSI-realizable decimators for sigma-delta analog-to-digital converters," *IEEE Proc. ISCAS'88*, pp. 1525-1528, June 1988.

[27] S. R. Norsworthy, R. Schreier, G. C. Temes, "An Overview of Basic Concepts," *Delta-Sigma Data Converters; Theory, Design, and Simulation*, IEEE Press, Piscataway, NJ, 1997.

[28] T. L. Brooks, D. H. Robertson, D. F. Kelly, A. Del Muro, S. W. Harston, "A cascaded sigma-delta pipeline A/D converter with 1.25 MHz signal bandwidth and 89 dB SNR," *IEEE J. Solid-State Circuits*, vol. 32, No. 12, Dec 1997.

[29] S. R. Norsworthy, R. Schreier, G. C. Temes, "Analog Circuit Design For ΔΣ ADCs," *Delta-Sigma Data Converters; Theory, Design, and Simulation*, IEEE Press, Piscataway, NJ, 1997.

[30] H. Inose, Y. Yasuda, "A Unity Bit Coding Method by Negative Feedback," *Proc. IEEE*, vol. 51, pp.1524-1535.

[31] M. W. Hauser, R. W. Broderson, "Circuit and Technology Considerations for MOS Delta-Sigma A/D Converters," *IEEE Proc. ISCAS'86*, pp. 1310-1315, May 1986.

[32] L. R. Carley, "A Noise-Shaping Coder Topology for 15+ bit Converters," *IEEE J. Solid-State Circuits*, vol. SC-24, pp. 267-273, April, 1989.

[33] Y. Sakina, "Multi-bit ΔΣ Analog-to-Digital Converters with Nonlinearity Correction using Dynamic Barrel Shifting," Electronics Research Laboratory, College of Engineering, University of California, Berkeley CA, Memorandum No. UCB/ERLM93/63, 1993.

[34] C. L. Keast, "An Integrated Image Acquisition, Smoothing and Segmentation Focal Plane Processor," M. I. T Thesis, Feb. 1992.

[35] Barbe, David F., "Imaging Devices using the Charge-Coupled Concept," *Proceedings of the IEEE*, vol. 63, no. 1, pp. 38-67, Jan. 1975.

[36] Bencuya, Selim S., Steckl, Andrew J., "Charge-Packet Splitting in Charge-Domain Devices," *IEEE Trans. Electron devices*, vol. DC-31, no. 10 , pp. 1494-1501, Oct.1984.

[37] Bencuya, S. S., Steckl, A. J., "Dynamic Packet Splitting in Charge Domain Devices," *IEEE Electron Device Letters*, vol. EDL-3, no. 9, pp. 268-270, Sep. 1982.

[38] Bencuya, Selim S., Steckl, Andrew J., Burkey, B.C., "Impact of Edge Effects on Charge-Packet-Splitting Accuracy," *Solid-State Electronics*, vol. 30, no. 3, pp. 299-305, 1987.

[39] Benyon, J. D. E., Lamb, D. R., "*Charge-Coupled Devices and their Applications*," McGraw-Hill Book Co., New York, 1980.

[40] Berglund, C., Boll H., "Performance Limitations of the IGFET Bucket-Brigade Shift Register," *IEEE Trans. Electron Devices*, vol. ED-19(7), pp. 852-860, 1972.

[41] Butler, W. J., Puckette, C. M., Smith, D. A., "Differential Mode of Operation for Bucket Brigade Circuits," *Electronics Letters*, vol. 19, pp. 520-522, 1971.

[42] Chandler, Charles E., Bredthauer, Richard A., "Sub-electron noise charge coupled devices," *SPIE Charge-Coupled Devices and Solid-State Optical Sensors*, vol. 1242, pp. 238-251, 1990.

[43] Colbeth, Richard E., Fossum, Eric R., "A 1-GHz Charge-Packet Replicator/Subtractor Circuit for GaAs CCD Signal Processing," *IEEE J. Solid-State Circuits*, vol. 25, no. 4, pp. 1016-1019, Aug. 1990.

[44] Colbeth, R. E., Doudoumopoulos, N. A., Eid, E-S., Kemeny, S.E., Montalvo, A., Fossum, E. R., "High Performance Charge Packet Comparison for Analog CCD Applications," unpublished document.

[45] Early, James M., "Charge Coupled Amplifier," U.S. Patent #3,806,772, Fairchild Camera and Instrument Corporation, Mountain View, CA, Apr. 23, 1974.

[46] Edwards, Nathan P., White, James M., "Non-Destructive Charge Transfer Device Differencing Circuit," U.S. Patent #4,239,983, IBM Corp., Armonk, NY, Dec. 16, 1980.

[47] Fossum, E. R., Kemeny, S. E., Bredthauer, R. A., LaShell, M., "Digitally Programmable Gain Control Circuit for Charge-Domain Signal Processing," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, Apr. 1991.

[48] Fossum, Eric R., "A Linear and Compact Charge-Coupled Charge Packet Differencer/Replicator," *IEEE Trans. Electron Devices*, vol. ED-31, no. 12, pp. 1784-1789, Dec. 1984.

[49] Fossum, Eric R., "Wire Transfer of Charge Packets for On-Chip CCD Signal Processing," *SPIE Charge Coupled Devices and Solid-State Optical Sensors*, vol. 1242, pp. 179-186, 1990.

[50] Fossum, Eric R., "Wire Transfer of Charge Packets using a CCD-BBD Structure for Charge-Domain Signal Processing," *IEEE Trans. Electron Devices*, vol. 38, no. 2, pp. 291-297, Feb. 1991.

[51] Hilsum, Cyril, Moss, T.S., "*Handbook on Semiconductors*," North-Holland Publishing Co., New York, 1981.

[52] Howes, M.J., Morgan, D.V., "*Charge-Coupled Devices and Systems*," John Wiley & Sons, New York, 1979.

[53] Krambeck, R. H., Walden, R. H., Pickar, K. A., "Implanted Barrier Two-Phase Charge Coupled Device," *Applied Physics Letters*, vol. 19, pp. 520-522, 1971.

[54] Maclennan, D. J., Mavor, J. J., Vanstone, G. F., Windlw, D. J., "Novel Tapping Technique for Charge-Coupled Devices", *Electronics Letters*, vol. 9, pp. 610-611, 1973.

[55] Maclennan, D. J., Mavor, J., "Novel Technique for the Linearization of Charge-Coupled Devices", *Electronics Letters*, vol. 11, pp. 222-223, 1975.

[56] Sangster, F. L. J., "Integrated MOS and Bipolar Analog Delay Lines using Bucket-Brigade Capacitor Storage," *IEEE International Solid-State Circuits Conference*, pp. 74-75, Feb. 1970.

[57] Sequin, Carlo H., Tompsett, Michael F., "*Charge Transfer Devices*," Academic Press, Inc., New York, 1975.

[58] Sequin, Carlo H., "Linearity of Electrical Charge Injection into Charge-coupled Devices," *IEEE J. Solid-State Circuits*, vol. SC-10, no. 2, pp. 81-92, Apr. 1975.

[59] Taylor, G. W., Tasch, A. F., "On the Measurement of the Channel Potential in Charge-Coupled-Device Structures," *Solid-State Electronics*, vol. 20, pp. 473-474, 1977.

[60] Thornber, K. K., "Incomplete Charge Transfer in IGFET Bucket-Brigade Shift Registers," *IEEE Trans. Electron Devices*, vol. ED-18, pp. 941-950, 1971.

[61] Tompsett, M. F., "Surface Potential Equilibration Method of Setting Charge in Charge-Coupled Devices," *IEEE Trans. Electronic Devices*, vol. ED-22, p. 305, 1975.

[62] Wen, D. D., Salsbury, P. J., "Analysis and Design of a Single-Stage Floating Gate Amplifier," *IEEE International Solid-State Circuits Conference*, pp. 154-155, Feb. 1973.

[63] Wen, David D., Early, James M., Kim, Choong-Ki, Amelio, Gilbert F., "A Distributed Floating-Gate Amplifier in Charge-Coupled Devices", *IEEE International Solid-State Circuits Conference*, pp. 24-25, Feb. 1975.

[64] Chang, W. C., Tredwell, T. J., Stevens, E. G., Nichold, D. N., "High-Density Solid-State Image Sensor," *SMPTE Journal*, pp. 1186-1188, Dec. 1987.

[65] S. R. Norsworthy, R. Schreier, G. C. Temes, "*Delta-Sigma Data Converters, Theory, Design, and Simulation*," IEEE Press, Piscataway, NJ, 1997.

[66] S. A. Paul, H-S. Lee, "A 9-b Charge-to-Digital Converter for Integrated Image Sensors," *IEEE J. Solid-State Circuits*, vol. 31, no. 12, pp. 1931-1938, Dec. 1996.

[67] S. A. Paul, H-S. Lee, "A 9-b Charge-to-Digital Converter for Integrated Image Sensors," *IEEE International Solid-State Circuits Conference*, Digest of Technical Papers, Feb. 1996.