

Copyright © 1995, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**LOW-POWER LOW-VOLTAGE ANALOG-TO-DIGITAL
CONVERSION TECHNIQUES USING PIPELINED
ARCHITECTURES**

by
Thomas Byunghak Cho

Memorandum No. UCB/ERL M95/23

2 April 1995

**LOW-POWER LOW-VOLTAGE ANALOG-TO-DIGITAL CONVERSION
TECHNIQUES USING PIPELINED ARCHITECTURES**

by

Thomas Byunghak Cho

Memorandum No. UCB/ERL M95/23

2 April 1995

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

Low-Power Low-Voltage Analog-to-Digital Conversion Techniques Using Pipelined Architectures

by

Thomas Byunghak Cho

Doctor of Philosophy in

Engineering - Electrical Engineering and Computer Sciences

University of California at Berkeley

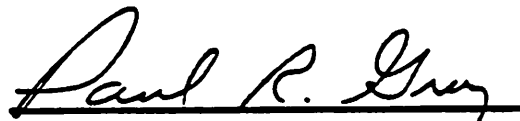
Professor Paul R. Gray, Chair

With a recent boom of portable applications from camcorders to wireless personal communication devices, power reduction in high speed A/D converters is becoming increasingly important to minimize battery drain for longer device operating hours. Low voltage operation is another key requirement for these A/D converters in order to be integrated in larger mixed signal IC's containing mostly digital blocks from DSP and control.

The main objective of the research is to examine the fundamental limitations to high speed CMOS A/D converters and to develop techniques to reduce power dissipation and to allow low voltage operation in a pipeline A/D converter architecture. Proposed circuit techniques for low power consumption are the use of fully dynamic comparators that dissipate no DC power, made possible by the digital correction within the pipeline, and optimum scaling of switched capacitor circuits down the pipeline. Also, to be compatible with low voltage mixed-signal system environments, a switched capacitor

circuit in each pipeline stages is implemented and operated at 3.3V with a new high speed low voltage operational amplifier, and dynamic gate voltage boosting technique is used to operate transmission gates with full analog voltage switching capability on a 3.3V supply.

To verify the effectiveness of the proposed techniques, a 10bit 20MS/s pipeline A/D converter is designed and fabricated in 1.2 μ m CMOS technology. The test results show that 59.1 dB of SNDR (signal-to-noise-plus-distortion ratio) can be achieved for the input frequency of 100kHz while the whole A/D converter dissipates only 35mW at 20MS/s. At 1MS/s and reduced bias current, the power dissipation is only 2.8mW with 58.0dB of SNDR.


Paul R. Gray, Chairman of Committee

Acknowledgment

I would like to express my sincere appreciation to my advisor, Prof. Paul Gray. My work here wouldn't be possible without his support and guidance throughout my graduate school years. His keen insights into IC design led me into the right direction of research, and I've learned many valuable lessons from him. I also thank Prof. Boser for being my qual-committee chairman and giving me valuable comments for my research work.

There are many colleagues who deserve recognitions here. Cormac Conroy, my ex-officemate who used to sit right next to me, was such a wonderful person to be with, not only for his abundant knowledge and experience but also for his humor and caring heart for people. During most of my graduate years, I've learned a lot from him on both research and some of his humor. I would like to also thank Gregory Uehara, now a professor at University of Hawaii, for his voluntary help even during his busy schedule. He's a man of big warm smile, never failing to provide our group with Hawaiian Macademia Nut Chocolate Candies. Also, I would like to acknowledge Dave Cline, Robert Neff, and Ken Nishimura for their endless help and valuable suggestions and discussions for my research throughout my graduate school years. I would like to recognize Edward Liu, Henry Chang, Gani Jusuf for their help during early days of my research. I also thank Dave Cline and Andy Abo for reviewing my thesis and the rest of analog IC group for their help and support.

I would like to also thank Korean EECS members for their support, and they are Hangyoo Kim, Youngchul Sim, Hyungchul Shin, and Byoungcho Lee. My thanks also go to my many ex-roommates and friends, Frank Yim, Gichoen Park, and many others for their friendship. Especially, I want to recognize my good friend, Frank Yim for his faithful friendship throughout years.

I would like to thank Berkland Baptist Church family members here. First, I would

like to thank Eun Sok Kim, now a professor at University of Hawaii, and his wife Jong Nam Kim. I am grateful to them for helping me during my early graduate school years in Berkeley. I am also grateful to Seongil Yim who was my Christian Big Brother. Throughout my graduate school years with him in Berkeley, I cannot forget his continual support and care. I also feel deeply thankful to Pastor Namha Lee for his love and care throughout my years in Berkeley. I would like to also recognize Hyunchul Park, Inho Kim, Soonhoi Ha, Seungjun Lee, Seogchul Kim, Bongyoon Han, and Yushin Kim for their continual support and care. Especially, I would like to recognize and thank Juyoung Lee, Sungjin Lee, and Kidae Lee, my church “cell” members who share their lives with me being true brothers in Christ.

I am grateful to my parents, Mr. Kwang Mook Cho and Mrs. Il Nam Cho for their never ceasing love and support during my whole life. Also without the love and support of my brothers, Byung Kee Cho, Byung Woon Cho, and Byung Suk Cho, my graduate school life wouldn't be possible. I also thank my parents-in-law, Mr. Sung Kyun Suk and Mrs. Bok Hee Suk for their love and encouragement.

I thank my wife, Okkyung, for her continual love and understanding. Being a busy graduate student herself, her continual support made my work possible. Also, without her, my life wouldn't be as meaningful as it is now.

Most of all, I am truly thankful to my Lord, Jesus Christ for His endless love from the moment He gives me life. I owe everything to Him who is a true shepherd of my life.

This research work was supported by the National Science Foundation, the Advanced Research Projects Agency, and SRC.

Table of Contents

Chapter 1 Introduction	1
1.1 Motivation.....	1
1.2 Thesis Organization	3
Chapter 2 Fundamental Limitations to Power Dissipation of CMOS Analog-to-Digital Converters	5
2.1 Introduction.....	5
2.2 Basic A/D Conversion Concepts	5
2.3 Sampling in MOS Technologies	7
2.3.1 Basic MOS Sample/Hold Circuit.....	7
2.3.2 Practical MOS S/H Amplifiers: Op-Amp-Based SC S/H Circuits and Limitations .	15
2.3.3 Additional Practical Design Considerations	28
2.4 Quantization.....	33
2.4.1 Limitations on Comparators	34
2.4.2 Practical Implementations of Comparators.....	38
2.5 Reference Potential Generation	40
2.5.1 Resistor String DAC(R-DAC)	41
2.5.2 Capacitor Array DAC(C-DAC)	43
2.5.3 Power Comparison: R-DAC vs. C-DAC	46
2.5.4 Practical Limitations	49
Chapter 3 CMOS High Speed A/D Converter Architectures	51
3.1 Introduction.....	51
3.2 Flash Architecture.....	51
3.3 2 Step Flash Architecture.....	53
3.4 Pipeline Architecture	58
3.5 Power Comparison.....	60
3.6 Other ADC Architectures	64
Chapter 4 A CMOS Pipeline A/D Converter Architecture	70
4.1 Introduction.....	70
4.2 Key Building Blocks.....	70
4.3 Stage Accuracy Requirements	76
4.4 Per-Stage Resolution & Power	81
4.4.1 No Scaling.....	84
4.4.2 Scaling	88

Chapter 5 Low Power Low Voltage Techniques in Pipeline A/D Converter Architecture	98
5.1 Introduction.....	98
5.2 1.5 bit/stage Pipeline Architecture.....	98
5.3 Power Reduction Techniques.....	102
5.3.1 Dynamic Comparators	103
5.3.2 Scaling of SC Circuits Through the Pipeline.....	105
5.4 Low Voltage Operation	107
5.4.1 High-speed 3.3 V Op Amp	107
5.4.2 Low Voltage Operation of SC Circuits	110
Chapter 6 Experimental Prototype and Measurement Results	114
6.1 Prototype	114
6.1.1 Floor Plan.....	114
6.1.2 Op Amp Bias Circuit and Clock Generation Circuit.....	115
6.2 Measurement Result	121
6.3 Power Comparison.....	124
Chapter 7 Conclusion	126
Appendix 1 Error Sources in MOS S/H Circuits	130
Appendix 2 kT/C Noise Calculation in SC Circuit	136
Appendix 3 Digital Error Correction	142
References	147

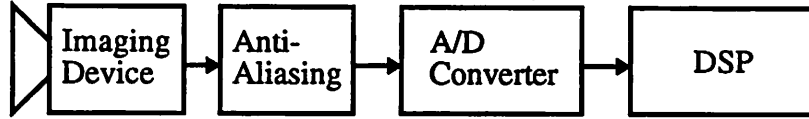
CHAPTER 1

Introduction

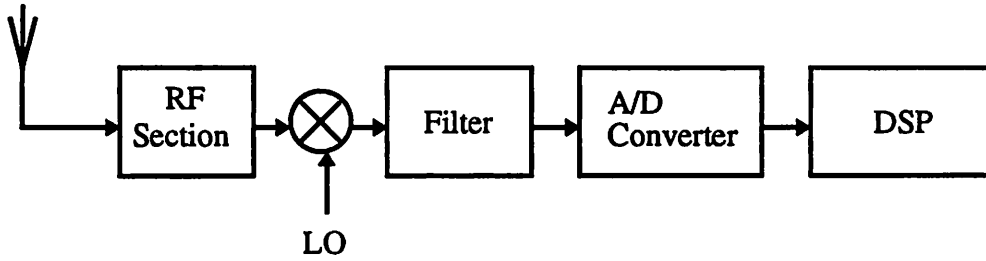
1.1 Motivation

Reduction of the power dissipation associated with high speed sampling and quantization is a major problem in many applications, including portable video devices such as camcorders, personal communication devices such as wireless LAN transceivers, in the read channels of magnetic storage devices using digital data detection, and many others as illustrated in Fig. 1. In the past, high-speed A/D converters required for these applications in the sampling rate range above 5 Msample/sec (MS/s) with 8 to 12 bit of resolutions have consumed large power ranging typically from 100 mW to 500 mW[7][8][9][11][12][13][14][15][17][18][19][22][23][24]. For battery-powered portable applications this level of power consumption may not be suitable, and further power reduction is essential for power-optimized A/D interfaces.

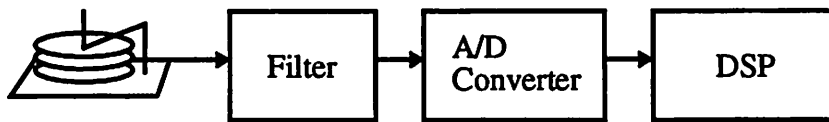
Low voltage operation is another important key factor in these portable A/D interface environments. With the trend that A/D interfaces are incorporated as a cell in complex mixed-signal ICs containing mostly digital blocks for DSP and control, the use of the same supply voltage for both analog and digital circuits can give advantages in reducing the overall system cost by eliminating the need of generating multiple supply voltages with DC-DC converters. Therefore, in order to be compatible with low-voltage systems, a new generation of A/D converters that can operate at supply voltage below 5 V



(a) Video-imaging systems



(b) Personal communication system



(c) Disk drive read channel

FIGURE 1. Examples of analog-to-digital interfaces.

is desired.

With recent improvements on higher speed and higher integration capability of the scaled technologies, a CMOS technology is becoming increasingly attractive as a cost-effective solution for many applications once reserved for bipolar or other fast technologies. This trend is expected to continue with scaled sub-micron CMOS technologies.

Among many types of CMOS A/D converter architectures, a pipeline architecture can achieve good high input frequency dynamic performances and as a high throughput as the flash ADC due to a S/H circuit in each stage of the pipeline for concurrent processing[2][5][6][8][9][12][13]. In this dissertation, both fundamental and practical limitations to the power dissipation in CMOS A/D converters are examined, and techniques to allow low power and low voltage operation of the pipeline architecture are described.

To verify the effectiveness of the techniques, a 10bit 20MS/s pipeline A/D converter is designed and fabricated in 1.2 μ m CMOS technology. The test results show that 59.1 dB of SNDR (signal-to-noise-plus-distortion ratio) can be achieved for the input frequency of 100kHz while the whole A/D converter dissipates only 35mW at 20MS/s. At 1MS/s and reduced bias current, the power dissipation is only 2.8mW with 58.0dB of SNDR.

1.2 Thesis Organization

In Chapter 2, fundamental limitations to the power dissipation in CMOS A/D converters are discussed by examining implementation issues on three key functions, sampling, quantization, and reference generation. Practical issues are also briefly discussed.

In Chapter 3, several high speed CMOS A/D converter architectures are reviewed. First, a flash A/D architecture is presented and its limitations are studied. Then, an attempt is made to present how ADC architectures have evolved to reduce power and area from the power dissipation point of view.

In Chapter 4, the pipeline architecture is presented in more detail from its basic operation to actual implementation of each pipeline stage.

In Chapter 5, techniques to reduce the power and to allow low-voltage operation of the pipeline architecture are presented.

An experimental prototype A/D converter has been fabricated, and its measurement results are presented in Chapter 6 along with discussions on key performances.

Finally, the conclusion is presented in Chapter 7.

CHAPTER 2

Fundamental Limitations to Power Dissipation of CMOS Analog-to-Digital Converters

2.1 Introduction

In electronic signal processing, the A/D conversion process involves *sampling* the applied analog input signal and *quantizing* it to its digital representation by comparing to *reference* voltages before further signal processing in subsequent digital systems. Depending on how these functions are combined, different A/D converter(ADC) architectures can be implemented with different requirements on each function. For instance, while the flash architecture requires many precision comparators, the pipeline architecture requires precision op amps. In order to implement power-optimized ADC functions, understanding the performance limitations of each function is important before discussing the system issues. In this chapter, the concept of the basic A/D conversion process and the fundamental limitation to the power dissipation of each key building block are presented.

2.2 Basic A/D Conversion Concepts

The basic concept of the A/D conversion process can be explained with a 3-bit

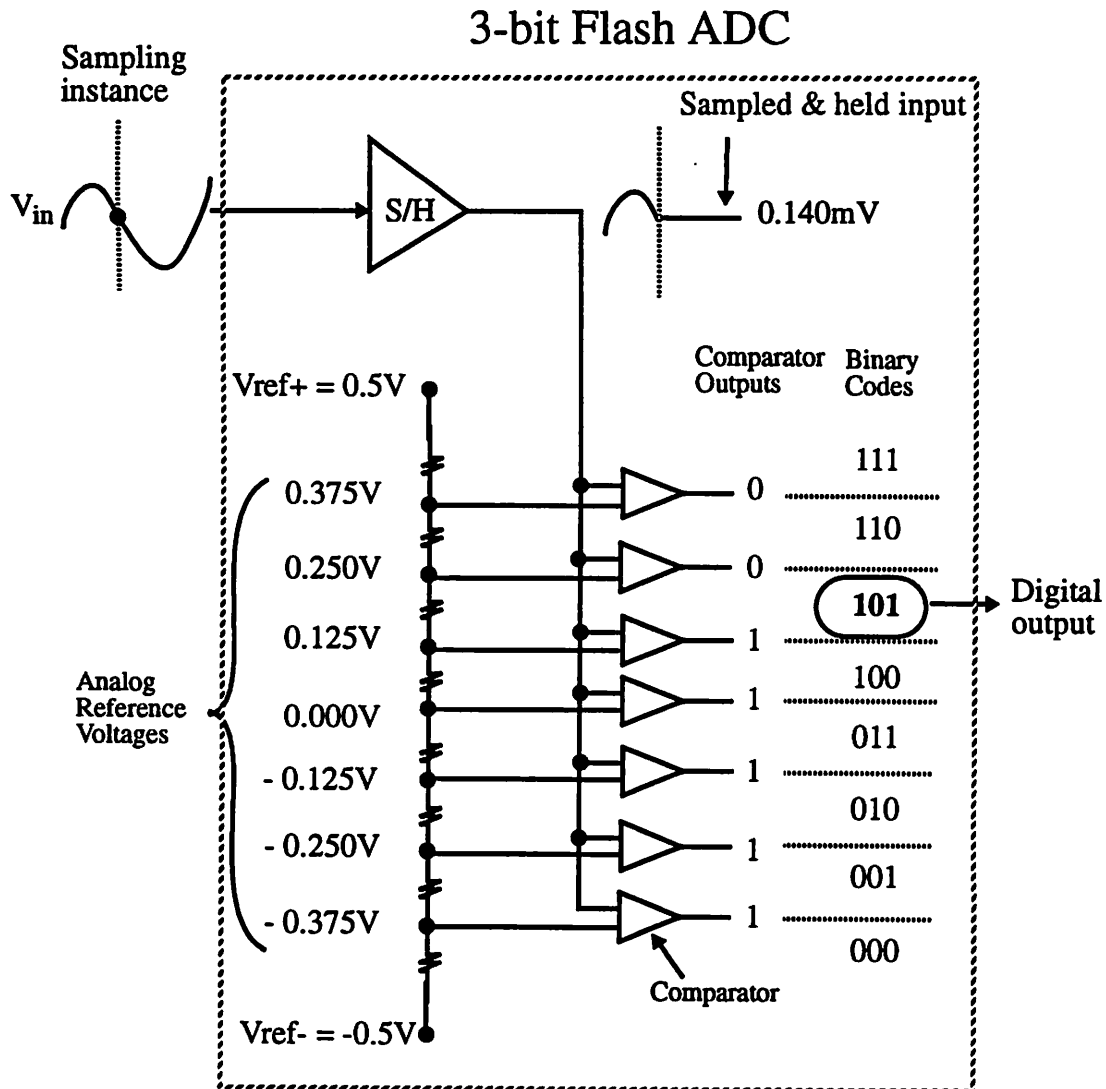


FIGURE 2. A/D conversion using a resistor string(N=3)

flash ADC shown in Fig. 2. When the continuous-time continuous-value input signal is applied, the input sample-and-hold (S/H) circuit¹ first samples the signal and holds the sampled amplitude constant for a period of time. During this time, comparators compare the held signal with reference voltages generated from the resistor string, and the resulting thermometer code from the comparator outputs is encoded into a digital binary

1. Not all ADC's have the input S/H circuit, and the limitations are discussed in section 3.2

representation. Then, the S/H circuit samples a new input voltage, and the whole conversion procedure repeats for the next sample.

Three key functions performed during this process are: *sampling*, *quantization*, and *reference generation*. The power dissipation associated with each function depends on its *accuracy* requirement. In the following sections, both fundamental and practical limitations on the performance of each function and how they relate to the power consumption are discussed.

2.3 Sampling in MOS Technologies

2.3.1 Basic MOS Sample/Hold Circuit

The function of the S/H circuit is to track/sample the analog input signal and to hold that value while subsequent circuitry digitizes it. In MOS technologies, this function is implemented by storing the input signal voltage on a sampling capacitor through a MOS transistor switch and holding the voltage for subsequent stages usually with some active circuitry such as op amps. Since the achievable precision of the S/H function is limited by the initial accuracy of the sampled signal, the fundamental accuracy is limited by the accuracy of the sampling circuit, not the active circuitry which holds the value.

The limitations of sampling can be studied with a simple MOS S/H circuit implemented with one MOS transistor and one capacitor as shown in Fig. 3. During the sampling phase of the clock, the voltage on the sampling capacitor C_S tracks the input voltage through the MOS transistor switch. Then, in the next clock phase when the clock V_g goes low, the transistor turns off and the input voltage is sampled and held on the capacitor for further processing.

In this simple MOS S/H circuit, a number of non-idealities produce errors, and

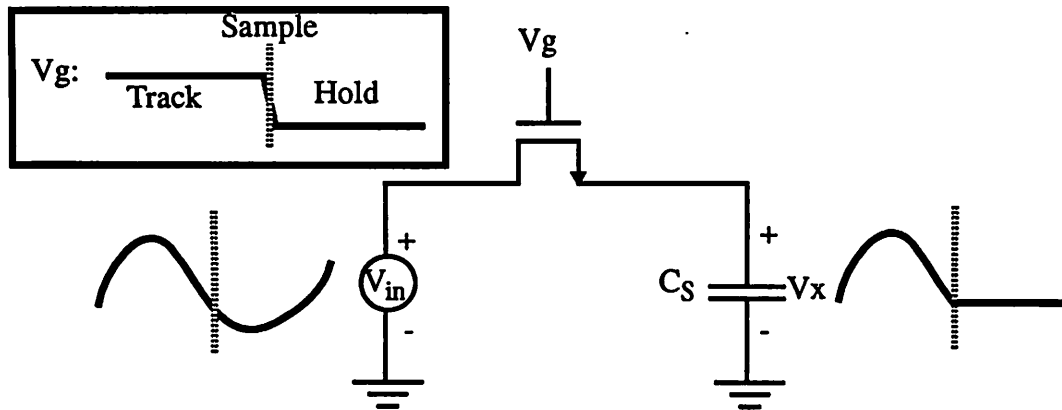


FIGURE 3. A simple MOS S/H Circuit

they can be categorized into two groups, deterministic components and random components. The term “deterministic component” refers to an error source whose relationship with the signal is known to be consistent from sample to sample, such as the finite bandwidth in the sample mode, the signal-dependant charge injection from the MOS transistor, clock feedthrough, etc. Various circuit techniques have been developed to cancel or to suppress these effects to achieve high sampling accuracy. In Table 1, error sources and possible solutions/techniques are shown. In [2], the accuracy up to 15 bits has been reported using these techniques. Therefore, deterministic components do not set the fundamental limit for the input sampling to the first order at least at resolutions in the 8-12 bit range¹. Brief discussions and references on the deterministic error components are

1. According to [10], it appears that there's no fundamental limit on the performance of the MOS S/H, especially at low input frequency. However, for high input/sampling frequency, various practical considerations (such as capacitive loading, offsets, etc.) set the limit to the achievable accuracy of the MOS S/H circuit, and further research is necessary in order to understand the device dynamics and to be able to design high-speed/high-resolution MOS S/H circuits.

presented in Appendix 1.

TABLE 1. Deterministic Error Components and Possible Solutions

Error Sources	Possible Solution/Techniques
Finite Bandwidth	Advanced technologies to lower the switch on-resistance Gate voltage bootstrapping[5][33]
Charge Injection	Bottom plate sampling[1] Dummy switch[45]
Clock Feedthrough	Differential signal path[1]

The other error components are “random errors”, errors that may be unpredictable from sample-to-sample, and the dominant source in the circuit of Fig. 3 is thermal noise. In conventional resistors, noise is generated due to the random thermal motion of electrons and is unaffected by the presence or absence of direct current[3]. Therefore, this noise appears as additive noise to the signal, and its mean-square value within the bandwidth Δf (in Hz) is given by

$$\overline{v^2} = 4kTR\Delta f , \quad (\text{EQ 1})$$

where k is Boltzmann’s constance and T is the temperature in Kelvin[3]. At room temperature $4kT = 1.66 \times 10^{-20}$ V-C.

Another noise source present in MOS transistor is the Flicker noise or “1/f noise” whose noise spectral density has a 1/f frequency dependence[3]. Because of its low frequency characteristics, there are techniques to suppress this noise especially for high

frequency sampling circuits [3][31][32], and the analysis from now on will be focused on thermal noise.

In the MOS S/H circuit, thermal noise is generated and added to the sampled signal due to the finite resistance of the MOS transistor switch. This is illustrated in Fig. 4. For

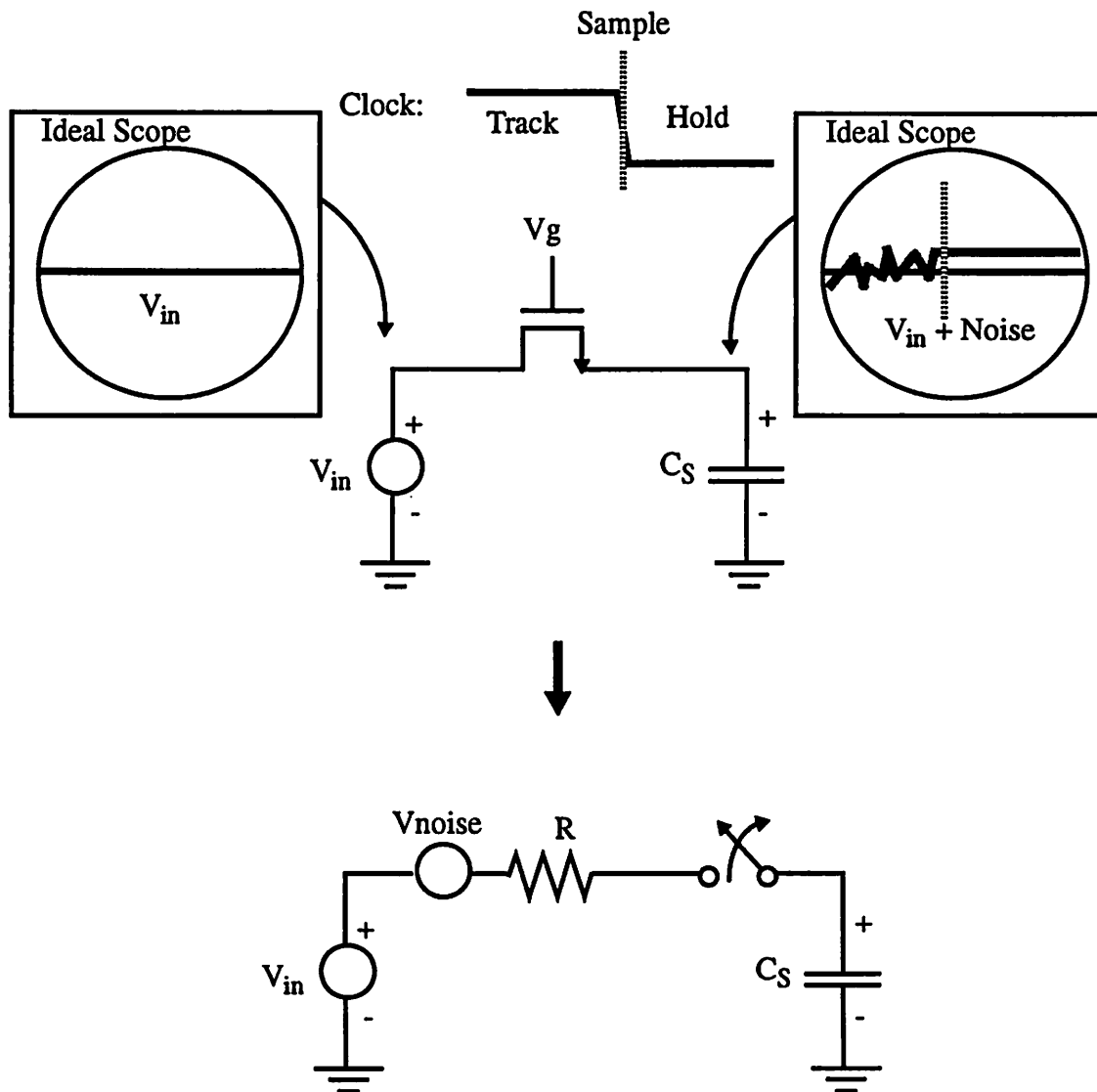


FIGURE 4. A simple MOS S/H circuit and its equivalent model for noise calculation.

illustration purpose, let's assume that the input signal is held at DC and the signal is sampled on a capacitor. In this case, the sampled voltage on the capacitor contains not only the signal component but also the thermal noise component at the instance of the sampling. Assuming single pole frequency response (R and C), the total noise variance can be found by integrating the noise spectral density over frequency and is given by

$$\sigma^2 = \int_0^{\infty} \frac{\overline{v^2}}{\Delta f} \cdot \frac{1}{\left|1 + \frac{jf}{f_{3dB}}\right|^2} \cdot df = \int_0^{\infty} 4kTR \cdot \frac{1}{\left(1 + \left(\frac{f}{2\pi RC_S}\right)^2\right)} \cdot df = \frac{kT}{C_S} \quad (\text{EQ 2})$$

where $f_{3dB} = \frac{1}{2\pi RC_S}$, R is the on-resistance of the MOS transistor, and C_S is the sampling capacitor value assuming that parasitic capacitance from the MOS switch is negligible compared to the sampling capacitor. Notice that this is independent of R because the increase in the mean-square value of the noise due to the increase in R value cancels the corresponding bandwidth reduction and the same expression results.

In the literature, this noise is often called “kT/C” noise (for the obvious reason), and Table 2 shows RMS values for the noise for different sampling capacitor values at room temperature.

TABLE 2. RMS values for the thermal noise for different sampling capacitor values at room temperature

C	$\sigma = \sqrt{kT/C}$
0.01pF	640μV
1pF	64μV
100pF	6.4μV

Due to the randomness from sample to sample, the error due to thermal noise cannot be corrected, and therefore it limits the achievable signal-to-noise-ratio(SNR) for a given sampling capacitor value. For instance, for C= 1pF and full scale input of $V_{FS} = +/- 1V$ sine wave, the SNR is given by

$$\text{SNR} = 10\log\left(\frac{V_{\text{FS}}^2/2}{\sigma^2}\right) = 10\log\left(\frac{1/2}{(64\mu\text{V})^2}\right) = 81\text{dB} , \quad (\text{EQ 3})$$

assuming that an infinite resolution ADC can resolve the held signal. In more realistic case, however, the resolution of the ADC is finite and its quantization noise dominates if thermal noise is much less than the quantization step. In that case, the quantization noise power[4] must be included in the noise term and the SNR is given by

$$\text{SNR} = 10\log\left(\frac{V_{\text{FS}}^2/2}{\frac{\Delta^2}{12} + \sigma^2}\right) = 10\log\left(\frac{V_{\text{FS}}^2/2}{\frac{(2V_{\text{FS}}/2^N)^2}{12} + \frac{kT}{C}}\right) , \quad (\text{EQ 4})$$

where Δ is the quantization step (magnitude of LSB) and N is the resolution of the ADC in bits. For $N=10$ and $C=1\text{pF}$, the SNR is 61.91dB in contrast to 61.96dB of noiseless ideal 10 bit ADC. In Fig. 5, the maximum achievable SNR is plotted for different sampling capacitor values at different resolution level. For a small sampling capacitor, thermal noise limits the SNR, and for a large sampling capacitor, the SNR is limited by the quantization noise and the curve flattens out.

In this case, the power dissipation associated with charging/discharging of the sampling capacitor depends on the input frequency, since the voltage on the capacitor varies at each sampling instance. The worst case power dissipation occurs when the input frequency is equal to the sampling frequency(f_s) and the maximum voltage (V_{DD} in this case) is sampled on the capacitor at every instance, and it is given by

$$\text{Power}_{\text{worst}} = C \cdot V_{\text{DD}}^2 \cdot f_s = C \cdot (2V_{\text{FS}})^2 \cdot f_s, \quad (\text{EQ 5})$$

assuming that $V_{\text{FS}} = V_{\text{DD}}/2$ and the capacitor must be completely discharged on each sampling period. Combining this result with (EQ. 4), the SNR can be related to the power

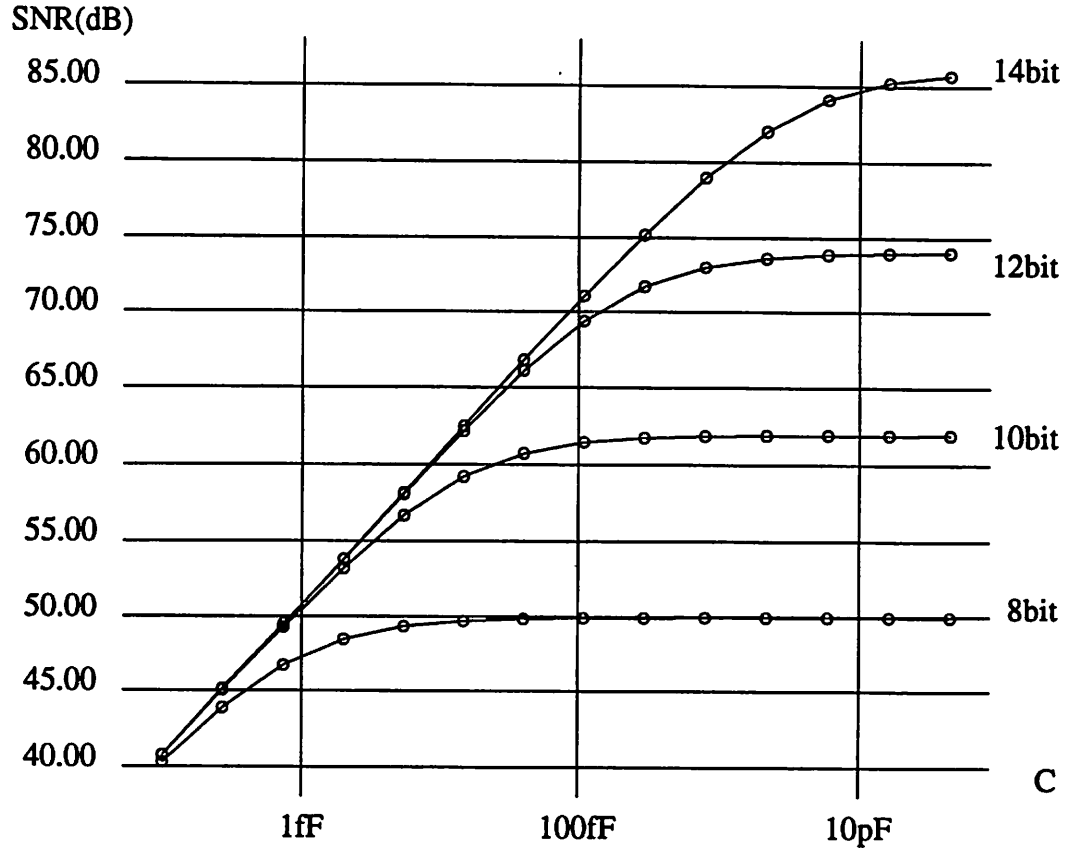


FIGURE 5. Maximum achievable SNR for different sampling capacitor values and resolutions.

dissipations by

$$\text{SNR} = 10 \log \left(\frac{1}{2} \cdot \frac{1}{\left(\frac{4}{12 \cdot 2^{2N}} + \frac{4kT \cdot f_s}{\text{Power}} \right)} \right), \quad (\text{EQ } 6)$$

and this relationship is plotted in Fig. 6.

According to (EQ. 6), the theoretical lower bound of the power dissipation for the simple S/H circuit in Fig. 3 is $0.2 \mu\text{W}/\text{MS/s}$ at 10 bit level when the SNR is degraded by 1dB due to kT/C noise relative to the quantization noise level.

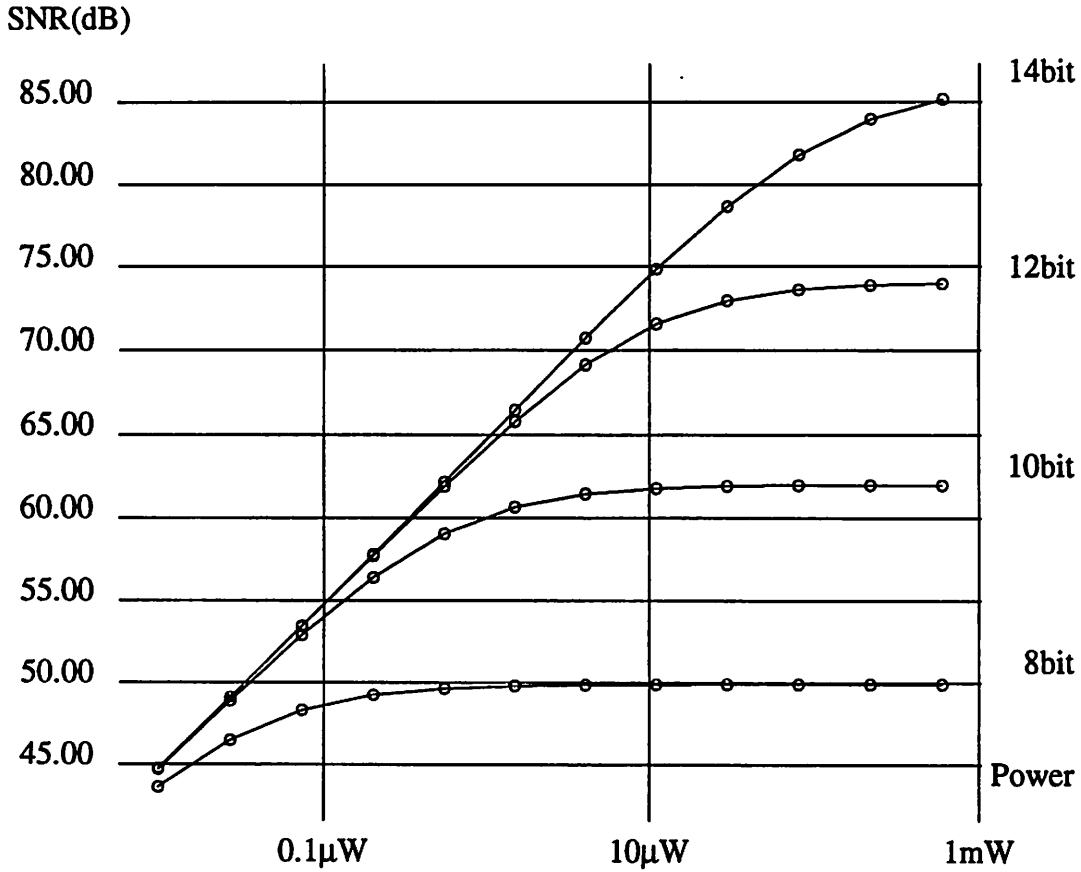


FIGURE 6. SNR vs. power dissipation according to (EQ. 6) for $f_s=10\text{MS/s}$.

Also notice that from (EQ. 6), the power dissipation is independent of the supply voltage. This is because the increase in the signal amplitude over the thermal noise level cancels out the increase in the power consumption due to the increased supply voltage. Therefore, from the fundamental point where the thermal noise limits the achievable SNR, the power dissipation in the S/H circuit does not depend on the supply voltage. When the RMS value for thermal noise degrades the SNR by 1dB over and above the quantization noise, the power dissipation of the S/H circuit from (EQ. 6) is given by

$$\text{Power} \approx 46.3 \cdot kT \cdot f_s \cdot 2^{2N} \text{W.} \quad (\text{EQ 7})$$

Therefore, if the input bandwidth is much higher than the signal bandwidth, the

power dissipated in the simple MOS S/H circuit is *linearly proportional to the sampling frequency* and it *quadruples for an additional bit*. At $f_s=10\text{MS/s}$, the power dissipations for 10 and 12 bit levels are about $2\mu\text{W}$ and $32\mu\text{W}$.

2.3.2 Practical MOS S/H Amplifiers: Op-Amp-Based SC S/H Circuits and Limitations

The power dissipation limit given by (EQ. 7) is about four orders of magnitude below the dissipation achieved in recently described high-speed A/D converters[5][8][9][11]. In practice, the S/H power is dominated by dissipation in the operational amplifier or buffer that drives the sampling capacitor in the sample and/or charge transfer modes. As a practical matter, power minimization in the sampling function translates to minimizing the power in the active circuitry driving the sampling capacitors whose kT/C noise limits the SNR of the converter.

In CMOS, the S/H circuits are usually implemented in an op-amp based switched capacitor(SC) circuit configuration, and minimizing the op amp power for the SC circuit involves the choice of the SC configuration and the op amp topology. Since there are a number of different configurations and op amp topologies, it is necessary to limit the scope of discussions here just on the key operation to present a bound for the actual performance limitations and to examine the general trend with different operating conditions, such as the choice of the supply voltage and the technology. A direct comparison between performances of different configurations involves many variables, and therefore it is omitted here. Detailed op amp topology comparison can be found in [43].

So, in this section three common SC S/H configurations are chosen and their basic operation and related key parameters are presented first assuming ideal op amps. Then, a basic configuration is identified, and discussions on its power dissipation are presented

using a single transistor op amp instead of the ideal op amp. In this way, a general conclusion on power dissipation can be made without introducing complicated op amp parameters.

A. Three Common Configurations for SC S/H Circuits.

Fig. 7, Fig. 8, and Fig. 9, show three common configurations for SC S/H circuits ([1][2][5][8][9][11][12][13][14][15], etc.). For simplicity, single-ended configurations are shown. Switch configurations shown in each figure are for the sampling phase, and the arrows indicate the switch configurations during the transfer (or hold) phase.

In all cases, the basic operations include sampling the signal on the sampling capacitor(s) and transferring the signal charge onto the feedback capacitor by using an op amp in the feedback configuration. In the configuration in Fig. 7, assuming an ideal op

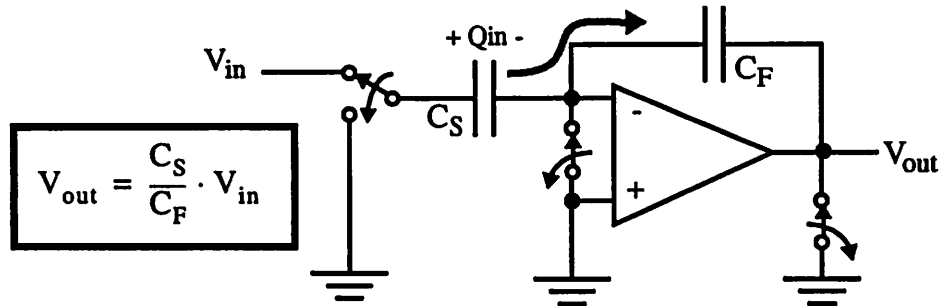


FIGURE 7. A SC circuit with separate C_S and C_F .

amp and switches, the op amp forces the sampled signal charge on C_S to transfer to C_F as indicated by the gray arrow. If C_S and C_F are not equal capacitors, the signal charge transferred to C_F will display the voltage at the output of the op amp according to

$$V_{out} = \frac{Q_{in}}{C_F} = \frac{C_S \cdot V_{in}}{C_F} = \left(\frac{C_S}{C_F} \right) \cdot V_{in}. \quad (\text{EQ 8})$$

In this way, both S/H and Gain functions can be implemented within one SC circuit, and examples can be found in [9][12].

In the configuration shown in Fig. 8, only one capacitor is used as both sampling capacitor and feedback capacitor. This configuration cannot implement the gain function, but it can achieve high speed because the feedback factor (the ratio of the feedback capacitor to the total capacitance at the summing node) can be much larger than that of the previous configuration, operating much closer to the unity gain frequency of the amplifier. This configuration is often used in the front-end input S/H circuit[8][13].

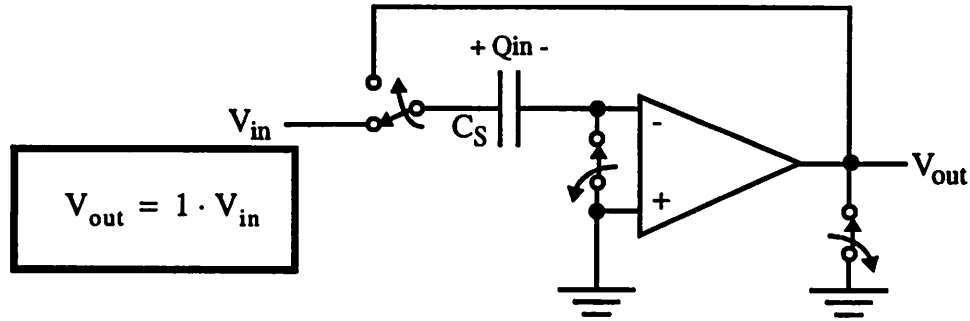


FIGURE 8. A SC circuit with one capacitor.

Fig. 9 shows another configuration which is a combined version of the configurations in Fig. 7 and Fig. 8. In this configuration, the signal is sampled on both C_S and C_F , and the resulted transfer function is

$$V_{out} = \frac{Q_{in}}{C_F} = \frac{C_F \cdot V_{in} + C_S \cdot V_{in}}{C_F} = \left(1 + \frac{C_S}{C_F} \right) \cdot V_{in}. \quad (\text{EQ 9})$$

In this configuration, C_F is also used as a sampling capacitor in order to improve

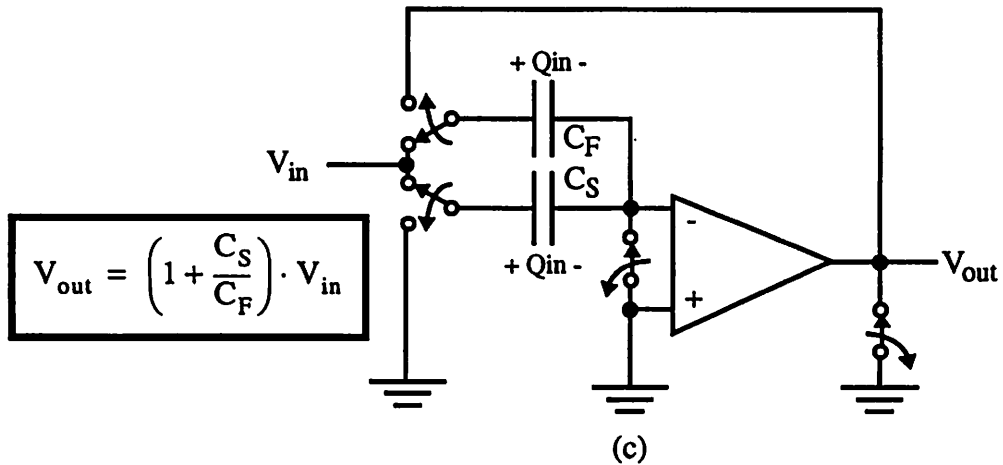


FIGURE 9. A SC circuit with C_F shared as a sampling capacitor.

the feedback factor. For instance, assuming that the closed loop gain is 2 and the op amp input capacitance (C_{opamp}) is ignored for simplicity, the feedback factor in this configuration is $C/2C = 0.5$, much larger than that of the configuration in Fig. 7 ($C/3C = 0.33$), which in turn results in 50% improvement in the SC circuit bandwidth[8].

Important parameters in determining the bandwidth of the SC circuit are G_m (transconductance of the op amp), feedback factor, and output load capacitance. In all of these three configurations, the bandwidth is given by:

$$BW = \frac{1}{\tau} = \frac{G_m}{C_{load}} \cdot f \quad (\text{EQ 10})$$

where C_{load} is the total capacitance seen at the op amp output.

Key parameters for three configurations are summarized in Table 3.

TABLE 3. Summary of key parameters for three common SC circuits.

Configurations	Transfer Function (V_{out}/V_{in})	Feedback Factor (f)
Fig. 7	$\frac{C_S}{C_F}$	$\frac{C_F}{C_S + C_F + C_{opamp}}$
Fig. 8	1	$\frac{C_S}{C_S + C_{opamp}}$
Fig. 9	$1 + \frac{C_S}{C_F}$	$\frac{C_F}{C_S + C_F + C_{opamp}}$

B. A Single Transistor Op Amp SC Circuits/Power Dissipation

Up to now, discussions presented for the SC configurations are based on the assumption that op amps are ideal. As mentioned previously, the power dissipation in a SC circuit is dominated by the op amp power, and a more realistic op amp must be considered in order to examine its power dissipation and its dependency on other parameters such as technology, supply voltage, etc.

Analysis for power involves a number of variables such as the SC configuration¹, op amp DC gain, op amp bandwidth, op amp input/output range, etc., and therefore is not trivial. Also, since there are a number of different op amp topologies (telescopic, folded-cascode, two-stage with Miller compensation, three stage with nested Miller compensation, class AB, etc.), it is more difficult to make direct comparisons on power among different topologies including all variables into considerations.

1. The analysis given here focuses on the SC gain configurations where C_S can be comparable to C_F . Detailed analysis on power dissipation of the general SC integrator case where C_F is much larger than C_S can be found in [46].

To a first order approximation, however, the trade-off between power, speed, and noise can be analyzed using a single transistor op amp, and it can give some insights on the power consumption of the SC circuit for different operating conditions. Several other assumptions are made to simplify this analysis.

1. A NMOS transistor is assumed for the op amp, and it is biased in the saturation mode with a square law I-V characteristic. The gate length of the device is assumed to be $1.2\mu\text{m}$ with $k_p=60\mu\text{A}/\text{V}^2$ and $C_{ox}=1.53\text{fF}/\mu\text{m}^2$. Based on this assumption, important parameters (such as g_m , C_{GS} , etc.) can be defined as a linear function of a bias current with fixed current density through the transistor. Detailed explanation on this will be presented later. For sub-micron devices whose I-V characteristic is more linear, the analysis can be easily repeated just by changing the proportionality constants according to velocity saturated I-V characteristic.
2. The signal swing is assumed to be equal to the supply voltage, the maximum voltage swing available in the system. Although the input range of many ADC's is only some fraction of the supply due to a limited op amp output swing, it is a reasonable assumption for the analysis here since the signal swing approximately scales with the supply voltage.
3. Ideal switches are assumed for simplicity. In reality, however, it will affect the bandwidth of the amplifier and must be considered. Therefore, this assumption gives theoretical upper bound on the achievable bandwidth. Brief discussion on this is presented in section 2.3.3.C.

4. It is assumed that the speed of the SC circuit is limited mainly by the time constant, and the slew rate constraint is not considered here. In reality this may or may not be true depending on the load condition. However, again it gives the basis for the theoretical upper bound on the achievable bandwidth.

With these assumptions, the SC configuration shown in Fig. 10 can be analyzed. In this case, the equivalent small signal model is shown in Fig. 11, and the time constant is given by:

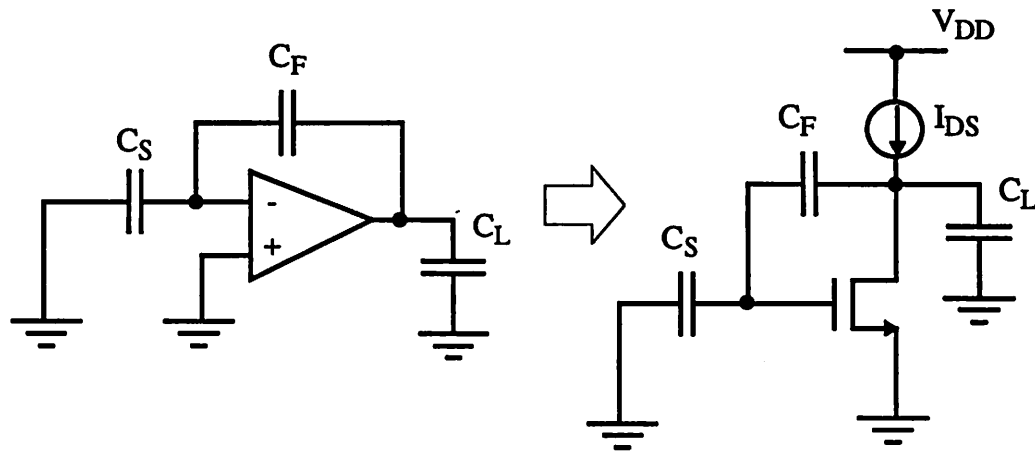


FIGURE 10. A SC circuit implemented with a single transistor op amp.

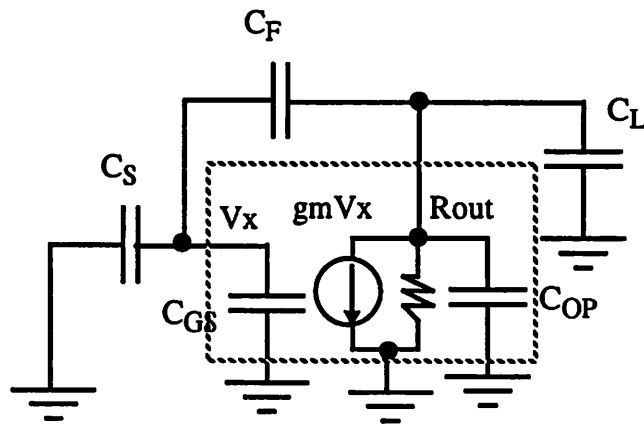


FIGURE 11. The equivalent small signal model for the circuit in Fig. 10.

$$\tau = \left(\frac{C_{LT}}{gm} \right) \cdot \frac{1}{f} = \left(\frac{C_L + C_{OP} + \frac{C_F \cdot (C_S + C_{GS})}{C_F + C_S + C_{GS}}}{gm} \right) \cdot \frac{C_S + C_{GS} + C_F}{C_F}, \quad (\text{EQ 11})$$

where C_{LT} is the total capacitance the op amp sees at its output, C_{OP} is the junction capacitance at the drain of the transistor op amp, and the C_{GS} is the input gate capacitance. The overlap capacitances are ignored for simplicity. The third term for C_{LT} is the capacitance loading at the output from the series combination of C_F and $C_S + C_{GS}$.

At this point, one important design consideration is the choice of V_{dsat} of the transistor. In real op amps, the output swing and the DC gain requirements set the maximum allowable V_{dsat} ($=V_{gs} - V_{th} = \sqrt{(2I_{ds}) / (k_p \cdot W/L)}$) for the transistor, and typical values are 200-300mV (can be higher depending on the available supply voltage and the swing requirement). For instance, if the DC gain greater than 50 is required from a single transistor with V_A (early voltage) of 10V, required V_{dsat} can be found according to

$$A_V = gm \cdot R_{out} = \frac{2I_{ds}}{V_{gs} - V_{th}} \cdot \frac{V_A}{I_{ds}} = \frac{2V_A}{V_{dsat}}, \quad (\text{EQ 12})$$

and the required value is 400mV. If the biased V_{ds} is larger than this value with extra 200-300mV, then the transistor will be safely biased in the saturation region to get the maximum transconductance. On the other hand, reducing V_{dsat} also means reducing the bandwidth of the device (operating close to subthreshold bias condition), and therefore, a careful choice of V_{dsat} is very critical.

Setting V_{dsat} is equivalent to setting the current density of the device, since the V_{dsat} is related to the current density $\rho (=I_{ds}/W)$ by:

$$V_{dsat} = \sqrt{\frac{2I_{ds}}{k_p \cdot W/L}} = \sqrt{\frac{2 \cdot \rho \cdot L}{k_p}}, \quad (\text{EQ 13})$$

and as a result, the choice of bias current sets the device size and other key parameters such as C_{GS} , C_{OP} (drain junction parasitic capacitance), and g_m by the following equations.

$$C_{GS} = C_{ox} \cdot W \cdot L = C_{ox} \cdot (I_{ds}/\rho) \cdot L \quad , \quad (EQ 14)$$

$$C_{OP} = \alpha \cdot C_{GS}, \quad (EQ 15)$$

$$\text{and } g_m = \sqrt{\frac{2 \cdot k_p}{\rho \cdot L}} \cdot I_{ds}. \quad (EQ 16)$$

Here, in (EQ. 15), the drain junction capacitance is assumed to be proportional to the input capacitance and its proportionality constant α depends on the geometry of the layout. The typical value for α ranges from 0.5 to 1 for a single MOS device to 2 or 3 for MOS amplifiers where the parasitic capacitance from both PMOS and NMOS has to be considered[16]. Again, for simplicity, α is assumed to be 1 for the analysis given here.

Now, having all the key parameters set for a given current density, the power dissipation of the single-transistor op amp SC circuit can be found by relating its bias current to the given settling requirement. From (EQ. 11), the achievable time constant can be found for the given sampling/feedback capacitor values, current density, and load capacitance. According to the detailed analysis carried out in [16], the closed-loop time constant is given by:

$$\tau = \frac{C_{GS}}{g_m} \cdot \left(2\sqrt{\alpha} \sqrt{\frac{C_L C_F + C_L C_S + C_F C_S}{C_F^2}} + 1 + \frac{C_L}{C_F} + \alpha \left(1 + \frac{C_S}{C_F} \right) \right), \quad (EQ 17)$$

where the optimum size input transistor for minimum settling time has an input gate capacitance of

$$C_{GSopt} = \sqrt{\frac{C_L C_F + C_L C_S + C_F C_S}{\alpha}} \quad (\text{EQ 18})$$

The above expression is rather complicated and does not provide immediate intuition as it is. By limiting the scope of the discussion to more realistic cases[5][6][8], let's consider one specific configuration where the closed loop gain is 2 with $C = C_S = C_F = C_L$. Also, ρ is assumed to be $1\mu\text{A}/1\mu\text{m}$ which corresponds to V_{dsat} of 200mV. Then, there is an optimum value for the sampling capacitor size for the minimum time constant (or maximum bandwidth) for the given bias current as illustrated in Fig. 12. The time constant

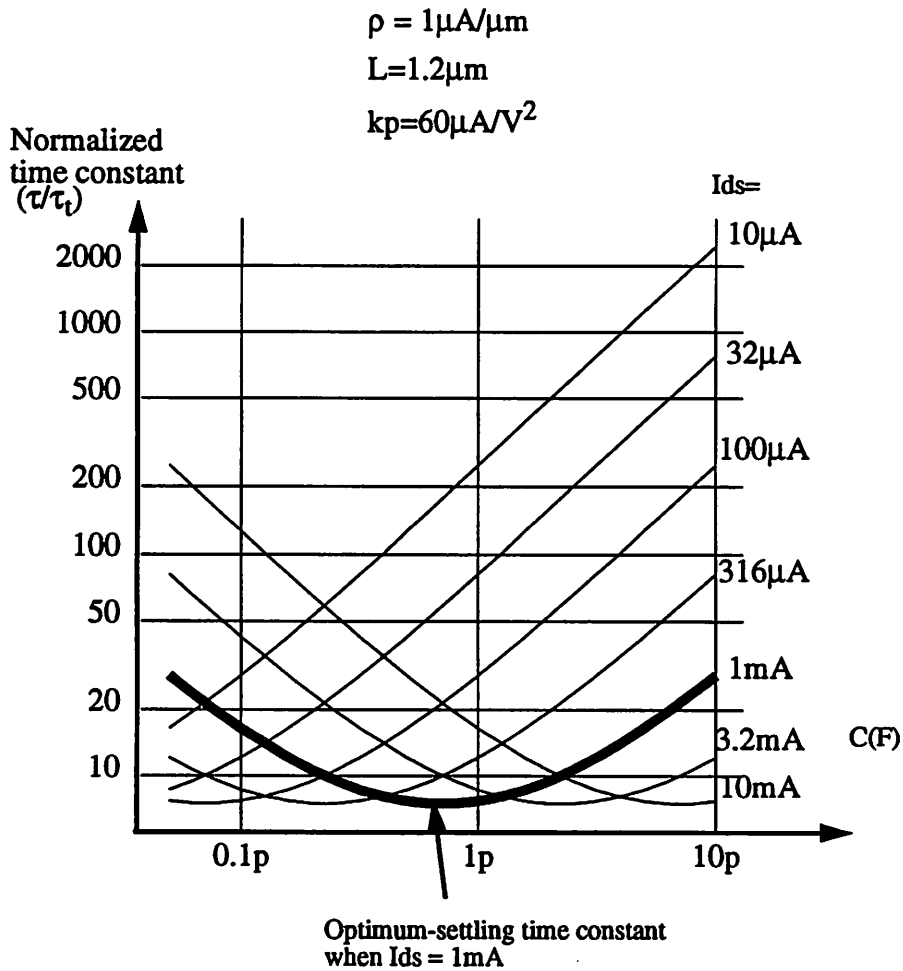


FIGURE 12. Closed-loop time constant vs. sampling capacitor value for different bias currents. The closed loop gain is 2, $\alpha = 1$, and $C = C_S = C_F = C_L$.

is normalized to the $\tau_i (=1/f_{t_{intrinsic}})$ of the device which is approximately $(C_{GS}/gm)^1$. For a given bias current, increasing the sampling capacitor value (and all other capacitors except C_{GS} according to the above assumption) increases the feedback factor $(=C/(2C+C_{GS}))$ since the device size (or C_{GS}) is given with a fixed current density, and in turn decreases the settling time. If C is too big, however, the output capacitance loading from C_L and the series combination of C_F and $(C_S + C_{GS})$ increases, and the settling time increases. The optimum time constant remains constant regardless of the SC circuit size

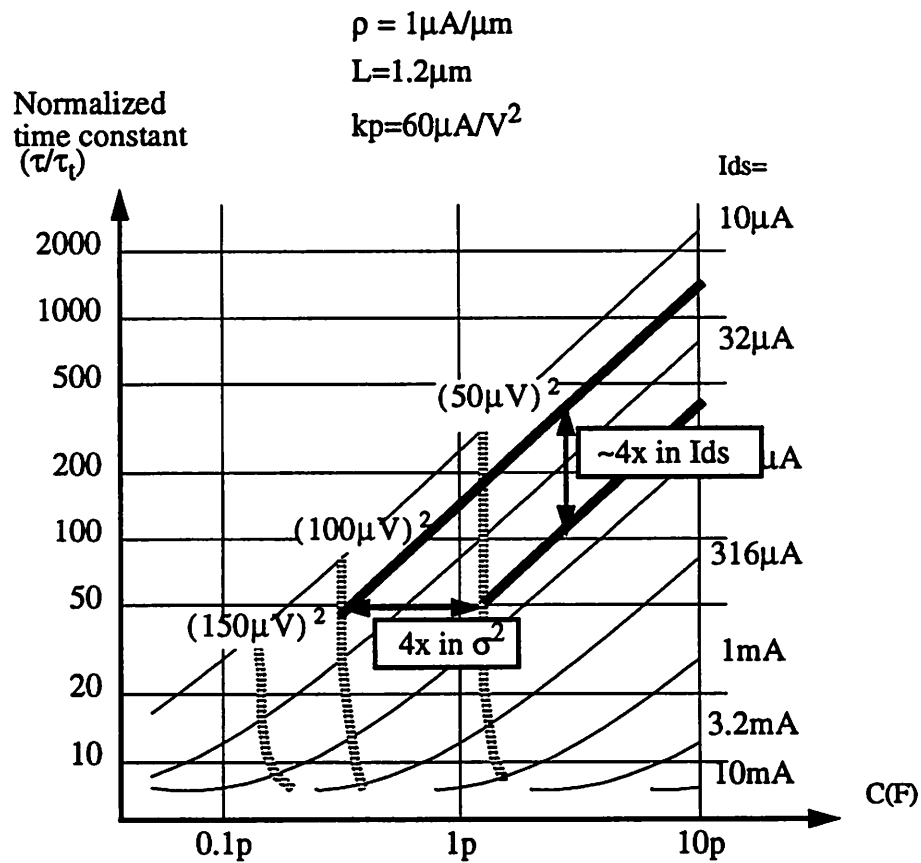


FIGURE 13. Closed-loop time constant vs. sampling capacitor value for different bias currents with noise contours. The closed loop gain is 2, $\alpha = 1$, and $C = C_S = C_F = C_L$.

(or I_{ds}) because C_L scales together with C_S and C_F . So, if speed is the only constraint, the

1. C_{GD} is ignored from previous assumption. If $C_{overlap}$ is large, then $(1/f_{t_{intrinsic}}) = (C_{GS} + C_{GD})/gm$.

power dissipation can be reduced by scaling down the capacitor size until the speed is limited by other practical considerations, such as layout, matching, etc. At each point on the curve, there are two possible bias conditions, and only the curve with a positive slope will be considered here on since it is the low power solution.

Now, including the noise requirement into consideration, the minimum allowable value for C_S can be set for a given bias current. Assuming that the dominant noise is the thermal noise, the input referred noise power is given by¹:

$$\overline{v_n}^2 = \frac{kT}{4C} \left(\frac{1}{f} + \frac{2}{3} \frac{2C + C_{GS}}{C_{LT}} \right) , \quad (\text{EQ 19})$$

where the first term is the noise sample on the sampling capacitors and the second term is the noise from the op amp, and the result is shown in Fig. 13. The dotted lines are contours for different noise levels.

From Fig. 13, it can be seen that 4x reduction in noise power corresponds to approximately 4x increase in power since capacitors need to be increased by the same ratio for a given time constant. Therefore, choosing the minimum size capacitor for a given noise requirement is essential for low power dissipation.

From the discussion presented in this section, the following can be concluded.

1. If noise is not a constraint, the sampling capacitor value that gives the minimum time constant can be chosen for minimum power dissipation. For instance, if $C = C_S = C_F = C_L$ and $\alpha=1$, the optimum value for C is $0.577C_{GS}$ (or $C_{GS}=1.73C$ conversely) according to (EQ. 18). In this case, because C_L scales with C_S and C_F , the optimum time constant remains constant regardless of the SC circuit size (or I_{ds}). Since the noise is not a

1. See Appendix 2 for derivation.

constraint and the minimum achievable time constant remains constant with the fixed ratio of C/C_{GS} , the SC circuit power can be reduced by scaling down both capacitors and an op amp (C_{GS}) until limited by other practical considerations, such as layout issues, matching issues, etc. More detailed analysis can be found in [16].

2. If noise is an important constraint (as in high resolution front end S/H circuits), an appropriate sampling capacitor size must be first chosen in order to reduce its kT/C noise level down below a given noise requirement. Then, the op amp size and its bias current can be determined for a given speed requirement and minimum power dissipation using τ -vs.- C curves as in Fig. 13(Fig. 13 is an example for a particular case where the closed loop gain is 2, $\alpha = 1$, and $C=C_S=C_F=C_L$). Notice that for low frequency operation (where τ/τ_t is large) the C_{GS} that achieves the minimum power dissipation for given settling time and noise requirements usually does not correspond to the minimum time constant point. This is because fixing the C/C_{GS} ratio of the SC circuit to the minimum time constant point(0.577 in this case) requires larger C_{GS} resulting in power increase and excessive bandwidth. Near the speed limit of the given technology (where the ratio τ/τ_t is small), however, the difference in power between the minimum power point and the minimum time constant point becomes smaller as the stringent settling time requirement forces the C/C_{GS} ratio to be at its optimum value to achieve the maximum bandwidth.

3. For a given speed requirement and signal swing, a 2x reduction in noise voltage (in σ) requires a 4x increase in the sampling capacitance value and the op amp size. Conversely, a 2x increase in the supply voltage and the signal swing results a 4x smaller SC circuit, and therefore, a 2x smaller op amp power dissipation. This means that the S/H circuit power *quadruples* for every additional bit resolved for a given speed requirement and supply voltage.

4. The only technology dependent term in (EQ. 19) is C_{GS} (the feedback factor is also a function of C_{GS}). If $(V_{gs}-V_{th})$ is fixed to meet an output swing requirement and g_m

(proportional to $C_{ox}(W/L)$) is fixed to meet a speed requirement, C_{GS} will be proportional to L^2 . Then, scaling of L with advanced CMOS technologies in the future will reduce C_{GS} and, (EQ. 19) will be bounded by

$$\overline{v_n}^2 \approx \frac{kT}{4C} \left(2 + \frac{4}{9}\right) = \frac{kT}{C} \frac{11}{18} \quad (\text{EQ 20})$$

when C_{GS} goes to 0. The result suggests that the noise is mostly from the sampling capacitor given its weak dependence on C_{GS} (in this case, with the ratio of 2 to 4/9), and further reduction on the capacitor value is not expected even with scaled technologies as a result.

2.3.3 Additional Practical Design Considerations

Other important requirements for the SC circuit include DC gain, slew rate, and switch resistances. Depending on the specifications set by the system, the S/H circuit design may be influenced more by these requirements than just minimizing noise and settling time constants.

A. DC Gain Requirement.

For the SC configuration in Fig. 7, if the op amp is ideal, the input/output transfer curve must be a straight line with a slope of C_S/C_F . However, in reality, the finite dc gain of an op amp introduces gain error in actual input/output transfer curve. Fig. 14 shows the transfer curves with and without gain error. The transfer curve can be related to the op amp dc gain by

$$\frac{V_{out}}{V_{in}} = \frac{C_S}{C_F} \cdot \frac{1}{1 + \frac{1}{A \cdot f}} \quad (\text{EQ 21})$$

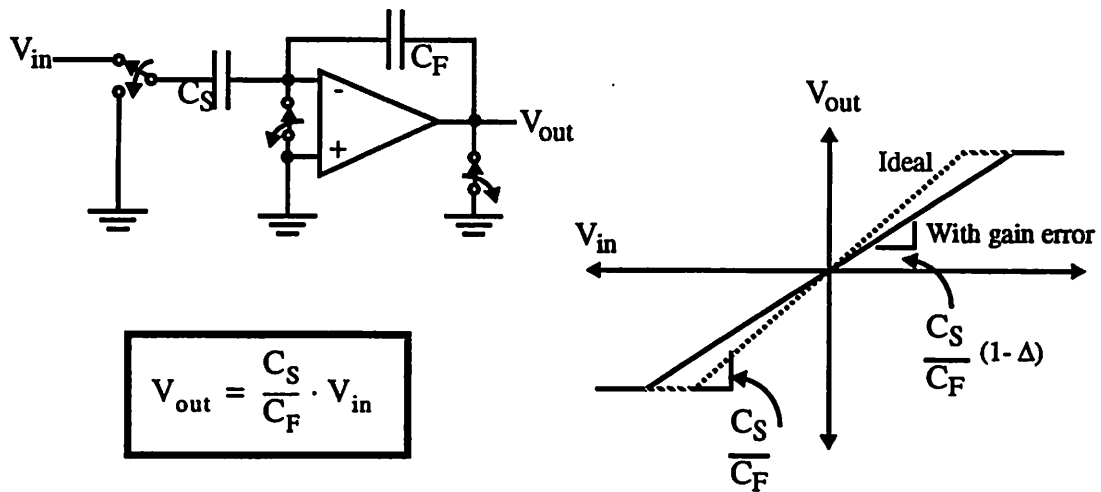


FIGURE 14. A transfer curve with gain error for the SC circuit in Fig. 7.

where f is the feedback factor and A is the DC gain of the amplifier. Therefore, if the product of Af which represents the loop gain of the feedback system is low, the gain will be less than the capacitor ratio of C_S/C_F . For the input S/H circuit, the gain error can be tolerated if the A/D conversion does not require absolute scale. This can be modelled as having a linear gain function in front of ideal S/H SC circuit as shown in Fig. 15. If the

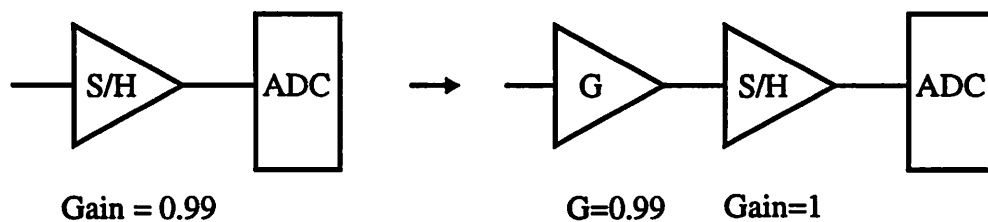


FIGURE 15. The front-end S/H circuit with a gain error does not affect the accuracy of the ADC as long as the gain is linear.

gain is linear, then it does not introduce any error except reducing the signal amplitude by a small portion. As an example, for the feedback factor of $1/3$, $C_S=C_F$ and $A=300$, the

gain is 0.99 instead of 1. Therefore, if the quantization is not based on the absolute scale, scaling of the input signal by 1% error is tolerable in most cases. However, there are applications where amplifiers need to have an exact gain to very high accuracy within +/- 0.1% of the nominal value, such as the front-end interstage amplifier in the high resolution(>8bit) pipeline A/D converters. In that case, the DC gain of the op amp must be larger than 60-80dB, or the capacitors must be trimmed to compensate the error due to insufficient op amp DC gain by changing the capacitor ratio. This can be done by having trim capacitor arrays in parallel with either C_S or C_F and switching them accordingly. The latter solution, however, usually requires extra high precision circuitry or some systematic methods to measure the relative capacitor values[5][9]. Therefore, the gain requirement may require a multi-stage op amp and can result in large power dissipation.

Another important factor is distortion. In actual op amps, the transfer curve is not a straight line but rather a line with some curvature and can cause harmonic distortions. In the SC circuit, the negative feedback around the op amp reduces its distortion by the loop gain, A_f , to obtain a highly linear transfer characteristic[3]. In an ADC, the distortion in the S/H circuit will result in large integral non-linearity error(INL) causing harmonic distortion and intermodulation distortion. For the high resolution ADC's with resolution greater than 10bits, the S/H circuit may require high dc op amp gain (~60 - 100dB) because of this reason.[1][2][5][8][9][11][13].

B. Slew Rate

A settling time of a typical op amp consists of two time periods, one limited by the time constant of the circuit and the other by the slew rate as illustrated in Fig. 16. When the SC circuit has to drive large load capacitance with a large amplitude signal, the bias current of the op amp must be chosen such that it produces large transconductance and necessary slew rate. For instance, if the circuit has to drive off-chip components with ~10 - 100pF, the required op amp bias current must be large enough in order to prevent

gain is 0.99 instead of 1. Therefore, if the quantization is not based on the absolute scale, scaling of the input signal by 1% error is tolerable in most cases. However, there are applications where amplifiers need to have an exact gain to very high accuracy within +/- 0.1% of the nominal value, such as the front-end interstage amplifier in the high resolution(>8bit) pipeline A/D converters. In that case, the DC gain of the op amp must be larger than 60-80dB, or the capacitors must be trimmed to compensate the error due to insufficient op amp DC gain by changing the capacitor ratio. This can be done by having trim capacitor arrays in parallel with either C_S or C_F , and switching them accordingly. The latter solution, however, usually requires extra high precision circuitry or some systematic methods to measure the relative capacitor values[5][9]. Therefore, the gain requirement may require a multi-stage op amp and can result in large power dissipation.

Another important factor is distortion. In actual op amps, the transfer curve is not a straight line but rather a line with some curvature and can cause harmonic distortions. In the SC circuit, the negative feedback around the op amp reduces its distortion by the loop gain, A_f , to obtain a highly linear transfer characteristic[3]. In an ADC, the distortion in the S/H circuit will result in large integral non-linearity error(INL) causing harmonic distortion and intermodulation distortion. For the high resolution ADC's with resolution greater than 10bits, the S/H circuit may require high dc op amp gain (~60 - 100dB) because of this reason.[1][2][5][8][9][11][13].

B. Slew Rate

A settling time of a typical op amp consists of two time periods, one limited by the time constant of the circuit and the other by the slew rate as illustrated in Fig. 16. When the SC circuit has to drive large load capacitance with a large amplitude signal, the bias current of the op amp must be chosen such that it produces large transconductance and necessary slew rate. For instance, if the circuit has to drive off-chip components with ~10 - 100pF, the required op amp bias current must be large enough in order to prevent

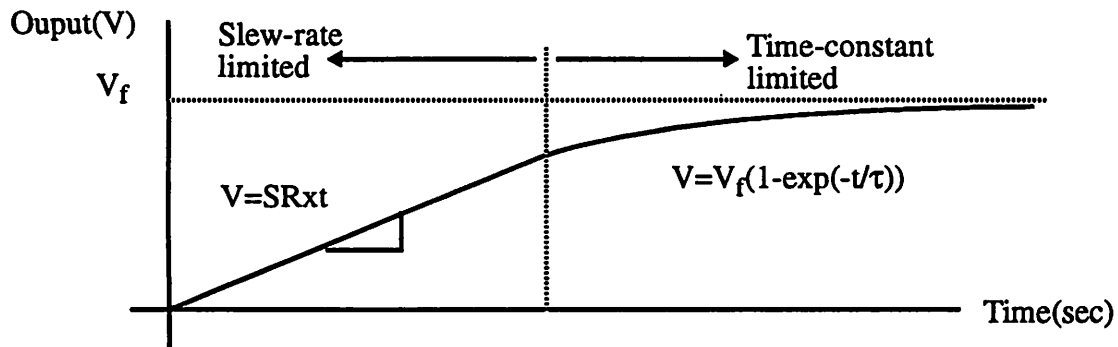


FIGURE 16. Voltage waveform at the output of a SC circuit.

excessive settling time due to slewing. In this kind of situation, a class A/B type of op amp may be more efficient than class A in terms of power. However, for cases driving on-chip circuitries where input capacitance is small (much less than $\sim 10\text{pF}$), a class A type of op amps usually achieve faster settling due to its simple architecture. In that case, a sufficiently large value of the op amp bias current is chosen to make the settling time limited more by the time constant rather than the slew rate; otherwise, signal-amplitude dependent settling will cause errors[1].

C. Switch Resistance

In actual implementations of SC circuits, on-resistance of the MOS switches can have a significant effect on the settling time of the circuit. High on-resistance in MOS switches can not only slow down the circuit but also make the feedback system poorly damped or unstable if it is in the feedback loop as shown in Fig. 17. This results from the increase in phase shift by increasing the delay and thereby reducing the phase margin. In order to avoid this type of situation, low enough on-resistances of the switches are required. Using a too large switch, however, adds significant amount of drain/source junction parasitic capacitance at the output reducing the overall bandwidth.

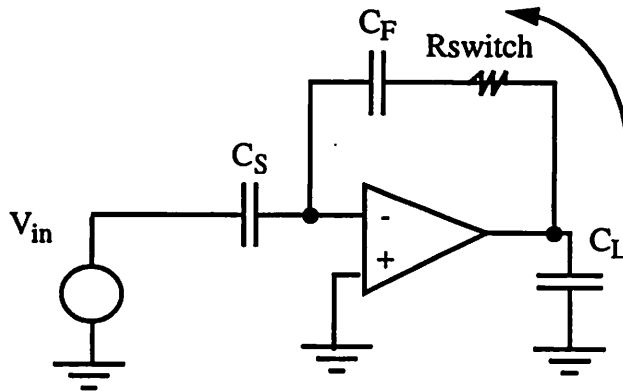


FIGURE 17. The switch resistance in the feedback loop can make the system unstable.

Other switches such as the summing switch(S1) and the input switches(S2 and S3) affect the sampling bandwidth of the input signal forming a RC network as shown in Fig. 18. A caution must be taken here to prevent excessive charge injection from large

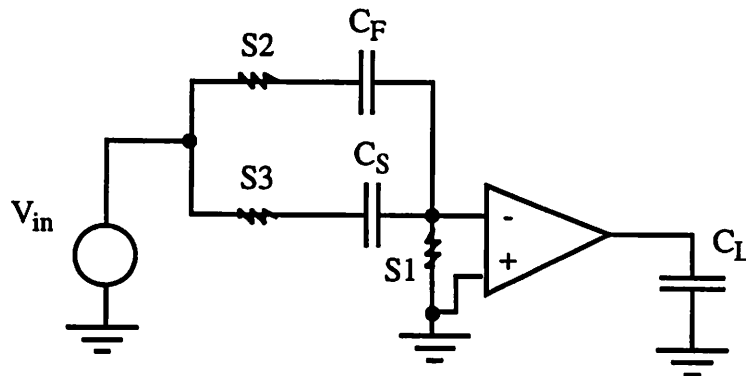


FIGURE 18. The on-resistances of S1, S2, and S3 limit the sampling bandwidth along with the sampling capacitors.

switches at the sampling instance. Charge injection is not a problem for S2 and S3 since they will be shorted out, usually to some reference voltage source. However, the amount of charge injection from S1 must be controlled, since it will change the input common mode voltage by $\Delta Q/C_T$ where C_T is the total capacitance at the summing node. For amplifiers which have limited input common mode range, such as a telescopic op amp, the

bias condition of the input device or current source can be changed affecting the overall operation as illustrated in Fig. 19. Since the sampling bandwidth considerations require a

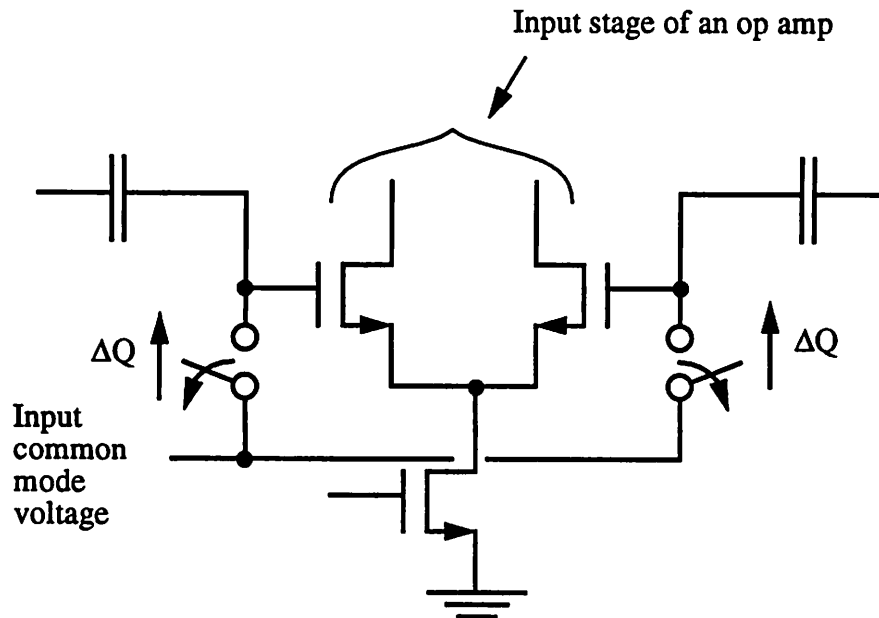


FIGURE 19. Input common mode voltage change due to charge injection from the summing node switches.

certain size for the input switch for a given capacitor size, the value of the input common mode voltage shift, $\Delta Q/C_T$, can be tens or hundreds of mV. For low voltage systems where the V_{DS} of the transistors for current sources is biased only 200-300mV above $V_{gs}-V_{th}$, these charge injection and headroom issues must be carefully considered.

2.4 Quantization

The next key function is quantization. The power dissipation associated with this process again depends on the accuracy requirement just like the S/H circuit. The uncertainty of the quantization comes from thermal noise and offsets.

2.4.1 Limitations on Comparators

The function of a comparator is to compare the applied input signal voltage to a reference voltage, and the simplest way to implement a comparator is to use simple regenerative cross-coupled inverters as shown in Fig. 20. In this case, internal nodes, X and Y are initialized with a voltage equal to or proportional to the difference of the input voltage to the reference voltage, and the circuit is placed in the regenerative mode to determine the polarity of the initialized voltage. Then, the signal is further amplified to the digital logic level for subsequent processing.

The key requirement which determines the power dissipation during this process is the accuracy; how accurately the comparator can make a decision in a given time period. For high resolution comparison where the signal needs to be resolved down to a couple of millivolts or less (typically 6 bit accuracy or higher), the comparator in Fig. 20 can not be

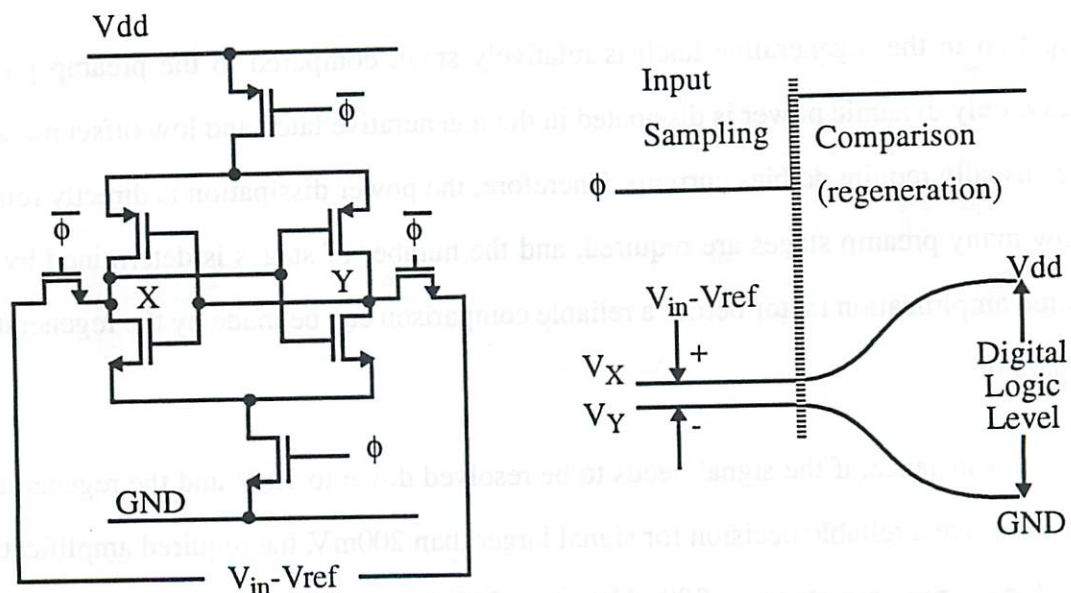


FIGURE 20. A simple regenerative comparator and its operation.

used as it is, because typical cross-coupled latch comparators exhibit offset voltages as large as 100mV. Instead several pre-amplifiers are placed before the regenerative latch to amplify the signal for accurate comparison as shown in Fig. 21. In this case, the power

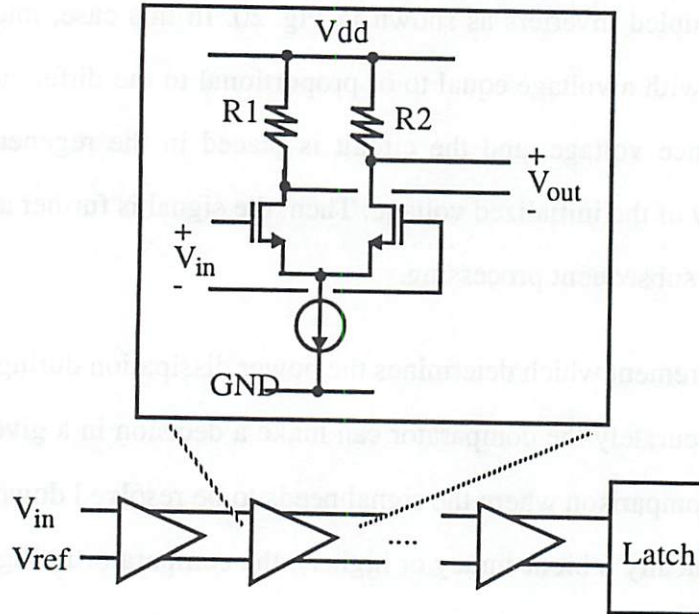


FIGURE 21. A typical preamp stage implemented with a source-coupled pair.

dissipation in the regenerative latch is relatively small compared to the preamp power, because only dynamic power is dissipated in the regenerative latch and low offset pre-amp stages usually require dc bias currents. Therefore, the power dissipation is directly related to how many preamp stages are required, and the number of stages is determined by the required amplification factor before a reliable comparison can be made by the regenerative comparator.

For instance, if the signal needs to be resolved down to 1mV and the regenerative latch can make a reliable decision for signal larger than 200mV, the required amplification factor for the pre-amp stages is $200\text{mV}/1\text{mV} = 200$. In order to get this high gain from a single stage preamp, the large value of the load resistor must be used and in turn slows down the amplification process with an increased RC-constant at the output. In situations like this, the gain is distributed among several cascaded low gain stages to speed up the

whole process[47].

During this process care must be also taken to design a low noise pre-amp stage since its own circuit noise is amplified through its gain. For instance, if the input signal is held constant close to the comparator threshold, the thermal noise from both circuits and input sampling switches, if there are any, is also amplified through the preamp gain. For multi-stage pre-amps, the noise requirement on the first stage is the most stringent and gets relaxed in later stages.

The preamp stage is usually implemented in some sort of a source-coupled pair, and its power-to-thermal-noise relationship is similar to that of the S/H circuit case where the key block is the high gain op amp, except that the preamp is usually in the open-loop configuration. Also, $1/f$ noise must be considered since it appears like a slowly varying offset of the comparator for high speed operation. Periodic offset cancellation at a rate much higher than the $1/f$ noise corner frequency, usually every clock period, can reduce this effect. The analysis for noise is omitted here since it is relatively straightforward compared to that of the amplifier in the feedback as in S/H circuits, and the reader can refer to [3].

Another major factor which affects the accuracy of the comparator is the offset voltage caused by the mismatches from process variations. For the circuit in Fig. 20, when the input signal is sampled on nodes X and Y, any mismatch between right and left half circuits will cause an offset voltage during its regenerative process. This includes charge injection mismatches from input switches, threshold and (W/L) mismatches between cross-coupled devices, and the offset voltage can be as high as 100mV easily. Due to this large offset present in this circuit, preamp stages are again required because the source coupled pair exhibits lower offset voltages. With careful layout (like common-centroid) of the input stage, the preamp stage can have the offset down to $\sim 1 - 10\text{mV}$ and about 8bit resolution can be achieved without calibration[34]. For higher resolution, however, the use

of preamp must be combined with offset nulling techniques to reduce the offset below ~1mV.

One technique of offset cancellation is to use a multi-stage configuration in which the offset voltage of the previous stage is nulled out by the subsequent stage as illustrated with a two stage comparator in Fig. 22. During the offset sampling period, the output of

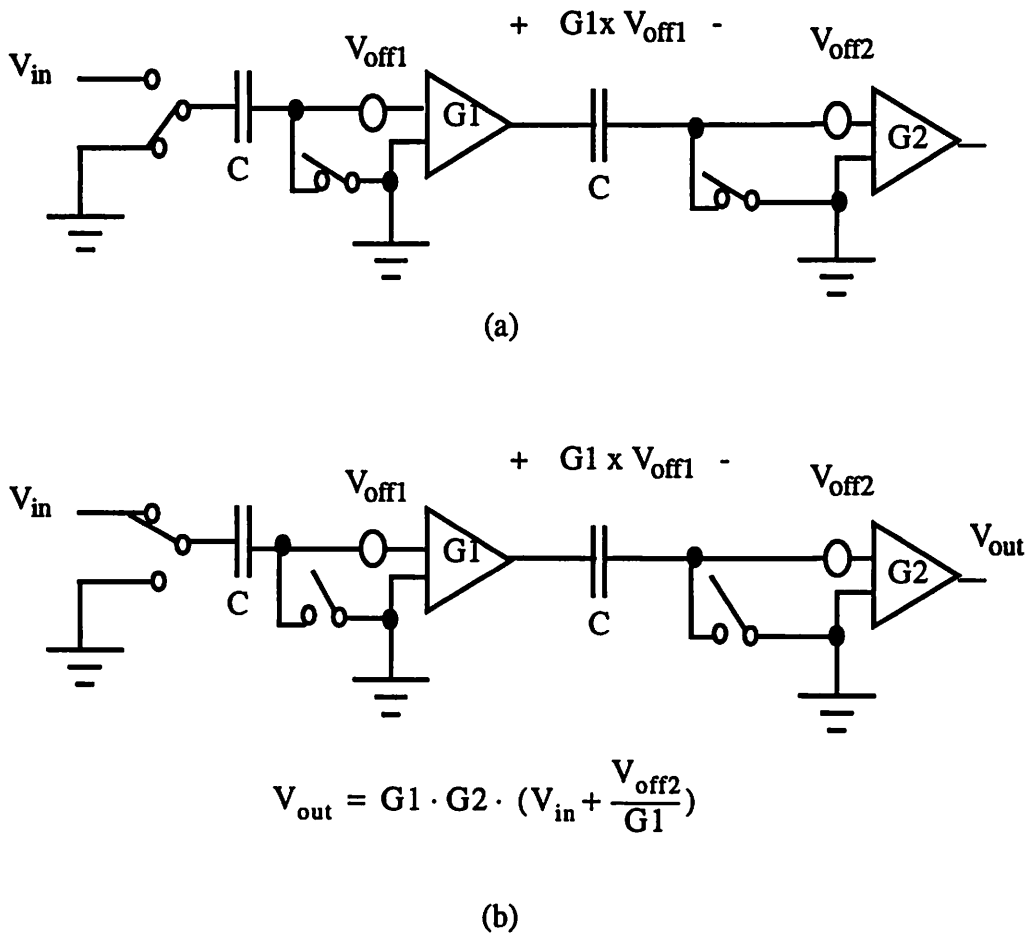


FIGURE 22. An offset cancellation technique (a)first stage offset sampling/cancellation (b)normal input comparison and the effective transfer function.

the first stage caused by its offset voltage is sampled on the sampling capacitor of the second stage. In the next clock phase, when the actual comparison is to be made, the stored voltage on the second stage sampling capacitor effectively cancels out the offset of the first amplifier, and a very accurate comparison can be made. For this cancellation

technique, notice that the gain of the first stage must be chosen relatively low so that the output voltage due to its offset does not rail out of the range (or supply).

One observation is that the offset voltage of the dynamic comparator circuit cannot be cancelled by this technique because the positive feedback amplifies even a small offset voltage to the supply rails and therefore no information on the offset voltage can be obtained at the output of the comparator. As a result, this technique requires a preamp with a DC bias current and therefore static power to reduce offset voltage.

Therefore, it can be (obviously) concluded that in typical MOS technologies, high-precision quantization function requires large power compared to the dynamic switching power of the cross-coupled latch because amplification and offset error cancellation require extra complex circuits which usually consume static power. As a result, ADC architectures which require many precision comparators for high resolution, large static power consumption must be expected.

2.4.2 Practical Implementations of Comparators

If an input signal is sampled on a capacitor before comparison, the capacitance value must be carefully chosen to reduce various non-idealities in addition to the kT/C noise. In many two-step flash type of ADC's, the comparator often has its own input sampling capacitor to eliminate the dedicated input S/H circuit, as shown in Fig. 23[20]. During the first clock phase, the left side of the sampling capacitor tracks the input voltage while the right side of the capacitor is precharged to V_m . In the next phase, the left side of the capacitor is connected to V_{ref} to compare the input voltage with V_{ref} . In this case, the minimum size of the sampling capacitor is determined by three factors: the charge injection from S1, signal attenuation, and the kT/C consideration.

As illustrated in Fig. 24, the charge injection from M_1 will cause an offset voltage, $\Delta V = \Delta Q / (C_S + C_P)$. By the same reason given in the section 2.3.3.C, requirement on the

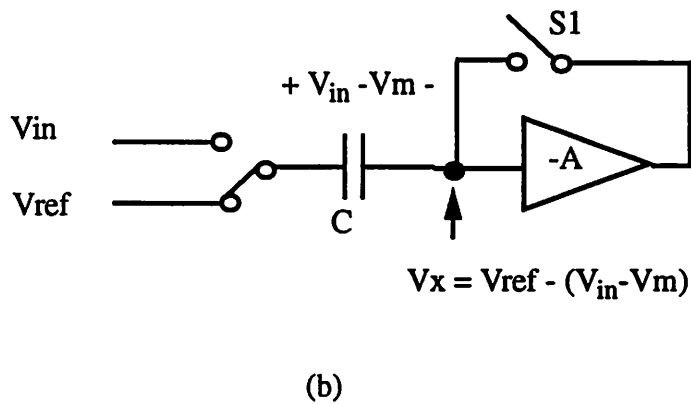
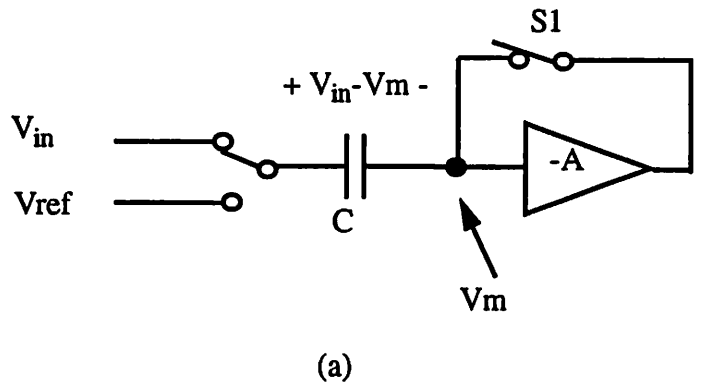


FIGURE 23. A comparator with its own sampling capacitor. (a) Input sampling (comparator threshold set at V_m), and (b) reference sampling for comparison.

input bandwidth sets the magnitude of ΔV , and the higher the sampling bandwidth is, larger the ΔV gets as a result. Reducing the gate channel length reduces this charge injection induced offset error.

Another important consideration for choosing C_S is the signal attenuation due to C_P . At the input of the amplifier, the input capacitance of the amplifier and the parasitic capacitance from the switch attenuates the input signal by $C_S/(C_S+C_P)$ and effectively reduces the amplification. Choosing a large C_S about 4 times larger than C_P gives this factor to about 80%.

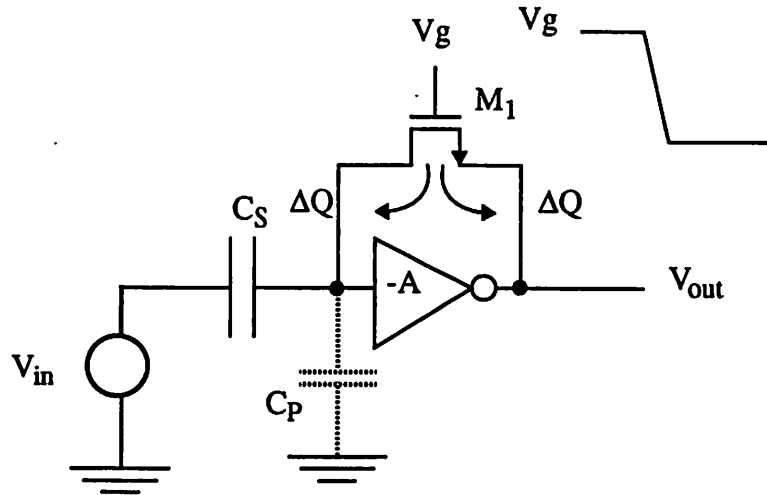


FIGURE 24. Offset error due to charge injection from M_1 .

kT/C noise must be also considered in determining the capacitor value. For instance, a 10 bit comparator with its input range of 1V must be able to resolve the signal down to 1mV, and C_S must be chosen around 100fF whose RMS value for the kT/C noise at room temperature is about $200\mu\text{V}$. If an additional one bit is to be resolved, the sampling capacitor value must be increased by 4 times to keep the kT/C noise power down relative to the reduced minimum detectable signal level for the same reason given in section 2.3.1.

Other examples on the error cancellation in comparators can be found in [20][33].

2.5 Reference Potential Generation

In all ADC's, the accuracy of the reference levels sets the limit to the achievable linearity of the A/D conversion process. For example, in the flash ADC in Fig. 2, if the input S/H circuit and the comparators are ideal, any error present in the reference resistor string will be directly translated into an A/D conversion error such as differential non-linearity error(DNL) and integral non-linearity error(INL), etc. Especially for high speed

applications where high speed switching noise may be coupled into the reference, settling time of the reference also limits its accuracy. In CMOS technologies, there are two common ways to generate reference voltages. One way is to use a resistor string (R-DAC) and the other is to use capacitor arrays (C-DAC), and these two are examined in this section.

2.5.1 Resistor String DAC(R-DAC)

The simplest way for reference generation is to use a resistor string. By having 2^N resistors connected in series between two reference voltages, the intermediate voltages can be generated according to the resistive voltage divider as shown in Fig. 25. This type of

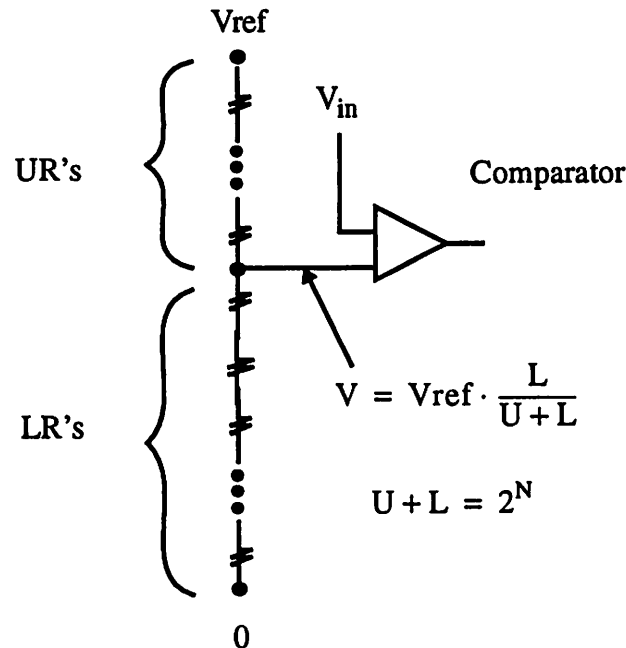


FIGURE 25. A resistor string for reference voltage generation.

resistor string is usually used in a flash or 2step-flash architecture where every reference level must be present at all times. Also, some pipeline ADC's uses R-DAC for resolution of about 8 - 9 bits[7][12].

The accuracy of the R-DAC is limited by settling and matching. For high speed

operation, various sampling capacitors can be switched to a tap point on R-DAC creating the voltage glitch. This transient has to settle out to the given accuracy within a given period of time, and the worst case settling occurs at the middle tap where the equivalent R-value is the total resistance divided by 2 plus the switch resistance as illustrated in Fig. 26.

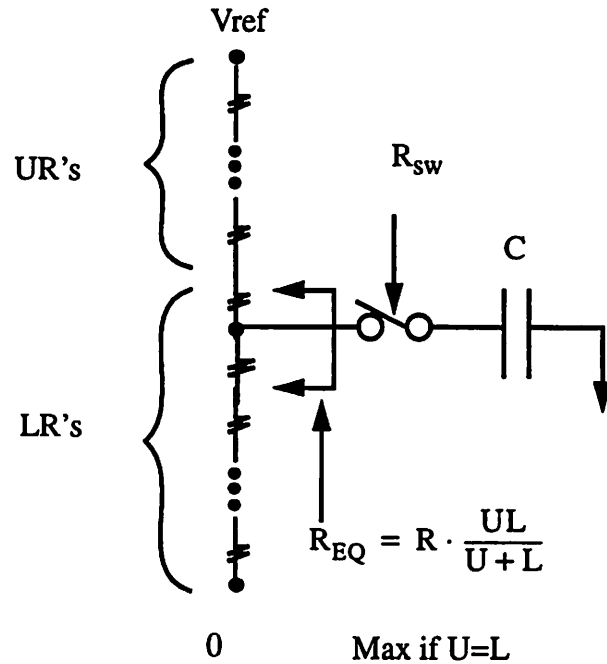


FIGURE 26. Location dependent settling of R-DAC.

This transient causes the signal dependent settling of the DAC and can translate into harmonic distortion. Therefore, the R-value must be small enough so that the worst case transient settles down within a given period of time. For a fixed value of Vref, this translates into large power dissipation in the resistor string since $Power = Vref^2/R$.

If the R-DAC settles fast and its settling error is negligible, mismatches of individual resistors due to process gradients, and random edge variations present in the process determine the overall accuracy of the generated reference voltages. A number of papers deal with this subject, and several experimental data and ADC performances also

suggest that R_DAC matching is good to about 8 - 9 bit linearity in today's technologies[28][20][21].

2.5.2 Capacitor Array DAC(C-DAC)

Reference levels can be also generated using capacitor arrays, and one way of using C-DAC can be found in a successive-approximation type of ADC where binary weighted capacitor array is used to generate the reference as shown in Fig. 27. The signal is first sampled on all capacitors, and then plates on one side are connected to the ground having the voltage at the top plate of the capacitors equal to the sampled input voltage in opposite polarity. If one of the capacitors, whose value is 4C, is connected to Vref instead of the ground, the voltage at node X is now given by:

$$V_X = -V_{in} + \frac{4C}{\left(\sum_0^{N-1} 2^i C + C\right)} V_{ref} = -V_{in} + \frac{4}{2^N} V_{ref} \quad (\text{EQ 22})$$

If both polarities of Vref are available, the reference voltage can be either added or subtracted to the signal. The above equation can be rewritten for capacitors being switched to either +Vref or -Vref

$$V_X = -V_{in} + \frac{\sum_{i=0}^{N-1} d_i \cdot C_i \cdot V_{ref}}{2^N \cdot C} \quad (\text{EQ 23})$$

where d_i is +1 for capacitors switched to Vref and -1 for capacitors switched to -Vref and C_i is $2^i \times C$. Since capacitor values are binary-weighted, reference values of N-bit resolution can be generated with N+1 capacitors.

A/D conversion in successive approximation ADC's is based on this idea. The signal is first sampled and then compared to some reference voltage to determine the

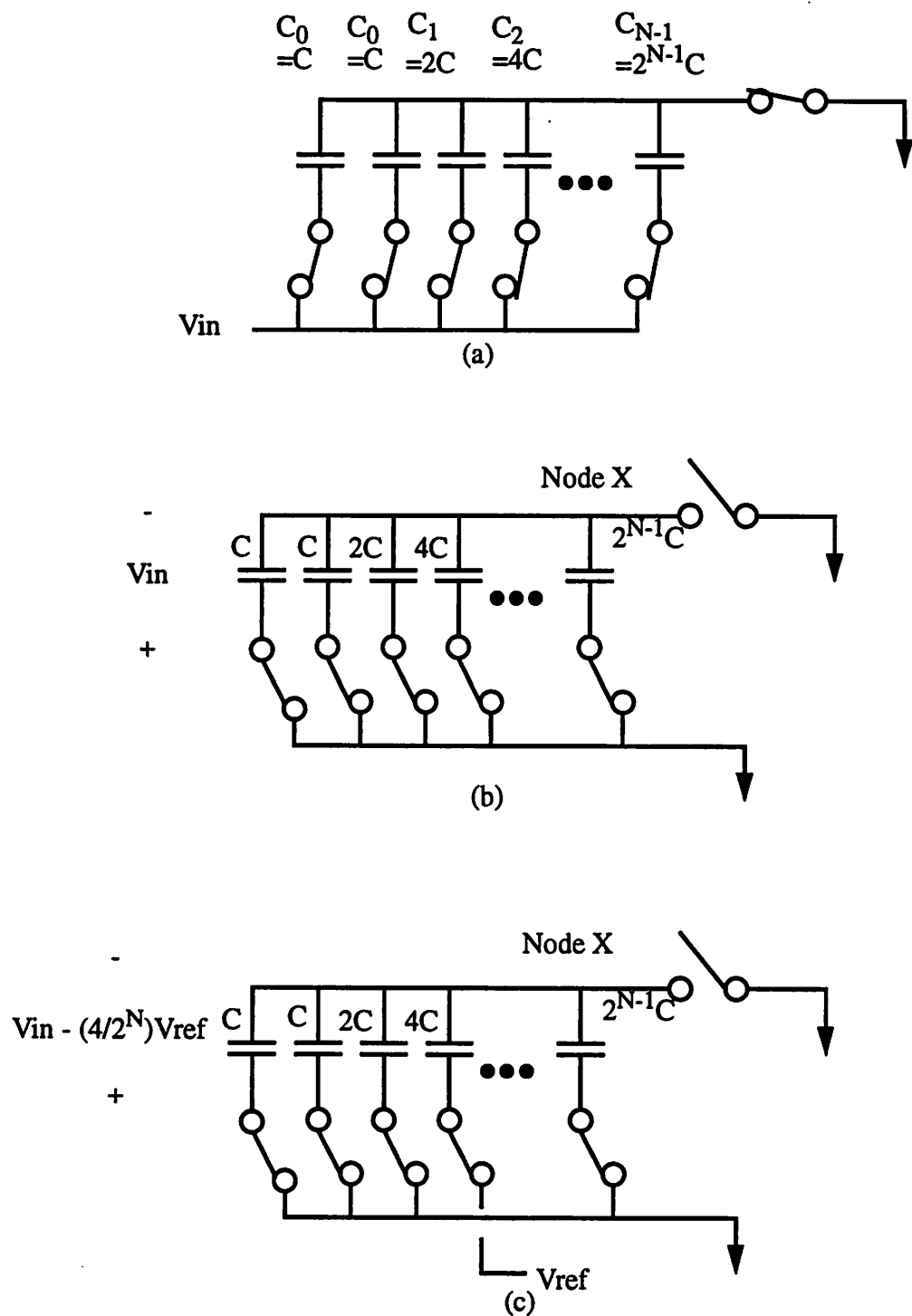


FIGURE 27. Binary weighed capacitor array for successive approximation A/D converter(a) input tracking (b) sampling and (c) charge subtraction.

MSB. Then, the quantized MSB is added or subtracted from the signal by switching the C_{N-1} to either V_{ref} or $-V_{ref}$, and the next significant bit is determined. By repeating this process N times, N bit A/D conversion is performed. This charge redistribution idea is also used for reference generation in other ADC architectures including pipeline, $\Sigma-\Delta$, etc.

In an array of binary weighted capacitors, an exact power of 2 ratio is required between capacitors, and a large size capacitor is required in order to obtain high accuracy matching. If N is not large, 2^N identical capacitors can be used instead to improve the monotonicity and linearity of the DAC.

As in the case of R-DAC, matching between capacitors determines the accuracy of the generated reference. For the same reasons as in R-DAC case, capacitors laid out identically show difference due to process gradient and random edge variations, and overall untrimmed capacitors show about 8-9 bit linearity without trimming[29]. However, these mismatch errors can be eliminated/reduced by various trimming/calibration techniques to achieve high resolution in ADC's[9][25][26][27].

Trimming requires very small size capacitors, sometimes below a few fF(10^{-15} F) in order to get high accuracy trimming. In a typical $1\mu\text{m}$ MOS technology, the minimum size capacitor is about a few fF's, and therefore, finer trimming down to 1/10th of fF's requires the use of series combination of capacitors[27]. If two identical capacitors are connected in series, then a capacitor of half the value can be realized. Therefore, more series combination means more attenuation in the capacitance value. This idea is illustrated in Fig. 28 where C_{att} is put in series to C_1 , C_2 , and C_3 to reduce their effective capacitance values in parallel to C . So, in the configuration shown in Fig. 28, C is trimmed to have the value of C plus attenuated C_1 and C_2 . More capacitance can be added to C by connecting C_3 in parallel with C_1 and C_2 . Other techniques to improve the accuracy include digital calibration in which actual capacitor values are measured[2][25][41][42] and capacitor averaging techniques with S/H circuits[13].

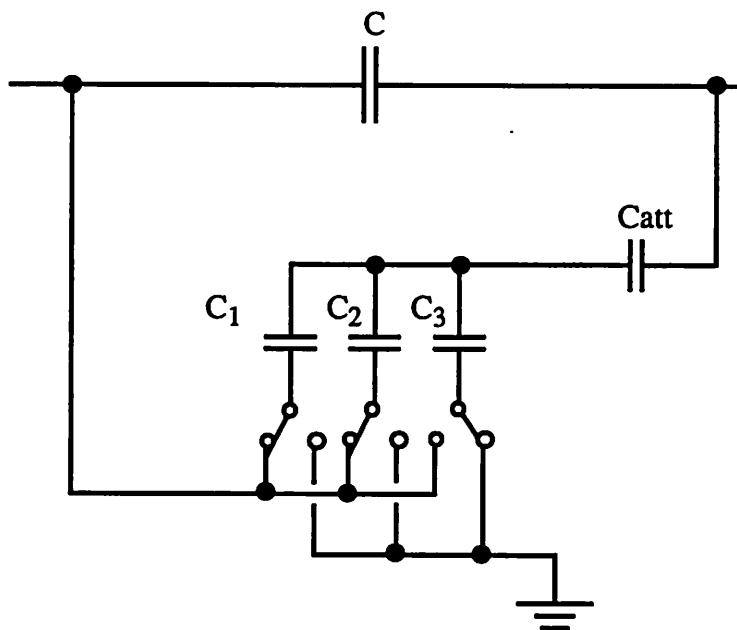


FIGURE 28. Capacitor trimming by using a small capacitor array.

In all cases with C-DAC, its basic operation is based on the charge redistribution and therefore requires no DC power dissipation. Therefore, if all reference voltages are not required to be present at all times (unlike the flash type of ADC's), C-DAC provides a low-power solution for reference generation, whose power dissipation is given by $\sim CV^2f$. Capacitance value must be chosen in this case according to the kT/C noise consideration as mentioned in section 2.3.1. Therefore, the minimum value for the total capacitance value quadruples as the resolution increases by one bit for a fixed signal amplitude and dynamic power increases by the same factor.

2.5.3 Power Comparison: R-DAC vs. C-DAC

Assuming that elements in both R and C DAC's can be calibrated out with some sort of circuit technique, theoretical minimum power dissipation in each case can be found considering just the thermal noise as the fundamental limit. To further simplify the power comparison without introducing more complicated issues such as architecture, etc., only the power dissipation per comparison (i.e., only one comparator is attached to the DAC) is

analyzed in each case. Again, the purpose of analysis here is to give a theoretical lower bound for the power dissipation and to study the general result as in the ideal S/H circuit case in section 2.3.1.

The power dissipation in a R-DAC can be found to the first order with the following assumptions. First, assume that the worst case RC time constant which occurs at the middle tap of R-string and the voltage transient is caused by switching of a sampling capacitor of some circuit such as a S/H or a comparator, whose size is determined by the kT/C noise consideration as shown in Fig. 29. Second, static power from the bias current

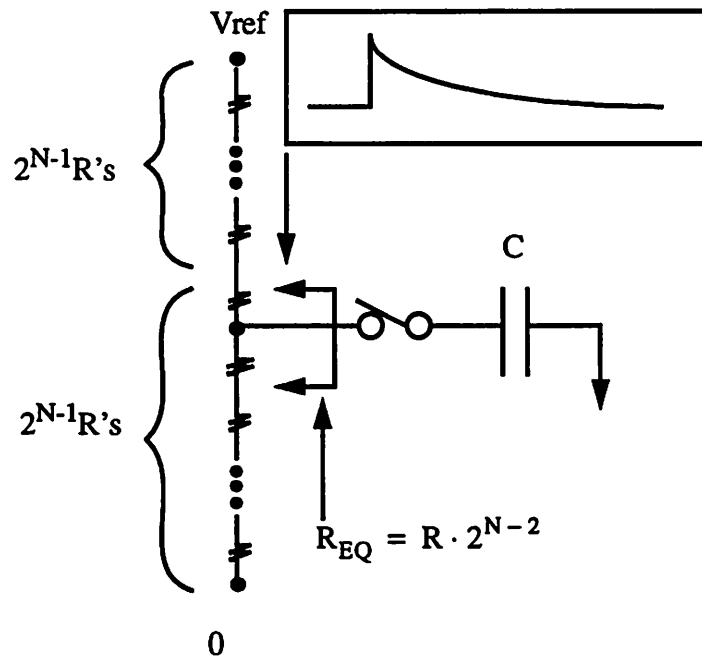


FIGURE 29. Settling of R-DAC at the middle tap.

through the R-string is assumed to be much larger than any switching power from the sampling capacitor. Third, the switch and interconnect parasitic capacitance and switch resistance are ignored for simplicity. In reality, they are important factors in determining the resistance value, and a distributed R-C model should be considered for better accuracy.

With above assumptions, the required time constant to settle to N bit accuracy in $1/f_s$ sec has to satisfy

$$t_{\text{settle}} = N \cdot \ln(2.0) \cdot \tau = N \cdot \ln(2.0) \cdot R_{\text{EQ}} \cdot C < \frac{1}{f_s}. \quad (\text{EQ 24})$$

Here, C is determined by kT/C noise consideration. So, let's assume that the noise level is some fraction of LSB, $1/\alpha$. Then,

$$\sigma = \sqrt{\frac{kT}{C}} = \frac{\text{LSB}}{\alpha} = \frac{V_{\text{REF}}}{\alpha \cdot 2^N}. \quad (\text{EQ 25})$$

Then, the minimum power dissipation can be found by:

$$\text{Power} \approx f_s \cdot \ln(2.0) \cdot \alpha^2 \cdot kT \cdot 2^{2N-2} \cdot N. \quad (\text{EQ 26})$$

For $N = 10$, $\alpha = 10$, $f_s = 10\text{MHz}$, the required minimum power is about $7.5\mu\text{W}$.

Power dissipation in C-DAC can be calculated with the following assumptions. First, the charge redistribution requires full scale (V_{REF} in this case) charging/discharging of the capacitor and therefore involves only dynamic power. Secondly, the parasitic capacitances are ignored in order to find theoretical minimum power in the DAC. Again, they are important factors in actual implementation of the DAC and should be considered as in the R-DAC case. With above assumptions, power is given by:

$$\text{Power} \approx C \cdot V_{\text{REF}}^2 \cdot f_s \quad (\text{EQ 27})$$

where C is determined by the same kT/C noise consideration. Then, (EQ. 27) becomes:

$$\text{Power} \approx f_s \cdot \alpha^2 \cdot kT \cdot 2^{2N}. \quad (\text{EQ 28})$$

This value is about $4.4\mu\text{W}$ for $N = 10$, $\alpha = 10$, $f_s = 10\text{MHz}$, and it is lower than that of R-DAC.

A couple of conclusions can be drawn from the above theoretical analysis. First, power dissipations in both cases are *independent of V_{REF}* . This is because increasing V_{REF} by 2x allows 4x decrease in required capacitor value and results in the same power dissipation as in the theoretical S/H circuit case in section 2.3.1. Also, power dissipations in both cases are linearly proportional to the operating clock frequency and temperature.

Secondly, *C-DAC dissipated less power at high resolution* and this can be seen from the ratio from (EQ. 26) and (EQ. 28):

$$\frac{\text{Power}_R}{\text{Power}_C} = \frac{\ln(2.0)}{4} \cdot N = 0.17 \cdot N. \quad (\text{EQ 29})$$

Under this set of assumptions, minimum required power in R-DAC is about 1.73 times larger at 10-bit level, and is 2.1 times larger at 12-bit level compared to that of C-DAC. This factor is independent of sampling frequency and temperature.

2.5.4 Practical Limitations

In actual implementations of ADC's, power dissipation in DAC is much larger than the theoretical minimum given above by several orders of magnitude and several practical issues come into play.

First, often more than one switch is connected to each tap of the R-DAC and each single capacitor of the C-DAC array, and the parasitic capacitance from interconnection and switches can be significant. So, for R-DAC, the resistance value must be reduced to compensate the increase in the capacitance, and for C-DAC, poly-poly capacitors, if available, should be used for small bottom plate parasitic capacitance. For this reason, careful layout techniques such as folded layout are developed to reduce the routing parasitics[21].

Also, for the R-DAC, large resistance value for low operating frequencies requires

a large area when poly resistors are used because its sheet resistance is typically low (~50 Ω /square) with 30-40% variation. For instance, 10k Ω resistor with 50 Ω /square requires 200 squares, and for matching reason non-minimum width is usually used. Because of these considerations, resistance values much smaller than the theoretically required value are often used for compact layout and therefore results in larger power dissipation. However, in C-DAC case, small capacitance much closer to the theoretically required value can be used since a smaller capacitor means smaller area and less power!

Another major factor in practical implementations is that only a fraction of the total clock period is allowed for the settling of the DAC, and again a smaller resistance value must be chosen to meet the requirement. The factors are about $\sim 1/(N+1)$ for successive approximation, $\sim 1/3$ to $1/4$ for two step flash, and $\sim 1/2$ for flash and pipeline ADC's, approximately equal to the reciprocal of the number of clock periods for A/D conversion. Also, since the reference voltage sources have finite impedances (especially if generated on chip), its RC-time constant is usually chosen several times smaller than the required value so that the combined settling time can meet the specification.

Matching is an important issue especially in C-DAC where large capacitors can be used to meet as low as 0.1% linearity requirement. Trimming can further improve the matching, but for medium resolution (~8 bit) large size capacitor and careful layout technique without trimming are much simpler.

CHAPTER 3

CMOS High Speed A/D Converter Architectures

3.1 Introduction

In the previous chapter, basic key functions are examined with special emphasis on the power dissipation associated with its implementation. In this section, several ADC architectures attractive for high speed sampling ($> 10\text{MS/s}$ in $\sim 1\mu\text{m}$ CMOS technology) are discussed in an attempt to illustrate how ADC architectures are evolved from flash to pipeline to reduce its power while increasing the performance. The order of presented ADC architectures is chosen according to author's convenience and may not reflect the actual chronological evolution of ADC architectures.

3.2 Flash Architecture

As shown in Fig. 2, the N bit A/D conversion can be performed in the flash ADC by comparing the applied input signal to the reference voltages generated from a resistor string with $\sim 2^N$ comparators. The advantage of this architecture is that only one clock cycle is required to perform the A/D conversion. However, the power consumption of this architecture increases exponentially as the resolution increases. For instance, while an 8 bit flash ADC requires 256 comparators, 10 bit requires 1024. In addition, the comparator

offset requirement becomes exponentially more stringent with the resolution; the offset of a 10bit comparator must be less than 1/4 the offset of an 8 bit comparator.

Another big disadvantage is that the input bandwidth is usually much lower than the sampling frequency without a dedicated input S/H circuit. Because the signal source has to drive many comparators implemented in parallel, any mismatch in the signal paths can cause wrong decisions as shown in Fig. 30. This error degrades the overall SNR for

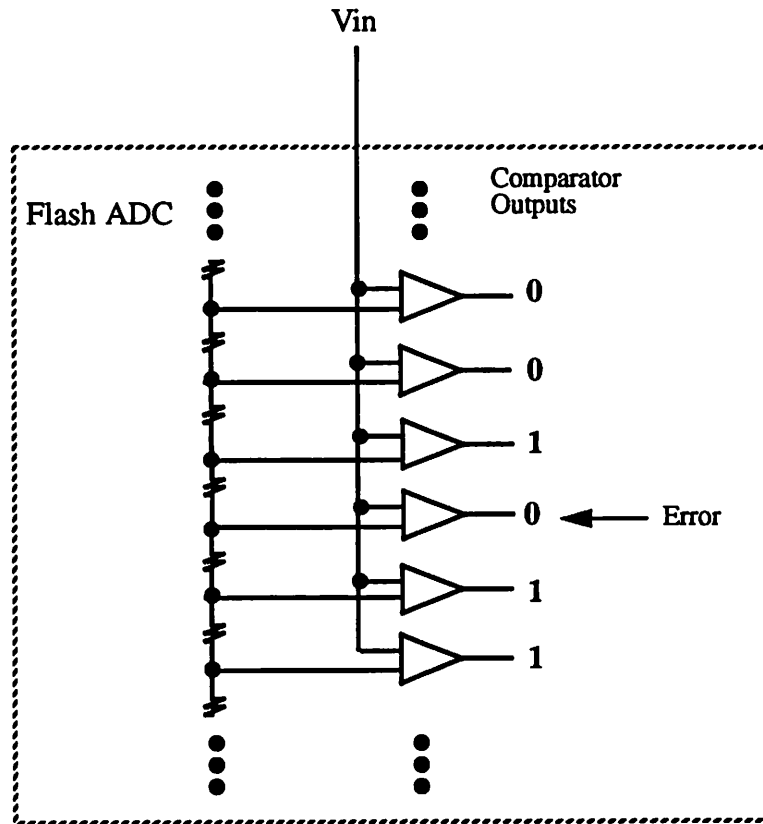


FIGURE 30. A possible error in a flash ADC due to mismatches in signal paths.

the high frequency input signal. At higher resolution, this problem becomes more severe since a large number of comparators laid out over a large area are more subject to process variation and the error budget gets tighter with smaller LSB size.

The most straightforward way to increase the input bandwidth is to use an input S/H circuit as mentioned. Since the stair-case output of the S/H circuit does not change as fast as the continuously varying input signal, the errors made by comparators can be greatly reduced (Fig. 31). The power dissipation of the S/H circuit however will be large in this case, since it has to drive a large input capacitance from many comparators.

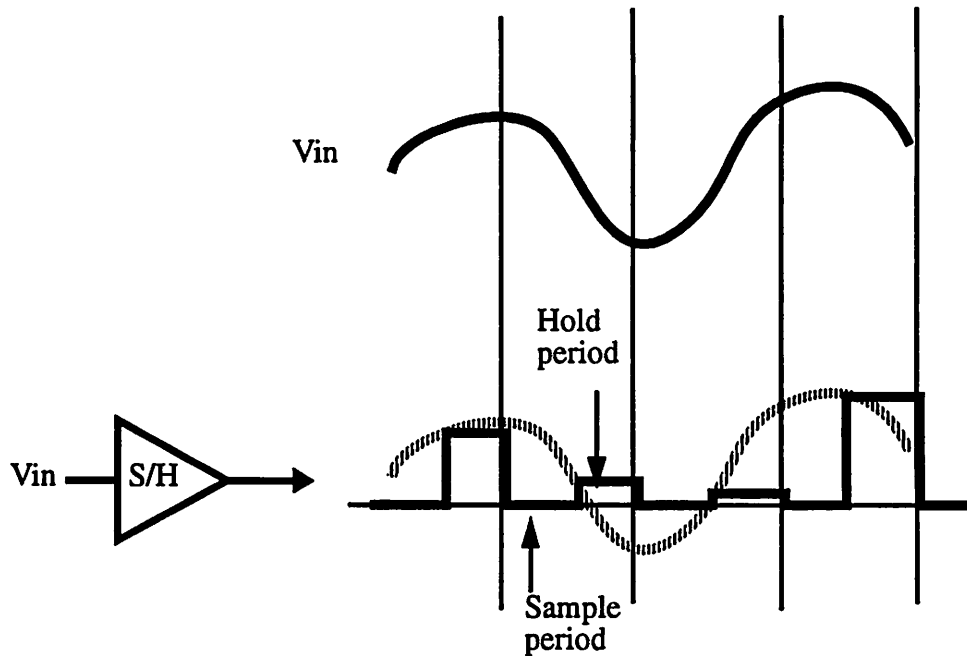


FIGURE 31. A S/H circuit to generate a stair case output.

Therefore, this architecture is only attractive for the low resolution (~6bits or less) applications with high throughput requirement, typically 100MS/sec or higher, as in the disk drive read channel[38][39].

3.3 2 Step Flash Architecture

One way to reduce the number of comparators in the flash ADC is to separate coarse and fine conversions into two time periods. For instance, if the total resolution is 10

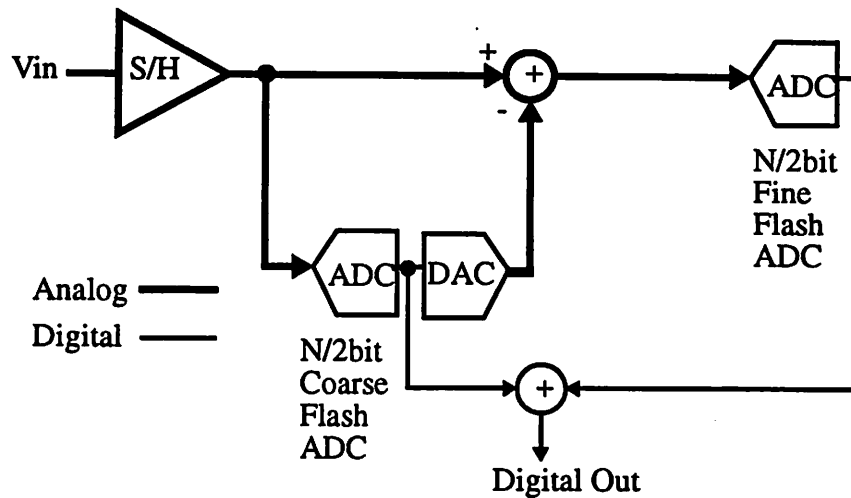


FIGURE 32. A 2 step flash architecture.

bits, the first 5 MSB's can be quantized in the first period and the next 5 MSB's in the next period. Since only 5 bits are quantized in each period, the required number of comparators is about 2^5 in each period, and the total number of comparators is $2 \times 2^5 = 64$ as opposed to 1024 in the straightforward 10bit flash ADC. In this way a substantial amount of comparator power can be saved at the expense of an extra clock cycle.

This architecture is called a "2 step flash¹", and its conceptual block diagram is shown in Fig. 32. The input signal is first sampled on the sampling capacitor of each comparator in both coarse and fine comparator banks. Then, the coarse conversion is performed by the N/2 bit coarse flash ADC. According to the outcome of the coarse conversion, the quantized signal is subtracted from the input signal, and the residual voltage is again quantized by the N/2bit fine flash ADC. By collecting bits from both

1. Or sometime it is called "subranging".

coarse and fine ADC's, corresponding digital output is generated. During this process, total three clock periods are required per sample for input sampling, coarse conversion, and fine conversion.

One practical implementation is shown in Fig. 33[18][19][21]. Comparators in

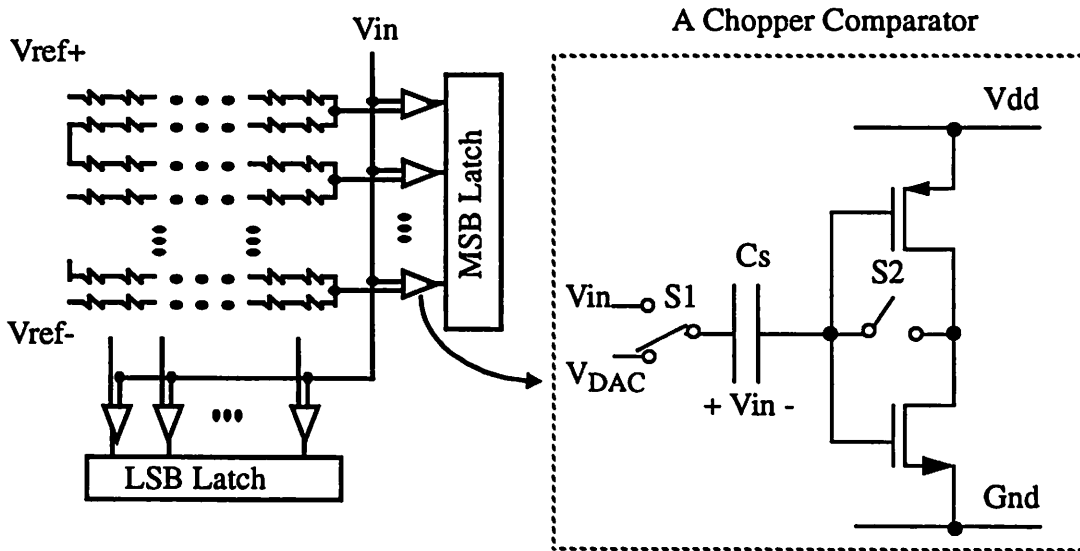


FIGURE 33. A practical implementation of 2 step flash architecture.

flash ADC sections are implemented with CMOS inverters for offset cancellation and compact layout, and reference levels are generated from a resistor string. Instead of using a dedicated S/H circuit as shown in Fig. 32, the S/H function is included within the comparator with the use of its own sampling capacitor, and the input signal is sampled on each and every comparator.

Although the number of comparators are greatly reduced from the flash architecture, path matching is still a major problem, and the input bandwidth is limited to relatively low frequency compared to the conversion rate[18][19][21]. Also, the comparator accuracy must still meet the full resolution requirement, and the offset voltage

of the comparator must be down to 1mV or less for 8-10 bit or higher resolution. As a result, multistage comparators may be required as discussed in the previous section, dissipating large power[18].

One way to relax the comparator accuracy requirement is to use digital error correction[7][8][35]. By making the fine flash ADC section capable of detecting the error due to the comparator offsets in the course ADC section, the coarse comparator requirements can be relaxed. This can be done by including extra comparators at both rails of the fine flash ADC sections as shown in Fig. 34. Therefore, if the comparators in the

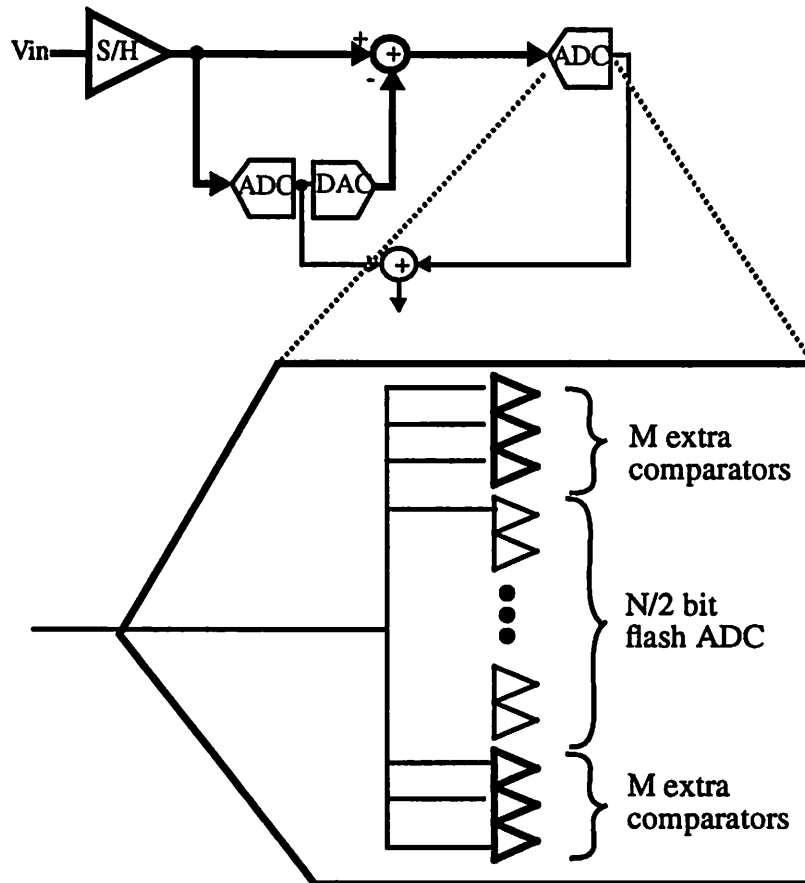


FIGURE 34. Digital error correction with 2M extra comparators in the fine flash ADC section.

coarse flash ADC makes an error and the input to the fine flash ADC goes out of the second stage nominal input range, then the extra comparators at either end detect the overflow level and correct the ADC digital output by digitally adding/subtracting the detected error. The correction range for the comparator offsets in the coarse flash ADC section is $\pm Mx$ LSB. So, the use of the digital correction can relax the comparator accuracy requirement of the coarse flash ADC section. However, the accuracy of the fine flash ADC section is still required to the full ADC resolution to make an error detection.

Fine comparators accuracy requirements can be relaxed by including an interstage gain amplifier to amplify the signal for the fine comparator bank as shown in Fig. 35

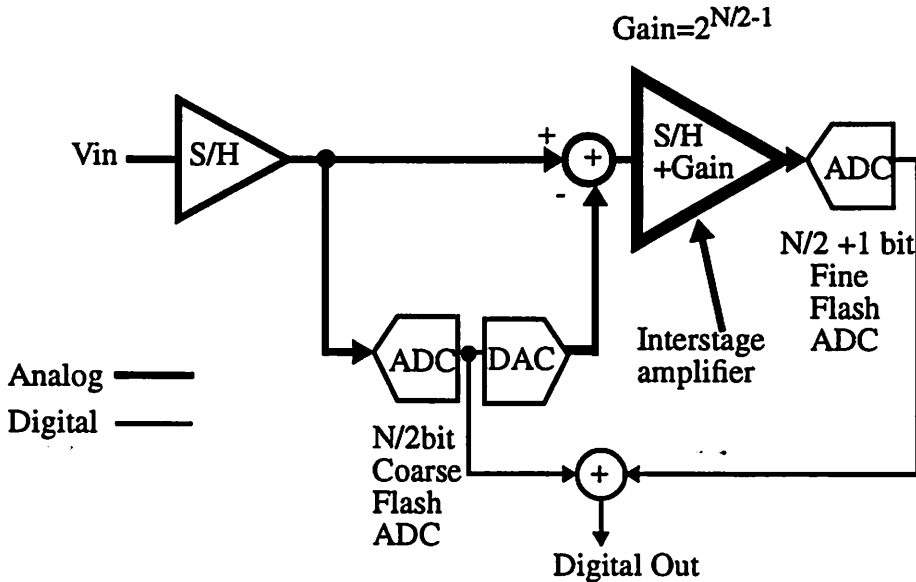


FIGURE 35. A 2 step flash ADC with an interstage amplifier.

[22][23]. By amplifying the signal, the accuracy requirements for fine conversion comparators are relaxed by the gain of the interstage amplifier, $2^{N/2-1}$. Here, the gain of $2^{N/2-1}$ is deliberately used instead of $2^{N/2}$ in order to prevent the over-range problem mentioned earlier, and the resolution of the fine flash ADC is increased by 1. Then, instead of using $N/2$ bit ADC with N bit offset requirements, an $N/2$ bit coarse flash ADC with $N/$

2 bit offset requirements ($N/2+1$ bit offset requirements for $N/2+1$ bit fine flash section) can be used[22].

Another advantage of this configuration is that the conversion steps can be pipelined due to the S/H interstage amplifier; while the first stage flash ADC works on the most recent sample, the second stage flash can concurrently work on the previous sample. As opposed to three clock periods in the previous scheme, only two clock periods are required for sampling and quantization, and in turn the throughput can be increased.

However, an op amp must be used for the S/H/Gain block and its power can be significant if fast output settling is required. While the input S/H function can be included in the comparator with the use of a sampling capacitor, the interstage amplifier must be implemented with a SC circuit which usually requires an op amp. Since it has to drive $2^{N/2+1}$ comparators in the fine flash ADC section, the op amp will dissipate large power if N is large.

3.4 Pipeline Architecture

In the 2step flash architecture with an interstage amplifier presented in the previous section, the accuracy requirements of the comparators are relaxed at the expense of the op amp power in the SC circuit. One interesting question to ask here is what happens if more interstage amplifiers are included to further relax the comparator requirements. This is the basic idea behind the pipeline architecture in relation to the power dissipation.

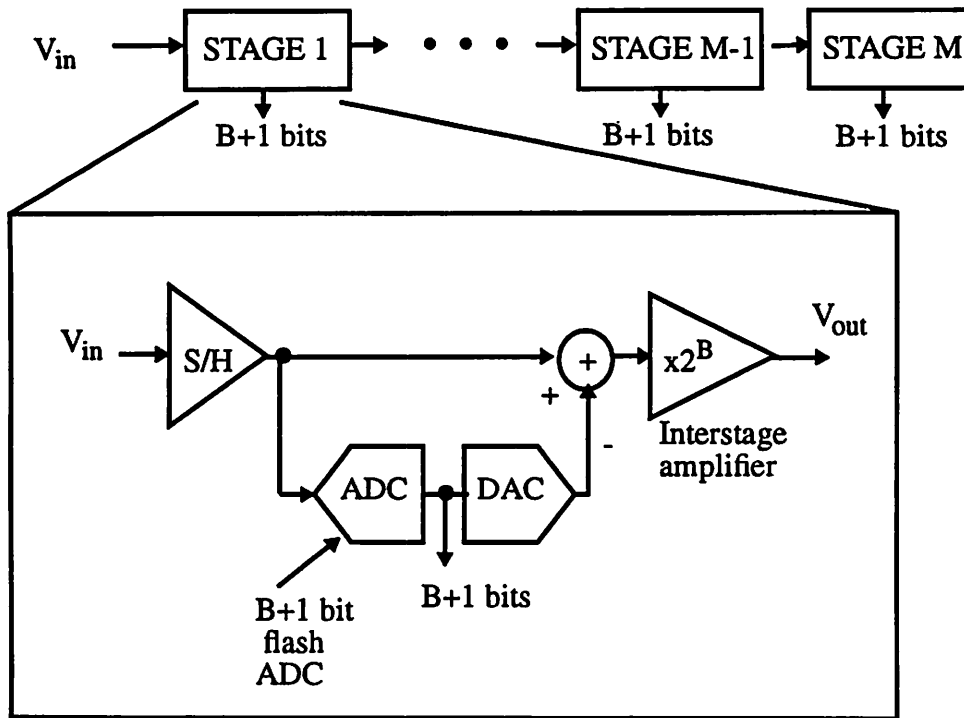


FIGURE 36. A typical pipeline architecture.

In Fig. 36, a basic schematic for a pipeline architecture is shown. Each stage samples the signal from the previous stage and it quantize to $B+1$ bits by the flash ADC section. Then, the quantized signal is subtracted and the residue is amplified through the interstage amplifier to be sampled by the subsequent stage. The same procedure is repeated in each stage down the pipeline to perform A/D conversion. The number of comparators required in this case is the number of stages times the number of comparators in each stage. From Fig. 36, it is roughly $(M \times 2^{B+1})^1$. The required number of stage is approximately the ADC resolution divided by effective per-stage resolution. Effective per-stage resolution here is denoted with B , and one extra bit is used for digital correction.

1. Number of comparators per stage can be even further reduced in actual implementation. Only general discussion is presented here from power dissipation perspective. For more detailed discussion on the pipeline architecture, see Chapter 5.

As discussed earlier, the flash ADC section in each stage has to meet only $B+1$ bit requirement due to the interstage gain and digital correction. Therefore, the lower B is, the more the comparator requirement gets relaxed.

Another observation is that both interstage amplifier and DAC requirements get relaxed down the pipeline. For instance, if the ADC resolution is 10bit and $B=1$, then while the first stage has to meet 10bit requirement, the requirement on the second stage is relaxed by 1bit. This implies potential power saving since the S/H circuits in later stages can be scaled down with smaller sampling capacitors due to relaxed accuracy requirements. The number of comparators are further reduced from 2step flash architecture at the expense of increased latency and required S/H circuits. Also, the circuit complexity grows approximately linear compared to exponential growth in flash and 2step flash architectures.

3.5 Power Comparison

Up to now, only general descriptions of three high speed ADC architectures in CMOS are reviewed in terms of the power dissipation. Detailed comparison of the power dissipation between different architectures is not easy because it involves a number of variables including resolution/sampling rate, the choice of technology, variations within the same architecture, etc. and many corresponding assumptions are needed in order to proceed with the analysis. Also, since there are many variations possible within each architecture itself, the result of the analysis based on the basic architecture may not apply to practical situations.

One way to look at the power consumption of different architectures is therefore to look at the *power factor*, meaning *what fraction of the circuits in the whole system has to meet what resolution requirement*. This is based on the assumption that the accuracy

requirement and the power consumption of a component are approximately proportional to each other, as discussed in previous sections.

For flash architecture, for example, the power factor of the architecture is 100%, since each and every comparator and DAC have to meet the full resolution requirement.

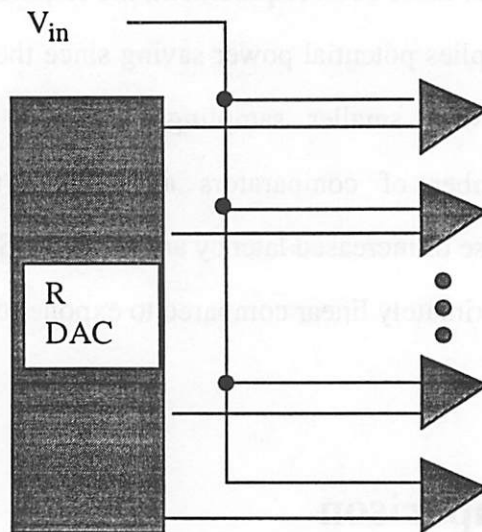


FIGURE 37. Flash architecture's power factor: the shaded region indicates full resolution requirements.

For a 2step flash architecture, the power factor of the architecture can be made less

than 100% as illustrated in Fig. 38. While both DAC and fine ADC section require full

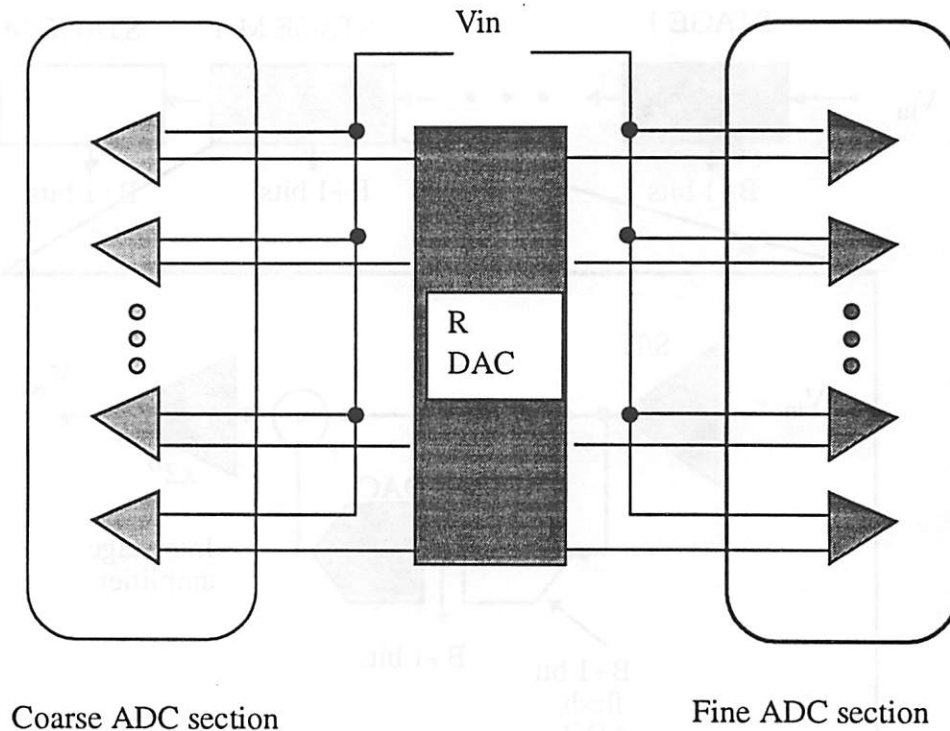


FIGURE 38. Illustration of the power efficiency of a 2 step flash architecture. Darker shaded region indicates more stringent accuracy requirement.

resolution requirement, the coarse ADC section requirement can be relaxed with the digital error correction. The level of error tolerance on the coarse ADC section depends on how much digital error correction range the fine ADC section can provide as explained in section 3.3, and the correction range varies from ± 3 LSB's in [24] to a much larger value in [22],[23] with a S/H interstage amplifier.

On the other hand, in pipeline architecture, the stage resolution decreases in later

stages as discussed earlier as illustrated in Fig. 39. While the flash section has a fixed

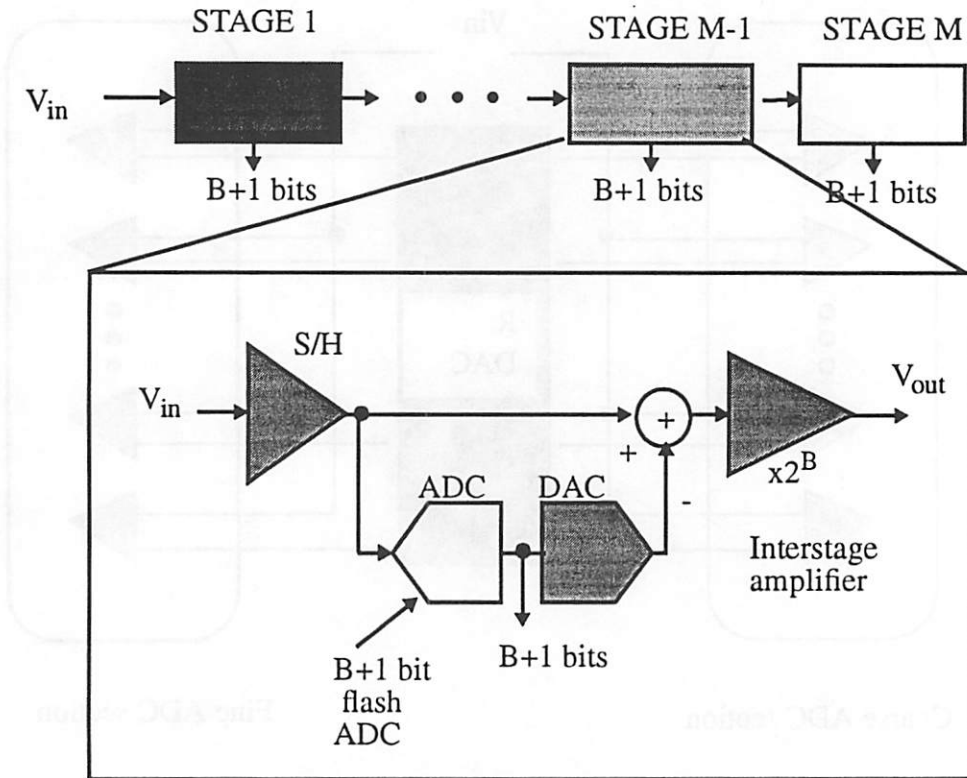


FIGURE 39. Illustration of the power factor of a pipeline architecture. Darker shaded region indicates more stringent accuracy requirement.

resolution requirement of $B+1$ bit, requirements on the interstage amplifiers and DAC section are relaxed in later stages. Among these three architectures, the power factor is therefore lowest. However, again, this power factor comparison presents a highly qualitative description of how each architecture is utilizing each key building block, not an absolute comparison.

3.6 Other ADC Architectures

Other widely used ADC architectures in CMOS technology are algorithmic, successive approximation, and Σ - Δ converters. All of these architectures are used for relatively low speed operation requiring many clock cycles to perform the A/D conversion, but their advantages are small area for the algorithmic converter and high dynamic range for successive approximation converter and Σ - Δ oversampled converter.

An algorithmic converter can be thought of as a pipeline A/D converter implemented in a recirculating manner as shown in Fig. 40. Input signal is first sampled at

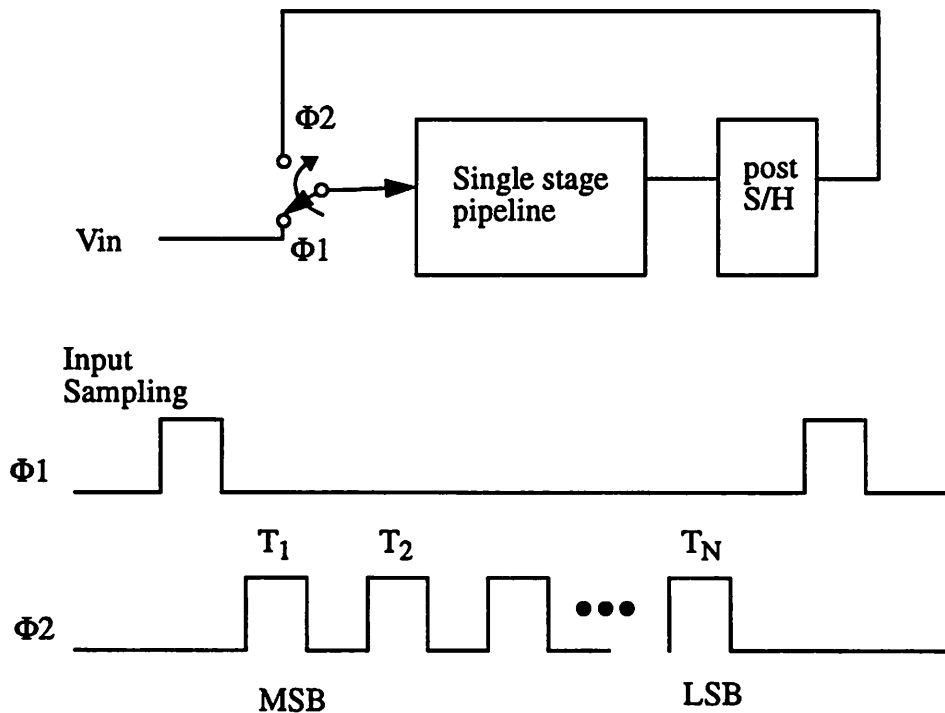


FIGURE 40. An algorithmic A/D converter.

the beginning of a clock cycle. Then, the A/D conversion is performed on the MSB. In the next clock period, the post S/H circuit samples/feeds back the residue from the output of the single pipeline stage back to its input. The pipeline stage then resolves the next

significant bit and the same procedure repeats till the last bit.

Power factor of this architecture, according to the definition presented in section 3.5, is close to 100%. Since the same SC circuit in the single pipeline stage is used repeatedly during all conversion periods from T_1 to T_N , it has to satisfy the most stringent accuracy requirements for the initial MSB conversion. However, since this architecture does not require many stages, it is good for the applications where small area is required with relatively low sampling rates of $\sim 1\text{MS/s}$. Although N N -bit parallel algorithmic converter stages can achieve the same throughput as a single N -bit pipeline A/D converter, the larger power consumption than a single N bit pipeline ADC running N times faster is expected due to its near 100% power factor even without considering power from the post S/H circuit and path matching. Examples can be found in [27][35].

A successive approximation converter shown in Fig. 41 usually dissipates a very

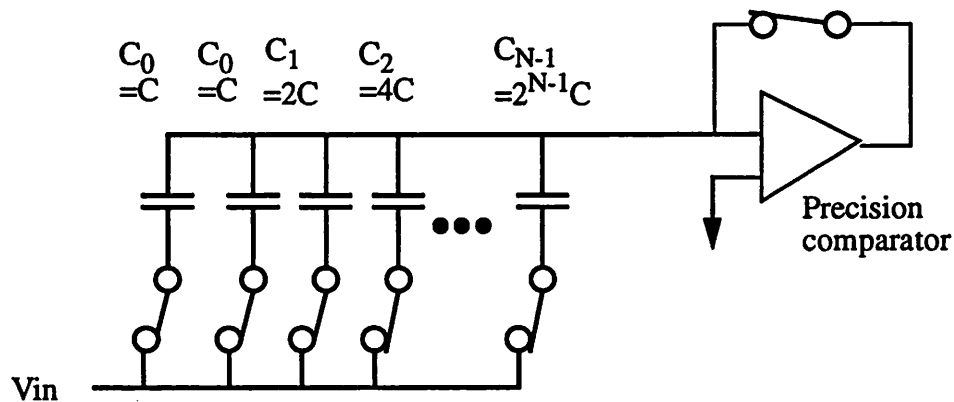


FIGURE 41. A successive approximation A/D converter core circuit.

low DC power mostly from the bias current of a single precision comparator; the rest of the power dissipation is purely dynamic from capacitive switching excluding the DAC power. The S/H function can be incorporated with the binary capacitor array requiring no

SC type of op amp based S/H circuit with DC power. Power factor of this converter is still close to 100%; the comparator offset and sensitivity has to be of full accuracy since the same comparator is used repeatedly, and the C-DAC during each clock period has to settle to a full accuracy.

The dynamic power in the C-DAC can be reduced by reducing the total capacitance by using T-network as shown in Fig. 42. By using a proper value of C_{att} ,

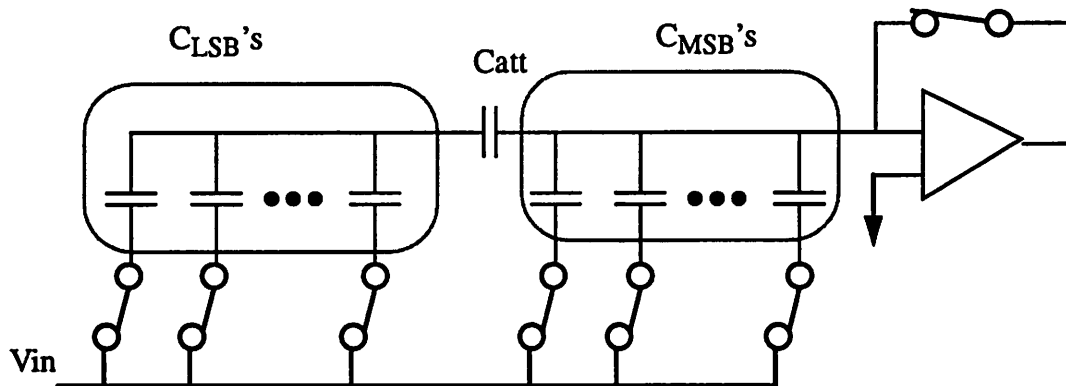
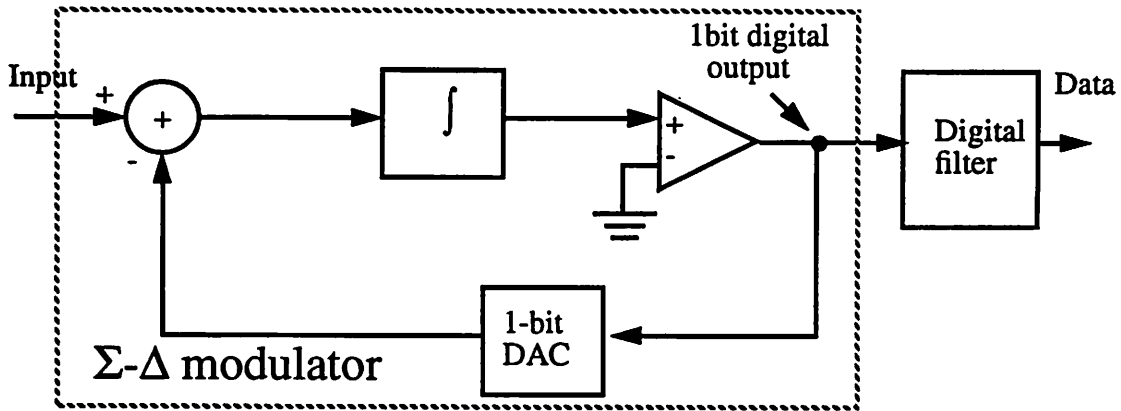


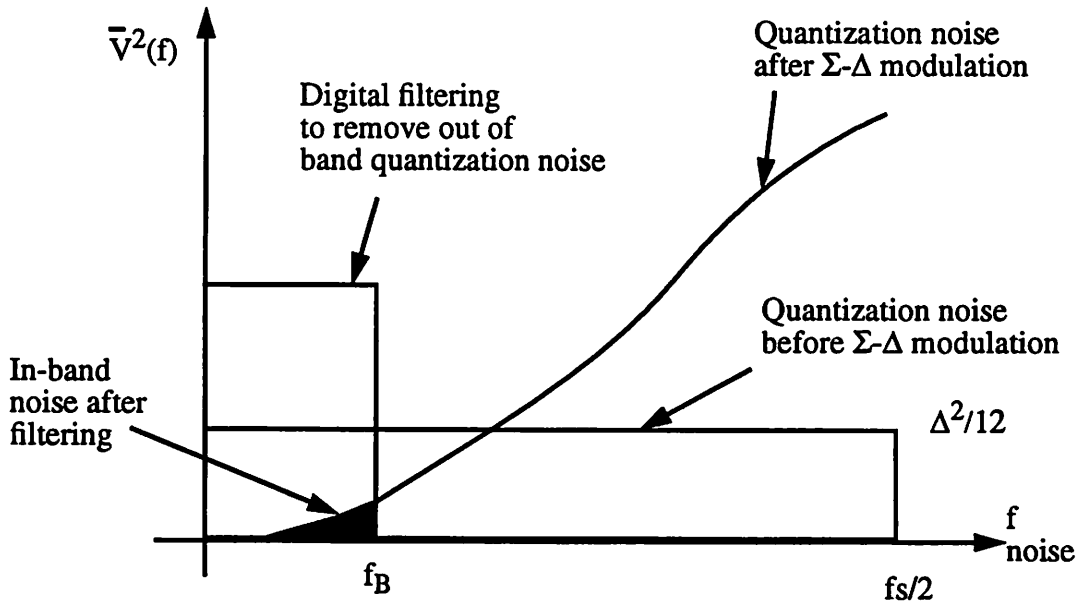
FIGURE 42. A T-network for C-DAC in a successive approximation ADC.

$C_{LSB}'s$ on the left side of C_{att} can be effectively attenuated by the series capacitance divider effect. If kT/C noise is not a concern, this technique allows to use smaller capacitors for C-DAC. However, reliably controlling C_{att} and other parasitic capacitance in as-fabricated state is almost impossible and careful calibration/trim is required[36]. According to the definition, the power factor of a T-network successive approximation ADC does not change.

A first order Σ - Δ oversampled A/D converter architecture is shown in Fig. 43. In



(a)



(b)

FIGURE 43. (a) Block diagram of a first-order Σ - Δ modulator (b) Modulator output spectrum.

this architecture, the frequency response of the quantization noise is reshaped in order to transfer most of its energy to higher frequencies by proper oversampling and negative feedback. Then, the noise is filtered out by the digital low pass filter leaving only a small portion of the quantization noise. The ratio of the sampling rate(f_s) to the signal bandwidth

(f_B) is called the oversampling ratio, and the SNR improvement with the increase in its oversampling ratio is 9dB per octave for quantization noise and 3dB per octave for thermal noise due to straightforward averaging.

The integrator in the first order Σ - Δ modulator can be implemented in a SC circuit configuration as shown in Fig. 44. Power dissipation of this SC circuit depends on the

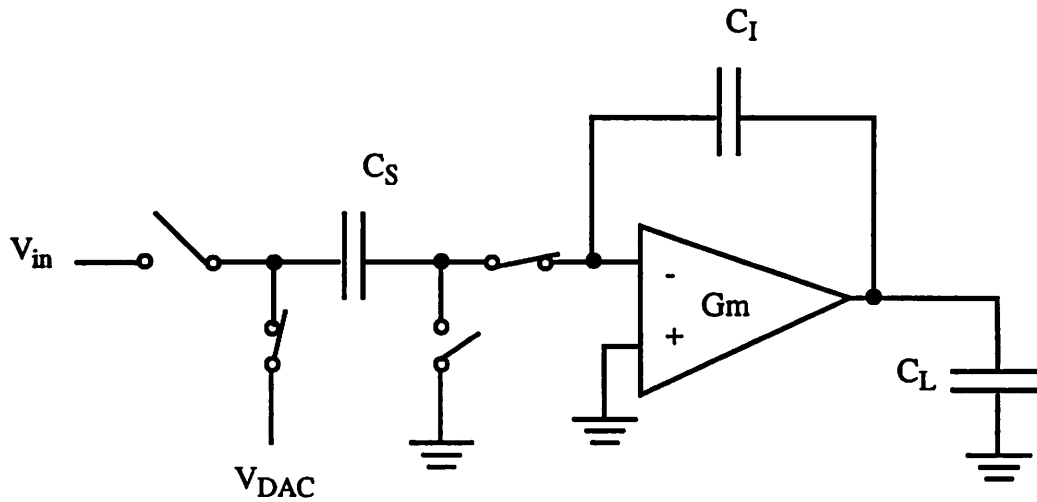


FIGURE 44. A simplified schematic for the SC implementation of the first order Σ - Δ modulator.

oversampling ratio and the size of the capacitor. If the oversampling ratio is M , the output of this integrator has to settle with a bandwidth, $(G_m \cdot f) / C_L$, where G_m is the overall transconductance of the op amp and C_L is the total output load capacitance. Since the kT/C noise on the sampling capacitor gets reduced by factor of M due to oversampling, the capacitors can be reduced by the same ratio if the same amount of kT/C noise for a Nyquist converter is to be allowed. However, since the SC circuit now has to operate M times faster, the same G_m is required as before, and as a result there's no net power saving compared to the front end S/H of the Nyquist converter if power from op amp bias current

is assumed to be approximately proportional to G_m .

Higher order loops can be used to reduce the oversampling ratio while achieving the same dynamic range. However, in addition to the increased number of modulator stages, more complex digital filter section is also required with possibly more power dissipation. Power factor is difficult to define here since the digital power dissipation must be included. However, it can be noted that the power dissipation in the Σ - Δ modulator is comparable but not much less than that of the S/H of the Nyquist converter, based on the argument given above.

In terms of practical considerations, tolerance on various component matching and relaxed requirement on the anti-aliasing filter make the Σ - Δ architecture more attractive for high resolution (above 12 bits) A/D converters than other architectures.

CHAPTER 4

A CMOS Pipeline A/D Converter Architecture

4.1 Introduction

In the previous chapter, several CMOS high speed ADC architectures are reviewed. The pipeline architecture can relax accuracy requirements on a large part of the system by reducing accuracy requirements in later stages of the pipeline yielding a high power-factor compared to other architectures. In this chapter, more detailed descriptions of the pipeline architecture are presented.

4.2 Key Building Blocks

A block diagram of a typical pipeline A/D converter is shown in Fig. 45. It consists of a cascade of N identical stages in which each stage performs a coarse quantization, a D/A function on the quantization result, subtraction, and amplification of the remainder. A S/H function in each stage allows all stages to operate concurrently, giving a throughput of one output sample per clock cycle. Fig. 45 illustrates the particular configuration of interest here in which the D/A, subtraction, amplification, and S/H functions are performed by a switched capacitor(SC) circuit, with a resolution of $B+1$ bits per stage and an interstage gain of 2^B . The D/A function is performed by a set of capacitors with a

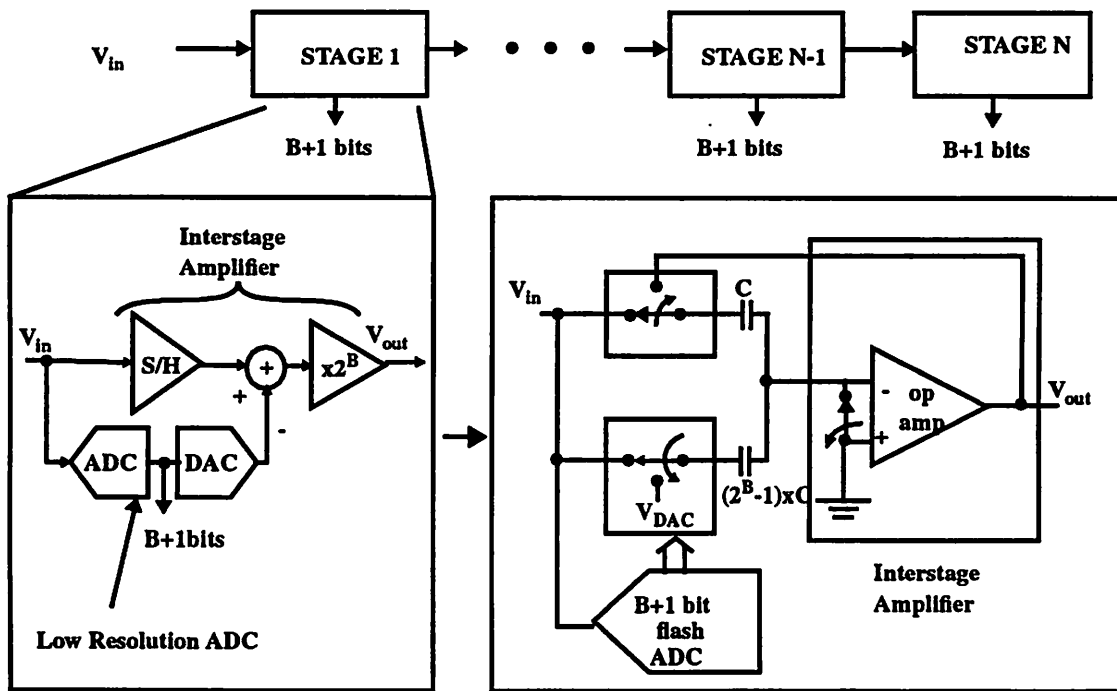


FIGURE 45. A typical pipeline architecture. Each pipeline stage can be implemented with a SC circuit and a low resolution flash ADC.

reference voltage source V_{DAC} . When the input signal is applied, each stage samples and quantizes the signal to its per-stage resolution of B+1 bits, subtracts the quantized analog voltage from the signal by connecting the bottom plate of capacitors $C_S (= (2^B - 1) \times C)$ to V_{DAC} , and passes the residue to the next stage with amplification for finer conversion. Then, B+1 bits from all stages are collected and produce a full resolution digital representation of the applied analog input.

The SC topology in Fig. 9 is chosen to implement the interstage amplifier for this particular pipeline example that will be described here. The feedback factor of the SC circuit is given by:

$$f = \frac{C_i}{2^B \cdot C_i + C_{opamp}}, \quad (\text{EQ 30})$$

and the total load capacitance at the output is:

$$C_{LT} = \frac{C_i [(2^B - 1) C_i + C_{opamp}]}{2^B C_i + C_{opamp}} + 2^B C_{i+1} + (2^{B+1} - 2) C_{comp} , \quad (\text{EQ 31})$$

where C_i is a unit capacitor of current stage i , C_{opamp} is the input capacitance of the op amp, C_{comp} is the comparator input capacitance in the flash A/D section. The first term in (EQ. 31) is the loading from the series capacitance of the current stage, and the second term is the sampling capacitors from the next stage. The last term is the total input capacitance of the flash ADC. Interconnection/switch parasitic capacitances are ignored for simplicity. One key observation here is that decreasing the per-stage resolution, B , is important if the speed of the SC circuit is to be maximized since small B allows the configuration with large feedback factor and low load capacitance. Table 4 shows values for the feedback factor and the load capacitance to illustrate the difference. Notice how

TABLE 4. Feedback factors and load capacitance for different B.

B	f	C_{LT}
1	$\frac{C_i}{2 \cdot C_i + C_{opamp}}$	$\frac{C_i [2C_i + C_{opamp}]}{2C_i + C_{opamp}} + 2C_{i+1} + 2C_{comp}$
2	$\frac{C_i}{4 \cdot C_i + C_{opamp}}$	$\frac{C_i [4C_i + C_{opamp}]}{4C_i + C_{opamp}} + 4C_{i+1} + 6C_{comp}$
3	$\frac{C_i}{8 \cdot C_i + C_{opamp}}$	$\frac{C_i [8C_i + C_{opamp}]}{8C_i + C_{opamp}} + 8C_{i+1} + 14C_{comp}$

quickly the feedback factor decreases and the load capacitance increases as the per-stage resolution increases. Therefore, low per-stage resolution is desirable if maximizing the speed of the SC circuit is the key constraint since its bandwidth is proportional to the feedback factor and inversely proportional to $C_{LT}(BW \sim (G_m \cdot f) / C_{LT})$.

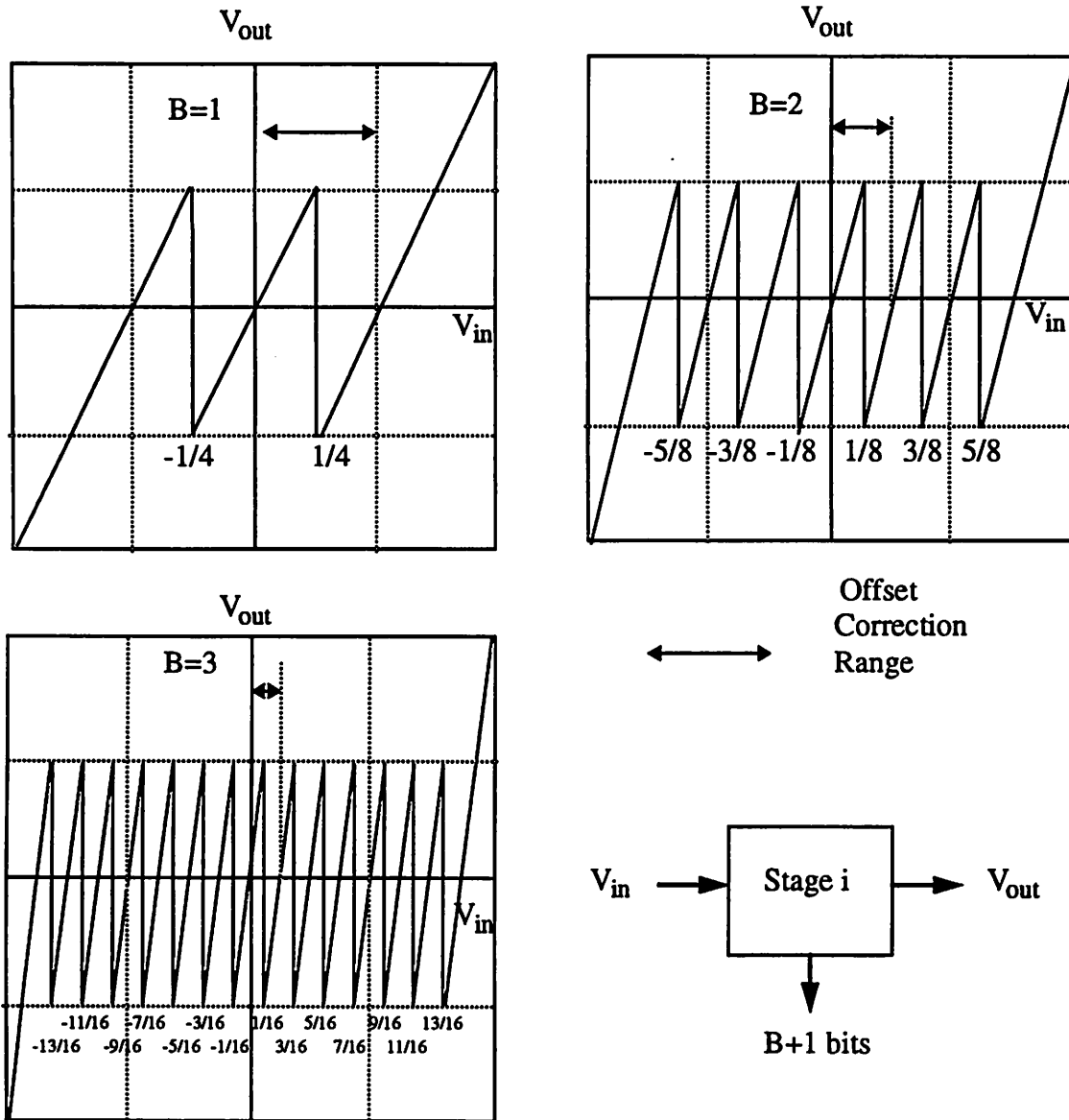


FIGURE 46. Saw-tooth input/output transfer characteristics for different per-stage resolutions($B=1,2,3$)

The flash ADC section is composed of $(2^{B+1}-2)^1$ comparators, and the digital error correction² can tolerate the offset voltage of each comparator up to $(\pm V_{ref}/2^{B+1})$. This is illustrated in Fig. 46 where the arrow indicates the offset correction range by the digital error correction for different per-stage resolutions. It is obvious that choosing the low per-

1. Top and bottom comparators are unnecessary. See Appendix 3.
2. Assuming zero interstage amplifier offset. See Appendix 3.

stage resolution allows less accuracy requirement for the comparator design in the flash ADC section. If B is small so that 100's of mV of offset voltage can be tolerated by digital correction, then even dynamic comparators can be used for the flash ADC section to eliminate its DC power[5].

DAC/subtraction functions can be implemented by using either a resistor string or multiple capacitors. Fig. 47 shows the DAC/subtraction functions implemented with a

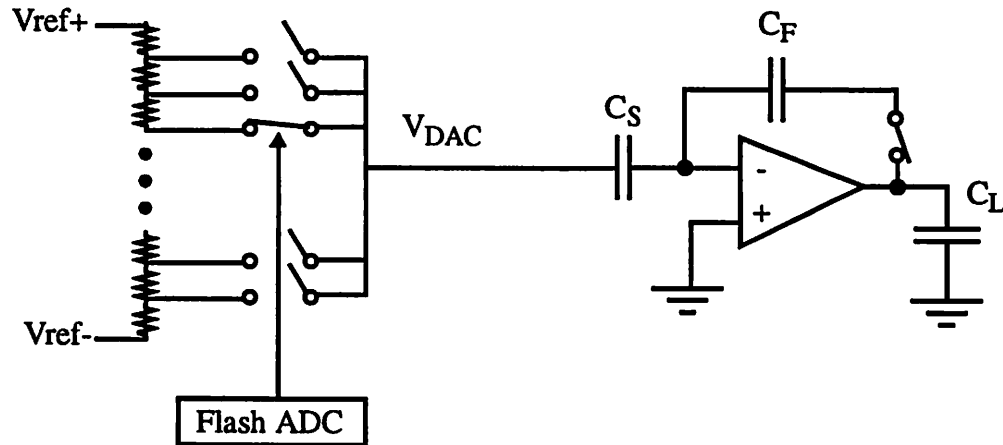


FIGURE 47. DAC/subtraction functions implemented with a resistor string.

resistor string. The transfer function of the SC circuit changes from $V_{out} = (1 + C_S/C_F)V_{in}$ in Fig. 9 to:

$$V_{out} = V_{in} + \frac{C_S}{C_F} \cdot (V_{in} - V_{DAC}) \quad (\text{EQ 32})$$

The second term in the equation is modified because among the signal charge sampled on C_S ($Q_{signal} = C_S \times V_{in}$) the charge corresponding the reference voltage ($C_S \times V_{DAC}$) does not transfer to C_F and effectively the subtraction function is performed as shown in the

equation.

The accuracy of this R-DAC (resistor string DAC) is determined by two factors, initial matching accuracy and fast R-C settling. In typical CMOS technologies, the R-DAC is implemented with poly resistors, and process variations limit the overall accuracy to about 8 bits as discussed in section 2.5.1. Trimming can be performed to improve the resistor matching, but often the process is very complicated. The second factor which affects the accuracy is the V_{DAC} settling. The R-DAC with the switch resistance forms a R-C network with sampling capacitors, and the settling of the V_{DAC} voltage according to this time constant limits the settling of the SC S/H/Gain function if it is slower than the SC circuit bandwidth. For this reason, low resistance resistors are used to set the RC-settling time much smaller than the SC circuit settling time and as a result large power is consumed.

Another way for DAC/subtraction is to use multiple capacitors. By sampling the input signal on the capacitor and redistributing the signal charge on the capacitor as mentioned in section 2.5.2, DAC/subtraction function can be performed. This is illustrated

in Fig. 48. In the charge redistribution process, the accuracy of this C-DAC (capacitive

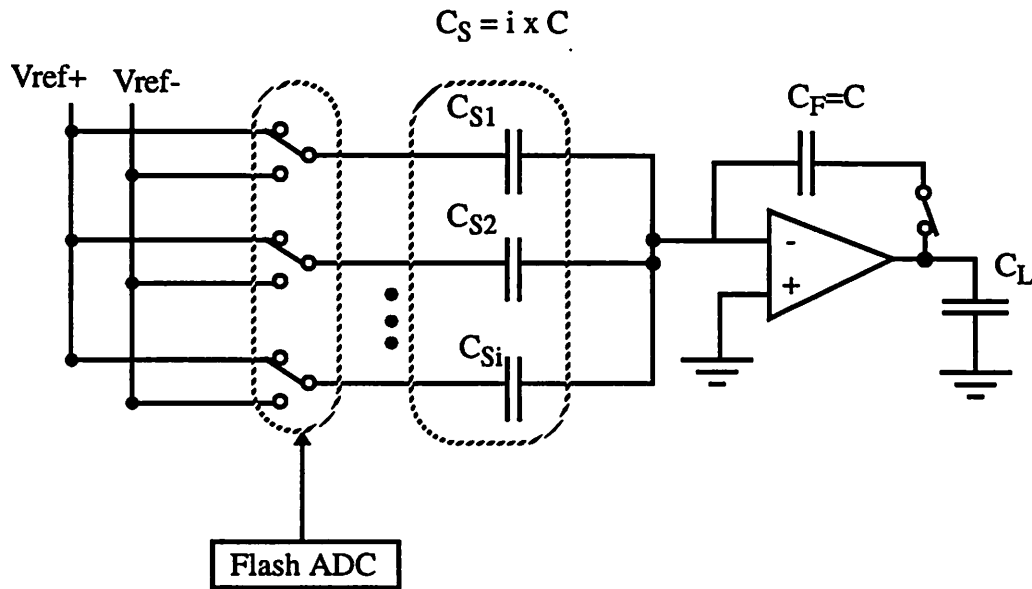


FIGURE 48. DAC/subtraction function using C-DAC.

DAC) depends on the capacitor matching. Typically, accuracy up to 8-9 bits can be obtained without calibration and without excessive area requirements[29], and matching to higher resolution typically requires trimming as in the R-DAC case. Without trimming or calibration, therefore, initial matching accuracies of both R-DAC and C-DAC are about 8-9bits maximum in typical current technologies[29]. However, the C-DAC requires less power than the R-DAC to settle faster since the whole process is dynamic with potentially less series resistance (only through a switch resistance to the reference voltage source) for smaller R-C time constant as described in section 2.5.3.

4.3 Stage Accuracy Requirements

In the pipeline architecture, the accuracy requirement on each stage is different

because the stage resolution decreases down the pipeline by the per-stage resolution. For instance, for a 10 bit ADC with the effective per-stage resolution of 1bit/stage, the first stage has to meet 10 bit accuracy requirement. However, the next stage has to meet only 9 bit accuracy requirement. Lower stage resolution means that design constraints are more relaxed, and in terms of pipeline stage design this translates into relaxed requirements on interstage amplifier gain accuracy, DAC accuracy, and thermal noise.

First, for a given total resolution of N bits and the per-stage resolution of B, the combined gain error in the interstage amplification should be less than $(1/2^{N-B})$ of the full scale input range where N-B is the resolution of the next stage[10]. Non-idealities of the interstage amplifiers include finite op amp gain, finite settling time, and capacitor matching as can be seen from

$$V_{out} = G \cdot (V_{in} - V_{DAC}), \quad (\text{EQ 33})$$

$$\text{and} \quad G = \left(1 + \frac{C_S}{C_F}\right) \cdot \left(1 - e^{-\frac{t}{\tau}}\right) \cdot \frac{1}{1 + \frac{1}{A \cdot f}} \quad (\text{EQ 34})$$

where $(1+C_S/C_F)$ is the voltage gain term, and the exponential term in the second bracket is for the settling time for the single pole op amp(τ is the time constant for the SC gain block configuration), and the third term is related to the finite op amp gain(A). In this case, the gain error term can be found from

$$\frac{\Delta G}{G} = \frac{G_{ideal} - G_{actual}}{G} = \frac{1}{2^B} \left(2^B - \left(1 + \frac{C_S}{C_F}\right) \cdot \left(1 - e^{-\frac{t}{\tau}}\right) \cdot \left(\frac{1}{1 + \frac{1}{A \cdot f}}\right) \right) \quad (\text{EQ 35})$$

$$\left| \frac{\Delta G}{G} \right| \approx \left| \left(\frac{(2^B - 1) \cdot \Delta C_F}{2^B \cdot C} - \sum_i \frac{\Delta C_{S,i}}{2^B \cdot C} + e^{-\frac{t}{\tau}} + \frac{1}{A \cdot f} \right) \right| \quad (\text{EQ 36})$$

where $C = \left(\sum_1^{2^B-1} C_{S,i} + C_F \right) / 2^B$, $\Delta C_F = C_F - C$, and $\Delta C_{S,i} = C_{S,i} - C$. In order to achieve N bit linearity, this normalized *combined* gain error term, $\Delta G/G$ should be less than $1/2^{N-B}$ (resolution of the next stage) to prevent any missing codes[10].

$$|\Delta G/G| < \frac{1}{2^{N-B}} \quad (\text{EQ 37})$$

From above, requirements on op amp DC gain and bandwidth, and capacitor matching can be determined. For op amp gain, $1/A_f$ should be $\ll 1/2^{N-B}$. Therefore, substituting the feedback factor term from (EQ. 30), the gain is given by

$$A > (2^{N-B}) \cdot \frac{(2^B \cdot C_i + C_{\text{opamp}})}{C_i} \quad (\text{EQ 38})$$

If C_{opamp} is ignored, above expression becomes 2^N . To tolerate any process variations in reality, the op amp gain must be at least two or three times larger than this value. For the same reason, the op amp output has to settle to $1/2^N$, and for the single pole system, the required number of time constants can be found from

$$N \cdot \ln(2.0) \cdot \tau < T_{\text{settle}} \quad (\text{EQ 39})$$

where T_{settle} is the allowed settling time, usually half the clock period($1/(2fs)$). For instance, for 10 and 12 bit settling, 6.9 and 8.3 τ 's are required respectively.

The matching requirement on the capacitors is determined by the allowable gain error and the required DAC accuracy if a C-DAC is used. (For the R-DAC, linearity of the resistor string must be also considered to analyze the DAC accuracy. Only C-DAC case will be analyzed here to derive the capacitor matching requirement.) So, for the C-DAC, if an ideal op amp is assumed, the transfer function becomes

$$V_{out} = 1 + \frac{\sum_1^{2^B-1} C_{s,i}}{C_F} \cdot V_{in} + \frac{\sum_1^{2^B-1} d_i \cdot C_{s,i}}{C_F} \cdot V_{ref}, \quad (\text{EQ 40})$$

$$V_{out} = G \cdot \left(1 + \frac{\Delta G}{G}\right) \cdot \left(V_{in} + \frac{\sum_1^{2^B-1} d_i \cdot C_{s,i}}{G \cdot C_F} \cdot V_{ref} \right), \quad (\text{EQ 41})$$

$$\text{and } V_{out} = G \cdot \left(1 + \frac{\Delta G}{G}\right) \cdot \left(V_{in} + \frac{1}{G} \cdot \left(\sum_1^{2^B-1} d_i \right) \cdot V_{ref} + \Delta V_{ref} \right) \quad (\text{EQ 42})$$

where d_i is either 0 or ± 1 , determined by the outcome of the flash ADC section, G is 2^B , $\left| \frac{\Delta G}{G} \right| \approx \left| \left(\frac{(2^B - 1) \cdot \Delta C_F}{2^B \cdot C} - \sum_i^{2^B-1} \frac{\Delta C_{S,i}}{2^B \cdot C} \right) \right|$, and $\Delta V_{ref} \approx \frac{1}{G} \cdot \left(\sum_1^{2^B-1} d_i \cdot \frac{\Delta C_{S,i}}{C} \right) \cdot V_{ref}$. For an N bit ADC, this means that ΔV_{ref} should be less than a LSB of the converter or $V_{ref}/2^N$ assuming other parameters are ideal. This sets the requirement on DAC of each stage, and $\Delta C/C$ of each capacitor must be less than $1/2^N$ to ensure that ΔV_{ref} is always less than a LSB.

In addition to the above deterministic error components, the other error source is the random component whose dominant source is thermal noise as discussed in previous sections. Assuming that thermal noise is additive Gaussian, the noise at the op amp output appears as being superimposed on the signal. The power of the thermal noise is then described by its variance, and the variance should be much less than the LSB in order to maintain sufficiently high SNR as discussed in section 2.3.2. So, for the pipeline stage with N bit accuracy requirement, the total input-referred noise should be much less than the LSB at N bit level.

$$\sigma_{total,i} \ll \frac{V_{FS}}{2^N} \quad (\text{EQ 43})$$

The total input referred noise at the input of stage i can be found by summing all the noise components from subsequent stages and is given by

$$\sigma_{\text{total},i}^2 = \sigma_i^2 + \left(\frac{1}{2^B}\right)^2 \sigma_{i-1}^2 + \left(\frac{1}{2^B}\right)^4 \sigma_{i-2}^2 + \dots \quad (\text{EQ 44})$$

where σ_i^2 is the variance of thermal noise from stage i and (2^B) is the interstage gain. Notice that the dominant source of noise is from the stage i and the noise contribution from subsequent stages is reduced by the interstage gain. Therefore, the noise requirement is most stringent on the first stage and gets relaxed down the pipeline.

The requirements on flash ADC sections (comparator offsets) remain the same throughout the pipeline stage since it is only related to the per-stage resolution, not the stage resolution. In Table 5, the stage requirements on building blocks are summarized

TABLE 5. Stage requirements for the i th pipeline stage with its stage resolution of N bit and per-stage resolution of B bits

Parameters	Minimum requirement
Op amp DC gain	$A > (2^{N-B})/f > 2^N$
Settling time (assuming single pole response)	$\tau < 1/(2 f_s N \ln(2.0))$
DAC accuracy	$ \Delta C/C < 1/2^N$
Noise	$\sigma_{\text{total},i} \ll \text{LSB}(=V_{\text{FS}}/2^N)$

4.4 Per-Stage Resolution & Power

From section 4.2, it is found that the power dissipation in flash ADC section and reference (C-DAC) can be dynamic leaving the major portion of the power dissipation from the op amp power from the SC S/H/Gain circuit. It is also pointed out that the bandwidth of the interstage amplifier can be maximized if lower per-stage resolution, which allows a SC configuration of low closed loop gain, is used. However, lower per-stage resolution requires more pipeline stages, and it is important to determine the optimum per-stage resolution to minimize the overall power consumption for given resolution and speed requirements.

In this section, the analysis for power dissipation in pipeline stages is carried out for different per-stage resolutions in order to examine the trade off between power and per-stage resolution. The analysis involves a number of variables, and the following key assumptions are made for simplification.

1. It is assumed that the power consumption is dominated by the static power from op amp bias currents for the SC interstage amplifiers. The comparators in the flash ADC sections are assumed to be non-critical components if per-stage resolution is 1-2 bits and therefore, can be implemented with some sort of dynamic circuits consuming no DC power.
2. It is assumed that a C-DAC is used for lower power consumption than the R-DAC case and the matching requirement can be met by various trimming/calibration methods. Again, the power in a C-DAC is dynamic and small, and therefore ignored.

3. Single transistor op amps are used for the analysis and the same assumptions on op amps are made as in section 2.3.2, except that the full scale input voltage range is assume to be 2V. For a transistor, $k_p=60\mu\text{A}/\text{V}^2$ and ρ (current density for the input device) $=1\mu\text{A}/\mu\text{m}$ are assumed.
4. Minimum size capacitors are assumed to be 50fF which corresponds to $10\mu\text{m} \times 10\mu\text{m}$ if $C_{\text{ox}} = 0.5\text{fF}/\mu\text{m}^2$. The minimum size capacitor value can be further reduced if the noise is the sole criterion. In reality, however, the minimum feature size and the matching issue limit the minimum capacitor value. The minimum width for the transistor in the op amp is also set to $10\mu\text{m}$. Technology is assumed to be $1.2\mu\text{m}$ CMOS.
5. Input capacitance of a single comparator is assumed to be 100fF.

The per-stage resolutions of interest here are 1 bit/stage and 2 bits/stage, the most widely used values[5][8][9][11][12][13]. Three bits or more per stage limits the sampling rate considerably because a large closed loop gain SC configuration for large interstage gain results in the reduced bandwidth, and therefore large per-stage resolution cases are not considered here for the analysis.

The flow chart for the power analysis is shown in Fig. 49. The analysis starts from

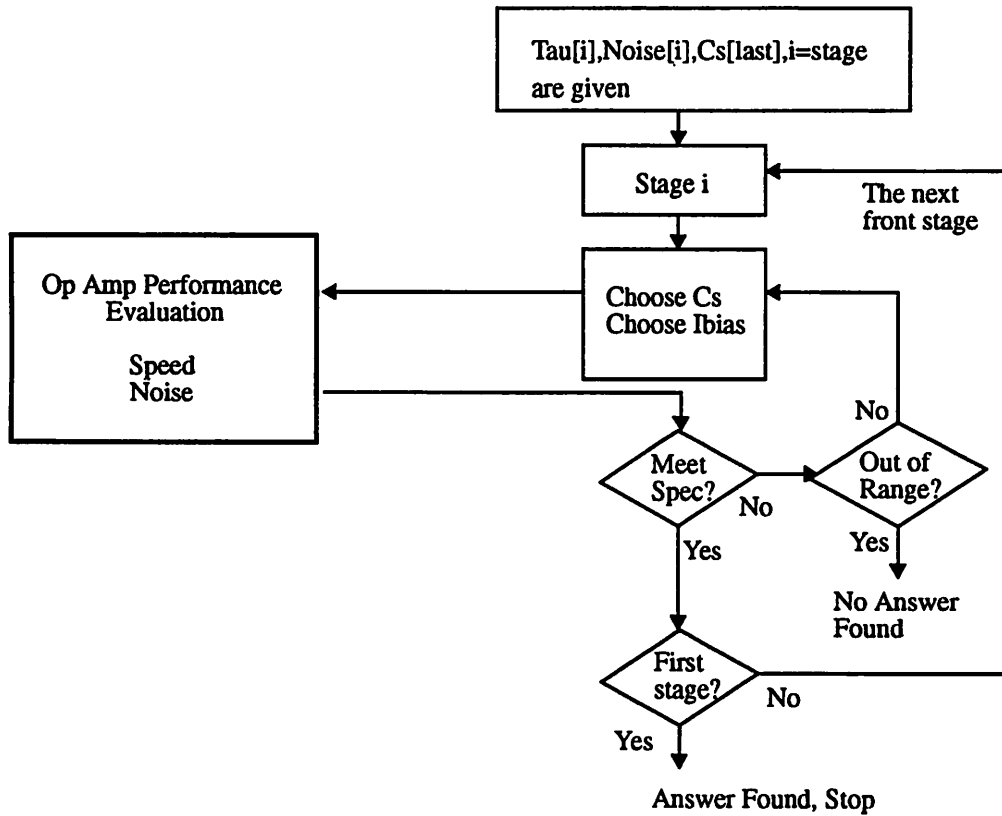


FIGURE 49. Flow chart for power estimation in pipeline architecture.

the last stage of the pipeline and works back toward the front end. For the last stage where the stage resolution is only 2 or 3 bits, the minimum values for the sampling capacitors and op amp size must be set to begin the analysis. Otherwise, the simulation will then give values too small to be drawn by the minimum feature size of the given technology. Once the last stage is designed, the analysis moves to the preceding stage, and the same procedure is repeated. In order to prevent any unreasonable values for capacitors and op amp size, the simulation stops if the values go out of the specified range.

Two cases will be examined in this section. First, the case where identical stages are implemented for all pipeline stages. This is the case when the design time is to be

minimized at the expense of the excessive power consumption due to over-design in later pipeline stages. Next, scaling is applied to each pipeline stage to minimize the power consumption in later part of the pipeline.

4.4.1 No Scaling

When design time is one of the key constraints, often identical stages are used for all pipeline stages[8]. In this case, a single stage is designed to meet the most stringent requirements of the first front stage and replicated for all subsequent stages. So, it is obvious that all later stages are over-designed and guaranteed to meet their specifications. The total power dissipation for identical pipeline stages is simply the power of each stage times the number of stages.

The power of the SC S/H/Gain stage can be simulated for the configuration shown in Fig. 50 where the load capacitance is $2^B C + C_{FLASH}$ where C_{FLASH} is the input

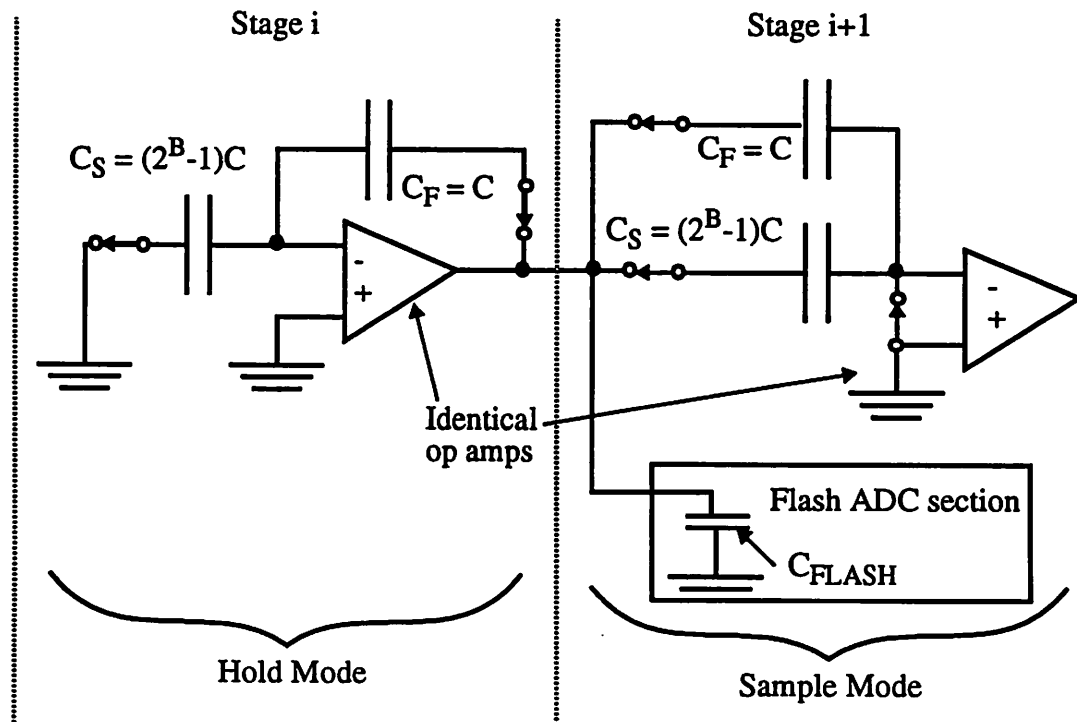


FIGURE 50. Operation of pipeline stages implemented with identical stages.

capacitance of the flash ADC section. The total number of op amps is approximately equal to the resolution of the converter divided by the B. The number of comparators required in the flash ADC section is also given by $(2^{B+1}-2)[7]$.

So, knowing the load capacitance and the interstage gain for the pipeline stage, the power dissipation can be found if a single stage power is known. Power estimation is carried out and the results on power dissipations of 8,10,12, and 14-bit converters are shown in Fig. 51 for B=1 and Fig. 52 for B=2. The time scale is normalized to the ft of the MOS transistor. In both cases, curves for 8 and 10 bit ADC power levels show a flat region at low sampling rate. This is because a certain minimum size SC circuit is assumed as an

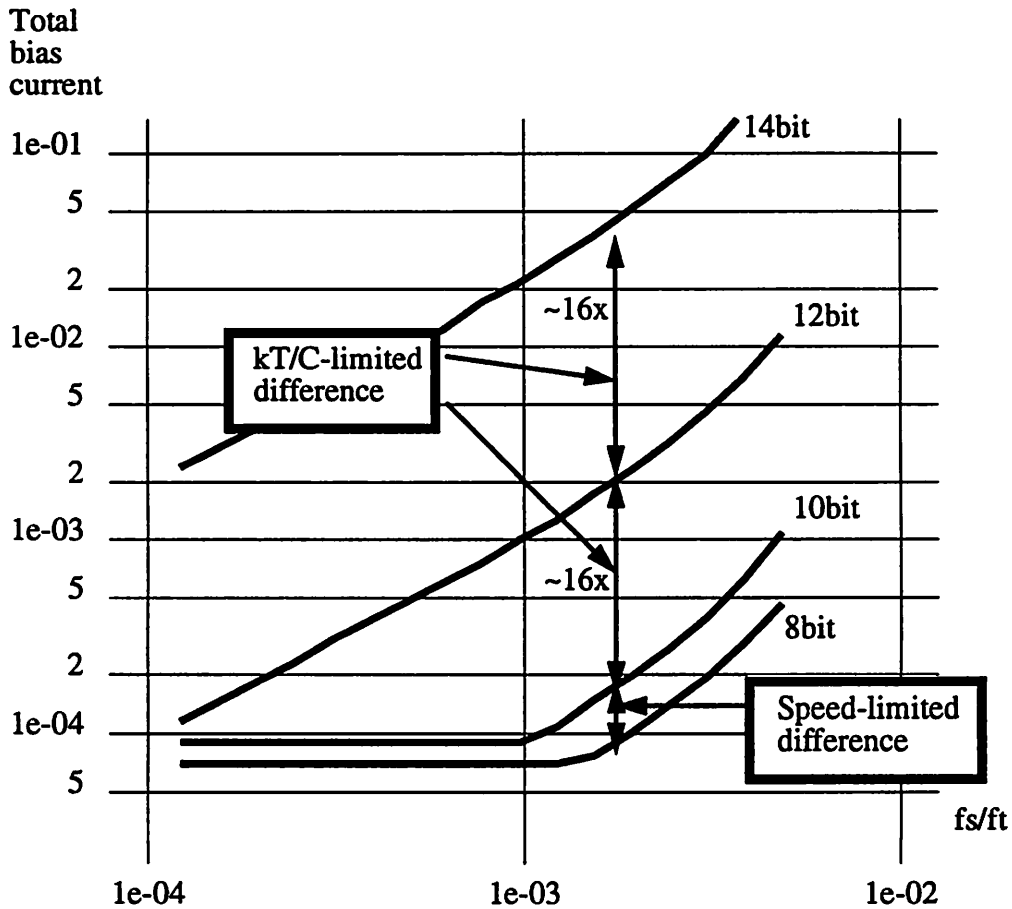


FIGURE 51. Power dissipation of an N-bit pipeline ADC for B=1 vs. fs/ft.

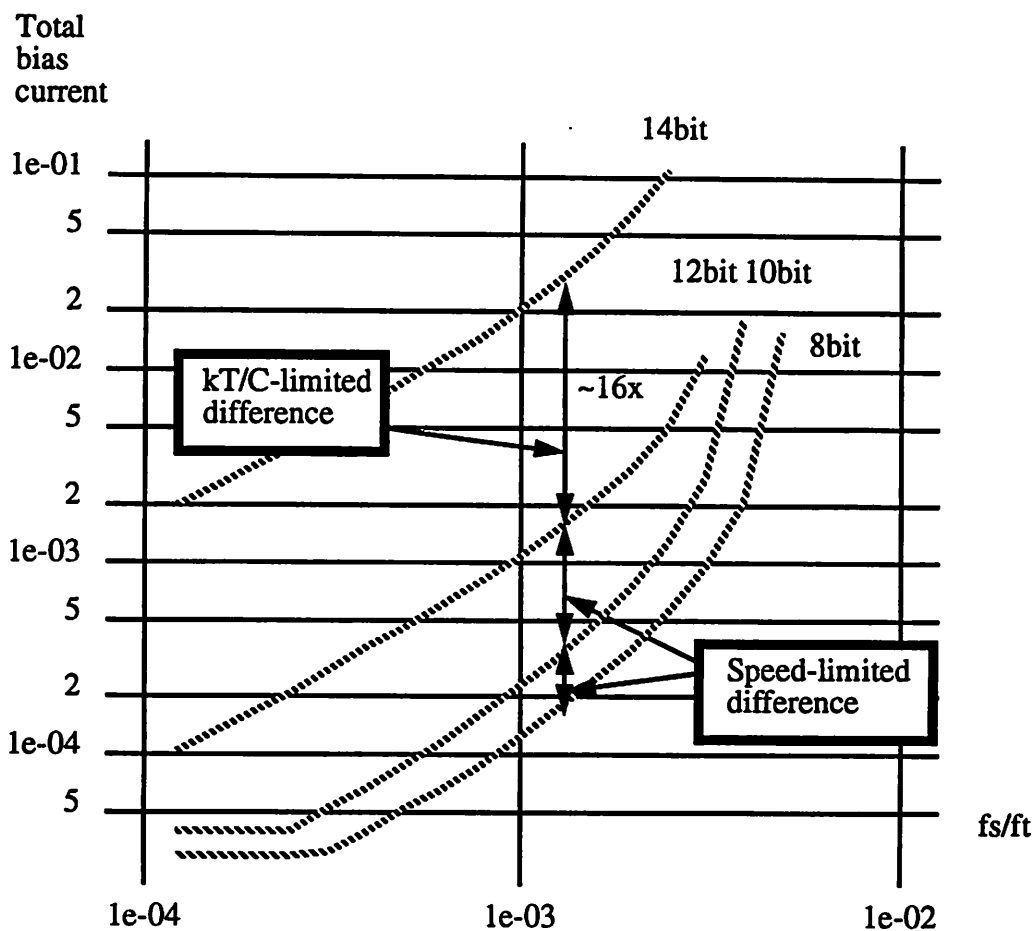


FIGURE 52. Power dissipation of an N-bit pipeline ADC for B=2 vs. fs/ft.

initial starting point for the analysis, and this minimum size is still larger than the value theoretically required for the particular resolution(8 and 10bits in this case). In other words, the pipeline stage for 8 and 10bit ADC's is basically over-designed because the theoretically minimum size op amps are too small to be drawn in practice (at least with 1.2 μ m CMOS technology used here) at low sampling rate. As the sampling frequency increases, the interstage bandwidth must be increased to meet the faster settling time requirement requiring larger bias currents, and as a result the power curves rise from the flat region. Also, notice that the power difference between 8 and 10bit curves is much

smaller than that predicted by kT/C noise consideration discussed in previous sections. This is again due to the fact that the noise is not much of an issue for these resolutions and therefore power is mainly determined by the settling time requirement. However, for higher resolutions, such as 12 and 14 bits, the power difference agrees with the prediction according to kT/C noise consideration; the 2bit difference between 12 and 14 bit ADC's shows $\sim 16x$ difference in power confirming that power dissipation of high resolution ADC's is mainly determined by kT/C noise level for a given speed.

Fig. 53 shows the comparison between $B=1$ and $B=2$ cases by overlapping Fig. 51 and Fig. 52. For low sampling rate and resolution of 8-10 bits, the minimum size pipeline

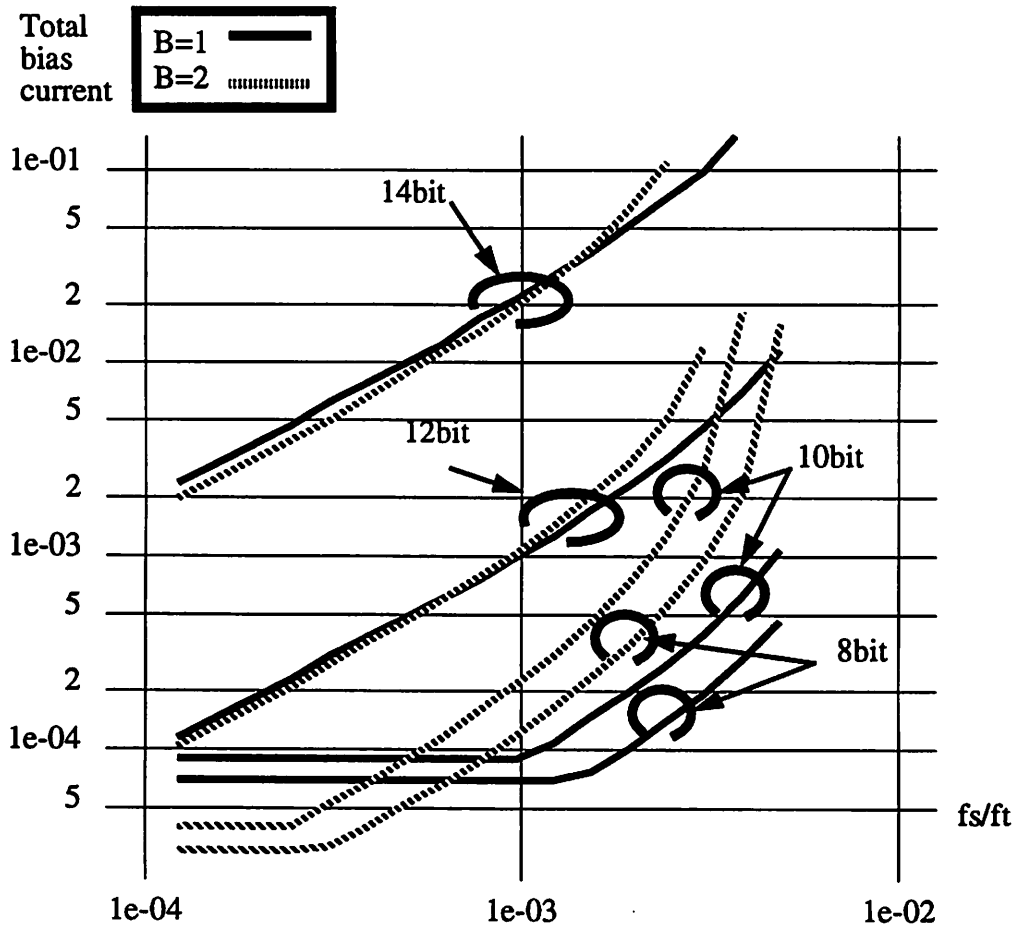


FIGURE 53. Fig. 51 and Fig. 52 on top of each other.

stage can easily meet both noise and speed requirements, and as a result its power is simply determined from the required number of stages; less number of stages (or larger per-stage resolution) requires less power. However, as the sampling frequency increases, settling time requirement becomes a dominant factor in determining the power dissipation, and lower per-stage resolution ($B=1$ in this case) results in lower power dissipation because the interstage amplifier operates much closer to its unity gain configuration with lower interstage gain and larger feedback factor as discussed in section 4.2. So, for these resolutions and high sampling rates where the settling time, not the noise, determines the power, lower per-stage resolution results in net power saving through its larger feedback factor and reduced load capacitance although more stages are required.

For low sampling rate and the resolution of 12-14 bits, power dissipations for both cases are about equal. Inherently the lower closed loop of $B=1$ case is faster than $B=2$ case, but the larger per-stage resolution is preferred since its large interstage gain attenuates the noise when it is referred to the input. As a result, power saving from lower closed loop gain for speed cancels with capacitance increase for noise consideration. Near the speed limit of the technology, $B=2$ case simply cannot make the settling time requirement, and only option becomes $B=1$ case.

Therefore, from comparing these two figures, the optimum per-stage resolution can be found; for low speed operation, since minimum size op amps are sufficient to meet both kT/C and speed requirements, larger per-stage resolution with less number of stages is better for low power. For a high speed operation near the limit of the given technology, lower per-stage resolution allows low closed loop and low load capacitance configuration for the interstage amplifier dissipating less power.

4.4.2 Scaling

Although the previous approach can save design time, later pipeline stages result

in excessive power dissipation since they are over-designed. Further reduction on power dissipation is possible by scaling each pipeline stage according to its stage requirements.

In Fig. 54, the normalized power dissipation of each pipeline stage is plotted for

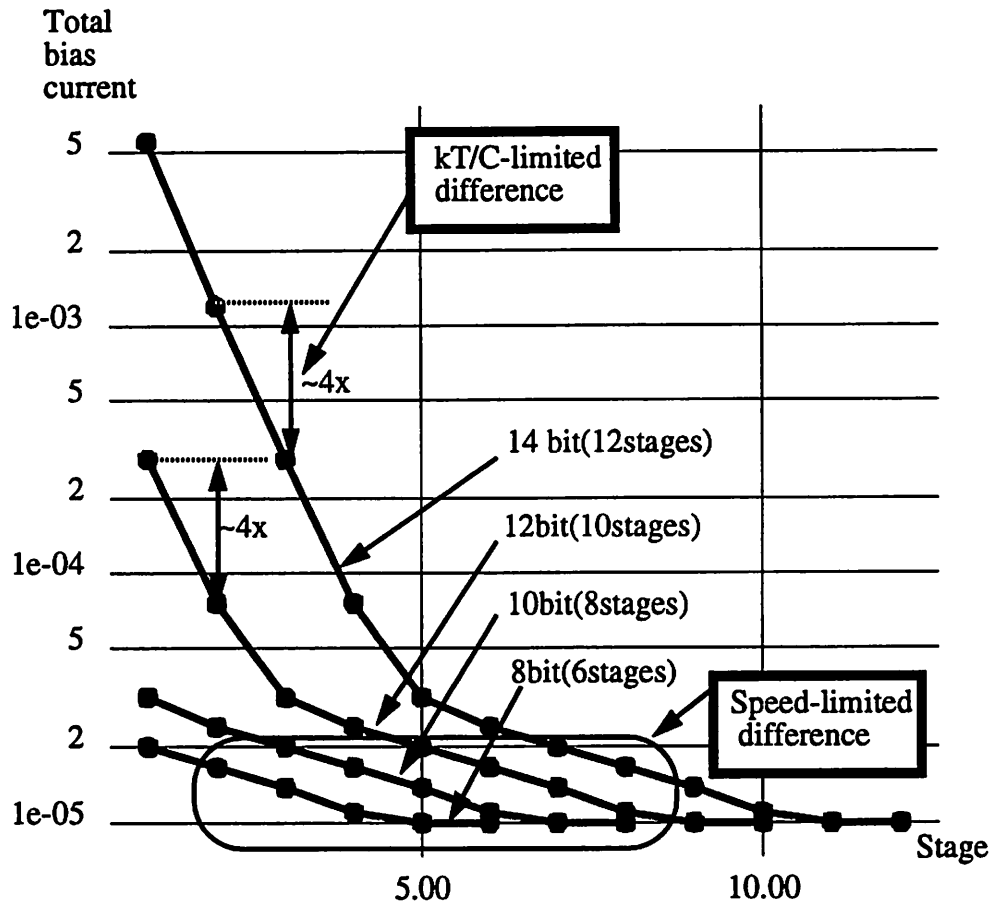


FIGURE 54. Scaling of pipeline stages for different resolutions for $f_s=20\text{MS/s}$ and $B=1$.

different resolutions at $f_s=20\text{MS/s}$. As each stage is placed before the previous one near the front end, the power of the front-end stage increases by factor of 4 dictated by kT/C noise consideration, and the first stage power is more than 50% of the total power for the resolution higher than 12 bits at this sampling rate. In later stages, the slope of the curve is less steep than the front, limited by the settling time requirement. In this case, later stages dissipate only small fraction of the total power and therefore can be implemented with

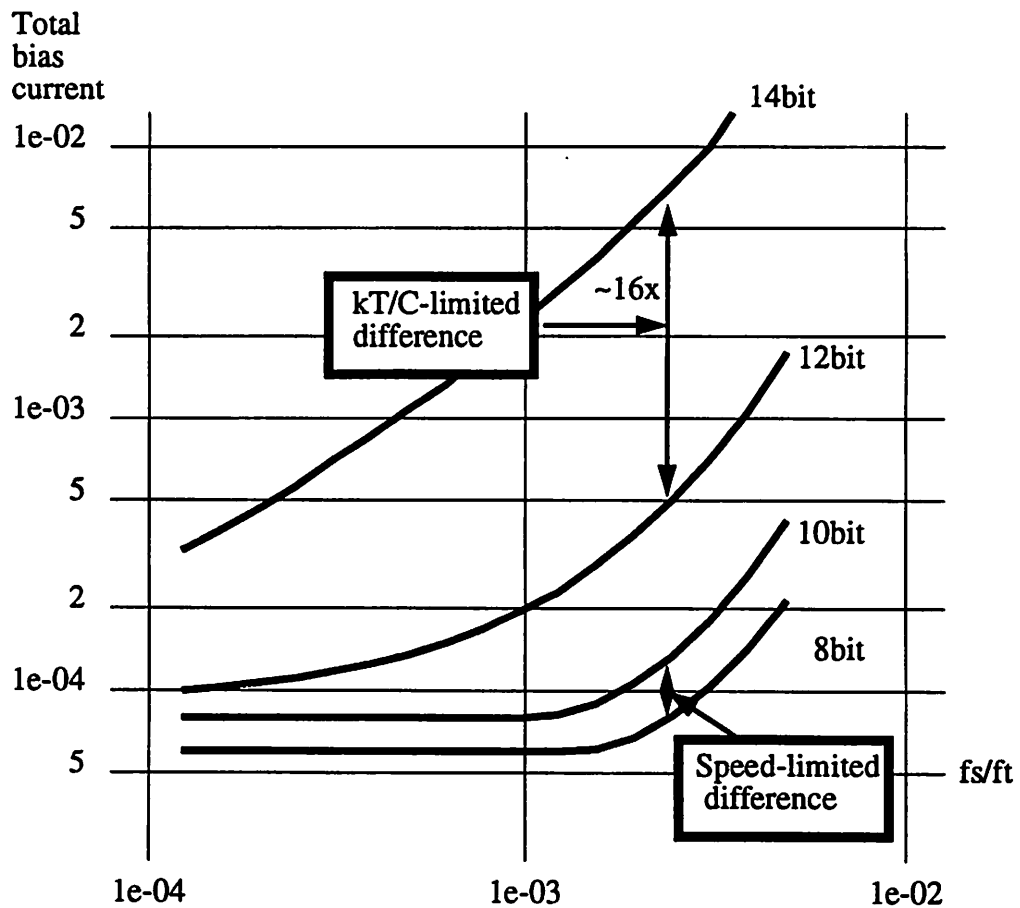


FIGURE 55. The total power dissipation vs. the sampling rate for B=1.

identical stages without much overall power increase.

In Fig. 55, power dissipations for different sampling rates and different resolutions are shown, and similar trends as in no-scaling case can be also found in this case.

Power estimation has been carried out for B=2 case, and their results are shown in Fig. 56 and Fig. 57. Again, similar trends as B=1 case are found except that the power increase near the front is 16x since two bits are resolved per-stage.

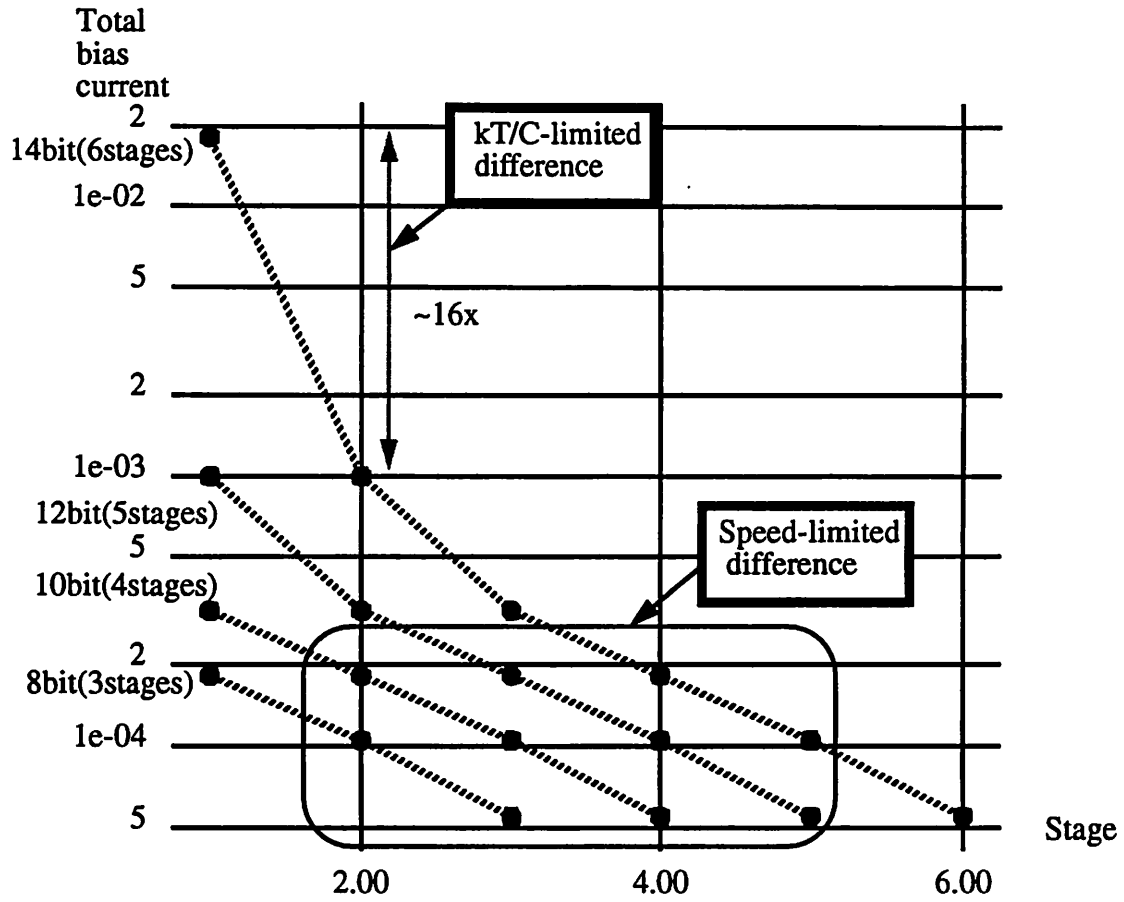


FIGURE 56. Scaling of pipeline stages for different resolutions for $f_s=20\text{MS/s}$ and $B=2$.

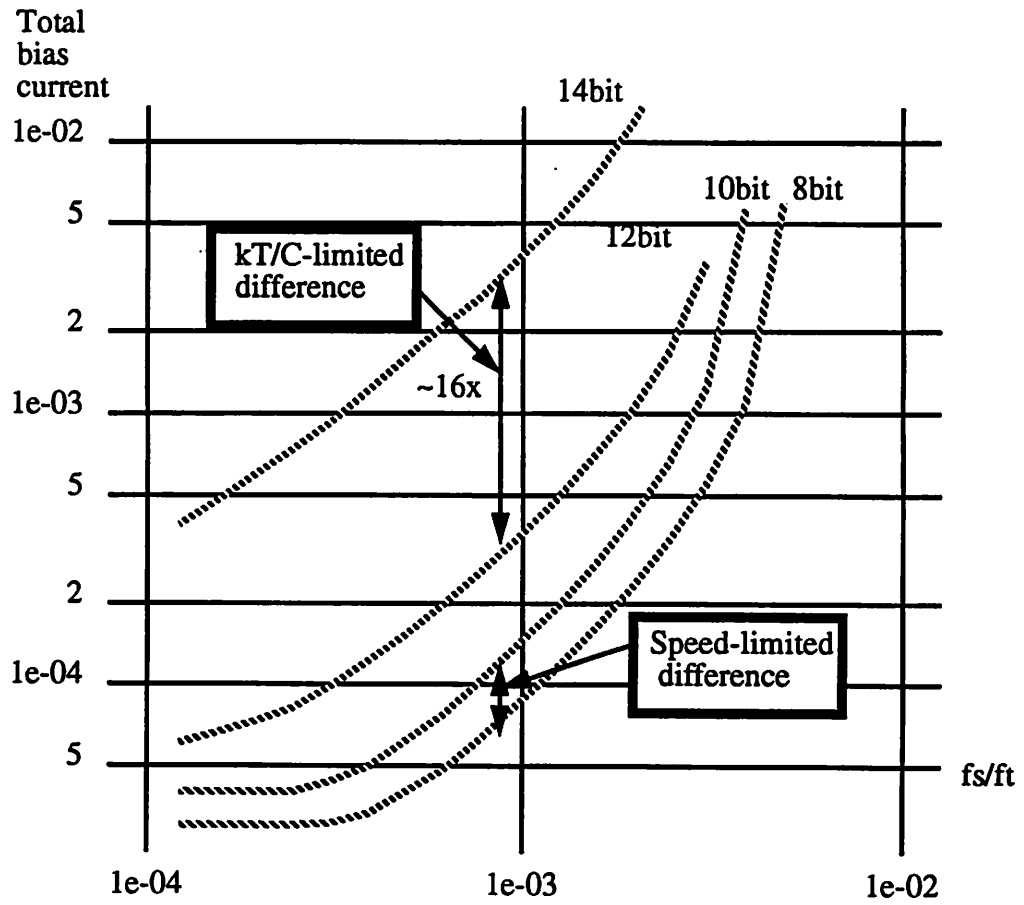


FIGURE 57. The total power dissipation vs. the sampling rate for B=2.

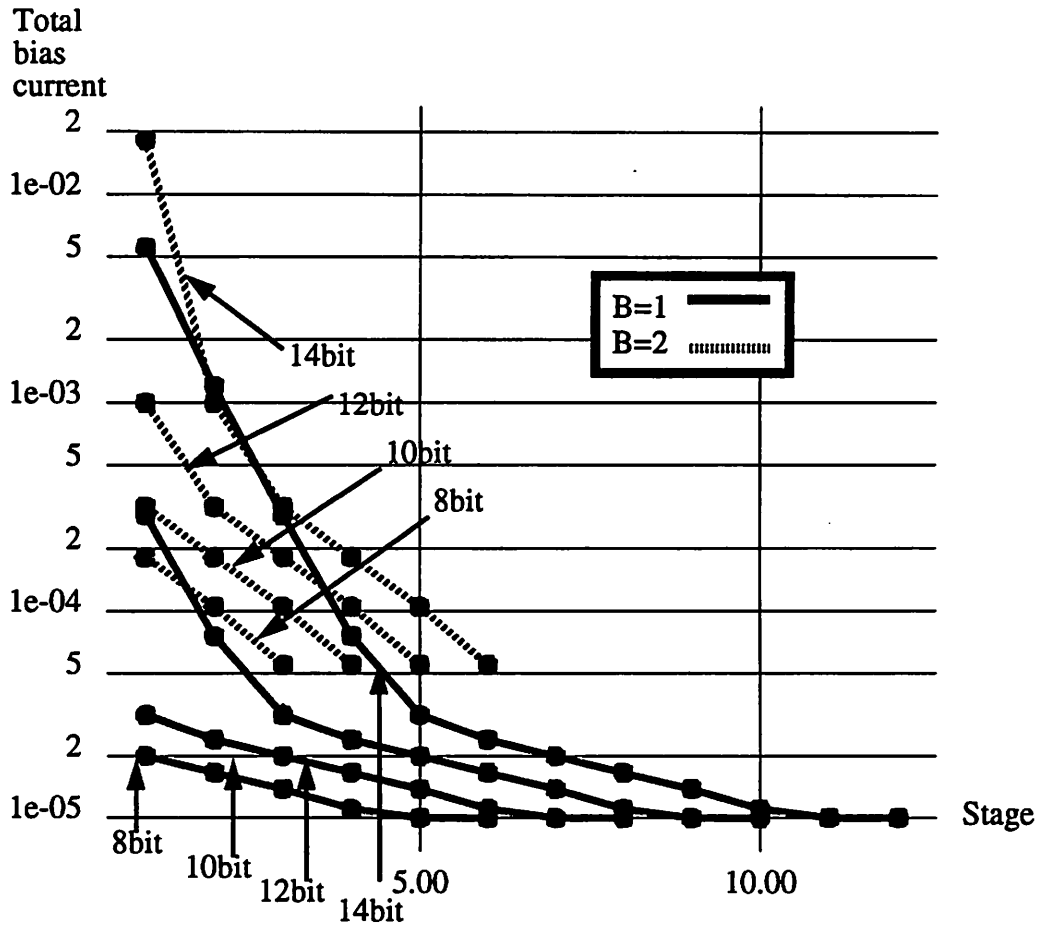


FIGURE 58. Scaling of pipeline stages for different resolutions for $f_s=20\text{MS/s}$ for B=1 and 2

Fig. 58 shows the comparison between B=1 and B=2 cases by overlapping curves in Fig. 54 and Fig. 56. It can be easily seen that the power per stage is much larger for larger per-stage case while it requires only half the number of stages as predicted.

Fig. 59 again shows power comparison for different sampling rates by overlapping curves in Fig. 55 and Fig. 57. Similar trends as no-scaling case still hold here. At low sampling rate, B=2 case dissipates less power simply due to a smaller number of required stages since the pipeline stage of minimum size is sufficient to meet both noise and speed requirements. At a high sampling rate, lower per-stage resolution allows faster operation of the interstage amplifier for a given power dissipation and results in lower power.

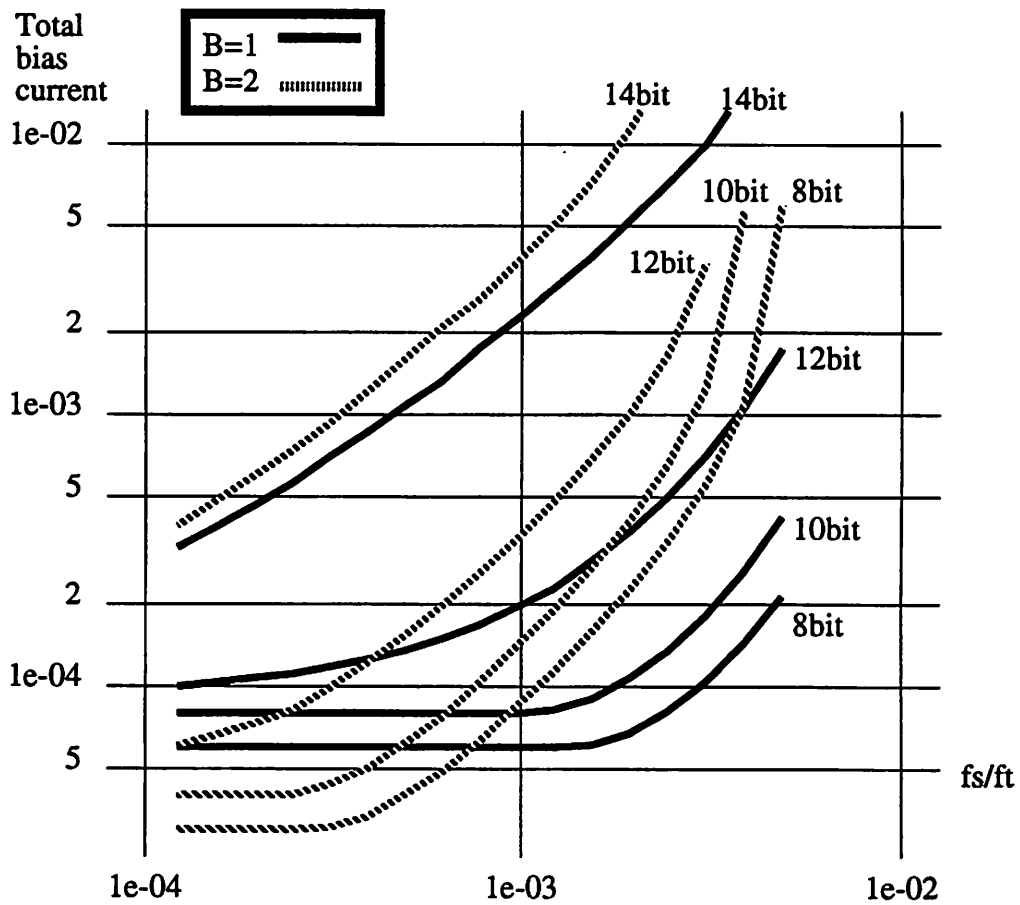


FIGURE 59. Normalized power of each pipeline stage vs. fs/ft for different resolutions and for $B=1$ and 2 .

In Fig. 60, the power dissipation of the scaling case is compared to that of the no-scaling case for $B=1$. As the sampling rate increases, the difference in power dissipation increases since the scaling case requires much less power in later stages. Without scaling, the power dissipation of a 10 bit converter approaches to that of a 12 bit converter with scaling. At high sampling rate, the difference is as high as 5x, indicating that considerable amount of power dissipation can be saved at the expense of increased design time for scaling.

In Fig. 61, the same comparison is made for $B=2$. Again, the power dissipation of

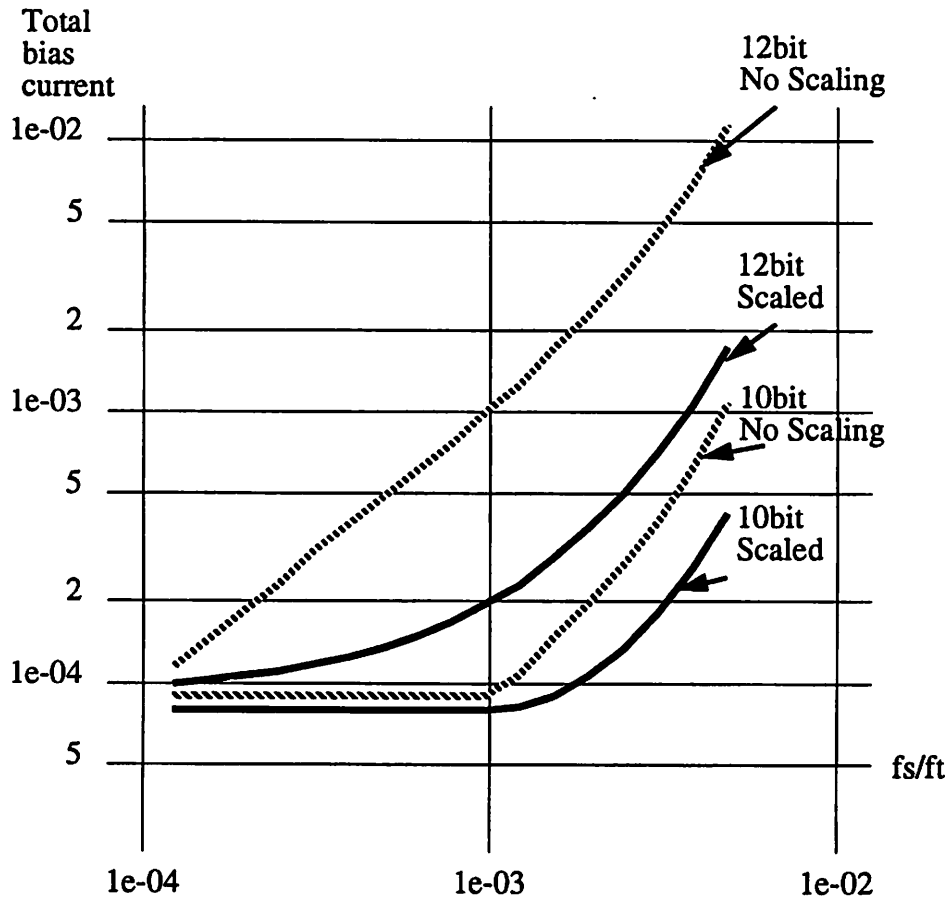


FIGURE 60. Power comparison between scaling case and no-scaling case for B=1.

the no-scaling case is much higher than that of the scaling case. However, the power difference is somewhat smaller than that of B=1 case; the difference is about 2 or 3x between scaling and no-scaling cases for this particular example. Power saving in B=1 case is more significant than B=2 case because lower per-stage resolution allows an interstage amplifier configuration for higher bandwidth allowing each pipeline stage design closer to the noise-limited minimum size.

Based upon the observations made up to now, the following can be concluded.

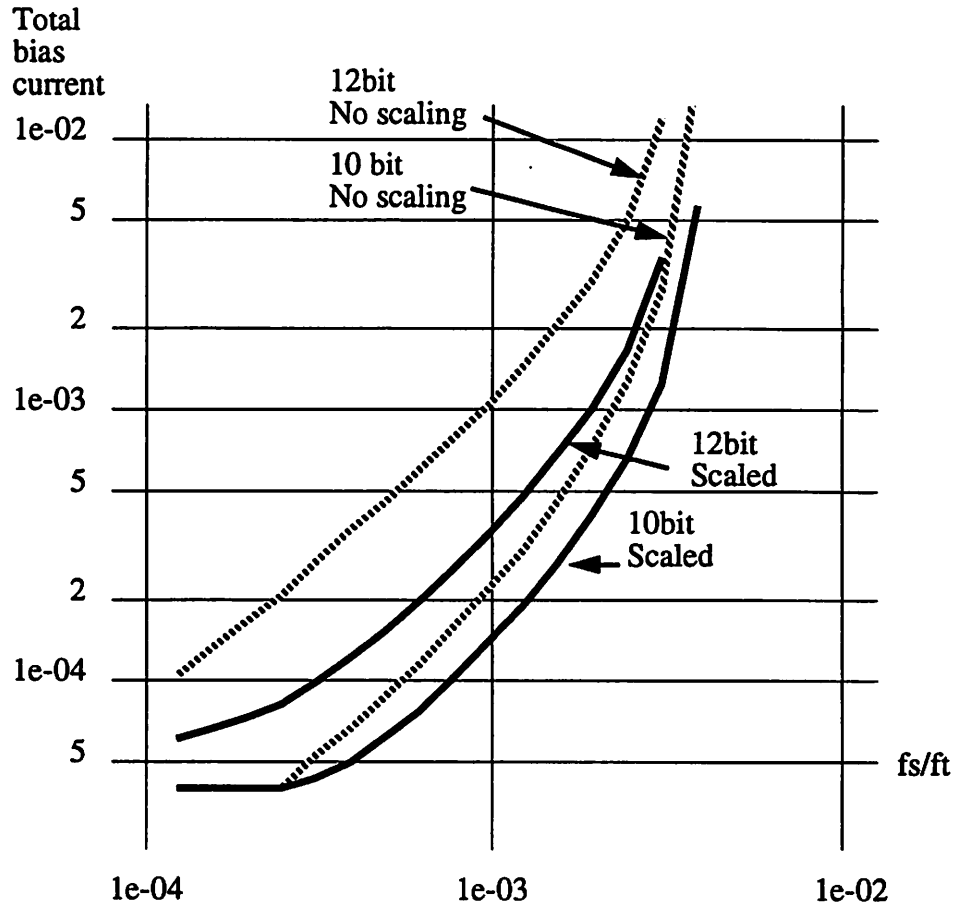


FIGURE 61. Power comparison between scaling case and no-scaling case for $B=2$.

1. Larger per-stage resolution yields lower power dissipation at low sampling rates for both scaling and no-scaling cases if a certain minimum size stage set by the layout issue of the technology is assumed. This is because the minimum size circuit is still larger than the value theoretically required for the particular resolution (8 and 10bits in this case).

2. At higher sampling rates, lower per-stage resolution dissipates less power dissipation overriding the fact that it requires more stages, because it allows each pipeline stage design much closer to the noise-limited minimum size. Larger per-stage resolution design is more limited by the settling time requirement and requires more power as a

result although it requires fewer stages.

3. As expected, scaling pipeline stages consumes considerably less power dissipation than identically implemented pipeline. Also, one can notice that in scaled pipeline stages, the power saving of a lower per-stage resolution to that of a larger per-stage resolution is greater compared to the same situation with identical stages, based on the analysis of this section.

CHAPTER 5

Low Power Low Voltage Techniques in Pipeline A/D Converter Architecture

5.1 Introduction

Up to now, basic discussions on key function implementations, various converter architectures, and a detailed operation of the pipeline converter are presented. In this chapter, techniques to reduce power dissipation and to allow low voltage operation of the pipeline A/D converter architecture are presented. The best way to illustrate the techniques is to present the prototype[5] implementation and its experimental results. In the prototype, per-stage resolution of 1.5 bit is chosen and its justifications are presented first. Then, techniques for low power and low voltage operations are presented with simulation and/or measurement results.

5.2 1.5 bit/stage Pipeline Architecture

A block diagram of a typical pipeline A/D converter is shown in Fig. 62. It consists of a cascade of N identical stages in which each stage performs a coarse quantization, a D/A function on the quantization result, subtraction, and amplification of the remainder. A S/H function in each stage allows all stages to operate concurrently, giving a throughput of

one output sample per clock cycle. Fig. 62 illustrates the particular configuration of

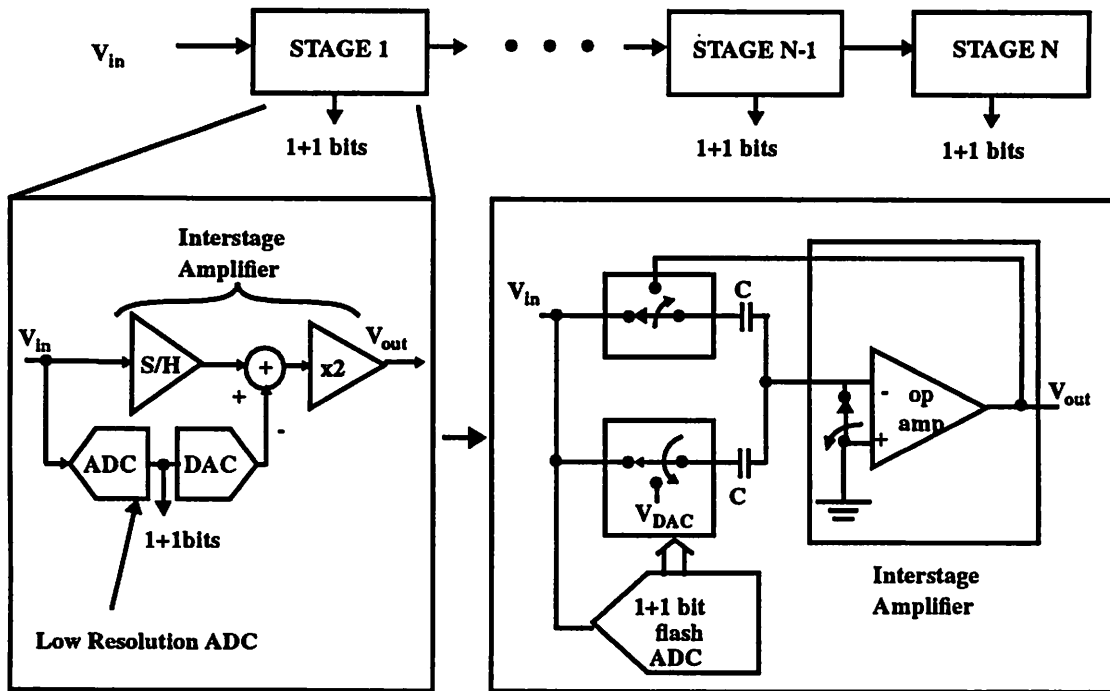


FIGURE 62. A 1.5 bit/stage pipeline A/D converter architecture.

interest here in which the D/A, subtraction, amplification, and S/H functions are performed by a switched capacitor(SC) circuit, with a resolution of 1.5 bits per stage and an interstage gain of 2. The D/A function is performed by two equal capacitors. When the input signal is applied, each stage samples and quantizes the signal to its per-stage resolution of 1.5 bits [8][11][6](i.e. 2 decision levels and 3 possible output codes, 00, 01 and 10 excluding 11), subtracts the quantized analog voltage from the signal by connecting the bottom plate of capacitor C_S to $V_{DAC}(\pm V_{ref}$ or 0), and passes the residue to the next stage with amplification for finer conversion. Then, 1.5 bits from all stages are collected and produce a full 10 bit representation of the applied analog input signal.

The resolution of 1.5 bits/stage is chosen in this pipeline implementation mainly for

the following two reasons. The first reason is to maximize the bandwidth of the S/H/Gain SC circuit which limits the overall conversion rate. In order to perform fast interstage signal processing, the output of operational amplifier in the SC circuit has to settle in half the clock period to the given accuracy of each stage prior to the next stage sampling instance. Since the bandwidth of the SC interstage amplifier depends on its interstage gain, choosing the per-stage resolution which allows the low closed-loop gain configuration for fast settling is essential. With the resolution of 1.5 bits/stage, the closed-loop gain of only 2 allows configuration for low load capacitance (composed of only two sampling capacitors of the next stage and input capacitance of two comparators in the flash A/D section) and large feedback factor (of about 1/3), and as a result a large interstage amplifier bandwidth can be achieved compared to that of larger per-stage resolution(2-3bits/stage).

Also, the resolution of 1.5 bits/stage allows large correction range for comparator offsets in the flash A/D section. Only two comparators are required in the flash A/D section of each stage, and the comparator offset up to $\pm V_{ref}/4$ can be tolerated without degradation of the overall linearity or SNR. This is illustrated with residue plots in Fig. 63. Input and output ranges of each stage are both $\pm V_{ref}$. Fig. 63(a) shows ideal case with zero comparator offsets, and in Fig. 63(b), the shifted residue plot due to the comparator offset ΔV is shown. With the use of digital correction algorithm in 1.5 bits/stage pipeline architecture, the overflow of present stage output from the input range of the following stage can be prevented even with the presence of a large comparator offset up to $\pm V_{ref}/4$, so that this offset error amplified down the pipeline can be detected for correction.

This large error correction range can also eliminate the dedicated input S/H circuit. Instead, the input signal can be sampled simultaneously by the switched capacitor amplifier and by the dynamic comparators of the flash A/D section in the first stage as illustrated in Fig. 64. This is made possible by the fact that digital correction allows

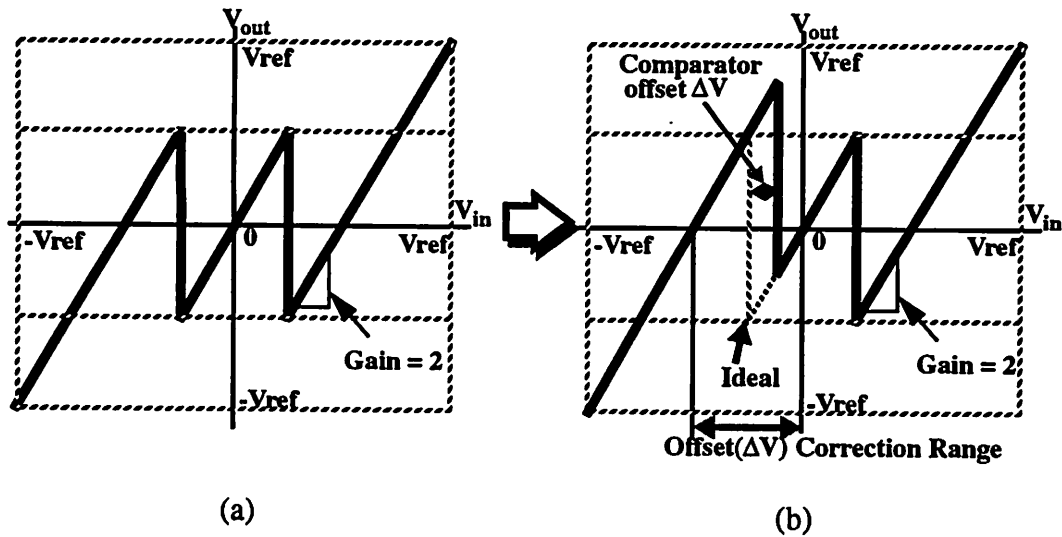


FIGURE 63. Residue plots (a) ideal (b) with comparator offset ΔV .

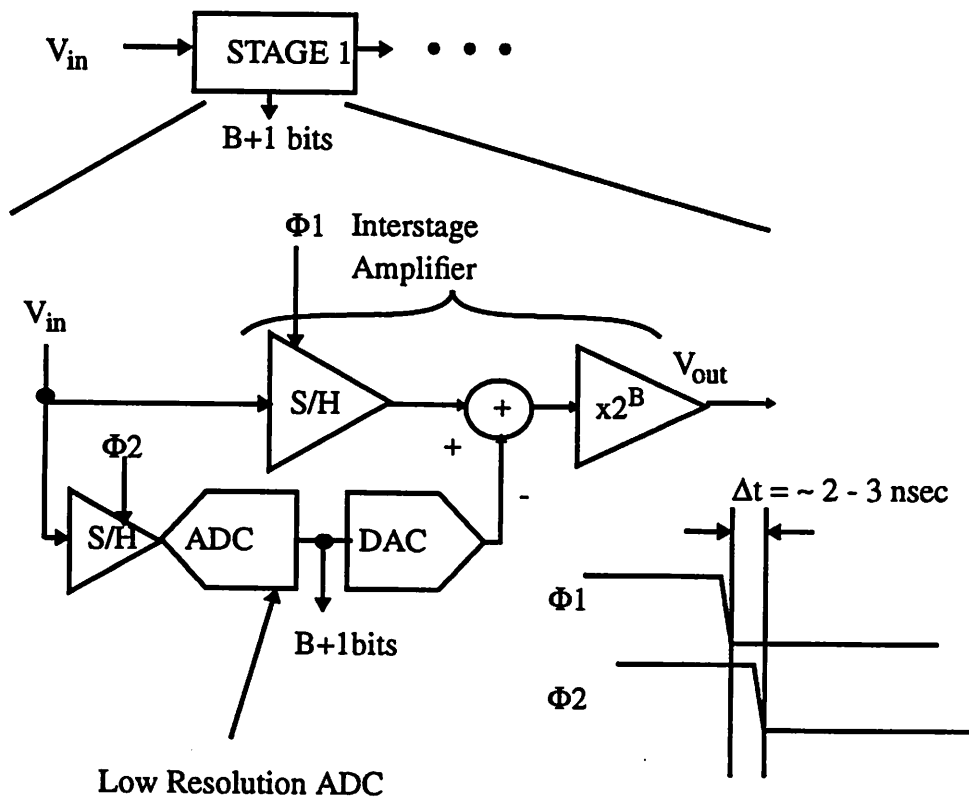


FIGURE 64. The first stage input sampling network.

comparator errors up to $\pm 1/4 V_{ref}$ without degradation of linearity or SNR with its input bandwidth given by

$$f_{input, BW} \approx \frac{V_{correction} - \sum V_{offset}}{A \cdot 2\pi \cdot \Delta t} \quad (EQ 45)$$

where $V_{correction}$ is the digital correction range, $\sum V_{offset}$ is the sum of comparator and interstage amplifier offset voltages, A is the input sinusoid amplitude, and Δt is the time difference between sampling instances as illustrated in Fig. 64. With $A=1V$, $V_{correction} = \pm 250mV$, $\Delta t = 3nsec$, and $\sum V_{offset} = 50mV$, input bandwidth is around 16MHz. For the sampling rate of 20MS/s in the prototype, this is larger than the Nyquist frequency of 10MHz and therefore acceptable for most video-applications whose signal bandwidth is around 5MHz[44]. Since Δt , which is created by simple inverter delays, scales with technology, the input bandwidth is expected to increase with faster scaled technologies for higher sampling ADC if this sampling scheme is used. In terms of power dissipation, eliminating the input S/H circuit saves about 20 - 30% of overall op amp power at a 10 bit level since the S/H circuit has to meet full resolution requirement potentially dissipating large power. Without the input dedicated S/H, the overall pipeline contains 9 2-bit flash quantizers and 8 interstage amplifiers.

5.3 Power Reduction Techniques

In pipeline A/D converters, a major portion of the total power dissipation is from the static power dissipated in analog circuit components that require DC bias currents, such as precision comparators and op amps. The charging/discharging of sampling capacitors, clock drivers and digital circuits contribute a relatively small amount to the overall power dissipation. Therefore, in this implementation, major effort is taken to reduce DC power dissipation. One effective method is to use dynamic comparators to implement the low-resolution flash A/D section in each stage, since the digital correction can relax the

comparator accuracy requirements as mentioned in section 5.2. In this way, static power dissipation of precision comparators can be eliminated. Also, a substantial power reduction can be achieved by using the minimum possible size of sampling capacitors at each point in the pipeline, as dictated by kT/C thermal-noise considerations. This is possible since later stages can tolerate more noise due to decreasing stage resolution down the pipeline and therefore can be made small to reduce power consumption. These design approaches are discussed in detail in the following sections.

5.3.1 Dynamic Comparators

In high resolution A/D converters, precision comparators consume DC power since low-offset pre-amp stages are required to amplify the signal before an accurate comparison is made. However, in the pipeline architecture, the error from a large comparator offset in flash A/D section of each pipeline stage can be easily compensated with digital correction. As mentioned above, for a 1.5 bit per-stage resolution, the comparator offset up to $\pm V_{ref}/4$ can be corrected. So, for example, with a reference voltage of 1V used in the prototype, comparator offset up to ± 250 mV can be tolerated. In typical dynamic cross-coupled inverter latches, process variations and mismatches can result in large offset voltage but they can still meet this offset requirement easily. So, without the use of a pre-amp, simple dynamic latches can implement the comparators in the low resolution flash A/D converter to remove DC power dissipation.

One implementation of a dynamic comparator is shown in Fig. 65. Here the lower set of NMOS devices operate in the triode region and they are connected to the input and the reference. As the upper cross-coupled inverter-latch regenerates when the latch clock goes high, the drain currents of the active switching NMOS devices are steered to obtain a final state determined by the mismatch in the total resistance. In this case, resistances (R_1 and R_2) or conductance ($G_1=1/R_1$ and $G_2=1/R_2$) of NMOS pairs biased in triode region are given to the first order by

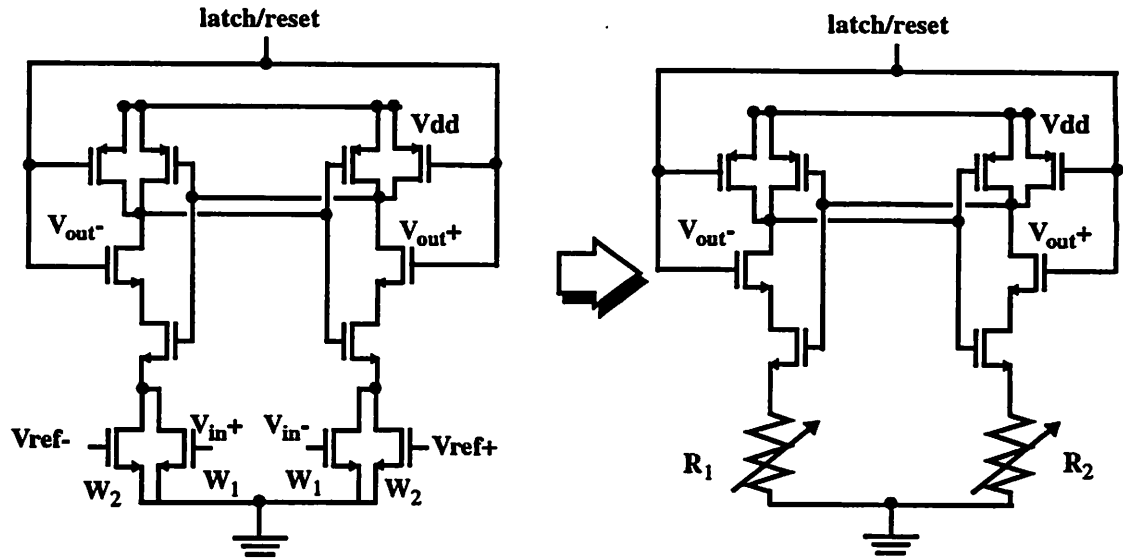


FIGURE 65. A dynamic comparator with built-in threshold level.

$$G_1 = \frac{1}{R_1} = kp \left[\frac{W_1}{L} \cdot (V_{in+} - V_{th}) + \frac{W_2}{L} \cdot (V_{ref-} - V_{th}) \right], \quad (\text{EQ 46})$$

$$G_2 = \frac{1}{R_2} = kp \left[\frac{W_1}{L} \cdot (V_{in-} - V_{th}) + \frac{W_2}{L} \cdot (V_{ref+} - V_{th}) \right]. \quad (\text{EQ 47})$$

The input voltage which causes G_1 equal to G_2 is the comparator threshold voltage. From EQ(46) and EQ(47), it is given by

$$V_{in}|_{\text{threshold}} = \frac{W_2}{W_1} \cdot V_{ref} \quad (\text{EQ 48})$$

where $V_{in} = V_{in+} - V_{in-}$ and $V_{ref} = V_{ref+} - V_{ref-}$.

Therefore, arbitrary noncritical thresholds can be set by properly ratioing the triode region device widths, (W_2/W_1) , without the use of any sampling capacitors or switches at

the input. In this implementation, the required (W_2/W_1) ratio is 1/4 to generate comparator threshold levels at $\pm V_{ref}/4$. Of the 35 mW total dissipation in the experimental A/D converter operating at 20MS/sec, it is estimated that only 3.5 mW was dissipated in 18 comparators.

5.3.2 Scaling of SC Circuits Through the Pipeline

A fundamental noise source present in A/D converters is thermal noise, and the magnitude of this noise is a function of the sampling capacitor size ($\sigma_{\text{thermal}}^2 \sim kT/C$). For instance if the input signal is sampled on a 1pF capacitor through a MOS transmission gate, the voltage sampled on the capacitor contains not only the signal but also the noise voltage whose RMS value is 64 μV at room temperature. Therefore, in this ideal case, the minimum achievable power dissipation in a MOS sample/hold circuit is set by the maximum allowable value of this kT/C noise to achieve the required signal-to-noise-ratio(SNR) before quantization. This sets the minimum sampling capacitor value, which in turn sets the minimum power dissipation for a given sample rate assuming the capacitor must be completely discharged on each sample period. At room temperature this limit corresponds to about 0.2 μW per MS/sec at the 10-bit resolution level, assuming the RMS thermal noise is set to cause 1dB decrease in overall SNR over and above that contributed by ADC quantization noise in the ideal case (see section 2.3.1), and that the signal swing is equal to the supply voltage. The required power dissipation quadruples for each additional bit of resolution and is independent of power supply voltage.

This limit is about four orders of magnitude below the dissipation achieved in recently described high-speed A/D converters. In practice, the S/H power is dominated by dissipation in the (usually class A) operational amplifier or buffer that drives the sampling capacitor in the sample and/or charge transfer modes. As a practical matter, power minimization in the overall A/D converter translates to minimizing the power in the active circuitry driving the sampling capacitor whose kT/C noise limits the SNR of the converter.

In the pipeline architecture, this again translates into minimizing the SC circuit power in each stage. In order to do so, the minimum allowable value of sampling capacitor must be used at each point of pipeline, since it becomes the load capacitance of the previous stage and the size of the amplifier is proportional to that of the capacitor for given speed. Thus, optimization of the power dissipation of each of the operational amplifier in the pipeline can be performed taking into account the source, load, and feedback capacitors seen by each one. Noting that the stage requirements on the speed and accuracy become less stringent as the stage resolution decreases down the pipeline, stages in the later part of the pipeline can be scaled down by using smaller sampling capacitors and op amps. In this case, the sizes of sampling capacitors and op amps near the front end are determined by the noise floor, and toward the end of the pipeline, parasitic capacitances begin to dominate, and so settling time requirements determine the size of each stage.

In the prototype, the optimization resulted in power dissipation ranging from 4.8mW in the first stage amplifier to 0.5mW in the last stage amplifier. In Fig. 66, normalized op

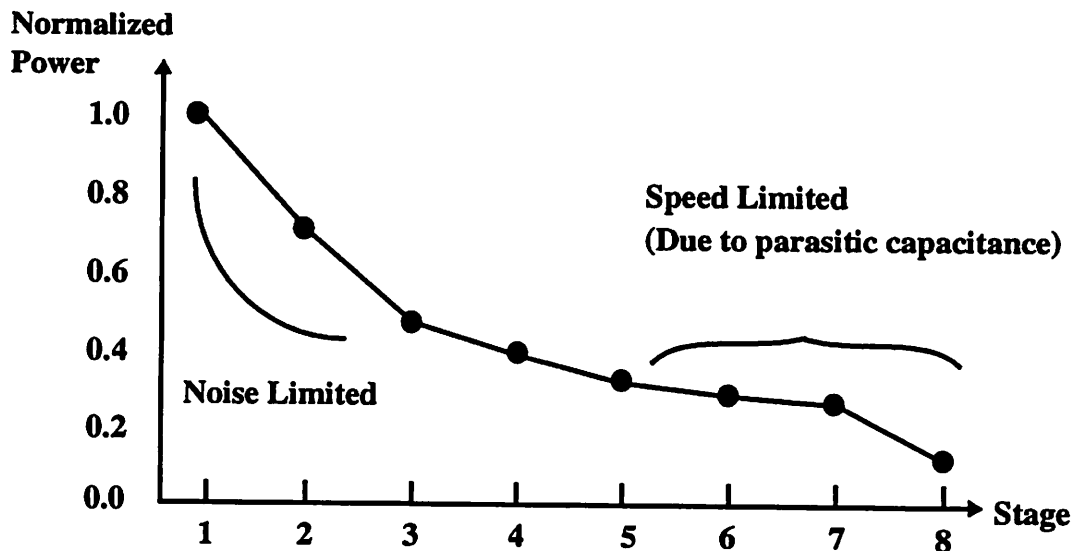


FIGURE 66. Scaling of pipeline stages.

amp bias currents of each stage are shown. Through this optimization process, the power

dissipation can be reduced by about 40 - 50% relative to the dissipation if all stages are identical.

One implication of the use of small capacitors in the first three stages is that the 0.1% matching in the D/A capacitors required for 10 bit INL will not be achieved in the as-fabricated state. Calibration circuitry has been incorporated into the first three stages to remove these mismatch errors[9]. This circuitry consists of a small T network of trim capacitors around the input sampling capacitor, and in the prototype, these are adjusted using external calibration logic control

5.4 Low Voltage Operation

For compatibility with low voltage digital IC systems, 3.3V supply was chosen for the prototype, and this requires the solution of two problems to operate SC circuits at low voltage. First, a high-speed 3.3V op amp with an output swing that is a large fraction of the supply voltage, and with large enough voltage gain for the desired resolution, is required. A second major problem in standard CMOS technologies is the fact that for 3.3 volt supplies transmission gates produce a high (or infinite) resistance region near the mid-supply voltage due to insufficient gate drive. Solutions to these two problems are described next.

5.4.1 High-speed 3.3 V Op Amp

In the pipeline architecture the most stringent requirement on the op amp is on the first stage where DC gain in excess of 60dB and 0.1% settling time under ~20nsec are required to implement an accurate SC S/H/Gain block for 10-bit 20MS/s operation. In a typical 1.2 μ m CMOS technology, designing such an op amp with power consumption of a few mW's is a difficult task. Especially with a 3.3V supply it becomes more challenging where the triple-cascode op amp structure, the simplest way to increase gain while

maintaining high speed as in [8][11], cannot be used due to a limited output swing. While a folded cascode op amp can be used at 3.3V, its folded implementation requires a slow PMOS transistor in the signal path and degrades achievable DC-gain due to reduced output resistance at the folding node compared to straightforward cascode. To improve the DC-gain, multi-stage configurations with pole-split compensation are attractive, but because of the non-dominant pole resulting from the load capacitance and the necessity of driving a compensation capacitor, a substantial degradation in achievable bandwidth and settling time at a given power dissipation can result.

For 10-bit 3.3V operation, however, another viable solution is to use a cascode stage with a low-gain, wideband pre-amplifier to increase the gain by a factor of about 2. While this does add another stage with its power dissipation, it has only NMOS transistors in its signal path and preserves the very desirable property of the cascode amplifier that the load capacitance is also the compensation capacitance. The low-gain preamplifier increases the effective g_m of the transconductance stage and provides necessary DC bias level shifting for the second stage input. However, a non-dominant pole is introduced due to the input capacitance of the transconductance stage. In this case, choosing the optimum value for preamplifier gain is important not to waste any achievable bandwidth. For example, if the preamplifier gain is too small, the “boosting” effect on the g_m of the transconductance stage will be sub-optimum. If it’s too large, then the non-dominant pole will be brought down and limit the bandwidth. Therefore, there will be an optimum value for the preamplifier gain for the given SC configuration which achieves minimum settling time.

This configuration was utilized in the prototype, and its circuit diagram is shown in Fig. 67. The optimum value for the preamplifier gain for the given SC configuration was about 1.75. In the particular technology used for the prototype, the output resistance of the PMOS transistor was much worse than that of the NMOS transistor for the same bias condition. Thus series feedback gain-boost amplifiers[40] are included in the PMOS

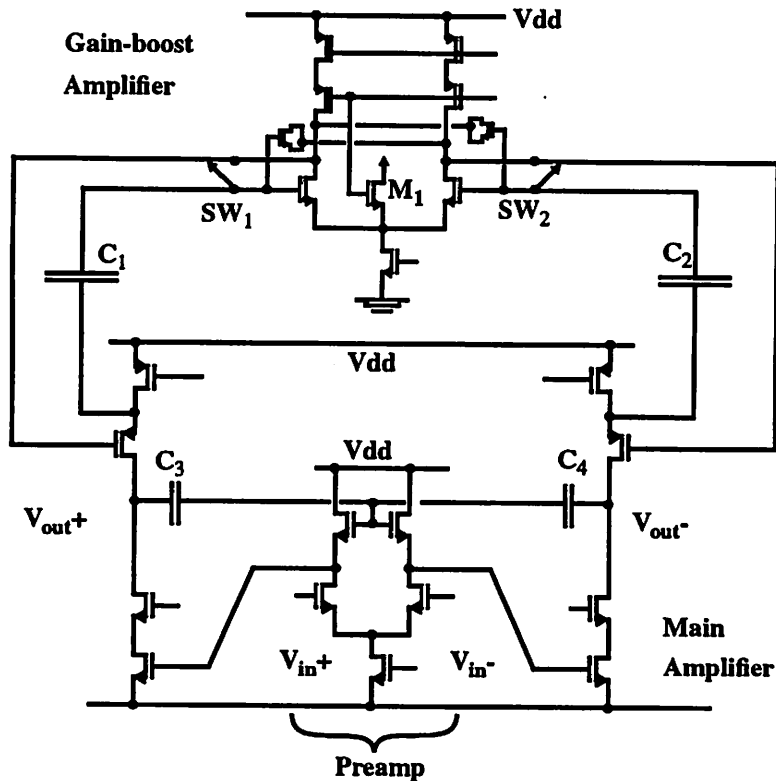


FIGURE 67. A 3.3V high-speed high-gain op amp.

current source to provide an adequate voltage gain as shown in Fig. 68. These amplifiers implemented in differential configuration are capacitively coupled into the signal path using level shift capacitors C_1 and C_2 which are initialized by closing switches SW_1 and SW_2 with transistor M_1 connected to desired input common mode level. The common-mode feedback of the main amplifier is also capacitive through C_3 and C_4 . This is also illustrated in Fig. 68.

Running on a single 3.3V supply, the first stage amplifier achieved a simulated 0.1% settling time of about 17 nsec with $C_S=C_F=0.39\text{pF}$ and external load of 1.8pF. Power dissipation of the first stage op amp is 4.8mW and from experimental results it can be deduced that the voltage gain is greater than 60dB. Gain-boost amplifiers are used only for first three stages, since DC gain requirements in later stages are relaxed.

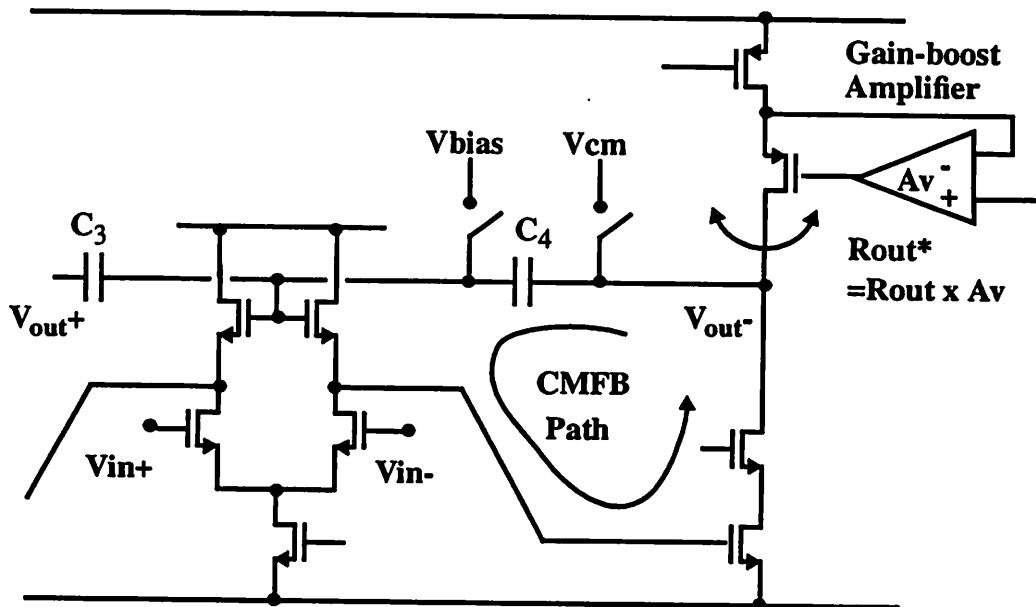
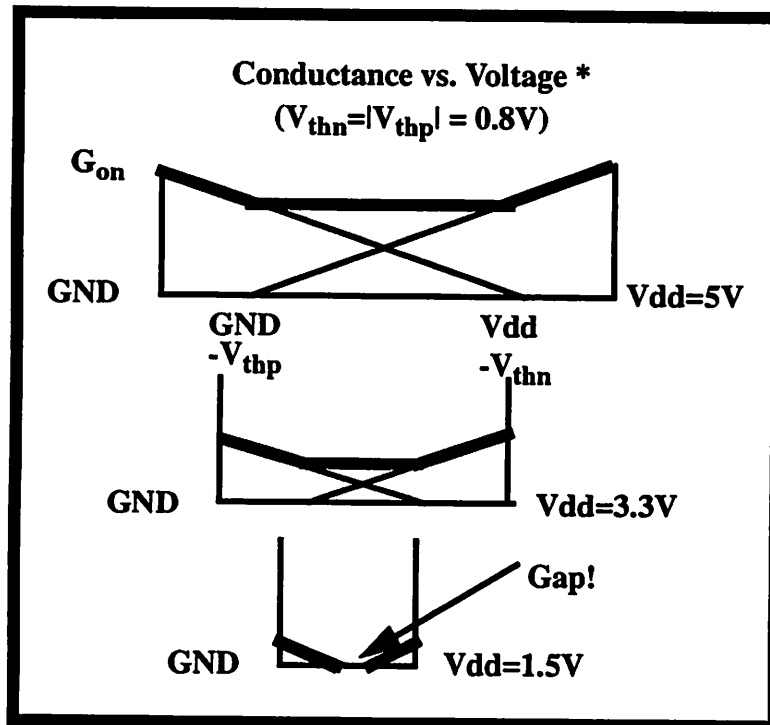
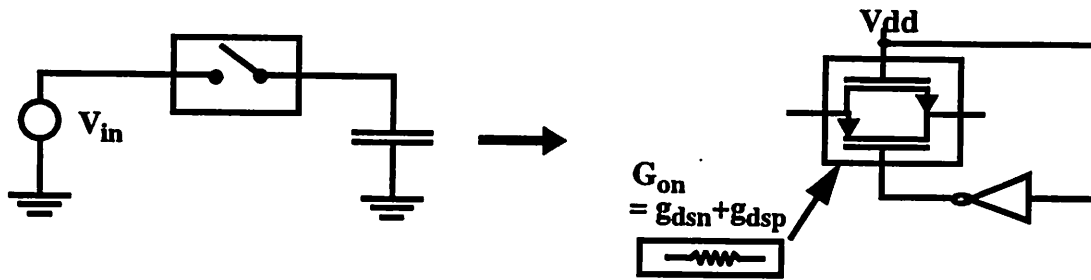


FIGURE 68. CMFB and gain-boost amplifier

5.4.2 Low Voltage Operation of SC Circuits

In standard CMOS technologies, the threshold voltage of MOS transistors (typically $\sim 0.8V$) does not scale with the supply voltage, and it becomes a large portion of the supply voltage leading to problems when MOS transistors are used as switches at low voltages as illustrated with conductance($1/\text{resistance}$) plots for different supply voltages in Fig. 69

For instance, assuming supply voltage of 3.3V, the input signal voltage at the mid-point of the supply, and threshold voltage of about 1.3V with body-effect, the gate voltage overdrive given by $(V_{gs}-V_{th})$ becomes 0.35V ($V_g-V_s-V_{th}=3.3V - 1.65V - 1.3V= 0.35V$). In this case, switch on-resistance can vary by 30-60% if threshold voltage changes by $\pm 100\text{-}200\text{mV}$ with process variations, resulting in speed degradation of the SC circuits. Although large transistor switches can be used for the worst case V_{th} design, the switch



*Integrated A/D&D/A Converter, OCATE, July 1991
Low Power ADC by Vlado Valencic

FIGURE 69. CMOS transmission gates conductances for different supply voltages.

parasitic capacitance can significantly overload the output of SC circuits, especially in later stages where they are small due to scaling. Therefore, in order to solve this problem, increasing $(V_{gs} - V_{th})$ is desirable to implement low on-resistance MOS switch without adding too much parasitic capacitance.

There are several possible ways to increase this gate voltage drive. One method is to

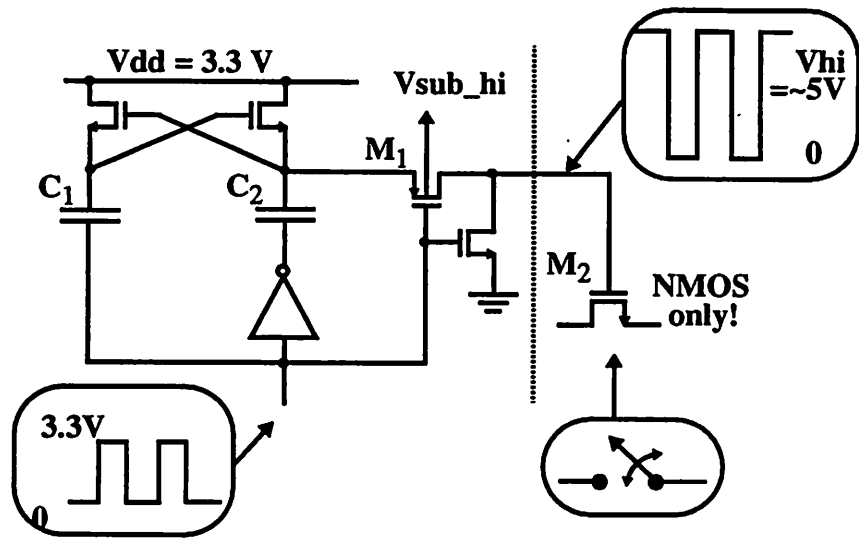
reduce V_{th} by including an extra low-threshold ($\sim 0.3V$) transistor in the process. However this adds process complexity. Another method is to increase V_{gs} by using one large 5V supply created from 3.3V chip supply to drive all switches on the chip, but potential problems of this method include possible cross-talk to some sensitive nodes through the shared supply and difficulty in estimating the total charge drain to drive all switches. Another viable solution is to simply use a dynamic circuit to locally boost the clock drive. In this case each individual charge pump circuit drives each transmission gate or set of transmission gates that use the same clock to avoid the problem of crosstalk through the clock line.

In the prototype, the last approach is used with the use of a high voltage generator circuit shown in Fig. 70(a). By applying a square wave input signal of 3.3 V, C_1 and C_2 are self-charged to 3.3 V through the cross-coupled NMOS transistors[34], and an inverted square wave output of ≈ 5 V is generated according to

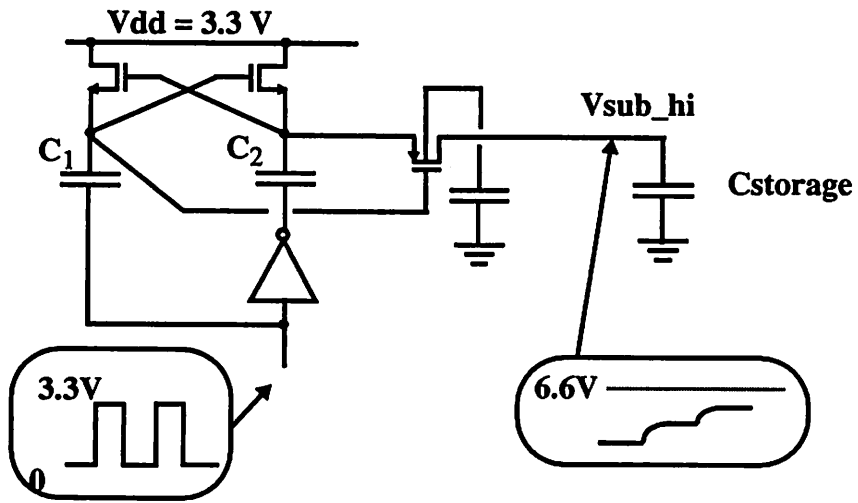
$$V_{hi} = 2V_{dd} \cdot \frac{C_2}{C_{gate,M2} + C_2 + C_{parasitic}}. \quad (EQ 49)$$

where $C_{gate,M2}$ is the gate capacitance of transistor M_2 . Because this gate voltage overdrive is much higher than the signal common-mode voltage ($\approx V_{dd}/2$), sampling switches are implemented with only NMOS transistors, and the parasitic capacitance from PMOS transistors is eliminated. Fig. 70(b) shows the bias voltage generator for the n-well of the PMOS transistor M_1 , which has been designed to prevent latch up during the initial power-up transient. Reliability is not a concern here since 5-volt-capable technology is

used at 3.3V



(a)



(b)

FIGURE 70. (a) A high voltage generator for switches and (b) a bias voltage generator for the well of M_1 to prevent latch up.

CHAPTER 6

Experimental Prototype and Measurement Results

6.1 Prototype

6.1.1 Floor Plan

A prototype A/D converter based on the above architecture was fabricated in a double-poly double-metal 1.2- μm CMOS technology. It consists of 8 pipelined stages and one flash A/D section in the end. Capacitors in the first three stages are calibrated with

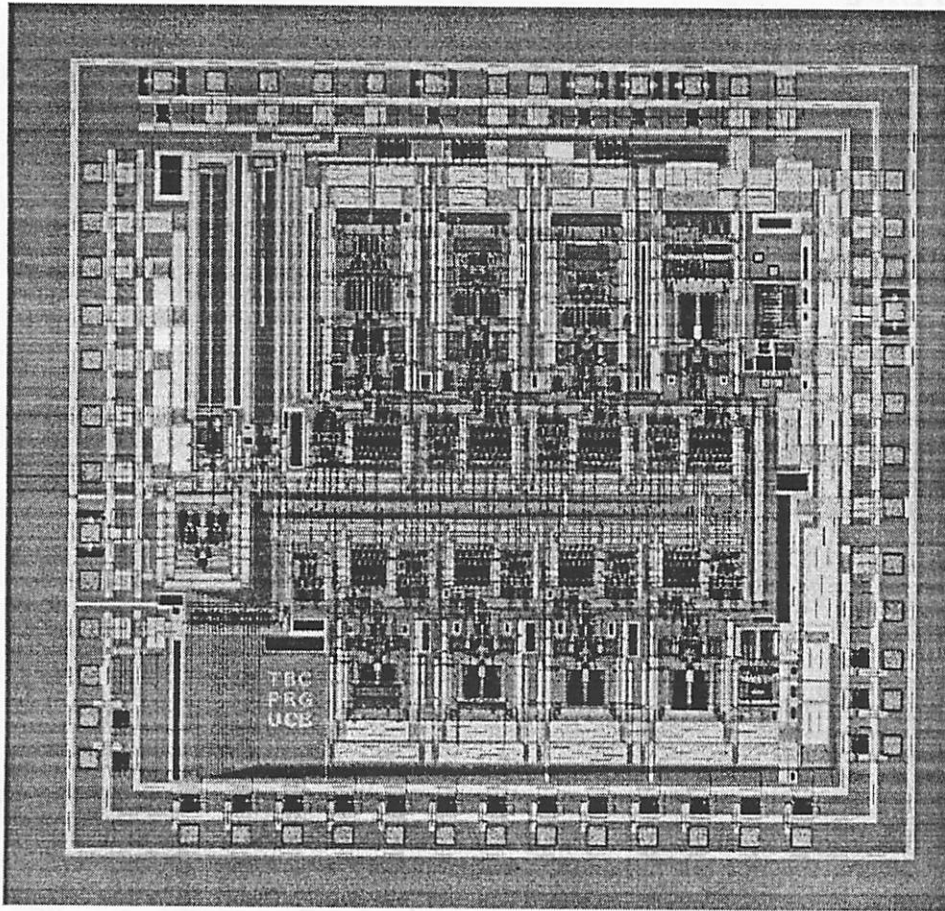


FIGURE 71. A die phot

trim capacitor arrays to achieve high accuracy. A die photo is shown in Fig. 71. Clock lines are routed in the middle, and the analog signal path is folded around to make the chip area square. Op amp bias circuits are shared between several op amps, and all bias currents are controlled by one external master bias current. Chip area not including the pad ring is $3.2 \times 3.3 \text{ mm}^2$.

6.1.2 Op Amp Bias Circuit and Clock Generation Circuit

In the prototype, one external master bias current produces several reference currents by current mirrors, and they are distributed to op amp bias circuits as shown in

Fig. 72. Currents are distributed instead of the gate voltage(V_{gs}) of the master current

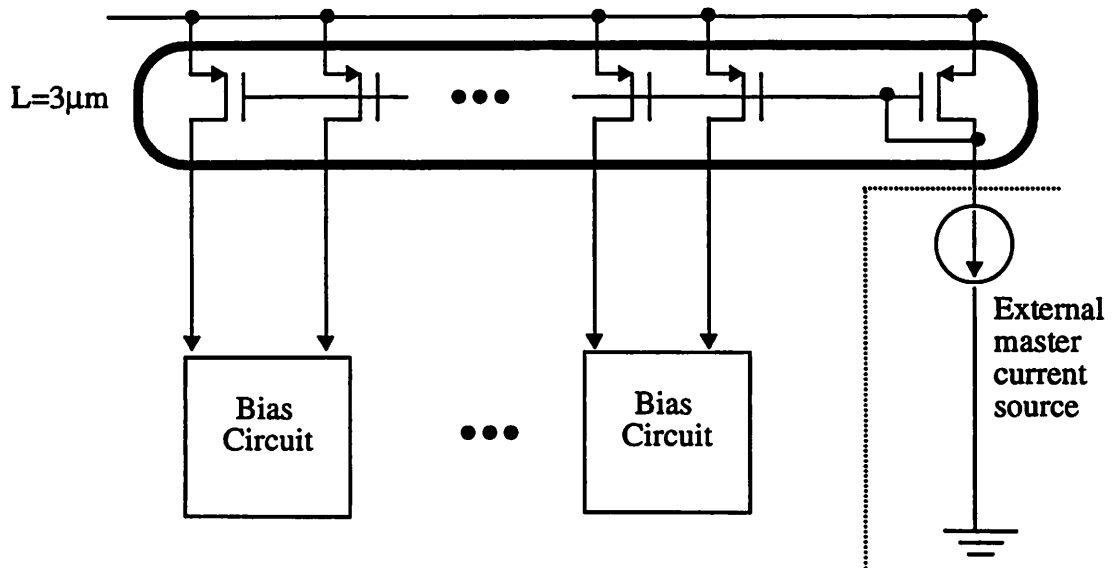


FIGURE 72. Distribution of bias currents from the external master current source.

source. This is because the finite resistance of the supply line can produce small voltage drop ($V=IR$) which can change the V_{gs} of the current source and in turn can alter the current value, if the gate voltage is distributed. For instance, if the voltage drop is 10mV and $V_{gs}-V_{th}$ of the current source is 100mV, then, the difference can cause about 20% change in current if square law device characteristic is assumed. If currents are distributed instead, this problem can essentially be eliminated by generating necessary gate voltages locally from the reference current, and the voltage drop can be minimized.

Fig. 73 shows the bias circuits used for op amps. Cascode bias voltage is generated

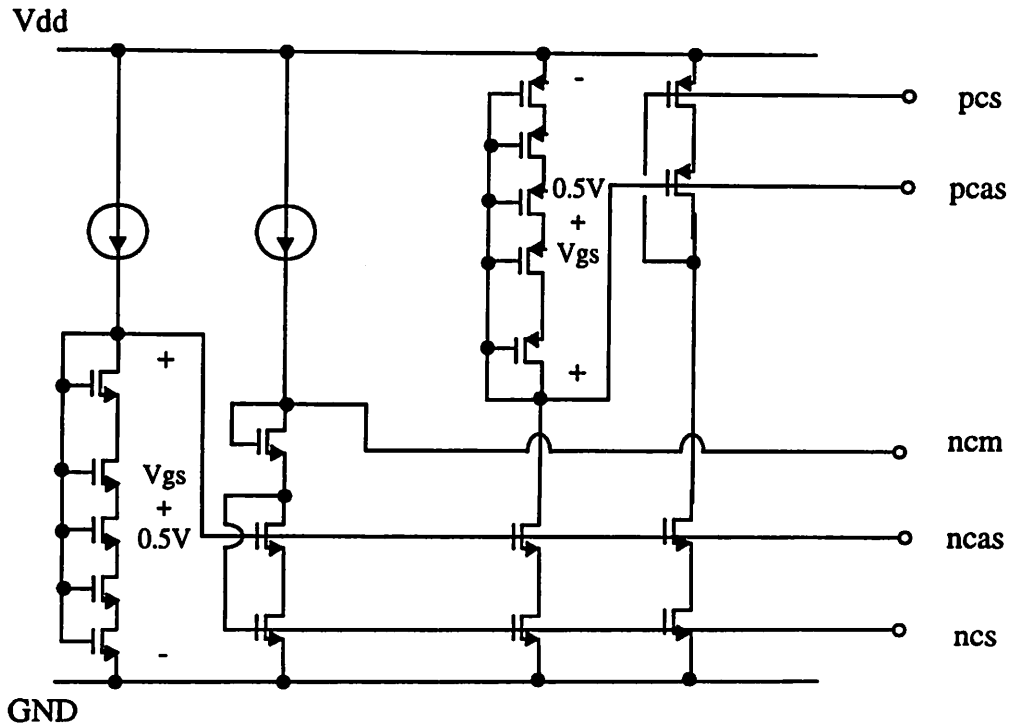


FIGURE 73. Op amp bias circuit

by using multiple transistors biased in triode region. Then, PMOS bias voltages are generated in slave to the NMOS side for better current matching and controllability. The bias voltage “ncm” is for the load device of the pre-amp stage in the op amp.

In the prototype, non-overlapping clock is required for reliable pipeline operations,

and required waveforms are shown in Fig. 74. Early and late clocks are needed for

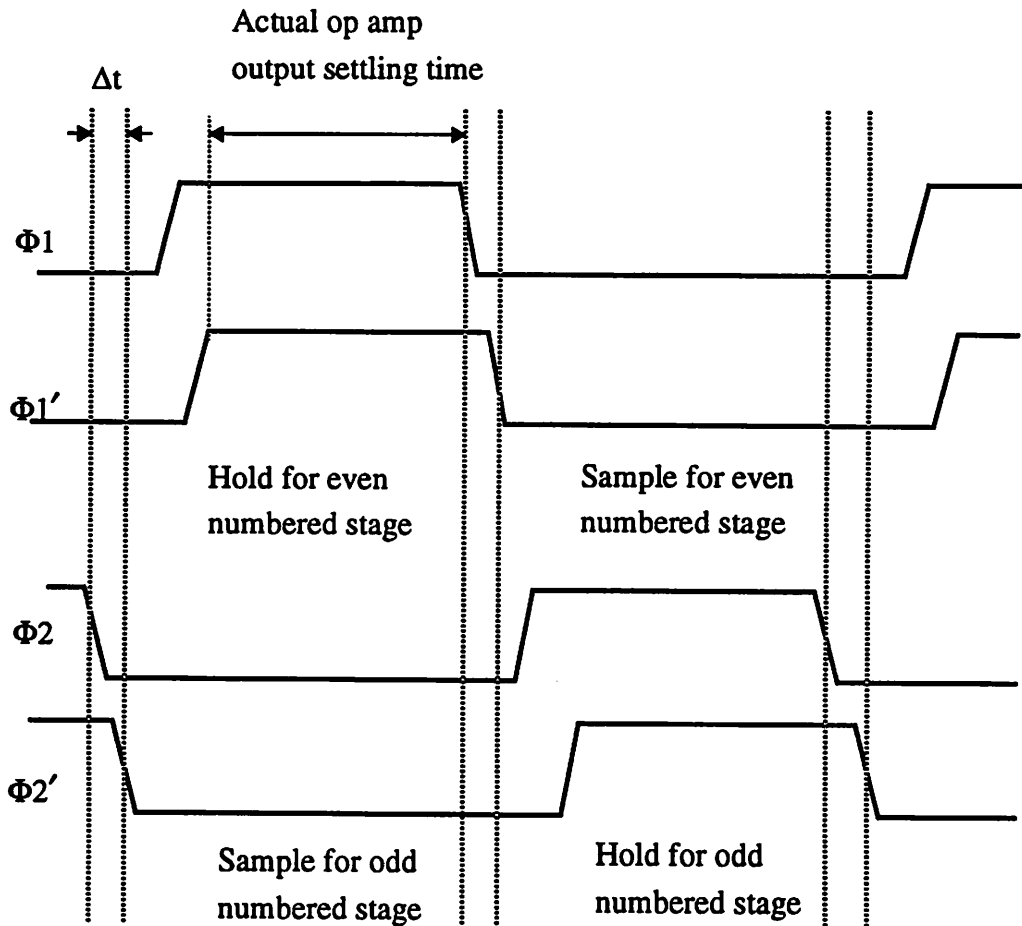


FIGURE 74. Clock waveforms for the pipeline operation.

sampling instance and other switching operations respectively, so that other transmission gates and digital switching noise does not couple into the accurate sampling operation. Due to these non-overlapping and delayed clock schemes, various clock rise/fall times add and reduces actual setting time for the interstage amplifier output. Especially, for the given $1.2 \mu\text{m}$ technology optimized for 5V operation, 3.3V operation increases the rise/fall

times and reduces available setting time even further.

Fig. 75 shows modified waveforms to allow more time for opamp settling. By using this clock scheme, extra 3-5 nsec is added to the opamp settling time.

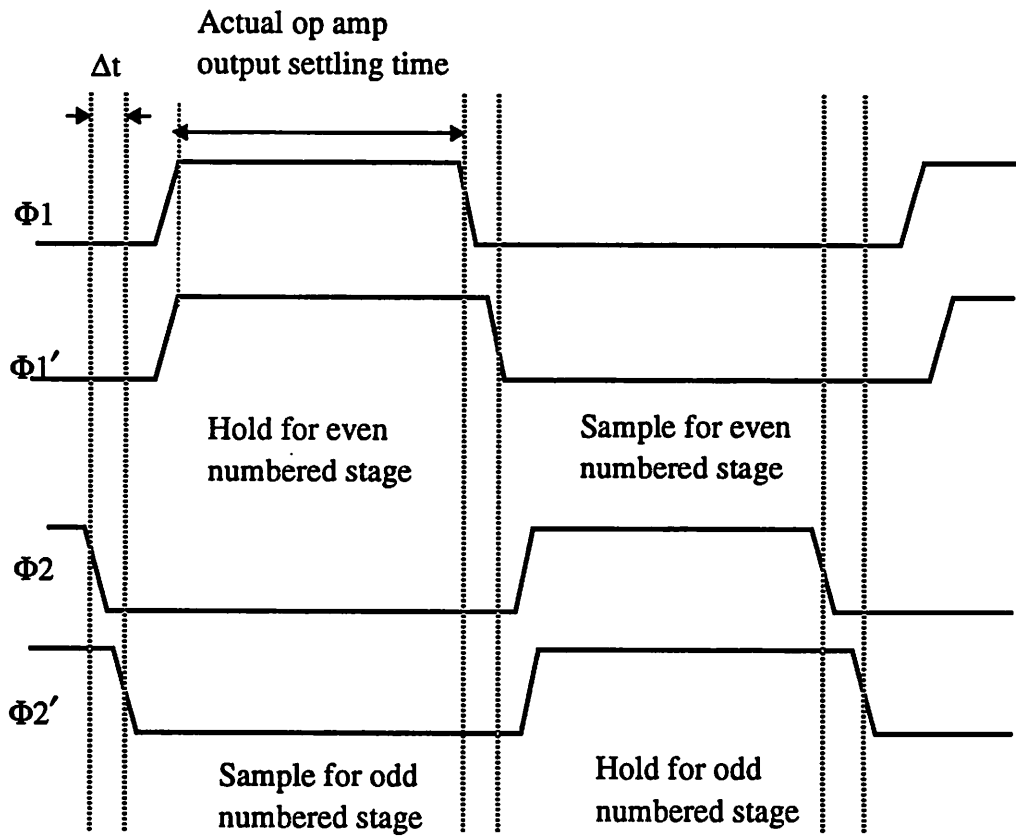


FIGURE 75. Modified clock waveforms to line up the rising edge of the early/late clocks

The new waveforms can be generated with circuits shown in Fig. 76. Parts of the

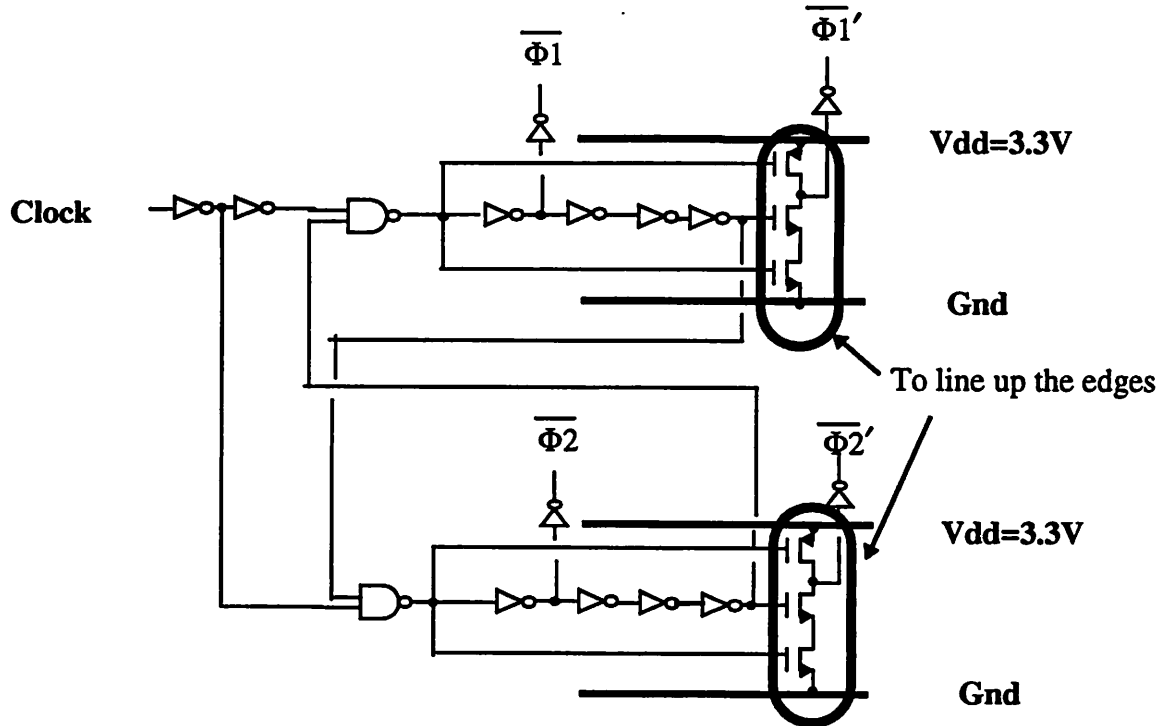


FIGURE 76. Clock generator circuits

inverter delay chains are modified so that the early and late clocks can be lined up together at the rising edge. This is done by bypassing the delay chains with transistors indicated with arrows to reset the internal node and in turn to line up the rising edges of the clocks. Falling edges are then created by the delay inverter chain action. Inverter clock waveforms are again inverted by the high voltage generator circuit shown in section 5.4.2 to create waveforms in Fig. 75.

6.2 Measurement Result

Fig. 77 shows SNDR vs. the input amplitude for 100 kHz and 10 MHz input frequencies at 20-MS/s conversion rate. The peak SNDR is 59.1 dB for 100 kHz input sine wave. At Nyquist sampling (10MHz input), the SNDR is 55.0 dB. Degradation in SNDR

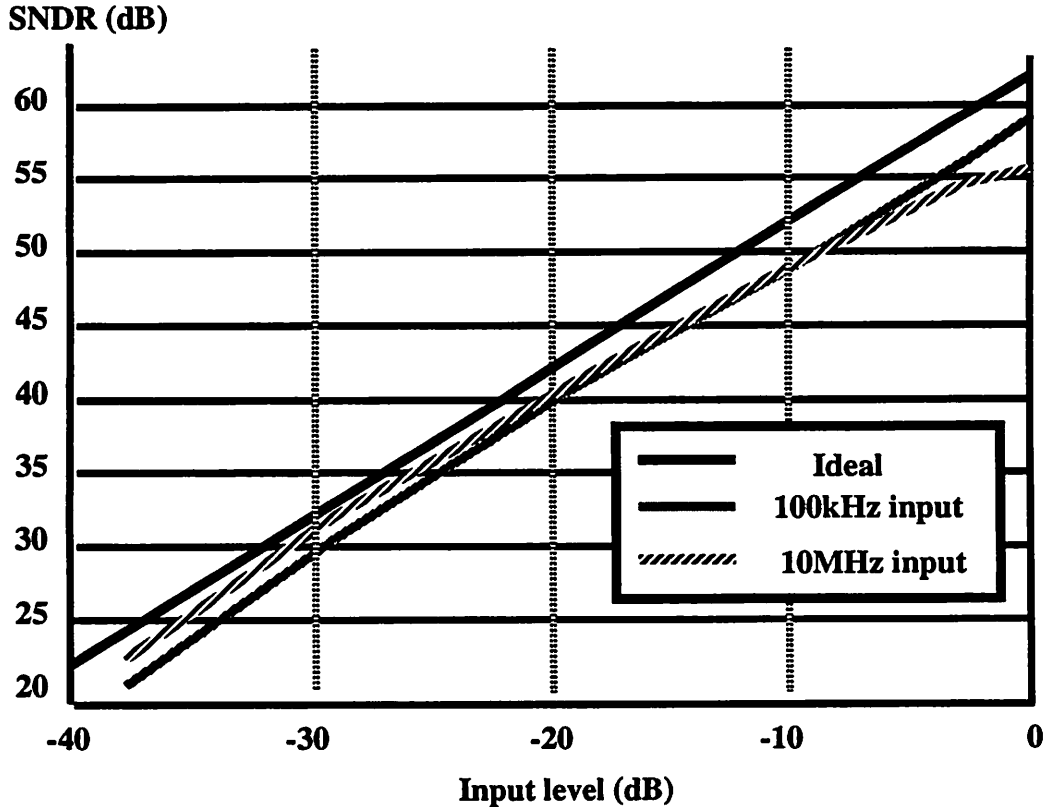


FIGURE 77. SNDR vs. the input amplitude for 100kHz and 10MHz input frequency at 20MS/s conversion rate.

at high input frequency is due to the fact that the input signal is sampled both on the first stage interstage amplifier and the flash ADC with some time difference in order to eliminate the dedicated S/H circuit for low power as described in section 5.2. In the

prototype, the time delay is set by 4 CMOS inverter delay and is about $\sim 3\text{-}4\text{ nsec}$ with 3.3V supply. According to (EQ. 45), this corresponds to the input bandwidth of around 10MHz, which agrees with the measurement result. If the time delay is reduced by having only 2 CMOS inverter delay, then a 2x improvement in its input bandwidth is expected.

In Fig. 78, differential nonlinearity (DNL) and integral nonlinearity (INL) vs. input code are plotted. The magnitude of the maximum DNL and INL are 0.5 LSB and 0.6 LSB, respectively.

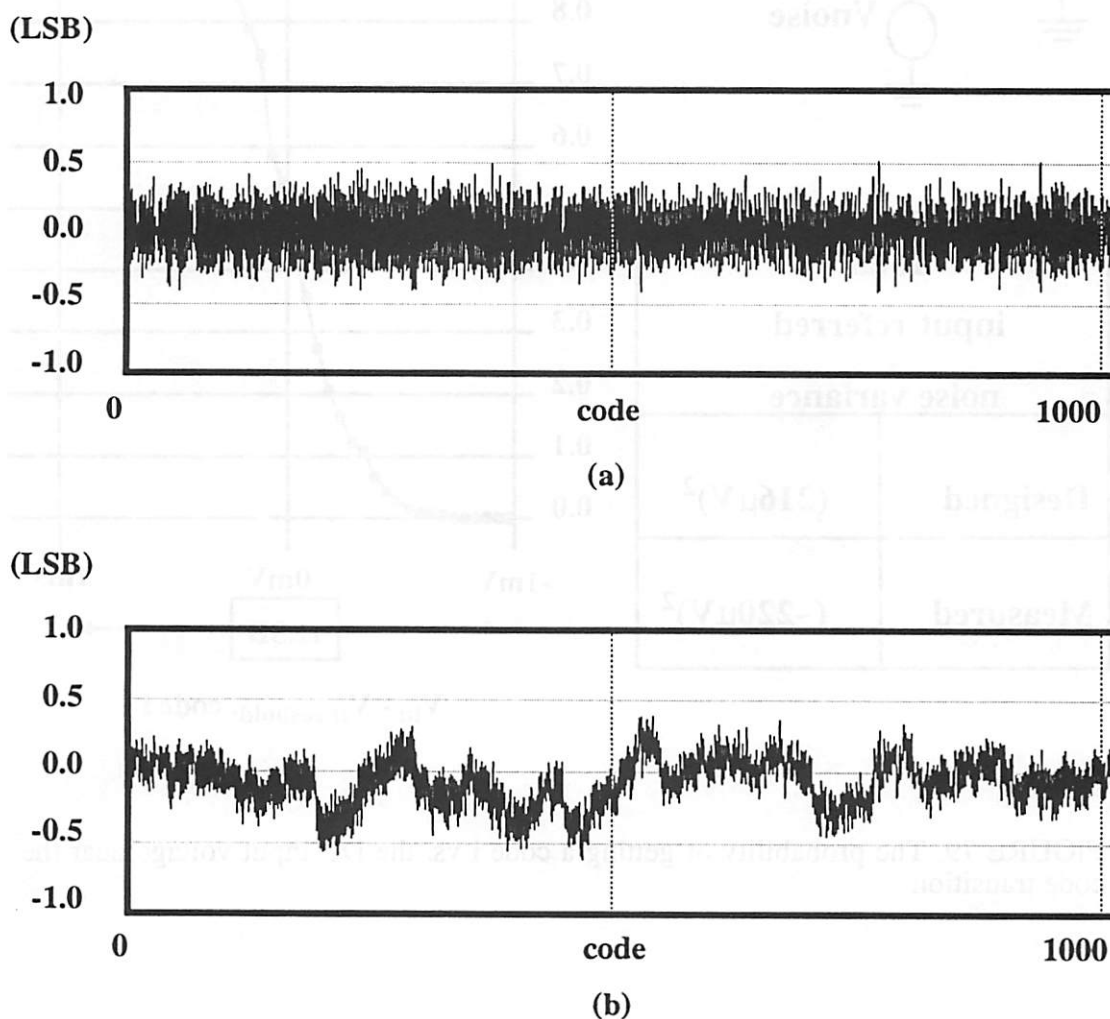


FIGURE 78. (a) Differential non-linearity error. (b) Integral non-linearity error.

Fig. 79 shows the probability of getting a code i vs. the DC input voltage near the code transition. The extracted total input-referred RMS noise voltage from this plot was $\sim 220\mu\text{V}$ while the designed value was $216\mu\text{V}$. This confirms the kT/C noise-limited design in the prototype.

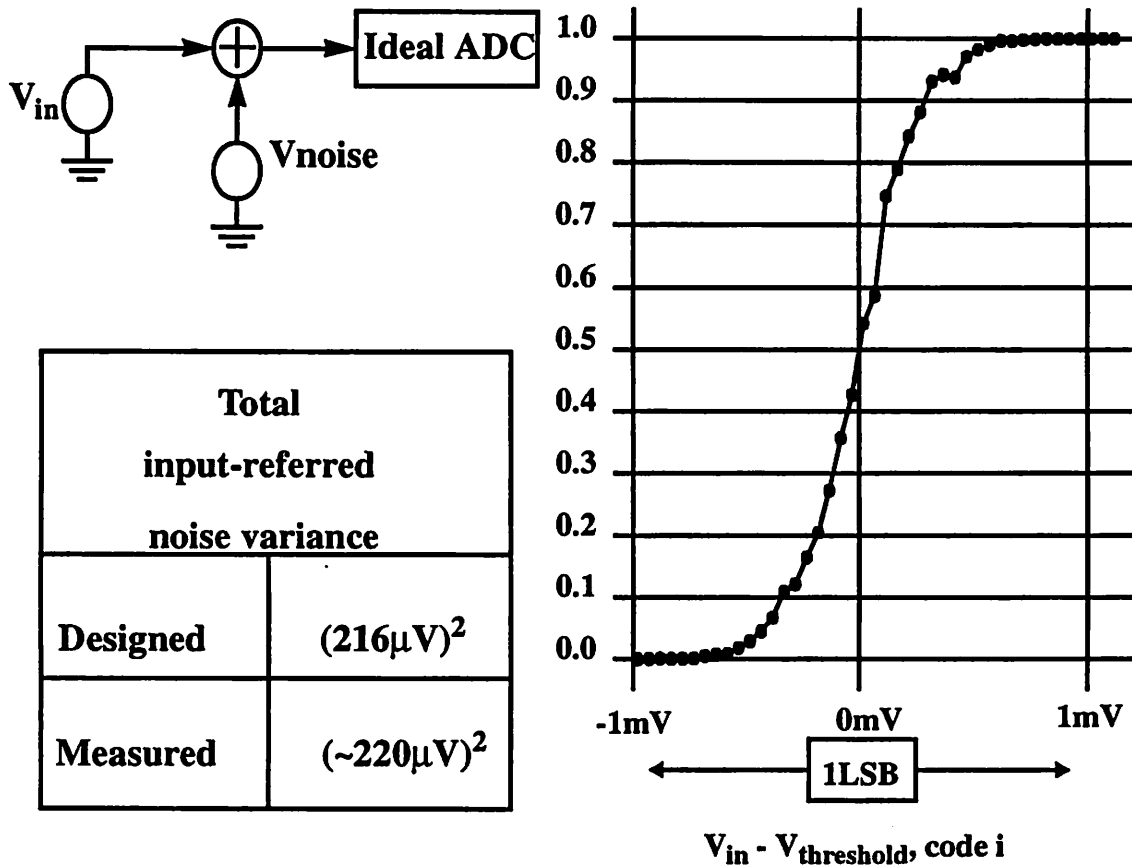


FIGURE 79. The probability of getting a code i vs. the DC input voltage near the code transition

Fig. 80 shows the measured power consumption vs. the sampling frequency on a log-log scale. Of 35mW of total power dissipation at 20MS/s , static power consumption was about 20mW . At a reduced bias current and a sampling frequency of 1MS/s , the

power consumption was 2.8 mW with peak SNDR of 58 dB.

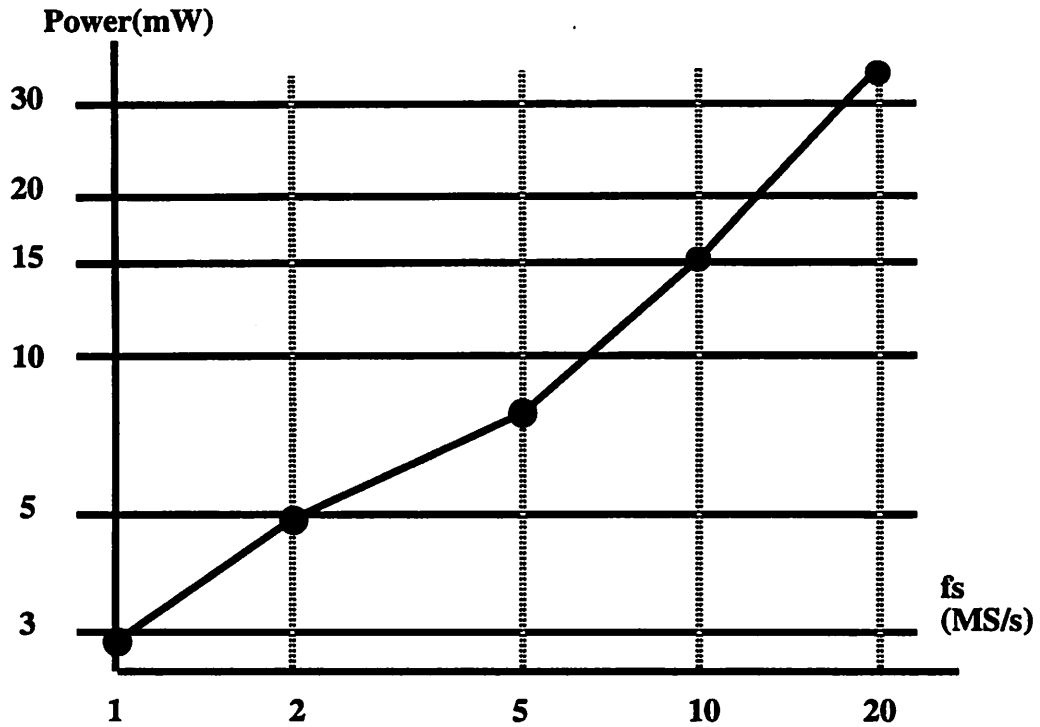


FIGURE 80. Measured power consumption vs. the sampling frequency

6.3 Power Comparison

One way to compare the power dissipation between ADC's is to compare the power dissipation normalized to its sampling frequency for different channel length as

shown in Fig. 81. Only 10 bit CMOS A/D converters are considered for the comparison in

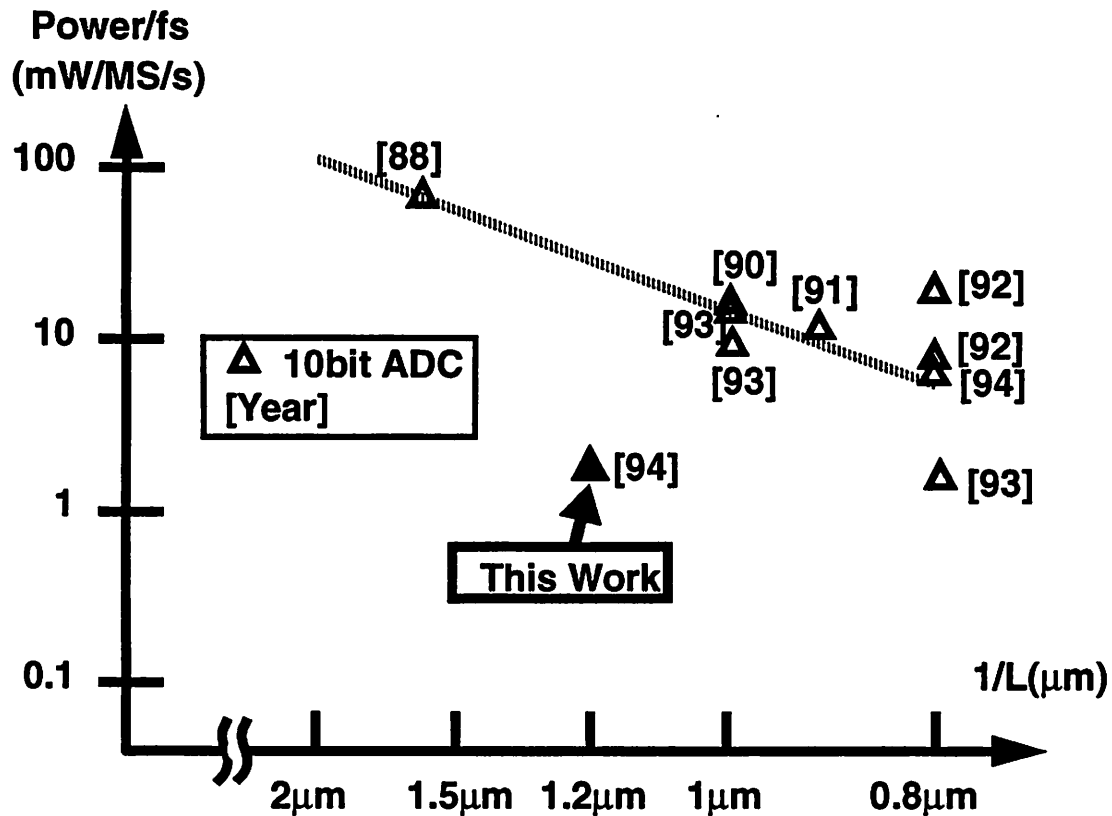


FIGURE 81. Power normalized to the sampling frequency vs. the channel length

this section. The data points shown in Fig. 81 suggest that the power dissipation of ADC's improves with advanced faster technologies, and even with 1.2 μm technology, the prototype ADC dissipates much less power than other ADC's. This confirms the effectiveness of the low power techniques used in the prototype.

CHAPTER 7

Conclusion

In this thesis, fundamental limitations to the power dissipation of key functions for high speed A/D converters are examined. Key functions are sampling, quantization, and reference generation, and required power dissipation in each case is dictated by accuracy consideration. The summary is as follows:

1. Theoretical limit on the power dissipation of the sampling in MOS technology is set by kT/C noise. Its power dissipation is purely dynamic and independent of the supply voltage for given sampling rate and resolution. In reality, however, the static power dissipation from active circuitries such as op amps in SC S/H circuits is three or four orders of magnitude higher power dissipation than the theoretical CV^2 limit. However, the power dissipation is still proportional to the size of the sampling capacitor dictated by the kT/C consideration.
2. Power dissipation associated with the quantization function is determined by accuracy requirements, which is fundamentally set by kT/C noise consideration. In reality, a broadband pre-amplifier(s) is required to minimize errors due to offsets and charge injection, and its static power dissipation again dominates over the dynamic power set by kT/C consideration.
3. Two common methods of generating reference levels are using R-DAC and C-DAC. The amount of allowable thermal noise again determines the power dissipation in each

DAC, and C-DAC usually achieves less power dissipation than R-DAC, because no DC bias current is required. Although theoretical analysis shows that C-DAC achieves smaller power than that of R-DAC by only a factor of 2 or 3, practical issues such as available resistor in the technology, DAC settling, matching, etc., makes this difference much larger.

4. Power dissipations of several ADC architectures are examined. From a “power factor” standpoint, a pipeline architecture presents itself as a potential candidate for low power consumption due to pipeline scaling and digital correction which allow the design of each stage close to kT/C limited minimum size.

5. In the experimental prototype, it has been demonstrated that digital correction and pipeline scaling techniques which allow the use of dynamic comparators and SC circuit scaling down the pipeline are effective in reducing the power dissipation. Also, for low voltage operation, 3.3V high-speed op amp and high voltage generators designed. This shows that low voltage, low power operation of pipeline A/D converters can be effectively achieved for video-rate applications.

Based on the analysis and experimental results given above, the following projection can be made for further reduction of power dissipation in CMOS pipeline A/D converters.

1. A new high speed op amp topology with larger output signal swing is needed in order to further reduce the sampling capacitor size allowing more kT/C noise in the system. Common source stages at the output is suggested for this, and op amp gain requirement may require an extra gain stage since cascoded output stages cannot be used, with possible reduction of its bandwidth from multi-stage configuration. Therefore, designing an op amp with both large swing and large bandwidth will be very challenging.

2. High speed class AB or B type operational amplifiers can be designed to reduce the static power dissipation. This may require an advanced scaled CMOS technology to

increase the bandwidth of the class AB or B type amplifiers.

3. Parallel ADC's can be implemented with auto-calibration techniques to eliminate path matching problems. For approximately square law MOS devices, parallelism can theoretically save power by factor of 2 at the expense of 2x area increase excluding ancillary circuit power such as clock driver power, etc.

4. A scaled implementation of the experimental prototype described in the thesis with an advanced submicron technology will again achieve substantial reduction of its power

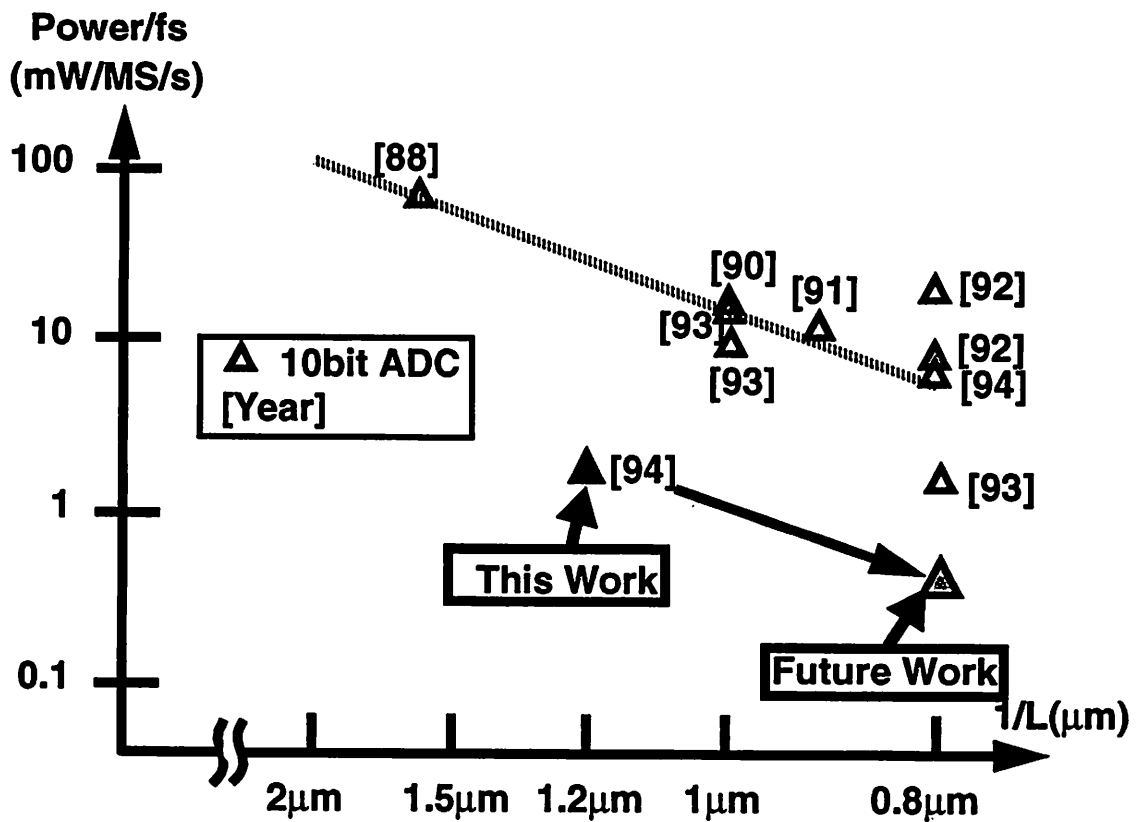


FIGURE 82. Projection of the power dissipation of the prototype ADC with a scaled technology.

dissipation. Projection from Fig. 82 predicts another ~2-3x reduction in power dissipation with 0.6 or 0.8 μm technology with a 3.3V supply voltage.

For low voltage operations, the following practical issues must be overcome.

1. A new low voltage op amp configuration is needed for lower supply voltages ranging from 1.5 to 2.8V. Especially for scaled submicron($< 0.5 \mu\text{m}$) technologies in the future, the supply voltage will be chosen by the thin gate oxide reliability issue and compatibility issue with low power/voltage DSP.
2. Efficient high voltage generator circuits or some means of providing low on-resistance transmission gates for SC circuits are also critical. Again, oxide reliability issue must be carefully considered for future thin oxide sub-micron devices.
3. Techniques to eliminate or minimize noise-coupling from DSP section on the same chip are necessary. Current trend is to include mixed signal circuits as a part of a larger digital sections, and minimizing noise coupling from digital circuit will be critical especially when signal swing is small with low supply voltage.

Appendix 1

Error Sources in MOS S/H Circuits

Three error sources in MOS S/H circuit are listed in Table 1. First one is the finite bandwidth. The transistor on-resistance and the sampling capacitance form a RC-network and set the signal bandwidth as illustrated in Fig. 83. The on-resistance of the MOS

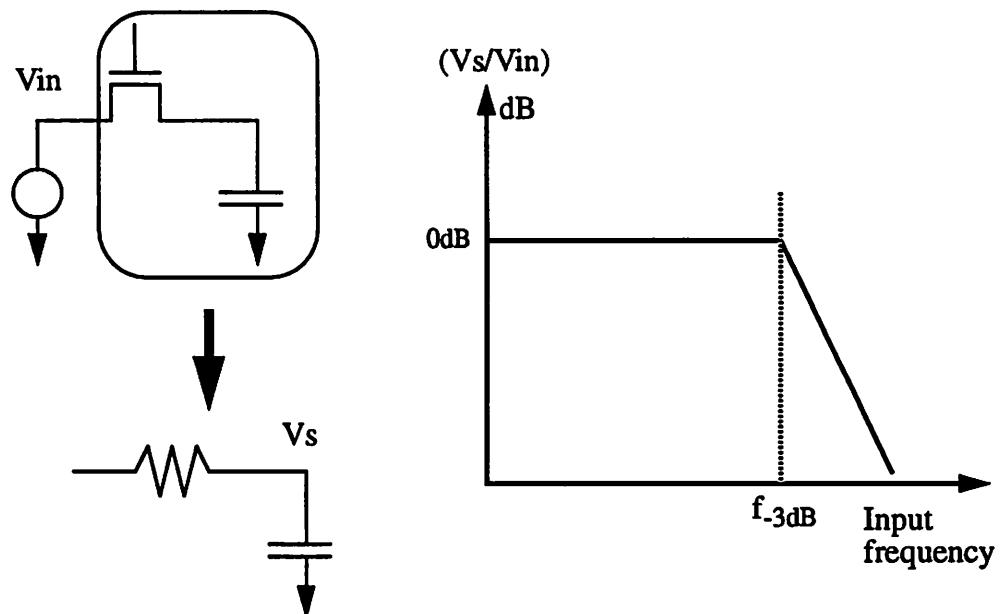


FIGURE 83. Input bandwidth of a simple MOS S/H circuit.

transistor in triode region is $1 / (k_p \cdot \frac{W}{L} \cdot (V_{gs} - V_{th}))$ if V_{ds} is small. Then the -3dB frequency is given by

$$f_{-3dB} = \frac{1}{2\pi} \cdot \frac{1}{RC} = \frac{1}{2\pi} \cdot \frac{k_p \cdot \frac{W}{L} \cdot (V_{gs} - V_{th})}{C_S} \quad (\text{EQ 50})$$

Assuming the worst-case $(V_{gs}-V_{th})$ of 1V, (W/L) of 10 and k_p of $60\mu\text{A/V}$, the sampling capacitor value of 1pF gives -3dB frequency of about 95MHz. For the given sampling capacitor size and $(V_{gs}-V_{th})$, the bandwidth of the device will increase if advance technology is used to reduce the gate channel length, L . Another way to increase the input bandwidth is to use a larger supply voltage in order to increase $(V_{gs}-V_{th})$. However, the trend with the advanced technology is to use lower and lower voltage for reliable operations of thin gate-oxide MOS devices and it sets the limit for the maximum allowed supply voltage. Therefore, it is not likely that the input bandwidth performance will improve at the same rate that the channel length is reduced, since the bandwidth depends on the product of $(k_p \cdot \frac{W}{L} \cdot (V_{gs} - V_{th}))$ for the given sampling capacitor value.

The second one is the charge injection error. When a MOS transistor is used as a switch as shown in Fig. 84, there is a finite amount of charge in the conducting channel, whose magnitude is approximately $C_{ox}(V_{gs}-V_{th})$. When input signal is sampled on a capacitor by turning off the transistor, this charge is pushed out from the channel to either direction, and part of it is dumped on C_S causing an error voltage of $\Delta Q/C_S$. The magnitude of ΔQ is a complex function of the falling time of the sampling clock edge and impedance level at drain/source, and in the fast switching-off conditions, the transistor conduction disappears quickly and ΔQ approaches 50% of $C_{ox}(V_{gs}-V_{th})$ [30].

This charge injection error can produce a large error in the S/H circuit because the magnitude of the error voltage is signal dependant(ΔQ is proportional to $C_{ox}(V_g - V_{in} -$

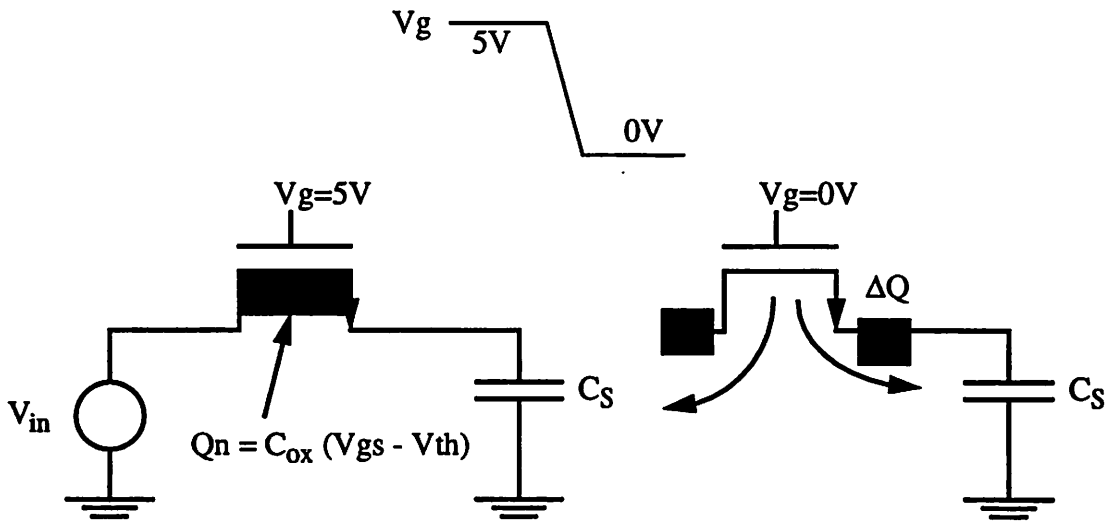


FIGURE 84. Charge in the conducting channel of MOS device.

V_{th}). If the input signal is close to V_g , then less charge is in the conducting channel and less error voltage is observed. However, when V_{in} is much less than V_g , the amount of charge in the channel increases and so the error voltage.

One method to reduce this charge injection error is to use a dummy switch as illustrated in Fig. 85. By absorbing the dumped charge with a half size transistor assuming

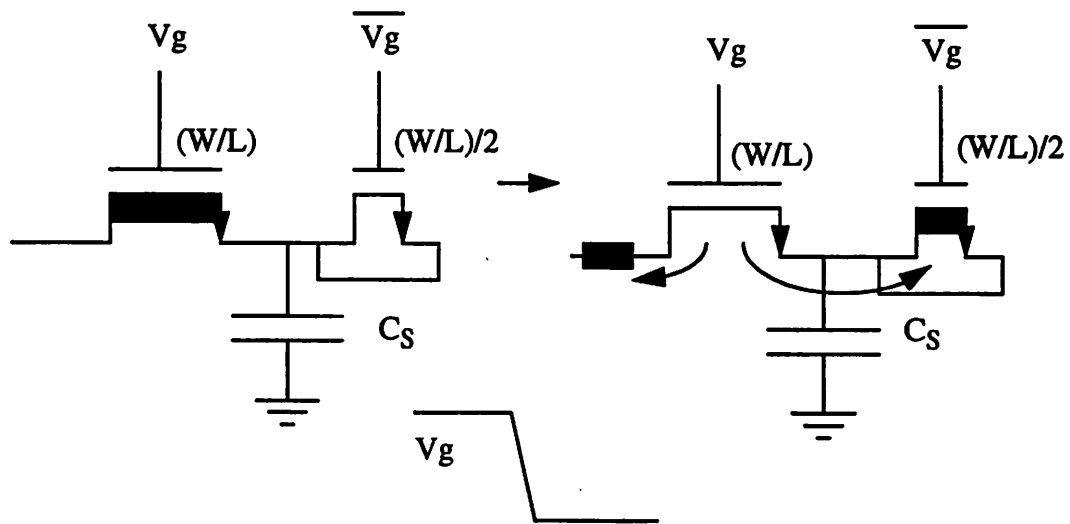


FIGURE 85. Charge injection cancellation technique using a half size dummy switch.

50% charge split in either direction, the charge injection error can be cancelled to a first order[45]. However, this technique relies on the matching of the transistor ratio, and 50% charge splitting ratio in fast switching-off condition, and the cancellation effect is reduced if the signal is driven from an external source whose impedance level is low and charge splitting becomes a function of the source impedance[31].

Another technique to eliminate the charge injection error is the bottom plate sampling scheme shown in Fig. 86. Two switches are used and the signal is sampled when

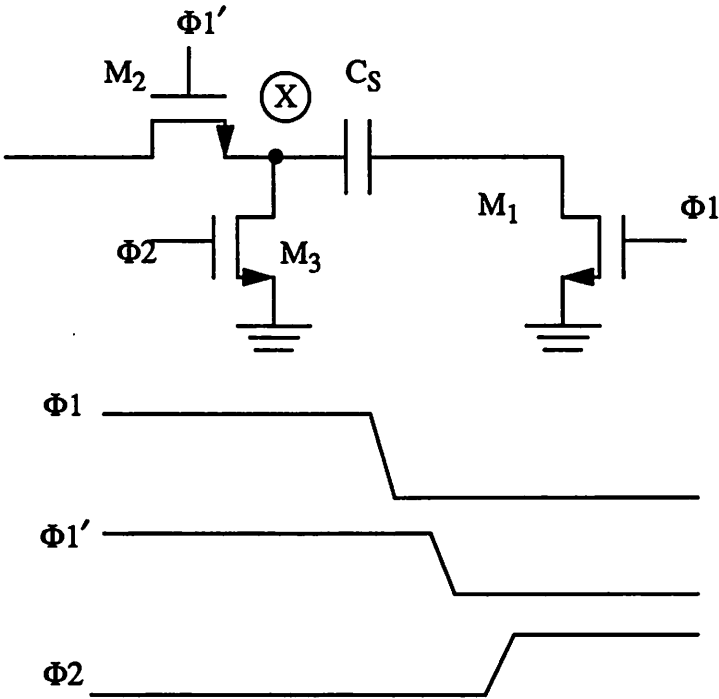


FIGURE 86. A bottom-plate sampling technique

M_1 turns off. Since the drain and source of M_1 are always at the same potential at every sampling instance, in this case at ground, the charge injection to the sampling capacitor will be constant all the time and its signal dependent characteristic can be eliminated. Charge injection from M_2 does not affect the sampled signal charge, since the injected

charge from M_2 will be shorted out by M_3 in the next clock phase.

Lastly, an error is caused by clock feedthrough as illustrated in Fig. 87. When M_1

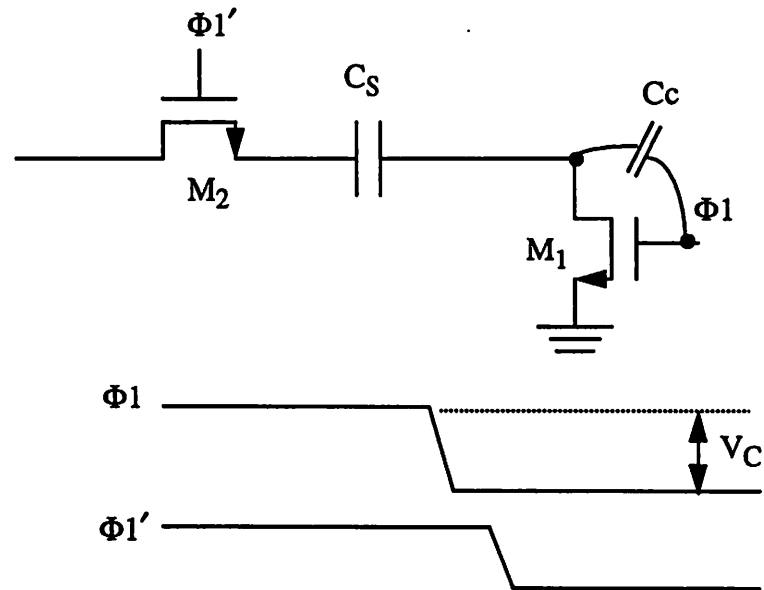


FIGURE 87. Clock feedthrough from M_1 .

turn off, the clock signal is coupled to the sampling capacitor through the gate-to-drain parasitic capacitance(C_c), and effectively causes an error voltage proportional to V_C and C_c . To the first order, this problem can be solved by using differential scheme as shown in Fig. 88. By having two identical stages and sampling the input signal differentially, the clock feedthrough can be cancelled if its magnitude on both sides are the same. This implementation can also cancel the charge injection error by the same mechanism. The level of cancellation by this technique depends on the matching between two stages as in the dummy transistor case, but at least the signal dependent error can be eliminated. In many reported ADC's([1][2][5][8][9][11][12][13] etc.), this scheme is widely used. Detailed performance limitations are described in [1][10].

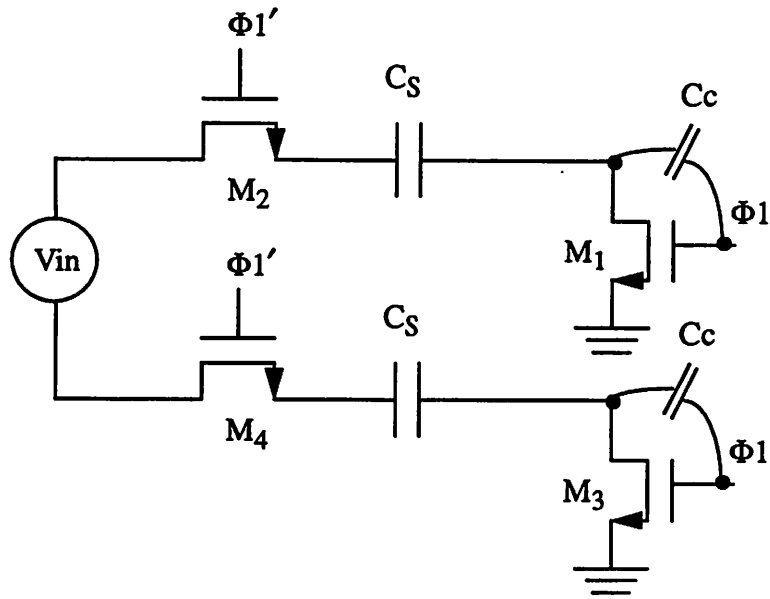


FIGURE 88. A differential bottom plate sampling network for clock feedthrough and charge injection cancellation.

Appendix 2

kT/C Noise Calculation in SC Circuit

1. kT/C noise calculation

Fig. 89 shows the detailed configurations for the switches in a SC circuit. During the input sampling process, kT/C noise is sampled on the capacitors along with the input signal(Fig. 89(a)(b)), and its magnitude is given by:

$$\overline{V^2} = \frac{kT}{C_S + C_F + C_{opamp}}. \quad (\text{EQ 51})$$

Then, in Fig. 89(c)(d), the sampled signal and noise charge gets transferred to C_F . Assuming $V_{in}=0$, the total noise charge is:

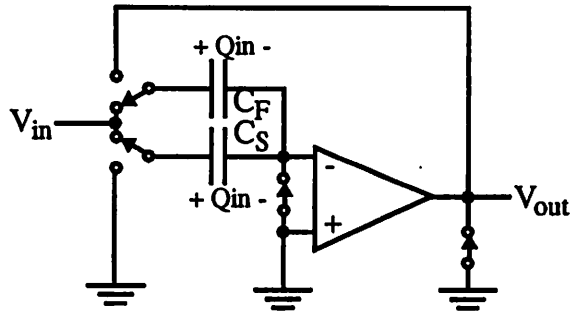
$$\overline{Q_n^2} = (C \cdot V)^2 = kT(C_S + C_F + C_{opamp}) \quad (\text{EQ 52})$$

and will cause an output voltage of

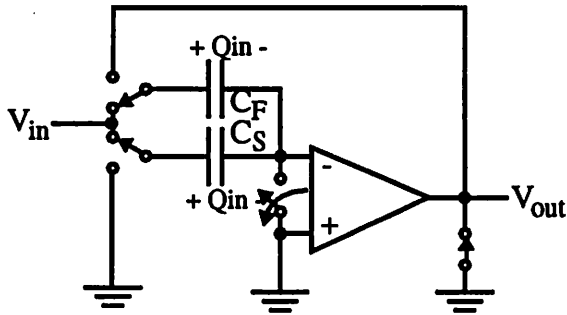
$$\overline{V_{out}^2} = \frac{\overline{Q_n^2}}{C_F^2} = kT \cdot \frac{(C_S + C_F + C_{opamp})}{C_F^2} = \frac{kT}{C_F} \cdot \frac{1}{f} \quad (\text{EQ 53})$$

where C_{opamp} is the input capacitance of the op amp.

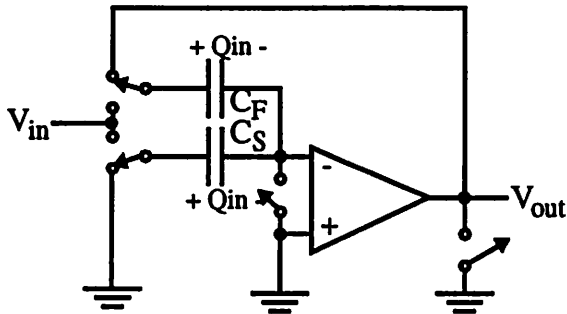
For differential implementation of the circuit, the noise power in (EQ. 53)



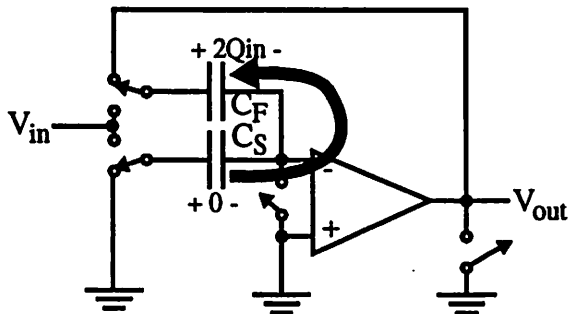
(a) C_F and C_S are connected to input voltage source.



(b) Both signal and kT/C noise are sampled.



(c) Other switches are opened.



(d) Charge transfers to C_F

FIGURE 89. SC switching configurations at different clock phases.

increases by a factor of 2 assuming no correlation between positive side and negative side, since the uncorrelated noise adds in power.

Input referred noise power can be found by dividing the output noise power by the square of the gain and is given by:

$$\overline{V_{in}^2} = \frac{\overline{V_{out}^2}}{G^2} = \frac{kT}{C_F} \cdot \frac{1}{f} \cdot \left(\frac{C_F}{C_S + C_F} \right)^2 = \frac{kT \cdot (C_S + C_F + C_{opamp})}{(C_S + C_F)^2}. \quad (\text{EQ 54})$$

2. Op amp noise calculation

Op amp noise can be calculated for a SC circuit by injecting a current noise from its generator as shown in Fig. 90. The noise generator in this case is the thermal noise

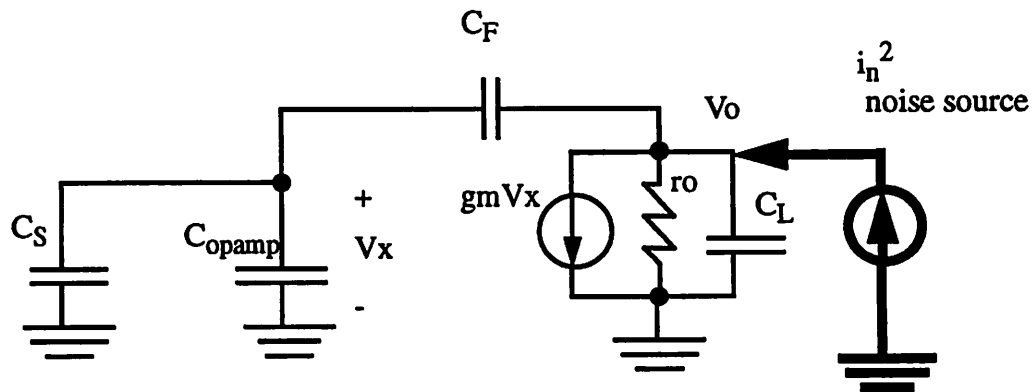


FIGURE 90. AC model for op amp noise calculation.

source from the transistor, whose magnitude is given by

$$\overline{i_n^2} = 4kT \cdot \left(\frac{2}{3} gm \right) \cdot \Delta f [3]. \quad (\text{EQ 55})$$

Then, noise power at the output can be found from

$$H(s) = \frac{V_o}{i_n} = \frac{r_o}{(1 + g_m \cdot r_o \cdot f) \cdot \left(1 + \frac{s C_{LT} r_o}{1 + g_m \cdot r_o \cdot f}\right)}, \quad (\text{EQ 56})$$

$$\overline{V_o^2} = \int_0^{\infty} (|H(s|_{j\omega})|^2 \cdot \overline{i_n^2}) \quad (\text{EQ 57})$$

where f is the feedback factor and C_{LT} is $C_L + f \cdot (C_S + C_{opamp})$, the capacitance loading at the output, giving the input referred noise variance of

$$\sigma^2 = \frac{2}{3} \cdot kT \cdot \frac{1}{f} \cdot \frac{1}{C_{LT}} \cdot \left(\frac{C_F}{C_S + C_F}\right)^2 V^2. \quad (\text{EQ 58})$$

For a single pole case, the noise voltage can be easily found by using the equivalent noise bandwidth concept given in [3] without having to go through above mathematics. However, for multiple pole/zero systems, the noise bandwidth depends on many variables such as relative pole locations, stage gain, etc., $H(s)$ must be found for each current generator present in the system to accurately find the noise voltage. For a two stage amplifier, for instance, the equivalent small signal model and its noise generators are shown in Fig. 91. Treating op noise generators, $i_{n,1}^2$ and $i_{n,2}^2$ as independent noise sources, the following procedure can be repeated to find the input referred noise voltage.

$$H_1(s) = \frac{V_o}{i_{n,1}} \quad (\text{EQ 59})$$

$$H_2(s) = \frac{V_o}{i_{n,2}} \quad (\text{EQ 60})$$

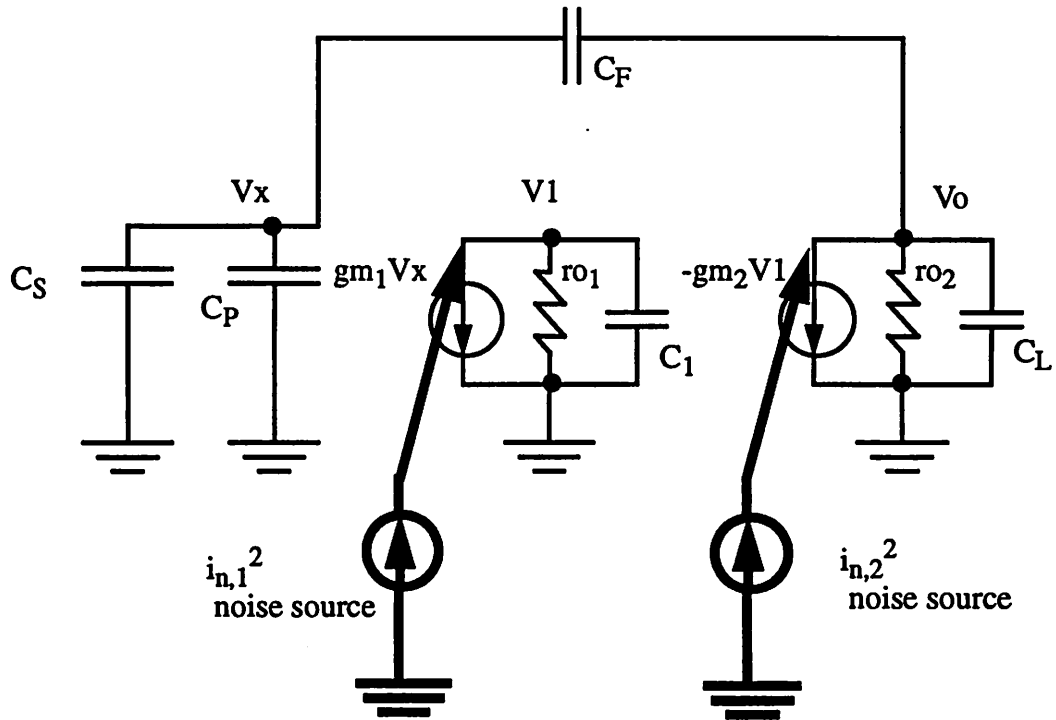


FIGURE 91. Op amp noise calculation method for a two stage amplifier.

$$\bar{V}_o^2 = \sum_{i=1}^2 \int_0^{\infty} (|H_i(s|j\omega)|^2 \cdot \bar{i}_{n,i}) \quad (\text{EQ 61})$$

$$\bar{V}_{in}^2 = \frac{\bar{V}_o^2}{G^2} \quad (\text{EQ 62})$$

where G is the voltage gain of the SC circuit defined in Table 3.

3. Total input-referred noise in (EQ. 19)

Now, kT/C noise and op amp noise can be added together to find the total input referred noise assuming two noise sources are uncorrelated. Using the result from (EQ. 54) and (EQ. 58), the total input referred noise power is given by

(EQ 63)

$$\sigma_{\text{Total}}^2 = \frac{kT \cdot (C_S + C_F + C_{\text{opamp}})}{(C_S + C_F)^2} + kT \cdot \frac{2}{3} \cdot \frac{1}{f} \cdot \frac{1}{C_{\text{LT}}} \cdot \left(\frac{C_F}{C_S + C_F} \right)^2$$

Now, substituting $C_S = C_F = C_L = C$, the assumption used in section 2.3.2, into this equation gives the result in (EQ. 19).

Appendix 3

Digital Error Correction

Multistage ADC configurations require the same input and output signal ranges for cascaded implementation. Therefore, the A/D conversion in each stage must perform subtraction of the quantized bits from the input signal so that the output of the current stage remains within the input range of the next stage. As discussed in section 3.3, offset voltages in the S/H and comparators in the flash ADC section can cause output signal larger than the maximum input range thereby losing the information. This is illustrated in Fig. 92.

In order to prevent over-range problem, the interstage gain of the amplifier can be reduced so that the residue signal with any offset-induced error can still remain within the input range of the next stage as illustrated in Fig. 93. This is a basic idea behind the digital error correction, and variations of this technique have been widely used in multi-stage A/D converters. Input/output characteristics of one such modified digital error correction scheme is shown in Fig. 94. In this case, offsets are purposely added to both flash ADC and DAC outputs in order to simplify the correction logic[7]. With this scheme, only adders are required while the previous one requires both adder and subtracter. Also, the top comparator can be eliminated without altering the transfer function. Further detail can be found in [7][35][10].

In the experimental prototype presented in Chapter 5, the modified digital

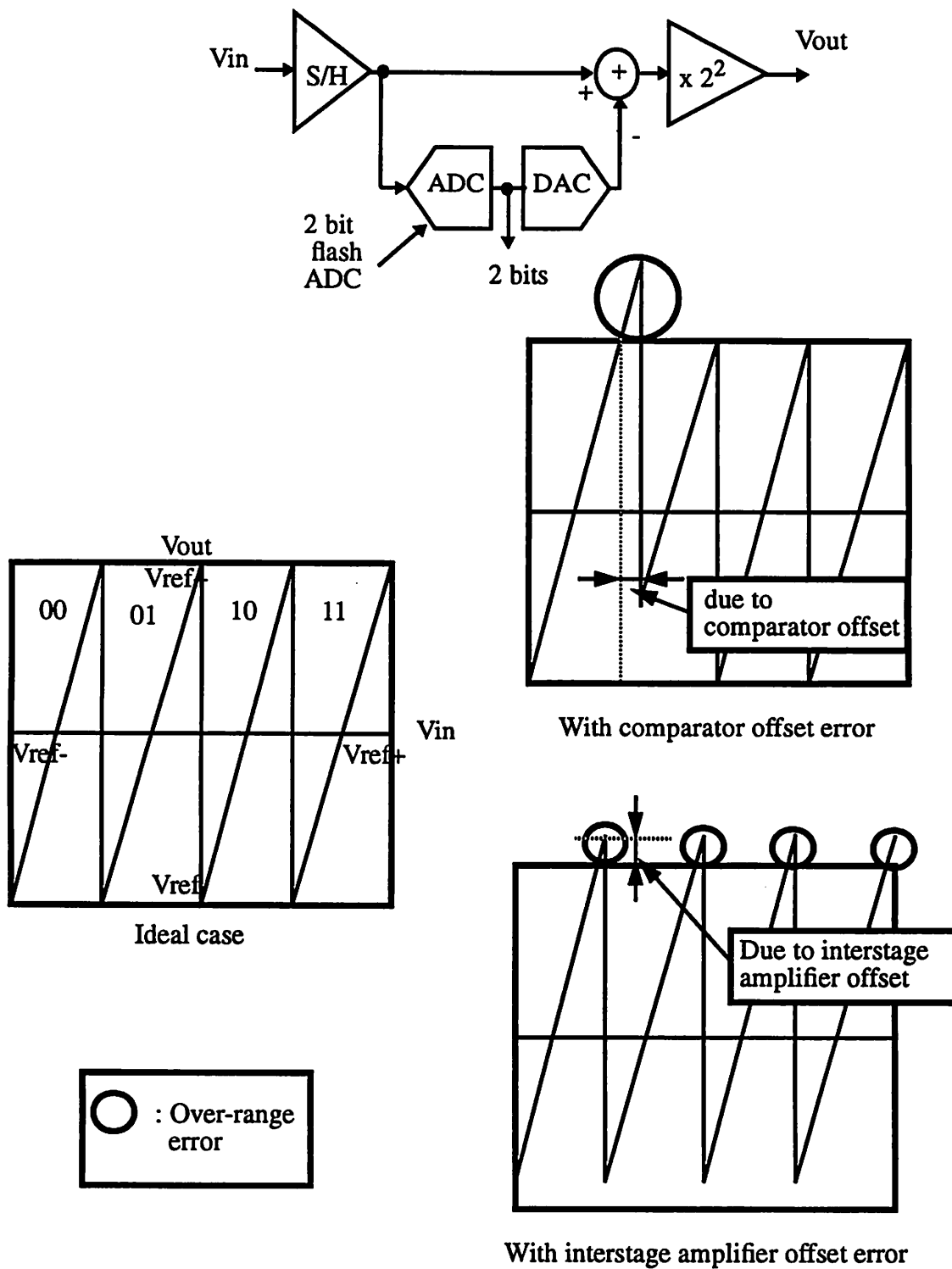


FIGURE 92. A straightforward implementation of a pipeline stage without digital correction, and its input/output transfer curves for ideal case and with offset errors present.

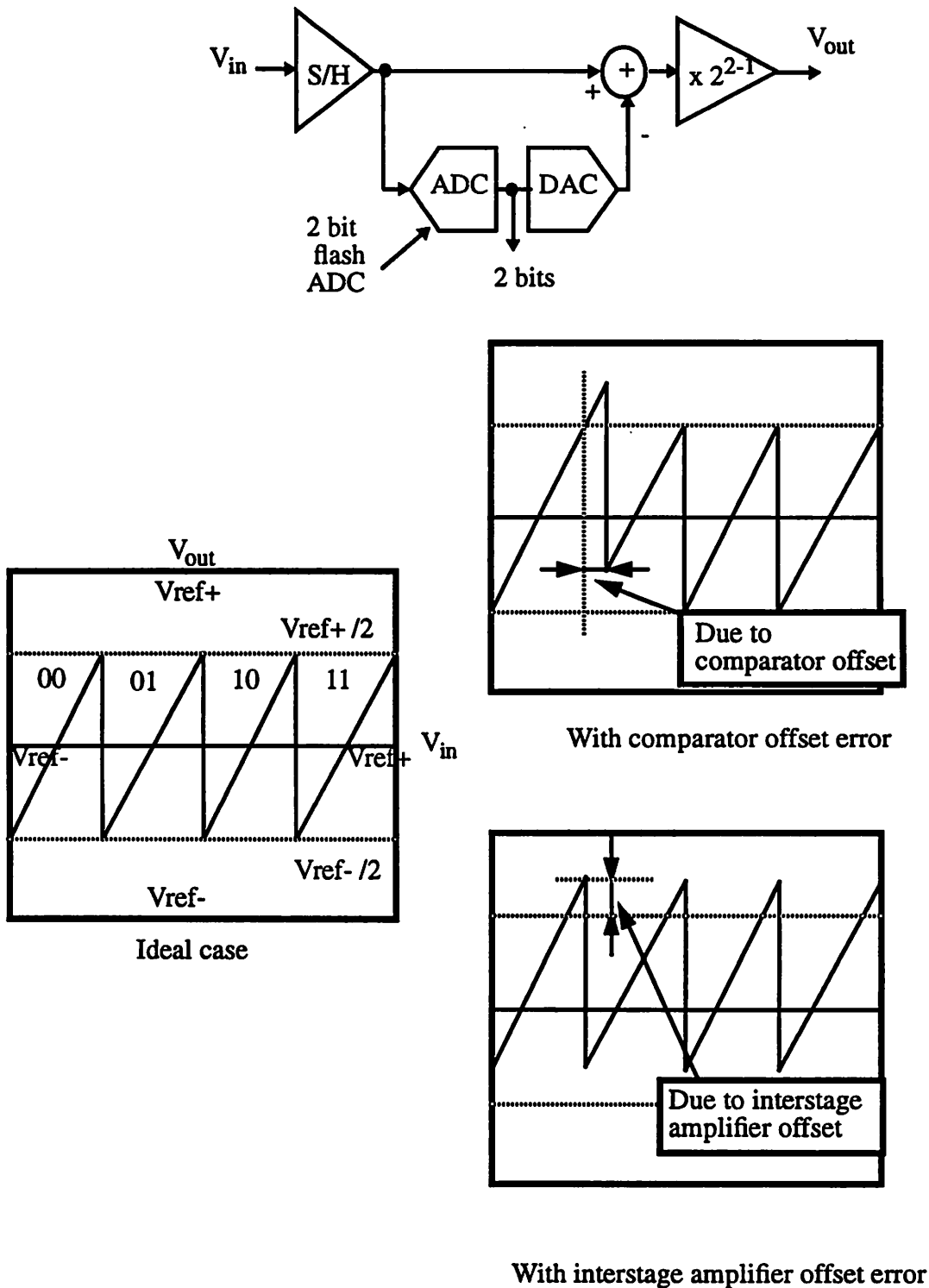


FIGURE 93. An implementation of a pipeline stage with digital correction, and its input/output transfer curves for ideal case and with offset errors present. Notice that the transfer curves with offsets still remain between V_{ref+} and V_{ref-} .

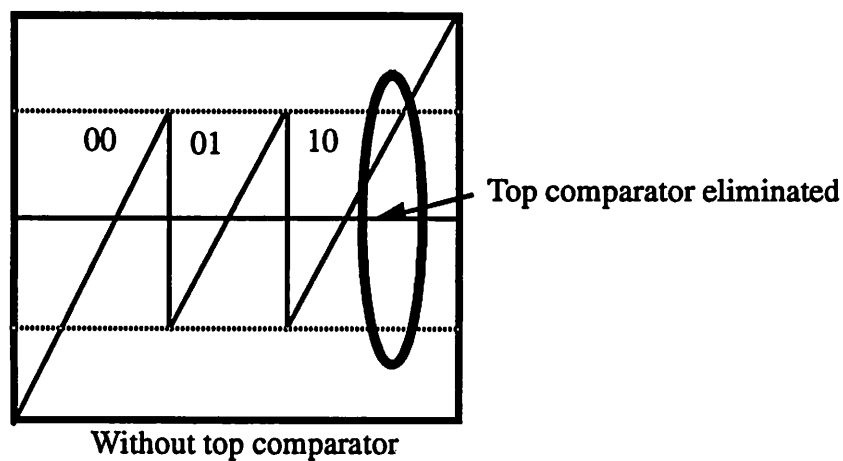
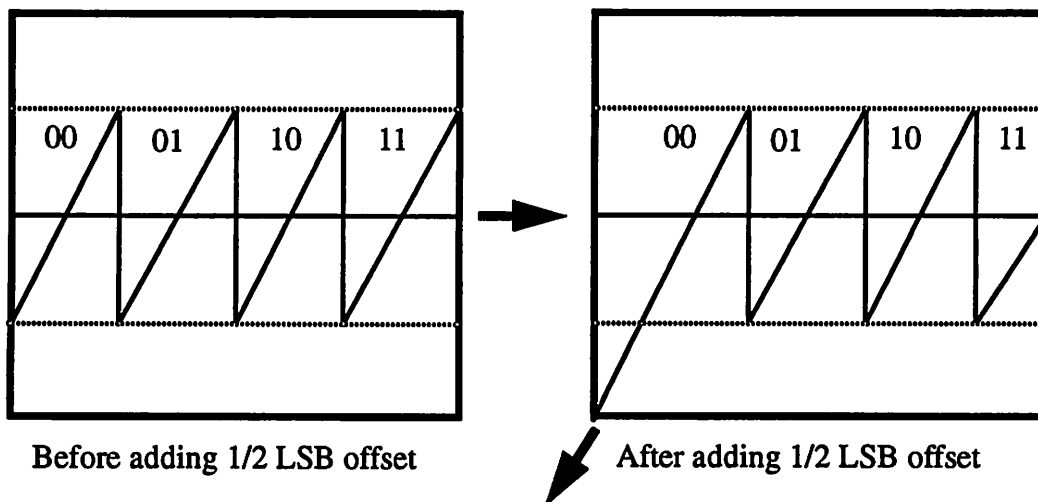
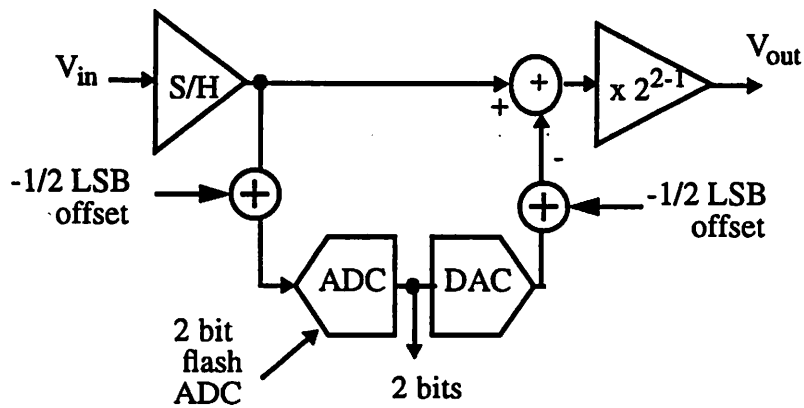


FIGURE 94. Modified digital error correction implementation to simplify the correction logic. The top comparator can be eliminated without altering the transfer function.

correction technique is used for simple correction logic as in [35][8](see section 5.2).

References

- [1] K. L. Lee, "Low Distortion Switched Capacitor Filters", *Electron. Res. Lab. Memo M86/12*, University of California, Berkeley, 1986.
- [2] A. N. Karanicolas, H. S. Lee, and K. L. Bacrania, "A 15b 1MS/s digitally self-calibrated pipeline ADC", in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 60-61.
- [3] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed., New York: Wiley, 1984.
- [4] W. R. Bennett, "Spectra of Quantized Signals", *Bell System Technical Journal*, pp. 446-472, July, 1948.
- [5] T. B. Cho and P. R. Gray, "A 10bit, 20MS/s, 35mW pipeline A/D converter," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1994, pp 23.2.1-23.2.4.
- [6] K. Nakamura et al., "A 85mW, 10bit 40Ms/s ADC with decimated parallel architecture," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1994, pp 23.1.1-23.1.4.
- [7] S. H. Lewis and P. R. Gray, "A pipelined 5-Msamples/s 9-bit analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. SC-22, pp. 954-961, Dec.,1987.
- [8] S. H. Lewis, et al., "10b 20Msamples/s analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. 27, pp. 351-358, March 1992.
- [9] Y. M. Lin, B. Kim, and P. R. Gray, "A 13b 2.5MHz self-calibrated pipelined A/D converter in 3- μ m CMOS," *IEEE J. Solid-State Circuits*, vol. 26, pp. 628-636, April 1991.
- [10] Y. M. Lin, "Performance limitations on high-resolution video-rate analog-digital interfaces," Memo. No. UCB/ERL M90/55, 19 June, 1990.

- [11] Matsuura, T., et al., "A 95mW, 10b, 15MHz low-power CMOS ADC using analog double-sampled pipelining scheme," *Symposium on VLSI Circuits Dig. Tech. Papers*, pp. 98-99, Jun. 1992.
- [12] C. S. G. Conroy, D. W. Cline, and P. R. Gray, "An 8-bit 85-MS/s parallel pipeline A/D converter in 1- μ m CMOS," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 447-454.
- [13] B. S. Song, M. F. Tompsett, and K. R. Lakshmikummar, "A 12-bit 1-Msample/s capacitor error-averaging pipelined A/D converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, December 1988, pp. 1324 - 1333.
- [14] M. Yotsuyanagi, E. Etoh, and K. Hirata., "A 10-b 50-MHz pipelined CMOS A/D converter with S/H," *IEEE J. Solid-State Circuits*, vol. 28, no. 3, March 1993, pp. 292-300.
- [15] C. W. Mangelsdorf, H. Malik, S.-H. Lee, S. Hisao, and M. Martin, "A two-residue architecture for multistage ADCs," in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 64-65.
- [16] C. S. G. Conroy, "A high-speed parallel pipeline A/D converter technique in CMOS," Memorandum No. UCB/ERL M94/9, Electronics Research Laboratory, U. C. Berkeley, February 1994.
- [17] B. Razavi and B. A. Wooley, "Design techniques for high-speed, high-resolution comparators," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, December 1992, pp. 1916-1926.
- [18] S. Hosotani, et al., "An 8bit 20MS/s CMOS A/D converter with 50mW power consumption," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, February 1990, pp. 167-172.
- [19] T. Matsuura, et al., "An 8b 50MHz 225mW submicron CMOS ADC using satura-

- tion eliminated comparators,” in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1990, pp 6.4.1-6.4.4.
- [20] K. Kusumoto et al., “A 10b 20MHz 30mW pipelined interpolating CMOS ADC,” in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 62-63.
- [21] M. Ito et al., “A 10b 20MS/s 3V supply CMOS A/D converter for integration into system VLSIs,” in *ISSCC Dig. Tech. Papers*, Feb. 1994, pp. 48-49.
- [22] B. Razavi and B. A. Wooley, “A 12b 5MSample/s two-step CMOS A/D converter,” *IEEE J. Solid-State Circuits*, vol. 27, no. 12, December 1992, pp. 1667-1678.
- [23] M. Ishikawa and T. Tsukahara, “An 8-bit 50-MHz CMOS subranging A/D converter with pipelined wide-band S/H,” *IEEE J. Solid-State Circuits*, vol. 24, no. 6, December 1989, pp. 1485-1491.
- [24] T. Matsuura, et al., “An 8-b 50-MHz 225-mW submicron CMOS ADC using saturation eliminated comparators,” in *Proc. IEEE Custom Integrated Circuits Conf.*, 1990, pp 6.4.1-6.4.4.
- [25] H-S, Lee, “Self-calibration techniques for successive approximation analog-to-digital converters,” Memorandum No. UCB/ERL M84/33, Electronics Research Laboratory, U. C. Berkeley, 10 April 1984.
- [26] K. B. Ohri and M. J. Callahan, JR., “Integrated PCM codec,” *IEEE J. Solid-State Circuits*, vol. 14, February 1979, pp. 38-46.
- [27] H. Ohara, et al., “A CMOS Programmable self-calibrating 13Bit eight-channel data acquisition peripheral,” *IEEE J. Solid-State Circuits*, vol. 22, no. 6, December 1987, pp. 930-938.
- [28] S. Kuboki et al, “Nonlinearity analysis of resistor string A/D converters,” *IEEE*

Transactions on Circuits and Systems, vol. 29, no. 6, June 1982, pp. 383-390.

- [29] J. Shyu, et al., "Random error effects in matched MOS capacitors and current sources," *IEEE J. Solid-State Circuits*, vol. 19, no. 6, December 1984, pp. 948-955.
- [30] J. Shieh, et al., "Measurement and analysis of charge injection in MOS analog switches," *IEEE J. Solid-State Circuits*, vol. 22, no. 2, April 1987, pp. 277-281.
- [31] R. C. Yen and P. R. Gray, "A MOS switched-capacitor instrumentation amplifier," *IEEE J. Solid-State Circuits*, vol. 17, no. 6, December 1982, pp. 1008-1013.
- [32] K. C. Hsieh, et al., "A low-noise chopper-stabilized differential switched-capacitor filtering technique," *IEEE J. Solid-State Circuits*, vol. 16, no. 6, December 1981, pp. 708-715.
- [33] R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for signal processing*, Wiley, 1986.
- [34] Y. Nakagome *et al.*, "Experimental 1.5-V 64-Mb DRAM," *IEEE J. Solid-State Circuits*, vol. 26, pp. 465-472, Apr. 1991.
- [35] G. Jusuf, "A 1-bit/cycle algorithmic analog-to-digital converter without high-precision comparators," University of California, Berkeley, Memo. UCB/ERL M90/69, Aug. 1990.
- [36] M. Wit, K. Tan, and R. Hester, "A low-power 12b analog-to-digital converter with on-chip precision trimming," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 455-461.
- [37] G. M. Yin, et al., "A high-speed CMOS comparator with 8-b resolution," *IEEE J. Solid-State Circuits*, vol. 27, NO.2, pp. 208-211, Feb. 1992.

- [38] K. J. McCall, et al., "A 6-bit 125MHz CMOS A/D converter," in *Proc. IEEE Custom Integrated Circuits Conf.*, 1992, pp 16.8.1-16.8.4.
- [39] G. T. Uehara and P. R. Gray, "A 100MHz output rate analog-to-digital interface for PRML magnetic-disk read channels in 1.2 μ m CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 1994, pp. 280-281.
- [40] K. Bult and G.J.G.M. Geelen, "A fast-settling CMOS opamp with 90-dB DC gain and 116 MHz unity-gain frequency," *ISSCC Dig. Tech. Papers*, pp. 108-109, Feb. 1990.
- [41] S. H. Lee and B. S. Song, "A code-error calibrated two-step A/D converter," *ISSCC Dig. Tech. Papers*, pp. 38-39, 1992.
- [42] D. W. Cline and P. R. Gray, "A power optimized 13bit 5MSamples/s pipelined Analog-to-digital converter in 1.2 μ m CMOS," unpublished work at the time of this thesis completion.
- [43] D. W. Cline's Ph D Thesis, "Noise/speed/power trade-offs in pipeline A/D converters," U. C. Berkeley, 1995.
- [44] K. A. Nishimura, "Optimal architectures for an integrated NTSC decoder," Memorandum No. UCB/ERL M90/63, Electronics Research Laboratory, U. C. Berkeley, June 1990.
- [45] L. A. Bienstman and H. J. DeMan, "An eight-channel 8bit microprocessor compatible nmos D/A converter with programmable scaling," *IEEE J. Solid-State Circuits*, vol. 15, pp. 1051-1059, Dec. 1980.
- [46] R. Castello and P. R. Gray, "Performance Limitations in Switched Capacitor Filters," *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 9, pp. 865-876, Sept. 1985.

- [47] D. Soo, "High-Frequency Voltage Amplification and Comparison in a One-Micron NMOS Technology," Ph.D. thesis, University of California, Berkeley, 1985.