NOISE, SPEED, AND POWER TRADE-OFFS IN PIPELINED ANALOG TO DIGITAL CONVERTERS

by

David William Cline

Memorandum No. UCB/ERL M95/94

22 November 1995

NOISE, SPEED, AND POWER TRADE-OFFS IN PIPELINED ANALOG TO DIGITAL CONVERTERS

by

David William Cline

Memorandum No. UCB/ERL M95/94

22 November 1995

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

Noise, Speed, and Power Trade-offs in Pipelined Analog to Digital Converters

by

David William Cline

Doctor of Philosophy in Engineering
Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Paul R. Gray, Chair

Power dissipation is becoming an increasingly important issue in the design of analog to digital converters as signal processing systems move into applications requiring either portability, or as in the case of some telecommunications applications, a high degree of parallelism. This research focuses on minimizing power dissipation in high resolution pipelined analog to digital converters, which are needed in applications requiring both high data rates and high resolution, such as medical imaging, high data rate digital radio receivers, and in some telecommunications systems. Power dissipation was minimized in this research through the appropriate choice of the per stage resolution, optimizing the distribution of the thermal noise budget among the various stages of the pipeline, the appropriate choice of opamp architecture, and through optimal sizing of opamps.

This thesis concludes that the optimum per stage resolution increases as the resolution specification increases. Per stage resolutions of less than one bit per stage are ideal for low resolution pipelines, but higher per stage resolutions are optimal for pipelines that have tighter noise constraints. This thesis also concludes that power is saved by scaling down the sizes of the sampling capacitors in later stages of the pipeline.

The trade-off between speed and power dissipation for opamps in switched capacitor circuits was also studied. It was found that for low gain high speed switched capacitor circuits, the single stage, telescopic cascode, and preamplifier driving a single stage opamp met the speed require-

ments with the least power dissipation. For lower speeds, the preamplifier driving a single stage opamp had the least power.

Paul R. Gray, Chair

Acknowledgments

First of all, I would like to express my appreciation of Professor Gray for his patient guidance of me during my years of graduate school here. I was truly very fortunate to have the opportunity to work under him as a student. In addition to providing guidance in circuit design issues, he also provided help in technical writing style and presentation style, and I found this guidance to be extremely valuable.

I would also like to express my appreciation to the other students in Professor Gray's group. I think that the atmosphere in this research group is very good. When I started working in this group, I immediately noticed and valued the high level of communication and the free flow of ideas here. I also found the environment to be very comfortable to work in, and people were very accepting. I feel very fortunate to have had the opportunity to work with the other students in this group.

I especially appreciate Cormac Conroy for his guidance and friendship during my time here at U. C. Berkeley. I worked with him closely on his parallel pipelined ADC and through working with him learned a great deal about circuit design. In addition to being a valuable resource, he was also a very good friend and provided a great deal of support to me through all the phases of my Ph.D. work.

I greatly valued sharing hiking stories with Eric Boskin, with whom I share an enthusiasm for the environment, about backpacking trips to the Sierra Nevada.

I also appreciate Robert Neff, who did a lot of work on ADC calibration and opamp compensation techniques that helped me get started on this project. I also enjoyed talking with him about a variety of things, including trips to the mountains and to the rivers.

Thomas Cho also was a valuable resource and friend to me during my research here at U. C. Berkeley. Despite his genius for circuit design, he dealt with people in a very humble manner, and I also appreciated him for listening to my ideas and providing feedback.

I would also like to thank Gani Jusuf for providing so much interesting conversation during the time we shared the office at 550J Cory Hall. In addition to doing our electronics research, I

think we solved all of the world's problems during our conversations! His cheerfulness and his presence definitely made work more enjoyable.

I would like to thank Todd Weigandt, with whom I also shared an office for a significant length of time. He was a tremendous source of encouragement, and I really enjoyed our nerdy attempts at applying technical principles to other aspects of life! I also received very valuable advice on a number of subjects, including the automobile purchase negotiation process.

Ken Nishimura was an extremely valuable resource during his time here at Berkeley. Whenever problems came up, he was somehow always a very dependable source of solutions to those problems. He was also very patient and always cheerfully provided help even when he was being nagged repeatedly with questions. I was amazed that in spite of Ken's tremendous service to the group, he somehow managed to rapidly complete his degree.

I would also like to thank Greg Uehara, who was often a source of very interesting discussion about fundamental circuits problems. I think I learned a lot from these discussions. Furthermore, Greg was often a source of encouragement to me.

Ole Bentz often got together with me for lunch, and talking with him eased the stress of graduate school and made it more enjoyable.

Chris Rudell, a recent officemate of mine, was always very exciting to have around. His general enthusiasm and sense of humor had a big impact on the atmosphere in the office and made work a lot more fun.

I also enjoyed the ski trip to Lake Tahoe and Reno with Weijie Yun. Weijie was amazing at the wacky wire. We returned to Berkeley with a car packed to the hilt with stuffed animals! Many thanks also to Keith Onodera, Ed Liu, Shoichiro Tada, Tom Truman, Caesar Wong, Ichabod Nguyen, and Kwan Chan for making the IC ski trip in December 1992 a lot of fun!

A corrupted directory caused by a disk head crash in spring 1995 threatened to cast my thesis into oblivion. Fortunately, Sekhar Narayanaswami performed some wizadry and rescued my dissertation from this horrible fate and saved me a month of work. I am very grateful for this.

I would also like to thank Andy Abo, Thomas Cho, Arnold Feldman, and Li Lin for their help in reading various drafts of my thesis, helping me find mistakes, and providing me suggestions on the content and style.

This work was supported by the National Science Foundation and the California Micro Program. The fabrication was performed by Orbit Semiconductor. Many thanks are also in order to Accurel and Sean Lee with FIB Lab. Their ion milling operations saved the project from some layout mistakes.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1
1.1 Motivation 1
1.2 Organization of the Dissertation 1
1.3 ADC Definition 2
1.4 ADC Characterization 4
1.4.1 Resolution 4
1.4.1.1 Nonlinearity 5
1.4.1.2 Signal to Noise Ratio 6
1.4.1.3 Signal to Noise + Distortion Ratio 7
1.4.1.4 Dynamic Range 8
1.4.1.5 Spurious Free Dynamic Range 8
1.4.2 Sampling Rate 9
1.4.3 Input Bandwidth 9
1.4.4 Power Supply Rejection Ratio 10
1.4.5 Input Capacitance 10
1.4.6 Input Signal Swing 10
1.4.7 Power Dissipation 10
CHAPTER 2 REVIEW OF ANALOG TO DIGITAL CONVERTER ARCHITECTURES 12
2.1 Flash ADC 12
2.2 Two Step Flash ADC 14
2.3 Subranging ADC 16
2.4 Successive Approximation ADC 17
2.5 Pipelined ADC 19
2.6 Recirculating ADC 21
2.7 Oversampled ADC 22
2.8 Serial ADC 23
2.9 Recent Performance Achievements 24
CHAPTER 3 DESIGN TECHNIQUES FOR PIPELINED ANALOG TO DIGITAL CONVERTERS 43
3.1 Introduction to the Concept of Pipelined ADCs 43
3.2 Switched Capacitor DAC and Residue Amplifier 47
3.3 Sources of Error in Pipelined Analog to Digital Converters 51

3.3.1 Thermal Noise....51
3.3.2 Comparator Offsets....52
3.3.3 Residue Amplifier Gain Error....56
3.3.4 Nonuniform Reference Levels (Nonlinear DAC)....61
3.3.5 Residue Amplifier Nonlinearity....68
3.3.6 Incomplete Settling of the Sample and Hold Amplifier Output....78
3.3.6.1 Settling Time of a Single Pole System....79
3.3.6.2 Settling Time of a Critically Damped Two Pole System....79
3.3.6.3 Settling Time of an Underdamped Two Pole System....80
3.3.6.4 Settling Time of an Overdamped Two Pole System....83
3.3.6.5 Tabulated Settling Times....85
3.3.7 Sample and Hold Tracking Nonlinearity....86
3.4 Error Correction Techniques....98
3.4.1 Analog Offset Correction....98
3.4.2 Digital Comparator Error Correction....103
3.4.3 Analog DAC/Gain Calibration....107
3.4.4 Capacitor Error Averaging....110
3.4.5 Digital DAC/Gain Calibration....113
CHAPTER 4 SAMPLE AND HOLD AMPLIFIER ARCHITECTURES AND OPTIMIZATION....123
4.1 MOSFET Models for Transient Analysis....123
4.1.1 Long Channel Model for the MOSFET....124
4.1.2 Vertical Field Mobility Degradation....125
4.1.3 Velocity Saturation....126
4.1.4 Subthreshold....130
4.1.5 Putting the Models Together....130
4.2 Settling Time Analysis of Switched Capacitor Gain Stages....134
4.2.1 Single Stage Single Pole Amplifier....137
4.2.1.1 Fixed Current Density....139
4.2.1.1.1 Minimum Power with Fixed Speed....141
4.2.2 Telescopic Cascode Amplifier....144
4.2.2.1 Optimization of Current Density to Minimize Power - Fixed Speed and Fixed Feedback Capacitance - Model Including Mobility Degradation, Velocity Saturation,

and Subthreshold....149
4.2.3 Wide-Band Preamplifier Driving a Single Stage Amplifier....155
4.2.3.1 Optimization to Minimize the Power....161
4.2.4 Two Stage Amplifier with Standard Miller Compensation....167
4.2.5 Two Stage Amplifier with Ahuja Style Compensation....175
4.2.6 Three Stage Amplifier with Nested Miller Compensation....187
4.2.7 Comparison of Topologies....198
Appendix....203
4.A.1 Single Stage Amplifier Optimizations....205
4.A.1.1 Minimum Power with Fixed Speed....206
4.A.1.2 Maximum Speed with Fixed Feedback Capacitance....209
4.A.1.3 Speed and Power Optimization with Variable Current Density-Long Channel Model....211
4.A.1.3.1 Speed and Power Optimization with Variable Feedback Capacitance....214
4.A.1.3.2 Special Case: No Output Parasitic Capacitance....219
4.A.2 Telescopic Cascode Amplifier Optimizations....220
4.A.2.1 Optimization of Feedback Capacitance to Minimize Power - Fixed Current Density and Speed....221
4.A.2.1.1 Results of Optimization Using the Long Channel Model....221
CHAPTER 5 OPTIMIZATION TECHNIQUES FOR PIPELINED ANALOG TO DIGITAL CONVERTERS....227
5.1 Pipelined Analog to Digital Converter Design in the Absence of Noise....227
5.1.1 Optimum Sampling Capacitor Size to Minimize Pipeline Power....227
5.1.2 Optimum Number of Bits Per Stage to Minimize Pipeline Power....236
5.2 Pipelined Analog to Digital Converter Design in the Presence of Noise....239
5.2.1 Thermal Noise in Switched Capacitor Gain Blocks....241
5.2.1.1 Thermal Noise Contribution of the Sampling Switches (kT/C Noise)....241
5.2.1.2 Thermal Noise Contribution of the Transconductance Amplifier....243
5.2.1.2.1 Thermal Noise of a Single Stage Amplifier....244
5.2.1.2.2 Thermal Noise of a Critically Damped Telescopic Cascode Amplifier....247
5.2.1.2.3 Thermal Noise of a Preamplifier

Driving a Single Stage Amplifier 248
5.2.1.2.4 Thermal Noise of a Two Stage Miller Compensated Amplifier 251
5.2.1.2.5 Thermal Noise of an Amplifier with Ahuja Style Compensation 254
5.2.2 Optimal Capacitor Sizing in High Resolution-Low Speed Pipelined Analog to Digital Converters 256
5.2.3 Optimal Closed Loop Gain of Interstage Gain Amplifiers in High Resolution-High Speed Pipelined Analog to Digital Converters 264
5.2.4 Optimum Closed Loop Gain of Interstage Gain Amplifiers in a Pipelined ADC with Parasitics Included 273
5.2.5 Optimum Supply Voltage for Power Dissipation in a Pipelined ADC 280
5.2.5.1 Thermal Noise Limits to Power Dissipation in a Pipelined ADC 281
5.2.5.2 Power Dissipation Trade-offs in Choosing the Supply Voltage 283
5.3 Scaling of Power and Speed in an Optimized Pipelined ADC 289
Appendix 295
CHAPTER 6 PROTOTYPE DESIGN AND DESCRIPTION 317
6.1 Design Goal 317
6.2 Architecture 317
6.2.1 Comparator Architecture 317
6.2.2 Encoding Network 323
6.2.2.1 Voting Error Correction 323
6.2.2.2 Digital to Analog Converter 324
6.2.3 Sample and Hold Amplifier Architecture 332
6.2.4 Operational Amplifier Architecture 334
6.2.5 Bias Circuit for the Operational Amplifier 336
6.2.6 Clock Generation 341
6.2.7 Layout Considerations 344
6.2.7.1 Sampling Capacitor Layout 344
6.2.7.2 Interconnect Layout 346
6.3 Power Supply Noise Issues 348
CHAPTER 7 EXPERIMENTAL RESULTS 357
7.1 Die Photograph 357
7.2 Code Density Test 358

7.3 Signal to Noise Ratio and Distortion 363
7.3.1 Results for Low Input Frequency (100kHz) 364
7.3.2 Results for High Input Frequency (2MHz) 365
7.3.3 Signal to Noise Ratio Versus Input Amplitude 366
7.3.4 Idle Channel Noise 368
7.4 Summary of Results 368
CHAPTER 8 CONCLUSIONS AND FUTURE WORK 371
8.1 Conclusions 371
8.1.1 Optimization of Pipelined ADCs 371
8.1.2 Opamp Architectures for Switched Capacitor Applications 373
8.2 Pipelined ADC with Low Swing Amplifiers and Extra Comparators 374

CHAPTER 1

INTRODUCTION

1.1 Motivation

This research is concerned with improving the efficiency of analog to digital converters (ADCs). For a long time, ADCs have been used widely in digital test equipment. Recently, the applications for ADCs have expanded widely as many electronic systems that used to be entirely analog have been implemented using digital electronics. Examples of such applications include digital telephone transmission, cordless phones, transportation, and medical imaging. Furthermore, ADCs have found their way into systems that would normally be considered as being entirely digital as these digital systems are pushed to higher levels of performance. Data storage is one example of such a system. As storage density in disk drive systems is increased, the signals handled by the read circuitry have become increasingly analog in character. Presently, 6-bit ADCs are commonly used in the read circuits of disk drives.

Frequently, ADCs are integrated with other functions on a single monolithic device. Because the ADC must share the power budget with other functions on the integrated circuit, power dissipation is often an important consideration. Furthermore, many new applications, such as cordless phones and cellular phones, require portability and battery operation. Because of these requirements, power dissipation is becoming increasingly important.

1.2 Organization of the Dissertation

This dissertation is divided into eight chapters. The first chapter introduces the concept of analog to digital conversion and discusses methods used to characterize ADCs. Chapter 2 is a survey of ADC architectures. Chapter 3 focuses on pipelined analog to digital converters. In particular, error sources and techniques for dealing with them are discussed.

The sample and hold amplifier is a key block in pipelined ADCs. Therefore, in chapter 4, switched capacitor circuits using several different operational transconductance amplifier architectures are discussed with an emphasis on optimally trading off speed and power dissipation. At the end of the chapter, the speed performance of each of these architectures is compared with the others.

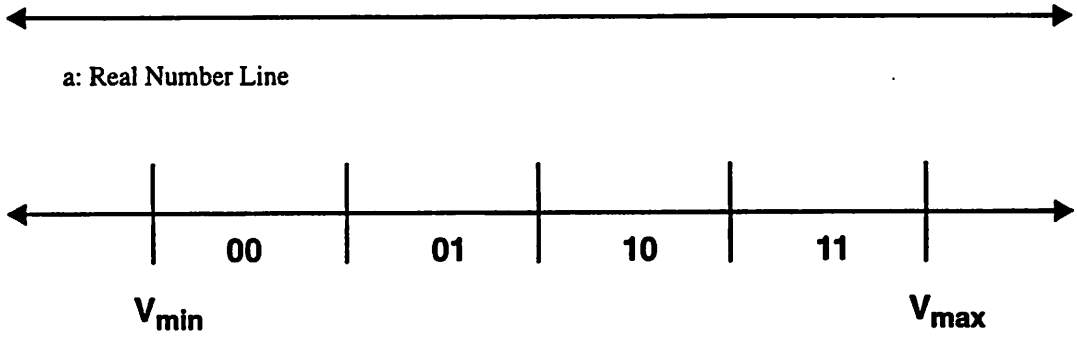
Chapter 5 discusses the optimization of the complete pipeline system to minimize power. A 14 bit prototype ADC is described in chapter 6, and the results of testing this ADC are presented in chapter 7.

Chapter 8 contains a summary of the conclusions of this research and presents a proposal for future research.

1.3 ADC Definition

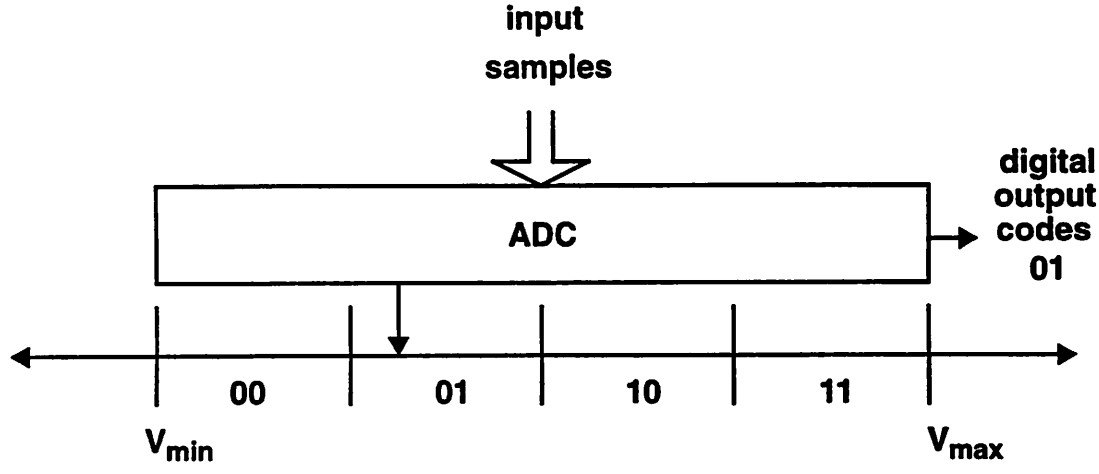
An analog to digital converter (ADC) is a device that converts real world (analog) signals into digital codes[Sheingold86]. Conceptually, an ADC works as shown in figure 1.1. Analog signals have a continuous range of values as do numbers on the real number line. An ADC takes a range of the real number line and divides it into smaller subranges. The size of each of the subranges is often referred to as the step size. These steps are usually uniform in size, but not always. A companding ADC for codecs is one example of an ADC having nonuniform step sizes. In this case the step sizes follow a logarithmic scale.

To each subrange or step a code is assigned. Then, during the conversion process input samples are taken and mapped onto this real number line. The ADC then decides which subrange corresponds to the sample and sends the appropriate digital code to the output.



a: Real Number Line

b: A range is selected and subdivided into smaller ranges.



c: ADC maps input samples to the real number line and decides which digital output code is appropriate.

Figure 1.1 Description of an ADC

To perform its task, an ADC always uses at least one comparator. A comparator is a device that compares two quantities and makes a decision based on which of the quantities is larger. The operation of a comparator is illustrated in figure 1.2. In general, the inputs to a comparator can be either analog or digital, but the output is always digital. In analog to digital conversion applications

the inputs to the comparator are analog, and the output is a binary digital quantity.

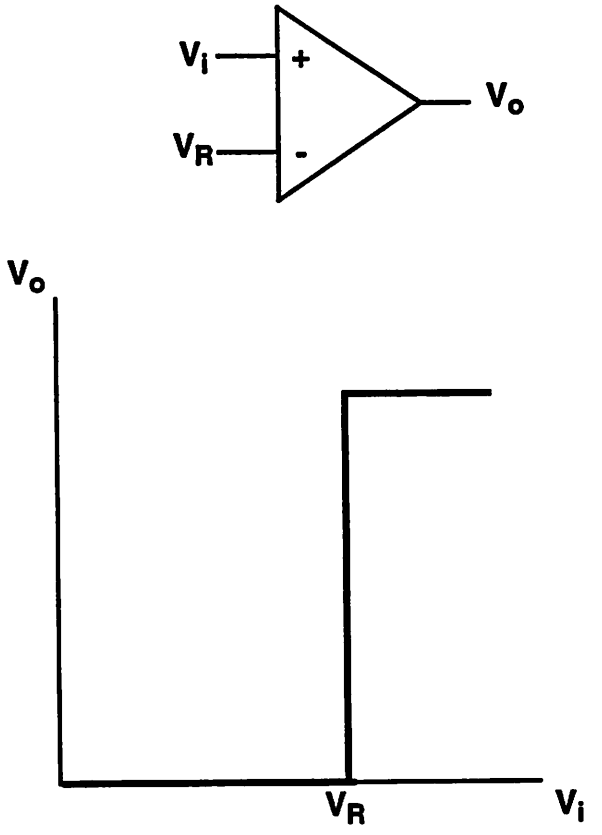


Figure 1.2 Comparator Circuit and Transfer Function

1.4 ADC Characterization

Analog to digital converters are characterized in a number of different ways to indicate the performance capability, cost, and ease of use. Some of the most important characteristics of ADCs are introduced below.

1.4.1 Resolution

Resolution describes the fineness of the quantization performed by the ADC. A high resolution ADC divides the input range into a larger number of subranges than a low resolution converter. Resolution is usually defined as the base 2 logarithm of the number of subranges the ADC input

range is divided into. This quantity is referred to as the number of bits resolved by the ADC. Thus, for a fixed full scale input range a high resolution ADC can resolve smaller signals than a low resolution ADC is able to resolve. Resolution is usually degraded by either noise or nonlinearity. Therefore, most techniques for characterizing the true resolution of an ADC measure either noise, nonlinearity, or both.

1.4.1.1 Nonlinearity

Some applications, such as telephone codecs, require an ADC that is intentionally nonlinear. However, most ADCs are intended to have a transfer characteristic that approximates a straight line. As the resolution increases, the input-output characteristic of the ADC better approximates a straight line. The transfer characteristic for an ideal version of such an ADC progresses from low to high in a series of uniform steps. Because of this fact, nonlinearity is present even in an ideal ADC. The transfer characteristic of a practical ADC contains steps which are not perfectly uniform, and this deviation generally contributes to further nonlinearity. Two types of nonlinearity are used to characterize this deviation. Differential nonlinearity (DNL) measures how far each of the step sizes deviates from the nominal value of the step size. Integral nonlinearity (INL) is the difference between the actual transfer characteristic and the straight line characteristic which the ADC is intended to approximate. DNL and INL are both plotted as a function of code. DNL and INL are generally expressed in terms of least significant bits (LSBs) of converter input. An LSB of converter input is equal to the full input range of the ADC divided by the number of steps. Figure 1.3 illustrates DNL and INL.

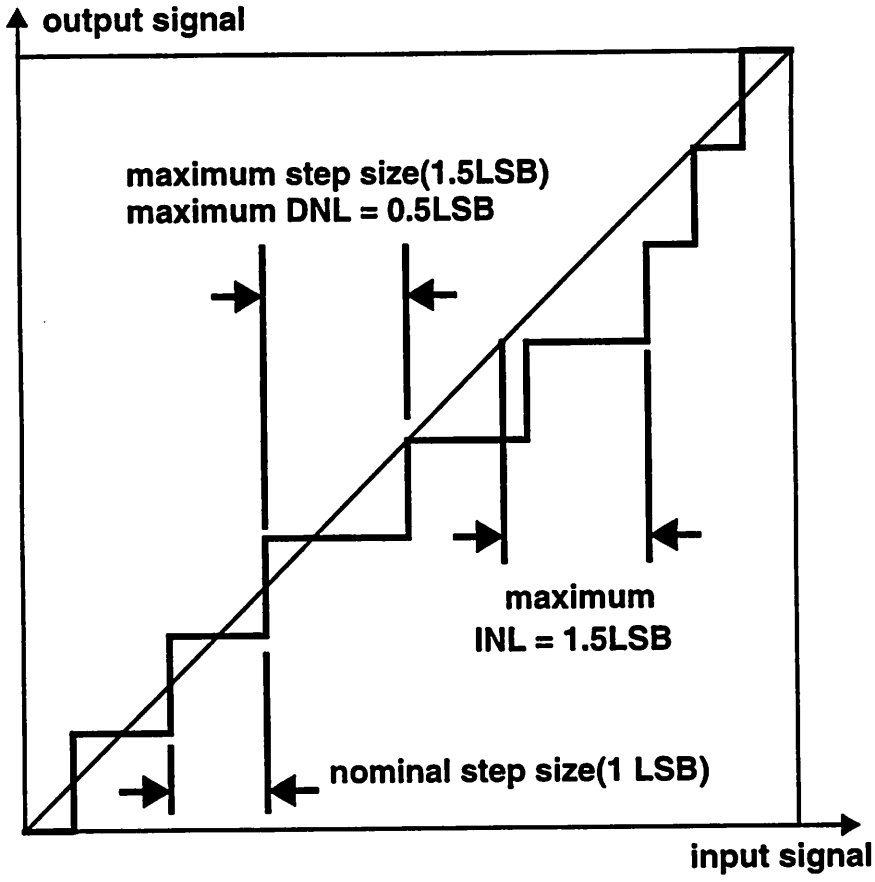


Figure 1.3 Example Transfer Characteristic of a 3 Bit ADC Showing DNL and INL

1.4.1.2 Signal to Noise Ratio

The signal to noise ratio (SNR) is the ratio of signal power to noise power in the output of the ADC. One way to measure SNR is to plot the spectrum of the output of the ADC. The SNR is calculated by measuring the difference between the signal peak and the noise floor and including a factor to adjust for the number of samples used to generate the spectrum as shown below.

$$\text{SNR (dB)} = \text{signal peak (dB)} - \text{noise floor(dB)} - 10\log N \qquad (1.1)$$

The last term in the above equation may be understood as follows. To generate an N point fast Fourier transform (FFT) of a signal, N samples of the signal are taken. Sampling the signal N times increases the signal energy by a factor of N^2 and the noise energy by a factor of N. Thus the ratio

of signal power to noise power is increased by a factor of N and the signal to noise ratio of the FFT is higher than the signal to noise ratio in one sample of the signal. The signal to noise ratio improvement in dB is 10logN. Thus, the noise floor in the FFT becomes lower relative to the signal as more samples are taken. This idea is illustrated in figure 1.4.

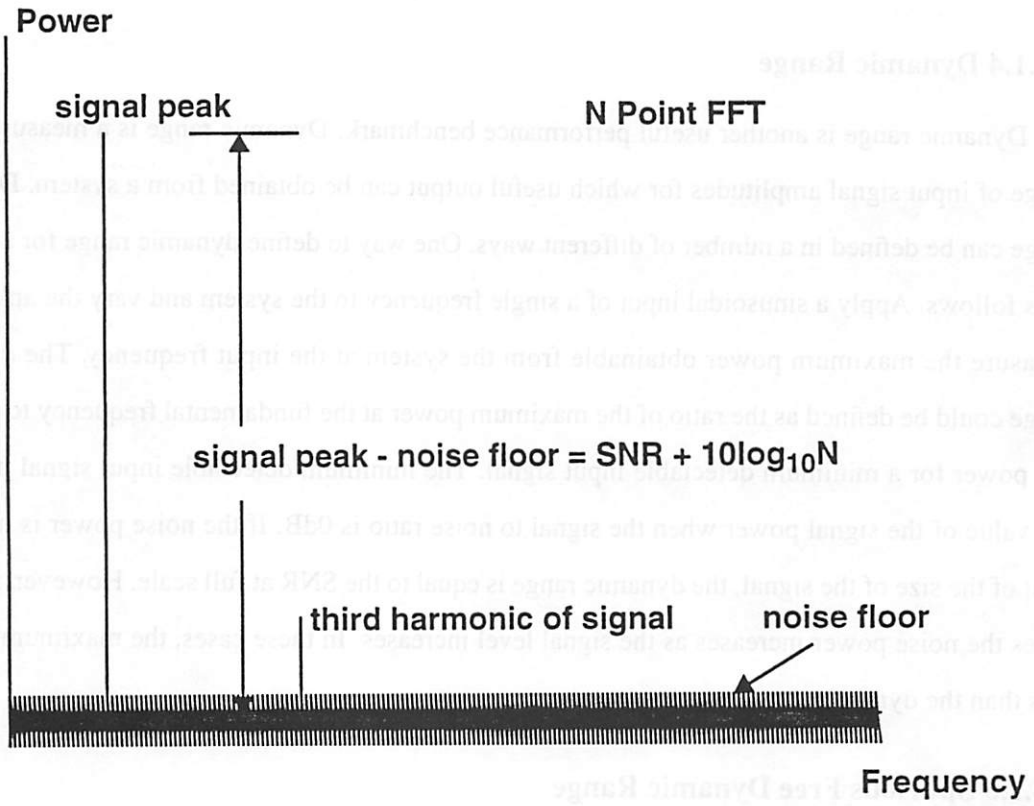


Figure 1.4 Procedure for Computing SNR from an N point FFT

1.4.1.3 Signal to Noise + Distortion Ratio

The signal to noise plus distortion ratio (SNDR) is often used to measure the performance of an ADC. It measures the degradation due to the combined effect of noise, quantization errors, and harmonic distortion. The SNDR of a system is usually measured for a sinusoidal input and is a function of the frequency and amplitude of the input signal. When a sinusoidal signal of a single frequency is applied to a system, the output of the system generally contains a signal component at the input frequency. Due to distortion, the output also contains signal components at harmonics of

the input frequency. An ADC usually samples an input signal at some finite rate. As a result, some of the harmonic distortion products are aliased down to lower frequencies. Furthermore, the ADC adds noise to the output, and this noise generally present to some degree at all frequencies. The SNDR of the ADC is defined as the ratio of the signal power in the fundamental to the sum of the power in all of the harmonics, all of the aliased harmonics, and all of the noise.

1.4.1.4 Dynamic Range

Dynamic range is another useful performance benchmark. Dynamic range is a measure of the range of input signal amplitudes for which useful output can be obtained from a system. Dynamic range can be defined in a number of different ways. One way to define dynamic range for a system is as follows. Apply a sinusoidal input of a single frequency to the system and vary the amplitude. Measure the maximum power obtainable from the system at the input frequency. The dynamic range could be defined as the ratio of the maximum power at the fundamental frequency to the output power for a minimum detectable input signal. The minimum detectable input signal power is the value of the signal power when the signal to noise ratio is 0dB. If the noise power is independent of the size of the signal, the dynamic range is equal to the SNR at full scale. However, in some cases the noise power increases as the signal level increases. In these cases, the maximum SNR is less than the dynamic range.

1.4.1.5 Spurious Free Dynamic Range

Another way to define dynamic range is the spurious free dynamic range. The spurious free dynamic range is the ratio of the input signal level for maximum SNDR to the input signal level for 0dB SNDR. This measure of dynamic range is useful because it indicates the amount of dynamic range that can be obtained before distortion becomes dominant over noise. Figure 1.5

indicates how to determine spurious free dynamic range from a plot of SNDR versus input level.

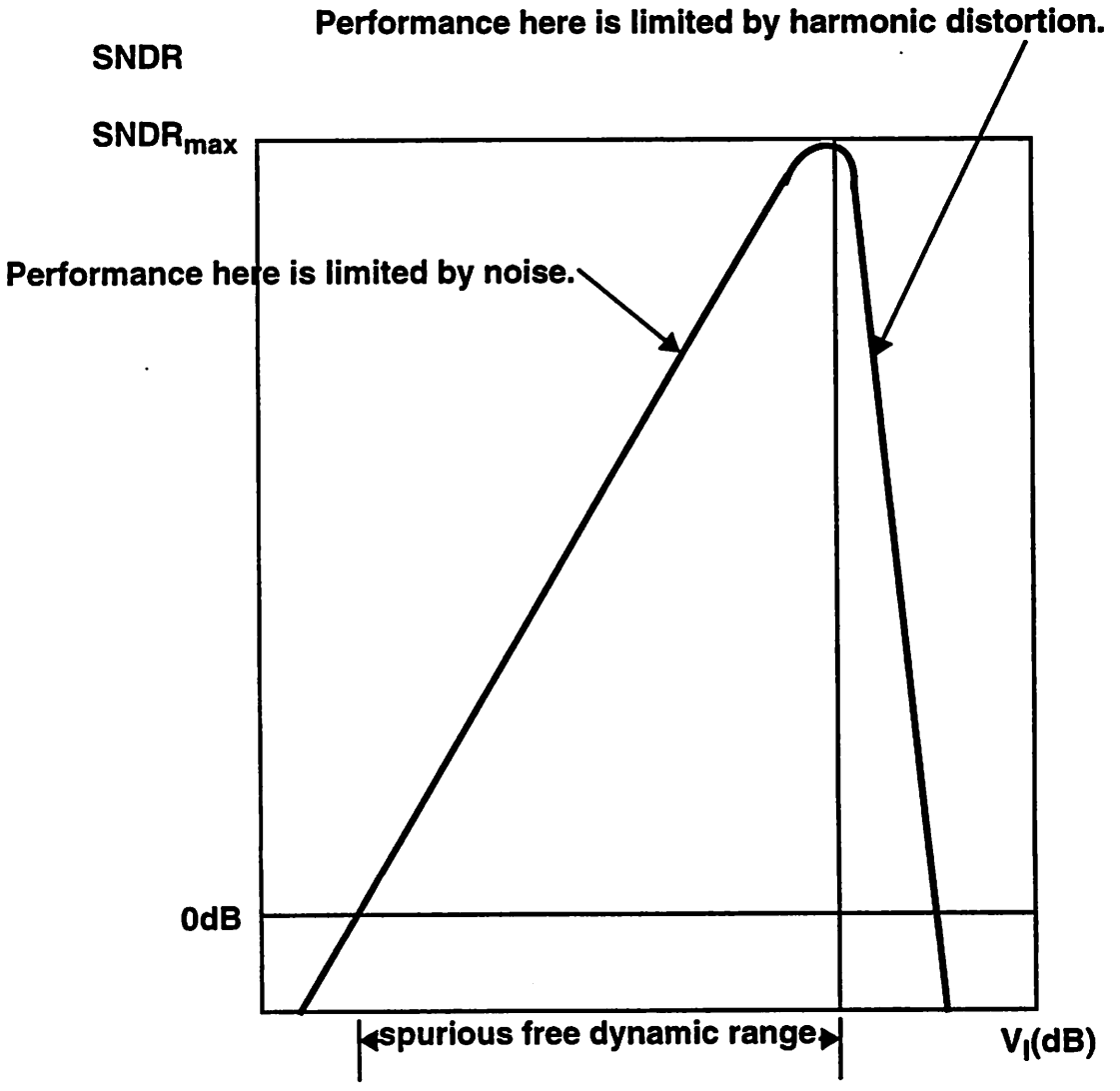


Figure 1.5 Typical SNDR versus Signal Level for an ADC

1.4.2 Sampling Rate

The sampling rate indicates the number times the input signal is sampled per second.

1.4.3 Input Bandwidth

ADC resolution is a function of the frequency of the input signal. At high input frequencies,

the SNDR of the ADC output can be reduced by a number of effects. The input bandwidth of the ADC is the input frequency at which the SNDR is 3dB below the maximum value.[van de Plassche94]

1.4.4 Power Supply Rejection Ratio

Noise on the power supply lines can couple into the output signal of an ADC. The power supply rejection ratio (PSRR) measures how well an ADC resists this tendency. It is the ratio of supply noise power to output noise due to the power supply noise.

1.4.5 Input Capacitance

Input capacitance is the capacitive load presented by the ADC to the circuit driving it. The input capacitance of an ADC is an important parameter because the input capacitance can load the circuit driving the ADC and degrade its performance.

1.4.6 Input Signal Swing

Input signal swing indicates the allowable range of values for the input. The input signal swing indicates the maximum and minimum values that the input signal may have without driving the ADC out of range or resulting in an unacceptable level of distortion.

1.4.7 Power Dissipation

Power dissipation is becoming an important ADC specification because many ADCs are being implemented in portable systems powered by a battery with limited energy. Reducing power dissipation can reduce system weight or improve battery life. Reducing power dissipation can also make it easier to keep the temperature of the ADC at a reasonable level.

References

[Sheingold86] Daniel H. Sheingold, ed., *Analog-Digital Conversion Handbook*, 3rd ed., Englewood Cliffs, New Jersey: Prentice-Hall, 1986, pp. 1-15.

[van de Plassche94] Rudy J. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, Boston, Mass.: Kluwer, 1994, pp. 74-75.

CHAPTER 2

REVIEW OF ANALOG TO DIGITAL CONVERTER ARCHITECTURES

Much research has been done on the implementation of analog to digital converters (ADCs). As a result, a number techniques for doing analog to digital conversion have been developed. In this chapter, some of the prominent architectural styles are introduced and compared. Each architecture has advantages and disadvantages, and each has a set of applications for which it is the best solution.

2.1 Flash ADC

The flash ADC architecture[Dingwall79][Sheingold86][Joy86][Kumamoto86][Peetz86][Mangelsdorf90][McCall92], also known as a fully parallel architecture, is fundamentally the fastest architecture. This architecture is also conceptually the easiest to understand. An n-bit flash ADC consists of an array of 2^n-1 comparators and a set of 2^n-1 reference values. Each of the comparators samples the input signal and compares the signal to one of the reference values. Each comparator then generates an output indicating whether the input signal is larger or smaller than the reference assigned to that comparator. The set of 2^n-1 comparator outputs that results is often referred to as a thermometer code. This is name is derived from the fact that if the comparator outputs are listed in a column and ordered according to the reference values associated with the comparator that produced them, the ones would all be at the bottom, and the zeros all at the top. The level of the boundary between ones and zeros would indicate the value of the signal, much as the level of mercury in a mercury thermometer indicates the temperature.

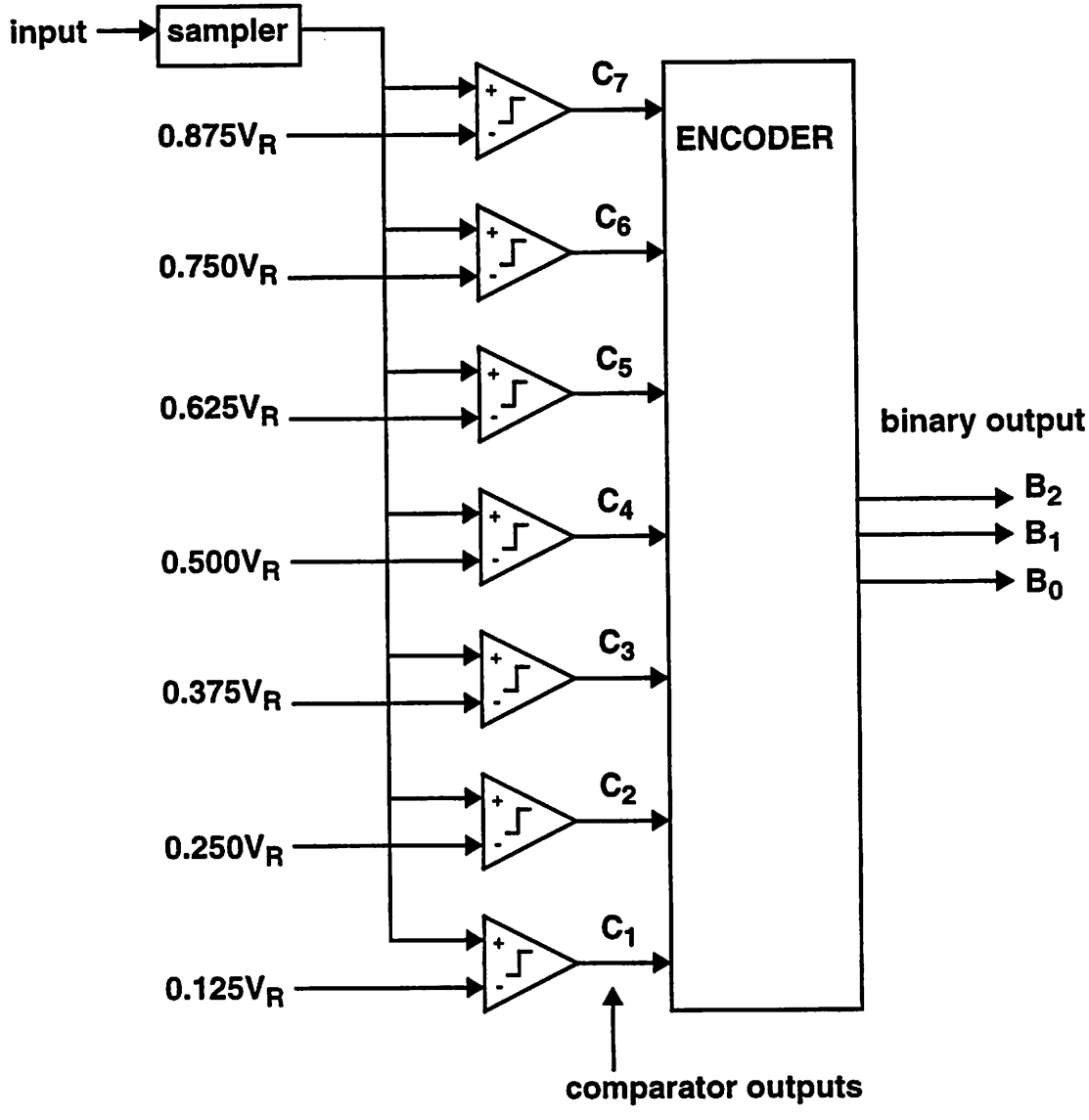


Figure 2.1 Simple 3-bit Flash ADC

A simple 3 bit flash ADC is shown in figure 2.1. The encoder converts the thermometer code produced by the comparators to a binary code as shown in the truth table in table 2.1. As seen from the figure, the comparators all operate in parallel. Thus, the conversion speed is limited only by the speed of the comparator or the sampler. For this reason, the flash ADC is capable of high speed.

Table 2.1 Truth Table for Flash ADC of Figure 2.1

	C_7	C_6	C_5	C_4	C_3	C_2	C_1	B_2	B_1	B_0
$V_I<0.125V_R$	0	0	0	0	0	0	0	0	0	0
$0.125V_R < V_I < 0.25V_R$	0	0	0	0	0	0	1	0	0	1
$0.25V_R < V_I < 0.375V_R$	0	0	0	0	0	1	1	0	1	0
$0.375V_R < V_I < 0.50V_R$	0	0	0	0	1	1	1	0	1	1
$0.50V_R < V_I < 0.625V_R$	0	0	0	1	1	1	1	1	0	0
$0.625V_R < V_I < 0.75V_R$	0	0	1	1	1	1	1	1	0	1
$0.75V_R < V_I < 0.875V_R$	0	1	1	1	1	1	1	1	1	0
$V_I>0.875V_R$	1	1	1	1	1	1	1	1	1	1

The two primary drawbacks to the flash ADC are the large hardware requirement and sensitivity to comparator offsets. As mentioned earlier, 2^n-1 comparators are required. For this reason, a high resolution flash ADC requires a large circuit area and dissipates high power. Furthermore, the large number of comparators present a large capacitance to the output of the sampling circuit. The required comparator offset voltage for a flash ADC with n bit resolution is less than $1/2^n$. At high resolutions, this required comparator offset becomes very small. Because comparators with small offsets are difficult to design and expensive to build and because so many comparators are required, ADCs with resolutions higher than 8 bits rarely use the flash architecture.

2.2 Two Step Flash ADC

A two step flash ADC[van de Plassche79][Shimizu89][Mayes89][Doernberg89][Razavi92] consists of two stages, each containing a flash ADC. The block diagram in figure 2.2 illustrates the structure of a two step flash ADC. In this type of ADC, the conversion does not happen all at once as in the flash ADC. Here, the conversion takes two steps. During the first step, the most significant bits of the digital output are determined by the first stage flash ADC. Then a DAC converts this digital result back to an analog signal to be subtracted from the input signal. This residue is then sent to the second stage flash ADC. The second stage flash then determines the least significant bits of the digital output.

The conversion time for a two step flash ADC is longer than for a simple flash, but it is still very fast. Furthermore, the two step flash ADC requires only $2 \bullet 2^{n/2}$ comparators, which is fewer comparators than required by a simple flash ADC. Thus, the two step flash ADC saves hardware. As a result, two-step flash ADCs are often used in the 10 bit resolution range.

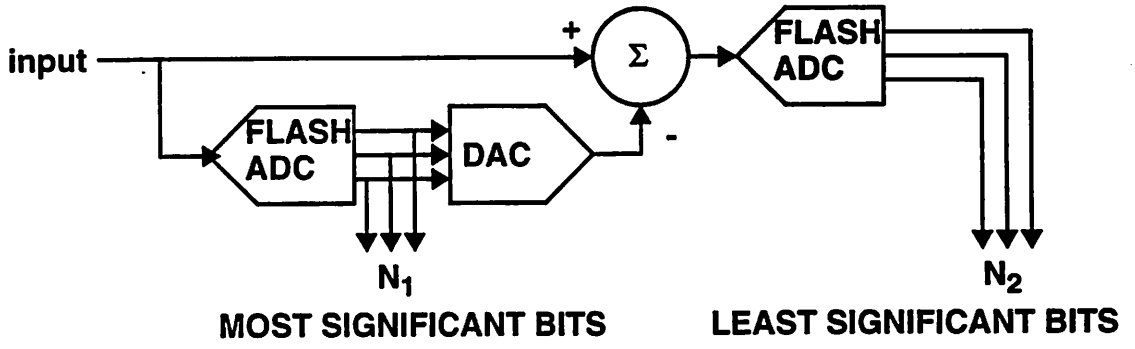


Figure 2.2 Two Step Flash ADC

The folding ADC architecture[van de Grift87][van de Plassche88][van Valberg92][Flynn95][Nauta95] is a popular subset of the two-step flash architecture. In this architecture, the DAC and differencing blocks are replaced by an analog preprocessing circuit called a folding circuit. The folding circuit has a triangular input-output characteristic. The output of this block can be used as the input to a fine ADC that determines the least significant bits of the digital output. A block diagram for this converter architecture is shown in figure 2.3.

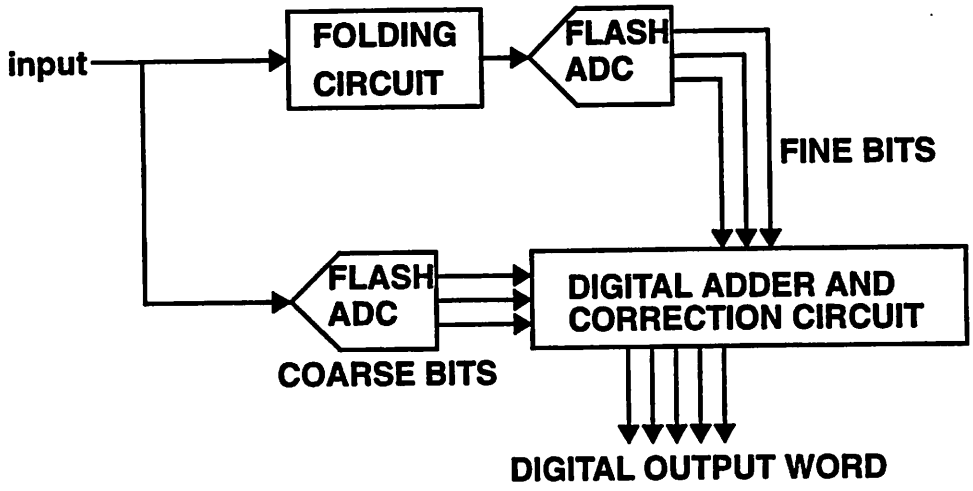


Figure 2.3 Block Diagram of a Folding ADC

2.3 Subranging ADC

A subranging ADC architecture[Dingwall85][Fernandes88][Ishikawa89][Kolluri89][Petschacher90][Mercer91] is a multi-step converter architecture that includes two-step flash ADCs and also includes ADCs that extend the concept of the two-step flash ADC to a larger number of steps. By breaking the conversion process into multiple steps, fewer comparators are required, but the conversion time is longer. The block diagram in figure 2.4 illustrates the structure of the subranging ADC. Each stage is responsible for resolving some part of the digital output word and delivering a residue to the following stage. The conversion time required increases with the number of stages while the hardware required decreases with the number of stages. Thus, there is a trade-off between speed and hardware. The comparators in the front stages need not be accurate, but the comparators in the last stage must be accurate to the full resolution of the ADC.

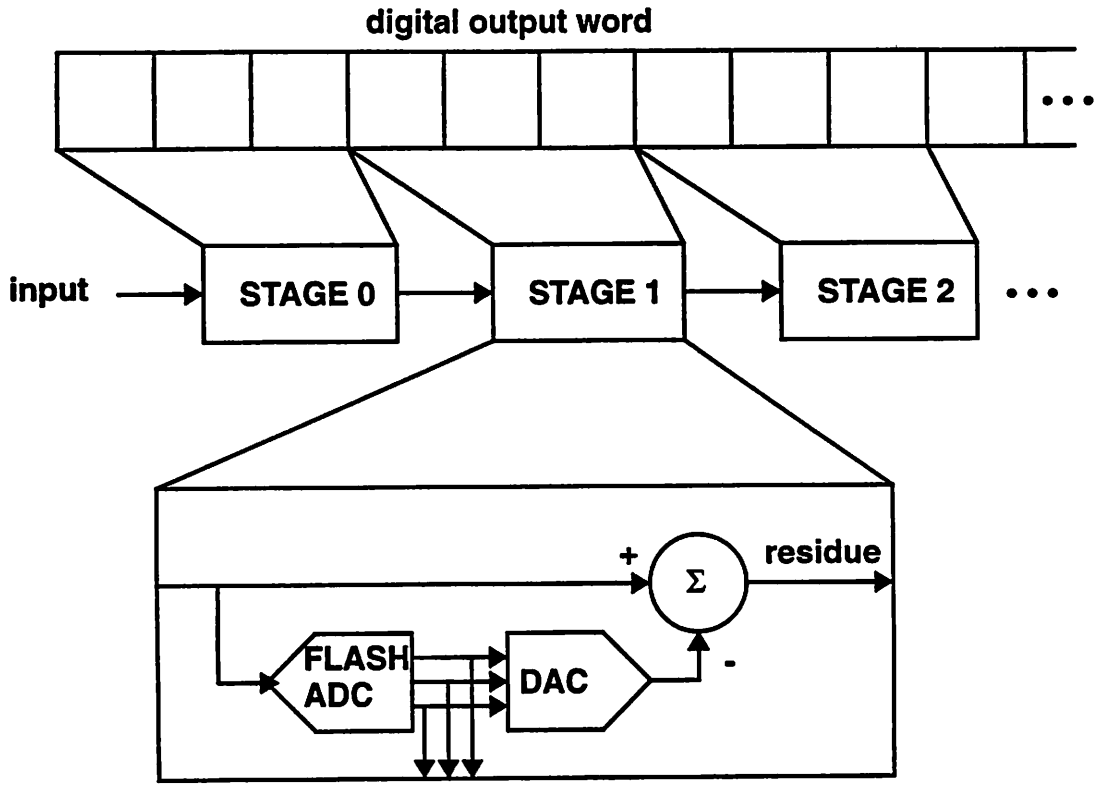


Figure 2.4 Subranging ADC

2.4 Successive Approximation ADC

A successive approximation ADC[McCreary75][Suarez75][Hamade78][Redfern79][Fotouhi79][Connolly80] [Timko80][Boyacigiller81][Saul81][Crolla82] [van de Plassche82][Lee84][Croteau86][Bacrania86] [Manoli89][Hester90][Hadidi90][Tan90][Fattaruso90][Satou94], also known as a binary search ADC, is a special type of subranging ADC that uses a DAC to produce an analog signal that approximates the input sample. By adjusting the DAC until the DAC output matches the input sample, a digital code representing the analog input can be generated. A successive approximation ADC consists of only one stage containing a sample and hold, an ADC, a DAC, and a digital processor that controls the DAC. An example of a successive approximation ADC is shown in the block diagram in figure 2.5. In the example shown, the ADC consists of a

single comparator. The operation of the successive approximation ADC is as follows. The control logic is initialized, and this initializes the output of the DAC. A sample of the input signal is taken by the sample and hold circuit, and the initial DAC output is subtracted from the input sample. The difference is quantized by the comparator which instructs the control logic to either increase or decrease the DAC output. The new DAC output is again subtracted from the input sample, and the process repeats until the desired accuracy is obtained. This single comparator successive approximation ADC resolves one bit per cycle.

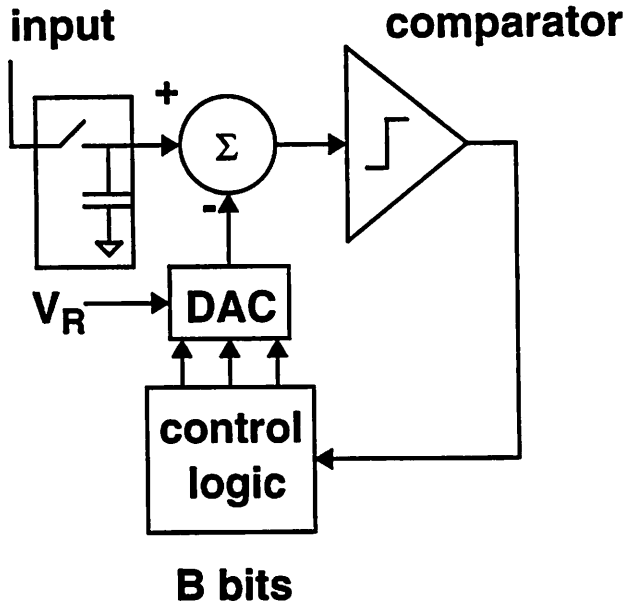


Figure 2.5 Successive Approximation ADC

This converter architecture has the advantage of using very little hardware. No amplifiers are required, and only a single comparator is required. By calibrating or trimming the DAC, very high resolution can be obtained. This technique has the disadvantage that the number of cycles required

per sample is proportional to the number of bits. Therefore, this converter architecture is slower than a flash.

2.5 Pipelined ADC

A pipelined ADC [Kyung80][Lewis87][Sutarja88][Lin91][Real91][Lewis92A][Lewis92B][Yotsuyanagi93][Conroy93][Sone93] [Colleran93] [Kusumoto93][Karanicolas93][Mercer94][Cho95][Nakamura95] is another type of subranging ADC that has features that improve the throughput rate and tolerance to comparator errors. The block diagram in figure 2.11 shows the structure of a pipelined ADC. The pipelined ADC is similar to the subranging ADC with the exception that a sample and hold circuit and amplifier has been added to each stage. The sample and hold circuit is used by the first stage to sample the input. Subsequent stages use a sample and hold to sample the residue from the previous stage. This feature allows each stage of the pipeline to begin processing a new sample as soon as its residue is sampled by the following stage. Thus, the throughput rate is independent of the number of stages in the pipeline. Because of this feature, pipelined ADCs can generally operate at much higher sampling rates than other subranging ADCs.

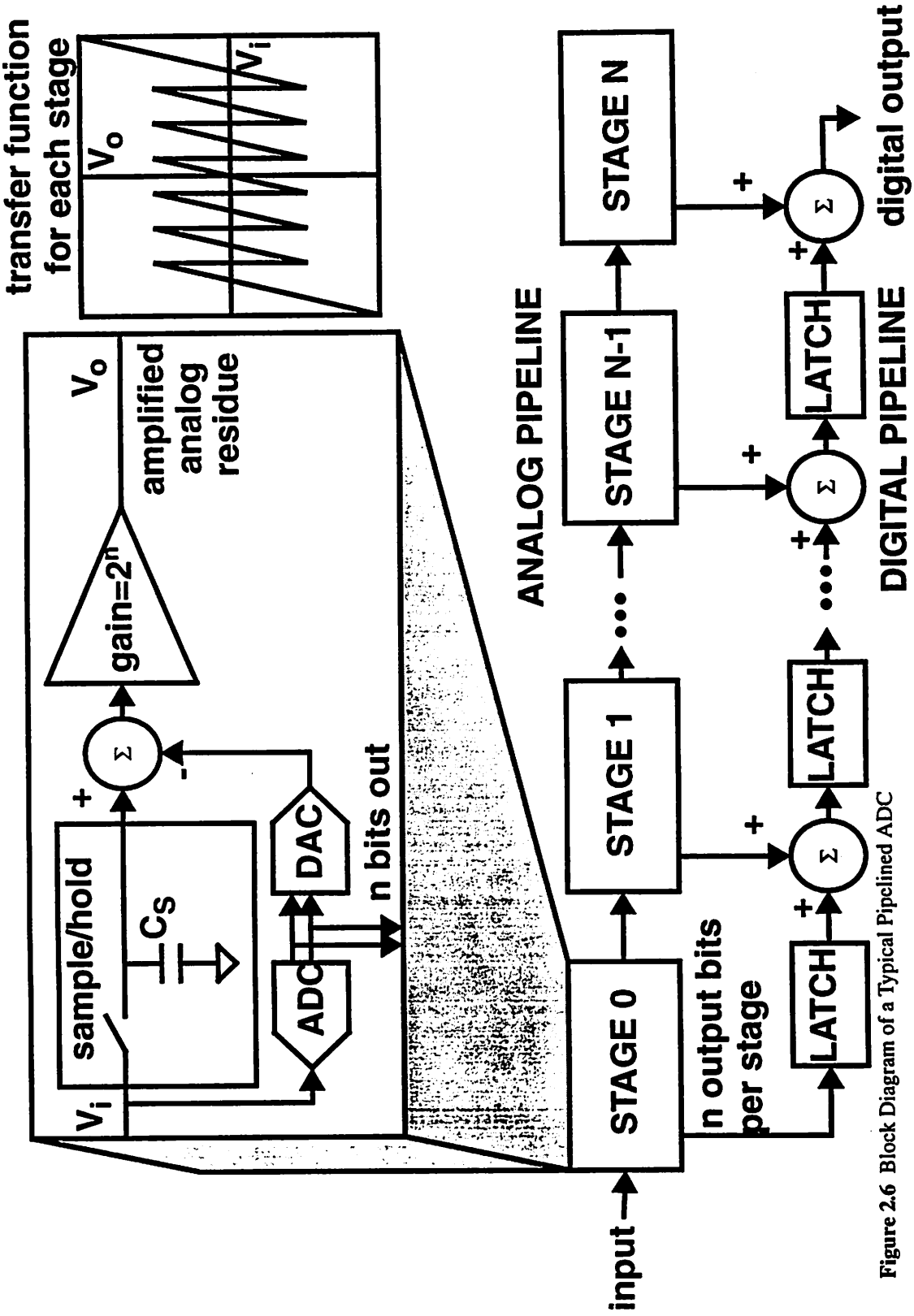


Figure 2.6 Block Diagram of a Typical Pipelined ADC

Although the throughput rate is independent of the number of stages in the pipeline, conversion time for any given sample is proportional to the number of stages in the pipeline. This is true because the signal must work its way through all of the stages before the complete output word is generated. This delay can be an issue if the pipelined ADC is part of a feedback system.

The amplifier is used to amplify the residue before passing it on to the next stage. By doing this, the resolution requirements for the following stages are relaxed. One significant implication of this is that the comparators in the last stages of the pipeline need not be accurate to the full ADC resolution as they are required to be in other subranging ADCs.

The disadvantage of adding the gain blocks is that they tend to be the dominant source of power dissipation in the ADC. Therefore, pipelined ADCs tend to dissipate more power than subranging ADCs. However, like the other subranging ADCs, pipelined ADCs can achieve high resolutions with relatively little hardware. Furthermore, mismatches can easily be eliminated as a limitation to resolution by self-calibration techniques.

Because of their tolerance to comparator offsets and the ability of the pipeline stages to operate in parallel, pipelined ADCs are well suited for high resolution applications where high speed is required.

2.6 Recirculating ADC

A recirculating ADC[Li84][Shih86][Ohara87][Onodera88][Song90] consists of a single pipeline stage with the output fed back to the input. The operation of a recirculating ADC is the same as for a pipeline except that one stage does all of the processing. A block diagram is illustrated in figure 2.7. The delay from input sample to complete digital output is the same as for a pipeline. However, the throughput rate is much less than for a pipeline because the entire digital word must be generated before a new sample can be taken. In spite of their low throughput rate, recirculating ADCs are very attractive for many applications because they use very little hardware and dissipate low power.

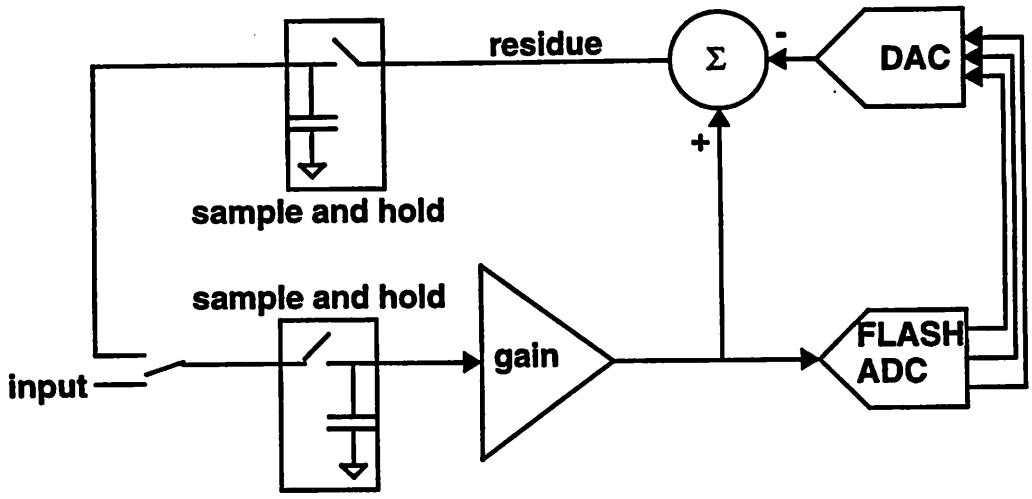


Figure 2.7 Recirculating ADC

2.7 Oversampled ADC

The oversampled ADC architecture[Hauser85][Candy85][Yamakido86][Koch86][Del Signore90][Brandt91A][Brandt91-B][Ribner91][Kerth92][Jantzi93][Fattaruso93][Ritoniemi94][Song95][van Gog95][Singor95] is another architecture that is capable of achieving high resolution with a small amount of hardware. A simple oversampled ADC is shown in figure 2.8.The operation is based on a sigma-delta modulator which repeatedly samples the input and performs a one bit quantization of the error between the signal and the estimate of the signal. By sampling the signal many times, errors due to the coarse quantization and noise are averaged out. This ADC architecture is used widely for audio applications, geophysical applications, and other low frequency applications requiring high dynamic range. Presently, the highest resolution ADCs use oversampling.

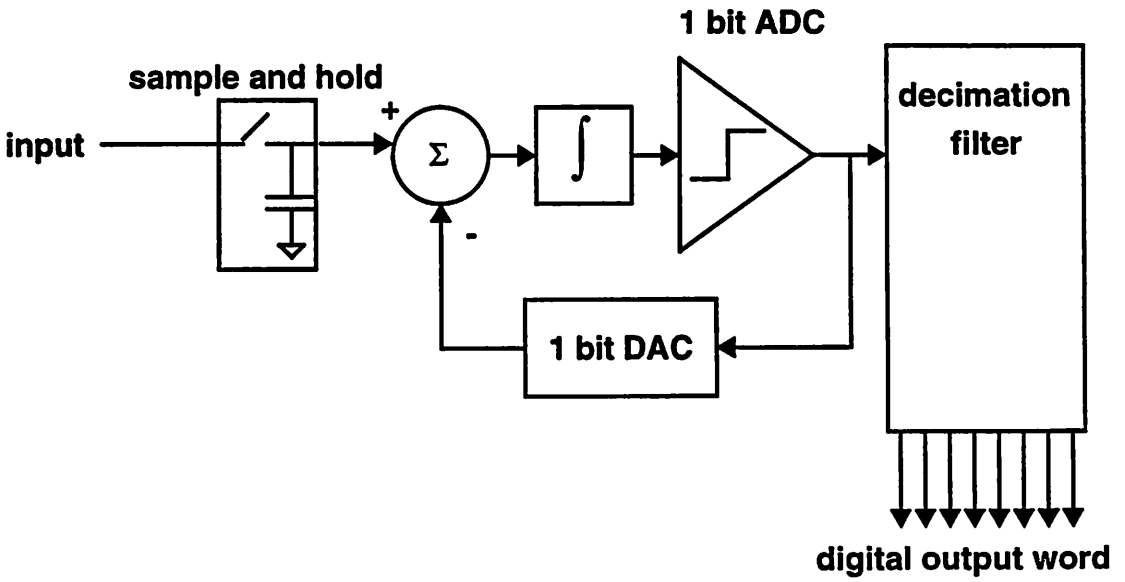


Figure 2.8 Simple Oversampled ADC

2.8 Serial ADC

The serial ADC architecture[Musa76][Smarandoiu76][Landsburg77][Masuda78] is the slowest type of ADC, but it is very simple and the resolution tends to be high. A simple serial ADC is shown in figure 2.9 This type of architecture compares the input signal to a ramp signal. A digital counter begins counting when the ramp crosses zero and then stops counting when the ramp becomes larger than the input signal. The final value for the counter is then equal to the digital code.

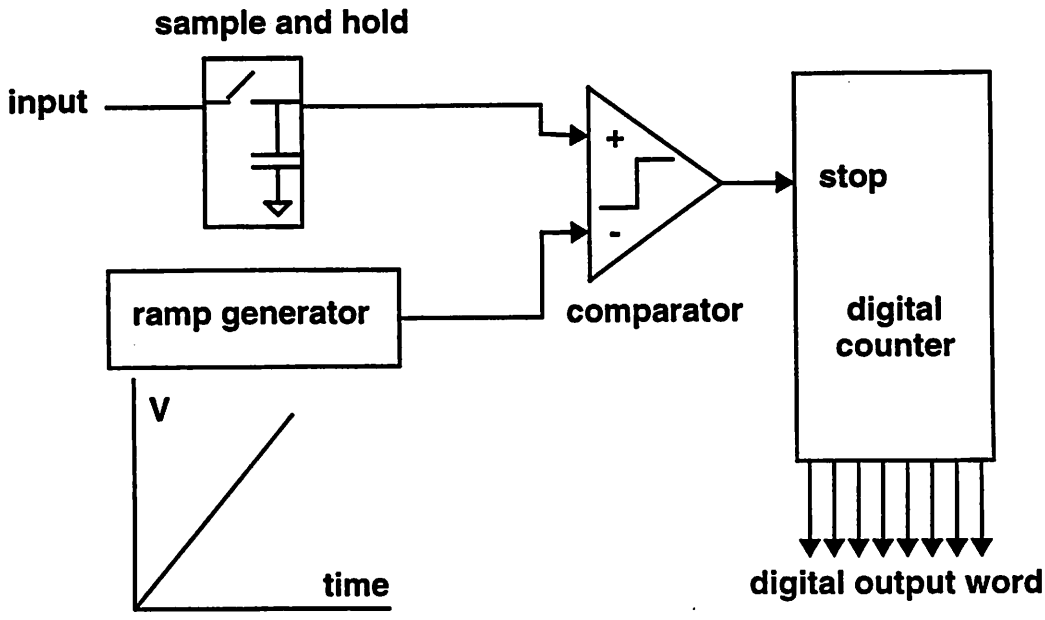


Figure 2.9 Serial ADC

2.9 Recent Performance Achievements

In this section, a number of recently published high resolution analog to digital converters are compared. The following table lists the major performance characteristics of each ADC.

The graph in figure 2.10 shows a scatter plot of normalized power versus dynamic range. This figure is of interest because it is an indicator of the inefficiency of analog to digital converters of a given resolution. As resolution increases, the power dissipation required to meet that resolution tends to increase because larger components are needed. It is assumed here that ADC power also tends to increase linearly with sampling rate and the minimum available channel length. In reality, this is a simplification because the power dissipation required becomes nonlinear as the speed capabilities of a process technology are pushed to the limit. Furthermore, ADCs are composed of a number of different types of circuits that scale differently in power as the process technology is advanced. For example, the power delay product in digital circuits potentially can scale with the

cube of the channel length if voltage is scaled with channel length.[Hodges88] The power delay product of an analog circuit that is limited by thermal noise issues, on the other hand, is likely not to scale with channel length. Therefore, it is difficult to estimate an appropriate relation between power dissipation, speed and channel length. However, if the ADC contains both digital circuits and noise limited analog circuits, and if an attempt is made to optimize the trade-off between the power in the analog circuits and the power in the digital circuits, then it is reasonable to expect that the ratio of power to sampling rate should scale at a rate comparable to the first power of the channel length. For the case of pipelined ADCs, an attempt to quantify some of these scaling issues is discussed in chapter 5. The conclusion of that analysis is that the ratio of power to sampling rate should scale approximately as the channel length to a power of 0.75.

In spite of these limitations to the comparison, this graph is still useful for observing some overall trends. The dynamic range used in this plot is the spurious free dynamic range reported in the publication of the ADC performance. In cases where this figure was unavailable, maximum SNR was used instead. From the plot, it can be seen that oversampled converters dominate the highest resolution applications. However, pipelines tend to be the most efficient at achieving a given resolution and sampling rate specification. Furthermore, because of the development of calibration techniques, pipeline resolution capabilities are improving and becoming competitive with oversampled converters in this respect.

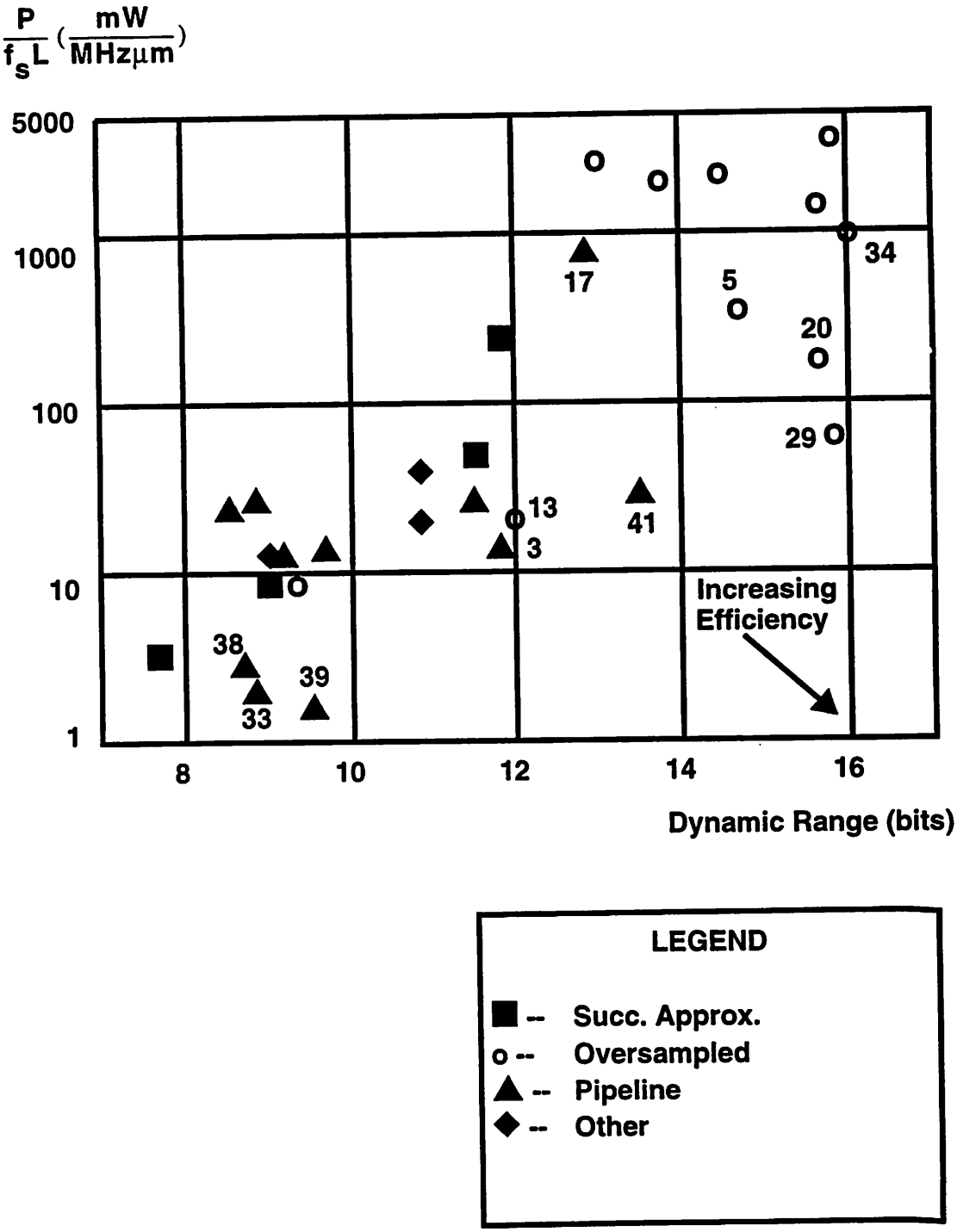


Figure 2.10 Normalized Power versus Dynamic Range

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
1	pipeline	Real	Analog Devices	JSSC Aug. 1991	2 µm BiCMOS	10	10	55		20	1000
2	sigma delta	Naus	Philips	JSSC July 1991				77	80		
3	pipeline	Lin	UC Berkeley	JSSC April 1991	3 µm CMOS	5	13	72	73	2.5	100
4	successive approx.	Ito	Mitsubishi	ISSCC Feb. 1994	1.2 µm CMOS	3	10	56		20	135
5	sigma delta	Dedic	Fujitsu	ISSCC Feb. 1994	1.2 µm CMOS		16	90		0.100	40

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
6	sigma delta	Alexander	National Semi	ISSCC Feb. 1994	1 µm BiCMOS	5		81	84.6	0.192	375
7	sigma delta	Matsuya	NTT	ISSCC Feb. 1994	0.5 µm CMOS	1	10	51	58	0.384	1.56
8	sigma delta	Ritoniemi	VLSI Solution	ISSCC Feb. 1994	1.2 µm BiCMOS	5	16	97		0.0441	180
9	2 step flash	Razavi	Stanford	JSSC Dec. 1992	1.0 µm CMOS	5	12	65	67	5	200
10	recirc.	Lee	Univ. of Illinois	JSSC Dec. 1992	2.0 µm CMOS	6.5	12	67		5	200

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
11	recirc.	Ginetti	Ginetti	JSSC July 1992	3.0 µm CMOS	10	13			0.025	45
12	pipeline	Lewis	AT&T Bell Labs	JSSC March 1992	0.9 µm CMOS		10	60		20	240
13	sigma delta	Brandt	Stanford	JSSC Dec. 1991	1.0 µm CMOS	5	12		74	2.1	41
14	successive approx.	Mercer	Analog Devices	JSSC Dec. 1991	2.0 µm BiCMOS	10	12	73		1.25	600
15	sigma delta	Ribner	General Electric	JSSC Dec. 1991	1.2 µm CMOS		15	90		0.080	

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
16	sigma delta	Lerch	Philips	JSSC Dec. 1991	3.0 µm CMOS	5	14		80	0.016	125
17	pipeline	Karanicolas	MIT	JSSC Dec. 1993	2.4 µm BiCMOS	8	15		79	1	1800
18	sigma delta	Fattaruso	Texas Instruments	JSSC Dec. 1993	1.0 µm CMOS	5			89	0.048	100
19	sigma delta	Nys	CSEM	JSSC July 1993	2.0µm CMOS	4	16			0.0001	0.4
20	sigma delta	Yin	Katholieke Univ.	JSSC June 1993	1.2 µm CMOS	5	16	93	96	0.320	65

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
21	sigma delta	Sarhang-Nejad	UCLA	JSSC June 1993	2.0 µm CMOS	5	16	95	96	0.0205	
22	flash	Kimura	Matsushita	JSSC April 1993	25 GHz bipolar	5.2	10	56		300	4000
23	successive approx.	deWit	Texas Instruments	JSSC April 1993	1.0 µm CMOS		12	71		0.200	10
24	pipeline	Yotsuyanagi	NEC	JSSC March 1993	0.8 µm CMOS	5	10	53		50	900
25		Zhang	Crystal	ISCAS 1994		2	12			1	60

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
26	pipeline	Lee	MIT	JSSC April 1994	1.6 µm CMOS	5	12			0.600	45
27	2 step flash	Miki	Mitsubishi	JSSC April 1994	0.8 µm BiCMOS	5	10	56		50	500
28	successive approx.	Yuan	Linkoping Univ.	JSSC Aug. 1994	1.2 µm CMOS	5	10	48		70	267
29	sigma delta	Yin	Katholieke Univ.	JSSC Aug. 1994	2.0 µm BiCMOS	5	16	87	97	1.5	180

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
30	pipeline	Macq	Alcatel-Bell	JSSC Aug. 1994	2.4 µm CMOS	5	10			0.550	20
31	pipeline	Sone	NEC	JSSC Dec. 1993	0.8 µm BiCMOS	5	10	57		100	950
32	pipeline	Colleran	UCLA	JSSC Dec. 1993	4 GHz Bipolar	10	10	59		75	800
33	pipeline	Kusumoto	Matsushita	JSSC Dec. 1993	0.8 µm CMOS	2.5	10	55		20	30

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
34	sigma delta	Williams	Stanford	CICC May 1993	1.0 µm CMOS	5	17	98		0.050	47
35		Shu	Univ. of Illinois	VLSI 1994	1.4 µm CMOS	5	13			10	350
36		Kumazawa	Sony	VLSI 1994		3	10			20	24
37		Chen	Univ. of Waterloo	VLSI 1994	1.2 µm CMOS	5	16	91	96	0.040	67.5
38	pipeline	Nakamura	Hitachi	CICC May 1994	0.8 µm CMOS	2.7	10	54		40	85

Table 2.2 Comparison of ADC Performance

ref	architecture	author	affil.	pub.	tech.	supply	res. bits	peak SNDR	dynamic range dB	sample rate MHz	power mW
39	pipeline	Cho	Univ. of California	CICC May 1994	1.2 µm CMOS	3.3	10	59.1		20	35
40	successive approx.	Satou	Toshiba	CICC May 1994	0.8 µm CMOS	3	12			1.0	1
41	pipeline	Cline	Univ. of California	CICC May 1995	1.2 µm CMOS	5	14	80	87	5	166

References

[Bacrania86] Kanti Bacrania, "A 12-Bit Successive-Approximation-Type ADC with Digital Error Correction," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 6, Dec. 1986, pp. 1016-1025.

[Brandt91A] Brian P. Brandt, Drew E. Wingard, and Bruce A. Wooley, "Second-Order Sigma-Delta Modulation for Digital-Audio Signal Acquisition," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, April 1991, pp. 618-627.

[Brandt91B] Brian P. Brandt and Bruce A. Wooley, "A 50-MHz Multibit Sigma-Delta Modulator for 12-b 2-MHz A/D Conversion," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, Dec. 1991, pp. 1746-1756.

[Boyacigiller81] Ziya G. Boyacigiller, Basil Weir, and Peter D. Bradshaw, "An Error-Correcting 14b/20μs CMOS A/D Converter," *ISSCC Dig. Tech. Papers*, Feb. 1981, pp. 62-63.

[Candy85] James C. Candy, "A Use of Double Integration in Sigma Delta Modulation," *IEEE Trans. Commun.*, vol. COM-33, no. 3, March 1985, pp. 249-258.

[Cho95] Thomas Byunghak Cho and Paul R. Gray, "A 10 b, 20 Msample/s, 35 mW Pipeline A/D Converter," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, March 1995, pp. 166-172.

[Colleran93] William T. Colleran and A. A. Abidi, "A 10-h, 75-MHz Two-Stage Pipelined Bipolar A/D Converter," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1187-1199.

[Connolly80] J. John Connolly, Thomas O. Redfern, Sing W. Chin, and Thomas M. Frederiksen, "A Monolithic 12b+Sign Successive Approximation A/D Converter," *ISSCC Dig. Tech. Papers*, Feb. 1980, pp. 12-13.

[Conroy93] Cormac S. G. Conroy, David W. Cline, and Paul R. Gray, "An 8-b 85-MS/s Parallel Pipeline A/D Converter in 1-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 447-454.

[Crolla82] Paul A. Crolla, "A Fast Latching Current Comparator for 12-Bit A/D Applications," *IEEE J. Solid-State Circuits*, vol. SC-17, no. 6, Dec. 1982, pp. 1088-1094.

[Croteau86] John Croteau, Don Kerth, and Dave Welland, "Autocalibration cements 16-bit performance," *Electronic Design*, Sept. 4, 1986, pp. 101-106.

[Del Signore90] Bruce P. Del Signore, Donald A. Kerth, Navdeep S. Sooch, Eric J. Swanson, "A Monolithic 20-b Delta-Sigma A/D Converter," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1311-1317.

[Dingwall79] Andrew G. F. Dingwall, "Monolithic Expandable 6 Bit 20 MHz CMOS/SOS A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-14, no. 6,

Dec. 1979, pp. 926-932.

[Dingwall85] Andrew G. F. Dingwall and Victor Zazzu, "An 8-MHz CMOS Subranging 8-Bit A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-20, no. 6, Dec. 1985, pp. 1138-1143.

[Doernberg89] Joey Doernberg, Paul R. Gray, David A. Hodges, "A 10-bit 5-Msample/s CMOS Two-Step Flash ADC," *IEEE J. Solid-State Circuits, vol. 24, no. 2, April 1989, pp. 241-249.*

[Fattaruso90] John W. Fattaruso, Michiel de Wit, Greg Warwar, Khen-Sang Tan, Richard K. Hester, "The Effect of Dielectric Relaxation on Charge-Redistribution A/D Converters," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1550-1561.

[Fattaruso93] John W. Fattaruso, Sami Kiriaki, Michiel de Wit, Greg Warwar, "Self-Calibration Techniques for a Second-Order Multibit Sigma-Delta Modulator," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1216-1223.

[Fernandes88] John Fernandes, Stephen R. Lewis, A. Martin Mallinson, Gerald A. Miller, "A 14-bit 10-μs Subranging A/D Converter with S/H," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1309-1315.

[Flynn95] Michael P. Flynn and David J. Allstot, " CMOS Folding ADCs with Current-Mode Interpolation," *ISSCC Dig. Tech. Papers*, Feb. 1995, pp. 274-275, 378.

[Fotouhi79] Bahram Fotouhi and David A. Hodges, "High-Resolution A/D Conversion in MOS/LSI," *IEEE J. Solid-State Circuits*, vol. SC-14, no. 6, Dec. 1979, pp. 920-926.

[Hadidi90] KH. Hadidi, Vincent S. Tso, Gabor C. Temes, "An 8-b 1.3-MHz Successive-Approximation A/D Converter," *IEEE J. Solid-State Circuits*, vol. 25, no. 3, June 1990, pp. 880-885.

[Hamade78] Adib R. Hamade, "A Single Chip All-MOS 8-Bit A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-13, no. 6, Dec. 1978, pp. 785-791.

[Hauser85] Max W. Hauser, Paul J. Hurst, Robert W. Brodersen, "MOS ADC-Filter Combination That Does Not Require Precision Analog Components," *ISSCC Dig. Tech. Papers*, Feb. 1985, pp. 80-81, 313.

[Hester90] Richard K. Hester, Khen-Sang Tan, Michiel de Wit, John W. Fattaruso, Sami Kiriaki, James R. Hellums, "Fully Differential ADC with Rail-to-Rail Common-Mode Range and Nonlinear Capacitor Compensation," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, Feb. 1990, pp. 173-183.

[Hodges88] David A. Hodges and Horace G. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2nd ed., New York: McGraw-Hill, 1988, pp. 102-104.

[Ishikawa89] Masayuki Ishikawa, Tsuneo Tsukahara, "An 8-bit 50-MHz CMOS Sub-

ranging A/D Converter with Pipelined Wide-Band S/H," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, Dec. 1989, pp. 1485-1491.

[Jantzi93] Stephen A. Jantzi, W. Martin Snelgrove, and Paul F. Ferguson, Jr., "A Fourth-Order Bandpass Sigma-Delta Modulator," *IEEE J. Solid-State Circuits*, vol. 28, no. 3, March 1993, pp. 282-291.

[Joy86] Andrew K. Joy, Robert J. Killips, and Peter H. Saul, "An Inherently Monotonic 7-Bit CMOS ADC for Video Applications," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 3, June 1986, pp. 436-440.

[Karanicolas93] Andrew N. Karanicolas, Hae-Seung Lee, and Kantilal L. Bacrania, "A 15-b 1-Msample/s Digitally Self-Calibrated Pipeline ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1207-1215.

[Kerth92] Donald A. Kerth and Douglas S. Piasecki, "An Oversampling Converter for Strain Gauge Transducers," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1689-1696.

[Koch86] Rudolf Koch, Bernd Heise, Franz Eckbauer, Eduard Engelhardt, John A. Fisher, and Franz Parzefall, "A 12-bit Sigma-Delta Analog-to-Digital Converter with a 15-MHz Clock Rate," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 6, Dec. 1986, pp. 1003-1010.

[Kolluri89] Madhav P. V. Kolluri, "A 12-bit 500-ns Subranging ADC," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, Dec. 1989, pp. 1498-1506.

[Kumamoto86] Toshio Kumamoto, Masao Nakaya, Hiroki Honda, Sotoju Asai, Yoichi, Akasaka, and Yasutaka Horiba, "An 8-bit High-Speed CMOS A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 6, Dec. 1986, pp. 976-982.

[Kusumoto93] Keiichi Kusumoto, Akira Matsuzawa, and Kenji Murata, "A 10-b 20-MHz 30-mW Pipelined Interpolating CMOS ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1200-1206.

[Kyung80] Chong-Min Kyung and Choong-Ki Kim, "Pipeline Analog-to-Digital Conversion with Charge-Coupled Devices," *IEEE J. Solid-State Circuits*, vol. SC-15, no. 2, April 1980, pp. 255-257.

[Landsburg77] George F. Landsburg, "A Charge-Balancing Monolithic A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-12, no. 6, Dec. 1977, pp. 662-673.

[Lewis87] Stephen H. Lewis and Paul R. Gray, "A Pipelined 5-Msample/s 9-bit Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 954-961.

[Lewis92A] Stephen H. Lewis, H. Scott Fetterman, George F. Gross, Jr., R. Ramachandran, T. R. Viswanathan, "A 10-b 20-Msample/s Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. 27, no. 3, March 1992, pp. 351-358.

[Lewis92B] Stephen H. Lewis, "Optimizing the Stage Resolution in Pipelined, Multi-

stage, Analog-to-Digital Converters for Video-Rate Applications," *IEEE Trans. Syst. II*, vol. 39, no. 8, Aug. 1992, pp. 516-523.

[Li84] Ping Wai Li, Michael J. Chin, Paul R. Gray, and Rinaldo Castello, "A Ratio-Independent Algorithmic Analog-to-Digital Conversion Technique," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 828-836.

[Lin91] Yuh-Min Lin, Beomsup Kim, and Paul R. Gray, "A 13-b 2.5-MHz Self-Calibrated Pipelined A/D Converter in 3-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, April 1991, pp. 628-636.

[Mangelsdorf90] Christopher W. Mangelsdorf, "A 400-MHz Input Flash Converter with Error Correction," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, Feb. 1990, pp. 184-191.

[Manoli89] Yiannos Manoli, "A Self-Calibration Metho for Fast High-Resolution A/D and D/A converters," vol. 24, no. 3, June 1989, pp. 603-608.

[Masuda78] Eiji Masuda, Chikara Sato, Tetsuya Iida, Yasoji Suzuki, Yasushi Agawa, and Toru Shima, "A Single-Chip C^2MOS A/D Converter for Microprocessor Systems," *ISSCC Dig. Tech. Papers*, Feb. 1978, pp. 134-135, 271.

[Mayes89] Michael K. Mayes and Sing W. Chin, "A Multistep A/D Converter Family with Efficient Architecture," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, Dec. 1989, pp. 1492-1497.

[McCall92] K. J. McCall, M. J. Demler, and M. W. Plante, "A 6-bit 125 MHz CMOS A/D Converter," in *Proc. CICC*, May 1992, pp. 16.8.1-16.8.4.

[McCreary75] James L. McCreary and Paul R. Gray, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques-Part I," *IEEE J. Solid-State Circuits*, vol. SC-10, Dec. 1975, pp. 371-379.

[Mercer91] Douglas A. Mercer, "A 12-b 750-ns Subranging A/D Converter with Self-Correcting S/H," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, Dec. 1991, pp. 1790-1799.

[Mercer94] Douglas A. Mercer, "A 14-b 2.5 MSPS Pipelined ADC With On Chip EPROM," *Proc. 1994 Bipolar/BiCMOS Circuits and Technology Meeting*, Sept. 1994, pp. 15-18.

[Musa76] Faud H. Musa and Robert C. Huntington, "A CMOS Monolithic 3.5-Digit A/D Converter," *ISSCC Dig. Tech. Papers*, Feb. 1976, pp. 144-145.

[Nakamura95] Katsufumi Nakamura, Masao Hotta, L. Richard Carley, David J. Allstot, "An 85 mW, 10 b, 40 Msample/s CMOS Parallel-Pipelined ADC," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, March 1995.

[Nauta95] Bram Nauta and Ardie G. W. Venes, " A 70Msample/s 110mW 8b CMOS Folding Interpolating A/D Converter," *ISSCC Dig. Tech. Papers*, Feb. 1995, pp. 276-277, 379.

[Ohara87] Harlan Ohara, Hung X. Ngo, Michael J. Armstrong, Chowdhury F. Rahim, and Paul R. Gray, "A CMOS Programmable Self-Calibrating 13-bit Eight-Channel Data Acquisition Peripheral," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 930-938.

[Onodera88] Hidetoshi Onodera, Tetsuo Tateishi, and Keikichi Tamaru, "A Cyclic A/D Converter That Does Not Require Ratio-Matched Components," *IEEE J. Solid-State Circuits*, vol. 23, no. 1, Feb. 1988, pp. 152-158.

[Peetz86] Bruce Peetz, Brian D. Hamilton, James Kang, "An 8-bit 250 Megasample per Second Analog-to-Digital Converter: Operation Without a Sample and Hold," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 6, Dec. 1986, pp. 997-1002.

[Petschacher] A 10-b 75-MSPS Subranging A/D Converter with Integrated Sample and Hold," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1339-1346.

[Razavi92] Behzad Razavi and Bruce A. Wooley, "A 12-b 5-Msample/s Two-Step CMOS A/D Converter," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1667-1678.

[Real91] Peter Real, David H. Roberson, Christopher W. Mangelsdorf, Theodore L. Tewksbury, "A Wide-Band 10-b 20-Ms/s Pipelined ADC Using Current-Mode Signals," *IEEE J. Solid-State Circuits*, vol. 26, no. 8, Aug. 1991, pp. 1103-1109.

[Redfern79] Thomas P. Redfern, Joseph J. Connolly, Jr., Sing W. Chin, and Thomas M. Frederiksen, "A Monolithic Charge-Balancing Successive Approximation A/D Technique," *IEEE J. Solid-State Circuits*, vol. SC-14, no. 6, Dec. 1979, pp. 912-920.

[Ribner91] David B. Ribner, Richard D. Baertsch, Steven L. Garverick, Donald T. McGrath, Joseph E. Krisciunas, and Toshiaki Fujii, "A Third-Order Multistage Sigma-Delta Modulator with Reduced Sensitivity to Nonidealities," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, Dec. 1991, pp. 1764-1774.

[Ritoniemi94] Tapani, Eero Pajarre, Seppo Ingalsuo, Timo Husu, Ville Eerola, and Tapio Saramaki, "A Stereo Audio Sigma-Delta A/D-Converter," *IEEE J. Solid-State Circuits*, vol. 29, no. 12, Dec. 1994, pp. 1514-1523.

[Satou94] Kouichi Satou, Kazuhiro Tsuji, Masayuki Sahoda, Hiroshi Otsuka, Kyoko Mori, and Totsuya Iida, "A 12 bit 1 MHz ADC with 1mW Power Consumption," *Proc. Custom Integrated Circuits Conf.,"* May 1994, pp. 23.6.1-23.6.4.

[Saul81] Peter H. Saul, "Successive Approximation Analog-to-Digital Conversion at Video Rates," *IEEE J. Solid-State Circuits*, vol. SC-16, no. 3, June 1981, pp. 147-151.

[Sheingold86] Daniel H. Sheingold, ed., *Analog-Digital Conversion Handbook*, 3rd ed.,

Englewood Cliffs, New Jersey: Prentice-Hall, 1986, pp. 218-219.

[Shimizu89] Toshihiko Shimizu, Masao Hotta, Kenji Maio, and Seiichi Ueda, "A 10-bit 20-MHz Two-Step Parallel A/D Converter with Internal S/H," *IEEE J. Solid-State Circuits, vol. 24*, no. 1, Feb. 1989, pp. 13-20.

[Shih86] Cheng-Chung Shih and Paul R. Gray, "Reference Refreshing Cyclic Analog-to-Digital and Digital-to-Analog Converters," *IEEE J. Solid-State Circuits*, vol. SC-21, no. 4, pp. 544-554.

[Singor95] Frank W. Singor and W. Martin Snelgrove, "Switched-Capacitor Bandpass Delta-Sigma A/D Modulation at 10.7 MHz," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, March 1995, pp. 184-192.

[Smarandoiu76] G. Smarandoiu, K. Fukahori, P. R. Gray, and D. A. Hodges, "An All-MOS Analog-to-Digital Converter Using a Constant Slope Approach," *IEEE J. Solid-State Circuits*, vol. SC-11, no. 3, June 1976, pp. 408-410.

[Sone93] Kazuya Sone, Yoshio Nishida, and Naotoshi Nakadai, "A 10-b 100-Msample/s Pipelined Subranging BiCMOS ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1180-1186.

[Song90] Bang-Sup Song, Seung-Hoon Lee, and Michael F. Tompsett, "A 10-b 15-MHz CMOS Recycling Two-Step A/D Converter," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1328-1338.

[Song95] Bang-Sup Song, "A 4th-Order Bandpass $\Delta\Sigma$ Modulator with Reduced Number of Opamps," *ISSCC Dig. Tech. Papers*, Feb. 1995, pp. 204-205, 367.

[Suarez75] Ricardo E. Suarez, Paul R. Gray, and David A. Hodges, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques-Part II," *IEEE J. Solid-State Circuits*, vol. SC-10, Dec. 1975, pp. 379-385.

[Sutarja88] Sehat Sutarja and Paul R. Gray, "A Pipelined 13-bit, 250-ks/s, 5-V Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1316-1323.

[Tan90] Khen-Sang Tan, Sami Kiriaki, Michiel de Wit, John W. Fattaruso, Ching-Yuh Tsay, W. Edward Matthews, Richard K. Hester, "Error Correction Techniques for High-Performance Differential A/D Converters," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1318-1327.

[Timko80] Michael P. Timko and Peter R. Holloway, "Circuit Techniques for Achieving High Speed-High Resolution A/D Conversion," *IEEE J. Solid-State Circuits*, vol. SC-15, no. 6, Dec. 1980, pp. 1040-1051.

[van de Grift87] Rob E. J. van de Grift, Ivo W. J. M. Rutten, and Martien van der Veen, "An 8-bit Video ADC Incorporating Folding and Interpolation Techniques," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 944-953.

[van de Plassche79] Rudy J. van de Plassche and Rob E. J. van der Grift, "A High-Speed 7 Bit

A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-14, no. 6, Dec. 1979, pp. 938-943.

[van de Plassche82] Rudy J. van de Plassche and Hans J. Schouwenaars, "A Monolithic 14 Bit A/D Converter," *IEEE J. Solid-State Circuits*, vol. SC-17, no. 6, Dec. 1982, pp. 1112-1117.

[van de Plassche88] Rudy J. van de Plassche and Peter Baltus, "An 8-bit 100-MHz Full-Nyquist Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1334-1344.

[van Gog95] Peter van Gog, Ben M. J. Kup, and Rob van Osch, "A Two-Channel 16/18b Audio AD/DA Including Filter Function with 60/40mW Power Consumption at 2.7V," *ISSCC Dig. Tech. Papers*, Feb. 1995, pp. 208-209, 368.

[van Valburg92] J. van Valburg and R. J. van de Plassche, "An 8-b 650-MHz Folding ADC," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1662-1666.

[Yamakido86] Kazuo Yamakido, Sigeo Nishita, Masaru Kokubo, Hirotoshi Shirasu, Ken'ichi Ohwada, Tatsuya Nishihara, "A Voiceband 15b Interpolative Converter Chip Set," *ISSCC Dig. Tech. Papers*, Feb. 1986, pp. 180-181.

[Yotsuyanagi93] Michio Yotsuyanagi, Toshiyuki Etoh, Kazumi 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.

CHAPTER 3

DESIGN TECHNIQUES FOR PIPELINED ANALOG TO DIGITAL CONVERTERS

This chapter discusses the operational principles of pipelined analog to digital converters (ADCs), sources of error in pipelined ADCs, and correction and calibration techniques for dealing with these errors. The calibration techniques are of particular interest because these techniques have relaxed the matching requirements of pipeline components. These relaxed matching constraints have made it possible to design pipelined ADCs with improved resolution. A second result of these calibration techniques is that it is now possible to design pipelined ADCs that are limited by thermal noise rather than matching considerations. As a result, it is possible to build pipelined ADCs with smaller components that dissipate relatively low power for a given sampling rate and resolution. Optimization techniques that take advantage of this fact and reduce power dissipation in thermal noise limited pipelined ADCs are discussed in chapter 5.

3.1 Introduction to the Concept of Pipelined ADCs

Pipelining is a method of speeding up high volume processes. Situations where pipelining is useful include consumer manufacturing systems that produce a high volume of products, such as automobiles, computers, and radios. Pipelining is also useful in sampled data signal processing systems that process a high volume of samples. High speed data acquisition is one example of an application.

When pipelining is applied to a task, the task is divided into a number of steps, each requiring an approximately equal length of time to perform. To each of these manufacturing or processing steps a performer is assigned to do that step. The steps are then ordered into a sequence. For each product to be manufactured or for each sample to be processed, these steps are performed in succession. When performer 1 completes step 1 of the sequence, that performer passes the product or sample to the next performer who performs step 2 of the pipelined process. At this point, per-

former 1 begins to perform step 1 on the next product or sample. Thus, the various processing steps can be performed concurrently, and this improves the rate at which products are produced, or in the case of signal processing systems, improves the rate at which samples are processed. On the other hand, a large delay exists from the time work is begun on a given product or sample until the work on that product or sample is complete. In other words latency is associated with pipelined systems, and this latency or delay is equal to the number of steps in the pipeline multiplied by the time required to execute the slowest step. An automobile assembly line is a good example of a pipelined manufacturing system.

A pipelined analog to digital converter is a good example of a pipelined signal processor. One example of a pipelined analog to digital converter is a pipelined subranging analog to digital converter. In a subranging ADC, the conversion operation is divided into a number of steps. During each step of the conversion, a certain number of bits of the digital output are resolved. The most significant bits are resolved in the first step, and the least significant bits are resolved in the last step. A pipelined subranging ADC is a subranging ADC which has a processor dedicated to each step of the conversion operation. In other words, a pipelined ADC contains a number of stages. Each stage of the pipeline is responsible for resolving some segment of the digital output word. In a pipelined ADC, the various stages of the pipeline operate concurrently. For example, as stage 3 processes sample 1, stage 2 processes sample 2, and stage 1 processes sample 3.

To further illustrate pipelined subranging analog to digital converters, the algorithm executed by a subranging analog to digital converter is very similar to the commonly known long division algorithm. This analogy will be used heavily in this chapter to help illustrate errors introduced by comparator offsets, gain errors, and nonuniform DAC levels. This analogy will also be used to describe techniques for correcting these errors.

In a long division problem, a dividend and divisor are given, and the goal is to determine the quotient. The procedure begins by comparing the most significant digit of the dividend with numbers in a multiplication table. This comparison allows the most significant digit of the quotient to be determined. Once this determination is complete, the appropriate entry from the multiplication table, in this case the largest entry smaller than the most significant digit of the dividend, is subtracted from the dividend to obtain a residue. This residue is then amplified (multiplied by 10 in a

decimal system) in order to obtain a value that lies within the range of the multiplication table. The comparison process is then repeated with the amplified residue in order to determine the second most significant digit of the quotient. The process continues until the residue is zero or until the desired precision is obtained.

A subranging analog to digital converter can be obtained by making the appropriate substitutions for elements in the above description. In a subranging analog to digital converter, the dividend is replaced by an analog input, the quotient by a digital output, and the multiplication table by a set of references. The subranging analog to digital conversion algorithm begins by comparing the analog input to the reference values in order to determine the most significant bits of the digital output. Once this decision has been made, the appropriate reference value is then subtracted from the analog input. This algorithm yields a residue as a result. This residue is then amplified so that the comparison process may again be done with the same set of references in order to determine the next most significant bits of the digital output. The process is repeated until the desired accuracy is obtained.

A typical pipelined analog to digital converter is a subranging converter having the basic comparison/subtraction/amplification block repeated. Thus, a pipelined analog to digital converter has one block for each time the processing step is repeated. Each of these blocks also includes a sample and hold circuit to hold the analog input signal or residue signal. This feature allows the first stage of the pipeline to perform a coarse quantization on a sample of the signal while the second stage processes the previous sample. In a subranging converter without pipelining, all the steps in quantizing a signal must be completely finished before the next sample can be taken. In a pipelined analog to digital converter, a higher throughput rate can be obtained because a new sample can be taken as soon as the first stage of the pipeline has finished processing the old sample.

Figure 3.1 shows a block diagram of a typical pipelined ADC

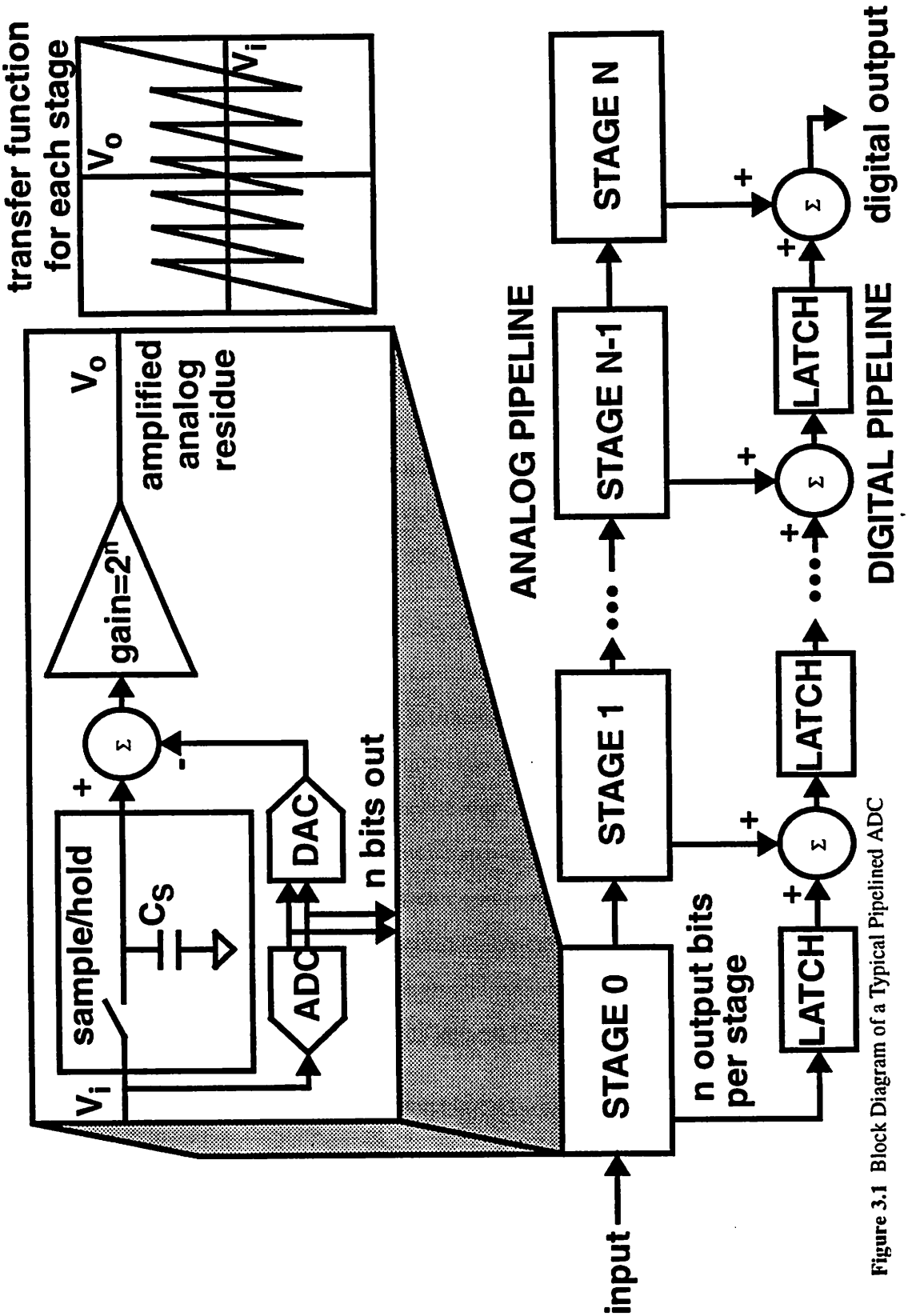


Figure 3.1 Block Diagram of a Typical Pipelined ADC

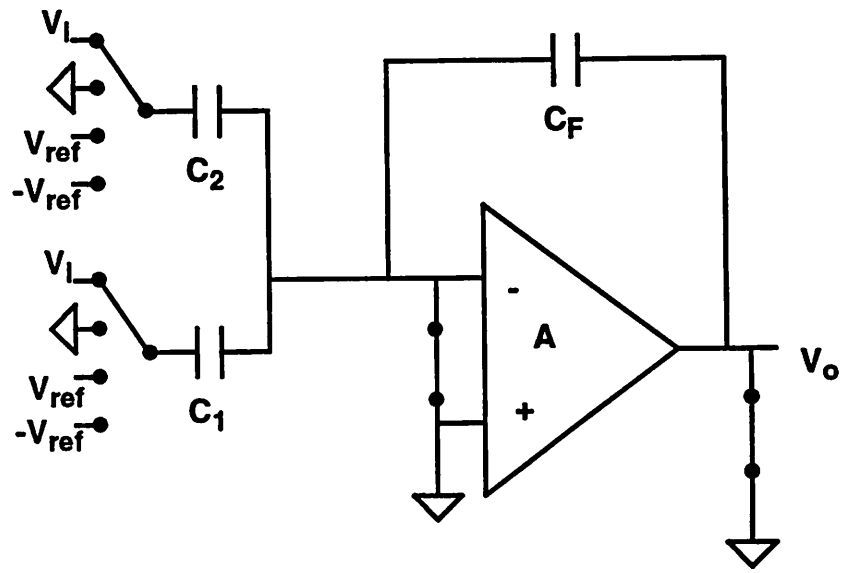
3.2 Switched Capacitor DAC and Residue Amplifier

This example illustrates one way of implementing the DAC and residue amplifier, a major component of the pipelined ADC shown in figure 3.1. A diagram of such a system is shown in figure 3.2. In the example shown here, the residue amplifier has a gain of two. The function of this circuit is threefold; to sample and hold the input signal, to generate a residue that is the difference between the input and some reference, and to amplify this residue.

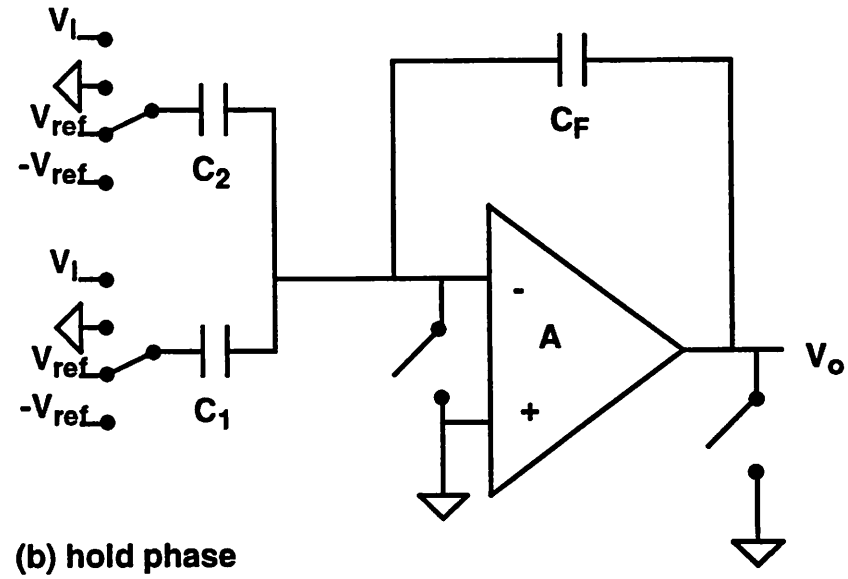
In this approach, the circuit operates on two phases, a sampling phase and a hold phase. During the sampling phase shown in figure 3.2a, the input signal is sampled onto the capacitors C_1 and C_2. During the hold phase the capacitors are then switched to one of three voltages, $+V_{ref}$, $-V_{ref}$, and ground. The voltage is chosen based on the digital output of the analog to digital converter block in figure 3.1. As the voltage is switched, the input voltage to the high gain amplifier, also known as the summing node voltage, tends to change. As it does, the output of the high gain amplifier changes a great deal. The negative feedback through the capacitor C_F drives this summing node voltage to zero. The result is that the charge initially stored on capacitors C_1 and C_2 is transferred to the capacitor C_F. For the case shown in the figure, the output voltage is a function of the input voltage and reference voltage.

$$V_o = \left(\frac{C_1 + C_2}{C_F}\right)V_I - \frac{C_1}{C_F}V_{ref1} - \frac{C_2}{C_F}V_{ref2} \tag{3.1}$$

V_{ref1} could be $+V_{ref}$, $-V_{ref}$, or 0 depending on where C_1 is connected during the hold phase. V_{ref2} is determined in a similar manner.



(a) sampling phase



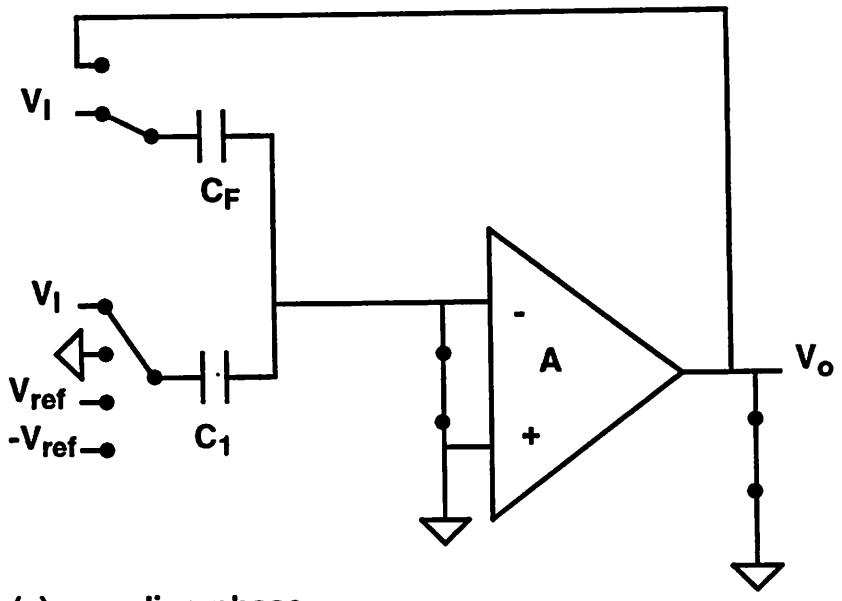
(b) hold phase

Figure 3.2 Operation of Switched Capacitor Sample/Hold Block with DAC, Subtraction, and Residue Amplifier Included

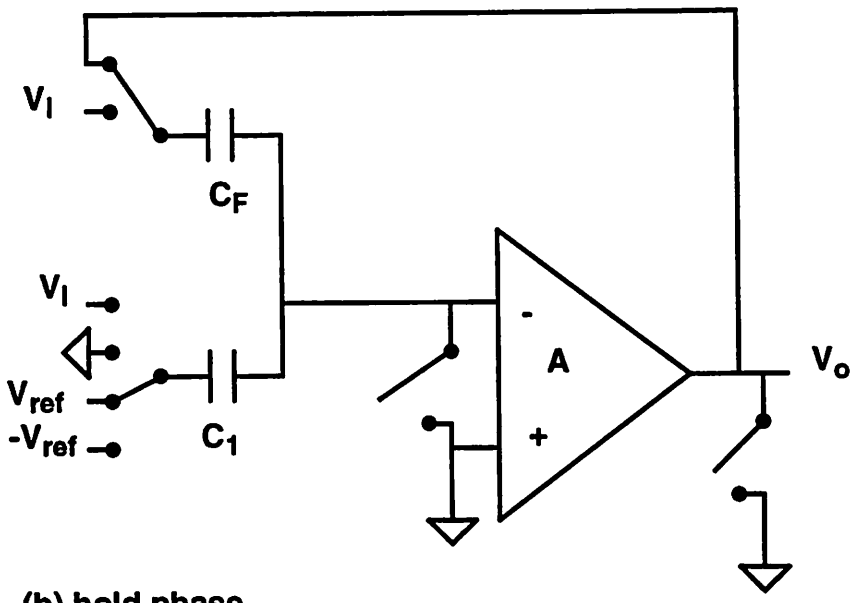
A modified version of this switched capacitor circuit is shown in figure 3.3. In this scheme, the feedback capacitor is used for sampling during the sampling phase since it would normally be idle at this time anyway. The modified circuit has the advantage that it uses fewer capacitors. Because there are fewer capacitors at the input of the amplifier, the feedback, and thus the speed, is improved. However, there is a drawback to this technique. During the hold phase, the switch in series with the feedback capacitor may degrade the settling speed. This degradation may cancel the improvement from the improved feedback factor.

The output voltage of this modified circuit is given by the following equation.

$$V_o = \frac{(C_1 + C_F)}{C_F} V_I - \frac{C_1}{C_F} V_{ref} \tag{3.2}$$



(a) sampling phase



(b) hold phase

Figure 3.3 Operation of Switched Capacitor Sample/Hold Block with Shared Feedback Capacitor

3.3 Sources of Error in Pipelined Analog to Digital Converters

In this section, some error sources affecting typical implementations of pipelined ADCs are discussed. These error sources have historically limited the performance of pipelined ADCs. These error sources may be divided into two categories; noise, which varies from sample to sample, and mismatches, which do not vary from sample to sample. This distinction has an important impact. Mismatch related errors can be corrected by calibration. Noise related errors, on the other hand, cannot be easily corrected by calibration.

The discussion in this section attempts to quantify the effect of some of the error sources on the performance of the ADC. In order to simplify the analysis, it is assumed in each of the sections below that a single error source acts alone in the absence of other errors. For example, when comparator offsets are discussed, noise and gain errors are assumed absent, and the DAC levels are assumed ideal. In a real pipelined ADC, a number of these errors could act simultaneously and this could result in compounding effects not predicted by this analysis.

3.3.1 Thermal Noise

Thermal noise is caused by the random motion of electrons. All particles at temperatures above absolute zero are in random motion. Since electrons carry charge, the thermal motion of electrons results in a random current that increases with temperature. This noise current is present in all circuits and corrupts any signals passing through. In a pipelined analog to digital converter, the sample and hold circuit is the most important source of noise. Within the sample and hold circuit, two noise sources are significant: the sampling switches and the sample and hold amplifier. The sampling switch is used to sample the input signal onto a sampling capacitor. As this happens, noise from the sampling switch is sampled with it onto the sampling capacitor. This source of thermal noise is commonly referred to as kT/C noise because the noise power is proportional to kT/C where C is the size of the sampling capacitor. The sample and hold amplifier also contributes thermal noise degradation to the signal being processed. The contribution of the sample and hold amplifier is also inversely proportional to a capacitance. In a single stage amplifier, it is inversely proportional to the load capacitance. In a Miller compensated amplifier it is inversely proportional to the compensation capacitance.

Thermal noise is perhaps the most fundamental source of error in a pipelined ADC. Because it is random from one sample to the next, it is not easily corrected by calibration. Thermal noise can be alleviated by using large components or by oversampling. However, for a fixed input bandwidth specification, both of these remedies increase the power dissipation. Thus, a fundamental trade-off exists between thermal noise, speed, and power dissipation. The impact of thermal noise on analog to digital converter design will be discussed further in chapter 5.

3.3.2 Comparator Offsets

The comparator is a fundamental part of an analog to digital converter. Its function is to make a comparison between two input signals and produces an output indicating which of the two inputs is larger. A comparator may be thought of as doing this by subtracting the two inputs and generating a binary output of 1 if the difference is positive and 0 if the difference is negative. Comparator nonidealities affect circuit performance in important ways. One of the most important nonideal characteristics is the offset of the comparator. When the comparator computes the difference of between the two input signals, an internal offset voltage is added to this difference. Thus, when the two inputs are close together, the comparator may make a wrong decision. When the comparator makes a wrong decision, the output code is wrong, and the wrong reference is subtracted from the input. The result is a residue that is out of range of the next stage of the pipeline when amplified. This idea may be illustrated by going back to the long division analogy. Consider an example where it is desired to divide the dividend 50 by the divisor 7. In the ideal case the process proceeds as shown below.

multiplication table

quotient digit	comparator thresholds	reference values
0	0	0
1	7	7
2	14	14
3	21	21
4	28	28
5	35	35
6	42	42
7	49	49
8	56	56
9	63	63

```
                    7.142857
                 7| 50.000000
subtracted ref     - 49
residue              1
amplified residue    1 0
subtracted reference - 7
residue                3
amplified residue      30
subtracted reference - 28
residue                 2
amplified residue       20
subtracted reference  - 14
residue                 6
amplified residue       60
subtracted reference  - 56
residue                 4
amplified residue       40
subtracted reference  - 35
residue                 5
amplified residue       50
subtracted reference  - 49
residue                 1
                        ...
```

Note that in long division the comparator threshold is equal to the reference value. Comparator offsets have the effect of moving the comparator thresholds. The example below illustrates long division with bad decision thresholds (comparator offsets). Note that in the multiplication table the reference values used in the subtractions are still correct, but the values used to make the decisions are wrong.

multiplication table

quotient digit	comparator thresholds	reference values
0	0	0
1	8	7
2	13	14
3	19	21
4	26	28
5	36	35
6	41	42
7	50	49
8	55	56
9	66	63

incorrect output digit

```
                    |
                 7.143000...
              7| 50.000000
subtracted ref  - 49
residue           1
amplified residue 1 0
subtracted reference - 7
residue             3
amplified residue   30
subtracted reference - 28
residue              2
amplified residue    20
subtracted reference - 21 <wrong reference subtracted
residue              - 1
amplified residue    - 10<amplified residue out of
                         range
subtracted reference  - 0
residue              -10
amplified residue    -100
subtracted reference  - 0
residue              -100
amplified residue    -1000
subtracted reference   - 0
residue              -1000
                        ...
```

Note that after the error is made, the amplified residue -10 is out of range of the multiplication table. The result is that the residue continues to get amplified at each subsequent stage with the same output code from each stage. The accuracy of the result is limited to the accuracy of the quotient at the stage where the error was made.

Figure 3.4 shows a plot of the amplified residue versus the input where comparator errors are present. This plot further illustrates the effect of these errors on the ADC output. The ideal comparator switching thresholds occur where the thin vertical lines are. The actual comparator switching thresholds are located at the heavy vertical lines. Note that the maximum amplified residue is greater than V_{max} for the case where comparator offsets are present. This overrange situation causes information about the signal to be lost.

Some techniques for reducing comparator offsets have been developed. Capacitive interpolation is one example of such a technique[Kusumoto93]. Furthermore, it is common practice to alleviate the effects of comparator offsets by adding extra comparators to each stage of the pipeline. This technique, called digital error correction, helps keep the maximum amplified residue in range and

is discussed further later in this chapter.

In addition to comparator offsets, the transfer characteristic in figure 3.4 has a second problem. The nominal transfer characteristic has a switching threshold at V_I=0. This trait is undesirable because it means that all of the ADC output bits are likely to change nearly simultaneously as the input signal makes a transition from a negative to positive value. This trait tends to result in a high degree of distortion and noise for small input signals that are biased closed to zero.

Figure 3.5 shows a modified transfer characteristic where extra comparators have been added to bring the maximum output signal in range. The comparator thresholds have also been shifted in this figure so that there is no threshold near V_I=0. This figure is also used to illustrate the effect of gain errors discussed in the next section.

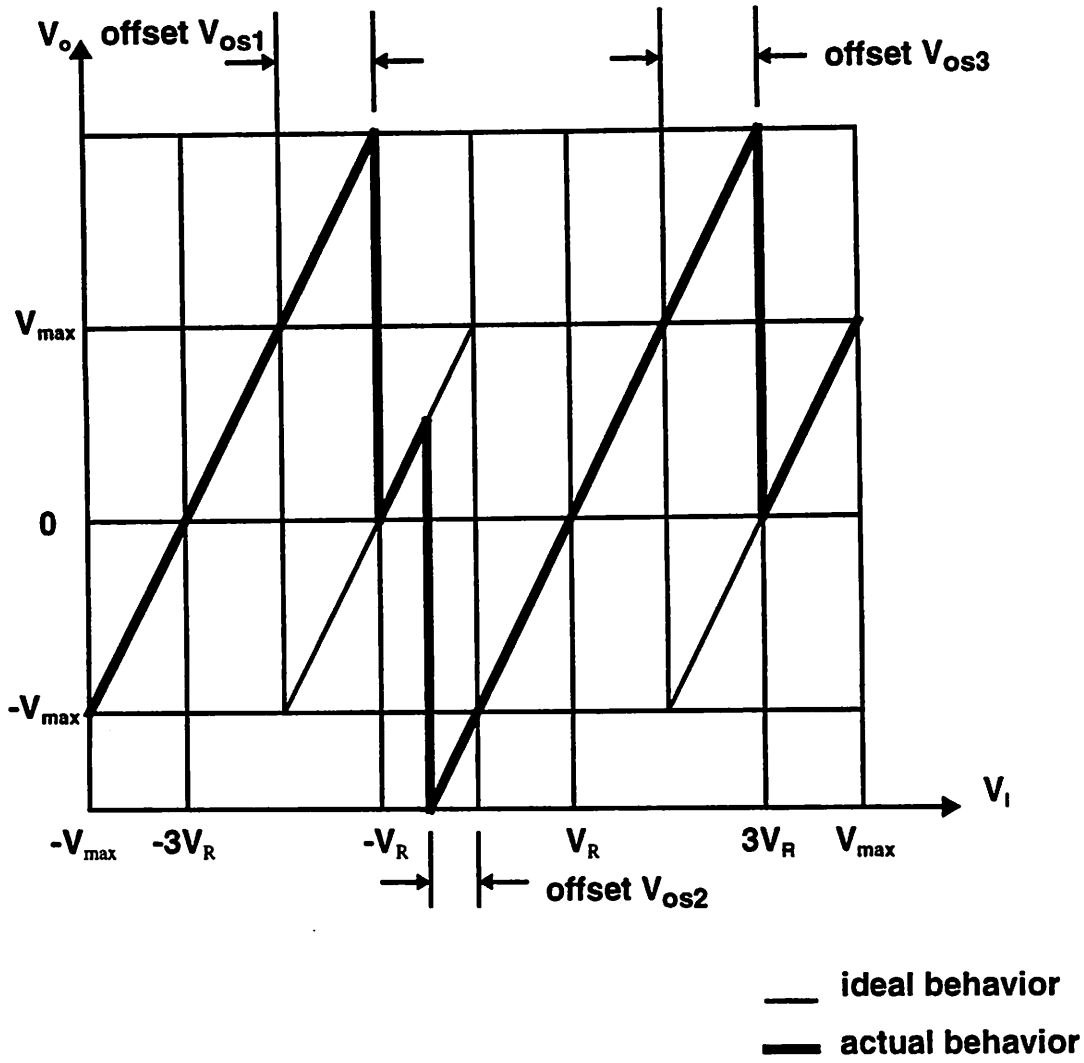


Figure 3.4 Example of the Effect of Comparator Offsets on the Transfer Characteristic of a Pipeline Stage, $V_{os1}=-V_R$, $V_{os2}=0.5V_R$, $V_{os3}=-V_R$

3.3.3 Residue Amplifier Gain Error

In a pipelined analog to digital converter, errors in the per stage gain can be caused by a variety of sources. In a pipeline using switched capacitor gain stages these error sources could include capacitor mismatches and noninfinite opamp gains. An error in the per stage gain causes nonlinearity in the transfer characteristic from input to output of the pipelined ADC. This effect may be

illustrated by thinking about the long division analogy. Pipelined analog to digital conversion with a gain error is analogous to long division where the residue of each step is multiplied by some number other than 10. In both cases, the multiplier is different from the assumed radix. If a plot of the quotient versus the dividend is made for the long division case where the multiplier is different from 10, discontinuities will show up in the plot. If the multiplier is greater than 10, the plot displays nonmonotonicity. If the multiplier is less than 10, some quotient values are never obtained. In an ADC, this is analogous to missing codes.

The effect of gain error in a pipelined ADC on the differential nonlinearity (DNL) may be determined graphically as follows. In this analysis, the comparators are assumed to be ideal, reference levels are assumed to be uniform, and the gain is assumed to be linear. Linear gain error is the only nonideality treated in this analysis. Figure 3.5 shows a plot of the amplified residue versus the input for one stage of the analog to digital converter. It is assumed that digital error correction is used to deal with the comparator offset problem. Therefore, this transfer characteristic has twice as many comparator thresholds as figure 3.4, and as a result the nominal output swing is smaller. The gain error shown produces jumps in the input output transfer characteristic of the ADC as shown in figure 3.6.

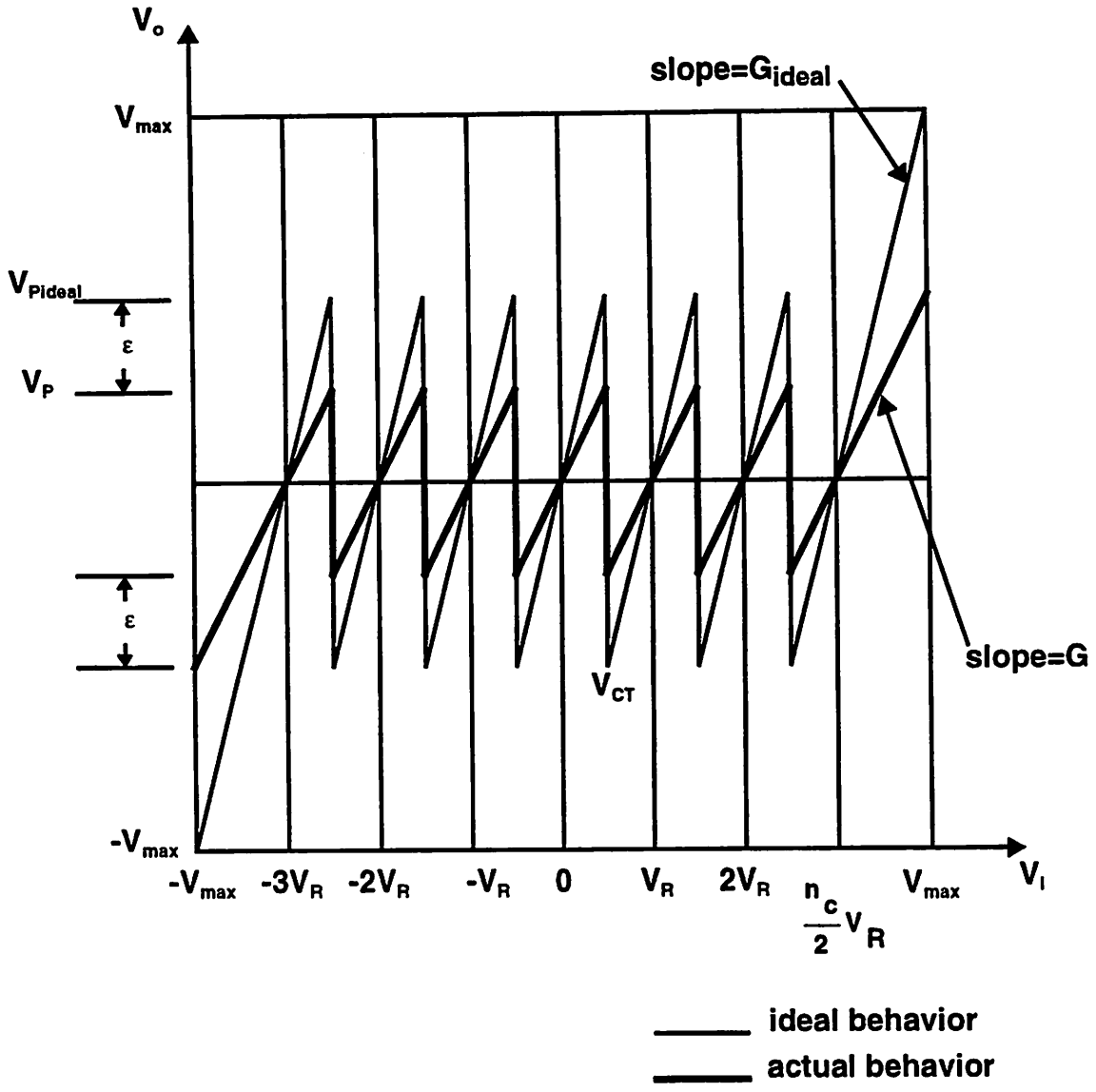


Figure 3.5 Effect of Residue Amplifier Gain Error on the Transfer Characteristic of a Pipeline Stage (In the example shown figure, the actual gain is smaller than the nominal value.)

By the following analysis, an algebraic relation between the differential nonlinearity (DNL) and the gain error can be determined[Conroy90]. It is assumed here that the input range is equal to the output range. This analysis also assumes uniform DAC reference levels as shown in figure 3.5.

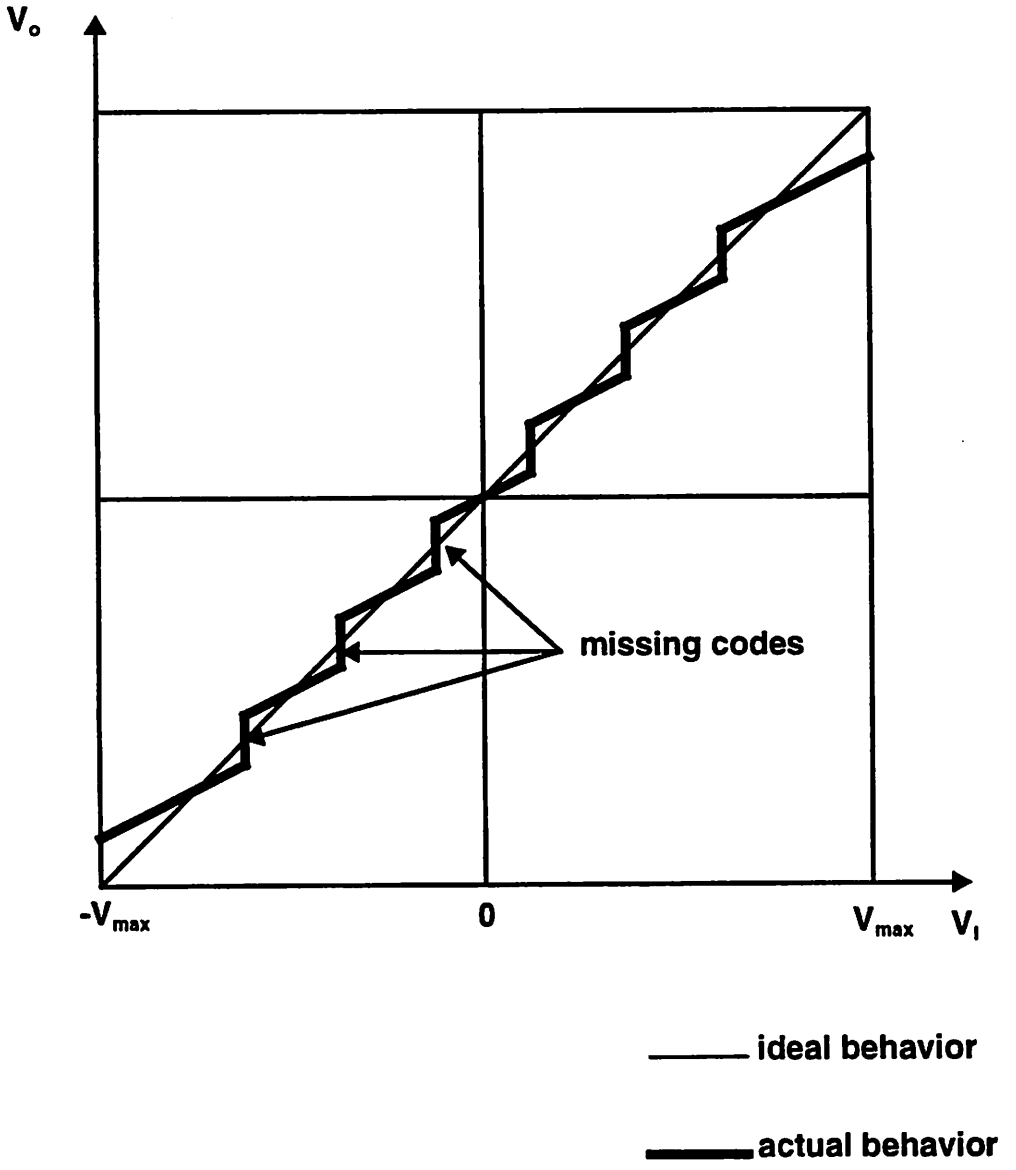


Figure 3.6 Effect of Residue Amplifier Gain Error on the ADC Transfer Characteristic
(In the figure, the actual gain is smaller than the nominal value.)

Furthermore, the comparator thresholds are also assumed here to be uniform. The implication of these assumptions of uniform reference levels and comparator thresholds is that the DNL is the same for each transition point.

First, the relationship between the reference level V_R and the number of comparators is determined. V_R is the DAC reference level subtracted from the input to obtain the residue. From figure

3.5,

$$V_{max} = G_{ideal}(V_{max} - \frac{n_C}{2}V_R) \tag{3.3}$$

$$n_C = \text{number of' comparators per stage} \tag{3.4}$$

V_{max} is the maximum peak swing of the input signal. Because the output swing is assumed to be equal to the input swing in this case, V_{max} is also the peak output swing. By rearranging equation (3.3), the following equation is obtained.

$$V_R = \frac{2}{n_C}(1 - \frac{1}{G_{ideal}})V_{max} \tag{3.5}$$

V_R can also be related to the peak value V_P of the residue by the gain by using figure 3.5 as shown below.

$$V_{Pideal} = G_{ideal}\frac{1}{2}V_R \tag{3.6}$$

$$V_P = G\frac{1}{2}V_R \tag{3.7}$$

The error voltage ε at a comparator threshold can be expressed as the difference between V_P and V_{Pideal} as follows.

$$\varepsilon = V_{Pideal} - V_P \tag{3.8}$$

Now substitute (3.5), (3.6), and (3.7) into (3.8) in order to obtain the following expression for the error voltage. as a function of gain error.

$$\varepsilon = \left(\frac{G_{ideal} - G}{G_{ideal}}\right)\left(\frac{G_{ideal} - 1}{n_C}\right)V_{max} \tag{3.9}$$

Now, to determine the DNL, note that the amplified residue just to the right of the threshold is too large by ε while the amplified residue just to the left of the threshold is too small by ε. Therefore, the magnitude of the mismatch between these two points is 2ε. To refer the error back to the input, the error is divided by the ideal gain of the stage with the error and the ideal gain of every stage proceeding it. Therefore, the input referred discontinuity in the transfer characteristic of the

ADC is given as follows.

$$\Delta = \frac{2\varepsilon}{G_{ideal}^{k+1}} = \left(\frac{G_{ideal} - G}{G_{ideal}}\right)\left(\frac{G_{ideal} - 1}{n_C}\right)\frac{2V_{max}}{G_{ideal}^{k+1}} \tag{3.10}$$

The parameter k is the number of gain stages preceding the one with the error. The magnitude of the DNL as a fraction of full scale can then be determined by dividing the input referred value of the discontinuity by the full scale swing of the input. The DNL in LSBs is then found by multiplying by the number of quantization levels in the ADC. In the result shown below, N is the number of bits resolved by the ADC.

$$\text{DNL (fraction of full scale)} = \frac{\Delta}{2V_{max}} \tag{3.11}$$

$$\text{DNL} = \left(\frac{G_{ideal} - G}{G_{ideal}}\right)\left(\frac{G_{ideal} - 1}{n_C}\right)\frac{2^N}{G_{ideal}^{k+1}} \tag{3.12}$$

The DNL caused by the first stage (k=0) is the largest and is given by the following formula.

$$\text{DNL (LSBs)} = \left(\frac{G_{ideal} - G}{G_{ideal}}\right)\left(\frac{G_{ideal} - 1}{n_C}\right)\frac{2^N}{G_{ideal}} \tag{3.13}$$

Note from the above equation that the effect of gain error on DNL can be reduced if the number of comparators is increased. It is also possible to use calibration to alleviate this problem. Calibration techniques capable of correcting the gain error are discussed later in this chapter.

An example illustrating the use of equation 3.4 is now shown to provide some numerical intuition. Commonly, pipelined ADCs are built such that $n_C=2(G_{ideal}-1)$. With this assumption, $G_{ideal}=4$, and $(G_{ideal}-G)/G_{ideal}=1/2^N$, the DNL is 0.125LSB.

3.3.4 Nonuniform Reference Levels (Nonlinear DAC)

Nonuniform spacing of the DAC reference levels also contributes to the nonlinearity of the analog to digital converter. The DAC reference levels are the signals subtracted from the input

sample to generate a residue. The DAC reference levels are often controlled by a resistor string or capacitor array. Nonuniformity is caused in this case by mismatches in the resistors or capacitors. The effect of this type of error on the amplified residue is shown graphically in figure 3.7. In this

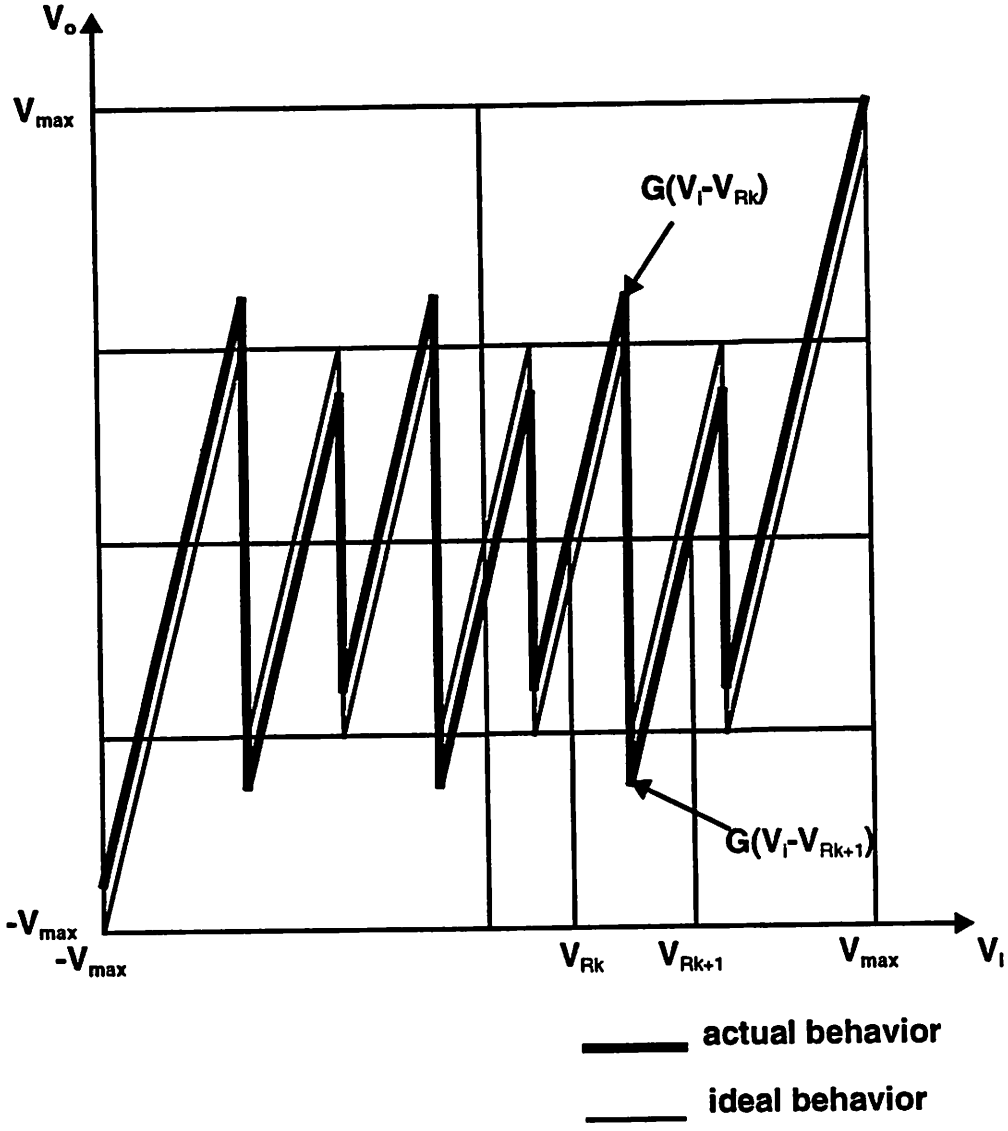


Figure 3.7 Effect of DAC Nonlinearity on the Transfer Characteristic of a Pipeline Stage

figure, the height of the kth discontinuity is determined by the reference values as shown below.

$$\text{step height}_k = G(V_{Rk+1} - V_{Rk}) \tag{3.14}$$

In the above equation, G is the residue amplifier gain discussed in the previous section.

The long division analogy may also be used here to demonstrate the effect of nonuniform reference values on the output. In this case, the threshold values in the multiplication table are correct, but the reference values subtracted from the dividend to generate the residue are not correct.

quotient digit	comparator thresholds	reference values
0	0	0
1	7	8
2	14	13
3	21	19
4	28	26
5	35	36
6	42	41
7	49	48
8	56	55
9	63	66

```
                    7.295555...
                 7| 50.000000
subtracted ref     - 48
residue              2
amplified residue    2 0
subtracted reference -1 3
residue                7
amplified residue      70
subtracted reference   - 66
residue                 4
amplified residue       40
subtracted reference    - 36
residue                 - 4
amplified residue       -40
subtracted reference    -36
residue                 - 4
amplified residue       - 40
subtracted reference     -36
residue                    4
amplified residue          40
subtracted reference       -36
residue                      4
                             ...
```

In the above example the use of incorrect references in the multiplication table leads to an incorrect quotient.

Reference level nonuniformity in an ADC is caused by component mismatches. In order to explain the source of reference nonuniformity in a circuit implementation of an ADC, examples of

techniques of generating the references will now be described.

One architecture, called a resistor string DAC, consists of a set of series connected resistors between two well defined reference voltages as shown in figure 3.8.

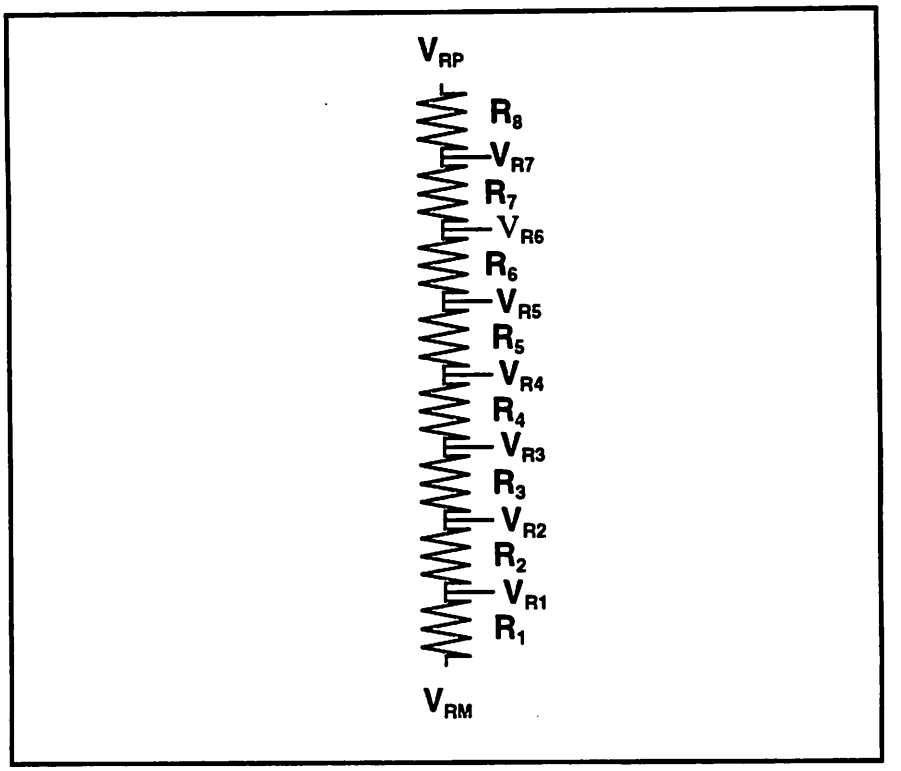


Figure 3.8 Resistor String DAC

In this architecture, the reference voltages are given as shown below.

$$R = \sum_{i=1}^{7} R_i \tag{3.15}$$

In the above equation, R is the total resistance of the resistor string. R_i is the resistance of the ith element of the string.

$$V_{Ri} = V_{RM} + (V_{RP} - V_{RM}) \frac{\sum_{j=1}^{i} R_j}{R} = V_{RM} + \sum_{j=1}^{i} w_j \tag{3.16}$$

$$w_j = (V_{RP} - V_{RM}) \frac{R_j}{R} \tag{3.17}$$

In the above equations, V_{RM} is the reference voltage at the negative end of the resistor string, and V_{RP} is the reference voltage at the positive end of the resistor string. The w_j in the above expression are the weights associated with each segment of the DAC, in this case each resistor in the resistor string. The step heights in the transfer characteristic shown in figure 3.7 are related to the DAC weights as shown below.

$$\text{step height}_j = Gw_j \tag{3.18}$$

G in the above expression is the gain of the residue amplifier.

Ideally each of these resistors would be identical, and thus each of these w_j would match. In reality, these resistors do not match perfectly due to process variations. The process variations cause the size and shape of the resistors to vary from one to the next. Voltage dependence of the resistor values can also contribute to this error. To correct errors due to mismatches in these resistors, laser trimming is frequently used to adjust the resistance of each resistor.

Another architecture often used to implement the reference voltages uses switched capacitors that can be switched between the input, a reference voltage, and ground as shown in figure 3.9. This circuit realizes the sample and hold, the reference voltages, and the subtractor circuit. Charge redistribution on the set of capacitors is used to generate a residue. The circuit performs its operation in two phases as follows. During the first phase, called the sample phase, the four switches on the left side connect the input to the left sides of the capacitors while the switch on the right ties V_S to ground. At this time, the charge $Q = -C_T V_I$ is stored at node V_S where $C_T = C_1 + C_2 + C_3 + C_4$. During the second phase, called the hold phase, the right hand switch is opened. This leaves V_S without a DC path to any other nodes in the circuit. During this phase, the left hand switches connect the left sides of the capacitors to either V_R or ground. Because V_S

has no DC path to ground, the charge at this node does not change, but remains $Q = -C_T V_I$. By charge conservation, the voltage V_S may be determined as follows.

$$V_S = \frac{-C_T}{(C_T + C_{IP})} \left(V_I - V_R \frac{\sum_{j=1}^{i} C_j}{C_T} \right) \tag{3.19}$$

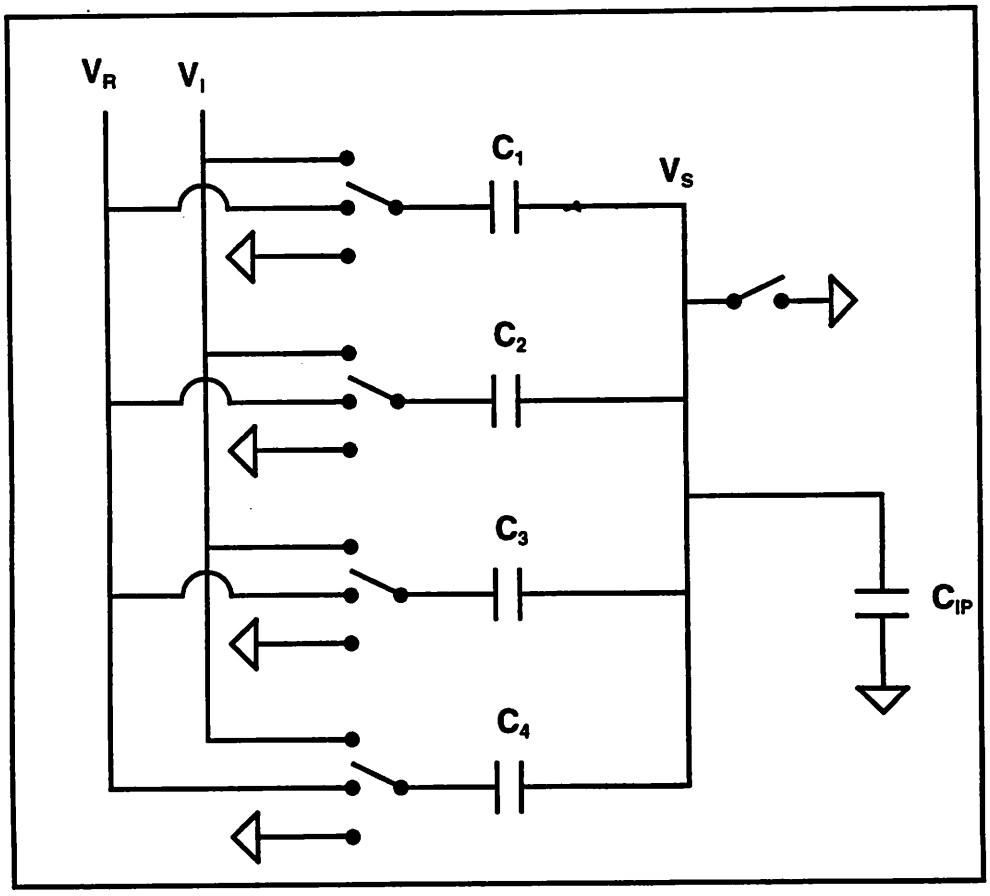


Figure 3.9 Switched Capacitor DAC

In this architecture, the reference voltages V_{Ri} and the segment weights w_j are given as fol-

lows.

$$V_{Ri} = V_R \sum_{j=1}^{i} \frac{C_j}{C_T} \tag{3.20}$$

$$w_j = V_R \frac{C_j}{C_T} \tag{3.21}$$

Now, a quantitative relationship between the nonlinearity of the reference levels and the DNL will be developed. In general, a reference voltage V_{Ri} can be represented as the sum of a set of segment weights as shown below.

$$V_{Ri} = \sum_{j=1}^{i} w_j \tag{3.22}$$

The error $\varepsilon_{V_{Ri}}$ may be expressed as the difference between the actual V_{Ri} and the value V_{Ri} would have if all the weights were equal.

$$\varepsilon_{V_{Ri}} = \sum_{j=1}^{i} w_j - \frac{i}{n} \sum_{j=1}^{n} w_j \tag{3.23}$$

In the above equation, n is the number of weights per pipeline stage.

$$\varepsilon_{V_{Ri-1}} = \sum_{j=1}^{i-1} w_j - \frac{(i-1)}{n} \sum_{j=1}^{n} w_j \tag{3.24}$$

The size ε of the discontinuity is equal to the difference in the errors in two adjacent references.

$$\varepsilon = \varepsilon_{V_{Ri}} - \varepsilon_{V_{Ri-1}} = \left(\frac{n-1}{n}\right) w_i - \frac{1}{n} \sum_{j \neq i} w_j \tag{3.25}$$

Thus, the standard deviation σ_ε of the discontinuity may be expressed by the following equation. To obtain this equation, it is necessary to assume that the variations in the weights are uncorrelated.

$$\sigma_\varepsilon^2 = \left(\frac{n-1}{n}\right) \sigma_w^2 + \left(\frac{n-1}{n}\right) \sigma_w^2 = 2\left(1 - \frac{1}{n}\right) \sigma_w^2 \tag{3.26}$$

The discontinuity may be referred to the input by dividing by the number of gain stages preceding the stage where the error is made. The DNL is then found by multiplying this input referred DNL by the ratio of the number 2^N of quantization levels to the full scale range $2V_{Imax}$. Thus, the standard deviation of the DNL is found as shown below.

$$\sigma_{DNL}(\text{LSBs}) = \frac{\sigma_\varepsilon}{G}\left(\frac{2^N}{2V_{Imax}}\right) = \frac{\frac{\sigma_w}{w} w\sqrt{2\left(1-\frac{1}{n}\right)}}{G^k}\left(\frac{2^N}{2V_{Imax}}\right) \tag{3.27}$$

The weight w in this case is equal to V_R as given by equation (3.5). Thus, the DNL can be rewritten as shown below. It is assumed here that the number of independent references is equal to the number of comparators divided by 2 because this is the case in a common implementation of pipelined ADCs.

$$\sigma_{DNL}(\text{LSBs}) = \frac{\frac{\sigma_w}{w}\sqrt{2\left(1-\frac{2}{n_c}\right)}}{G^k}\left(\frac{2^N}{n_c}\right)\left(1-\frac{1}{G_{ideal}}\right) \tag{3.28}$$

Note that for a given $\frac{\sigma_w}{w}$ increasing the number of comparators n_c reduces the DNL. However, increasing the number of comparators also has the disadvantage of increasing the area of each pipeline stage.

3.3.5 Residue Amplifier Nonlinearity

Ideally, the residue amplifier produces an output that is proportional to the input. In a real implementation, the amplifier is not perfectly linear, but introduces some distortion to the signal. Often, this nonlinearity takes the form of a saturating characteristic or "S curve". The qualitative effect of such a nonlinearity on the relation between the amplified residue and the input is shown in figure 3.10. In the example shown, the actual residue matches the ideal residue at the comparator threshold points. This is the case in a pipelined ADC with digital calibration of the gain. Digital calibration is discussed later in this chapter.

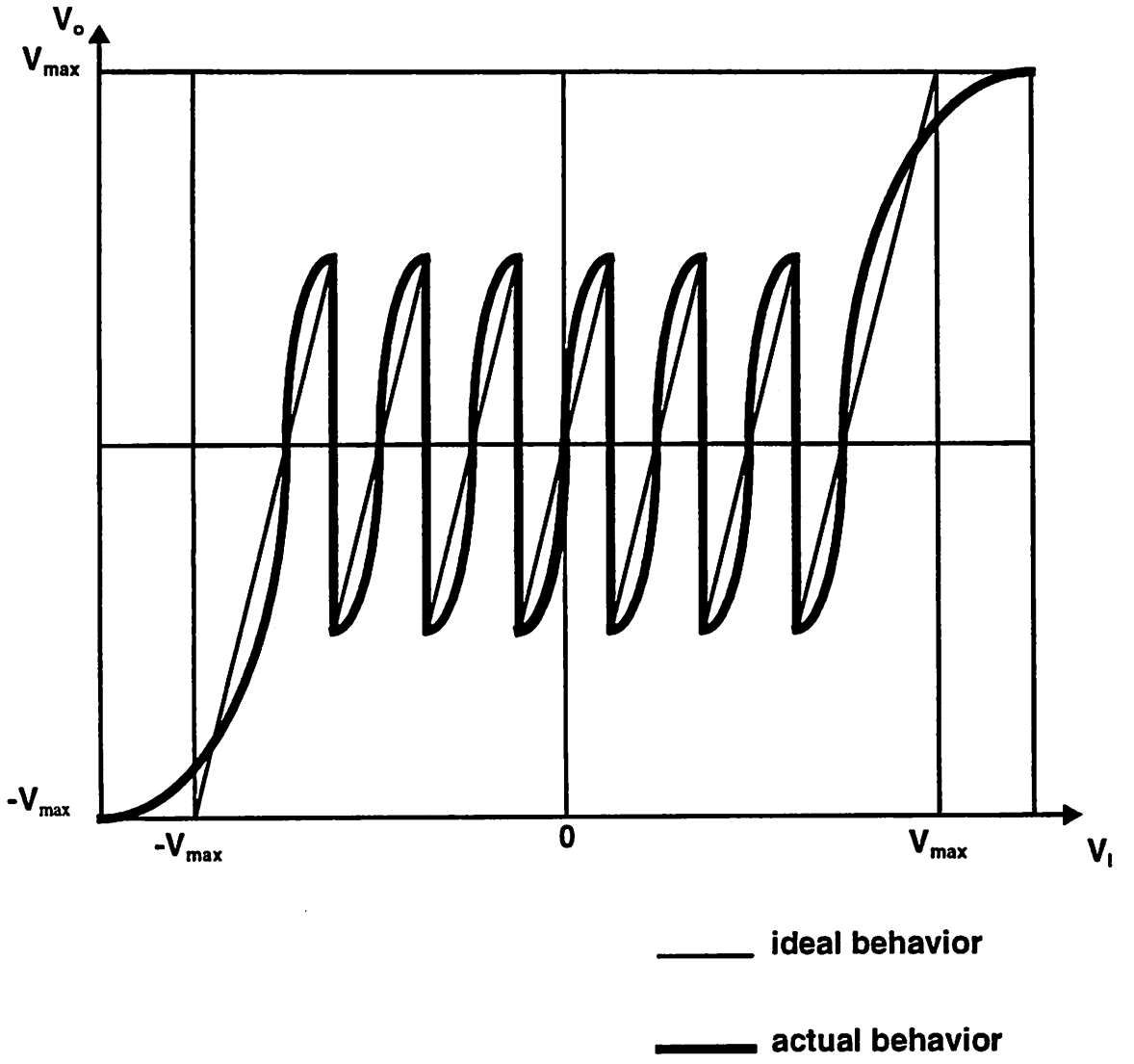


Figure 3.10 Effect of Residue Amplifier Nonlinearity on the Transfer Characteristic of a Pipeline Stage

Figure 3.11 shows the effect of the amplifier nonlinearity on the overall input-output transfer

characteristic of the ADC. Note that very little differential nonlinearity is present at the comparator

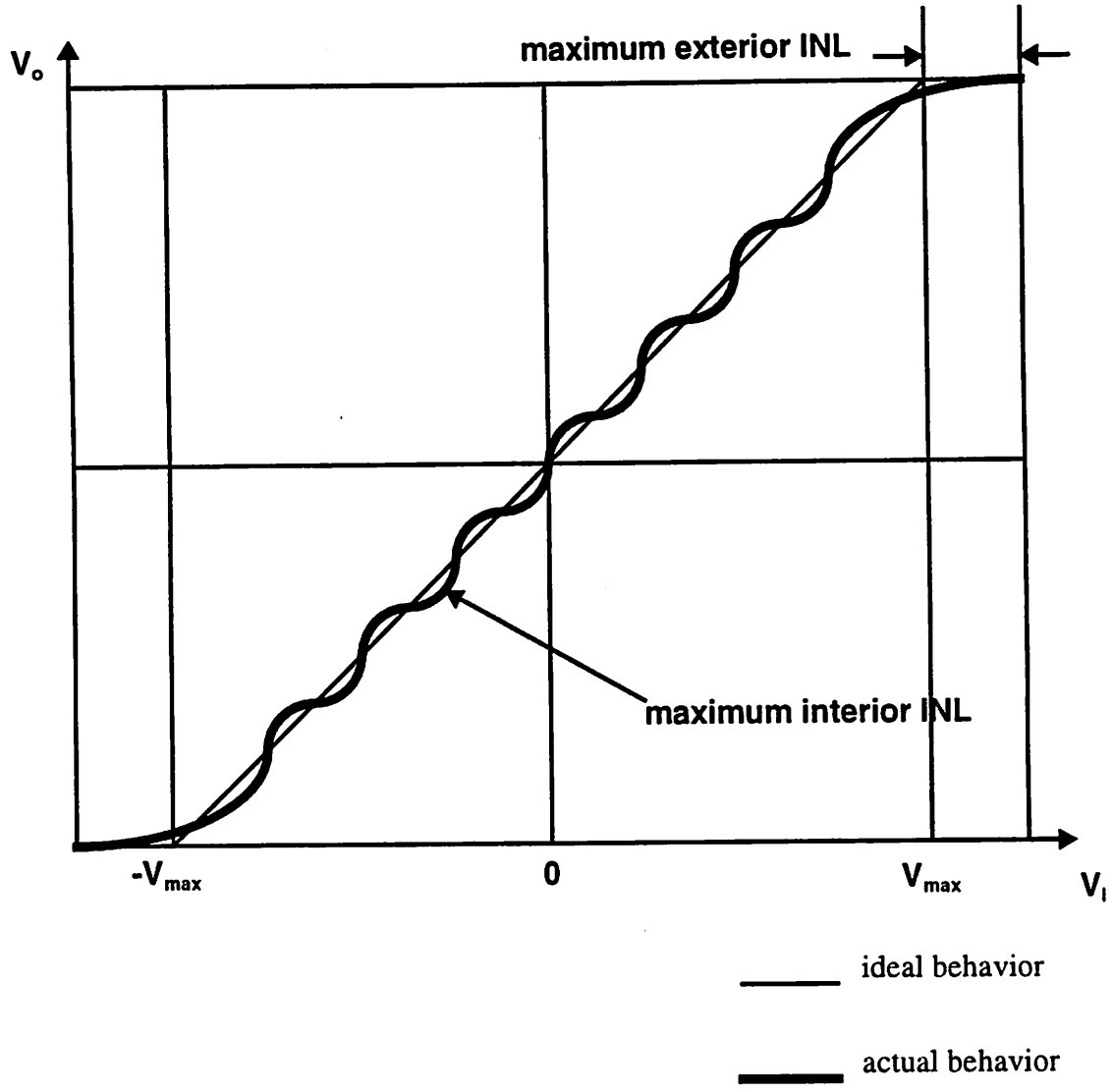


Figure 3.11 Effect of Residue Amplifier Nonlinearity on the ADC Transfer Characteristic

thresholds in the example shown above because at the threshold, the actual amplifier output matches the ideal at the comparator thresholds. However, the integral nonlinearity (INL) is not zero, and the INL contributed by the nonlinearity of the amplifier will next be calculated. In the figure the exterior INL refers to the INL that occurs for an input greater than the maximum comparator threshold or smaller than the minimum comparator threshold. Interior INL is defined here

as INL that occurs for inputs smaller than the maximum comparator threshold and greater than the minimum comparator threshold.

The feedback system in figure 3.12 is used to model the residue amplifier in order to determine its contribution to the INL. In the figure, V_I represents the input to a stage of the pipeline, and V_E represents the summing node voltage. If the gain block $A(V_I)$ is linear, the ratio of output to input

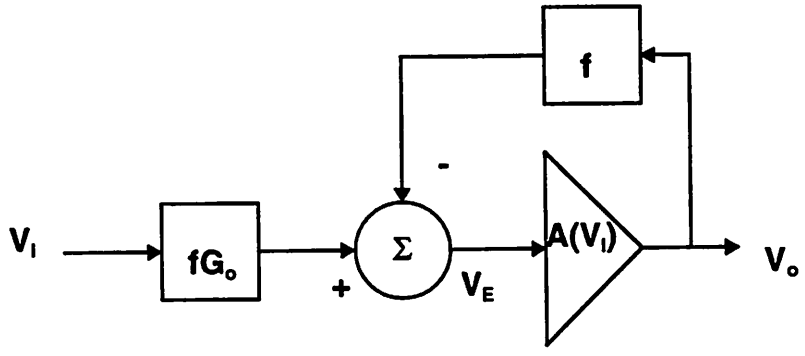


Figure 3.12 Model used to Analyze Effect on Opamp Nonlinearity on DNL

is given as follows.

$$\frac{V_o}{V_I} = \frac{G_o}{1 + \frac{1}{Af}} \tag{3.29}$$

G_o in the above equation is the value of the closed loop gain G that is obtained if the loop gain Af is infinite.

If the gain is nonlinear, then a more in depth analysis is necessary. The error voltage V_E in the above figure is given as follows.

$$V_E = fG_oV_I - fV_o \tag{3.30}$$

In order to compute the INL, the deviation of the input for a known output is computed. There-

fore, the above equation is solved for the input V_I as shown below.

$$V_I = \frac{V_o}{G_o} + \frac{V_E}{fG_o} \tag{3.31}$$

The error ε is found by subtracting the ideal input value $\frac{V_o}{G_{ideal}}$ from the input given by the above equation.

$$\varepsilon = V_I - \frac{V_o}{G_{ideal}} = \frac{V_E}{fG_o} + V_o\left(\frac{1}{G_o} - \frac{1}{G_{ideal}}\right) \tag{3.32}$$

The error ε_i referred to the input of the ADC is found by dividing the above error by the gains preceding the amplifier being studied.

$$\varepsilon_i = \frac{\varepsilon}{G^k} \tag{3.33}$$

In the above equation, k is the number of gain stages preceding the stage being studied.

The INL in LSBs is then computed by dividing this value by the full scale range and multiplying by the number of quantization levels as shown below.

$$INL\,(LSBs) = \frac{2^N}{2V_{max}}\varepsilon_i = \frac{2^N}{2V_{max}G_{ideal}^k}\left(\frac{V_E}{fG_o} + V_o\left(\frac{1}{G_o} - \frac{1}{G_{ideal}}\right)\right) \tag{3.34}$$

In the above equation, N represents the number of bits resolved by the ADC.

EXAMPLE 1:

An example is worthwhile to make the above analysis more specific. In this example it is assumed that the output to input characteristic of the amplifier has a hyperbolic tangent shape as shown below. This transfer characteristic corresponds to that of an emitter coupled pair with infinite output resistance.

$$V_o = V_{max}\tanh\left(\frac{A_o V_E}{V_{max}}\right) \tag{3.35}$$

In this amplifier, the maximum swing is +/-V_{max}, and the maximum small signal gain is A_o. In

order to use the analysis just discussed, it is necessary to get the error voltage V_E as a function of the output voltage V_o.

$$V_E = \frac{V_{max}}{A_o} \operatorname{atanh}\left(\frac{V_o}{V_{max}}\right) \tag{3.36}$$

The above relation is used to replace V_E in equation [3.34] as shown here.

$$INL = \frac{2^N}{2V_{max}G_{ideal}^k}\left(\frac{V_{max}}{fA_oG_o}\operatorname{atanh}\left(\frac{V_o}{V_{max}}\right) + V_o\left(\frac{1}{G_o} - \frac{1}{G_{ideal}}\right)\right) \tag{3.37}$$

Now, the ideal gain G_{ideal} must be determined. This is calculated by assuming that the actual transfer function crosses the ideal transfer function at $V_o=0.5V_{max}$. This situation will occur if the ADC is calibrated to be continuous when the input is equal to a comparator threshold. This calibration technique is discussed later in this chapter.

$$V_I\left(\frac{1}{2}V_{max}\right) = \frac{\frac{1}{2}V_{max}}{G_{ideal}} = \frac{\frac{1}{2}V_{max}}{G_o} + \frac{V_{max}}{fG_oA_o}\operatorname{atanh}\left(\frac{1}{2}\right) \tag{3.38}$$

This equation can then be solved for G_o as a function of G_{ideal} as shown below.

$$G_o = G_{ideal}\left[1 + \frac{2}{fA_o}\operatorname{atanh}\left(\frac{1}{2}\right)\right] \tag{3.39}$$

In order to find the output voltage V_{om} at which INL is maximum, the derivative of the INL in [3.37] is set to 0.

$$\left.\frac{dINL}{dV_o}\right|_{V_o = V_{om}} = 0 \tag{3.40}$$

The result is that V_{om} is given as follows.

$$V_{om} = V_{max}\sqrt{1 - \frac{1}{2\operatorname{atanh}\left(\frac{1}{2}\right)}} \tag{3.41}$$

The maximum INL is then given by the following formula.

$$INL_{max} = \frac{2^N}{2fA_oG_{ideal}^k}\left(\frac{1}{1+\frac{2}{fA_o}\text{atanh}\left(\frac{1}{2}\right)}\right)\left\{\text{atanh}\left(\left(\sqrt{1-\frac{1}{2\,\text{atanh}\left(\frac{1}{2}\right)}}\right)\right)-2\,\text{atanh}\left(\frac{1}{2}\right)\sqrt{1-\frac{1}{2\,\text{atanh}\left(\frac{1}{2}\right)}}\right\} \tag{3.42}$$

$$INL_{max} = \frac{2^N}{2G_{ideal}^k fA_o}\left(\frac{0.02004}{1+\frac{1.09861}{fA_o}}\right) \tag{3.43}$$

Now, in order to get a numerical idea of the gain required to achieve a linearity to a certain number of bits, a few more assumptions are made. It is assumed that 0.5LSB INL is desired. It is also assumed that the first gain block contributes the most to the nonlinearity, so k=0. Also fG_{ideal} is assumed to be 0.5. With these assumptions, equation (3.43) can be solved for the open loop gain with the following result.

$$A_o = \frac{2^N}{25} - \frac{1.1}{f} \tag{3.44}$$

This above result gives the required gain to obtain half LSB linearity in the interior region of the ADC transfer characteristic. In other words, the maximum interior INL is 0.5LSB for the gain shown above.

It turns out that the maximum exterior INL is much worse than the interior INL because the amplifier must swing to full scale at the endpoints of the transfer characteristic. The generalized INL due to nonlinearity in an emitter coupled pair is given by the following formula.

$$INL\,(LSBs) = \left(\frac{\frac{2^N}{fA_oG_{ideal}^{k+1}}}{1+\frac{2}{fA_o}\text{atanh}\left(\frac{1}{2}\right)}\right)\left\{\text{atanh}\left(\frac{V_o}{V_{omax}}\right)-2\left(\frac{V_o}{V_{omax}}\right)\text{atanh}\left(\frac{1}{2}\right)\right\} \tag{3.45}$$

Now, if k=0, N is large, fG_{ideal}=0.5, and A_o is large, the following table may be generated.

Table 3.1 Required Gain in order to Achieve .5LSB Linearity at N Bits

$\frac{V_o}{V_{omax}}$	A_o
0.6	$0.136 \cdot 2^N$
0.7	$0.393 \cdot 2^N$
0.8	$0.879 \cdot 2^N$
0.9	$1.934 \cdot 2^N$
0.95	$3.152 \cdot 2^N$

Note that the open loop gain required to achieve adequate linearity is surprisingly small in the interior region, but quite high near the endpoints.

EXAMPLE 2:

In this example it is assumed that the opamp is a MOS source coupled pair with a large signal characteristic shown below.

$$V_o = A_o V_E \sqrt{1 - \frac{A_o^2}{4}\left(\frac{V_E}{V_{max}}\right)^2} \tag{3.46}$$

In this amplifier, the maximum swing is +/-V_{max}, and the maximum small signal gain is A_o. In order to use the analysis just discussed, it is necessary to get the error voltage V_E as a function of the output voltage V_o.

$$V_E = \left(\frac{V_{max}\sqrt{2}}{A_o}\right)\sqrt{1 - \sqrt{1 - \left(\frac{V_o}{V_{max}}\right)^2}} \tag{3.47}$$

The above relation is used to replace V_E in equation (3.34) as shown here.

$$INL = \frac{2^N}{2V_{max}G_{ideal}^k}\left\{\frac{V_{max}\sqrt{2}}{fA_oG_o}\sqrt{1 - \sqrt{1 - \left(\frac{V_o}{V_{max}}\right)^2}} + V_o\left(\frac{1}{G_o} - \frac{1}{G_{ideal}}\right)\right\} \tag{3.48}$$

Now, the ideal gain G_{ideal} must be determined. This is calculated by assuming that the ideal actual transfer function crosses the ideal transfer function at V_o=0.5V_{max}. This situation will occur if the ADC is

calibrated to be continuous when the input is equal to a comparator threshold. This calibration technique is discussed later in this chapter.

$$V_I(\frac{1}{2}V_{max}) = \frac{\frac{1}{2}V_{max}}{G_{ideal}} = \frac{\frac{1}{2}V_{max}}{G_o} + \frac{V_{max}\sqrt{2}}{fG_oA_o}\sqrt{1-\sqrt{\frac{3}{4}}} \tag{3.49}$$

This equation can then be solved for G_o as a function of G_{ideal} as shown below.

$$G_o = G_{ideal}\left[1 + \frac{2\sqrt{2}}{fA_o}\sqrt{1-\sqrt{\frac{3}{4}}}\right] \tag{3.50}$$

In order to find the output voltage V_{om} at which INL is maximum, the derivative of the INL in (3.37) is set to 0.

$$\left.\frac{dINL}{dV_o}\right|_{V_o = V_{om}} = 0 \tag{3.51}$$

The result is that V_{om} is given as follows.

$$V_{om} = 0.2966V_{max} \tag{3.52}$$

The maximum INL is then given by the following formula assuming A_o is large.

$$INL_{max} = \frac{2^N}{fA_oG_{ideal}^{k+1}}(3.524 \bullet 10^{-3}) \tag{3.53}$$

Now, in order to get a numerical idea of the gain required to achieve a linearity to a certain number of bits, a few more assumptions are made. It is assumed that 0.5LSB INL is desired. It is also assumed that the first gain block contributes the most to the nonlinearity, so k=0. Also fG_{ideal} is assumed to be 0.5. With these assumptions, equation (3.43) can be solved for the open loop gain with the following result.

$$A_o = \frac{2^N}{70.95} \tag{3.54}$$

This above result gives the required gain to obtain half LSB linearity in the interior region of the ADC transfer characteristic. In other words, the maximum interior INL is 0.5LSB for the gain

shown above.

It turns out that the maximum exterior INL is much worse than the interior INL because the amplifier must swing to full scale at the endpoints of the transfer characteristic. Therefore, the INL in this region will be considered, as it was for the bipolar case. The generalized INL due to nonlinearity in an emitter coupled pair is given by the following formula.

$$INL(LSBs) = \left(\frac{\frac{2^N\sqrt{2}}{2fA_oG_{ideal}^{k+1}}}{1+\frac{2\sqrt{2\left(1-\sqrt{\frac{3}{4}}\right)}}{fA_o}} \right) \left\{ \sqrt{1-\sqrt{1-\left(\frac{V_o}{V_{max}}\right)^2}} - 2\left(\frac{V_o}{V_{max}}\right)\sqrt{1-\sqrt{\frac{3}{4}}} \right\} \quad (3.55)$$

Now, if k=0, N is large, fG_{ideal}=0.5, and A_o is large, the following table may be generated.

Table 3.2Required Gain in order to Achieve .5LSB Linearity at N Bits

$\frac{V_o}{V_{omax}}$	A_o
0.6	$\frac{2^N}{124.8}$
0.7	$\frac{2^N}{31.78}$
0.8	$\frac{2^N}{15.09}$
0.9	$\frac{2^N}{7.663}$

Table 3.2Required Gain in order to Achieve .5LSB Linearity at N Bits

1.00	$\frac{2^N}{2.639}$

From the above two examples, the MOS source coupled pair might be expected to be more linear than the bipolar emitter coupled pair. However, this analysis neglects the effects of nonlinear output resistance and multiple gain stages. Therefore, the linearity performance of the source coupled pair is most likely similar to the emitter coupled pair.

3.3.6 Incomplete Settling of the Sample and Hold Amplifier Output

Errors introduced by incomplete settling of the sample and hold amplifier are important because these errors limit the maximum sampling rate of the ADC. Therefore, this section includes an analysis of opamp settling behavior for both first order and second order linear systems. This type of analysis is frequently performed by pipelined ADC designers in an effort to determine the time constant required in order to meet a given sampling rate specification. A discussion of nonlinear settling of first order systems that includes slew rate effects is performed by Chuang.[Chuang82]

During hold phase, the switched capacitor sample and hold amplifier behaves as a feedback circuit with a unit step applied to the input. The output response of this circuit is a step response that requires a finite amount of time to settle to a given accuracy. Failure to settle accurately that is caused by inadequate hold time results in errors that degrade the performance of ADCs and other systems using switched capacitor sample and hold amplifiers. If the settling is linear, the error is proportional to the input, and the result is a fixed gain error just as described previously. If the settling is nonlinear, the effect is a signal dependent error.

In many pipelined ADCs the available settling time is somewhat less than half the sampling rate. The following analyses for first order and second order linear systems relate the required settling time to the time constant and required accuracy. With this information, a relation between the time constant and the maximum sampling rate can be determined. Slewing effects are not treated here.

3.3.6.1 Settling Time of a Single Pole System

The time domain unit step response of a single pole system with time constant τ and final value y_∞ is given by the following formula.

$$y(t) = y_\infty (1 - e^{-t/\tau}) \tag{3.56}$$

The relative error ε is the difference between y and y_∞ normalized to y_∞ as shown below.

$$\varepsilon = \left(\frac{y - y_\infty}{y_\infty}\right) = -e^{-t/\tau} \tag{3.57}$$

For settling to N bit accuracy, $|\varepsilon|$ should be equal to $\frac{1}{2^N}$. Therefore, the settling time t_s of the single pole system is given by the formula below.

$$t_s = N\tau \ln 2 \tag{3.58}$$

3.3.6.2 Settling Time of a Critically Damped Two Pole System

The s domain transfer function H(s) of a critically damped system is given by the following formula.

$$H(s) = \frac{1}{(1 + \tau s)^2} \tag{3.59}$$

The s domain unit step response Y(s) of the critically damped system is found by multiplying the above H(s) by the s domain unit step input X(s) as shown below.

$$Y(s) = H(s)X(s) = \frac{1}{s(1 + \tau s)^2} = \frac{1}{s} - \frac{1}{(s + \frac{1}{\tau})} - \frac{\frac{1}{\tau}}{(s + \frac{1}{\tau})^2} \tag{3.60}$$

The time domain unit step response is then given by the following formula.

$$y(t) = (1 - e^{-t/\tau} - \frac{t}{\tau} e^{-t/\tau})\, u(t) \tag{3.61}$$

The relative error in the output is the difference between y and the value y would have after an infinite amount of time normalized to the final value of y. For this case the relative error is given as

follows.

$$\varepsilon = y - 1 = -(1 + \frac{t}{\tau}) e^{-t/\tau} \tag{3.62}$$

3.3.6.3 Settling Time of an Underdamped Two Pole System

The s domain transfer function H(s) of an underdamped system is given by the following formula.

$$H(s) = \frac{1}{\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1} \tag{3.63}$$

In this system, the two poles p_1 and p_2 form a complex conjugate pair with the locations shown below.

$$p_1 = -\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2} = -\frac{1}{\tau} - j\omega_n\sqrt{1-\zeta^2} \tag{3.64}$$

$$p_2 = -\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2} = -\frac{1}{\tau} + j\omega_n\sqrt{1-\zeta^2} \tag{3.65}$$

The s domain unit step response Y(s) of the underdamped system is found by multiplying the above H(s) by the s domain unit step input X(s) as shown below.

$$Y(s) = H(s)X(s) = \frac{1}{s\left(\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1\right)} \tag{3.66}$$

To find the time domain unit step response, the s domain unit step response is rearranged as shown below.

$$Y(s) = 1 - \frac{(s + \zeta\omega_n)}{(s^2 + 2\zeta\omega_n s + \omega_n^2)} - \frac{\zeta\omega_n}{(s^2 + 2\zeta\omega_n s + \omega_n^2)} \tag{3.67}$$

The time domain unit step response y(t) is then given by the following formula where

$\omega_d = \omega_n\sqrt{1-\zeta^2}$.

$$y(t) = 1 - e^{-\zeta\omega_n t}\left[\cos(\omega_d t) + \frac{\zeta\omega_n}{\omega_d}\sin(\omega_d t)\right] \tag{3.68}$$

The relative error ε is then given as shown below.

$$\varepsilon = y - y_\infty = -e^{-\zeta\omega_n t}\left[\cos(\omega_d t) + \frac{\zeta\omega_n}{\omega_d}\sin(\omega_d t)\right] \tag{3.69}$$

Figure 3.13 shows a plot of the ratio of settling time to time constant for an underdamped second order system. The settling time in the plot is the time required to guarantee that the magnitude of the relative error ε is less than or equal to $1/2^n$ where n is given in bits. The time constant τ is $1/(\zeta\omega_n)$. Note that in some cases it is theoretically possible to reduce settling time by carefully controlling the damping factor so that a minimum point in settling time is reached. In practice it is usually not very useful to do this because process variations make it too difficult to control the damping factor accurately enough to ensure that the settling time will be at a minimum point. However, reducing the damping factor from 1 to 0.7 for a fixed time constant can significantly improve the settling time regardless of whether or not this corresponds to a minimum point. If stability considerations allow, it may therefore be beneficial to design a system with a damping factor smaller than 1. These second order settling issues are further treated by Yang[Yang90].

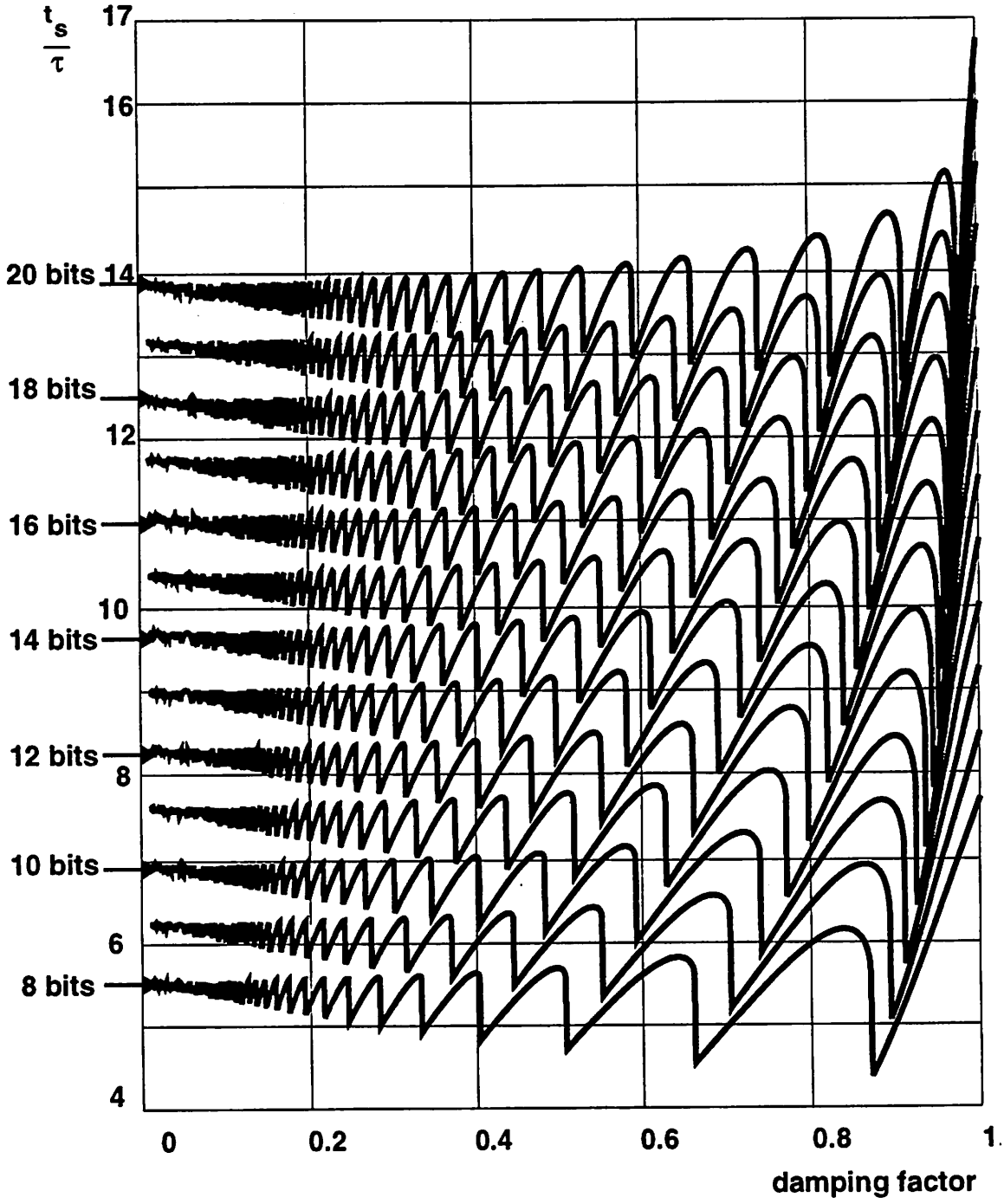


Figure 3.13 Settling Time versus Damping Factor for a Two Pole Underdamped System for Different Settling Accuracies

3.3.6.4 Settling Time of an Overdamped Two Pole System

In a second order system with a transfer function described by equation (3.63) with $\zeta > 1$, both poles are real and given by the following formulas.

$$\omega_1 = \omega_n(\zeta - \sqrt{\zeta^2 - 1}) \tag{3.70}$$

$$\omega_2 = \omega_n(\zeta + \sqrt{\zeta^2 - 1}) \tag{3.71}$$

The unit step response y(t) is given by the following formula.

$$y(t) = 1 - \left(\frac{\omega_2}{\omega_2 - \omega_1}\right)e^{-\omega_1 t} + \left(\frac{\omega_1}{\omega_2 - \omega_1}\right)e^{-\omega_2 t} \tag{3.72}$$

It is convenient to define an overdamping factor c_o in terms of the damping factor as follows.

$$c_o = \sqrt{1 - \frac{1}{\zeta^2}} \tag{3.73}$$

The error in the unit step response is expressed by the following formula.

$$\varepsilon = \frac{-e^{-\omega_1 t}}{2c_o}\left[1 + c_o - (1 - c_o)\, e^{-2\omega_1 t\left(\frac{c_o}{1 - c_o}\right)}\right] \tag{3.74}$$

Figure 3.14 shows a plot of $\omega_1 t_s$ versus ζ where t_s is the settling time. Again, the plot shows a number of different settling accuracies. For large values of damping factor $\omega_2 >> \omega_1$. For these cases the overdamped system is well approximated as a first order system, and equation (3.58) gives a good prediction of the relation between the settling time and the time constant. Note that the ratio of settling time to time constant is maximum for the critically damped case (ζ=1). This might lead one to believe that it is nonoptimal to design for critical damping. If the time constant is held fixed, this is true. However, in many cases designing for critical damping makes it possible to reduce the time constant to a value not possible in an overdamped system. This is usually the case in a two pole opamp with fixed power dissipation. As a result, critical damping often does offer a higher speed solution than overdamping, especially if the power of the system is fixed.

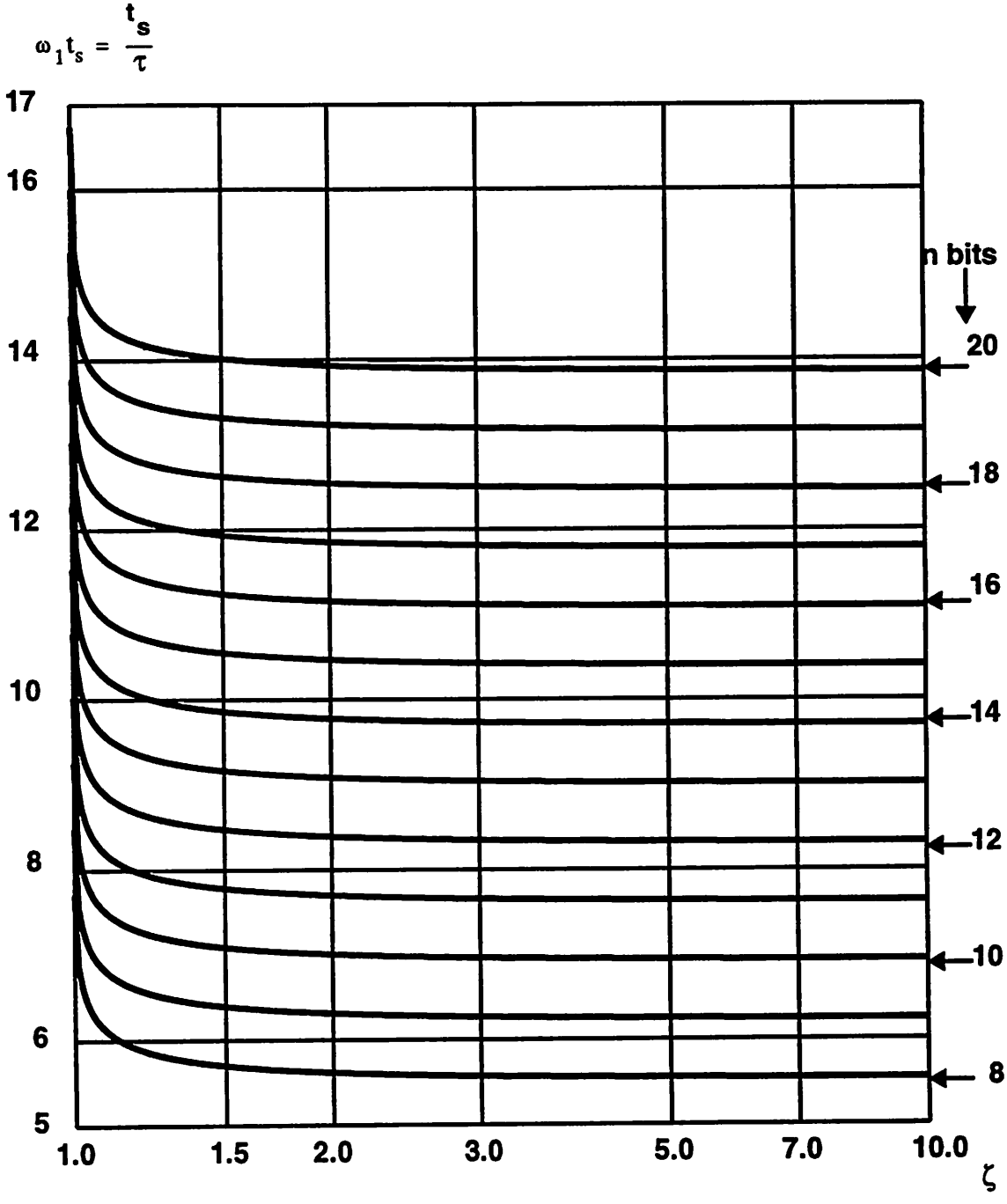


Figure 3.14 Settling Time versus Damping Factor for a Two Pole Overdamped System for Different Settling Accuracies

3.3.6.5 Tabulated Settling Times

Now the settling time t_s required to reach a certain number of bits of accuracy is listed for the single pole and two pole systems.

Table 3.3 Ratio of Settling Time to Time Constant Versus Resolution in bits

one pole or two pole →		one	two	two	two	two	two	two	two	two	two
ζ→			1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2
N bits	error ε	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$	$\frac{t_s}{\tau}$
8	$\frac{1}{256}$	5.55	7.71	5.02	5.97	4.97	5.61	5.68	5.63	5.27	5.31
9	$\frac{1}{512}$	6.24	8.49	5.23	6.43	6.46	5.91	6.13	6.00	6.17	5.96
10	$\frac{1}{1024}$	6.93	9.26	7.58	6.79	7.22	6.11	6.36	7.02	6.97	6.61
11	$\frac{1}{2048}$	7.62	10.03	8.44	7.06	7.65	7.80	7.73	7.38	7.28	7.26
12	$\frac{1}{4096}$	8.32	10.79	9.14	7.25	7.95	8.16	8.05	8.41	8.18	7.91
13	$\frac{1}{8192}$	9.01	11.54	9.74	9.52	8.16	8.40	8.23	8.75	9.04	9.03
14	$\frac{1}{16384}$	9.70	12.29	10.26	10.14	10.04	9.93	9.69	9.79	9.27	9.70

Table 3.3 Ratio of Settling Time to Time Constant Versus Resolution in bits

15	$\frac{1}{32768}$	10.40	13.04	10.72	10.60	10.56	10.39	9.94	10.13	10.19	10.37
16	$\frac{1}{65536}$	11.09	13.78	11.10	10.97	10.91	10.67	11.23	11.18	11.08	11.03
17	7.6e-6	11.78	14.53	11.41	11.24	11.16	11.83	11.62	11.50	11.27	11.68
18	3.8e-6	12.48	15.27	11.65	11.43	12.76	12.60	11.82	12.56	12.19	12.34
19	1.9e-6	13.17	16.00	13.65	13.68	13.44	12.93	13.23	12.88	13.09	12.99
20	9.5e-7	13.86	16.74	14.65	14.30	13.85	13.14	13.52	13.94	13.91	13.64

3.3.7 Sample and Hold Tracking Nonlinearity

Tracking nonlinearity refers to the distortion created by the sample and hold circuit as it tracks the input signal. This type of distortion tends to increase as the input frequency increases. In a pipelined ADC, this effect is usually an issue in the first stage, where a high frequency input is present. In a switched capacitor sample and hold circuit the nonlinearity is primarily caused by the nonlinear resistance of the MOS switch and the nonlinear junction capacitance associated with the source and drain diffusions. A thorough analysis of this problem in a mixer application is presented by Keys[Keys94]. In this section, a simplified analysis of the distortion caused by the nonlinearity of the switch resistance is performed. This analysis might be useful when trying to choose the switch sizes, the switch on voltage, or the signal swing. This analysis neglects the nonlinear capacitance.

The switched capacitor sample and hold circuit during the sampling phase of operation is basically a MOS switch in series with a capacitor. This circuit may be thought of as an impedance divider with the output taken across the capacitor. This impedance divider consists of a frequency dependent impedance (the capacitor) and a nonlinear impedance (the resistor). At low frequencies, the output across the capacitor tracks the input signal closely. However, at higher frequencies, the output is attenuated and distorted.

The analysis presented here is based on the model shown in figure 3.15. The model consists of a fully differential sample and hold consisting of a linear resistor (R_L) in series with a CMOS

switch and a linear capacitor. Because of the fully differential architecture, the second harmonic component of distortion is largely reduced. However, a significant third harmonic component of distortion still remains. The purpose of this analysis is to estimate how this third harmonic distortion varies with various parameters such as signal amplitude, switch size, switch matching.

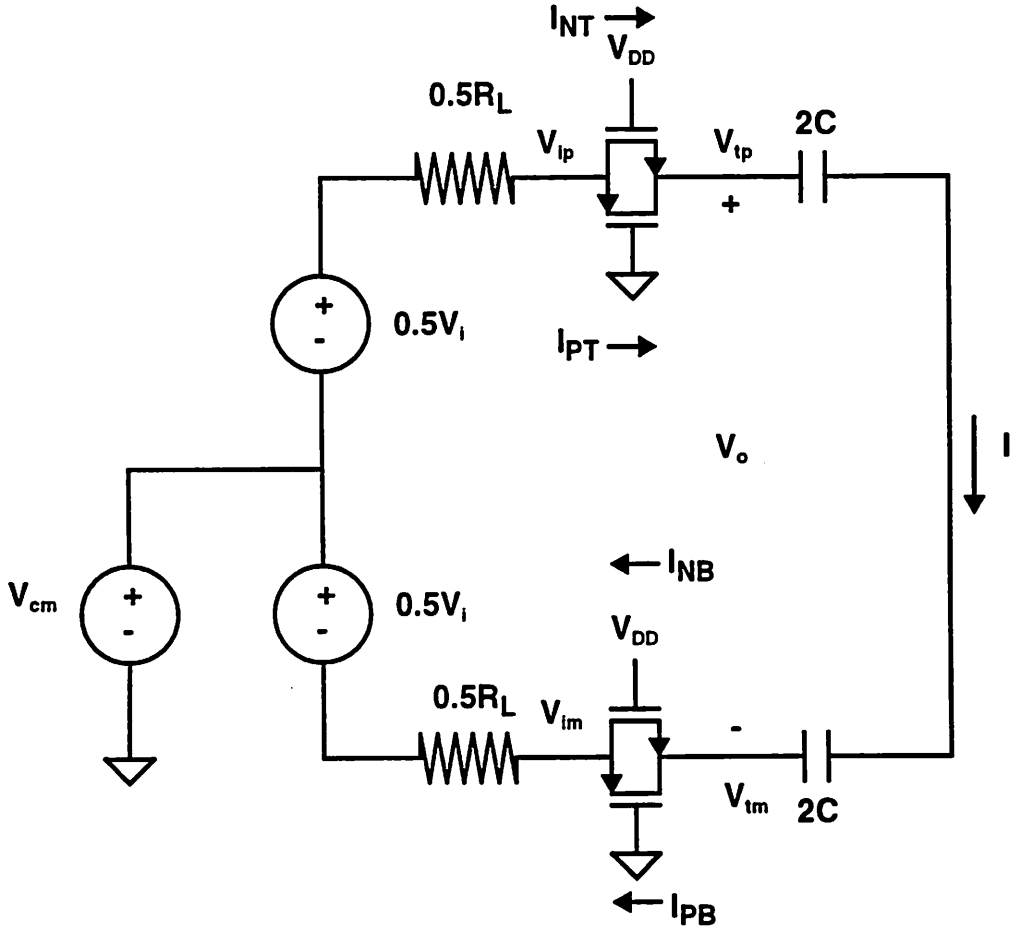


Figure 3.15 Model of Sample and Hold Circuit in the Sampling Phase of Operation

To begin the analysis, the current I is expressed in terms of the voltages at the terminals of the MOS devices. The long channel model for the MOS devices is used with body effect included. Mobility degradation and velocity saturation effects are ignored.

During the sampling phase the MOS transistors are usually in the triode region, so the current may be expressed by the following equations.

$$I = I_{NT} + I_{PT} = I_{NB} + I_{PB} \tag{3.75}$$

current I_{NT} in the top NMOS switch:

$$I_{NT} = \mu_n C_{ox} \frac{W_N}{L_N} \left[(V_{DD} - V_{tp} - V_{TN0} - \gamma_N \sqrt{V_{tp} + 2\phi_F} + \gamma_N \sqrt{2\phi_F}) (V_{ip} - V_{tp}) - \frac{1}{2} (V_{ip} - V_{tp})^2 \right] \tag{3.76}$$

current I_{PT} in the top PMOS switch:

$$I_{PT} = \mu_p C_{ox} \frac{W_P}{L_P} \left[(V_{DD} - V_{TP0} - \gamma_P \sqrt{V_{DD} - V_{ip} + 2\phi_F} + \gamma_P \sqrt{2\phi_F}) (V_{ip} - V_{tp}) - \frac{1}{2} (V_{ip} - V_{tp})^2 \right] \tag{3.77}$$

current I_{NB} in the bottom NMOS switch:

$$I_{NB} = \mu_n C_{ox} \frac{W_N}{L_N} \left[(V_{DD} - V_{im} - V_{TN0} - \gamma_N \sqrt{V_{im} + 2\phi_F} + \gamma_N \sqrt{2\phi_F}) (V_{tm} - V_{im}) - \frac{1}{2} (V_{tm} - V_{im})^2 \right] \tag{3.78}$$

current I_{PB} in the bottom PMOS switch:

$$I_{PB} = \mu_p C_{ox} \frac{W_P}{L_P} \left[(V_{tm} - V_{TP0} - \gamma_P \sqrt{V_{DD} - V_{tm} + 2\phi_F} + \gamma_P \sqrt{2\phi_F}) (V_{tm} - V_{im}) - \frac{1}{2} (V_{tm} - V_{im})^2 \right] \tag{3.79}$$

It is assumed here that the voltage drop from drain to source of each of the switches is small. This means that the distortion is a small perturbation of the signal. This assumption is valid for frequencies below the bandwidth of the sample and hold, but it breaks down for higher frequencies. With this assumption, the following expressions may be written.

$$V_{ip} \approx V_{tp} \approx V_{cm} + \frac{1}{2} V_i \tag{3.80}$$

$$V_{im} \approx V_{tm} \approx V_{cm} - \frac{1}{2}V_i \tag{3.81}$$

With these assumptions, the conductances of each switch may be expressed as follows.

conductance of the top NMOS switch:

$$G_{TN} \approx \mu_n C_{ox} \frac{W_N}{L_N}\left[V_{DD} - V_{TN0} - V_{CM} + \gamma_N\sqrt{2\phi_F} - \frac{1}{2}V_i - \gamma_N\sqrt{V_{CM} + 2\phi_F + \frac{1}{2}V_i}\right] \tag{3.82}$$

conductance of the top PMOS switch:

$$G_{TP} \approx \mu_p C_{ox} \frac{W_P}{L_P}\left[V_{DD} - V_{TP0} + \gamma_P\sqrt{2\phi_F} + \frac{1}{2}V_i - \gamma_P\sqrt{V_{DD} + 2\phi_F - V_{cm} - \frac{1}{2}V_i}\right] \tag{3.83}$$

conductance of the bottom NMOS switch:

$$G_{BN} \approx \mu_n C_{ox} \frac{W_N}{L_N}\left[V_{DD} - V_{TN0} - V_{CM} + \gamma_N\sqrt{2\phi_F} + \frac{1}{2}V_i - \gamma_N\sqrt{V_{CM} + 2\phi_F - \frac{1}{2}V_i}\right] \tag{3.84}$$

conductance of the bottom PMOS switch:

$$G_{TP} \approx \mu_p C_{ox} \frac{W_P}{L_P}\left[V_{cm} - V_{TP0} - \frac{1}{2}V_i + \gamma_P\sqrt{2\phi_F} - \gamma_P\sqrt{V_{DD} + 2\phi_F - V_{cm} + \frac{1}{2}V_i}\right] \tag{3.85}$$

total nonlinear conductance G:

$$G = \frac{(G_{TN} + G_{TP})\,(G_{BN} + G_{BP})}{G_{TN} + G_{TP} + G_{BN} + G_{BP}} \tag{3.86}$$

Because the distortion is assumed to be a small perturbation, the circuit model may be simpli-

fied to the single ended equivalent circuit shown in figure 3.16. In this circuit, R_N represents the

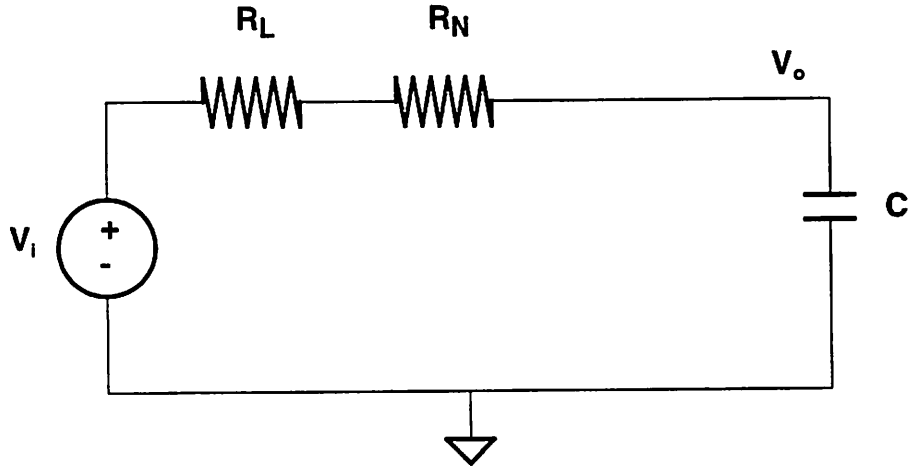


Figure 3.16 Simplified Circuit Model for the Sample and Hold Circuit During the Sampling Phase

combined resistance of the MOS switches. R_N is related to the previous analysis by the following expressions.

$$R_N = \frac{1}{G} = \frac{G_{TN} + G_{BN} + G_{TP} + G_{BP}}{G_{TN}G_{BN} + G_{TN}G_{BP} + G_{TP}G_{BN} + G_{TP}G_{BP}} \tag{3.87}$$

$$G_{TN} + G_{BN} = k_N\left[2V_{yN} - \gamma_N\sqrt{V_{xN}}\left(\sqrt{1 + \frac{\frac{1}{2}V_i}{V_{xN}}} + \sqrt{1 - \frac{\frac{1}{2}V_i}{V_{xN}}}\right)\right] \tag{3.88}$$

$$G_{TP} + G_{BP} = k_P\left[2V_{yP} - \gamma_P\sqrt{V_{xP}}\left(\sqrt{1 + \frac{\frac{1}{2}V_i}{V_{xP}}} + \sqrt{1 - \frac{\frac{1}{2}V_i}{V_{xP}}}\right)\right] \tag{3.89}$$

$$G_{TN}G_{BN} = k_N^2\left(V_{yN} - \frac{1}{2}V_i - \gamma_N\sqrt{V_{xN}\left(1 + \frac{\frac{1}{2}V_i}{V_{xN}}\right)}\right)\left(V_{yN} + \frac{1}{2}V_i - \gamma_N\sqrt{V_{xN}\left(1 - \frac{\frac{1}{2}V_i}{V_{xN}}\right)}\right) \tag{3.90}$$

$$G_{TP}G_{BP} = k_P^2\left(V_{yP} - \frac{1}{2}V_i - \gamma_P\sqrt{V_{xP}\left(1 + \frac{\frac{1}{2}V_i}{V_{xP}}\right)}\right)\left(V_{yP} + \frac{1}{2}V_i - \gamma_P\sqrt{V_{xP}\left(1 - \frac{\frac{1}{2}V_i}{V_{xP}}\right)}\right) \tag{3.91}$$

$$G_{TN}G_{BP} = k_N k_P\left(V_{yN} - \frac{1}{2}V_i - \gamma_N\sqrt{V_{xN}\left(1 + \frac{\frac{1}{2}V_i}{V_{xN}}\right)}\right)\left(V_{yP} - \frac{1}{2}V_i - \gamma_P\sqrt{V_{xP}\left(1 + \frac{\frac{1}{2}V_i}{V_{xP}}\right)}\right) \tag{3.92}$$

$$G_{BN}G_{TP} = k_N k_P\left(V_{yP} + \frac{1}{2}V_i - \gamma_P\sqrt{V_{xP}\left(1 - \frac{\frac{1}{2}V_i}{V_{xP}}\right)}\right)\left(V_{yP} + \frac{1}{2}V_i - \gamma_P\sqrt{V_{xP}\left(1 - \frac{\frac{1}{2}V_i}{V_{xP}}\right)}\right) \tag{3.93}$$

$$k_N = \mu_n C_{ox}\frac{W_N}{L_N} \tag{3.94}$$

$$k_P = \mu_p C_{ox}\frac{W_P}{L_P} \tag{3.95}$$

$$V_{xN} = V_{CM} + 2\phi_F \tag{3.96}$$

$$V_{xP} = V_{DD} + 2\phi_F - V_{CM} \tag{3.97}$$

$$V_{yN} = V_{DD} - V_{TN0} - V_{CM} + \gamma_N\sqrt{2\phi_F} \tag{3.98}$$

$$V_{yP} = V_{CM} - V_{TP0} + \gamma_P\sqrt{2\phi_F} \tag{3.99}$$

In order to simplify the algebra, some more assumptions were made. First of all, $V_{CM}=0.5V_{DD}$ was assumed. Secondly, it was assumed that $\gamma_N=\gamma_P=\gamma$ and that $V_{TP0}=V_{TN0}=V_{T0}$. These assumptions

result in the following simplifications to the above equations.

$$V_{xP} = V_{xN} = V_x = \frac{1}{2}V_{DD} + 2\phi_F \tag{3.100}$$

$$V_{yP} = V_{yN} = V_y = \frac{1}{2}V_{DD} - V_{T0} + \gamma\sqrt{2\phi_F} \tag{3.101}$$

$$G_{TN} + G_{BN} = k_N\left[2V_y - \gamma\sqrt{V_x}\left(\sqrt{1+\frac{\frac{1}{2}V_i}{V_x}} + \sqrt{1-\frac{\frac{1}{2}V_i}{V_x}}\right)\right] \tag{3.102}$$

$$G_{TP} + G_{BP} = k_P\left[2V_y - \gamma\sqrt{V_x}\left(\sqrt{1+\frac{\frac{1}{2}V_i}{V_x}} + \sqrt{1-\frac{\frac{1}{2}V_i}{V_x}}\right)\right] \tag{3.103}$$

$$G_{TN}G_{BN} = k_N^2\left(V_y - \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1+\frac{\frac{1}{2}V_i}{V_x}\right)}\right)\left(V_y + \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1-\frac{\frac{1}{2}V_i}{V_x}\right)}\right) \tag{3.104}$$

$$G_{TP}G_{BP} = k_P^2\left(V_y - \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1+\frac{\frac{1}{2}V_i}{V_x}\right)}\right)\left(V_y + \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1-\frac{\frac{1}{2}V_i}{V_x}\right)}\right) \tag{3.105}$$

$$G_{TN}G_{BP} = k_Nk_P\left(V_y - \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1+\frac{\frac{1}{2}V_i}{V_x}\right)}\right)^2 \tag{3.106}$$

$$G_{BN}G_{TP} = k_Nk_P\left(V_y + \frac{1}{2}V_i - \gamma\sqrt{V_x\left(1-\frac{\frac{1}{2}V_i}{V_x}\right)}\right)^2 \tag{3.107}$$

These expressions are simplified still further by using the following approximations that are valid for small x.

$$\sqrt{1-x}+\sqrt{1+x}\approx 2-\frac{1}{4}x^2 \tag{3.108}$$

$$\sqrt{1-x}-\sqrt{1+x}\approx -x \tag{3.109}$$

With these approximations, the above expressions may be rewritten as shown below.

$$G_{TN}+G_{BN}+G_{TP}+G_{BP}\approx (k_N+k_P)\left\{2V_y-2\gamma\sqrt{V_x}+\frac{1}{16}\gamma\frac{V_i^2}{V_x^{3/2}}\right\} \tag{3.110}$$

$$G_{TN}G_{BN}+G_{TP}G_{BP}\approx (k_N^2+k_P^2)\left\{(V_y-\gamma\sqrt{V_x})^2+\left(-\frac{1}{4}-\frac{\gamma}{4\sqrt{V_x}}-\frac{\gamma^2}{8V_x}+\frac{\gamma V_y}{16V_x^{3/2}}\right)V_i^2\right\} \tag{3.111}$$

$$G_{TN}G_{BP}+G_{BN}G_{TP}\approx k_Nk_P\left\{2(V_y^2-V_y\gamma\sqrt{V_x}+\gamma^2V_x)+\left(\frac{1}{2}+\frac{1}{8}\frac{V_y\gamma}{V_x^{3/2}}+\frac{\gamma}{2\sqrt{V_x}}\right)V_i^2\right\} \tag{3.112}$$

Finally, the nonlinear resistance may be expressed by the following formula.

$$R_N\approx\frac{R_o(1+bV_i^2)}{1+aV_i^2} \tag{3.113}$$

$$R_o=\frac{2}{(k_N+k_P)(V_y-\gamma\sqrt{V_x})} \tag{3.114}$$

$$a=\frac{1}{4}R_o^2\left\{\frac{\frac{1}{16}V_y\gamma}{V_x^{3/2}}(k_N+k_P)^2-\frac{\gamma^2}{8V_x}(k_N^2+k_P^2)-\frac{1}{4}(k_N-k_P)^2\left(1+\frac{\gamma}{\sqrt{V_x}}\right)\right\} \tag{3.115}$$

$$b=\frac{\frac{1}{32}\gamma}{V_x^{3/2}(V_y-\gamma\sqrt{V_x})} \tag{3.116}$$

Now that an expression for the nonlinear resistance has been obtained, a nodal analysis of figure 3.16 is performed as shown below.

$$V_i - V_o = (R_L + R_N) C \frac{dV_o}{dt} \tag{3.117}$$

Now, equation (3.113) is substituted into the above equation to yield the following result.

$$V_i + aV_i^3 = V_o + aV_i^2 V_o + (R_L + R_o) C \frac{dV_o}{dt} + (aR_L + bR_o) CV_i^2 \frac{dV_o}{dt} \tag{3.118}$$

Now assume that the input signal is a sinusoid given by the following equation.

$$V_i = V_P \cos(\omega t) \tag{3.119}$$

The output voltage V_o is expressed by the Fourier series as shown below.

$$V_o = \sum_{k=0}^{\infty} \{ V_{ck} \cos(k\omega t) + V_{sk} \sin(k\omega t) \} \tag{3.120}$$

The third harmonic distortion HD_3 is a good indicator of the degradation. By plugging equations (3.119) and (3.120) into (3.118), the third harmonic distortion is given approximately by the following relation.

$$HD_3 = \frac{V_{S3}}{V_{C1}} = \frac{\frac{1}{4} V_P^2 (b - a) R_o C\omega}{1 + \frac{3}{4} aV_P^2 + \frac{1}{16} (aV_P^2)} \tag{3.121}$$

Note that to first order, the linear resistance R_L has no effect on the distortion. Also, the third harmonic distortion is proportional to the input frequency and the square of the signal amplitude. Thus, a fundamental trade-off between distortion and noise can be seen. Increasing signal swing relaxes noise constraints but makes the distortion worse. If dynamic range is fixed, a factor of 2

increase in signal swing results in a reduction in required capacitance by a factor of 4. However, the distortion saved by reducing the capacitance is cancelled to first order by the increase in signal swing. Therefore, to fundamentally improve distortion behavior it is necessary to either improve the linearity of the switch resistance or reduce the switch resistance. Linearity can be improved to some extent by matching the PMOS conductance to the NMOS conductance. Switch resistance can be reduced by using wider switches. The use of wider switches to limit distortion is eventually limited by the fact that as the switch is widened parasitic capacitances at the source and drain are also increased. If the switch is large enough, this parasitic capacitance dominates the capacitance C of equation (3.121). As a result, any improvement in distortion obtained by reducing switch resistance is cancelled out by the associated increase in capacitance.

Another way of reducing switch resistance is to increase the voltage swing on the clocks driving the switch, either by raising the supply voltage or by using a charge pump to boost the voltage. This technique is effective, but usually technology constraints such as breakdown voltage limit the maximum voltage that can be used.

Example:

This example is included to give a sense of the kind of performance degradation that can be expected.

Table 3.4 Example Parameters

supply voltage V_{DD}	5V
substrate doping N_{sub}	$8*10^{16}$ cm^3
gate oxide capacitance C_{ox}	1.535 $fF/\mu m^2$
nominal threshold voltage V_{T0}	0.9V
$2\phi_F$	0.6V
body effect parameter γ	$0.613V^{0.5}$
N channel mobility μ_n	524 $cm^2/V/s$
P channel mobility μ_p	160.5 $cm^2/V/s$

Table 3.4 Example Parameters

sampling capacitance C	6pF
input frequency f	2 MHz
W_N/L_N	512/1.2
W_P/L_P	512/1.2
a	-0.0909 V^{-2}
b	0.00352 V^{-2}
R_o	179 Ω
V_P	1.65 V
HD_3	$2.65*10^{-4}$ = -71.5dB

The distortion is also graphed versus channel width for a number of different supply voltages in figure 3.17. In this graph the signal amplitude is proportional to the supply voltage.

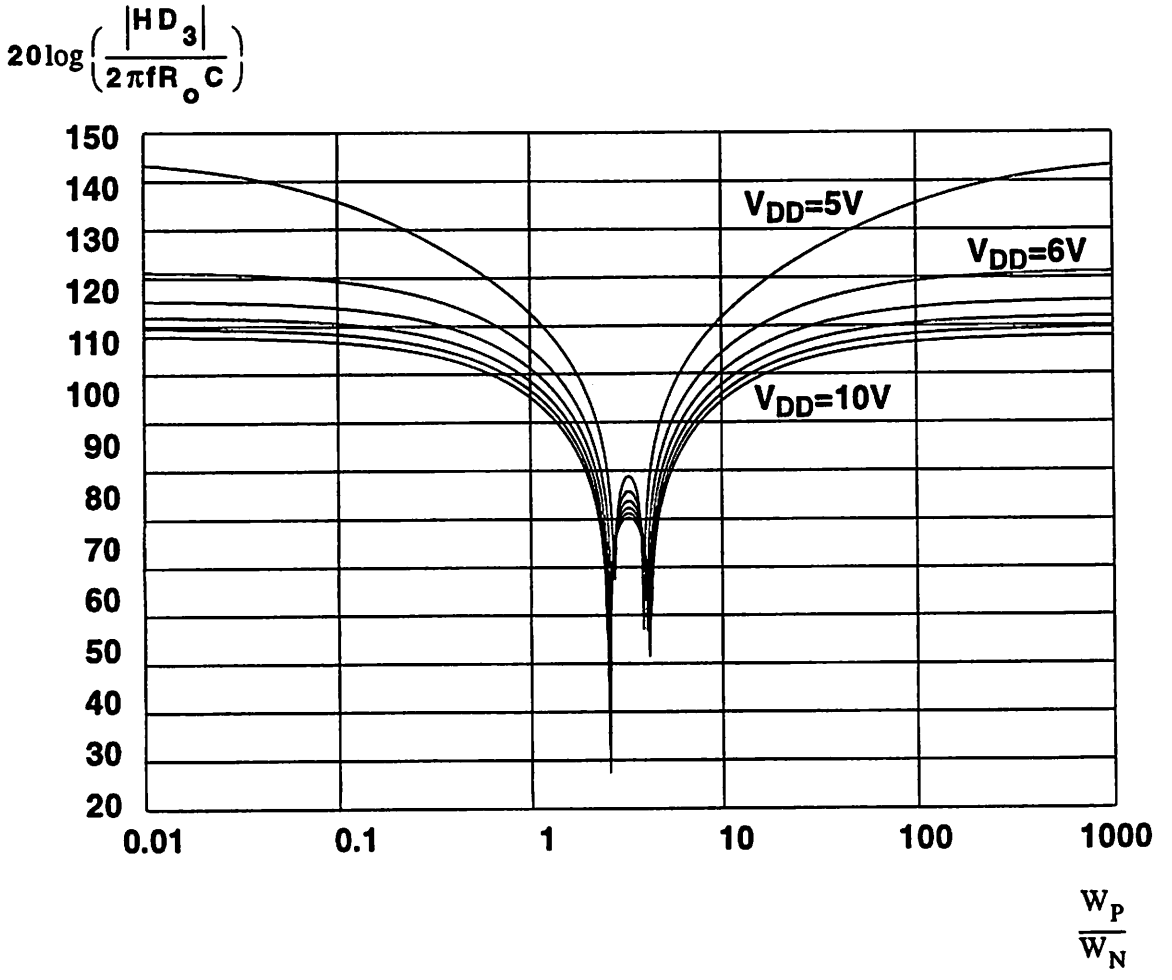


Figure 3.17 HD_3 versus W_P/W_N, W_N=512μm, V_P=0.2V_{DD}

Note that the distortion is can be reduced by increasing the clock voltage, even when the signal amplitude is increased in proportion to the supply voltage. The graph also indicates that the distortion can be greatly reduced by appropriately adjusting the ratio of the width of the PMOS device to the width of the NMOS device. In reality, other effects not included in the model used here would probably limit the extent to which the distortion could be reduced.

3.4 Error Correction Techniques

A number of error correction techniques have been developed to make high resolution analog to digital conversion possible in spite of the sources of error mentioned earlier in this chapter. Some of the techniques are based on adding circuit enhancements to reduce the error to a tolerable level. Analog offset cancellation, analog calibration, and capacitor error averaging are all examples of this type of technique. Other techniques do not attempt to fix the error but instead are based on design changes that make the error more tolerable. Digital error correction and digital calibration are examples of techniques that work this way. These error correction techniques are important because they have resulted in an improvement of the resolution capability of pipelined ADCs. This improvement has made it possible to design pipelines with resolutions that are limited by thermal noise rather than matching constraints.

3.4.1 Analog Offset Correction

As mentioned previously, comparator offsets are sources of error that can potentially limit the resolution of an analog to digital converter. The technique discussed here is one example of a technique to reduce the input referred offset voltage of an amplifier.[Degrauwe85][Ohara87] It is useful because a comparator is sometimes composed of a latch driven by an amplifier. In this type of design the offset is usually dominated by the offset of the amplifier. Also, amplifiers are used in pipelined analog to digital converters as residue amplifiers. The offset of the residue amplifier can sometimes be a problem. For example, the offset of the residue amplifier is a source of error in an analog to digital converter that uses parallel pipelines. An example of a parallel pipelined ADC is presented by Conroy[Conroy93].

In analog offset cancellation, the amplifier offset is sampled and stored at the input of an auxiliary amplifier that shares the same output as the main amplifier. The operation is illustrated conceptually in figure 3.18. Two phases of operation are illustrated. During the offset sampling phase, the main amplifier is connected in an open loop configuration, and the auxiliary amplifier is connected in a unity gain configuration. The result is that the offset voltage of the main amplifier scaled up by $\frac{g_m}{g_{mA}}$ is stored on the capacitor at the input of the auxiliary amplifier. During normal operation, the auxiliary amplifier inverts and amplifies the offset stored at its input and adds the result to the amplified offset from the main amplifier. The result is that the offset is cancelled out

by the auxiliary amplifier.

An analysis is now done in order to relate the effective offset voltage after correction to the input offset voltages of the main amplifier and the auxiliary amplifier and to the transconductance of each of the amplifiers. In order to perform this analysis, the system is modeled as shown in figure 3.19. First, the offset sampling phase is analyzed in order to determine the voltage V_C stored on the capacitor C_{AZ} at the input of the auxiliary amplifier.

$$V_C = g_m R_o V_{OSM} + g_{mA} R_o (V_{OSA} - V_C) \tag{3.122}$$

In the above equation, V_{OSM} is the offset voltage of the main amplifier, and V_{OSA} is the offset voltage of the auxiliary amplifier. The above equation is then solved for V_C.

$$V_C = \frac{g_m R_o V_{OSM}}{1 + g_{mA} R_o} + \frac{g_{mA} R_o V_{OSA}}{1 + g_{mA} R_o} \tag{3.123}$$

Now the block diagram corresponding to normal operation is analyzed in order to determine the output voltage in the presence of an input and the offset voltages.

$$V_o = g_m R_o (V_I + V_{OSM} - f V_o) + g_{mA} R_o (V_{OSA} - V_C) \tag{3.124}$$

In the above equation, f is the feedback factor. Now, equation (3.123) is substituted into equation (3.124), and the equation is solved for V_o.

$$V_o = \frac{g_m R_o V_I}{1 + f g_m R_o} + \frac{g_m R_o V_{OSM}}{(1 + g_{mA} R_o)(1 + f g_m R_o)} - \frac{g_{mA} R_o V_{OSA}}{(1 + g_{mA} R_o)(1 + f g_m R_o)} \tag{3.125}$$

The output voltage can be expressed as a function of the offset voltage as follows.

$$V_o = \frac{g_m R_o}{(1 + f g_m R_o)} (V_I + V_{OSeff}) \tag{3.126}$$

Then the effective input referred offset voltage V_{OSeff} is found by equating the above two equations to obtain the following result.

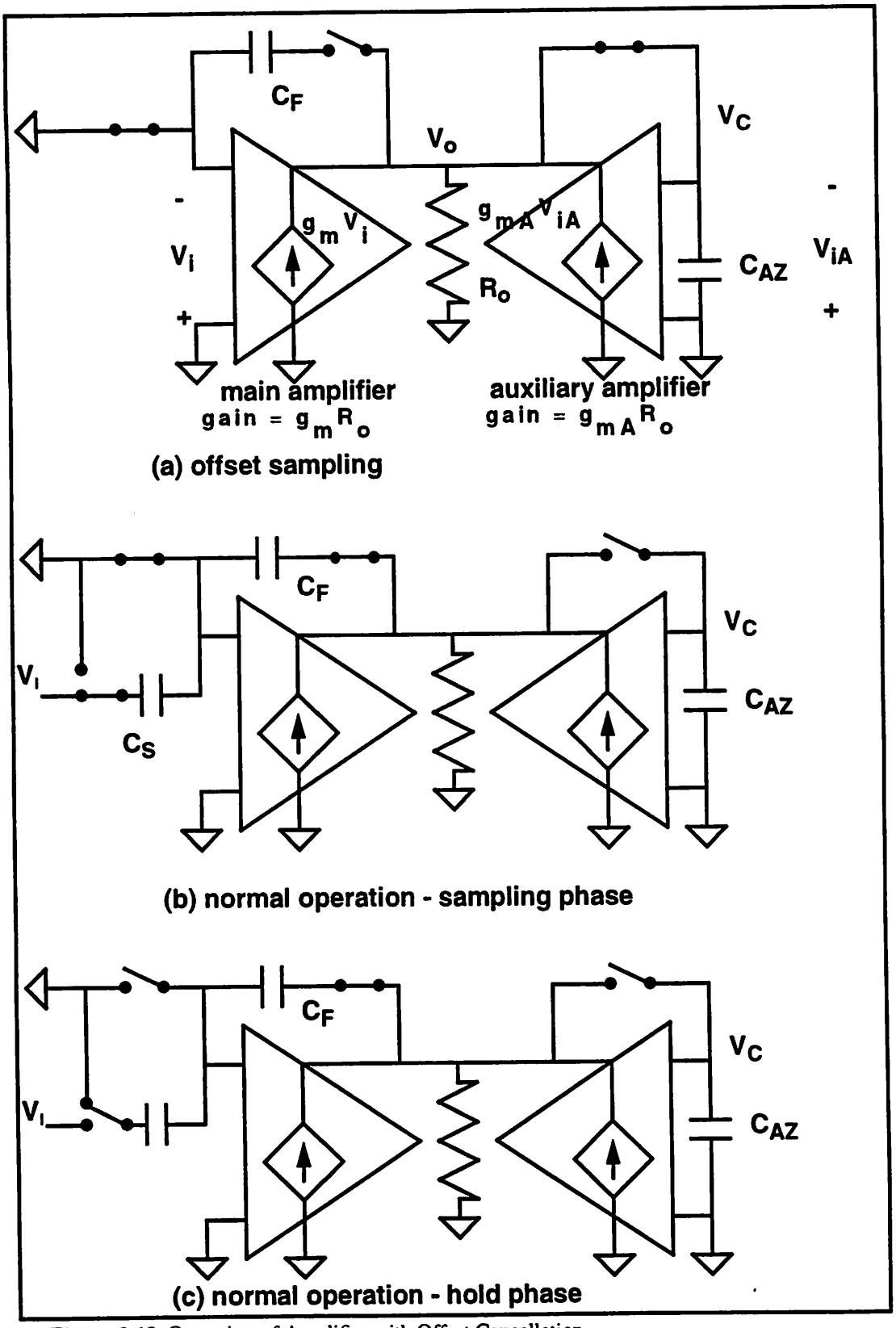
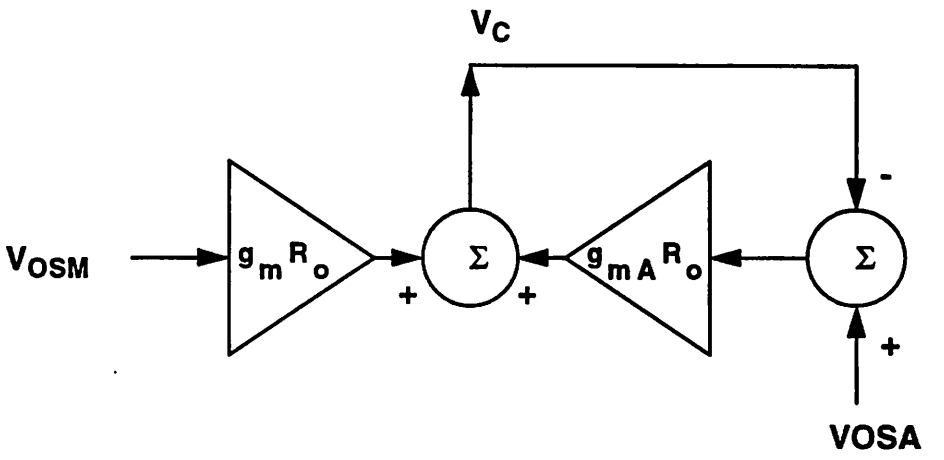
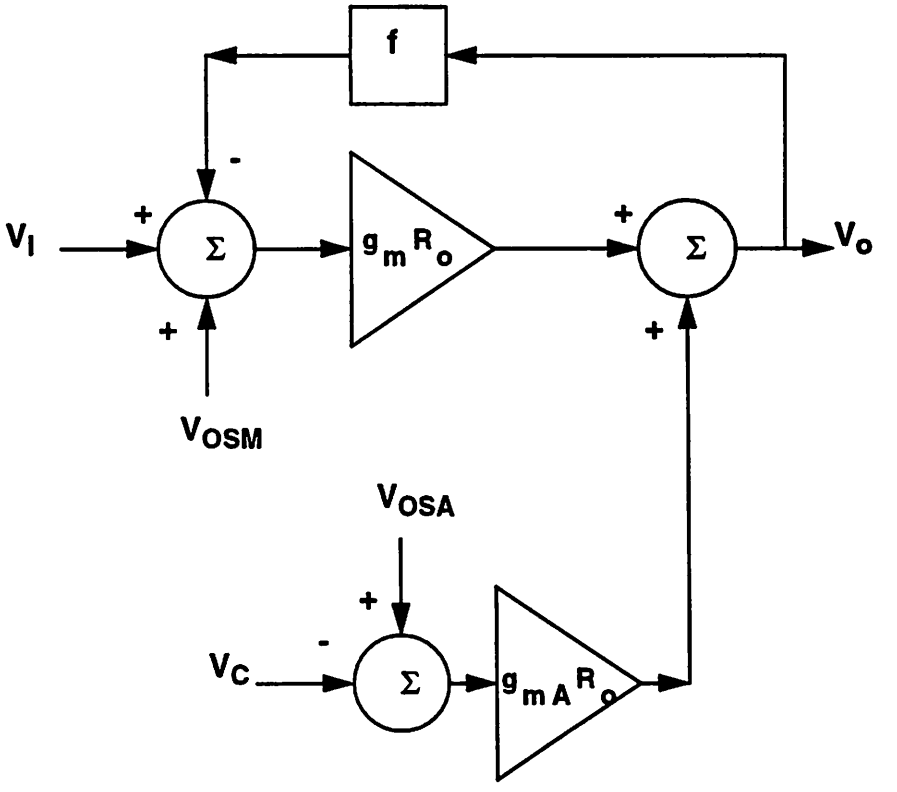


Figure 3.18 Operation of Amplifier with Offset Cancellation



(a) offset sampling



(b) normal operation

Figure 3.19 System Diagram of the Operation of Offset Corrected Amplifier

$$V_{OSeff} = \frac{V_{OSM}}{(1 + g_{mA}R_o)} + \frac{\frac{g_{mA}}{g_m}V_{OSA}}{(1 + g_{mA}R_o)} \tag{3.127}$$

Thus, the offset voltage of the main amplifier has been largely cancelled out. Unfortunately, the offset voltage of the auxiliary amplifier has been introduced as a new contributor to the effective offset voltage. Therefore, the degree of improvement obtainable by this technique is ultimately limited by the input offset voltage of the auxiliary amplifier.

The following analysis is performed to determine to analytically the minimum offset voltage this correction technique can achieve. If the offset voltages are approximately known, the offset cancellation circuit can be designed to minimize the effective offset as follows. First, it is useful to observe that the output resistance drops as the auxiliary amplifier becomes larger. Therefore, the output resistance R_o is rewritten as a function of the transconductance of the auxiliary amplifier as follows.

$$R_o = \frac{A_o}{g_m + g_{mA}} \tag{3.128}$$

A_o in the above equation is the gain the amplifier would have if the main amplifier were present alone with no auxiliary amplifier. Now equation (3.127) and equation (3.128) are combined to obtain an expression with R_o eliminated.

$$V_{OSeff} = \frac{g_m V_{OSM} + g_{mA}(V_{OSM} + V_{OSA}) + \frac{g_{mA}^2}{g_m}V_{OSA}}{g_m + (A_o + 1)\,g_{mA}} \tag{3.129}$$

Now to optimize the design, equation (3.129) is differentiated with respect to g_{mA}, the result is set to zero, and then the resulting equation is solved for g_{mA}.

$$\left.\frac{dV_{OSeff}}{dg_{mA}}\right|_{g_{mA} = g_{mAopt}} = 0 \tag{3.130}$$

$$g_{mAopt} = \left(\frac{g_m}{A_o + 1}\right)\left[-1 + \sqrt{1 + (A_o + 1)\left(\frac{A_o V_{OSM}}{V_{OSA}} - 1\right)}\right] \tag{3.131}$$

Now if $V_{OSA}=V_{OSM}$, the optimized result simplifies to the following formula for the transconductance of the auxiliary amplifier.

$$g_{mAopt} = \left(\frac{A_o - 1}{A_o + 1}\right) g_m \tag{3.132}$$

For this case, the optimized offset voltage is given by the following formula.

$$V_{OSeffopt} = \frac{2\left[(A_o + 1)V_{OSM} + (A_o - 1)V_{OSA}\right]}{(A_o + 1)^2} \tag{3.133}$$

Thus, the minimum obtainable offset voltage is approximately twice the sum of the offset voltages of the main and auxiliary amplifiers divided by the gain. In most cases, it is not attractive to design for this optimum case because the auxiliary amplifier would have to have a transconductance equal to that of the main amplifier. Doing this would add a significant amount of power dissipation assuming the speed is fixed. Furthermore, the open loop gain of the opamp is degraded if a large auxiliary amplifier is used.

3.4.2 Digital Comparator Error Correction

Digital comparator error correction[Taylor78][Lewis87] is a technique used to prevent comparator offsets from limiting the resolution of an analog to digital converter. In this technique, the comparator offsets are not corrected. Instead, the ADC is designed in a way that is tolerant of comparator offsets. Without digital error correction, the comparator offset must be no more than the least significant bit of the ADC. With digital error correction, larger offsets can be tolerated. This technique is attractive because it allows the use of simplified comparators. This can potentially save hardware and power. This technique also allows analog to digital converters to achieve resolutions that would not be possible without it.

To show how this works, the long division analogy is again useful. Analog to digital conversion with comparator offsets is analogous to long division with incorrect decision levels in the multiplication table. The result of the incorrect decision levels is that sometimes a large residue outside of the range of the multiplication table is generated. To get around this problem, the range of the multiplication table can be extended so that the large residues can be accommodated. The resulting quotient does not look the same as before, but no information has been lost. If an attempt is made to reconstruct the dividend from the residue and the quotient, the correct answer is

obtained as shown.

multiplication table

quotient digit	comparator thresholds	reference values
-9	-66	-63
-8	-55	-56
-7	-50	-49
-6	-41	-42
-5	-36	-35
-4	-26	-28
-3	-19	-21
-2	-13	-14
-1	-8	-7
0	0	0
1	8	7
2	13	14
3	19	21
4	26	28
5	36	35
6	41	42
7	50	49
8	55	56
9	66	63

```
                          incorrect output digit
                                  |
                          7.143(-2)57 ...
                      7| 50.0000   00
subtracted ref         - 49
residue                  1
amplified residue        1 0
subtracted reference     - 7
residue                    3
amplified residue          30
subtracted reference     - 28
residue                    2
amplified residue          20
subtracted reference     - 21 <wrong reference subtracted
residue                  - 1
amplified residue        - 10
subtracted reference      +14
residue                     4
amplified residue             40
subtracted reference        -35
residue                       5
amplified residue             50
subtracted reference        -49
residue                       1
                            ...
```

reconstruction of the dividend

$$\text{dividend} = \frac{1}{10^6} + \frac{49}{10^6} + \frac{35}{10^5} - \frac{14}{10^4} + \frac{21}{10^3} + \frac{28}{10^2} + \frac{7}{10} + 49 = 50 \tag{3.134}$$

As shown above, the out of range residues were anticipated, and extra decision levels were added to accommodate them. Thus, no information was lost.

To apply this technique to analog to digital conversion, extra comparators are added beyond the number that would be needed if the comparators were free of offsets. The full scale linear range of the amplifier must also be extended to accommodate residues that are larger than expected. The following plot of the input-output characteristic of one pipeline stage helps to further illustrate how digital error correction works. This plot is useful for deriving a relation between the number of comparators and the maximum allowable comparator offset.

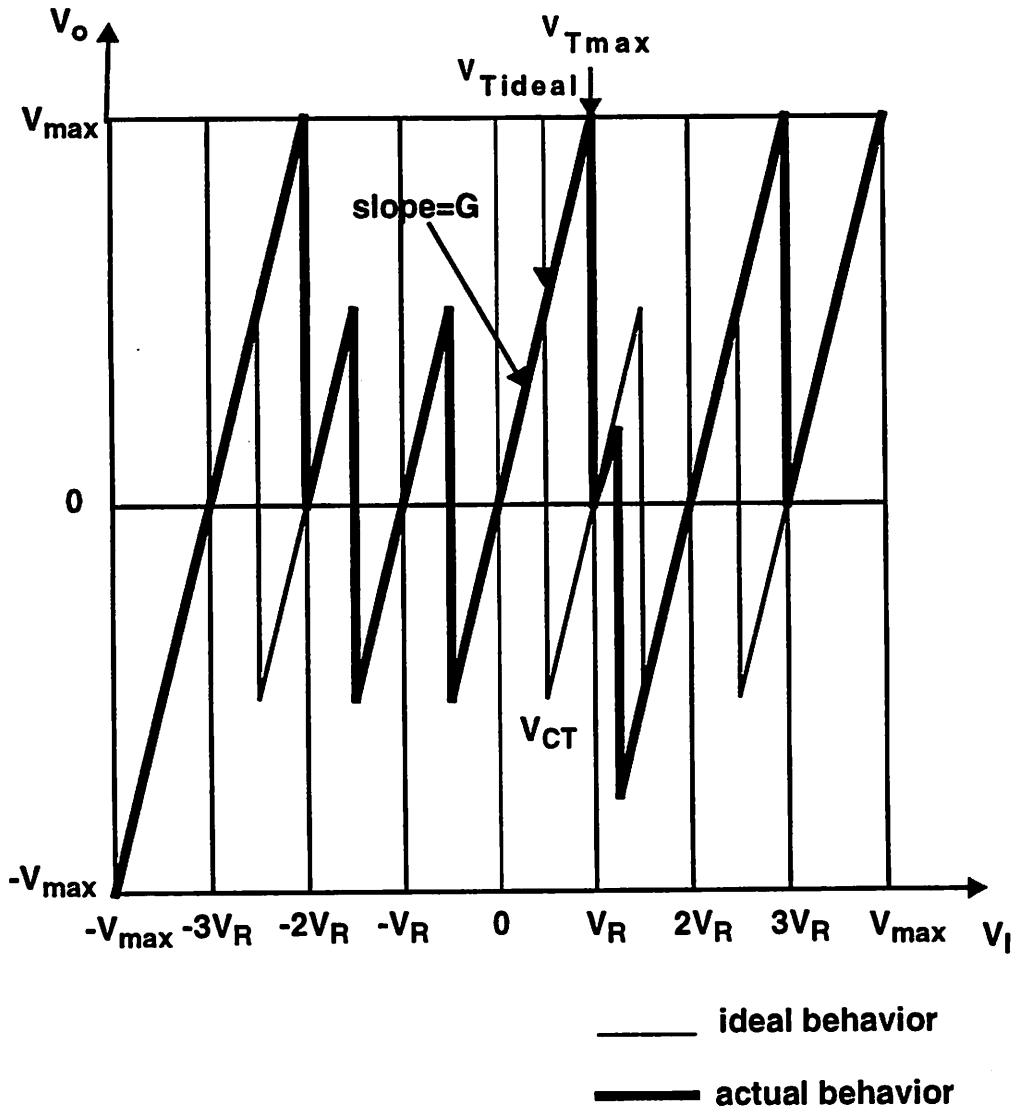


Figure 3.20 Transfer Function from Input to Output of a Pipeline Stage with Comparator Offsets

Note that in this example the value of the amplified residue is never more than $0.5V_{max}$ in the absence of comparator offsets. If comparator offsets are present, the amplified residue is still within range provided that the comparator offset does not go beyond a certain limit. This maximum allowable comparator offset will now be derived. From figure 3.20 the following relation can

be inferred.

$$GV_{Tmax} = V_{max} \tag{3.135}$$

The maximum allowable comparator offset $V_{coffsetmax}$ at DC is equal to the difference between the maximum comparator threshold given by the above equation and the ideal comparator threshold. For higher input frequencies, the maximum allowable comparator offset is reduced if there is a timing mismatch Δt between the comparator sampling instant and the sample and hold sampling instant[Cho95].

$$2\pi f_{max} V_{max} \Delta t + V_{coffsetmax} = V_{Tmax} - V_{Tideal} = \frac{V_{max}}{G} - \frac{1}{2}V_R \tag{3.136}$$

f_{max} in the above equation is the maximum allowable input frequency. The reference value V_R in the above equation is given by equation (3.5). Substituting for V_R results in the following formula for the maximum comparator offset.

$$2\pi f_{max} \Delta t V_{max} + V_{coffsetmax} = \frac{V_{max}}{G}\left(1 - \frac{1}{n_C}(G-1)\right) \tag{3.137}$$

From this formula it is evident that the maximum tolerable comparator offset is increased by reducing the interstage gain (residue amplifier gain) and by increasing the number of comparators.

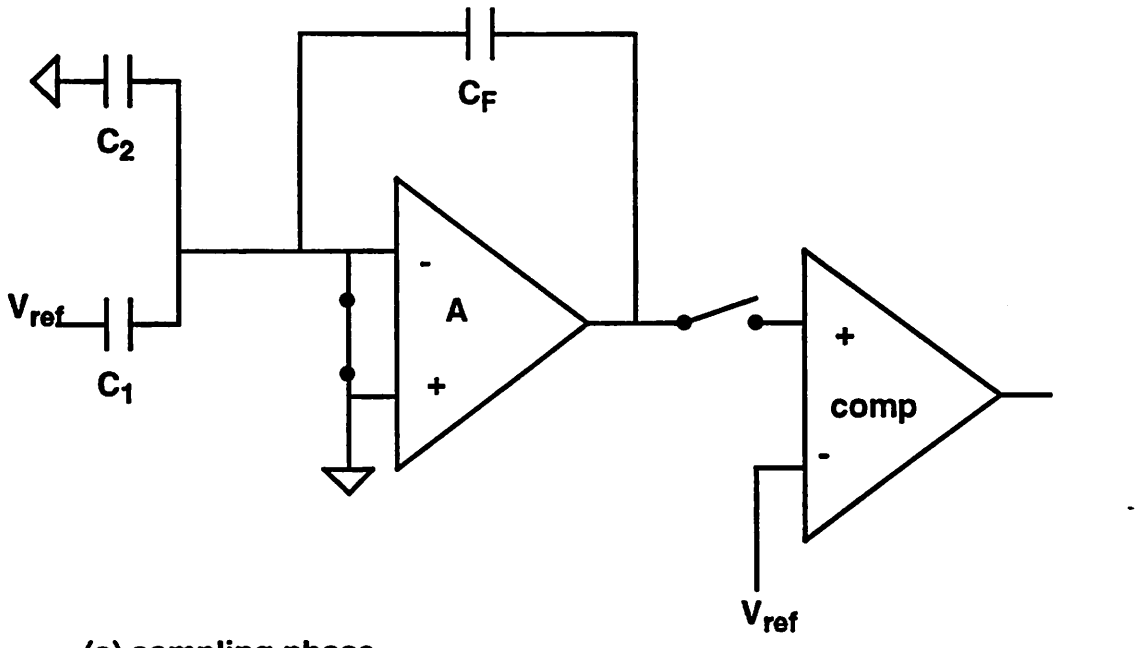
3.4.3 Analog DAC/Gain Calibration

As noted in section 3.3.4, mismatched components contribute to nonlinearity in digital to analog converters. An analog calibration technique has been developed by Lin[Lin91] to correct this problem. This technique also corrects the gain error in the residue amplifier. In this technique, the matching between two components is tested. Then based on the result of the test, one of the components is sized up or down. The test is then run again. This process continues until the components are well matched.

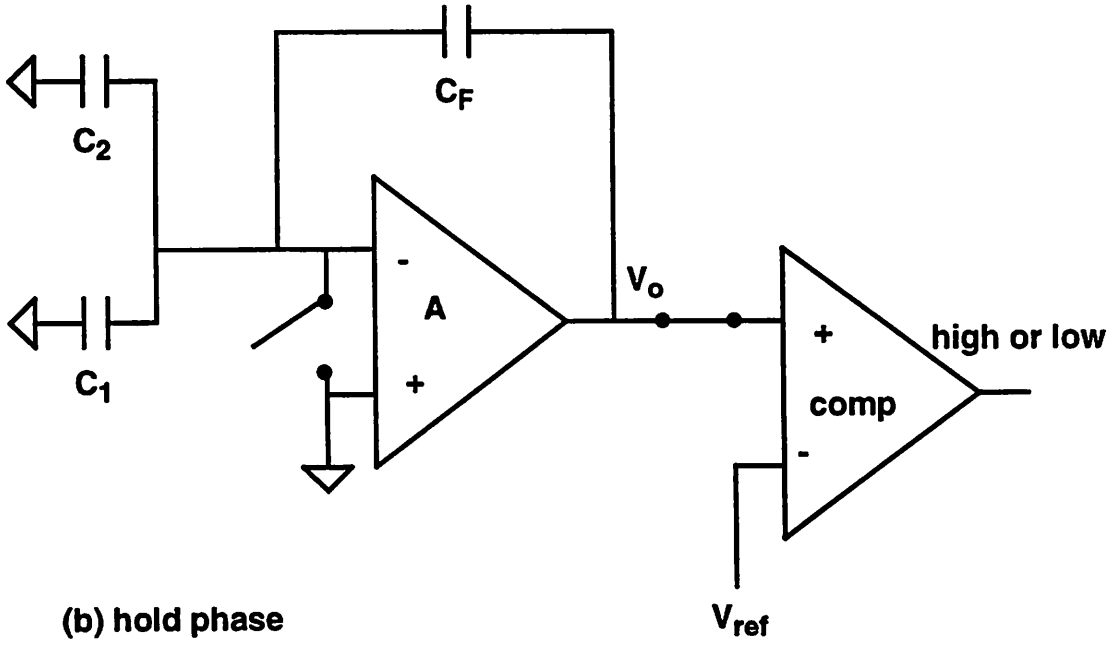
How this technique works is now described further by explaining how it is applied to a switched capacitor DAC/residue amplifier implementation such as the one discussed in section 3.2. Nominally all the capacitors are the same size. However, in a real implementation, the capacitors do not match perfectly. To correct for this problem, small trimming capacitors are put in parallel with capacitors C_1 and C_2 and switched in or out as necessary to match these capacitors to the

feedback capacitor C_F.

Figure 3.21 illustrates how the capacitor sizes are compared to determine whether C_1 is too big or too small. During the sampling phase, the reference voltage is sampled onto capacitor C_1. Then, during the hold phase C_1 is switched to ground. The reference voltage is amplified by C_1/C_F and sampled by the comparator. Then a decision is made as to whether the amplified output is larger or smaller than the reference voltage. If the amplified output is larger, C_1 is made smaller by switching out a trimming capacitor. Otherwise C_1 is made larger.



(a) sampling phase



(b) hold phase

Figure 3.21 Analog DAC/Gain Calibration

Note that this scheme is sensitive to the input referred offset voltage of the amplifier. If an offset is present, the comparator will think there is a mismatch in capacitor sizes when there is none. Therefore, an offset cancellation technique such as that described in section 3.4.1 must be used to make this technique effective.

3.4.4 Capacitor Error Averaging

Capacitor error averaging[Song88] is another technique for achieving a precise gain of 2 in the residue amplifier. In this technique, calibration is not used. Instead, the residue is amplified twice, each time with a different feedback capacitor. By averaging these two results together, the error introduced by capacitor mismatch can be averaged out. The disadvantage of this technique is that it requires the use of two amplifiers rather than one for each sample and hold stage, and an extra clock cycle is needed since the signal is amplified twice. Figure 3.22 illustrates what happens during each phase. During the sampling phase, the input signal V_I is sampled onto capacitors C_1 and C_2. These two capacitors do not match perfectly. Therefore, using either capacitor as the feedback capacitor will result in a gain that is too large or too small. Therefore, there are two hold phases, one where C_1 is used for feedback and the other where C_2 is used for feedback. During the first hold phase, C_2 is used for feedback, and the output of the first amplifier is sampled onto capacitor C_4 while its inverse is sampled onto capacitor C_3. During the second hold phase, C_1 is used for feedback, and the new output is connected to C_3.

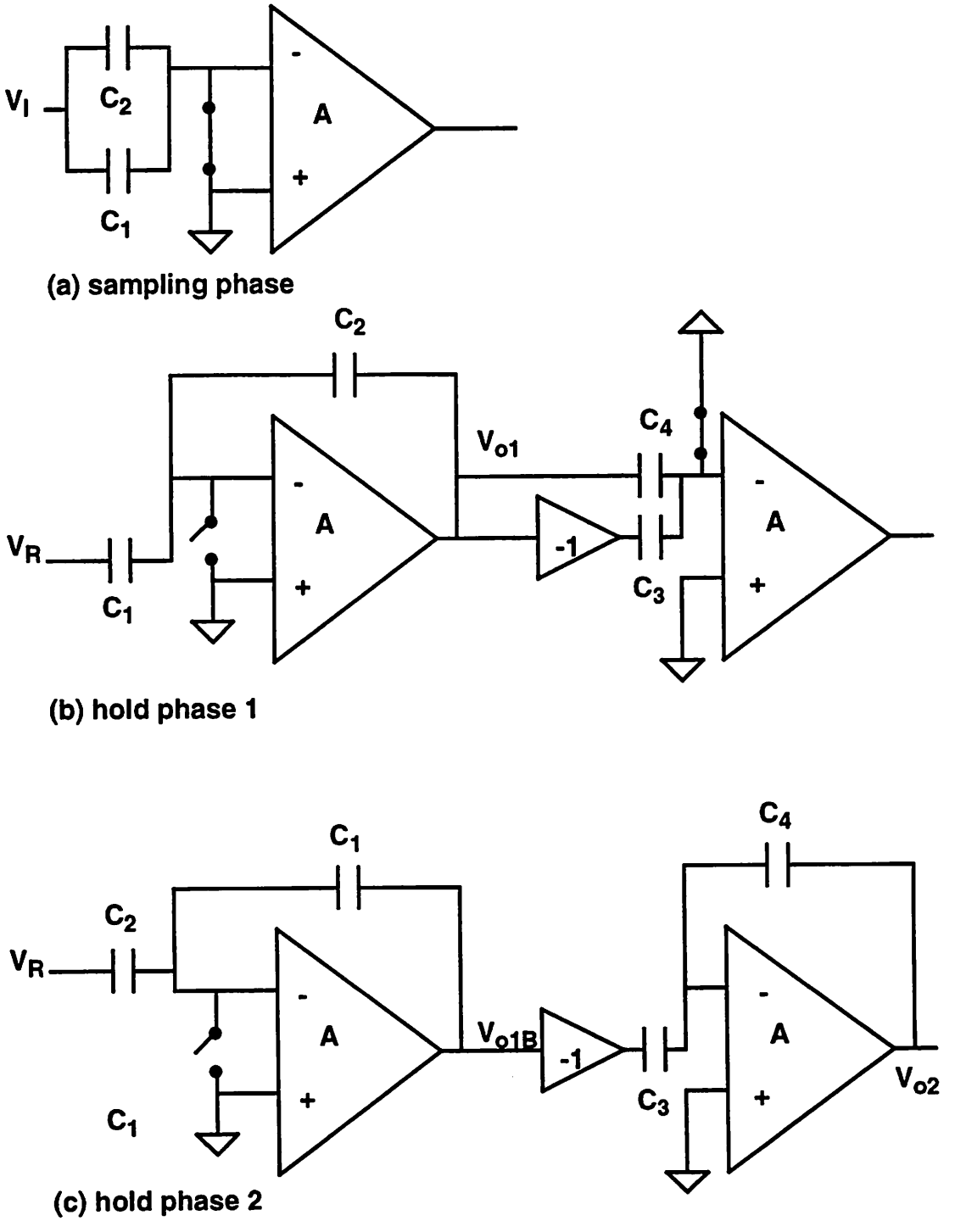


Figure 3.22 Capacitor Error Averaging

Now the above figure is analyzed to get the output voltage V_{o2} as a function of the input voltage V_I and the reference voltage V_R. From the sampling phase and hold phase 1, the following expression is obtained.

$$V_{o1} = \left(\frac{C_1 + C_2}{C_2}\right)V_I - \frac{C_1}{C_2}V_R \tag{3.138}$$

From the sampling phase and hold phase 2, the following expression is obtained.

$$V_{o1B} = \left(\frac{C_1 + C_2}{C_1}\right)V_I - \frac{C_2}{C_1}V_R \tag{3.139}$$

Finally, from hold phases 1 and 2, the following expression is obtained.

$$V_{o2} = \frac{(C_4 - C_3)\,V_{o1}}{C_4} - \frac{C_3}{C_4}V_{o1B} \tag{3.140}$$

Now, combining equations (3.138), (3.139), and (3.140), the following formula for V_{o2} is obtained.

$$V_{o2} = \left(1 + \frac{C_1}{C_2}\right)V_I + \frac{C_3}{C_4}\left(\frac{C_2^2 - C_1^2}{C_1C_2}\right)V_I - \frac{C_1}{C_2}V_R + \frac{C_3}{C_4}\left(\frac{C_1^2 - C_2^2}{C_1C_2}\right)V_R \tag{3.141}$$

Now C_1 and C_2 are rewritten in terms of a nominal capacitance C and an error capacitance δC.

$$C_1 = C - \frac{1}{2}\delta C \tag{3.142}$$

$$C_2 = C + \frac{1}{2}\delta C \tag{3.143}$$

By combining equations (3.141), (3.142), and (3.143), and using the following approximation, the effect of mismatches on the output may be seen.

$$\frac{1}{1 + \frac{\delta C}{C}} \approx 1 - \frac{\delta C}{C} \quad \text{for} \quad \frac{\delta C}{C} \ll 1 \tag{3.144}$$

$$V_{o2} \approx \left(2 - \frac{\delta C}{C} + \frac{2C_3}{C_4}\frac{\delta C}{C}\right)V_I - \left(1 - \frac{\delta C}{C} + \frac{2C_3}{C_4}\frac{\delta C}{C}\right)V_R \tag{3.145}$$

Note that if $C_3 = 0.5C_4$, the mismatches are well cancelled. Even if C_3 is not exactly $0.5C_4$, the

mismatch cancellation is very good.

3.4.5 Digital DAC/Gain Calibration

Digital DAC/gain calibration is a technique developed by Karanicolas[Karanicolas93]. A generalized discussion of this technique is also discussed by Conroy[Conroy94]. Digital calibration is similar to the analog calibration technique discussed in section 3.4.3 in that it is a technique used to eliminate errors caused by the nonlinearity of the DAC and by the gain error of the residue amplifier. However, unlike in the analog calibration technique where components are measured and then corrected until they match, in the digital calibration technique component ratios are measured, quantized by following stages of the ADC, and then stored digitally. No effort is made to correct the components to make them match. The results of the measurements are later used with the digital output of the analog to digital converter to generate a corrected digital output with improved linearity. This technique has the disadvantage that it requires the use of digital adders not required by the analog technique. However, the technique greatly relaxes the design requirements for some of the analog components and results in a design that is very reliable and robust.

The concept of digital calibration may be better understood by going back to the long division analogy. In the example discussed in section 3.3.4, the quotient is wrong because the multiplication table is wrong. In other words, the values subtracted from the dividend or previous residue to generate a new residue are wrong. However, all of the information about the dividend is contained in the incorrect quotient and incorrect multiplication table. Therefore, none of the information about the dividend is lost. The incorrect quotient was simply obtained by a nonlinear transformation of the dividend. If the inverse of the nonlinear transformation is done, the dividend can be recovered fully. This fact is illustrated below for the example discussed in section 3.3.4. Recall the incorrect multiplication table shown below.

quotient digit	comparator thresholds	reference values
0	0	0
1	7	8
2	14	13
3	21	19
4	28	26
5	35	36
6	42	41
7	49	48
8	56	55

9	63	66

This multiplication table was used to generate the incorrect quotient of 7.295555 and a residue of 4. If an attempt is made to reconstruct the dividend by simply multiplying this quotient by 7, the result is 51.0688... This result is off by 2.1%. However, the dividend can be reconstructed by using the quotient and the incorrect multiplication table as shown below.

$$\text{dividend} = 48 + \frac{13}{10} + \frac{66}{100} + \frac{36}{1000} + \frac{36}{10^4} + \frac{36}{10^5} + \frac{36}{10^6} + \frac{4}{10^6} = 50.000 \qquad (3.146)$$

In a calibrated pipelined ADC, the procedure works as follows. In each stage of the pipeline, the comparators divide the range of possible signal values into a set of smaller ranges or bins. Associated with each bin is a digital code and DAC weight. Each stage of the ADC assigns the signal to a certain bin and generates a residue signal. When the conversion is complete, the signal has a bin assignment for each stage of the ADC. Thus, the signal is assigned a digital code and weight by each stage of the ADC. To reconstruct a linear digital representation of the input signal, the weights are all summed. This summation of weights to linearly reconstruct the input is analogous to the sum in equation (3.146). In that case there is a term in the sum associated with each digit of the quotient. In the pipelined ADC, there is a term in the sum for each stage of the pipeline.

To see how this might be implemented in a real circuit, the following example is considered. The circuit consists of two comparators, a switched capacitor DAC, and a residue amplifier, and is followed by an analog to digital converter. In this example, the ADC is the remainder of the pipeline. A circuit diagram is shown in figure 3.23, and a hypothetical plot of the amplified residue versus the input is shown in figure 3.24.

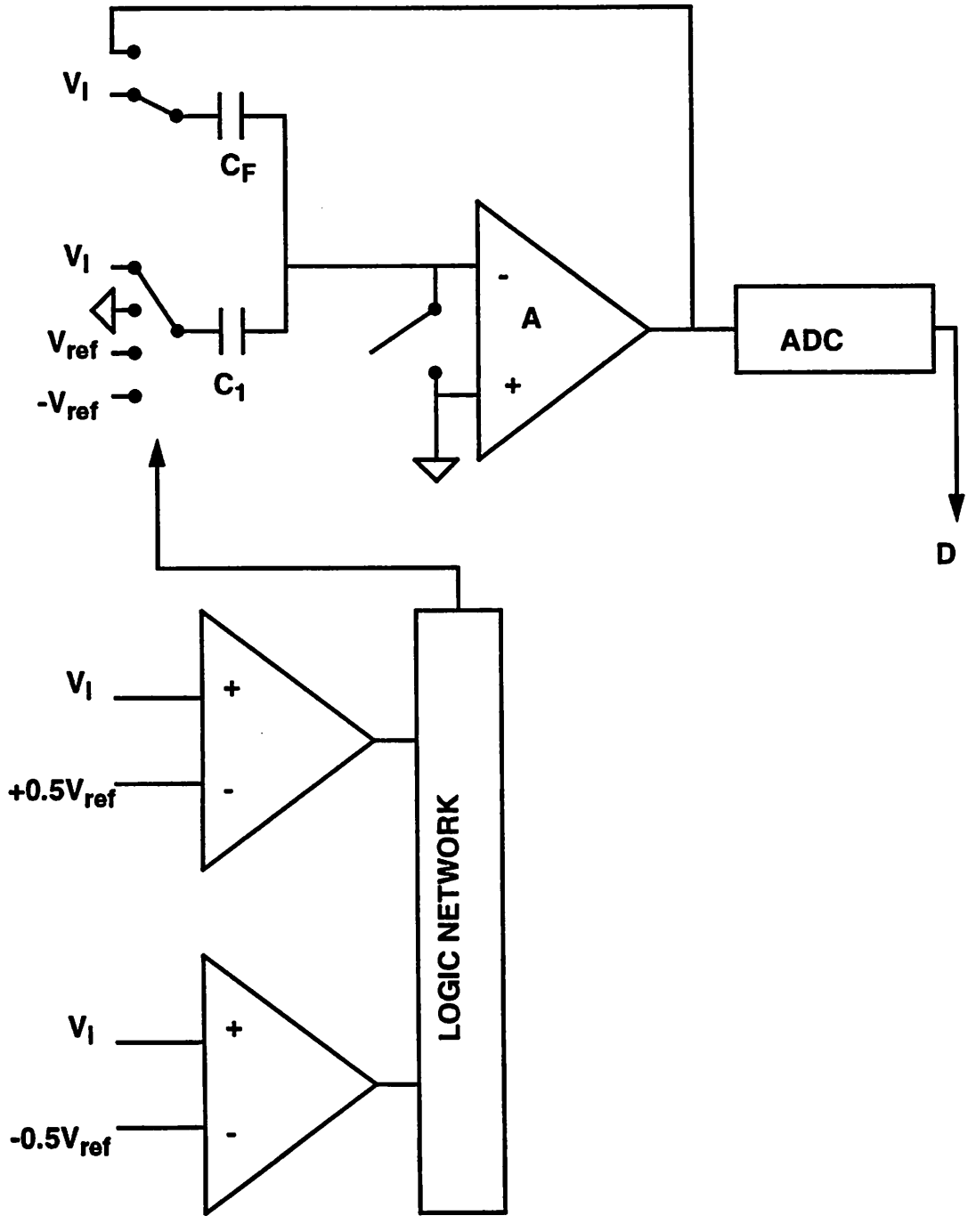


Figure 3.23 Example of a Pipeline Stage Containing Two Comparators and Requiring Calibration

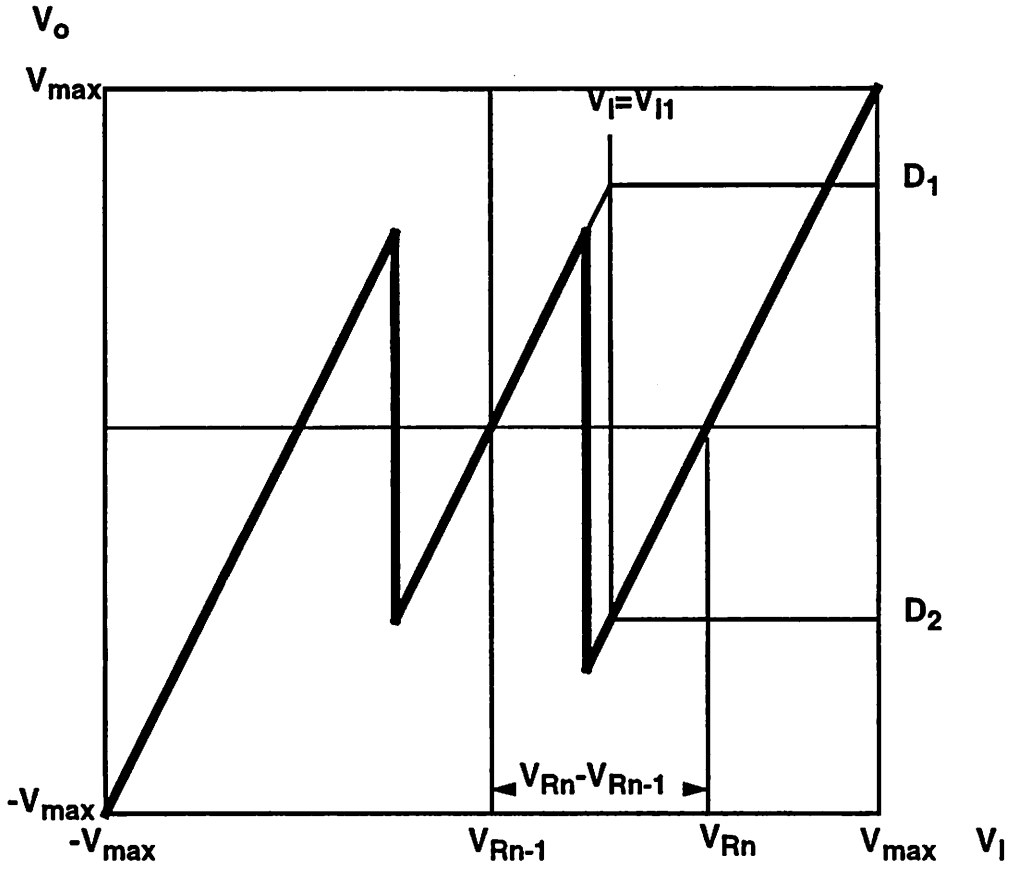
The height of the discontinuity in figure 3.24 is related to the bin weights in the following way. Imagine two signals on opposite sides of the discontinuity but infinitely close to it. When the linear reconstruction of the inputs is done, the results for both signals should match each other. If w is used to denote the weights from the stage of interest, and D is used to denote the code resulting from the rest of the ADC, then

$$code_1 = w_1 + D_1 \tag{3.147}$$

$$code_2 = w_2 + D_2 \tag{3.148}$$

It is desired that the codes match. Therefore,

$$w_2 - w_1 = D_1 - D_2 \tag{3.149}$$



$$w_n - w_{n-1} = D_1 - D_2 = G\,(V_{Rn} - V_{Rn-1})$$

Figure 3.24 Transfer Characteristic of a Pipeline Stage with Two Comparators and Requiring Calibration

$D_1 - D_2$ is the height of the discontinuity. Thus, the height of the discontinuity is the difference in the weights corresponding to the bins on either side of the discontinuity. Therefore, the weights are measured by measuring each of the discontinuities.

The method for measuring the discontinuities is as follows. Imagine, that a signal V_{I1} is applied to the input of the pipeline stage of interest. This signal is close to the threshold for the discontinuity as shown in figure 3.24. Two measurements are then made with this signal. In the first

case, the DAC level immediately below the comparator threshold is subtracted from the input signal. The resulting residue for this case is then amplified and quantized by the rest of the pipelined ADC. Therefore, the residue output D_1 is obtained from the ADC. Then, the input signal is kept the same, and the DAC level immediately above the comparator threshold is subtracted from the input signal. The residue for this case is also amplified and quantized by the rest of the pipeline. For this case, the residue output D_2 is obtained. D_1-D_2 is then computed to obtain the result for the weight difference. This measurement is done at each threshold. Thus, for the circuit shown here, two measurements are made. However, three bins are present, and three weights need to be assigned. Therefore, it seems that some information is still lacking. Actually, the information from the two measurements performed is sufficient to ensure the linearity of the ADC. The assignment of one of the weights can be arbitrary. For instance, in this example the middle weight can be assigned the value zero. With this assignment made, the two measurements provide enough information.

The accuracy of the calibration is dependent on the accuracy of the ADC used to perform the calibration. DNL in the ADC used to measure the weights corrupts the measurement, and the new ADC that includes the measured pipeline stage and the rest of the ADC will have DNL at the comparator thresholds. This presents a problem if the pipelined ADC is being used to calibrate itself, because one might wonder how an n bit DNL would be realized if, the ADC being used in the measurement has a resolution that is good only to n-x bits. Fortunately, the DNL caused by the ADC during this measurement is scaled down because the measurement occurs on the amplified residue. Therefore, if the DNL of the ADC without the calibrated stage is x, then the DNL of the ADC with the calibrated stage is x/G where G is the gain of the residue amplifier. Thus, it is possible to calibrate a pipelined ADC starting near the tail end and working forward toward the front. The weights for the stage near the tail end are measured, and these weights are used when the weights for the stage before it are measured.

Some accumulation of errors does occur during the calibration process. However, this accumulation of errors can be kept to a minimum if the pipeline is designed in such a way that small variations in the input signal only affect the code produced by the end stages of the pipeline. In the example shown this is accomplished in the following way. The input signal is close to a comparator threshold during calibration. This results in the output of the first stage being close to either

$+0.5V_{max}$ or $-0.5V_{max}$. When $0.5V_{max}$ or $-0.5V_{max}$ is quantized by the following stage of the pipeline the output of the following stage of the pipeline is close to zero. This zero is passed to the subsequent stages of the pipeline, each of which also produces an output near zero. Eventually, when the signal reaches the last stages of the pipeline it gradually drifts away from zero either due to mismatches or to noise. The result is that the codes produced by many of the pipeline stages are identical for both the high case and the low case. This means that any errors in the calibration of these stages will not contribute to the error in the stage currently being calibrated. As a result, the only errors accumulated in the calibration come from the tail end of the pipeline, and the error contribution of these stages is very small.

Figure 3.1 shows a block diagram of how a digitally calibrated pipelined ADC might operate during normal operation. The operation is similar to the operation of an uncalibrated pipeline with the following exception. Instead of incorporating the ADC output bits from each stage directly into the digital output word, the bits from each stage are instead used as the address to a lookup table. A decoder translates the address into a signal that selects the appropriate DAC weight from the lookup table. This DAC weight is then added to the digital result from the previous stage. At the end of the pipeline, the final digital output word is the sum of each of the selected DAC weights, one from each stage of the pipeline.

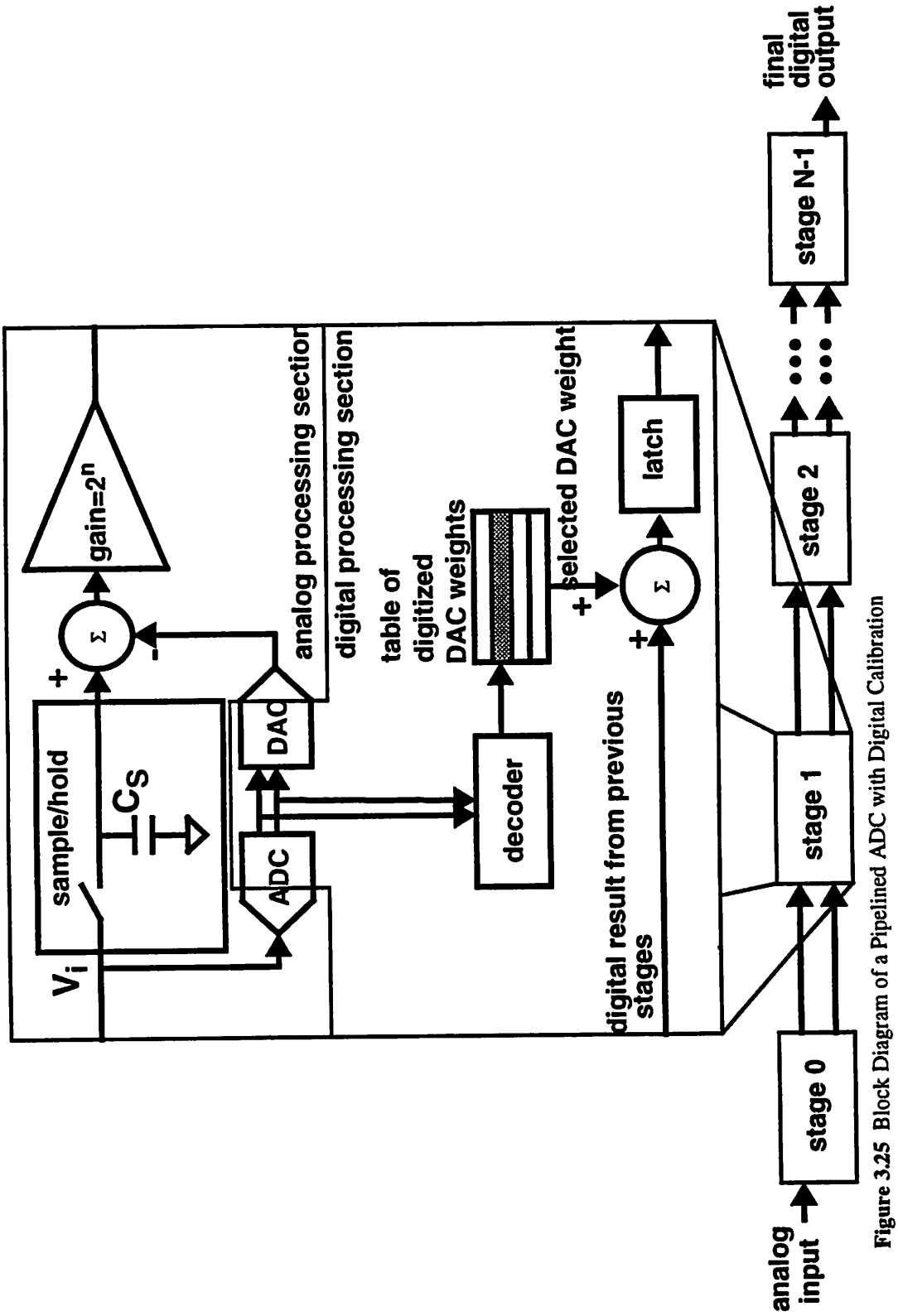


Figure 3.25 Block Diagram of a Pipelined ADC with Digital Calibration

References

[Cho95] Thomas B. Cho, "Low-Power Low-Voltage Analog-to-Digital Conversion Techniques Using Pipelined Architectures," Memorandum No.UCB/ERL M95/23, Electronics Research Laboratory, U. C. Berkeley, April 1995, pp. 98-102.

[Chuang82] C. T. Chuang, "Analysis of the Settling Behavior of an Operational Amplifier," *IEEE J. Solid-State Circuits*, vol. SC-17, no. 1, Feb. 1982, pp. 74-80.

[Conroy90] Cormac S. G. Conroy, private communication.

[Conroy93] Cormac S. G. Conroy, David W. Cline, and Paul R. Gray, "An 8-b 85-MS/s Parallel Pipeline A/D Converter in 1-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 447-454.

[Conroy94] Cormac 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, pp. 133-179.

[Degrauwe85] M. Degrauwe, E. Vittoz, and I. Verbauwhede, "A Micropower CMOS-Instrumentation Amplifier," *IEEE J. Solid-State Circuits*, vol. SC-20, no. 3, June 1985, pp. 805-807.

[Karanicolas93] Andrew N. Karanicolas, Hae-Seung Lee, and Kantilal L. Bacrania, "A 15-b 1-Msample/s Digitally Self-Calibrated Pipeline ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1207-1215.

[Keys94] Cynthia D. Keys, "Low-Distortion Mixers for RF Communications," Ph.D. Dissertation, University of California, Berkeley, Dec. 1994, pp. 40-67.

[Kusumoto93] Keiichi Kusumoto, Akira Matsuzawa, and Kenji Murata, "A 10-b 20-MHz 30-mW Pipelined Interpolating CMOS ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1200-1206.

[Lewis87] Stephen H. Lewis and Paul R. Gray, "A Pipelined 5-Msample/s 9-bit Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 954-961.

[Lin91] Yuh-Min Lin, Beomsup Kim, and Paul R. Gray, "A 13-b 2.5-MHz Self-Calibrated Pipelined A/D Converter in 3-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, April 1991, pp. 628-636.

[Ohara87] Harlan Ohara, Hung X. Ngo, Michael J. Armstrong, Chowdhury F. Rahim, and Paul R. Gray, "A CMOS Programmable Self-Calibrating 13-bit Eight-Channel Data Acquisition Peripheral," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 930-938.

[Song88] Bang-Sup Song, Michael F. Tompsett, and Kadaba R. Lakshmikumar, "A

12-bit 1-Msample/s Capacitor Error-Averaging Pipelined A/D Converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1324-1333.

[Taylor78] Stewart S. Taylor, "High Speed Analog-to-Digital Conversion in Integrated Circuits," Ph.D. Dissertation, University of California, Berkeley, Dec. 1978, pp. 30-42.

[Yang90] Howard C. Yang and David J. Allstot, "Considerations for Fast Settling Operational Amplifiers," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 3, March 1990, pp. 326-334.

CHAPTER 4

SAMPLE AND HOLD AMPLIFIER ARCHITECTURES AND OPTIMIZATION

The sample and hold amplifier is a key component of pipelined analog to digital converters. It is also an important building block in other electronic systems, such as switched capacitor filters. Settling time and power dissipation are key issues in the design of these amplifiers. Therefore, issues of speed and power optimization for these amplifiers are studied here for several of the most important amplifier architectures. This chapter begins with a discussion of techniques used to model the current and transconductance in MOS devices since these models directly affect speed estimates. Then, these models are used in the later part of this chapter to assist in the transient analysis of a number of different amplifier architectures. At the end of this chapter, the settling performance of the various amplifier architectures is compared. Throughout this chapter, the amplifiers are assumed to operate in a switched capacitor environment. Furthermore, only a linear settling analysis is performed here. Slewing effects are neglected in this study.

4.1 MOSFET Models for Transient Analysis

Before beginning the transient analysis of sample and hold amplifiers, it is worthwhile to discuss the circuit modeling of the metal oxide semiconductor field effect transistor (MOSFET). The modeling of the transconductance of this device is of key importance to the transient analysis of amplifiers using MOSFETs. The models used here were chosen because they model some of the most important second order effects in MOS devices without adding too much complexity. These models and other more accurate (and more complex) models are discussed by Tsividis.[Tsividis87-A][Tsividis87B][Tsividis87C] Figure 4.1 shows a schematic representation for the MOSFET to be used throughout this document.

(a) N channel MOSFET (b) P channel MOSFET

Figure 4.1 MOSFET Schematic Symbols Used in this Document

4.1.1 Long Channel Model for the MOSFET

The long channel model for the MOSFET is a useful tool for getting a basic description of device behavior. Because of its simplicity, it is especially useful in hand analysis, and it can be used to obtain analytic results to optimization problems. These analytic results are useful because they are simple enough to provide a high level of intuition.

In the long channel model the saturation drain current I_D of the MOSFET is given by the following function of the gate to source voltage V_{GS}[Muller86].

$$I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS}-V_T)^2 \tag{4.1}$$

In the above equation, μ represents the mobility, C_{ox} represents the capacitance per unit area between the gate and channel, and V_T represents the threshold voltage of the transistor. The width of the device is represented by W and the channel length by L. The transconductance g_m is then given by differentiating the drain current with respect to V_{GS} as follows.

$$g_m = \frac{dI_D}{dV_{GS}} = \mu C_{ox}\frac{W}{L}(V_{GS}-V_T) \tag{4.2}$$

It is often useful to express the transconductance in terms of the drain current. This can be done by combining equations (4.1) and (4.2). The result is given by the following formula.

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I_D} \tag{4.3}$$

The unity gain frequency using the long channel model is given by the following formula.

$$\omega_T = \frac{g_m}{C_{ox}WL} = \sqrt{\frac{2\mu}{L^2}\left(\frac{I_D}{C_{G1}}\right)} \tag{4.4}$$

4.1.2 Vertical Field Mobility Degradation

The IC industry is being driven toward the use of smaller transistors by the desire for higher speed circuits that dissipate less power. Small transistors with short channel lengths and thin gate oxides also allow increased levels of integration on a single integrated circuit (IC) and thus allows the production of more complex ICs. As the size of the transistor drops, the long channel model becomes less accurate in predicting the behavior of the transistor. Therefore, it is sometimes necessary to use more sophisticated models for the MOSFET.

The long channel model tends to overestimate the transconductance of a MOSFET. Two primary effects responsible for this inaccuracy are vertical field mobility degradation and velocity saturation. Vertical field mobility degradation is treated in this section, and velocity saturation is treated in the next section.

To model this problem, the MOSFET is assumed to be oriented such that the channel lies in a horizontal plain. When a voltage is applied to the gate of the MOSFET, a vertical electric field between the gate and the channel is set up. This electric field extends down into the semiconductor, and it is responsible for attracting charge carriers to the surface of the semiconductor and creating a conducting channel there. A side effect of this vertical electric field is a reduction in the mobility of the charge carriers. For circuit analysis, this effect is often conveniently modeled by the following equation[Tsividis87A][Hu88].

$$\mu_{eff} = \frac{\mu}{1+\theta(V_{GS}-V_T)} \tag{4.5}$$

The parameter θ in the above equation is called the vertical field degradation factor. A modified drain current can then be found by replacing the mobility in equation (4.1) with the effective mobility in equation (4.5). The result is shown below.

$$I_D = \frac{1}{2}\left(\frac{\mu C_{ox}}{1+\theta(V_{GS}-V_T)}\right)\frac{W}{L}(V_{GS}-V_T)^2 \tag{4.6}$$

From the above equation, it can be seen that at high gate voltages, the transconductance saturates. This result differs from the long channel model, which predicts that the transconductance is proportional to V_{GS}-V_T. The maximum value of the transconductance predicted by this model is given as follows.

$$g_{mlimit} = \frac{\mu C_{ox} \frac{W}{L}}{2\theta} \tag{4.7}$$

4.1.3 Velocity Saturation

When a voltage is applied between the drain and source of a MOSFET, a horizontal electric field is set up that moves charge carriers from the source toward the drain. As the electric field is increased, the average velocity (drift velocity) of the charge carriers increases. For low electric fields, the drift velocity v_d is proportional to the electric field as shown below where μ_{eff} is the mobility.

$$v_d = \mu_{eff} E \tag{4.8}$$

As the horizontal electric field is increased, scattering mechanisms responsible for slowing the charge carriers become more significant, and eventually the velocity of the charge carrier reaches a maximum called the saturation velocity v_{max}. For electrons in silicon, this saturation velocity is about 10^5 m/s. In order to account for velocity saturation, the drift velocity is often modeled by the following formula[Tsividis87B][Hu88].

$$v_d = \frac{\mu_{eff} E}{1 + \frac{E}{E_{sateff}}} \tag{4.9}$$

Thus, the maximum velocity v_{max} is $\mu_{eff} E_{sateff}$. E_{sateff} is the saturation electric field. The maximum velocity is independent of the gate to source voltage V_{GS}. Since μ_{eff} is a function of V_{GS}, E_{sateff} must also be a function of V_{GS} as shown below.

$$v_{max} = \mu E_{sat} = \mu_{eff} E_{sateff} \tag{4.10}$$

$$E_{sateff} = \frac{\mu}{\mu_{eff}} E_{sat} = E_{sat}(1 + \theta(V_{GS} - V_T)) \tag{4.11}$$

In the above equations, E_{sat} is the value of saturation electric field when V_{GS}=0.

A formula for the saturation drain current is now derived for the case where velocity saturation is present. The analysis refers to the MOSFET shown in figure 4.2.

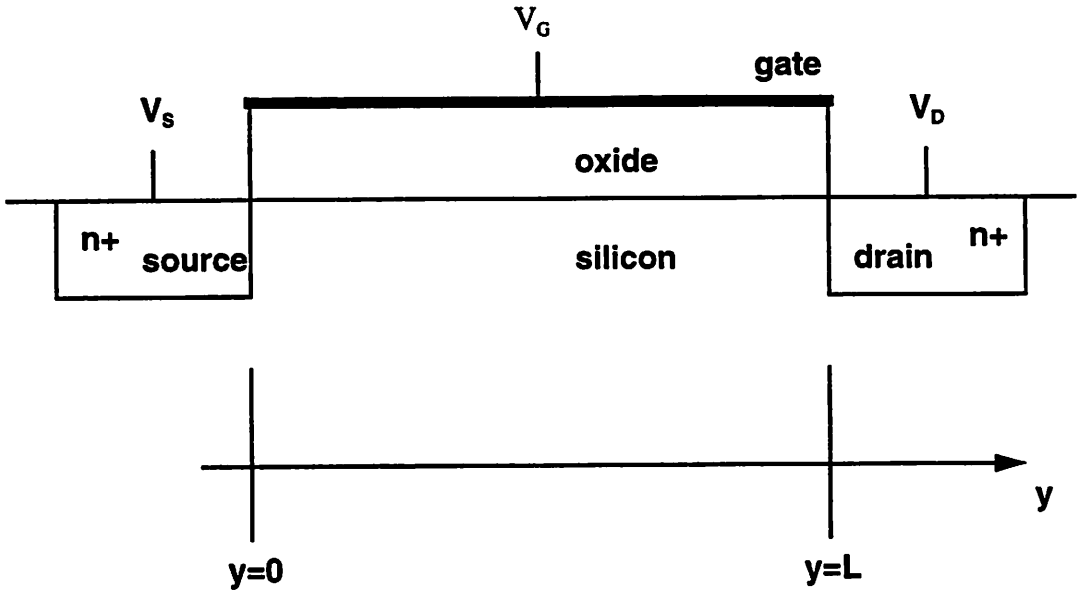


Figure 4.2 Cross Section of a MOSFET Showing the Source, Drain, and Channel

The voltage at position y along the channel is represented by the variable $V_C(y)$. The charge per unit area in the transistor channel is then given by the following formula.

$$Q_n(y) = C_{ox}(V_{GS} - V_T - V_C) \tag{4.12}$$

The current I_D is equal to the product of the charge per unit area, the transistor width, and the drift velocity as shown below.

$$I_D = Q_n W v_d \tag{4.13}$$

Now combine equations (4.9), (4.12), and (4.13) to get the result shown below.

$$I_D = \frac{C_{ox}(V_{GS} - V_T - V_C) W \mu_{eff} E}{1 + \frac{E}{E_{sat}}} \tag{4.14}$$

The electric field here is defined to be the derivative of the channel voltage with respect to the channel position y as shown below. Thus, the sign of the electric field as defined here is the opposite of the sign of the electric field as defined in most electromagnetic books. This change was made here simply to reduce the number of minus signs in the equations.

$$E = \frac{dV_C}{dy} \tag{4.15}$$

Now combine equations (4.14) and (4.15) to obtain the following differential equation.

$$E = \frac{dV_C}{dy} = \frac{I_D}{\mu C_{ox} W (V_{GS} - V_T - V_C) - \frac{I_D}{E_{sat}}} \tag{4.16}$$

Now integrate equation (4.16) to obtain the following relationship between drain current, channel position, and channel voltage. The fact that $V_C(y=0)=V_S$ is also used. The source voltage V_S is assumed here to be zero.

$$I_D = \frac{C_{ox} W \mu_{eff} (V_{GS} - V_T - 0.5V_C) V_C}{y + \frac{V_C}{E_{sat}}} \tag{4.17}$$

At y=L, the channel voltage V_C is equal to the drain voltage V_D. This fact leads to the following formula for the drain current.

$$I_D = \frac{C_{ox} W \mu_{eff} (V_{GS} - V_T - \frac{V_D}{2}) V_D}{L + \frac{V_D}{E_{sat}}} \tag{4.18}$$

To obtain the saturation current, it is somewhat arbitrarily assumed that the drift velocity at the drain is equal to the half the maximum velocity when saturation is reached. Thus, the saturation current is given by the following formula.

$$I_{Dsat} = Q_n W \frac{1}{2}\mu_{eff} E_{sateff} = \frac{1}{2}\mu_{eff} E_{sateff} W C_{ox} (V_{GS} - V_T - V_{Dsat}) \tag{4.19}$$

By equating the above two formulas for the drain current, the saturation drain voltage may be determined as shown below.

$$V_{Dsat} = \frac{(V_{GS} - V_T) E_{sateff} L}{V_{GS} - V_T + E_{sateff} L} \tag{4.20}$$

When this result for the saturation drain voltage is substituted into equation (4.19), the following equation for the saturation drain current is obtained.

$$I_{Dsat} = \frac{\frac{1}{2}\mu_{eff} E_{sateff} C_{ox} W (V_{GS} - V_T)^2}{V_{GS} - V_T + E_{sateff} L} \tag{4.21}$$

Now equations (4.5) and (4.11) are inserted into equation (4.21) to obtain the following equation for I_{Dsat}.

$$I_{Dsat} = \frac{\frac{1}{2}\mu E_{sat} C_{ox} W (V_{GS} - V_T)^2}{E_{sat} L + (1 + \theta E_{sat} L)(V_{GS} - V_T)} \tag{4.22}$$

For an n channel MOSFET, E_{sat}=1.5V/μm is a typical number.

From the above equation for drain current, the transconductance is given by the following formula.

$$g_m = \frac{\frac{1}{2}\mu E_{sat} C_{ox} W (V_{GS} - V_T)[2E_{sat} L + (1 + \theta E_{sat} L)(V_{GS} - V_T)]}{[E_{sat} L + (1 + \theta E_{sat} L)(V_{GS} - V_T)]^2} \tag{4.23}$$

The unity gain angular frequency ω_T is then given by the following formula.

$$\omega_T = \frac{g_m}{C_{ox} W L} = \frac{\frac{\frac{1}{2}\mu E_{sat}}{L}(V_{GS} - V_T)[2E_{sat} L + (1 + \theta E_{sat} L)(V_{GS} - V_T)]}{[E_{sat} L + (1 + \theta E_{sat} L)(V_{GS} - V_T)]^2} \tag{4.24}$$

The maximum unity gain frequency (corresponding to $V_{GS} \to \infty$) is then given by the following formula.

$$f_{Tmax} = \frac{\omega_{Tmax}}{2\pi} = \frac{(\frac{1}{2\pi})\frac{\mu}{2\theta L^2}}{1 + \frac{1}{\theta E_{sat} L}} \tag{4.25}$$

4.1.4 Subthreshold

Sometimes, it is desirable to bias the MOSFET device at very low current densities (current per unit width) to reduce power dissipation. A device operating in such a way is said to be biased in the subthreshold region. A transistor with a low current density has the disadvantage of a low unity gain frequency. However, the ratio of transconductance to current is high. This fact makes the use of low current density attractive in cases where low power dissipation but not high speed is required.

In the subthreshold region, also called weak inversion, the amount of charge in the conducting channel is not linearly related to the gate voltage as in the strong inversion case described by equation (4.12). Instead, the charge and therefore the current is exponentially related to the gate voltage. A simple way to model the current in the subthreshold region is given by the following formula.[Tsividis87C]

$$I_D = I_o e^{\frac{qV_{GS}}{2k_B T}} \tag{4.26}$$

In the above formula, q represents the magnitude of charge on one electron (1.602×10^{-19}C), k_B represents Boltzmann's constant, and T represents temperature. Using this model for the current, the transconductance is given by the following formula.

$$g_m = \frac{qI_D}{2k_B T} \tag{4.27}$$

Note that in equation (4.27), the transconductance is proportional to the current. This behavior differs from the strong inversion behavior given by equation (4.3). In that case the transconductance is proportional to the square root of the current.

4.1.5 Putting the Models Together

Sometimes it is nice to have a model that is applicable to all the regions of operation of the MOSFET. While the complexity of this model makes it impractical for hand calculations, it is still reasonable for use with a computer.

The goal of the combined model is to calculate a ratio of transconductance to current (g_m/I) as a function of current density (I/W) that is valid for very small current densities and very large current densities. To accomplish this, equation (4.22) is first rearranged to find V_{GS}-V_T as a function of

current density in the strong inversion region. This model incorporates vertical field mobility degradation and velocity saturation, but not subthreshold. In the subthreshold region, V_{GS}-V_T is not particularly meaningful, so subthreshold will be incorporated later. The formula for V_{GS}-V_T is given as follows.

$$V_{GS} - V_T = \left(\frac{1 + \theta E_{sat} L}{\mu E_{sat} C_{ox}}\right)\frac{I_D}{W}\left\{1 + \sqrt{1 + \frac{2\mu E_{sat} C_{ox} E_{sat} L}{(1 + \theta E_{sat} L)^2 (\frac{I_D}{W})}}\right\} \tag{4.28}$$

Now, equation (4.22) is rearranged again as shown below.

$$\sqrt{\frac{2I_D}{\mu C_{ox}\frac{W}{L}}} = \frac{(V_{GS} - V_T)}{\sqrt{1 + (\frac{1}{E_{sat} L} + \theta (V_{GS} - V_T))}} \tag{4.29}$$

Now subthreshold is included by adding an extra term to equation (4.29) as shown below.

$$\frac{2k_B T}{q} \ln\left(\frac{I_D}{I_o}\right) + \sqrt{\frac{2I_D}{\mu C_{ox}\frac{W}{L}}} = \frac{(V_{GS} - V_T)}{\sqrt{1 + (\frac{1}{E_{sat} L} + \theta)\ (V_{GS} - V_T)}} \tag{4.30}$$

Now the transconductance is calculated by differentiating the above equation with respect to V_{GS}. The result for transconductance is shown below.

$$g_m = \frac{1 + \frac{1}{2}(\frac{1}{E_{sat} L} + \theta)\ (V_{GS} - V_T)}{\left[1 + (\frac{1}{E_{sat} L} + \theta)\ (V_{GS} - V_T)\right]^{3/2}\left(\frac{2k_B T}{q I_D} + \sqrt{\frac{1}{2\mu C_{ox}\frac{W}{L} I_D}}\right)} \tag{4.31}$$

In the above formula, V_{GS}-V_T is given by equation (4.28). Figure 4.3 shows a plot of g_m/I ver-

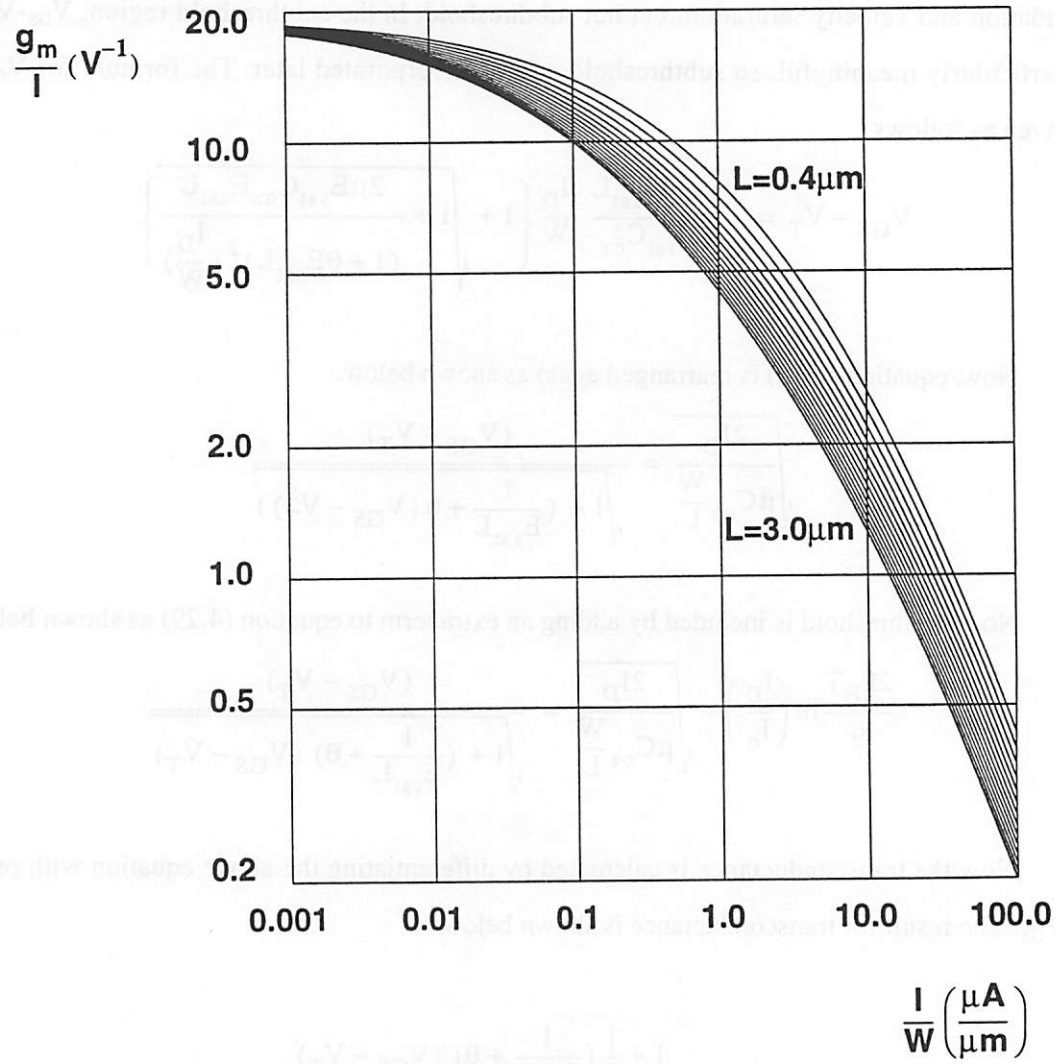


Figure 4.3 Ratio of Transconductance to Bias Current versus Current Density

sus current density, and figure 4.4 shows a plot of unity gain frequency versus current density. The

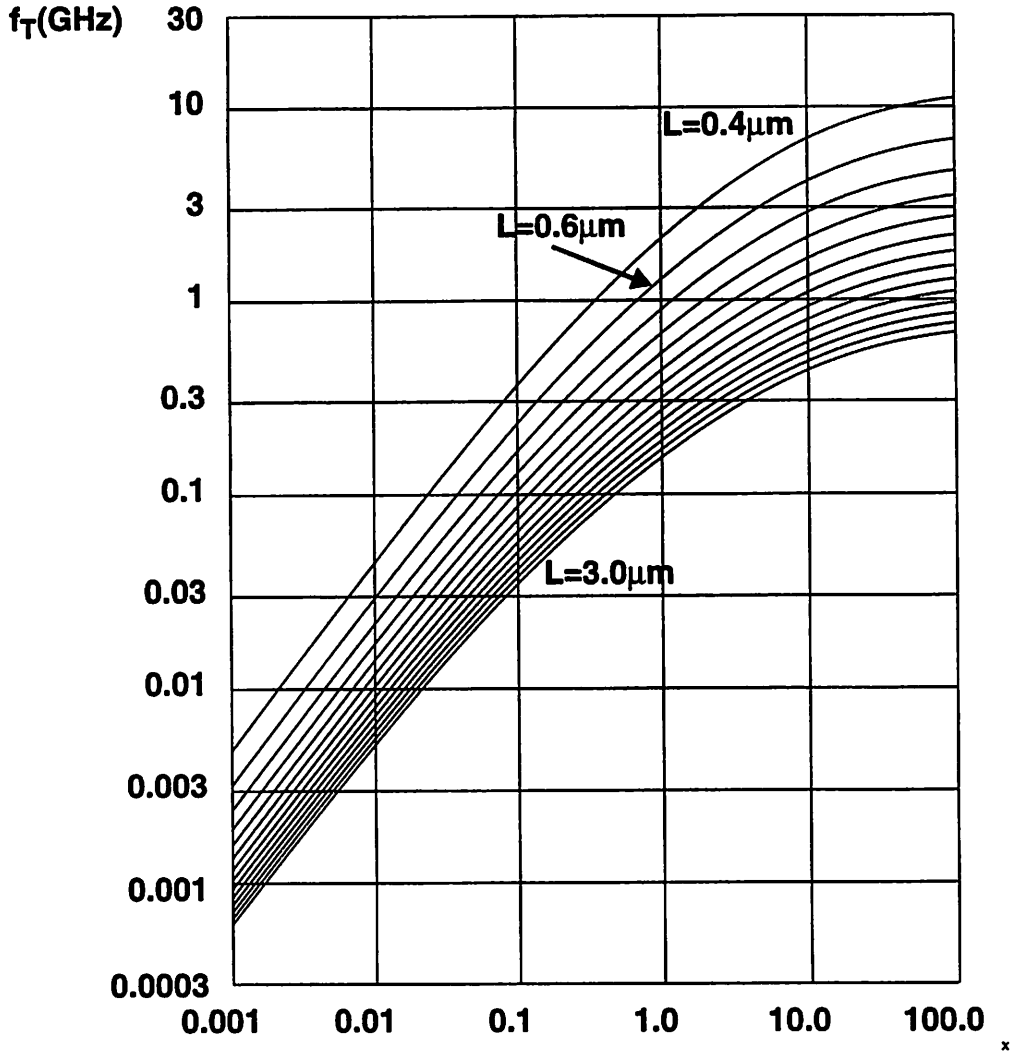


Figure 4.4 NMOS Unity Gain Frequency versus Current Density

following assumptions were made about the process parameters.

Table 4.1 Process Parameter Assumptions

Temperature (T)	300K
Gate Oxide Capacitance (C_{ox})	1.535 fF/μm²

Table 4.1 Process Parameter Assumptions

θ	$0.4092V^{-1}$
E_{sat}	1.5V/μm
Low Field Mobility (μ)	524 $cm^2/V/s$

4.2 Settling Time Analysis of Switched Capacitor Gain Stages

A transient analysis of switched capacitor gain stages will be presented here now that a few transistor models have been introduced. First, the simplest amplifier architectures will be analyzed, and then the analysis will proceed on to more complex architectures. Figure 4.5 shows the operation of a sample and hold circuit to be studied here. During the sampling phase, the input is sampled onto capacitors C_S and C_F. During the hold phase, an amplified version of the signal is presented at the output. The closed loop gain of the system is approximately $1+C_S/C_F$.

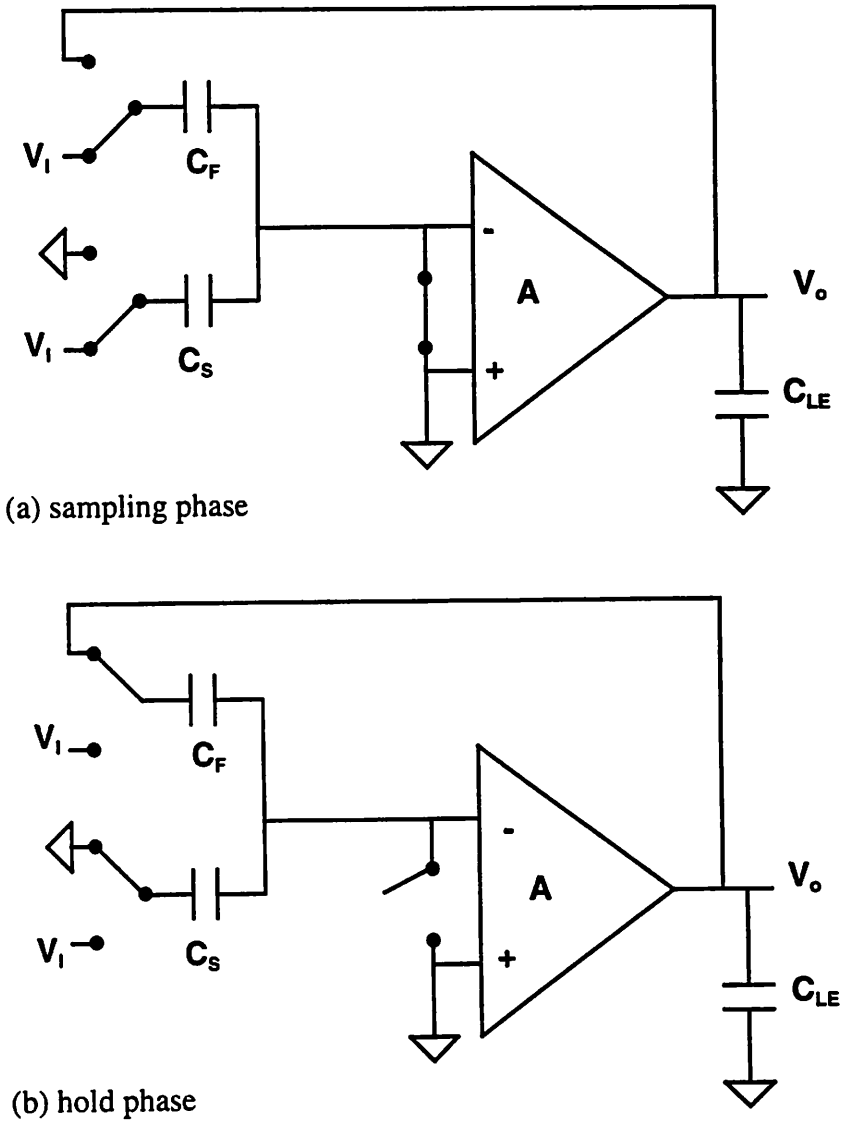


Figure 4.5 Sample/Hold Operation

Figure 4.6 shows a general circuit diagram modeling the sample and hold circuit during the hold phase. During this phase of operation, the sample and hold circuit is a feedback amplifier with a unit step applied to the input. The output response is therefore the step response of this circuit. This response requires a finite amount of time is to settle to a given accuracy, as discussed in chap-

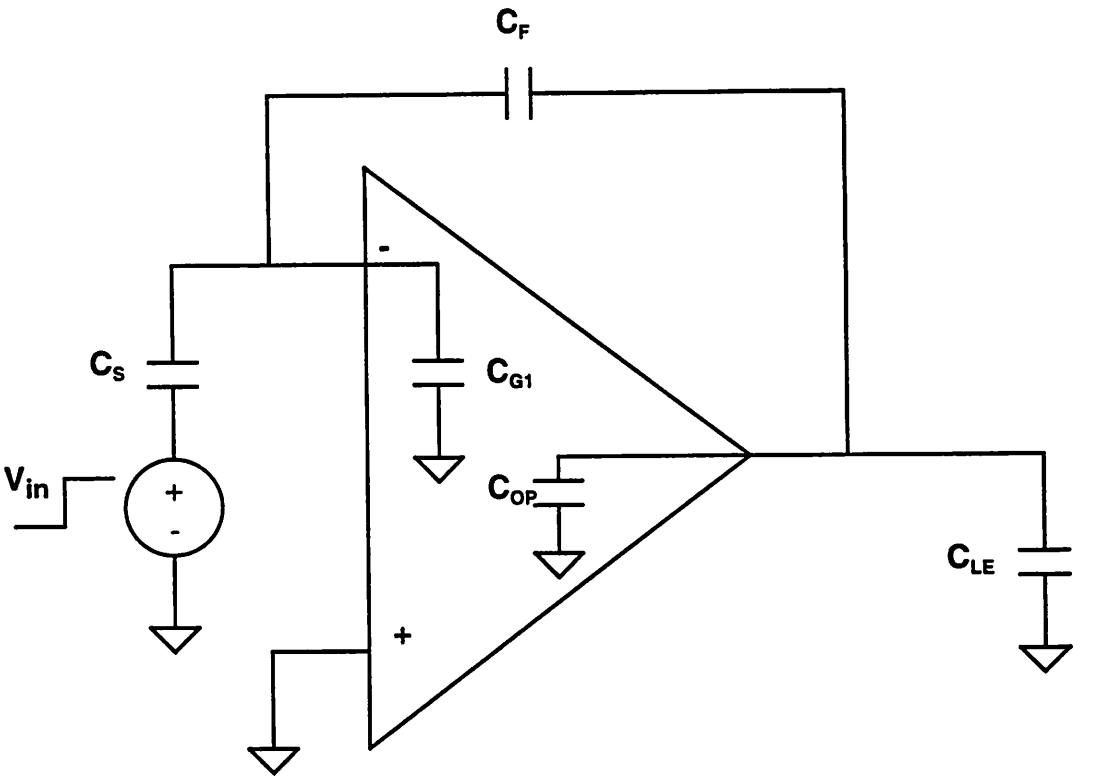


Figure 4.6 General Model for the Switched Capacitor Sample and Hold Amplifier During Hold Phase

ter 3. The length of time required for this settling during the hold phase of operation is the focus of this chapter.

Figure 4.6 applies to all of the amplifier architectures discussed here. The amplifier is represented by the triangular region. Five capacitors are shown. The function of C_S and C_F is described above. C_{LE} represents the load capacitance connected to the output of the amplifier that is not related to the feedback network or the output parasitic capacitance of the amplifier. C_{OP} is the output parasitic capacitance of the amplifier. C_{G1} is the input capacitance of the amplifier. In reality both the sampling capacitor and feedback capacitors have parasitic capacitances to ground associated with them. These parasitics are not considered here. However, they could be modeled without

changing the analysis by lumping these extra parasitic capacitances in with C_S. The circuit shown here is single ended for simplicity, but the single ended analysis also applies to differential circuits.

In reality, the switches shown in figure 4.5 are resistive, and this switch resistance adds poles to the response characteristic during the hold phase of operation. However, this resistance is not included in the model shown in figure 4.6 and is not included in the analysis of this chapter.

The ratio of sampling capacitance C_S to feedback capacitance C_F is represented in this discussion by the letter G.

$$G = \frac{C_S}{C_F} \tag{4.32}$$

The above ratio has a strong influence on the feedback factor. The feedback factor is defined as the ratio of the input voltage v_i to the output voltage v_o in the closed loop circuit.

$$f = \frac{v_o}{v_i} = \frac{1}{1 + \frac{C_S}{C_F} + \frac{C_{G1}}{C_F}} = \frac{1}{1 + G + \frac{C_{G1}}{C_F}} \tag{4.33}$$

A number of amplifier architectures for switched capacitor circuits are analyzed in this chapter. At the end of this chapter, the efficiencies with which each of these architectures achieves a settling time for a given value of G are compared. In this comparison G=1 was chosen, because if the feedback capacitor function is shared with the sampling capacitor function, a switched capacitor circuit with a closed loop gain of 2 results. The switched capacitor amplifier with a closed loop gain of 2 is a very useful circuit in many pipelined ADC designs because it often results in a very simple pipeline.

4.2.1 Single Stage Single Pole Amplifier

The single stage amplifier having 1 pole is the simplest architecture imaginable. Because of its simplicity, it can be extensively analyzed and optimized by hand analysis. The single pole amplifier is also of a great deal of interest because of its inherent stability. The simple single pole architecture makes this amplifier the most attractive for high speed applications with low closed loop gain.

Figure 4.7 shows a simplified model of the sample and hold circuit using a single stage amplifier. In a switched capacitor application the input voltage V_{in} is a unit step function.

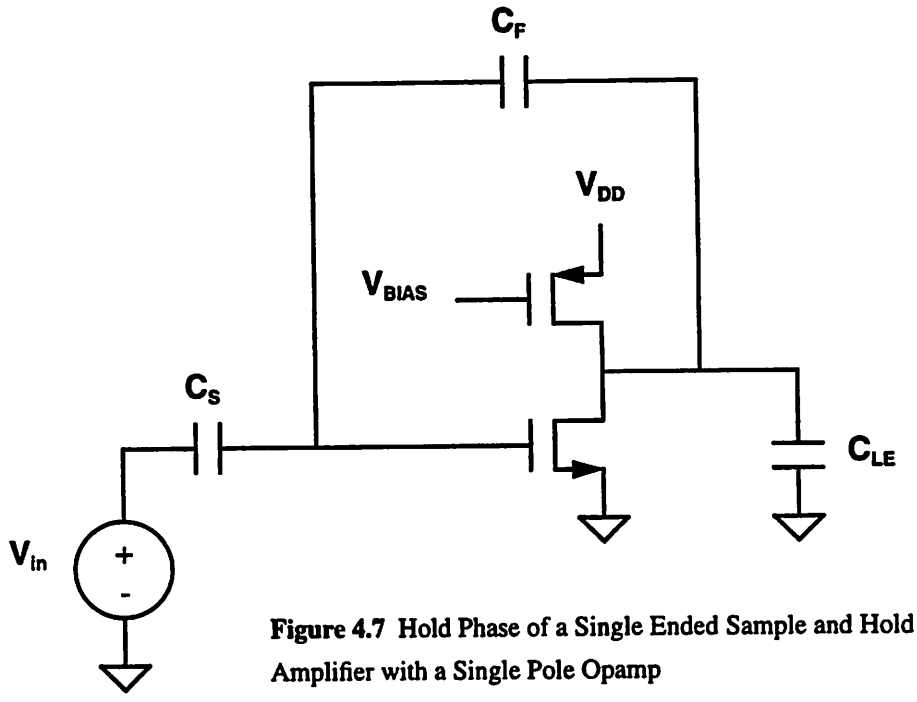


Figure 4.7 Hold Phase of a Single Ended Sample and Hold Amplifier with a Single Pole Opamp

Figure 4.8 shows a circuit diagram used to model the small signal transient behavior of the above circuit. Note that the output parasitic capacitance C_{OP} is assumed to be proportional to the input capacitance. The parameter α is used to relate the output parasitic capacitance to the input capacitance. In this model, the output resistances of the transistors are assumed to be infinite. Finite output resistance in actual transistor circuits has the effect of moving the poles somewhat to the left of what is predicted here. The effect is small if the product $g_m r_o$ is high. The output parasitic capacitance of the amplifier is assumed to be proportional to the input capacitance as shown in figure 4.8.

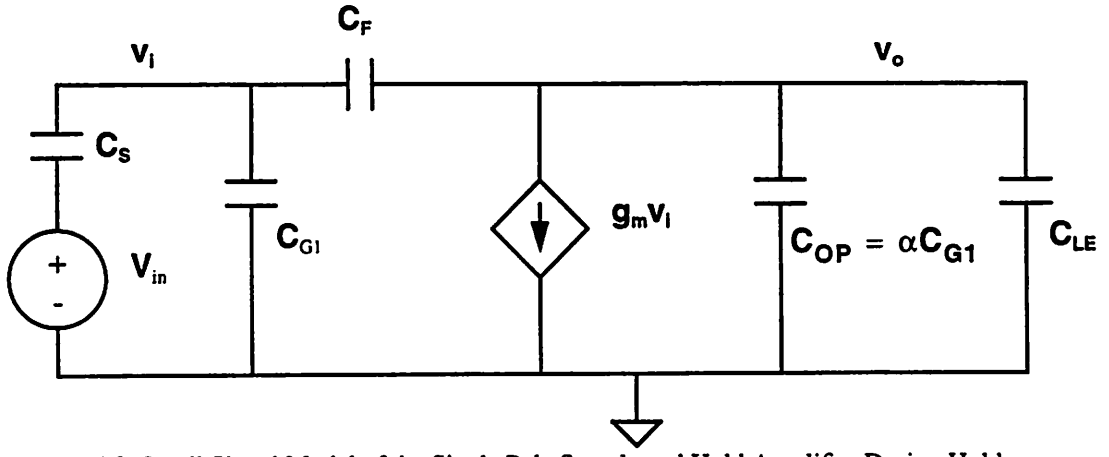


Figure 4.8 Small Signal Model of the Single Pole Sample and Hold Amplifier During Hold Phase

This circuit will now be analyzed and optimized using a number of different sets of assumptions in order to suit different applications.

4.2.1.1 Fixed Current Density

In this analysis, the current density is assumed to be fixed. This assumption implies that the unity gain frequency of the transistor in the critical path is also fixed. This assumption also implies that power dissipation is proportional to the input capacitance C_{G1}. Another feature of this assumption is that it leads to an analysis that is technology independent. The analysis is general enough to apply to bipolar transistors as well as MOSFETs. For more information, this type of analysis is treated carefully by Conroy[Conroy94].

To begin this analysis, the transconductance g_m is related to the input capacitance C_{G1} by the unity gain frequency ω_T as shown below.

$$g_m = \omega_T C_{G1} \tag{4.34}$$

In the analysis here, the ratio G of sampling capacitance to feedback capacitance is assumed to be constant. A nodal analysis of figure 4.8 yields the following formula for the time constant τ.

$$\tau = \frac{C_{LE} + C_F(1-f) + \alpha C_{G1}}{f g_m} \tag{4.35}$$

Now equations (4.33), (4.34), and (4.35) are combined to eliminate f and g_m. The result is shown below.

$$\tau = \frac{1}{\omega_T C_{G1}}\left[\left(1 + G + \frac{C_{G1}}{C_F}\right)(C_{LE} + C_F + \alpha C_{G1}) - C_F\right] \tag{4.36}$$

The above equation can be solved for the input capacitance of the amplifier. The result is shown below.

$$B = [\omega_T \tau - \alpha(1+G) - 1]\frac{C_F}{C_{LE}} - 1 \tag{4.37}$$

$$C_{G1} = \frac{B}{2\alpha}\left\{1 - \sqrt{1 - \frac{4\alpha}{B^2}\frac{C_F}{C_{LE}}\left(1 + G + G\frac{C_F}{C_{LE}}\right)}\right\} \tag{4.38}$$

A graph of input capacitance versus feedback capacitance is shown in figure 4.9. From the graph, it is seen that for every speed an optimum sized feedback capacitance exists to minimize the input capacitance. This is significant because minimizing input capacitance minimizes power dissipation. The value of the optimum feedback capacitance is discussed in the appendix.

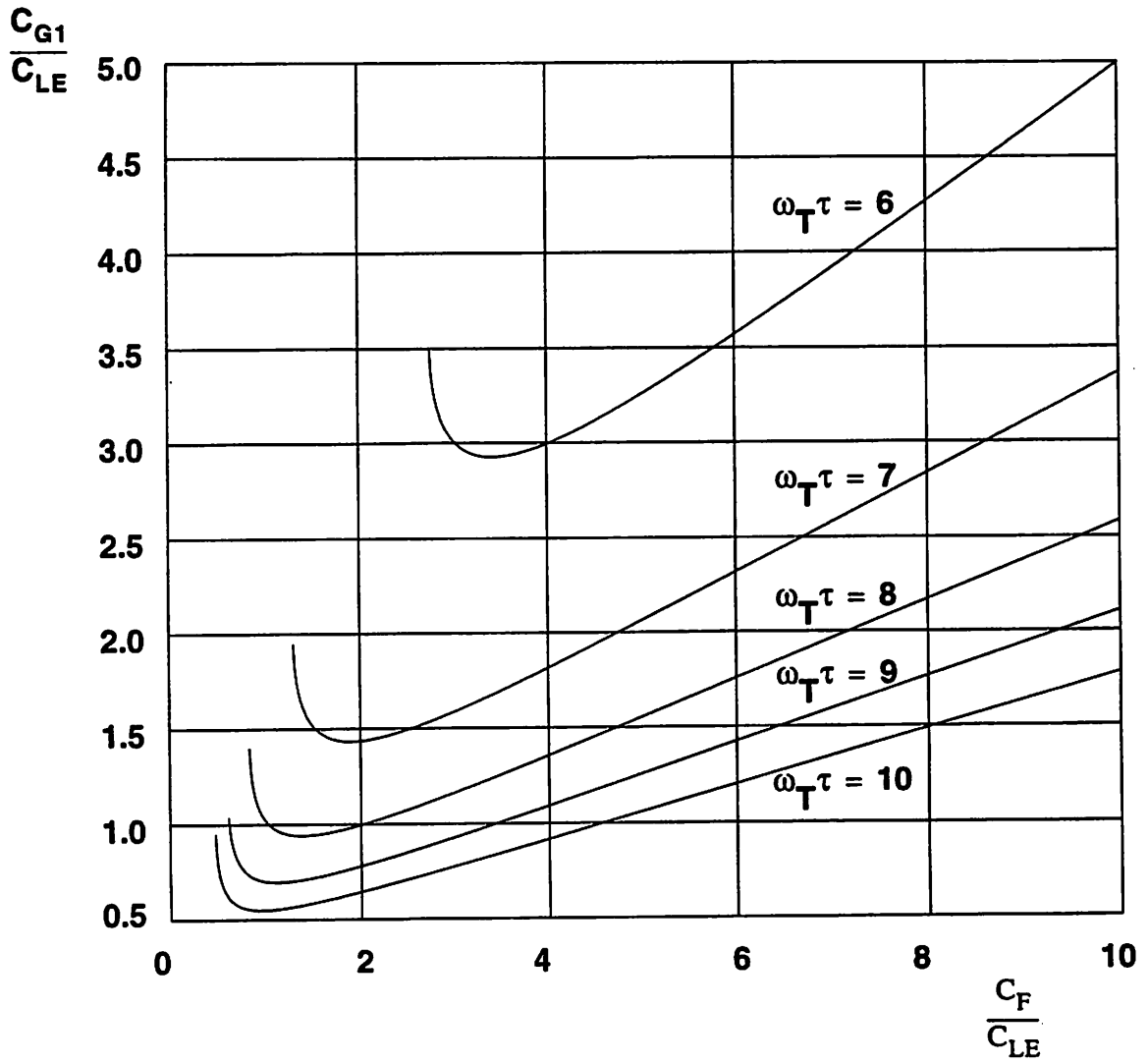


Figure 4.9 Input Capacitance versus Feedback Capacitance for a Single Pole Amplifier, α=1, G=1

4.2.1.1.1 Minimum Power with Fixed Feedback Capacitance

The analysis discussed in the above section was performed for a range of current densities with velocity saturation, mobility degradation, and subthreshold effects considered. The resulting input capacitance and current are plotted in figures 4.10 and 4.11. In the graphs, L=1.2μm. Also, the process parameters of table 4.1 were used, and the settling time indicated is the time to settle to 10 bits of accuracy.

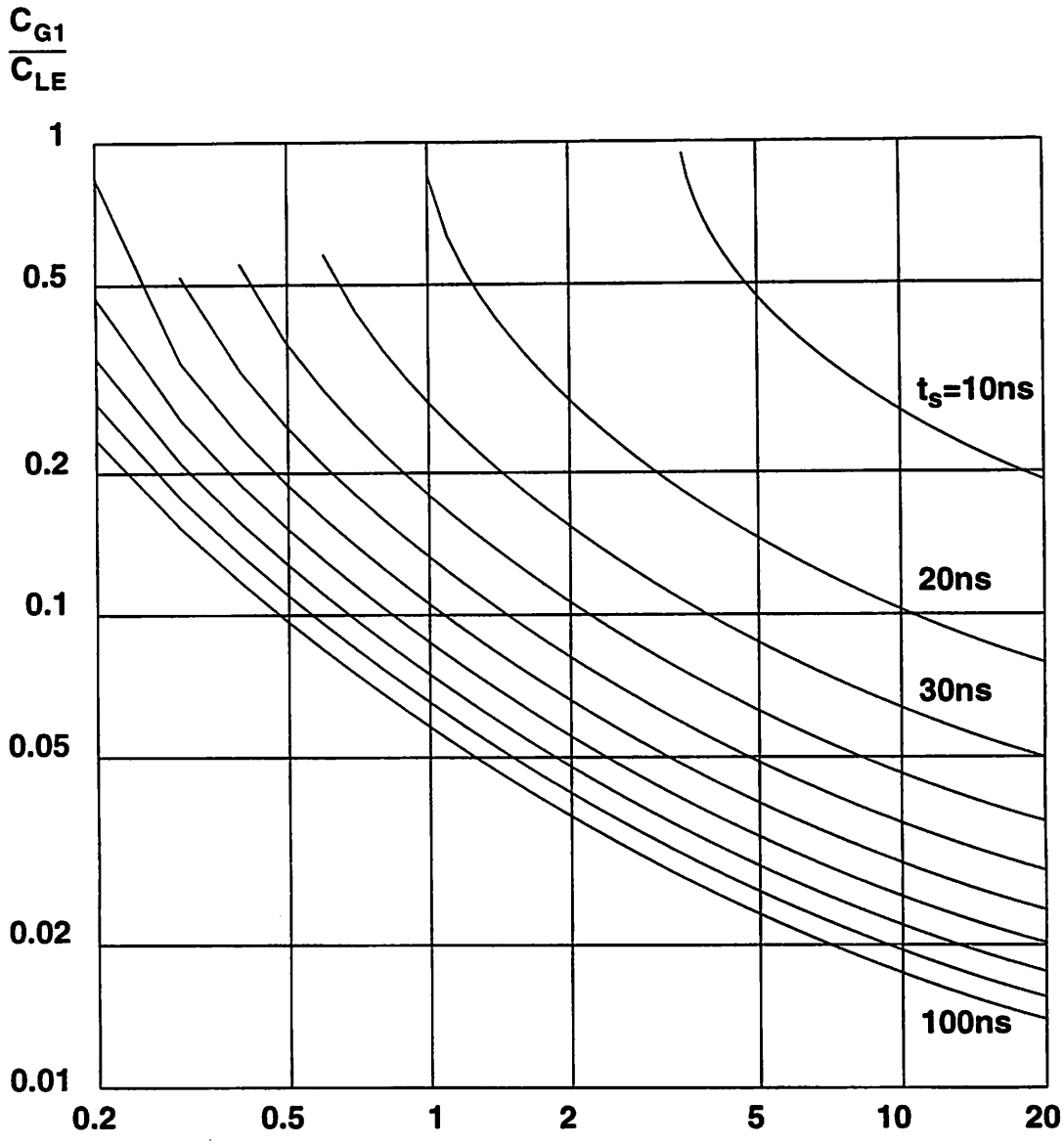


Figure 4.10 Input Capacitance Versus Current Density for a Single Stage Amplifier, α=1, G=1

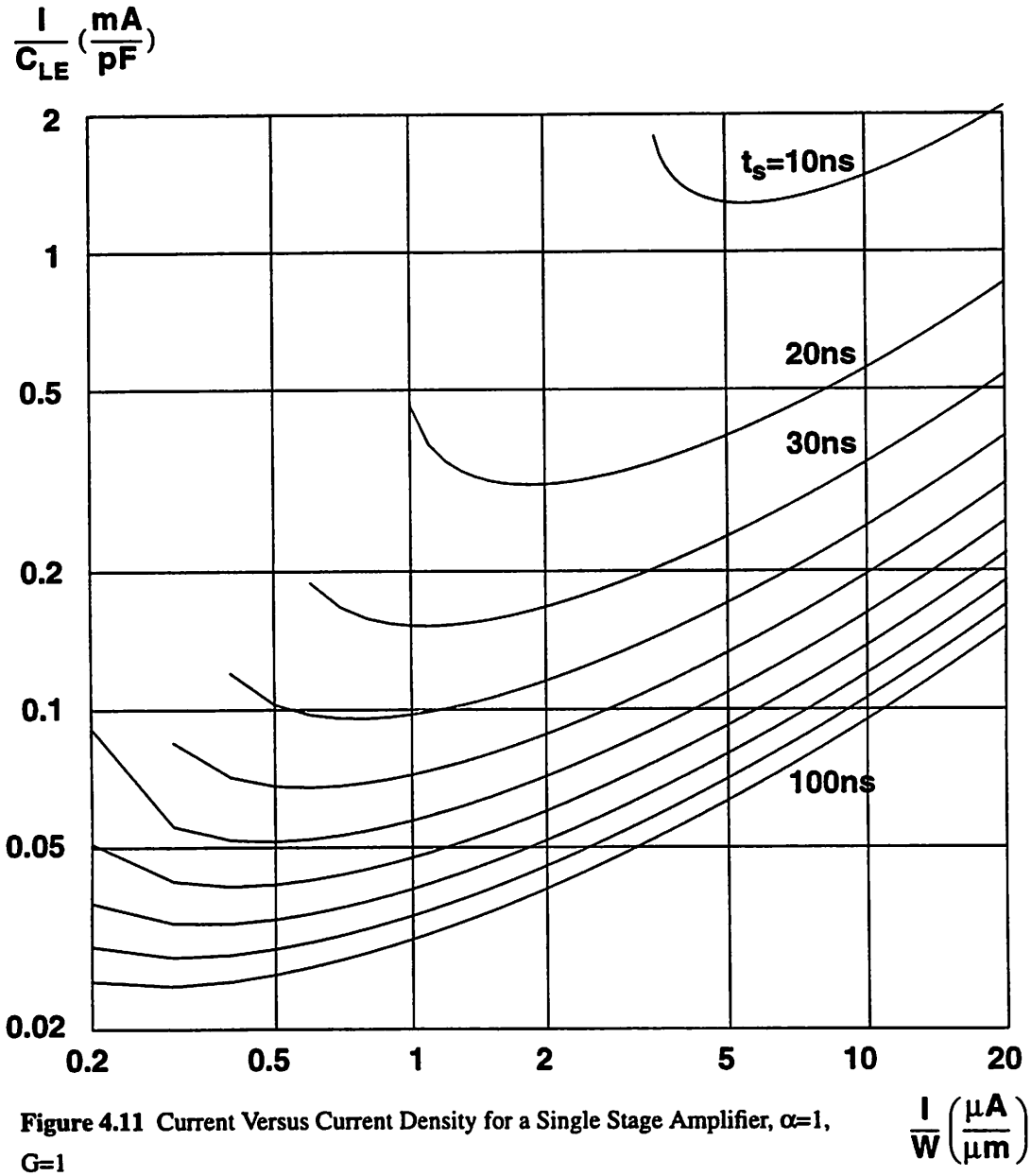


Figure 4.11 Current Versus Current Density for a Single Stage Amplifier, α=1, G=1

It can be seen from the graphs that, as expected, settling time reduction requires an increase in power dissipation. It is also evident from the graphs that the optimum current density to minimize power dissipation increases as the required settling time is reduced. This makes sense because increasing the current density in a transistor tends to increase the unity gain frequency of the transistor.

4.2.2 Telescopic Cascode Amplifier

The gain of the single transistor amplifier described in section 4.2.1 is limited. For a large number of applications, this gain is inadequate. The gain of the single stage amplifier can be increased without significantly degrading the bandwidth by using the cascode architecture shown in figure 4.12. This architecture has the added advantage that the effect of the parasitic feedback capacitance between the drain and gate of the input transistor is reduced.

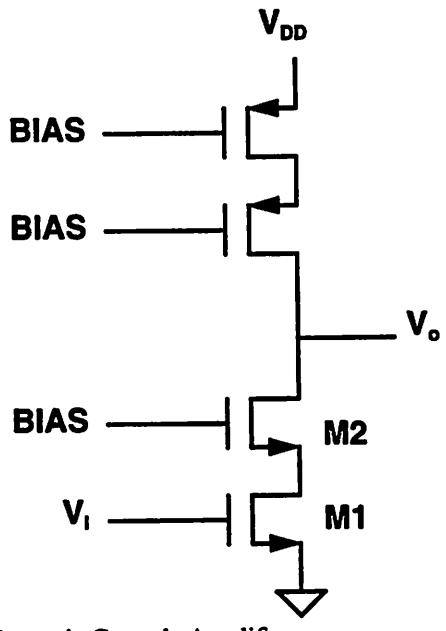


Figure 4.12 Single Ended Telescopic Cascode Amplifier

This circuit is often modeled as a single pole amplifier because the second pole due to the cascode transistor can often be at a much higher frequency than the dominant pole. In that case the previously described analysis for the single pole amplifier can be used. In this section, the effect of the nondominant pole is considered and the implications on design are evaluated. The circuit is modeled as shown in figure 4.13.

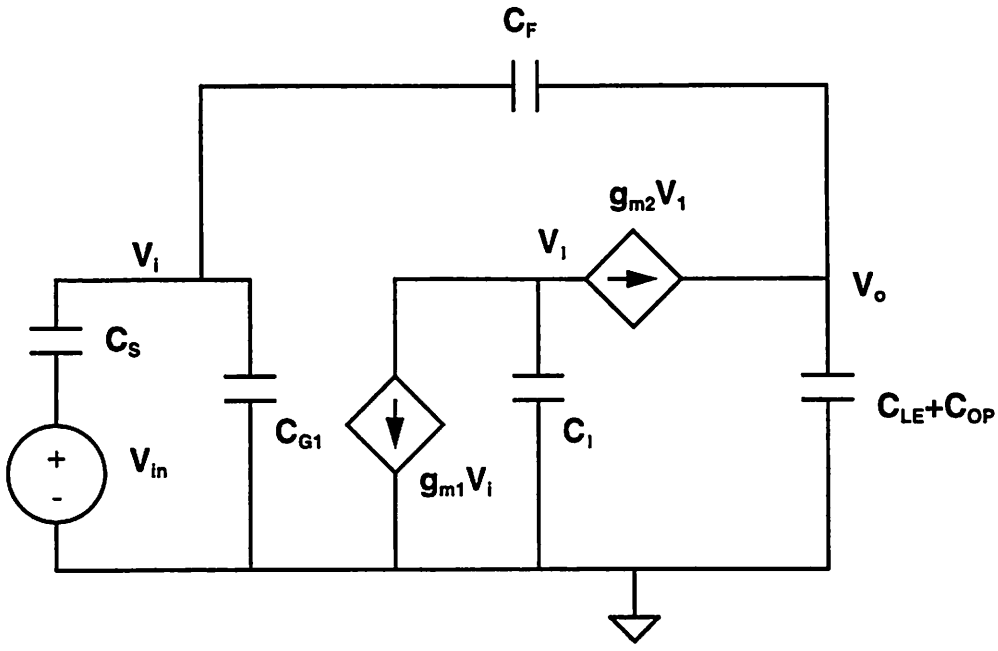


Figure 4.13 Small Signal Model for a Telescopic Cascode Amplifier

If the feedforward path through the feedback capacitor is neglected, then the simplified circuit model shown in figure 4.14 can be used. In this model, the loading effect due to the feedback network is included in the capacitor C_2.

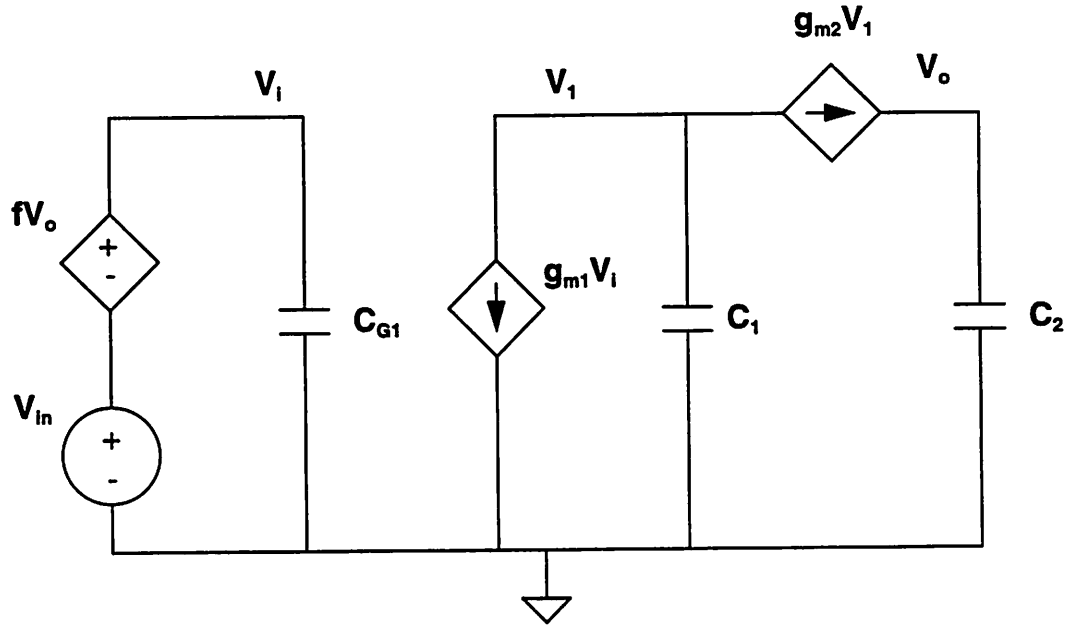


Figure 4.14 Simplified Small Signal Model for the Telescopic Cascode Amplifier (Feedforward Path through C_F Neglected

In the above figure, the following symbols are used:

g_{m1} the transconductance of the first stage of the amplifier

g_{m2} the transconductance of the Cascode device

C_1 the capacitance loading node V_1

C_2 the capacitance loading the output of the amplifier

A nodal analysis of the above circuit in the s domain yields the following formula for the closed loop gain A_{CL}.

$$A_{CL}(s) = \frac{V_O(s)}{V_I(s)} = \frac{N_{CL}(s)}{D_{CL}(s)} = \frac{\frac{g_{m2}g_{m1}}{C_1C_2}}{s^2 + \frac{g_{m2}}{C_1}s + \frac{fg_{m1}g_{m2}}{C_1C_2}} \tag{4.39}$$

In the above expression $N_{CL}(s)$ represents the numerator, which in general is a polynomial in s. The denominator $D_{CL}(s)$ in the above expression is the characteristic polynomial for the system. If it is written in the form shown below, the settling behavior of the system can be related to the component values in the circuit.

$$D_{CL}(s) = s^2 + \frac{g_{m2}s}{C_1} + \frac{fg_{m1}g_{m2}}{C_1C_2} = s^2 + 2\zeta\omega_n s + \omega_n^2 \tag{4.40}$$

The above parameters are related to the positions of the poles of the closed loop system as shown in figure 4.15 assuming complex conjugate poles. ζ is the damping factor of the system, ω_{CL} is the real part of the pole position, and τ is the time constant associated with the decay of the envelope of the response. The figure assumes complex poles, but if $\zeta>1$, both poles will be real.

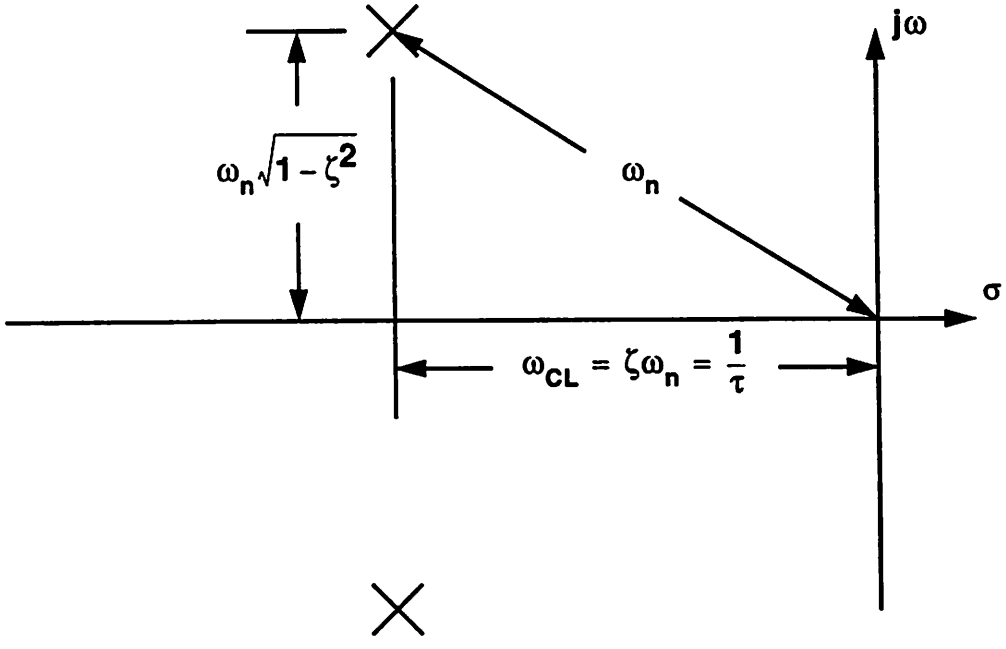


Figure 4.15 s Plane Pole Positions for a System with a Characteristic Polynomial Given by Equation (4.55)

It is now worthwhile to relate the parameters τ and ζ, which define the transient response of the closed loop amplifier, to the size of the load, the sizes of the input and cascode transistors, and to the size of the feedback capacitor C_F. From equation (4.55), the following two relations for the time constant and the damping factor can be obtained.

$$\tau = \frac{2C_1}{g_{m2}} \tag{4.41}$$

$$\tau^2 = \frac{C_1 C_2}{\zeta^2 f g_{m1} g_{m2}} \tag{4.42}$$

The above two equations will now be used as a starting point for the design and optimization of the sample and hold circuit. The various conductances g can be related to input capacitances by the unity gain frequencies of the various devices as shown below.

$$g_{m1} = \omega_{T1} C_{G1} \tag{4.43}$$

$$g_{m2} = \omega_{T2} C_{G2} \tag{4.44}$$

In the above equations, C_{G1} is the gate capacitance of M_1, and C_{G2} is the gate capacitance of M_2. A conceptually useful new variable is now introduced. The ratio r between C_{G2} and C_{G1} is defined as shown below.

$$r = \frac{C_{G2}}{C_{G1}} \tag{4.45}$$

Now the capacitor C_1 includes the drain capacitance of M_1, the drain capacitance of M_2, and the gate capacitance of M_2. Thus, C_1 may be expressed as follows.

$$C_1 = \alpha_1 C_{G1} + (1+\alpha_2) C_{G2} = [(1+\alpha_2) r + \alpha_1] C_{G1} \tag{4.46}$$

In the above expression, the variable α_1 is the ratio of the drain capacitance of M_1 to the gate capacitance of M_1, and α_2 is the ratio of the drain capacitance of M_2 to the gate capacitance of M_2. In a similar manner, the capacitance C_2 loading the second stage is the sum of the capacitance due to the external load capacitance C_{LE}, the output capacitance of the amplifier, and the capacitive loading of the feedback network. Therefore, C_2 may be expressed as follows.

$$C_2 = C_{LE} + C_F(1-f) + \alpha_2 C_{G2} = C_{LE} + C_F(1-f) + \alpha_3 r C_{G1} \tag{4.47}$$

In the above equation, α_3 represents the ratio of the drain capacitance at the output to the gate capacitance of M_2.

4.2.2.1 Optimization of Current Density to Minimize Power - Fixed Speed and Fixed Feedback Capacitance - Model Including Mobility Degradation, Velocity Saturation, and Subthreshold

In this section, the amplifier optimization is performed using the model discussed in section 4.1.5. The channel length is assumed to be 1.2μm. Furthermore, the damping factor ζ is allowed to vary. In order to allow comparisons between designs with different damping factors, the optimization was based on minimizing power dissipation for a fixed settling time. The settling time was assumed to be given by the following equations depending on the damping factor and the desired settling error ε. It is assumed here that the amplifier has a two pole response as discussed in chapter 3.

strongly underdamped case: $\zeta<\zeta_T$

In this case, the response function oscillates rapidly. Therefore, small changes in circuit or device parameters can produce rapid changes in the settling time. It is usually undesireable to produce a design that is sensitive to small changes. Therefore, in this case the settling time is assumed to be approximately equal to the time required for the envelope to settle. The settling time of the envelope of the response is much less sensitive to small changes than the actual response.

$$t_s = \frac{-1}{\zeta\omega_n \ln(\varepsilon\sqrt{1-\zeta^2})} \tag{4.48}$$

threshold damping factor: ζ_T

$$\frac{\zeta_T}{\sqrt{1-\zeta_T^2}} = 1 + \frac{1}{\ln\left[\frac{1}{\varepsilon\sqrt{1-\zeta_T^2}}\right]} \tag{4.49}$$

weakly underdamped case: $\zeta_T<\zeta<1$

In this case, the envelope of the response is not a good predictor of the actual settling time. Therefore, the settling time is approximated as shown below.

$$t_s \approx \frac{1}{\zeta\omega_n \ln\left[\frac{1}{\varepsilon}\left(\frac{1}{\zeta} + \omega_n t_s\right)\right]} \tag{4.50}$$

overdamped case: $\zeta > 1$

In this case a coefficient of overdamping c is introduced. c can have a value between 0 and 1. Zero corresponds to critical damping, and 1 corresponds to a single pole response.

$$c = \sqrt{1 - \frac{1}{\zeta^2}} \tag{4.51}$$

$$\omega_1 = \zeta\omega_n (1 - c) \tag{4.52}$$

$$t_s = \ln\left\{\left(\frac{1+c}{2c\varepsilon}\right)\left[1 - \left(\frac{1-c}{1+c}\right) e^{-2\omega_1 t_s \left(\frac{c}{1-c}\right)}\right]\right\} \tag{4.53}$$

Because of the complex nature of the algebra, the optimization is not carried out analytically. Instead, a computer was used, and the results are graphed here in figures 4.16, 4.17, 4.18, and 4.19.

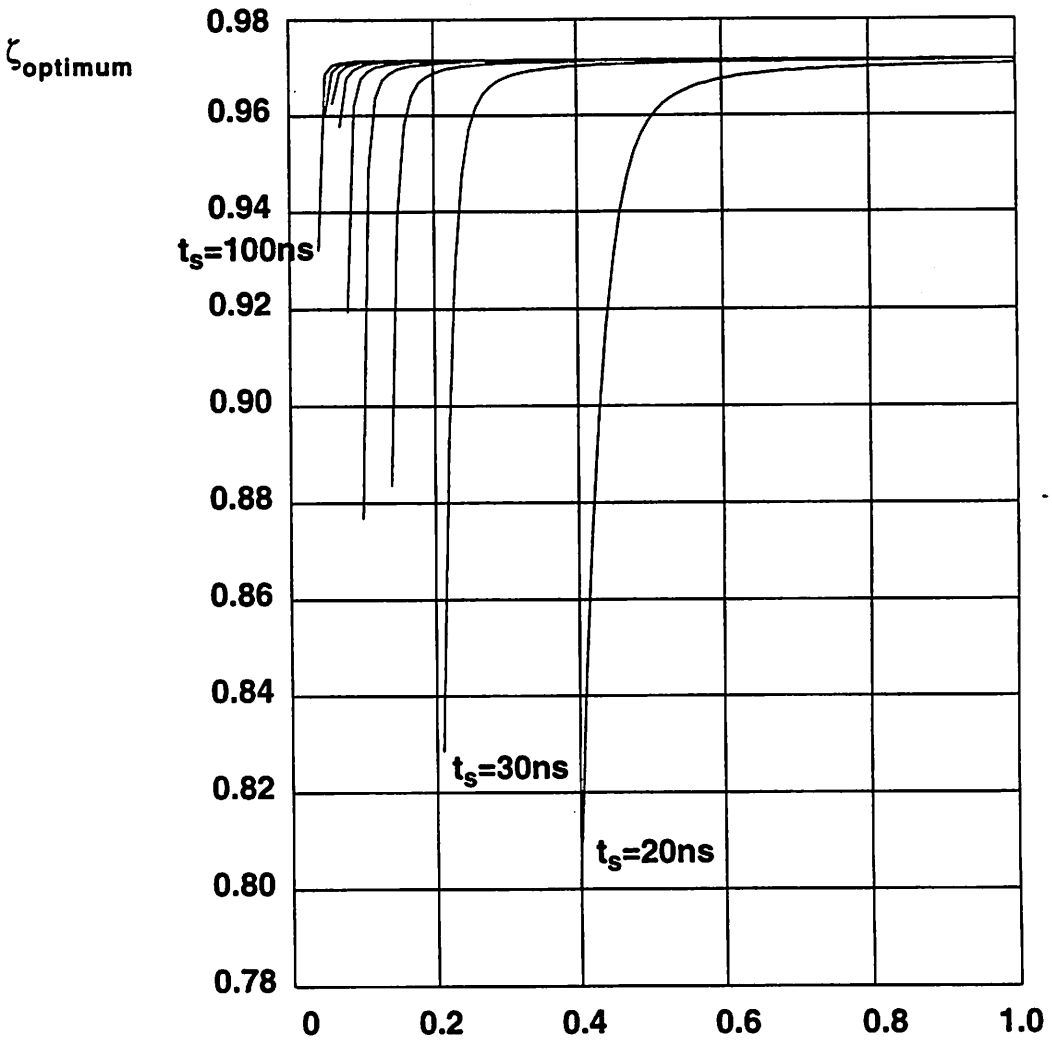


Figure 4.16 Damping Factor versus Current Density for Optimized Telescopic Cascode Opamp Design, $\alpha_1=\alpha_2=1$, G=1

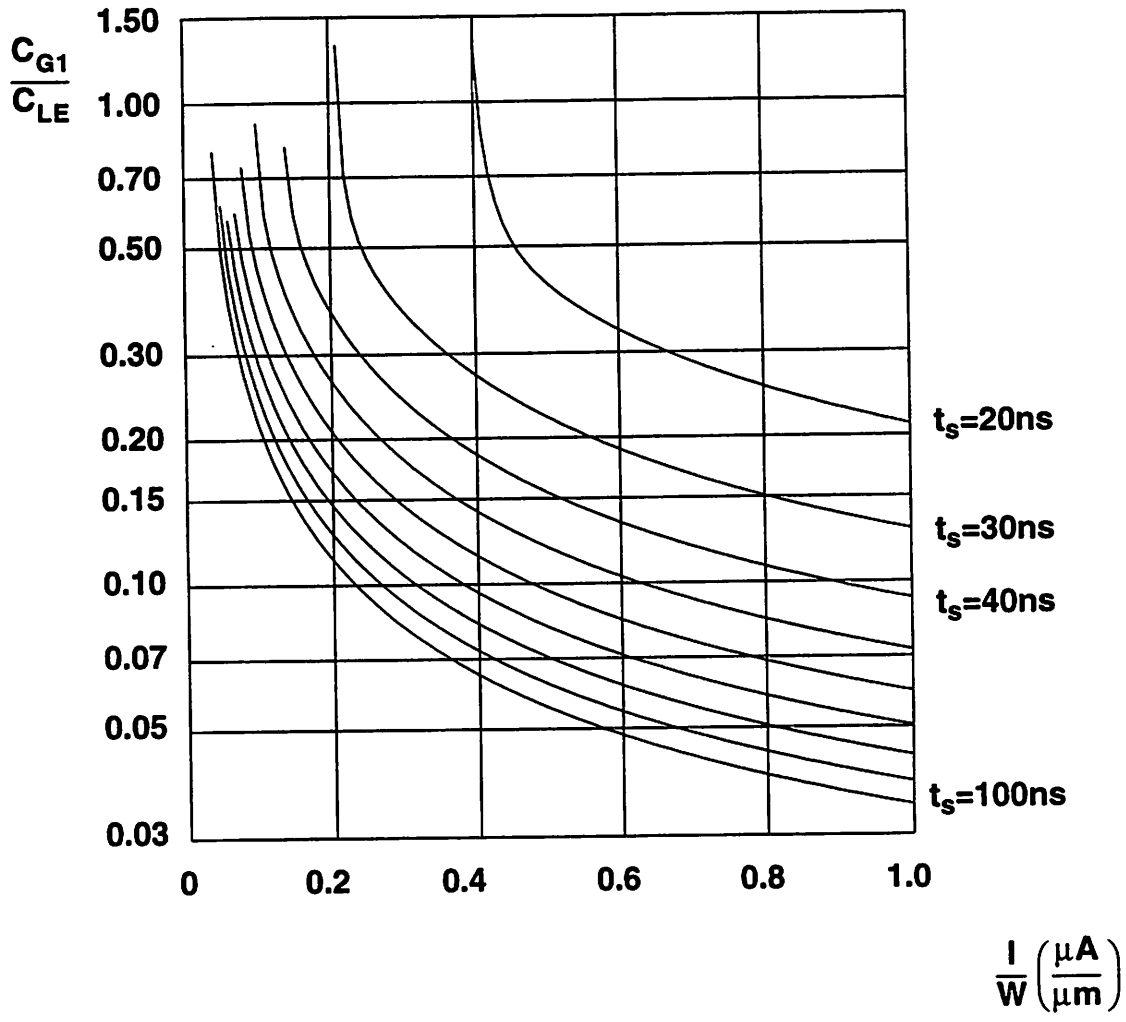


Figure 4.17 Input Capacitance versus Current Density for Optimized Telescopic Cascode Opamp Design, $\alpha_1=\alpha_2=1$, G=1

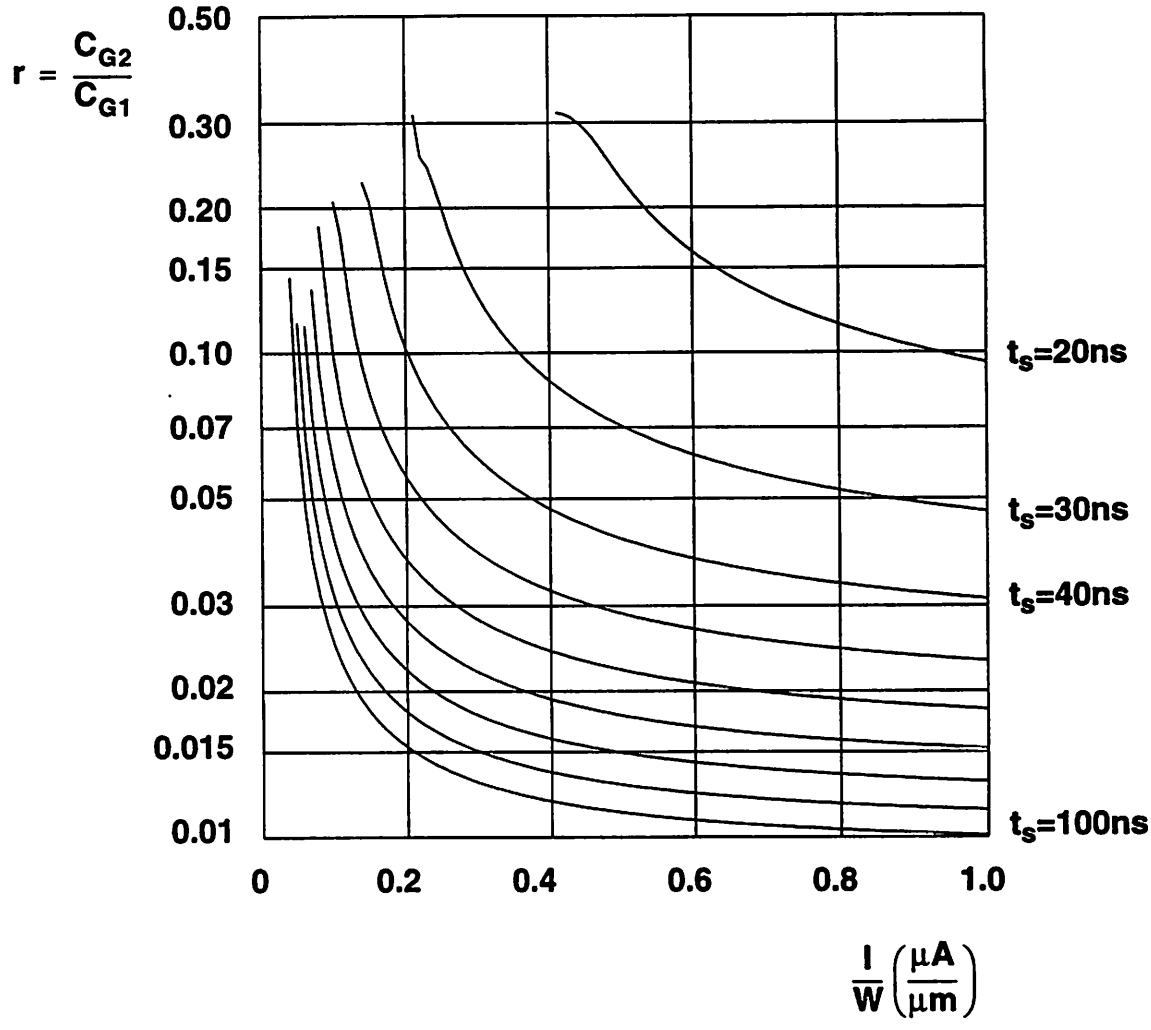


Figure 4.18 Ratio of Cascode Transistor Size to Input Transistor Size versus Current Density for Optimized Telescopic Cascode Opamp Design, $\alpha_1=\alpha_2=1$, G=1

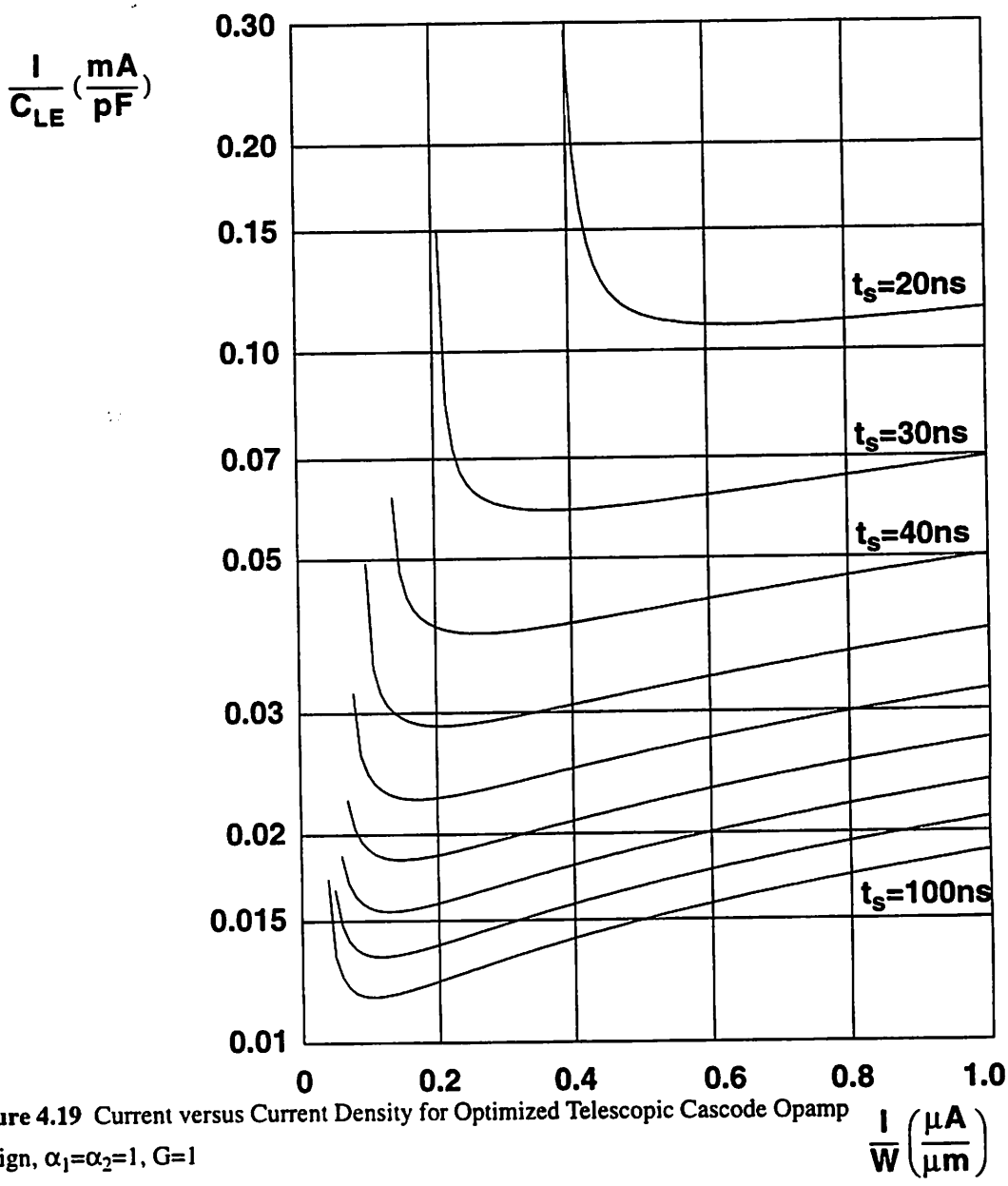


Figure 4.19 Current versus Current Density for Optimized Telescopic Cascode Opamp Design, $\alpha_1=\alpha_2=1$, G=1

A comparison of figure 4.11 to figure 4.19 leads to an interesting conclusion. As expected, a simple single stage amplifier is capable of higher speed than the telescopic cascode. However, for a settling time of 20ns or more the telescopic cascode dissipates less power than a simple single stage opamp. This is somewhat surprising since the telescopic architecture introduces a second pole to the response of the amplifier. The telescopic cascode amplifier can save power over a single stage amplifier because the output parasitic capacitance of the amplifier can be reduced relative to

the input parasitic capacitance by making the cascode transistor smaller than the input transistor. As shown by figure 4.18, using a relatively small cascode transistor to reduce output capacitance is often optimal, particularly in cases where the load capacitance is smaller than or comparable to the sampling capacitance.

4.2.3 Wide-Band Preamplifier Driving a Single Stage Amplifier

In some applications where the supply voltage is low or the gain required is very high, a single stage amplifier is not adequate. Therefore, it is worthwhile to study multistage topologies. In this section, a simple two stage amplifier with no compensation capacitors is considered[Cho95A][Cho95B]. An s domain analysis is performed. The results of this analysis are then used to optimize the trade-off between speed and power. Because the nondominant pole frequency of the amplifier is inversely proportional to the output resistance of the first stage, it is generally desirable to keep the first stage output resistance low. As a result, this type of amplifier is typically realized by making the first stage a low gain, low output impedance preamplifier. The second stage is typically a high output impedance amplifier. An example of a single ended version of an amplifier with this topology is shown in figure 4.20.

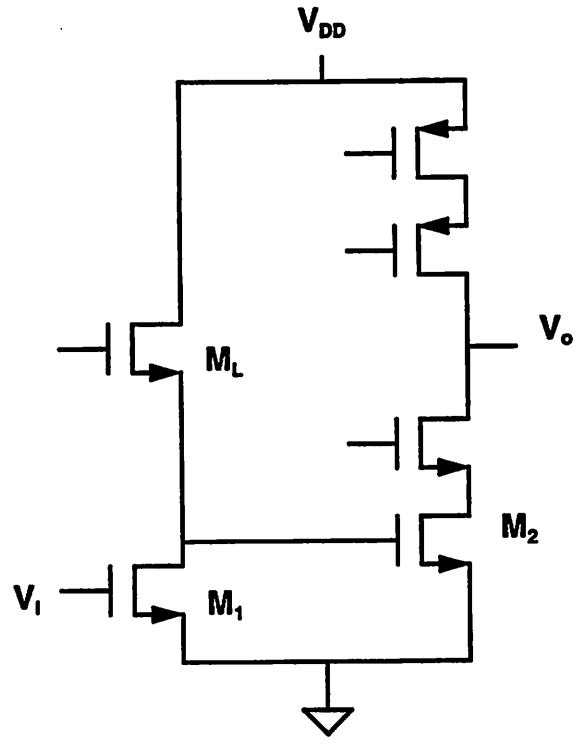


Figure 4.20 Example of a Single Ended Implementation of a Preamplifier Driving a Single Cascode Stage

Figure 4.21 shows a small signal model for the single stage amplifier driven by a wide-band preamplifier. The output resistance of the second stage is assumed to be infinite. The circuit is further simplified in figure 4.22. In this model, the forward path through the feedback capacitor is ignored. The loading effect of the feedback network is lumped with C_2. The feedback provided by the feedback capacitor is represented by the voltage source fV_o.

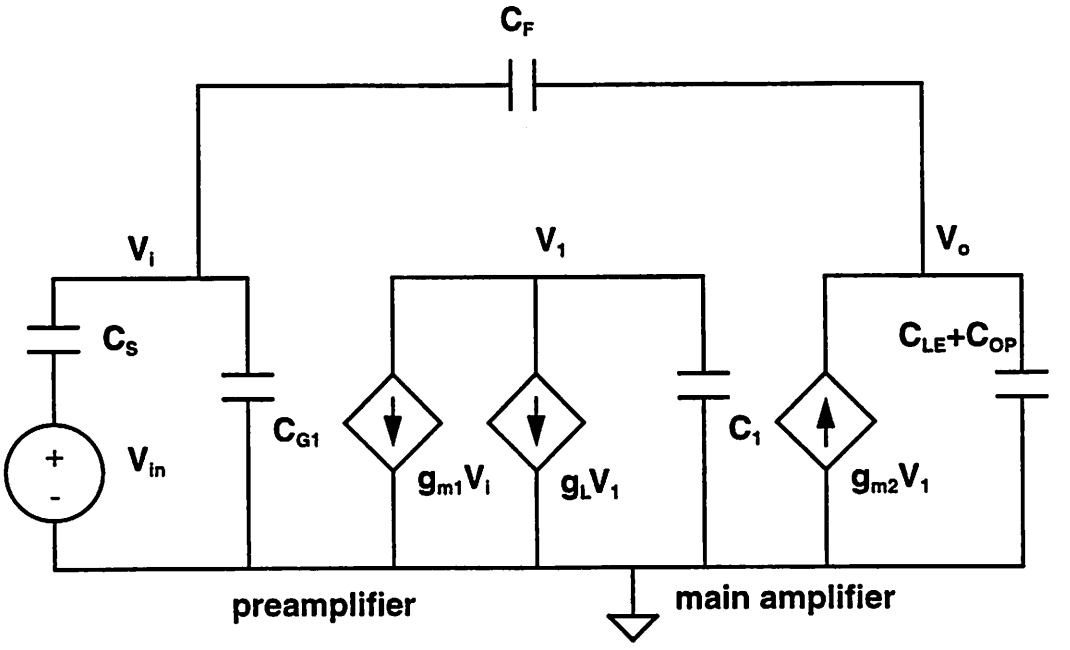


Figure 4.21 Small Signal Circuit Model for a Preamplifier Driving a Single Stage Opamp

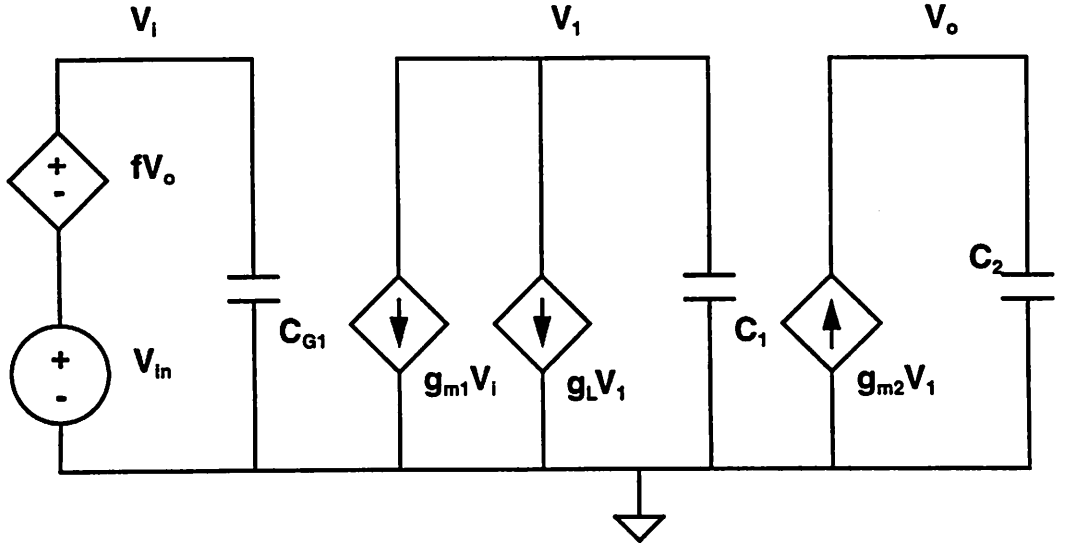


Figure 4.22 Simplified Small Signal Circuit Model for a Preamplifier Driving a Single Stage Opamp (Feedforward Path Neglected)

In the above figure, the following symbols are used:

g_{m1} the transconductance of the first stage of the amplifier

g_L the conductance of the load to the first stage of the amplifier

g_{m2} the transconductance of the second stage of the amplifier

C_1 the capacitance loading the first stage of the amplifier

C_2 the capacitance loading the second stage of the amplifier

A nodal analysis of the above circuit in the s domain yields the following formula for the closed loop gain A_{CL}.

$$A_{CL}(s) = \frac{V_O(s)}{V_I(s)} = \frac{N_{CL}(s)}{D_{CL}(s)} = \frac{\frac{g_{m2}g_{m1}}{C_1C_2}}{s^2 + \frac{g_L}{C_1}s + \frac{fg_{m1}g_{m2}}{C_1C_2}} \tag{4.54}$$

In the above expression $N_{CL}(s)$ represents the numerator, which in general is a polynomial in s. The denominator $D_{CL}(s)$ in the above expression is the characteristic polynomial for the system. If it is written in the form shown below, the settling behavior of the system can be related to the component values in the circuit.

$$D_{CL}(s) = s^2 + \frac{g_L s}{C_1} + \frac{fg_{m1}g_{m2}}{C_1C_2} = s^2 + 2\zeta\omega_n s + \omega_n^2 \tag{4.55}$$

The above parameters are related to the positions of the poles of the closed loop system as shown in figure 4.23 assuming complex conjugate poles. ζ is the damping factor of the system, ω_{CL} is the real part of the pole position, and τ is the time constant associated with the decay of the envelope of the response.

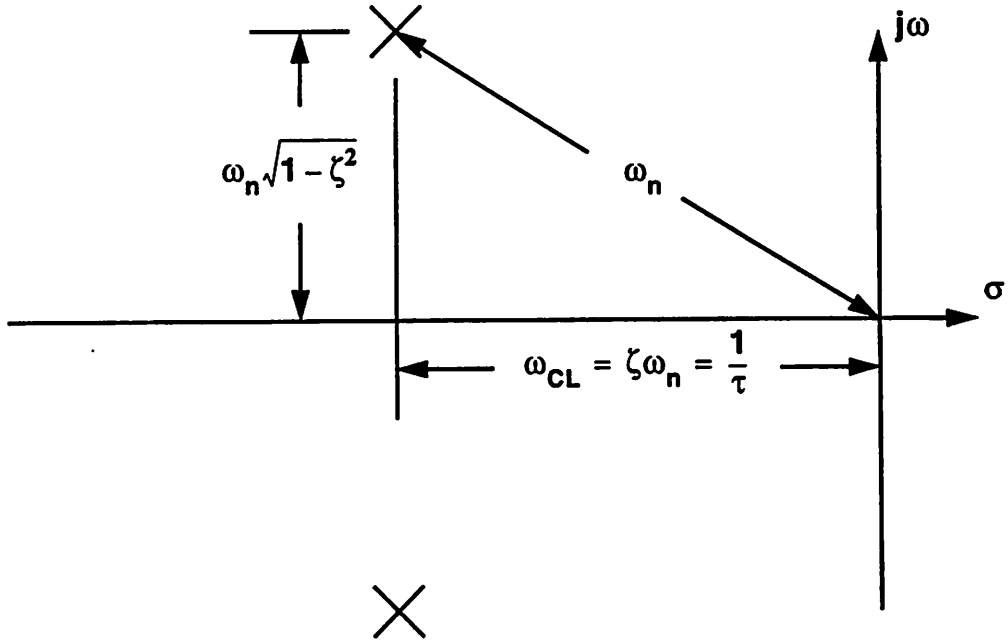


Figure 4.23 s-Plane Pole Positions for an Underdamped Two Pole System

It is now worthwhile to relate the parameters τ and ζ, which define the transient response of the closed loop amplifier, to the size of the load, the sizes of the two amplifier stages, and to the size of the feedback capacitor C_F. From equation (4.55), the following two relations for the time constant and the damping factor may be obtained.

$$\tau = \frac{2C_1}{g_L} \tag{4.56}$$

$$\tau^2 = \frac{C_1 C_2}{\zeta^2 f g_{m1} g_{m2}} \tag{4.57}$$

The above two equations will now be used as a starting point for the design and optimization of the sample and hold circuit. The various conductances g can be related to input capacitances by the unity gain frequencies of the various devices as shown below.

$$g_{m1} = \omega_{T1} C_{G1} \tag{4.58}$$

$$g_{m2} = \omega_{T2} C_{G2} \tag{4.59}$$

$$g_L = \omega_{TL} C_{GL} \tag{4.60}$$

In the above equations, C_{G1} is the gate capacitance of M_1, C_{G2} is the gate capacitance of M_2, and C_{GL} is the gate capacitance of M_L. Some conceptually useful new variables are now introduced. The gain A_{PRE} of the preamplifier is given by the following ratio.

$$A_{PRE} = \frac{g_{m1}}{g_L} \tag{4.61}$$

The ratio r_c between the gate capacitance C_{GL} of M_L and the gate capacitance C_{G1} of M_1 is defined as follows.

$$r_c = \frac{C_{GL}}{C_{G1}} \tag{4.62}$$

The ratio x between the gate capacitance C_{G2} of M_2 and the gate capacitance C_{G1} of M_1 is defined as follows.

$$x = \frac{C_{G2}}{C_{G1}} \tag{4.63}$$

Now the capacitor C_1 is the capacitance loading the preamplifier. It includes the drain capacitance of M_1, the gate and drain capacitance of M_L, and the gate capacitance of M_2. Thus, C_1 may be expressed as follows.

$$C_1 = \alpha_1 C_{G1} + (1+\alpha_L) C_{GL} + C_{G2} = [(1+\alpha_L) r_c + \alpha_1 + x] C_{G1} \tag{4.64}$$

In the above expression, the variable α_1 is the ratio of the drain capacitance of M_1 to the gate capacitance of M_1, and α_L is the ratio of the drain capacitance of M_L to the gate capacitance of M_L. In a similar manner, the capacitance C_2 loading the second stage is the sum of the capacitance due to the external load capacitance C_{LE}, the output capacitance of the amplifier, and the capacitive loading of the feedback network. Therefore, C_2 may be expressed as follows.

$$C_2 = C_{LE} + C_F(1-f) + \alpha_2 C_{G2} = C_{LE} + C_F(1-f) + \alpha_2 x C_{G1} \tag{4.65}$$

In the above equation, α_2 represents the ratio of the drain capacitance to the gate capacitance of M_2.

4.2.3.1 Optimization to Minimize the Power

The goal of the procedure discussed in this section is to choose the preamplifier gain, preamplifier input capacitance, and second stage input capacitance to meet a settling time specification with minimum power. The analysis starts by substituting equations (4.60) and (4.64) into (4.56) for the time constant. The result is shown below.

$$\tau = \frac{2\left[\alpha_1 + x + r_c\left(1+\alpha_L\right)\right] C_{G1}}{\omega_{TL} C_{GL}} \tag{4.66}$$

Now equation [4.62] is used to eliminate C_{GL} to obtain the following equation.

$$\tau = \frac{2\left[\alpha_1 + x + r_c\left(1+\alpha_L\right)\right]}{\omega_{TL} r_c} \tag{4.67}$$

The above equation is then solved simultaneously with equation (4.57) and optimized using a computer. The model including velocity saturation, mobility degradation, and subthreshold effects discussed earlier in this chapter is used. The results are plotted in figures 4.24, 4.25, 4.26, 4.27, and 4.28.

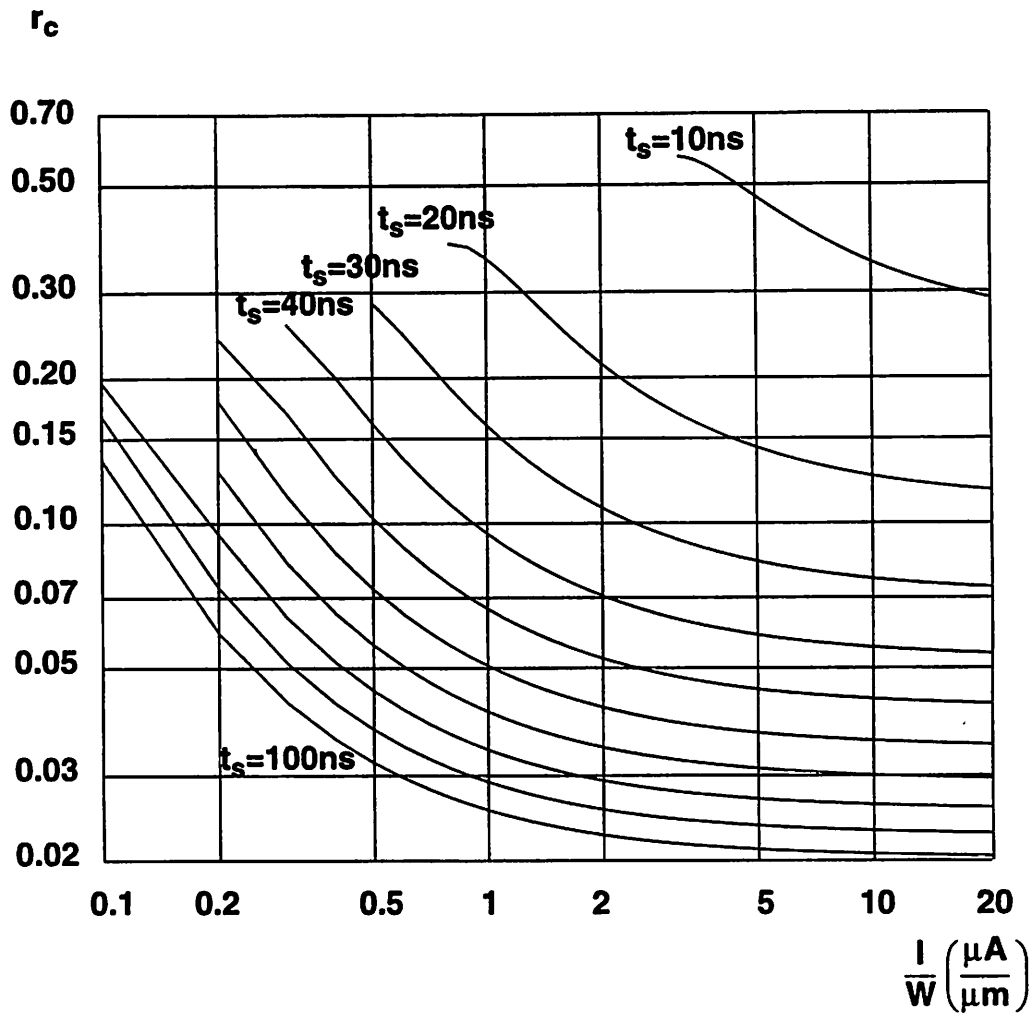


Figure 4.24 Ratio of Preamp Load Transistor Size to Input Transistor Size versus Current Density, $\alpha_1=\alpha_2=\alpha_L=1$, $\zeta=1$, G=1

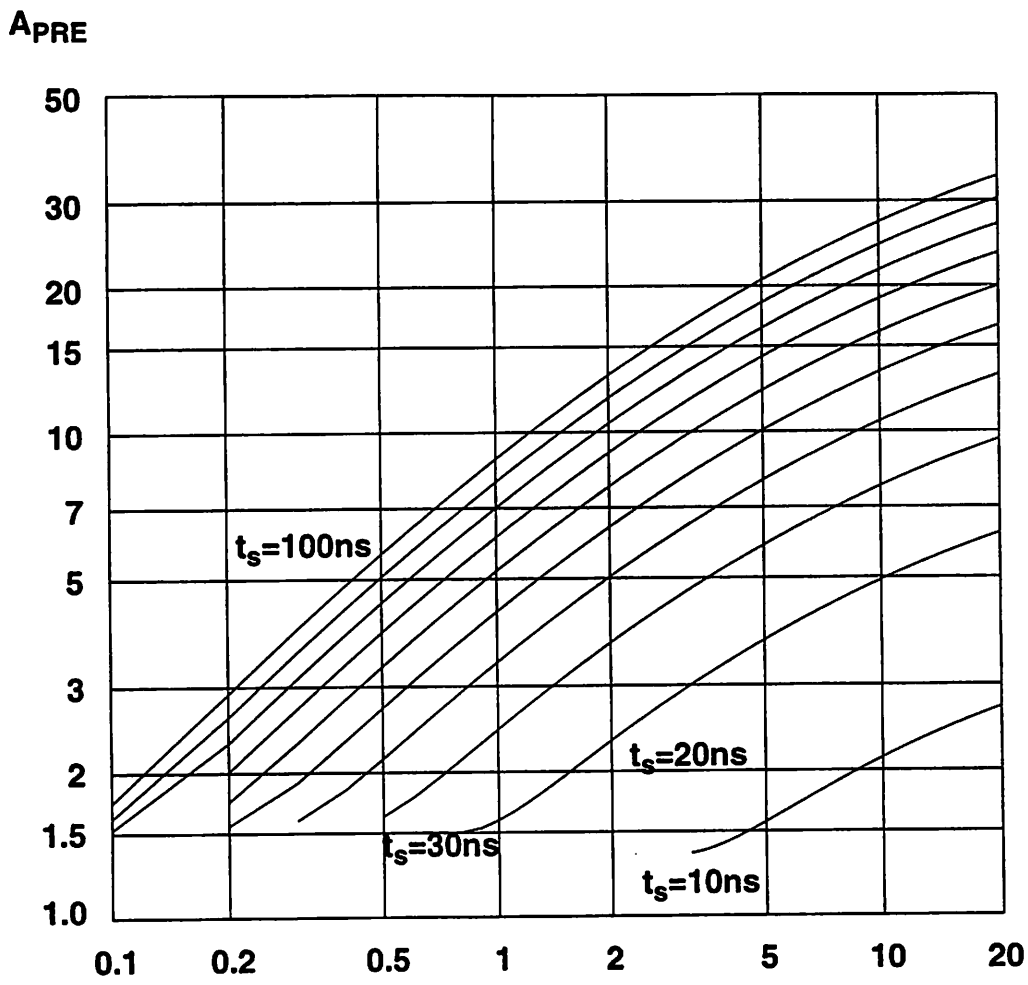


Figure 4.25 Optimized Preamp Gain versus Current Density, $\alpha_1=\alpha_2=\alpha_L=1$, $\zeta=1$, G=1

Figures 4.24 and 4.25 show that the optimal preamplifier gain A_{PRE} tends to decrease as the required settling time is decreased (required speed is increased). This result makes sense because an increase in preamplifier gain implies a reduction in the unity gain frequency of the load transistor in the preamplifier. As the speed requirement increases, the required value for the unity gain frequency increases. As a result, the optimum preamplifier gain is reduced.

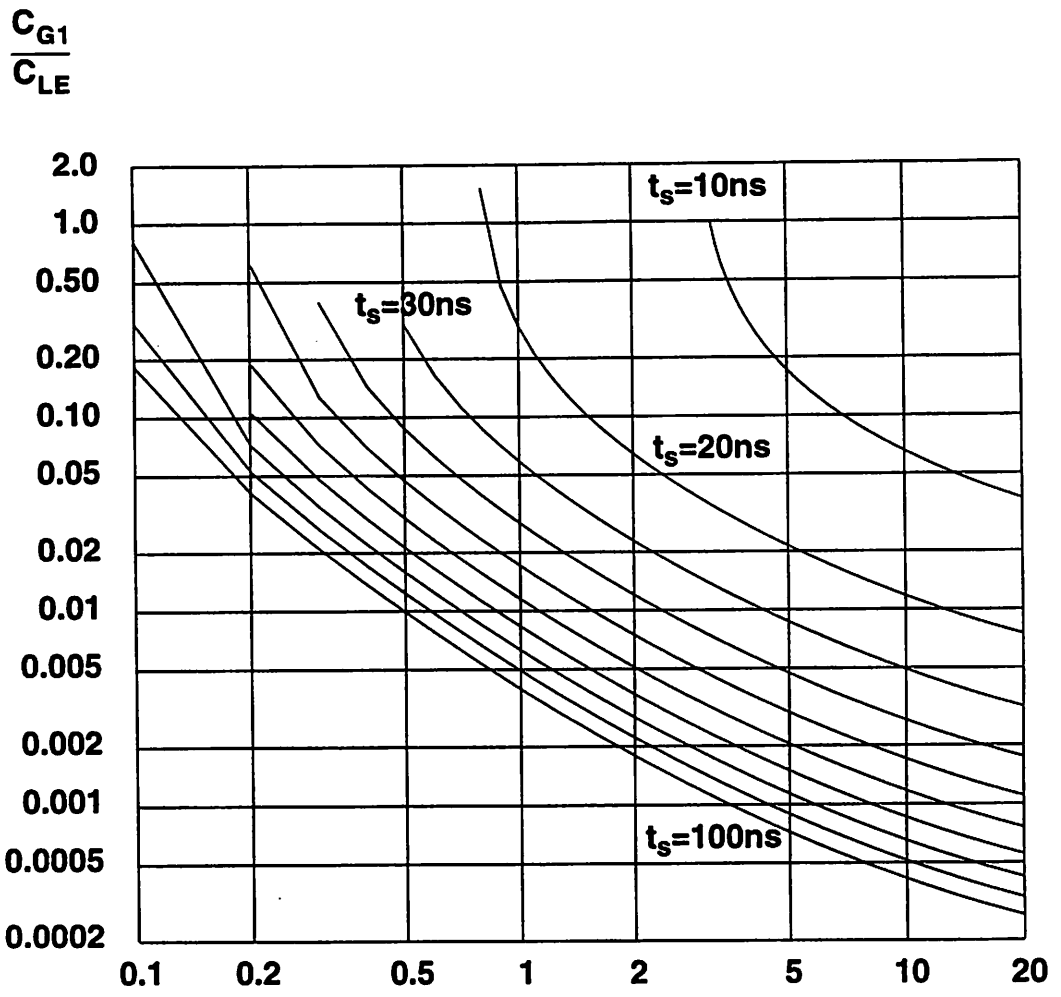


Figure 4.26 Optimized Input Capacitance versus Current Density, $\alpha_1=\alpha_2=\alpha_L=1$, $\zeta=1$, G=1

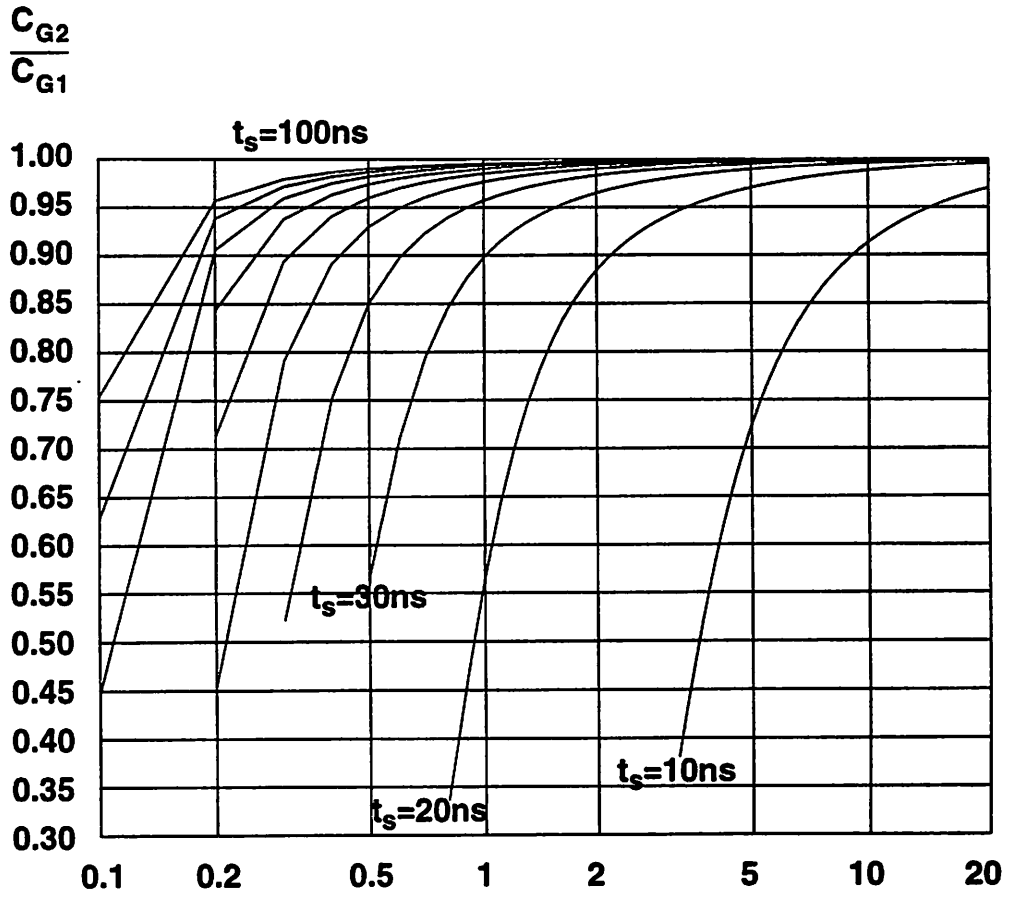


Figure 4.27 Optimized Second Stage Input Capacitance versus Current Density, $\alpha_1=\alpha_2=\alpha_L=1$, $\zeta=1$, G=1

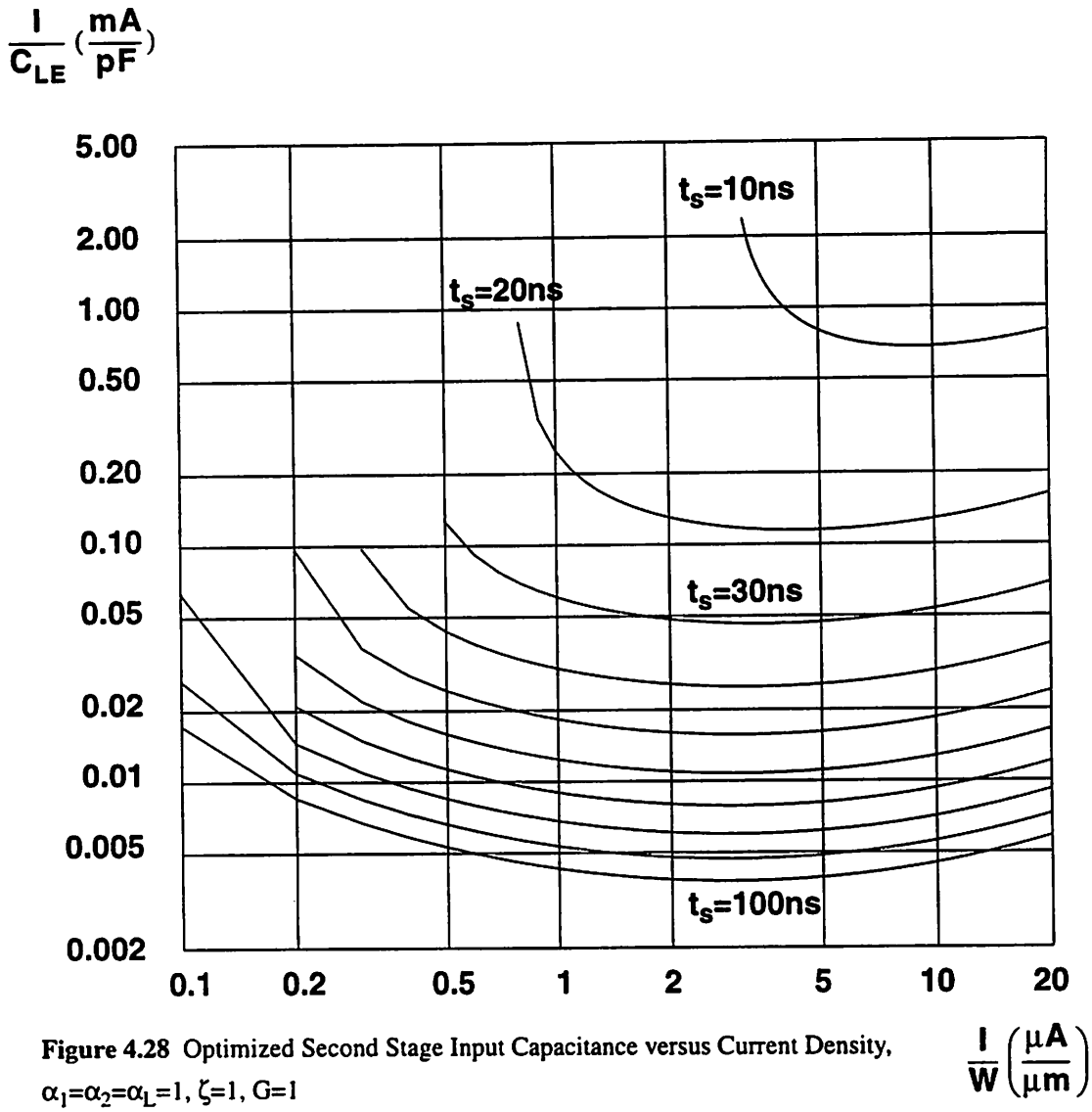


Figure 4.28 Optimized Second Stage Input Capacitance versus Current Density, $\alpha_1=\alpha_2=\alpha_L=1$, $\zeta=1$, G=1

Figure 4.28 shows that, as for the single stage amplifiers, the optimum current density for the preamplifier driving a single stage amplifier increases as the required speed increases. Again, this tendency is due to the fact that increasing the current density increases the unity gain frequency. It is also worth noting that for long settling times the preamplifier driving a single stage amplifier is actually more efficient in terms of power dissipation than the single stage amplifier or the telescopic cascode.

4.2.4 Two Stage Amplifier with Standard Miller Compensation

In some applications, especially those using small supply voltages, gain and signal swing requirements may force the designer to use an amplifier consisting of two high gain stages. An example of a single ended version of such an amplifier is shown in figure 4.29. A compensation capacitor applies negative feedback around the output stage. Without this capacitor, the amplifier has two low frequency poles. When feedback is applied around that amplifier, the real parts of the pole positions remain at low frequency while the imaginary parts become very large. This behavior implies poor stability and long settling times. With the compensation capacitor, sometimes called a pole splitting capacitor, the amplifier has a low frequency pole and a high frequency pole. When feedback is applied around the amplifier, the poles move toward each other along the real axis until they meet and move into the complex plane. Using the compensation capacitor it is possible to design the amplifier such that the poles remain real even after feedback is applied. Thus, the settling time is reduced and stability is improved. Miller compensation is further discussed by Gray and Meyer[Gray93].

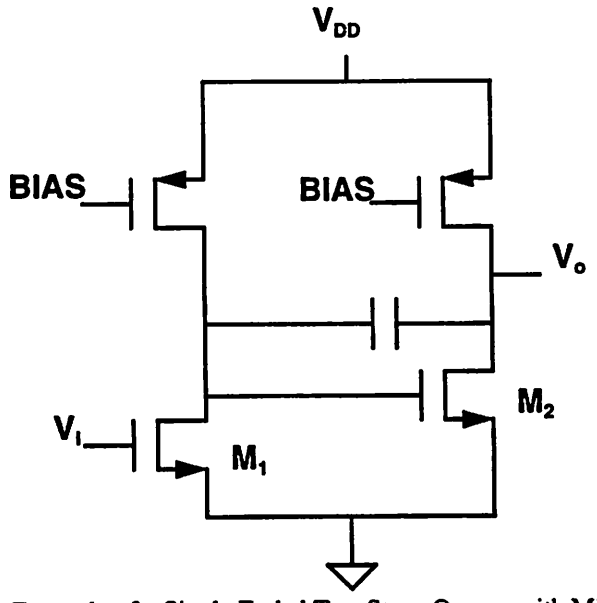
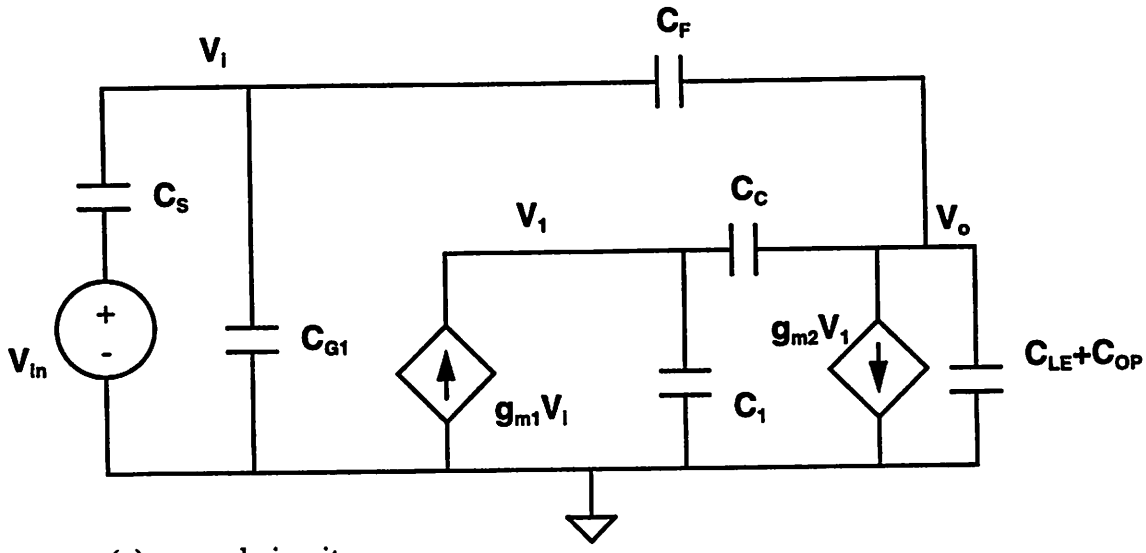
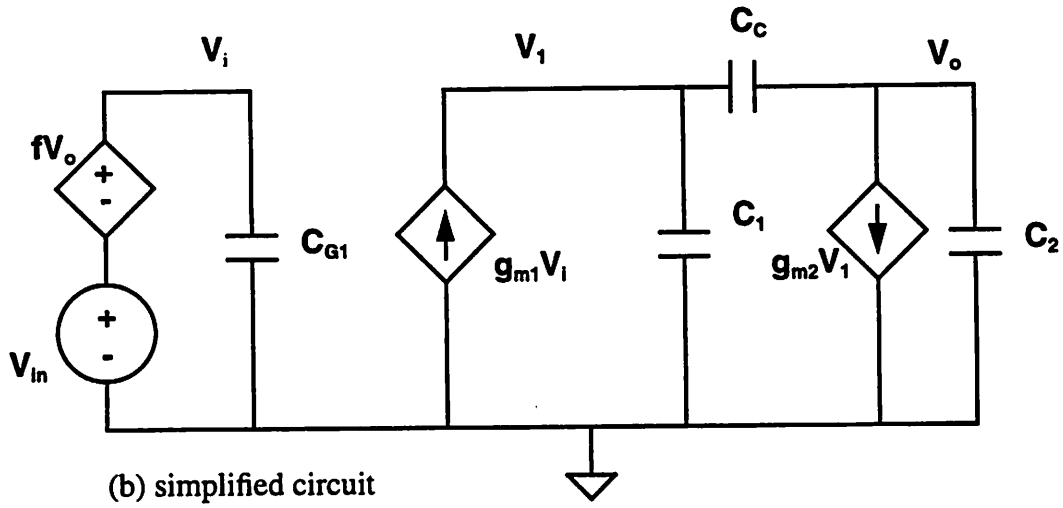


Figure 4.29 Example of a Single Ended Two Stage Opamp with Miller Compensation

The amplifier is modeled as shown in figure 4.30 with the usual simplification to the feedback network.



(a) general circuit



(b) simplified circuit

Figure 4.30 Small Signal Model for the Hold Phase of a Switched Capacitor Sample and Hold Circuit Using a Two Stage Opamp with Miller Compensation

The capacitor C_1 represents the load capacitance on the first stage and is the sum of the input capacitance C_{G2} of stage 2 and the parasitic capacitance $\alpha_1 C_{G1}$ contributed by stage 1.

$$C_1 = C_{G2} + \alpha_1 C_{G1} \tag{4.68}$$

C_2 is the capacitance loading stage 2 and is composed of the external load, the loading presented by the feedback network, and an output parasitic capacitance.

$$C_2 = C_{LE} + C_F(1-f) + \alpha_2 C_{G2} \tag{4.69}$$

By nodal analysis of the circuit of figure 4.30B, the following closed loop transfer function is obtained for the amplifier.

$$A_{CL}(s) = \frac{N_{CL}(s)}{D_{CL}(s)} = \frac{\frac{g_{m1}C_C}{C_T^2}\left(\frac{g_{m2}}{C_C} - s\right)}{s^2 + \frac{(g_{m2} - fg_{m1})C_C s}{C_T^2} + \frac{fg_{m1}g_{m2}}{C_T^2}} \tag{4.70}$$

where

$$C_T^2 = C_1C_2 + C_1C_C + C_2C_C \tag{4.71}$$

The closed loop poles and right half plane zero are shown in figure 4.31.

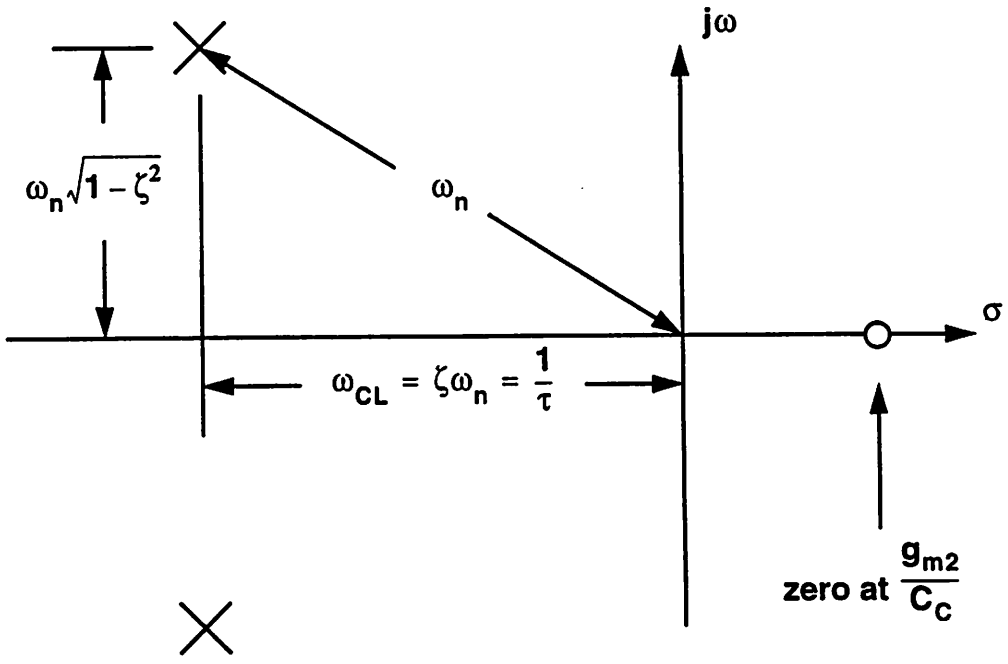


Figure 4.31 S-Plane Pole and Zero Positions for a Switched Capacitor Sample and Hold Circuit Using a Miller Compensated Two Stage Opamp

As before, it is useful to rewrite the denominator of the closed loop gain expression into a general form for the characteristic polynomial of a second order system.

$$D_{CL} = s^2 + \frac{(g_{m2} - fg_{m1})C_C s}{C_T^2} + \frac{fg_{m1}g_{m2}}{C_T^2} = s^2 + 2\zeta\omega_n s + \omega_n^2 \tag{4.72}$$

From the above equation, the following relations may be obtained.

$$\omega_{CL} = \frac{1}{\tau} = \frac{(g_{m2} - fg_{m1})C_C}{2C_T^2} = \sqrt{\frac{fg_{m1}g_{m2}}{C_T^2}} \tag{4.73}$$

The transconductances can be related to the input capacitances by the following formulas.

$$g_{m1} = \omega_{T1}C_{G1} \tag{4.74}$$

$$g_{m2} = \omega_{T2}C_{G2} \tag{4.75}$$

The quantity r_c is now introduced to represent the ratio of the sizes of the two amplifier stages.

$$r_c = \frac{C_{G2}}{C_{G1}} \tag{4.76}$$

Now equations (4.68), (4.71), and (4.73) through (4.76) can be used to obtain the following equation for the input capacitance of the first stage amplifier.

$$C_{G1} = \frac{2\omega_{CL}C_2}{[\omega_{T2}r_c - f\omega_{T1} - 2\omega_{CL}(r_c + \alpha_1)]}\left[1 + \frac{(r_c + \alpha_1)\omega_{CL}(\omega_{T2}r_c - f\omega_{T1})}{2f\zeta^2 r_c \omega_{T1}\omega_{T2}}\right] \tag{4.77}$$

By substituting (4.33), (4.69), and (4.76) into (4.77), a third order equation in C_{G1} is obtained. The analysis will not be shown here. However, the graphical results of the computer optimization are shown in figures 4.32, 4.33, 4.34 and 4.35.

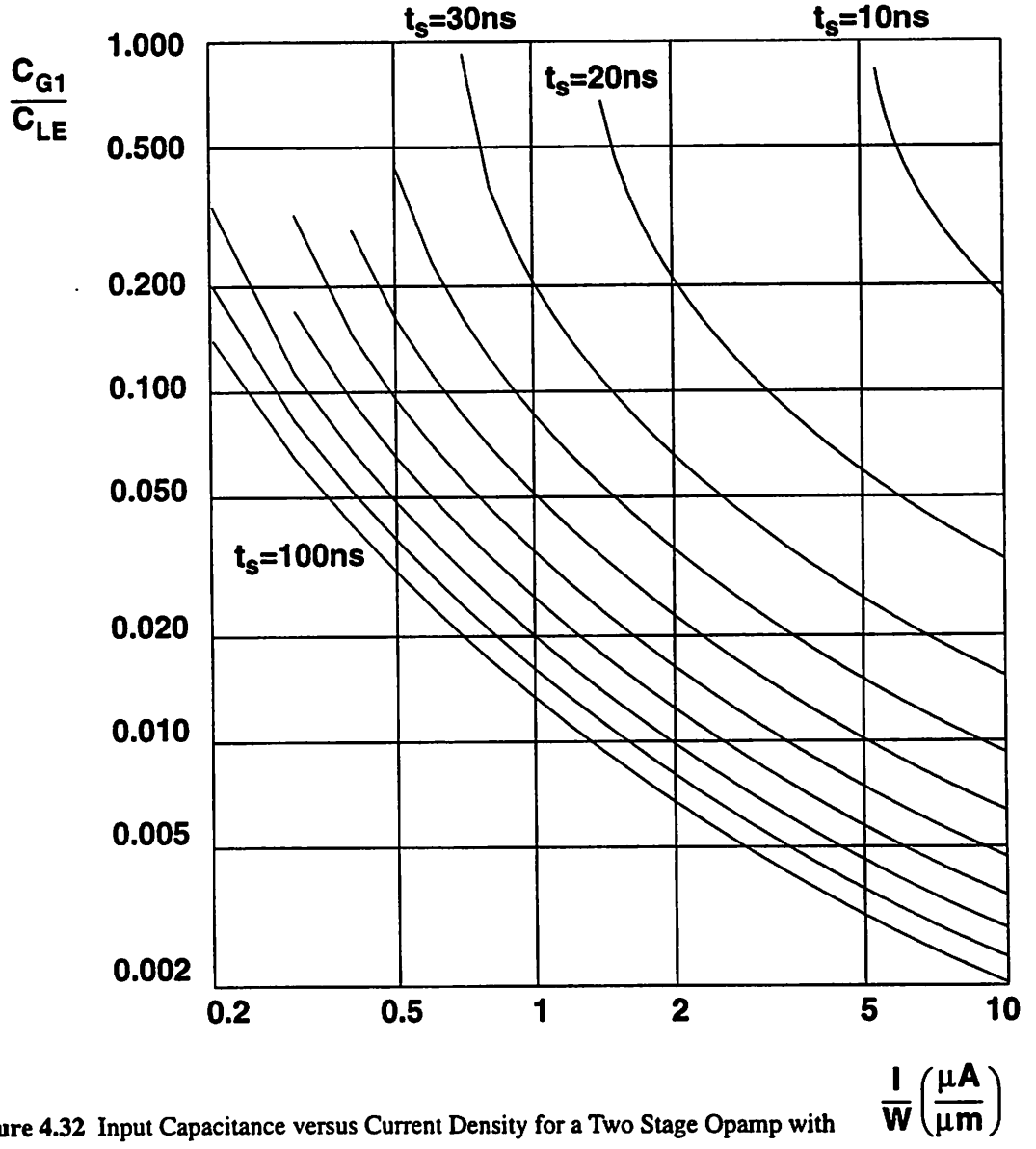


Figure 4.32 Input Capacitance versus Current Density for a Two Stage Opamp with Miller Compensation, ζ=1, $\alpha_1=\alpha_2$=1, C_F/C_{LE}=0.5, G=1

Figure 4.32 shows that the optimum input capacitance gets smaller as current density is increased. This makes sense because increasing the current density of a transistor allows the same transconductance to be obtained with a smaller device.

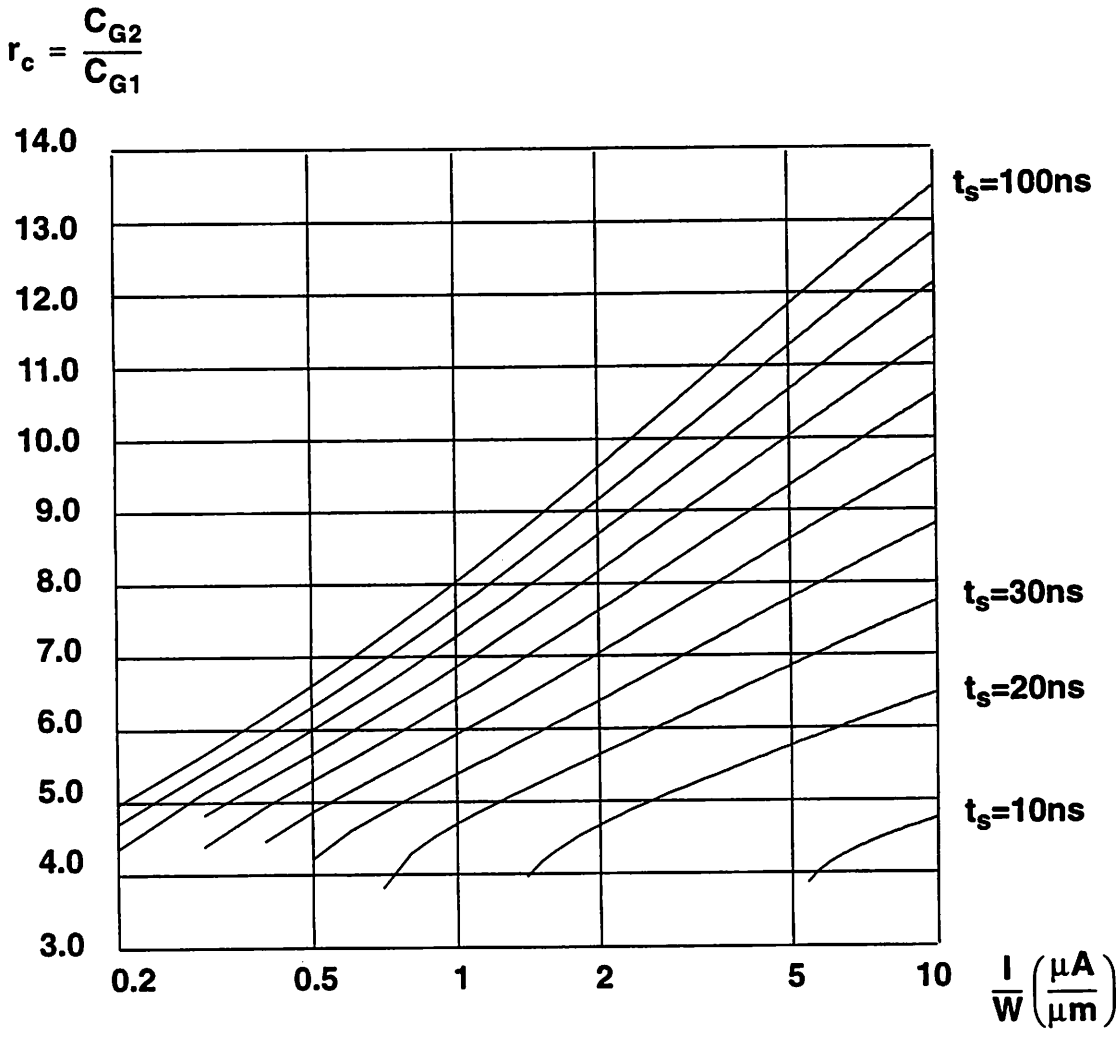


Figure 4.33 r_c versus Current Density for a Two Stage Opamp with Miller Compensation, ζ=1, $\alpha_1=\alpha_2$=1, C_F/C_{LE}=0.5, G=1

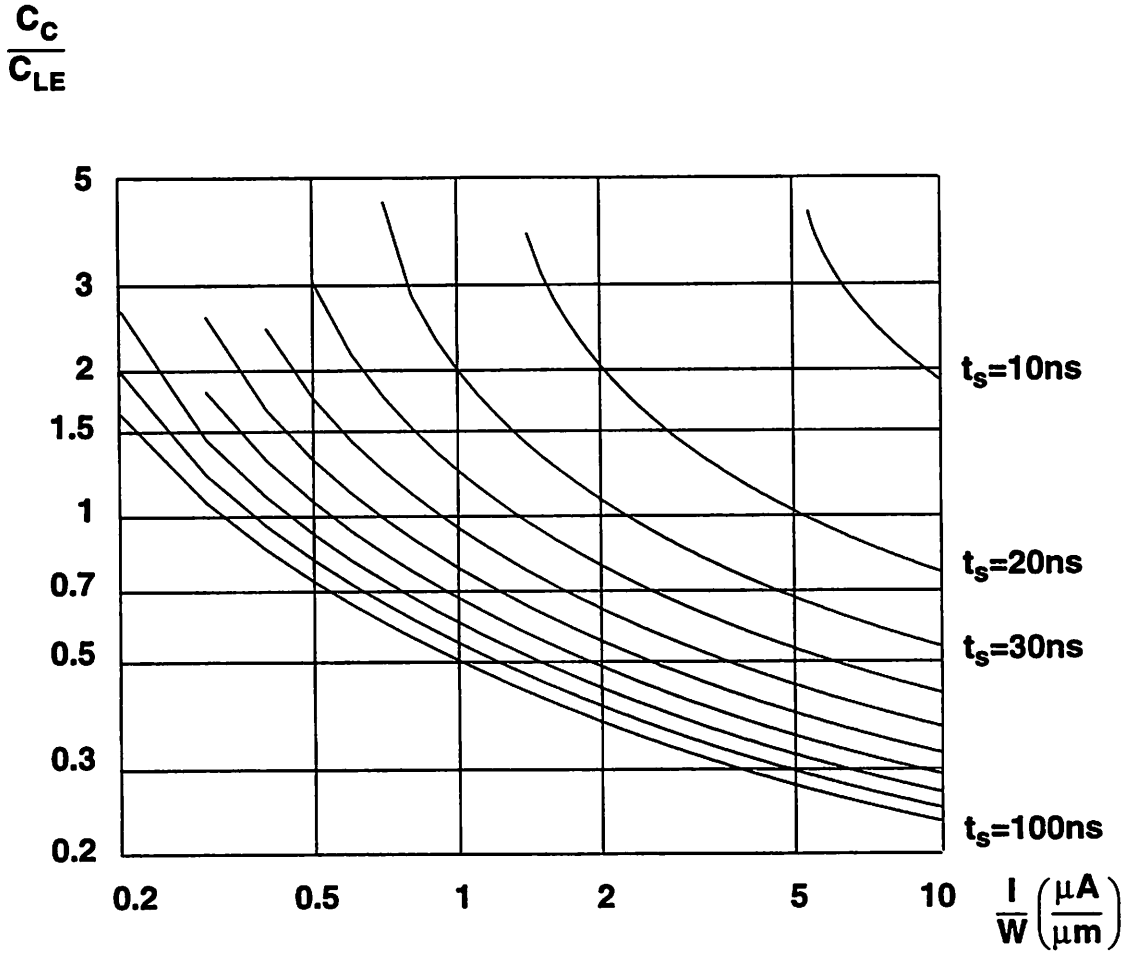


Figure 4.34 Compensation Capacitance versus Current Density for a Two Stage Opamp with Miller Compensation, ζ=1, $\alpha_1=\alpha_2$=1, C_F/C_{LE}=0.5, G=1

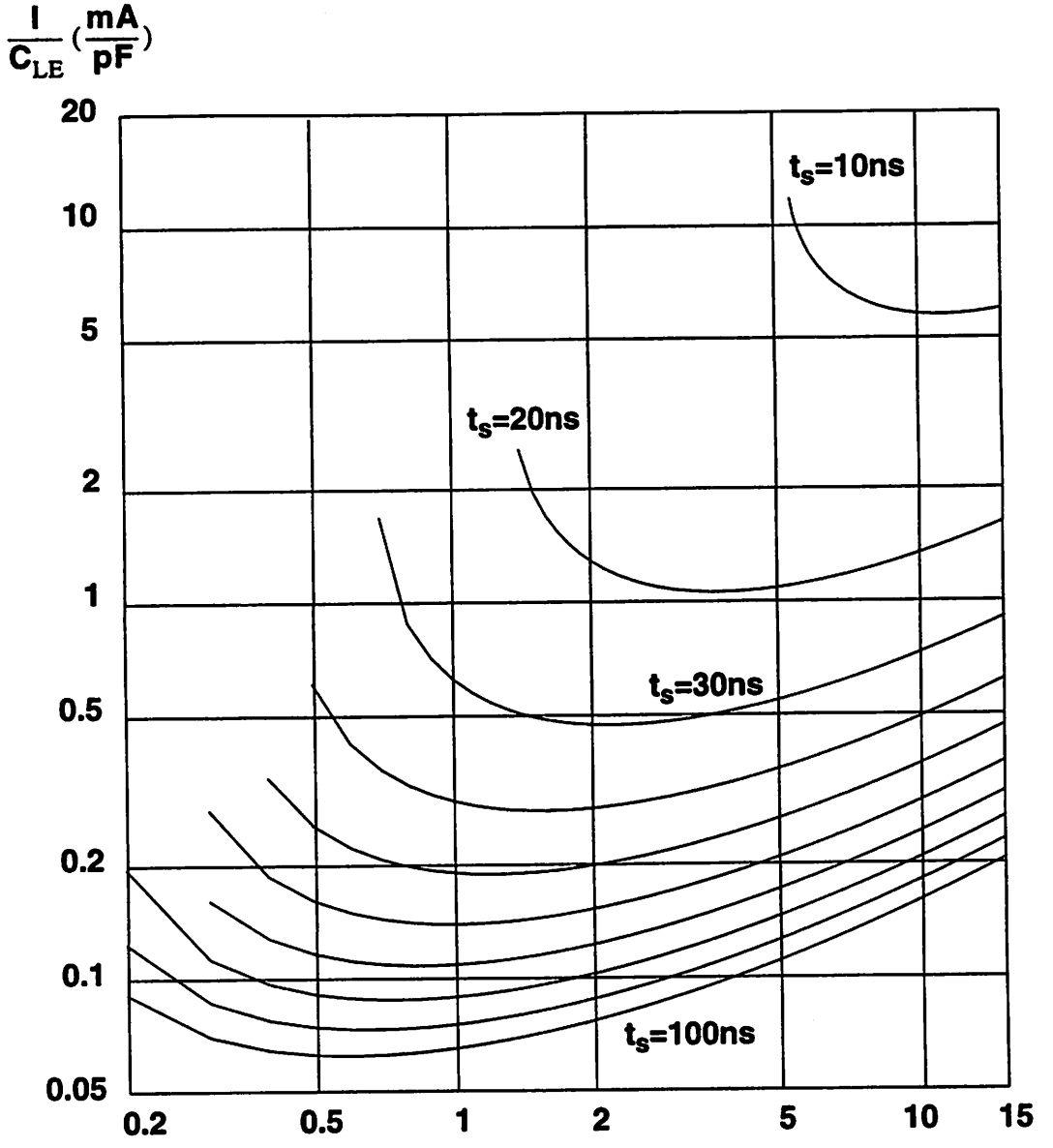


Figure 4.35 Total Bias Current versus Current Density for Two Stage Amplifier with Miller Compensation, ζ=1, α_1=α_2=1, C_P/C_{LE}=0.5, G=1

Figure 4.35 shows that the optimum current density increases as the required speed increases. This result matches that for the other architectures. This occurs because increasing the required speed increases the required unity gain frequency, and thus the required current density.

4.2.5 Two Stage Amplifier with Ahuja Style Compensation

The two stage amplifier with standard Miller compensation described in the previous section has the disadvantage of having a right half plane zero in the response due to the feedforward path through the compensation capacitor. This right half plane zero tends to pull poles toward the right half plane in the closed loop system. As a result, speed and stability are degraded. Techniques for dealing with this problem have been proposed. One such technique involves putting a resistor in series with the compensation capacitor[Gray93]. This has the effect of moving the zero farther into the right half plane where its effect is reduced. Another technique involves moving the input side of the compensation capacitor[Ahuja83][Ribner84]. An example of an amplifier using this second technique is shown in figure 4.36. In this configuration, the compensation capacitor is connected from the output node of the second stage to the cascode node of the first stage. Thus, the capacitor closes a feedback loop around a two pole system. This contrasts with the standard Miller compensation technique where the capacitor is connected from the output of the second stage to the output of the first stage and only closes a one pole feedback loop.

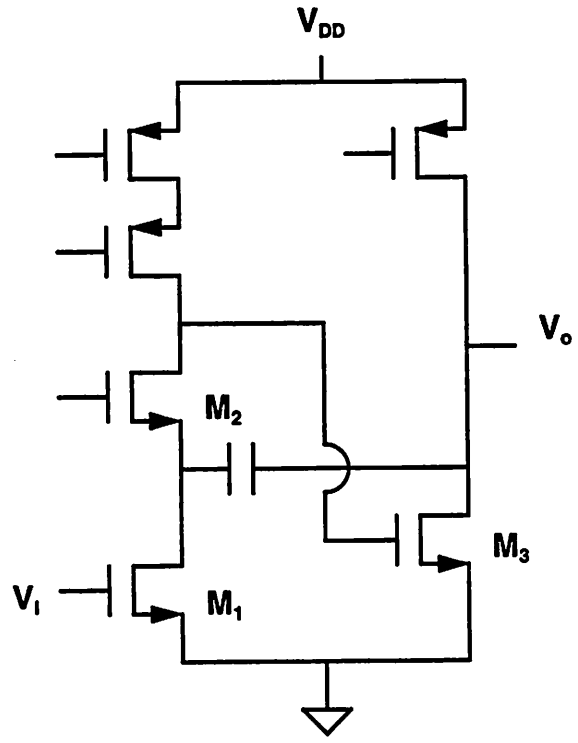
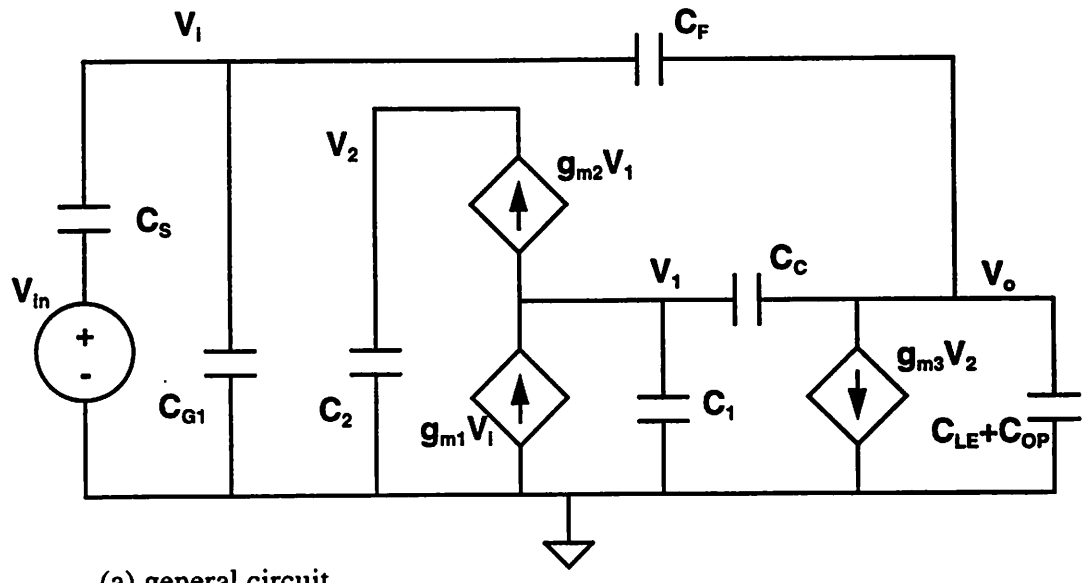
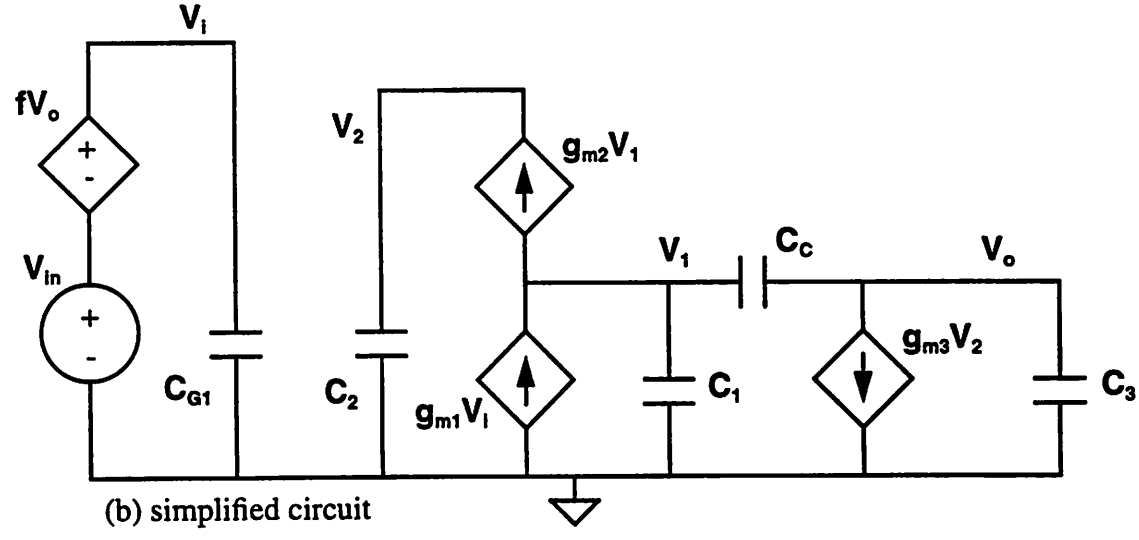


Figure 4.36 Two Stage Opamp with Ahuja Style Compensation

A general circuit model and a simplified small signal circuit model showing closed loop operation are shown in figure 4.37.



(a) general circuit



(b) simplified circuit

Figure 4.37 Small Signal Model for a Two Stage Opamp with Ahuja Style Compensation

The behavior of this amplifier may be understood in the following way. Consider the pole-zero constellation for the open loop amplifier with no compensation shown in figure 4.38a. The system has two low frequency poles and one pole at a very high frequency. If feedback is added to the sys-

tem without adding any compensation, the two low frequency poles will come together and split off while the high frequency pole will go still higher. The result is a system with two poles close to the jω axis. A system with this configuration will settle very slowly.

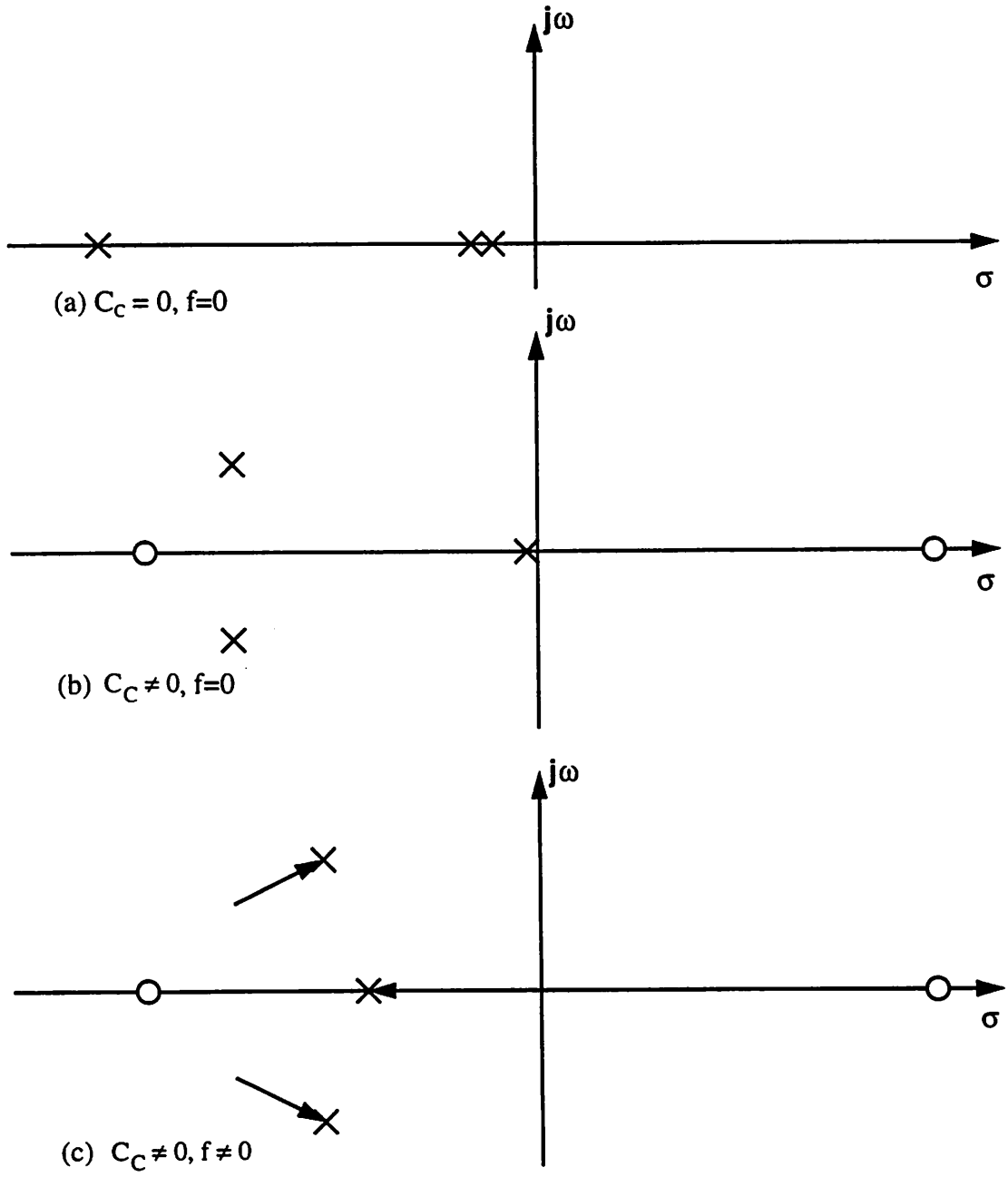


Figure 4.38 s-Plane Pole Zero Positions for a Two Stage Opamp with Ahuja Compensation for Various Levels of Compensation and Feedback

Now, if the compensation capacitor is added to the open loop system as shown in figure 4.36, the pole-zero constellation will look like the one in figure 4.38b. Note that two zeros have been added to the system, one right half plane zero and one left half plane zero. The right half plane zero is approximately the same distance from the jω axis as the left half plane zero. The locations z of the zeros are given by the following formula.

$$z = \pm\sqrt{\frac{g_{m2}g_{m3}}{C_2C_C}} \tag{4.78}$$

Also note that the high frequency pole present in figure 4.38a has moved to the right while one of the low frequency poles has moved to the left. The poles then meet and split off the real axis to form a complex conjugate pair of poles with highly negative real parts as shown in figure 4.38b. The low frequency pole that does not move out to meet the high frequency pole moves closer to the jω axis.

Finally, when feedback is added to the system, the constellation of figure 4.38c results. The low frequency pole moves left along the real axis while the two higher frequency complex poles move to the right.

The transfer function from input to output for the closed loop system is given by the following formula obtained by nodal analysis on the circuit of figure 4.37b.

$$A_{CL}(s) = \frac{N_{CL}(s)}{D_{CL}(s)} \tag{4.79}$$

$$A_{CL}(s) = \frac{\frac{g_{m1}}{C_2C_T^2}(g_{m2}g_{m3} - C_2C_Cs^2)}{s^3 + \frac{[g_{m2}C_2(C_3+C_C) - fg_{m1}C_2C_C]}{C_2C_T^2}s^2 + \frac{g_{m2}g_{m3}C_C}{C_2C_T^2}s + \frac{fg_{m1}g_{m2}g_{m3}}{C_2C_T^2}} \tag{4.80}$$

where

$$C_T^2 = C_1C_3 + C_1C_C + C_3C_C \tag{4.81}$$

With the assumption that the closed loop pole-zero constellation looks like the one in figure 4.38c, the characteristic polynomial D_{CL} is forced to take the following form.

$$D_{CL}(s) = (s+\omega_{CL})(s^2+2\zeta\omega_n s+\omega_n^2) \tag{4.82}$$

By equating the D_{CL} of equation (4.80) to the D_{CL} of equation (4.82), the following relations may be obtained.

$$\omega_{CL}+2\zeta\omega_n = \frac{g_{m2}(C_3+C_C)}{C_T^2} - \frac{fg_{m1}C_C}{C_T^2} \tag{4.83}$$

$$2\zeta\omega_n\omega_{CL}+\omega_n^2 = \frac{g_{m2}g_{m3}C_C}{C_2C_T^2} \tag{4.84}$$

$$\omega_n^2\omega_{CL} = \frac{fg_{m1}g_{m2}g_{m3}}{C_2C_T^2} \tag{4.85}$$

As before, each transconductance can be related to the gate capacitance of the transistor it is associated with by the unity gain frequencies as shown below.

$$g_{m1} = \omega_{T1}C_{G1} \tag{4.86}$$

$$g_{m2} = \omega_{T2}C_{G2} \tag{4.87}$$

$$g_{m3} = \omega_{T3}C_{G3} \tag{4.88}$$

The capacitors C_1, C_2, and C_3 are related to the three gate capacitances as follows. C_1 is the sum of the gate capacitance of M_2, the drain capacitance of M_2, and the drain capacitance of M_1.

$$C_1 = \alpha_1C_{G1}+C_{G2}+\alpha_CC_{G2} \tag{4.89}$$

C_2 is the sum of the gate capacitance of M_3 and the output capacitance of the first stage amplifier.

$$C_2 = \alpha_2C_{G2}+C_{G3} \tag{4.90}$$

C_3 is the sum of the external load capacitance, the feedback loading capacitance, and the output capacitance of the second stage of the amplifier.

$$C_3 = C_{LE}+C_F(1-f)+\alpha_3C_{G3} \tag{4.91}$$

The following relations are also defined in order to relate the sizes of the three transistors.

$$r_{C2} = \frac{C_{G2}}{C_{G1}} \tag{4.92}$$

$$r_{C3} = \frac{C_{G3}}{C_{G2}} \tag{4.93}$$

The analysis of this amplifier is not carried further by hand. However, a computer was used to optimize the amplifier to minimize power given a fixed speed. The results are shown in figures 4.39, 4.40, 4.41, 4.42, and 4.43. It was assumed here for simplicity that $\omega_{CL} = \omega_n$ and ζ=1. If these two assumptions had not been made, it was found that $\omega_{CL} = 1.1\omega_n$ and ζ=0.7 would have been approximately optimal.

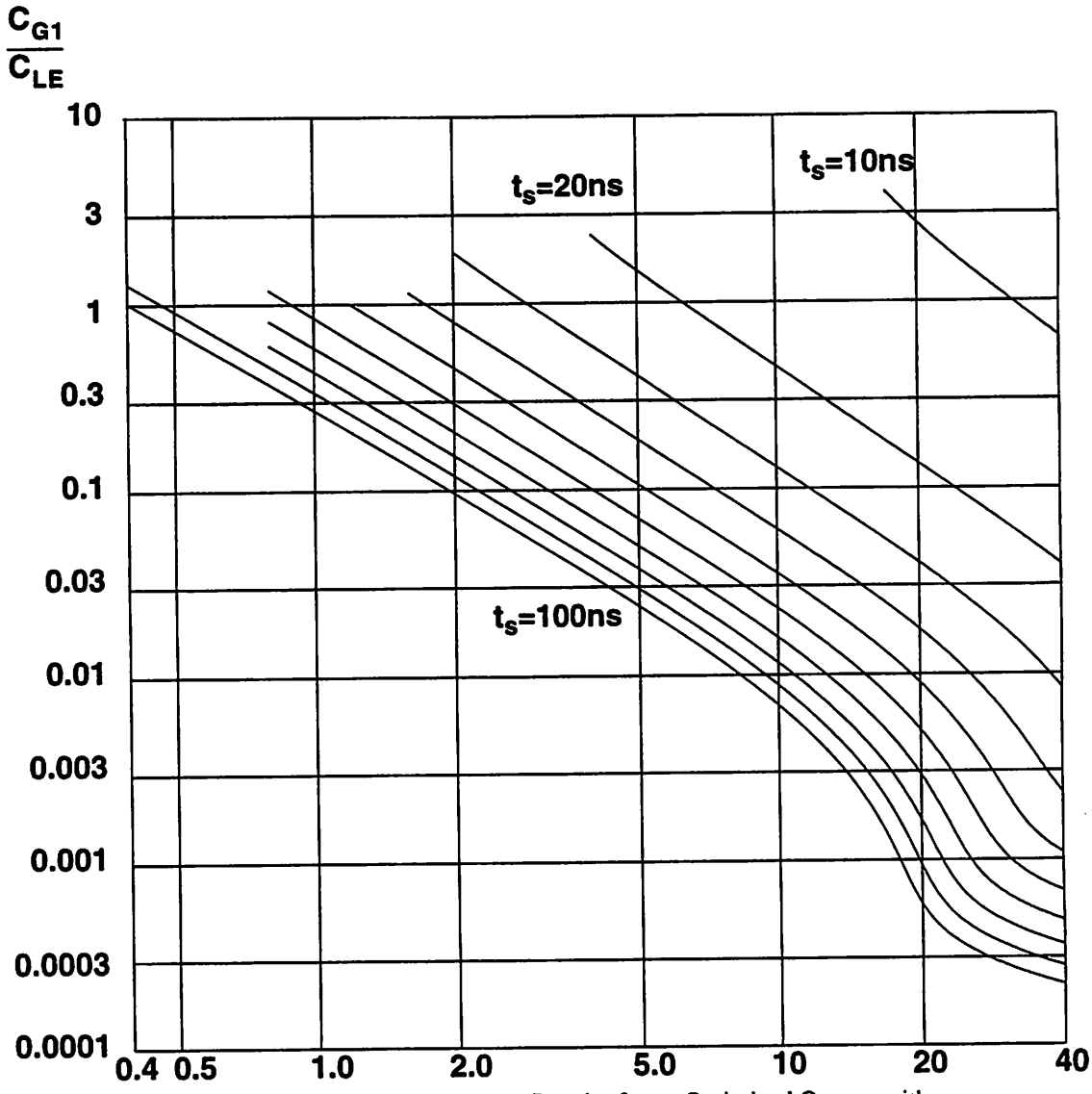


Figure 4.39 Input Capacitance versus Current Density for an Optimized Opamp with Ahuja Style Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_P/C_{LE}=0.5$, G=1

$$r_{c2} = \frac{C_{G2}}{C_{G1}}$$

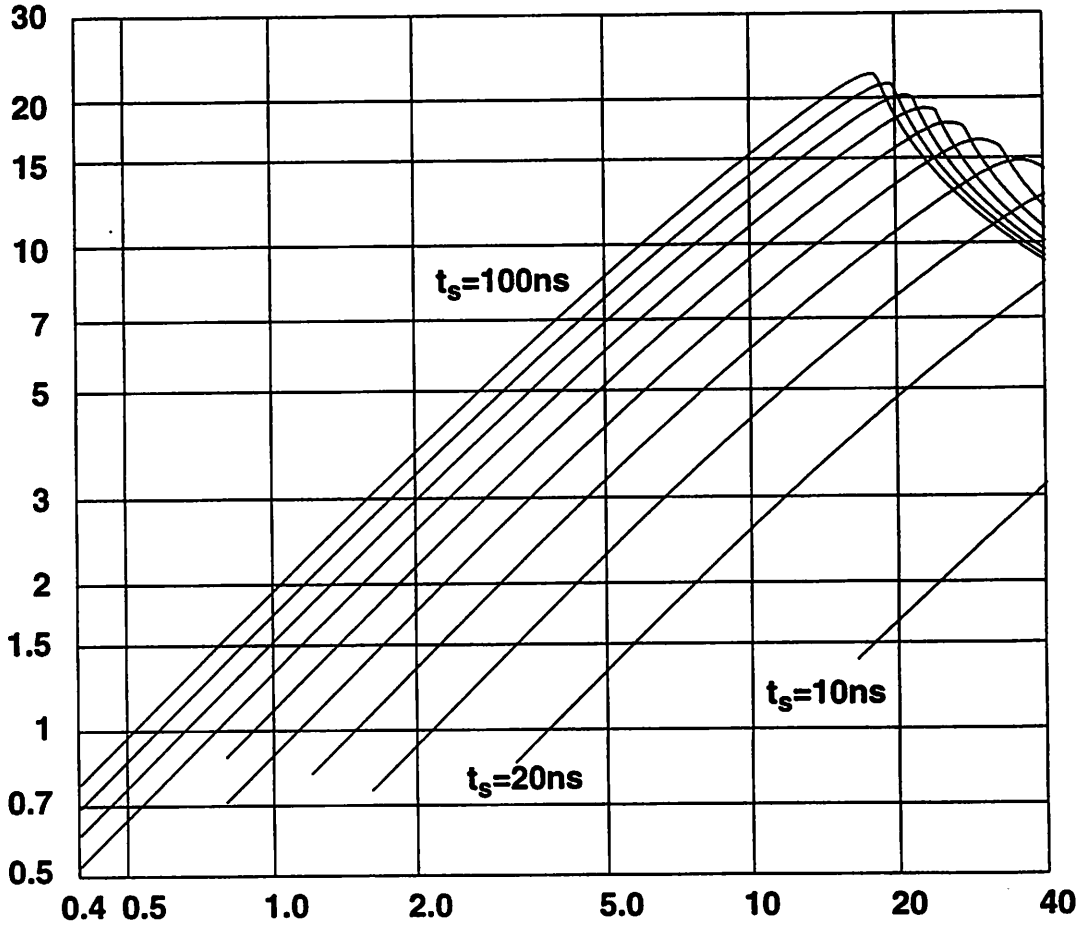


Figure 4.40 r_{c2} versus Current Density for an Optimized Opamp with Ahuja Style Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

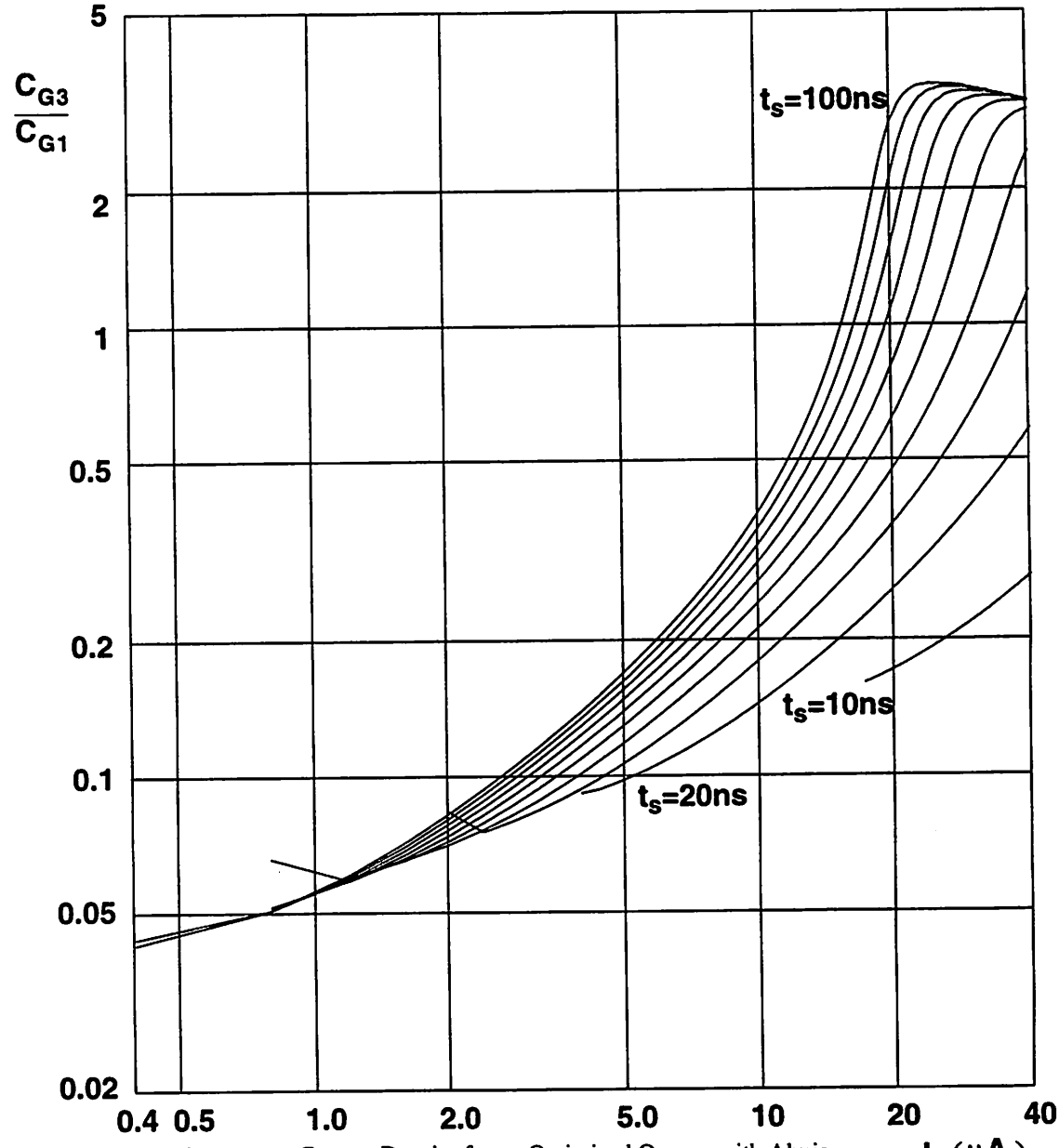


Figure 4.41 C_{G3} versus Current Density for an Optimized Opamp with Ahuja Style Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

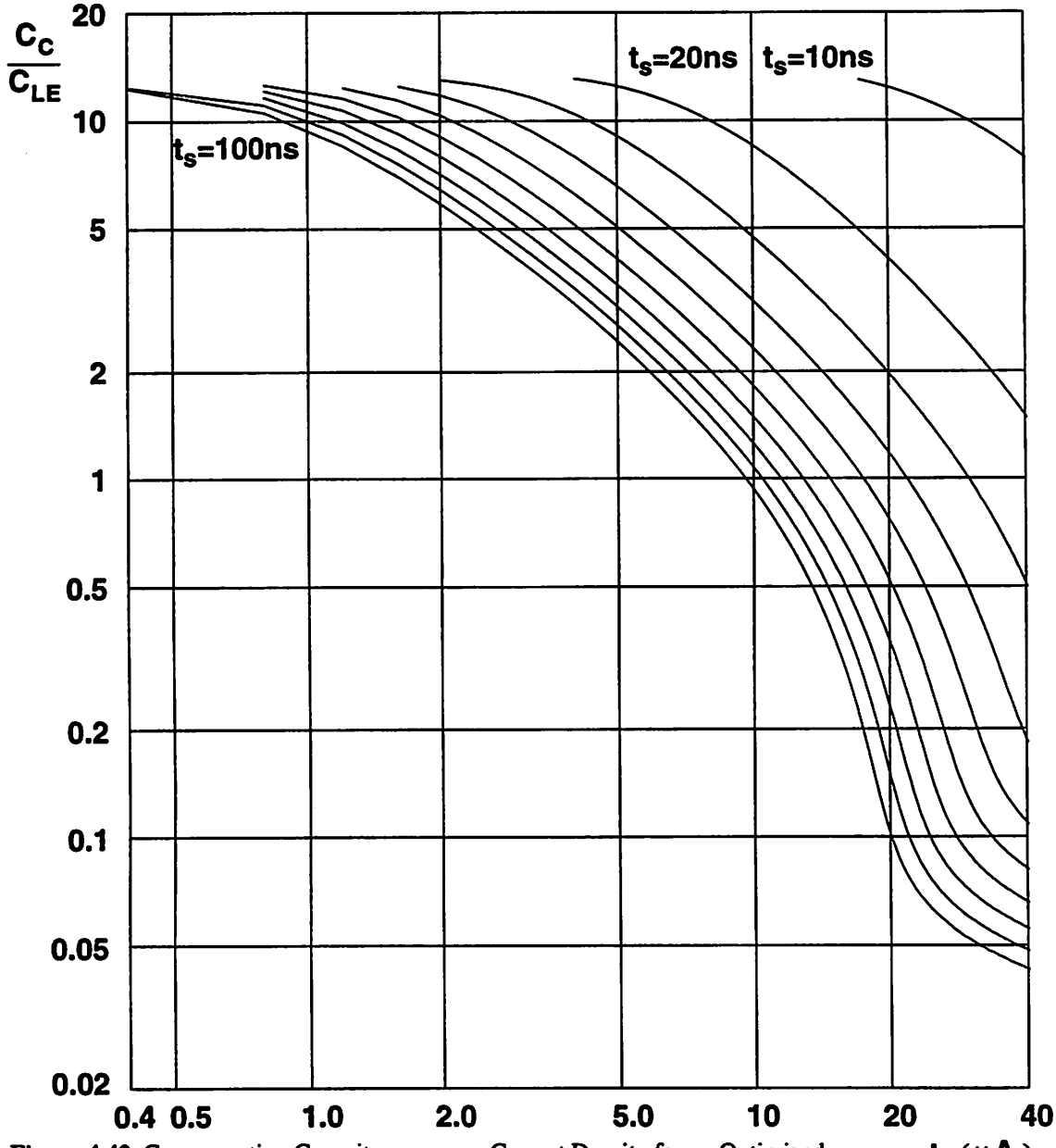


Figure 4.42 Compensation Capacitance versus Current Density for an Optimized Opamp with Ahuja Style Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_P/C_{LE}=0.5$,

Figure 4.42 shows that the optimal compensation capacitance decreases as the current density increases. This result makes sense because the compensation capacitor works by providing feedback around part of the amplifier. Increasing the current density tends to reduce the required sizes of the transistors in the amplifier, and thus reduces the parasitic capacitances in the amplifier. With

the parasitic capacitances reduced, the same feedback can be provided with a smaller compensation capacitor.

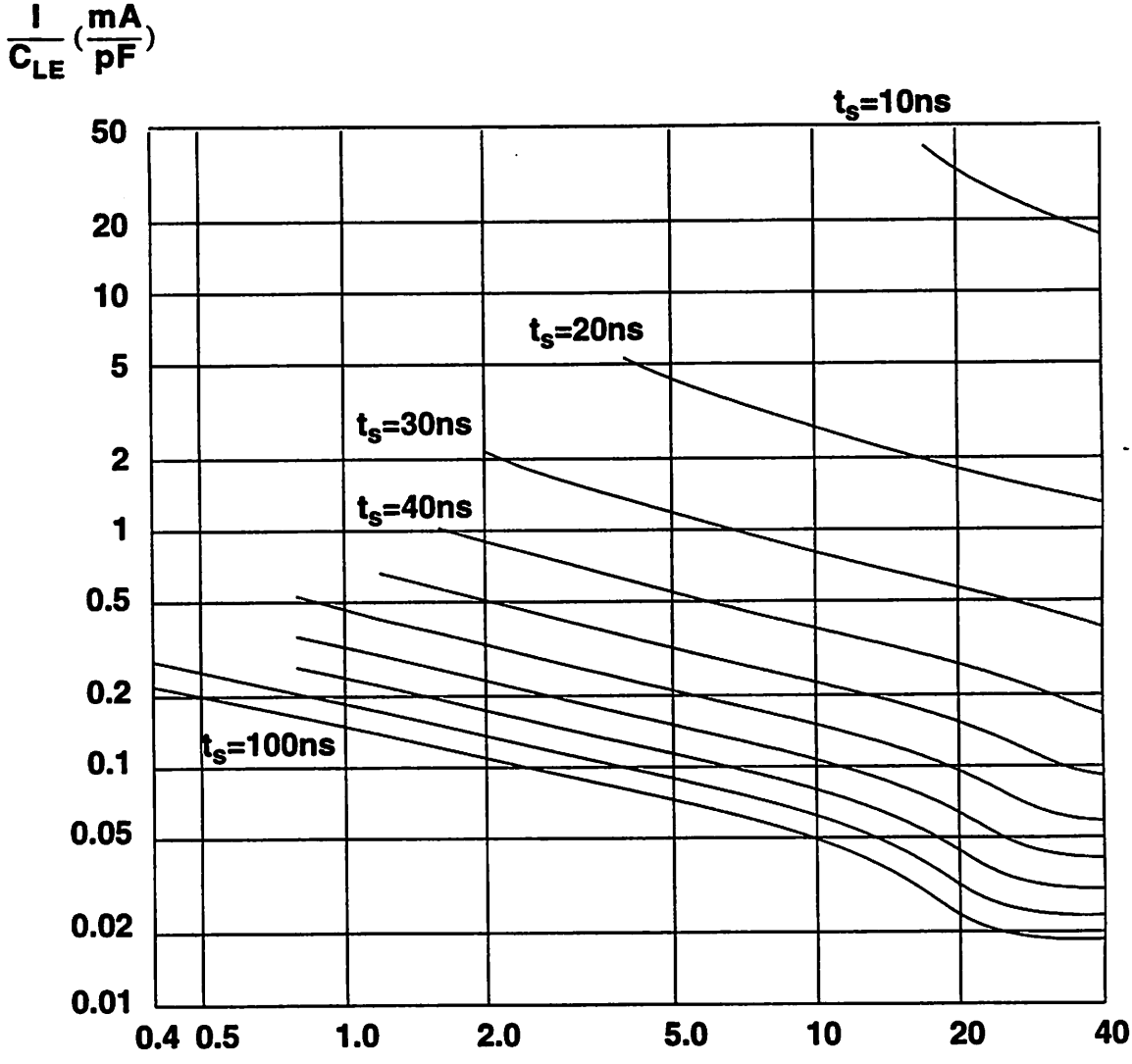


Figure 4.43 Total Bias Current versus Current Density for an Optimized Opamp with Ahuja Style Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$,

Figure 4.43 shows that increasing current density tends to reduce power dissipation for this type of amplifier architecture. This happens because increasing the current density allows the use of a smaller compensation capacitor. Using smaller compensation capacitors tends to improve efficiency because it reduces the load to be driven by the different pieces of the amplifier.

4.2.6 Three Stage Amplifier with Nested Miller Compensation

In applications where the available power supply voltage is low, a two stage amplifier may still not be enough to meet a gain specification. Three stage amplifiers are of interest for these applications. A three stage amplifier with two compensation capacitors has been proposed by Escauzier[E-schauzier92]. A single ended example of such an amplifier is shown in figure 4.44, and a small signal model is shown in figure 4.45.

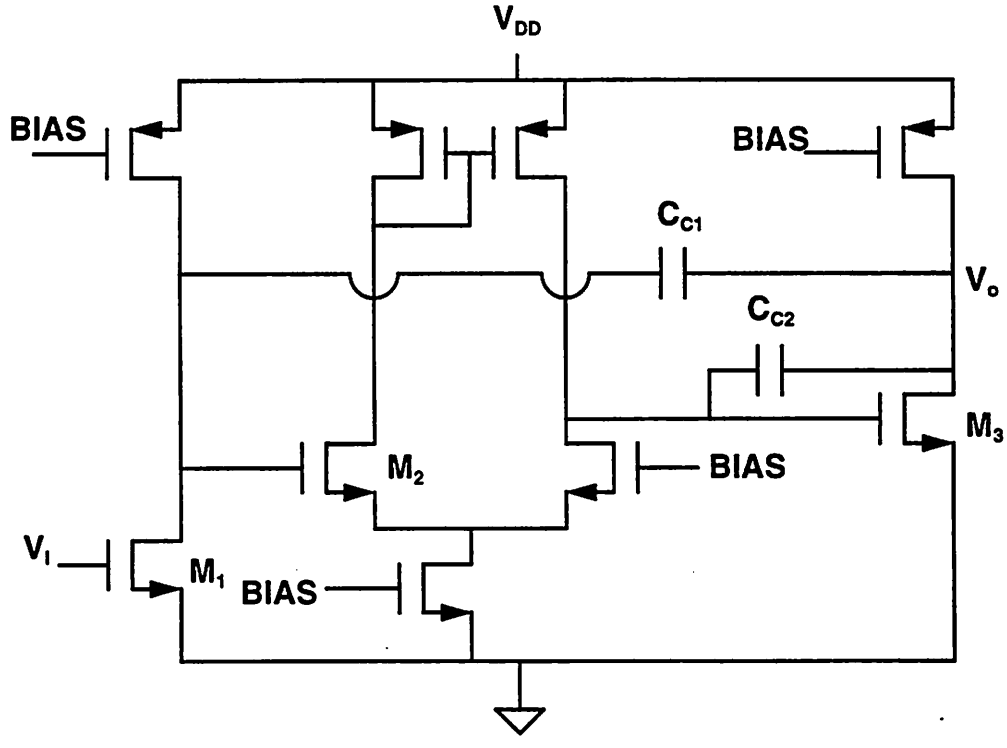


Figure 4.44 A Three Stage Opamp with Nested Miller Compensation

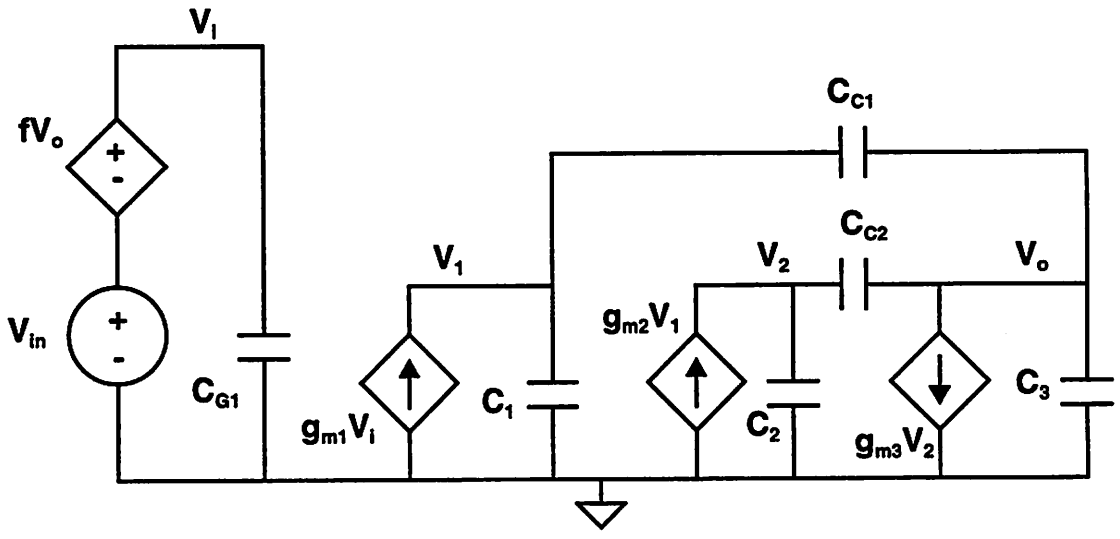


Figure 4.45 Simplified Small Signal Equivalent Circuit of a Three Stage Opamp with Nested Miller Compensation in a Switched Capacitor Gain Circuit (Feedforward Path through the Feedback Capacitor Neglected)

The pole-zero constellation for the open loop uncompensated system is shown in figure 4.46a. It consists of three poles, all at low frequencies. If feedback is placed around this amplifier without compensation, one pole will move along the real axis to the left, and the other two poles will move toward the right half plane. The closed loop amplifier will be unstable without compensation.

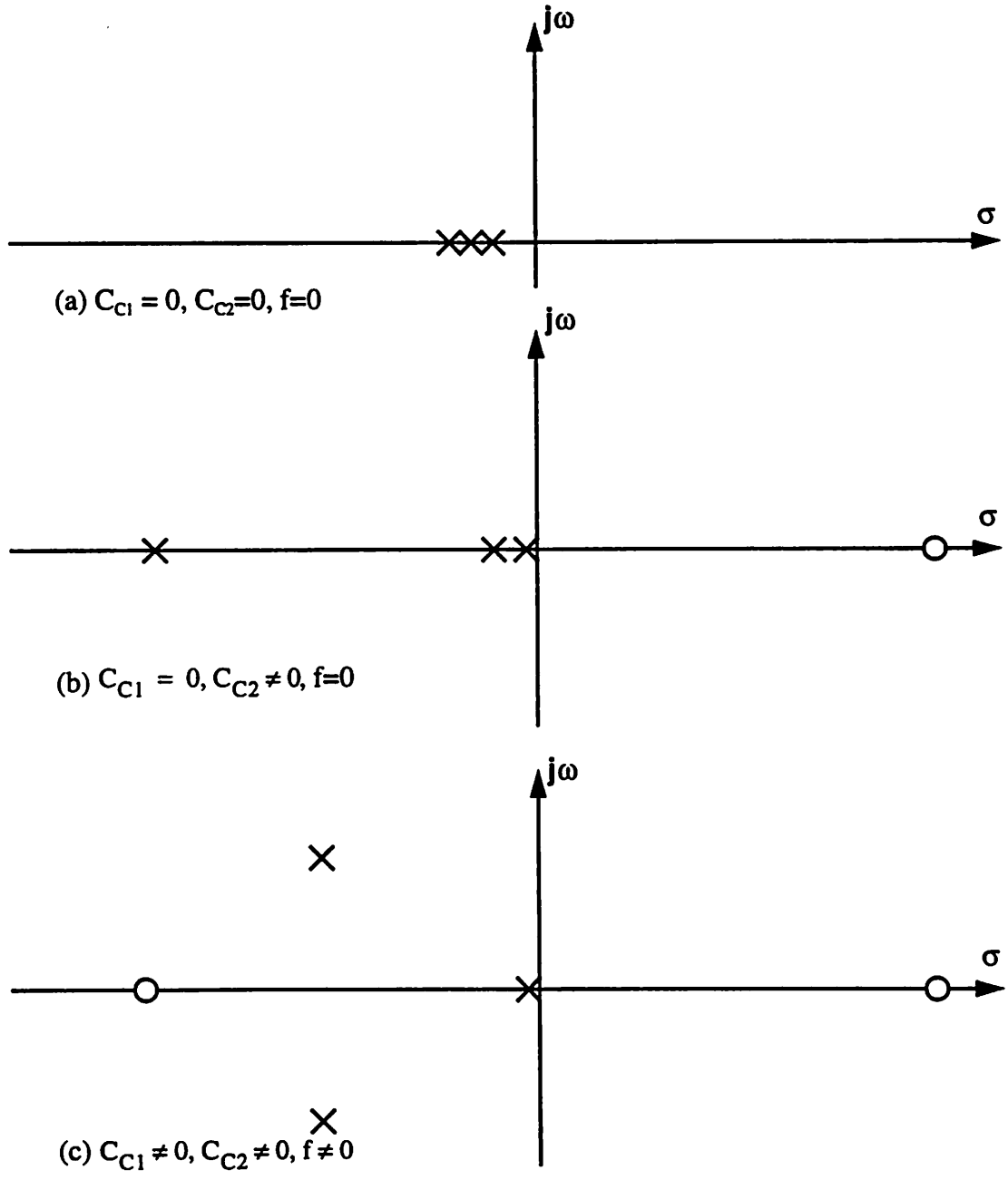


Figure 4.46 s-Plane Pole Positions for a Three Stage Opamp with Nested Miller Compensation in an Open Loop Configuration

Now consider the effect of C_{C2} alone. The pole-zero constellation corresponding to this case is shown in figure 4.46b. A right half plane zero has been added, and one of the poles has been pushed to the left. The pole positions now resemble those of figure 4.46b.

Now, the other compensation capacitor is added. This change results in the pole-zero constellation of figure 4.46c. A zero has been added to the left half plane, and the high frequency pole has been pulled in to meet one of the low frequency poles. Both of these poles then move off the real axis to form a pair of complex poles. When feedback is added to this system, it behaves like figure 4.47.

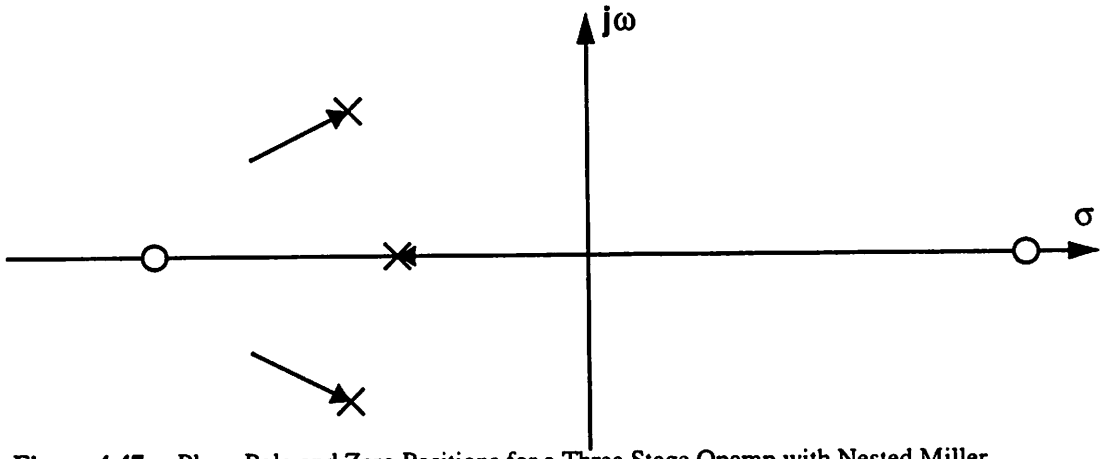


Figure 4.47 s-Plane Pole and Zero Positions for a Three Stage Opamp with Nested Miller Compensation During the Hold Phase in a Switched Capacitor Gain Circuit

To quantitatively describe the transient behavior of the amplifier, nodal analysis is applied to figure 4.45 yielding the following closed loop transfer function.

$$A_{CL} = \frac{N_{CL}}{D_{CL}} = \frac{\frac{g_{m1}}{C_T^3}\left[C_{C1}(C_2+C_{C2})s^2 + g_{m2}C_{C2}s - g_{m2}g_{m3}\right]}{D_{CL}} \tag{4.94}$$

where

$$D_{CL} = s^3 + \left[g_{m3}C_{C2}(C_1+C_{C1}) - g_{m2}C_{C1}C_{C2} - fg_{m1}C_{C1}(C_2+C_{C2})\right]\frac{s^2}{C_T^3} \tag{4.95}$$

$$+ (g_{m2}g_{m3}C_{C1} - fg_{m1}g_{m2}C_{C2})\frac{s}{C_T^3} + \frac{fg_{m1}g_{m2}g_{m3}}{C_T^3}$$

and

$$C_T^3 = C_3(C_1+C_{C1})(C_2+C_{C2}) + C_1C_{C1}(C_2+C_{C2}) + C_2C_{C2}(C_1+C_{C1}) \tag{4.96}$$

The zeros are given as follows.

$$z = \frac{-g_{m2}C_{C2}}{2C_{C1}(C_2+C_{C2})} \pm \sqrt{\left(\frac{g_{m2}C_{C2}}{2C_{C1}(C_2+C_{C2})}\right)^2 + \frac{g_{m2}g_{m3}}{C_{C1}(C_2+C_{C2})}} \tag{4.97}$$

Now, the characteristic polynomial in (4.95) is forced to be equal to the following characteristic polynomial corresponding to one real pole and two complex poles.

$$D_{CL} = (s+\omega_{CL})(s^2+2\zeta\omega_n s+\omega_n^2) \tag{4.98}$$

By equating equations (4.95) and (4.98), the following relations are obtained.

$$\omega_{CL}+2\zeta\omega_n\omega_{CL} = \frac{g_{m3}C_{C2}(C_1+C_{C1}) - g_{m2}C_{C1}C_{C2} - fg_{m1}C_{C1}(C_2+C_{C2})}{C_T^3} \tag{4.99}$$

$$2\zeta\omega_n\omega_{CL}+\omega_n^2 = \frac{g_{m2}g_{m3}C_{C1}}{C_T^3} - \frac{fg_{m1}g_{m2}C_{C2}}{C_T^3} \tag{4.100}$$

$$\omega_n^2\omega_{CL} = \frac{fg_{m1}g_{m2}g_{m3}}{C_T^3} \tag{4.101}$$

Now, as before, for each stage, the transconductance is assumed to be proportional to a corresponding input capacitance for that stage as shown below.

$$g_{m1} = \omega_{T1}C_{G1} \tag{4.102}$$

$$g_{m2} = \omega_{T2}C_{G2} \tag{4.103}$$

$$g_{m3} = \omega_{T3}C_{G3} \tag{4.104}$$

The parasitic capacitance C_1 is the sum of the input capacitance of the second stage and the output capacitance of the first stage. Similarly, C_2 is the sum of the input capacitance of the third stage and the output capacitance of the second stage. C_3 is the sum of the output capacitance of the third stage, the feedback load, and the external load. These relations are expressed mathematically below.

$$C_1 = C_{G2}+\alpha_1 C_{G1} \tag{4.105}$$

$$C_2 = C_{G3}+\alpha_2 C_{G2} \tag{4.106}$$

$$C_3 = C_{LE} + C_F(1-f) + \alpha_3 C_{G3} \tag{4.107}$$

The following ratios are introduced to describe the relative sizes of the amplifier stages.

$$r_{C2} = \frac{C_{G2}}{C_{G1}} \tag{4.108}$$

$$r_{C3} = \frac{C_{G3}}{C_{G2}} \tag{4.109}$$

The analysis is not carried further by hand here. However, a computer was used to obtain a minimum power solution that meets a speed specification. The results are graphed in figures 4.48, 4.49, 4.50, 4.51, 4.52, and 4.53. It was found that the optimum design corresponded approximately to $\omega_{CL}=1.1\omega_n$ and $\zeta=0.7$. However, to simplify the analysis $\omega_{CL}=\omega_n$ and $\zeta=1$ was assumed here.

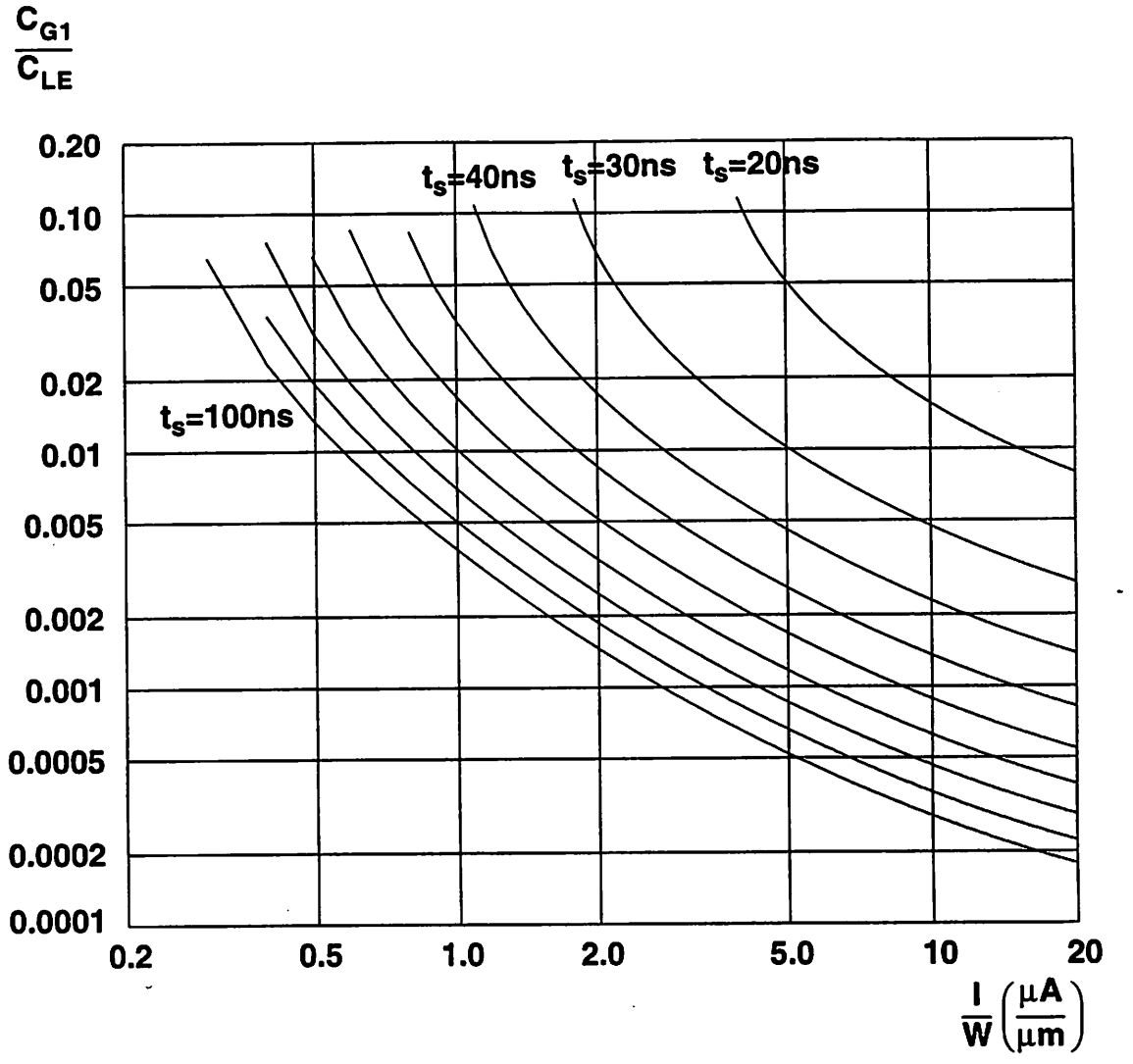


Figure 4.48 Input Capacitance of an Optimized Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=1$, $C_F/C_{LE}=0.5$, G=1

Figure 4.48 shows that the optimum input capacitance decreases as the current density increases. This effect is present in all of the amplifier architectures discussed in this document. This effect is due to the fact that increasing the current density reduces the transistor size required to achieve a specific transconductance.

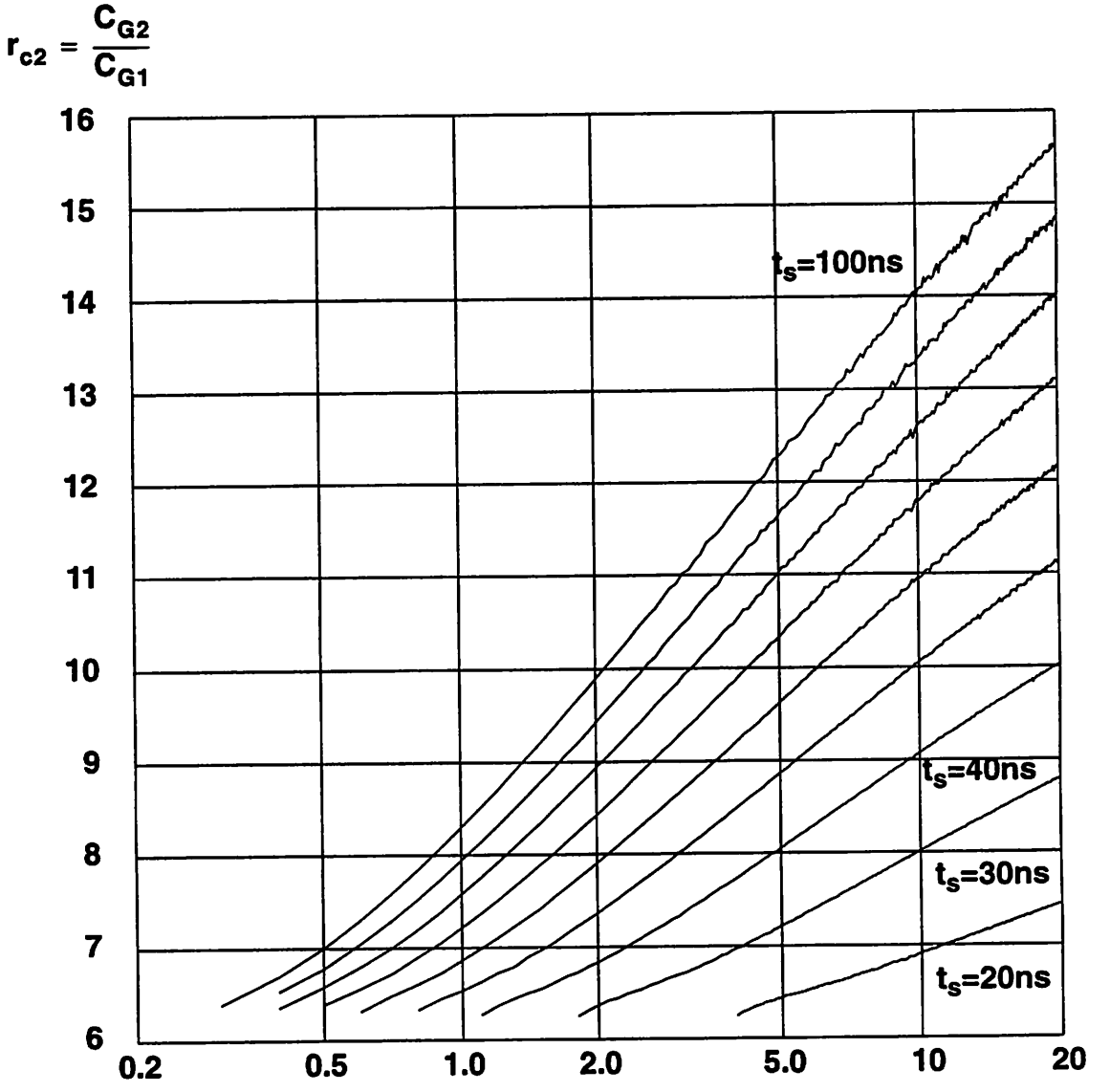


Figure 4.49 Gate Capacitance of Stage 2 versus Current Density for a Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

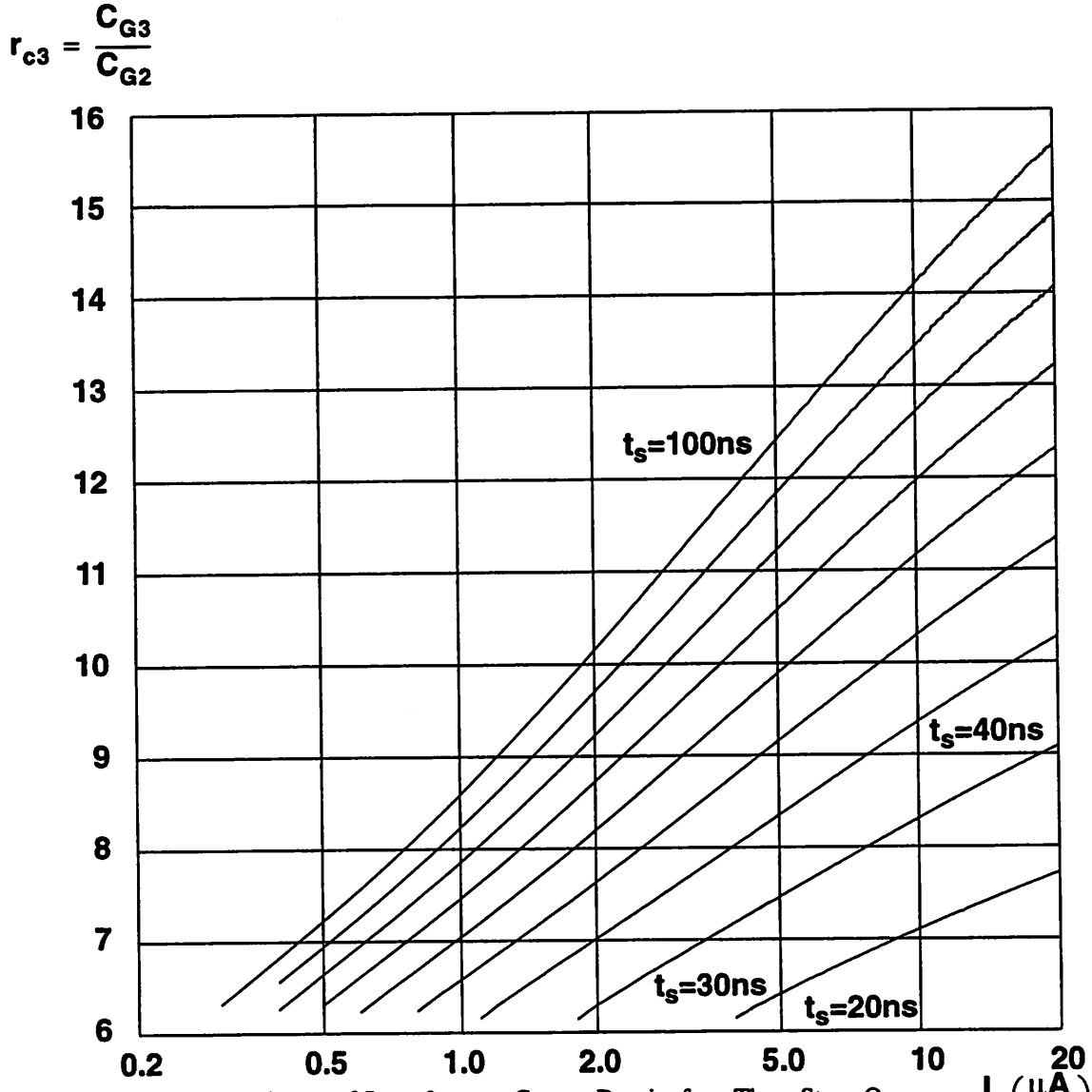


Figure 4.50 Gate Capacitance of Stage 3 versus Current Density for a Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_P/C_{LE}=0.5$, G=1

Figures 4.49 and 4.50 show that the optimum values of r_{c2} and r_{c3} increase with current density. In other words, it is optimal to make the last stage of the amplifier much larger than the first stage when current density is high. These figures also show that the optimal values for r_{c2} and r_{c3} increase as the required settling time is relaxed. Both of these effects may be understood in the following way. As current density is increased or settling time is increased, a capacitor of a certain size can be driven at a certain speed using a smaller transistor. In a three stage amplifier, the load

capacitance to be driven is the input capacitance of the following stage of the amplifier. Therefore, if the speed requirement is relaxed, the input stage can be reduced in size relative to the size of the second stage.

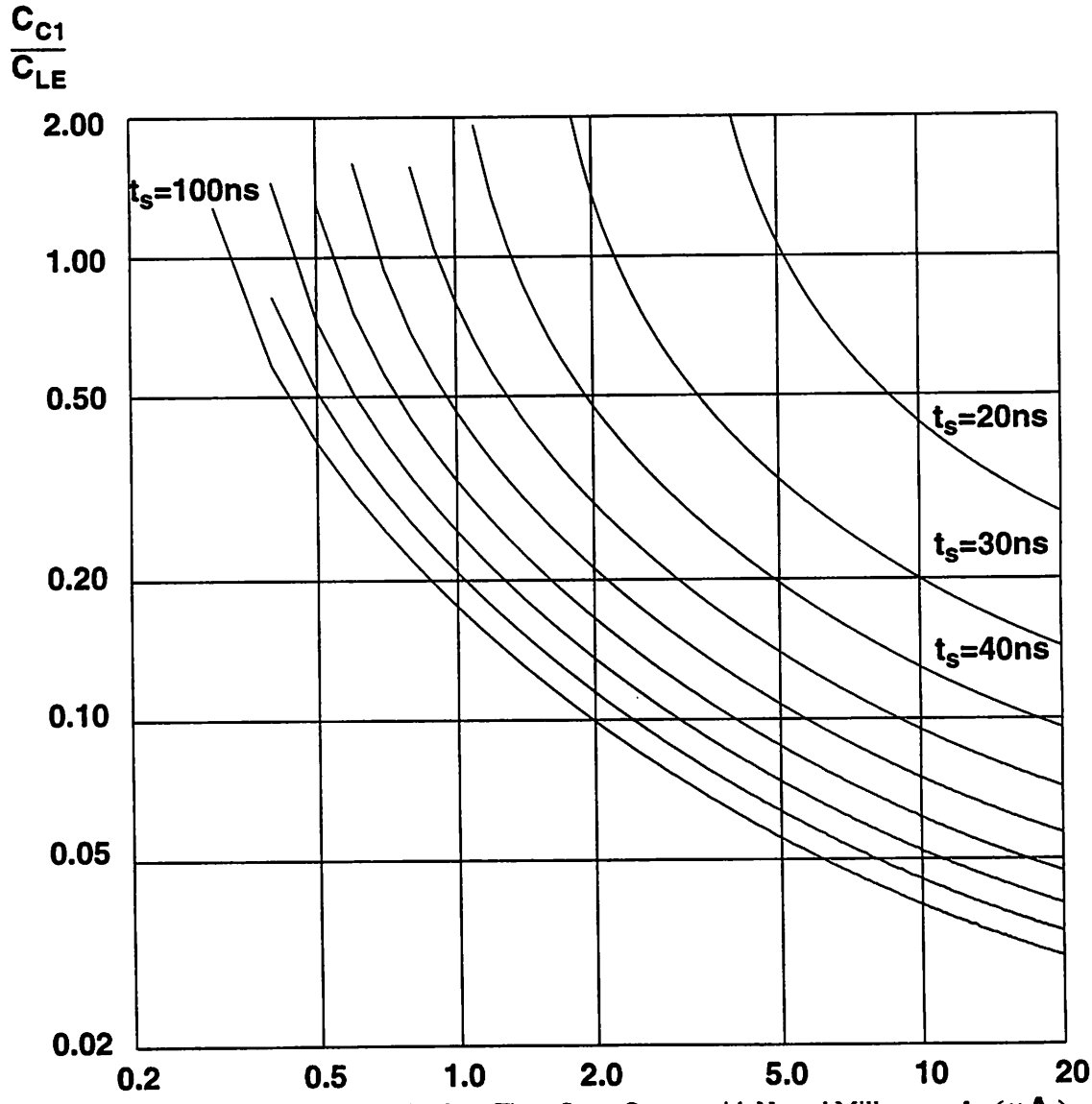


Figure 4.51 C_{C1} versus Current Density for a Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

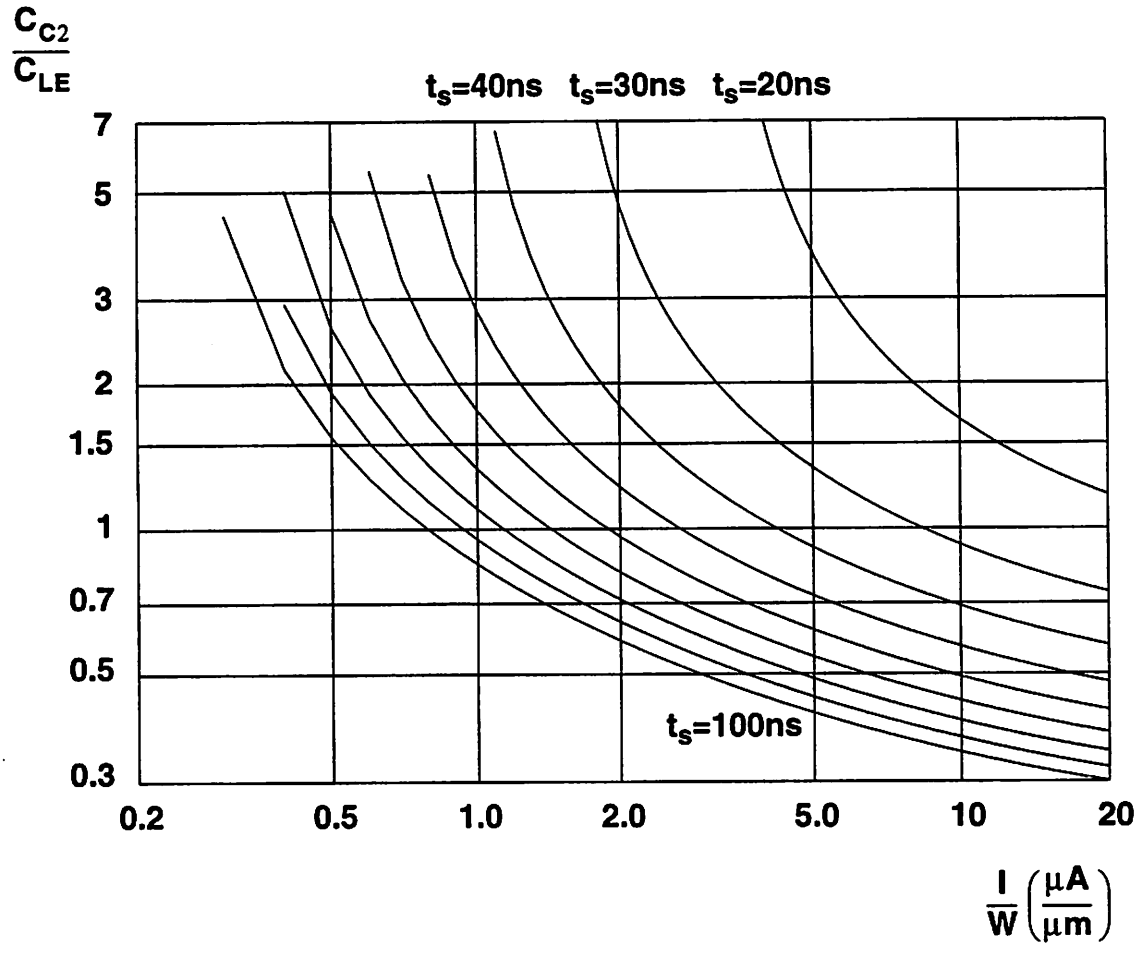


Figure 4.52 C_{C2} versus Current Density for a Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$, $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

Figures 4.51 and 4.52 show that the optimal sizes of the compensation capacitors decrease as current density increases and as the settling time requirement is relaxed. This effect is due to the fact that Miller compensation works by applying feedback around some part of the amplifier. Since both increasing current density and increasing settling time allow the use of smaller transistors, the same amount of feedback can be obtained with a smaller compensation capacitor.

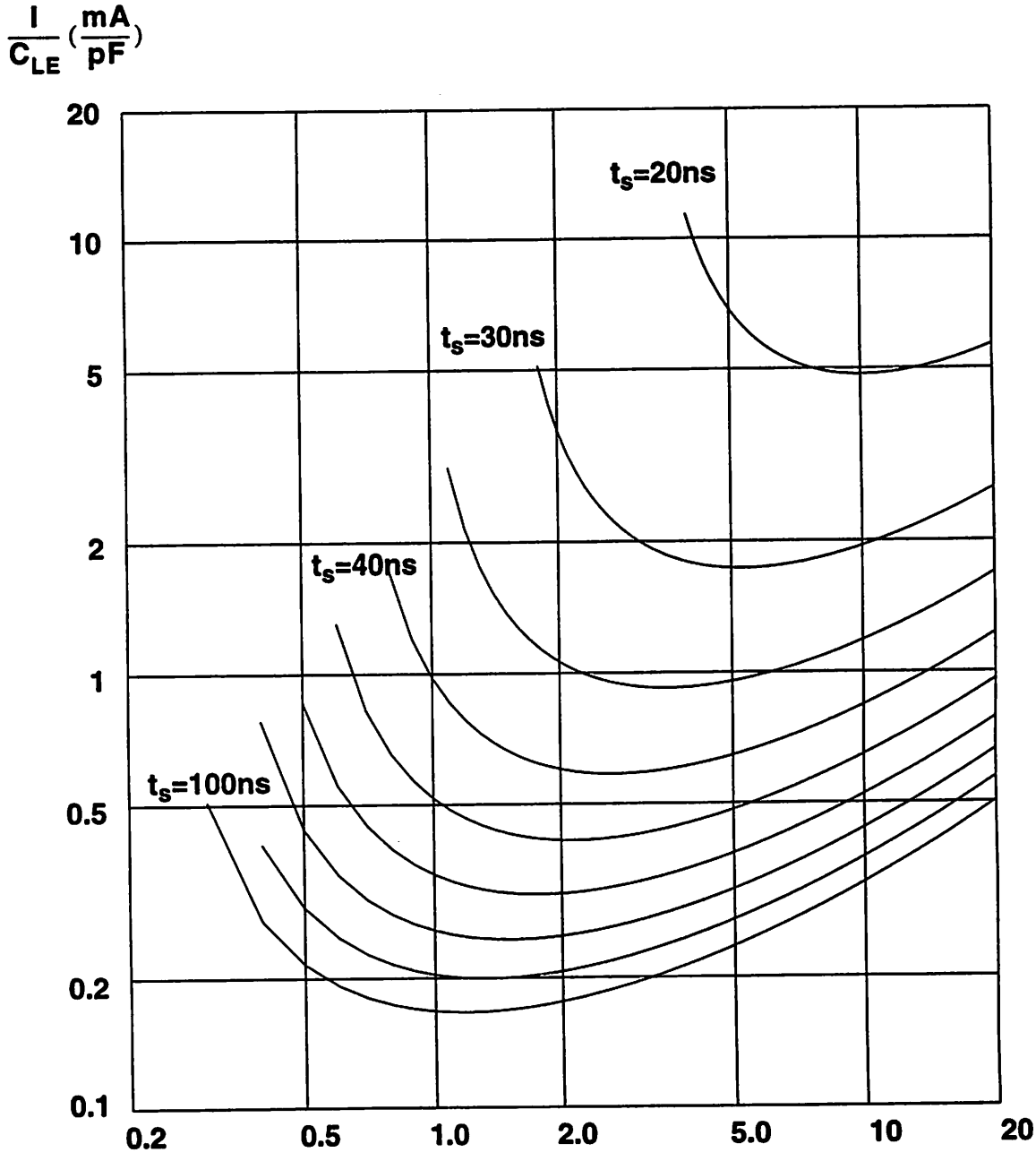


Figure 4.53 Total Bias Current versus Current Density for a Three Stage Opamp with Nested Miller Compensation, $\omega_{CL}=\omega_n$. $\zeta=1$, $\alpha_1=\alpha_2=\alpha_3=\alpha_c=1$, $C_F/C_{LE}=0.5$, G=1

4.2.7 Comparison of Topologies

In this section, a brief summary comparing the previously discussed amplifier topologies is presented. The graph in figure 4.54 compares the power dissipation of the different architectures at

a variety of speeds. The assumptions used to generate this graph are summarized in tables 4.2 and 4.3. The settling error ε is defined the same way here as it is in chapter 3.

Table 4.2 Assumptions used in the Opamp Comparison

$C_S = C_F = 0.5C_{LE}$
$L = 1.2\mu m$
settling error ε=1/1024

Table 4.3 Assumptions Used in the Opamp Comparison

single pole amplifier	Cascode amplifier	folded Cascode amplifier	preamp driving a 1 stage amplifier	2 stage amplifier with standard Miller compensation	2 stage amplifier with Ahuja compensation	3 stage amplifier with nested Miller compensation
$\alpha = 1$	$\alpha_1 = 1$ $\alpha_2 = 1$ $\alpha_3 = 1$	$\alpha_1 = 1$ $\alpha_2 = 1$	$\alpha_1 = 1$ $\alpha_L = 1$ $\alpha_2 = 1$	$\alpha_1 = 1$ $\alpha_2 = 1$	$\alpha_1 = 1$ $\alpha_C = 1$ $\alpha_2 = 1$ $\alpha_3 = 1$	$\alpha_1 = 1$ $\alpha_2 = 1$ $\alpha_3 = 1$
$\mu = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 524$ $\mu_2 = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 160.5$ $\mu_2 = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 524$ $\mu_L = 524$ $\mu_2 = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 524$ $\mu_2 = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 524$ $\mu_2 = 524$ $\mu_3 = 524$ $\frac{cm^2}{Vs}$	$\mu_1 = 524$ $\mu_2 = 524$ $\mu_3 = 524$ $\frac{cm^2}{Vs}$

Table 4.3 Assumptions Used in the Opamp Comparison

$E_{sat} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 1.5$ $E_{sat2} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 5.19$ $E_{sat2} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 1.5$ $E_{satL} = 1.5$ $E_{sat2} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 1.5$ $E_{sat2} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 1.5$ $E_{sat2} = 1.5$ $E_{sat3} = 1.5$ $\frac{V}{\mu m}$	$E_{sat1} = 1.5$ $E_{sat2} = 1.5$ $E_{sat3} = 1.5$ $\frac{V}{\mu m}$
			$\zeta = 1$	$\zeta = 1$	$\zeta = 1$	$\zeta = 1$
		$\frac{I_2}{W_2} = 2\frac{I}{W}$	$\frac{I_2}{W_2} = \frac{I_1}{W_1}$	$\frac{I_2}{W_2} = \frac{I_1}{W_1}$	$\frac{I_3}{W_3} = \frac{I_1}{W_1}$	$\frac{I_2}{W_2} = \frac{I_1}{W_1}$ $\frac{I_3}{W_3} = \frac{I_1}{W_1}$

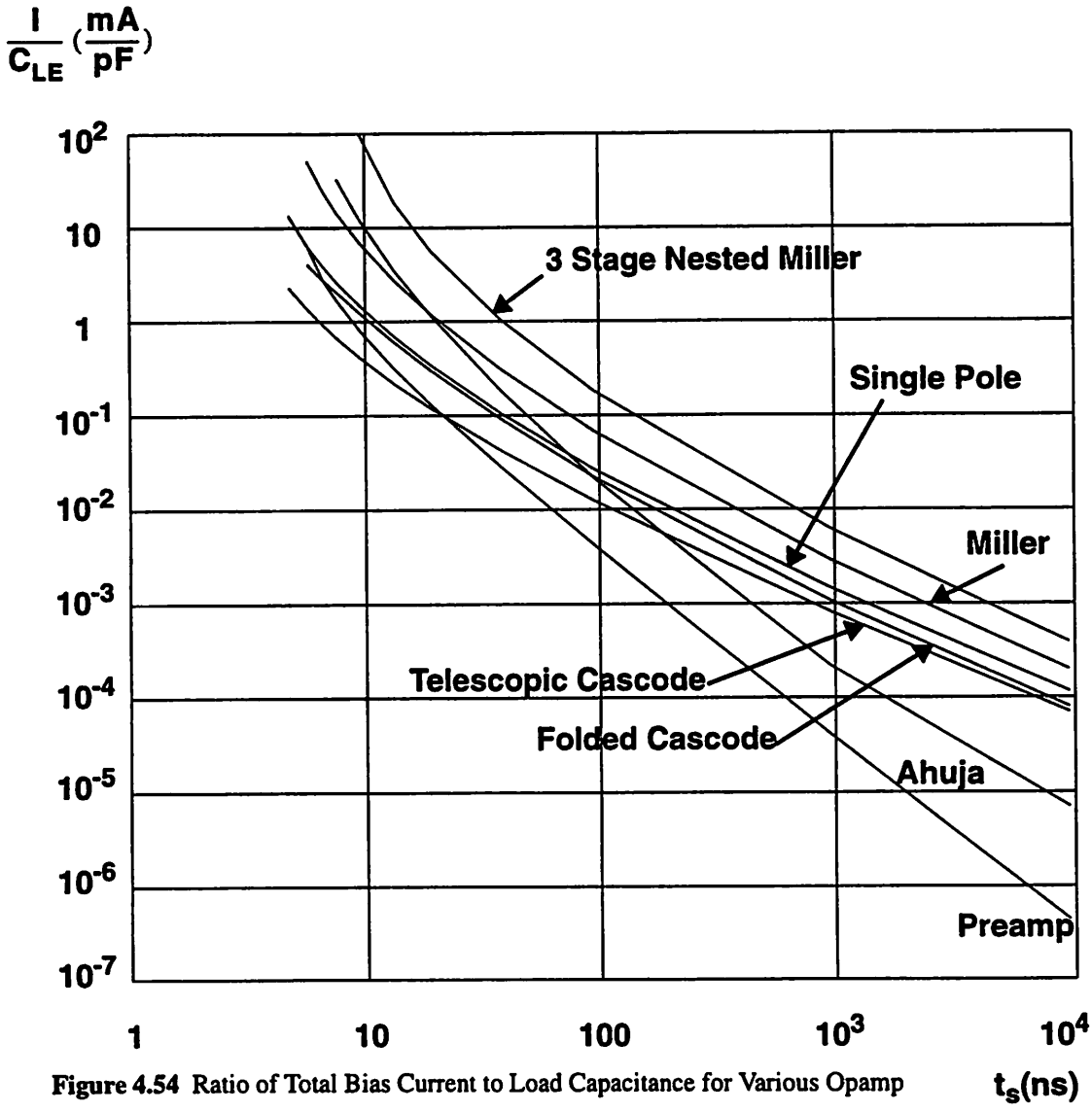


Figure 4.54 Ratio of Total Bias Current to Load Capacitance for Various Opamp Architectures

Note that at all speeds the telescopic cascode amplifier dissipates less power than the single transistor amplifier. In a telescopic cascode amplifier, the output transistor (cascode) can be sized smaller than the input transistor. This allows the output capacitance to be smaller for the telescopic cascode amplifier than it is for the single transistor amplifier.

. From the graph, it can be seen that the power of the preamplifier driving a single stage amplifier is inversely proportional to the square of the time constant. This is true because the transcon-

ductance of the second stage of the amplifier is boosted by the preamplifier. This fact makes the preamplifier driving a single stage amplifier to be the most efficient of the topologies discussed at low speeds.

This idea can be extended to multiple preamplifiers driving a single stage amplifier. If the amplifier is designed such that the closed loop configuration has all the poles at the same frequency, then the required power dissipation P at low speeds is given by the following formula.

$$P = \frac{KC_{LE}}{\tau^{n-1}} \tag{4.110}$$

In the above formula, K is a constant, and n is the number of preamp stages driving the high gain stage.

It should be noted that in this analysis, slew rate limitations are neglected. In most practical cases at low speed, slew rate requirements would prevent the power dissipation from dropping as the square of the settling time. A linear decrease is expected in the slew rate limited case. However, if the voltage swing can be reduced, slewing effects can be avoided.

The graph in figure 4.54 compares the power dissipations of the different architectures for a fixed closed loop gain of 2. It is also interesting to compare the speed performance of the different amplifier architectures as the closed loop gain specification is varied. In this graph, it is assumed that in the multipole amplifiers the poles are all at the same frequency. When interpreting the results of this graph, it is important to keep in mind that for the same time constant the multipole amplifiers take somewhat longer to settle. However, this graph indicates some interesting trends. The most obvious trend is that the multistage amplifiers perform much better than the single stage amplifier for high closed loop gain. This performance improvement is possible because of the buffering action provided by the extra stages. However, the high performance of the telescopic cascode amplifier across the range shows that many of the benefits of the multistage amplifiers can be obtained by simply adding cascode transistors to the single stage amplifier.

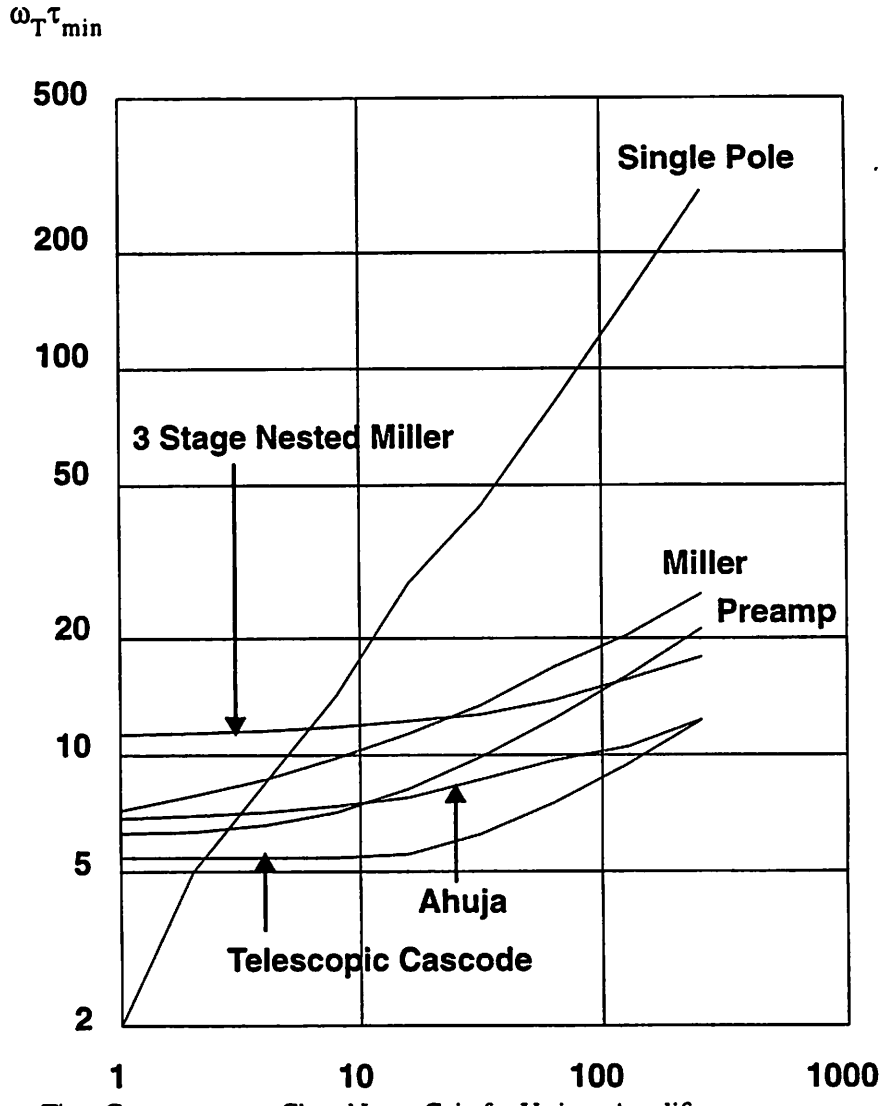


Figure 4.55 Minimum Time Constant versus Closed Loop Gain for Various Amplifier Architectures

APPENDIX

In this appendix, a number of amplifier optimizations are performed with a variety of assumptions. These different optimizations are included to address different applications. For example, in some applications the speed requirement and sampling capacitance might be fixed with the goal of minimizing power dissipation. A low noise application where the required sampling capacitance is determined by thermal noise is an example of such a case. In some applications, thermal noise may

not be an important issue. In these cases, the designer might be free to vary the size of the sampling capacitor. In these cases, a different optimization procedure is appropriate. The optimizations included here are summarized in table 4.4.

Table 4.4 Summary of Switched Capacitor Amplifier Analysis

Section	Architecture	Fixed Parameters	Modeling Assump.	Independent Variables	Optimized Variables	Example Application
4.A.1.1	Single - Stage Single Pole	C_{LE} $G=C_S/C_F$ I/W ω_T speed (τ)		C_F	C_{G1} I	power minimizing switched capacitor gain stage with a fixed speed but no noise constraint
4.A.1.2	Single - Stage Single Pole	C_{LE} $G=C_S/C_F$ I/W ω_T C_F		C_{G1}	speed (τ)	speed maximizing switched capacitor gain stage with a fixed noise spec.

Table 4.4 Summary of Switched Capacitor Amplifier Analysis

Section	Architecture	Fixed Parameters	Modeling Assump.	Independent Variables	Optimized Variables	Example Application
4.A.1.3	Single - Stage Single Pole	C_{LE} $G=C_S/C_F$ C_F	long channel	C_{G1}	$I\tau^2$	optimizing speed and power of switched capacitor gain stages with fixed noise constraints
4.A.1.3.1	Single - Stage Single Pole	C_{LE} $G=C_S/C_F$	long channel	C_F	$I\tau^2$	
4.A.1.3.2	Single - Stage Single Pole	C_{LE} $G=C_S/C_F$	long channel $C_{OP}=0$	C_{G1} C_F	$I\tau^2$	
4.A.2.1	Two Pole Telescopic Cascode	C_{LE} $G=C_S/C_F$ $r=C_{G2}/C_{G1}$ I/W ω_T speed (τ)		C_F	C_{G1} I	high speed moderate accuracy switched capacitor gain stages
4.A.2.1.1	Two Pole Telescopic Cascode	$G=C_S/C_F$ $r=C_{G2}/C_{G1}$	long channel	C_F	$\frac{2\mu}{L^2}\left(\frac{I\tau^2}{C_{LE}}\right)$	

4.A.1 Single Stage Amplifier Optimizations

This section deals with single stage single pole amplifiers. These amplifiers are conceptually

the simplest, and these also are the most easily analyzed.

4.A.1.1 Minimum Power with Fixed Speed

The goal of this analysis is to meet a speed requirement with minimum power while the feedback capacitance is allowed to vary. This type of analysis might be useful in a low resolution pipelined ADC where noise constraints are not an important issue. In the front end stages of a high resolution pipelined ADC, where noise constraints are important, freedom to vary the size of the feedback capacitance is limited. In this case the analysis presented above is applicable.

A trade-off exists in choosing the feedback capacitance because as the feedback capacitance is increased the load capacitance and the feedback factor both tend to increase. Increased load capacitance tends to slow things down, but increased feedback factor tends to speed things up.

Because the speed is fixed, the time constant τ is fixed, and because the current density is fixed, the current and thus the power is proportional to the input capacitance C_{G1}. Therefore, minimizing the input capacitance minimizes the power. To minimize the power, equation (4.36) is differentiated with respect to C_F, and the derivative of C_{G1} with respect to C_F is set to zero as shown below.

$$\frac{dC_{G1}}{dC_F} = 0 \tag{4.111}$$

The above operation results in the following relation between C_{G1} and C_F.

$$C_{G1} = \frac{(G+1)C_{LE} + 2GC_F}{\omega_T\tau - \alpha(G+1) - 1} \tag{4.112}$$

When equations (4.36) and (4.112) are combined, the following relations are obtained.

$$\frac{C_F}{C_{LE}} = \frac{\omega_T\tau + \alpha(G+1) - 1 + [\omega_T\tau - \alpha(G+1) - 1]\sqrt{\frac{\omega_T\tau(G+1) - 1}{G}}}{[\omega_T\tau - \alpha(G+1) - 1]^2 - 4\alpha G} \tag{4.113}$$

$$\frac{C_{G1}}{C_{LE}} = \frac{[\sqrt{\omega_T\tau(G+1) - 1} + \sqrt{G}]^2 - \alpha(G+1)^2}{[(\omega_T\tau - \alpha(G+1) - 1)]^2 - 4\alpha G} \tag{4.114}$$

From equation (4.114) it is possible to determine the maximum possible speed (minimum time constant τ_{min}) by setting the denominator to zero. The result is shown below.

$$\omega_T \tau_{min} = \alpha (G+1) + \sqrt{4\alpha G} + 1 \tag{4.115}$$

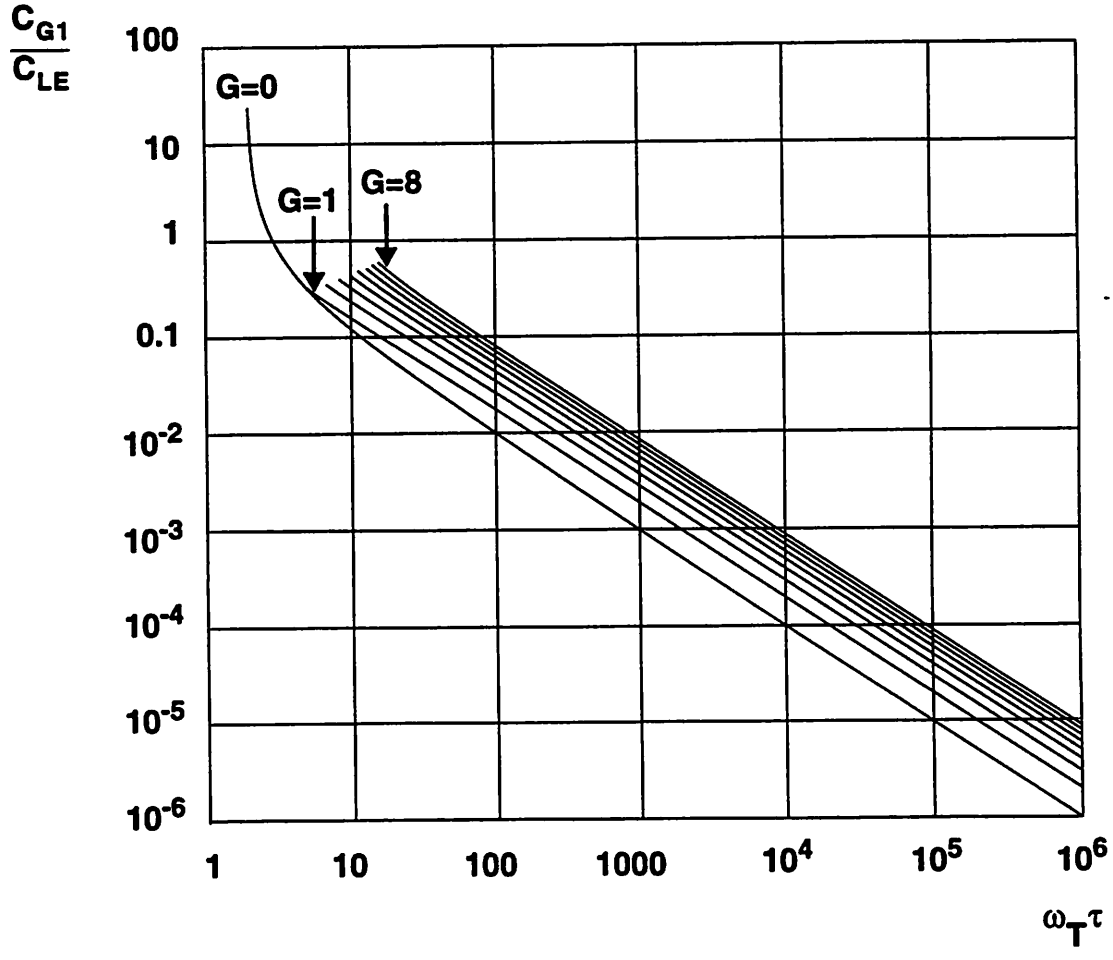


Figure 4.56 Input Capacitance Versus Normalized Time Constant for Single Pole Amplifier with Optimized Feedback Capacitance, α=1

Figure 4.56 shows a plot of the optimized input capacitance versus speed as given by equation (4.114). A plot of the optimum feedback capacitance is shown in figure 4.57. For time constants below the minimum time constant τ_{min}, no solution is possible. However, it is interesting to note that for G>0 the power does not approach infinity as the limit τ_{min} is approached.

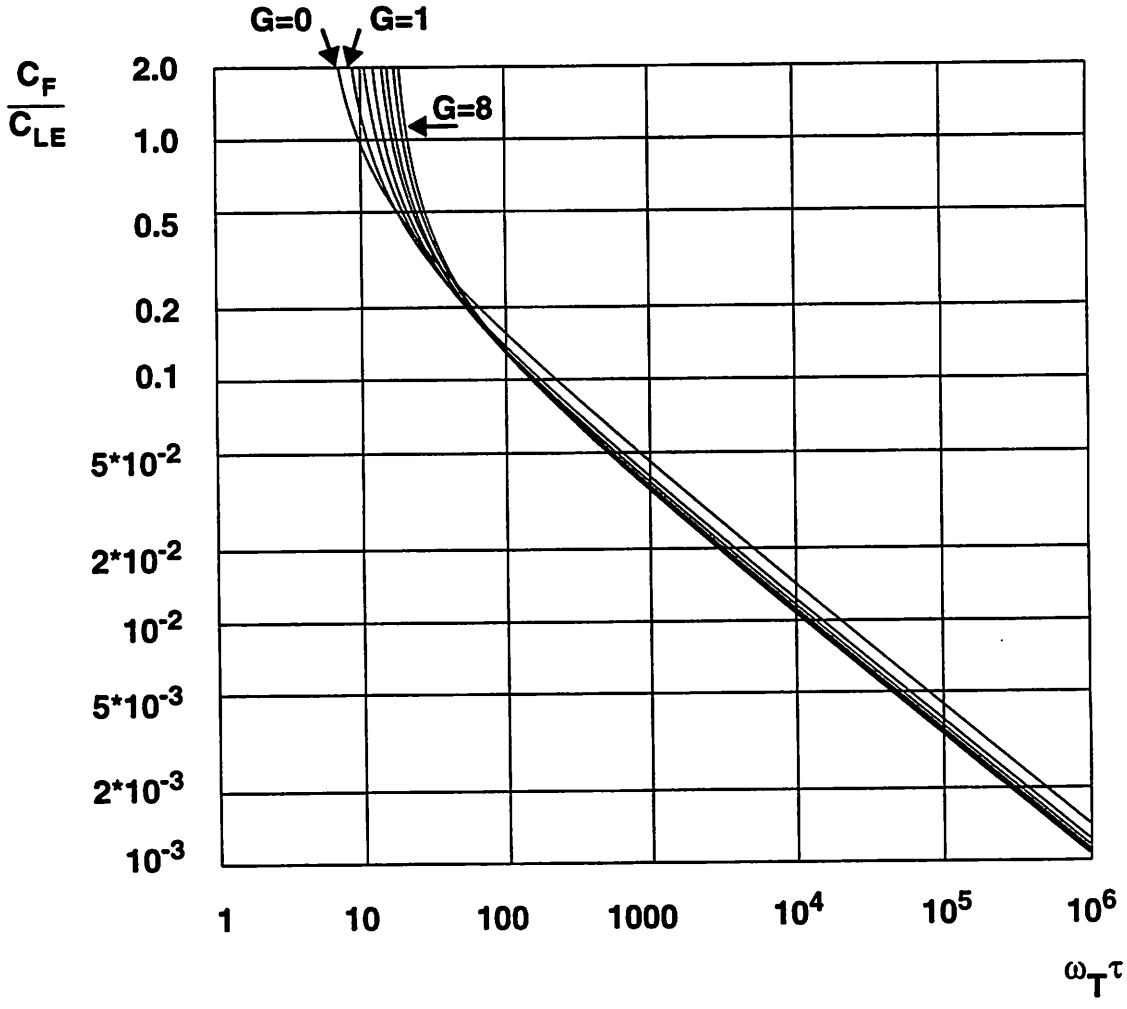


Figure 4.57 Optimum Feedback Capacitance to Minimize Power versus Normalized Time Constant for Single Pole Amplifier, α=1

4.A.1.2 Maximum Speed with Fixed Feedback Capacitance

In this section, the feedback capacitance is fixed, and the input capacitance is adjusted to maximize the speed. The analysis for this situation is discussed extensively by Conroy[Conroy94]. To accomplish this goal, equation (4.36) is differentiated with respect to C_{GI}, and the derivative of the time constant with respect to C_{GI} is set to zero. The result is shown below and graphed in figures 4.58 and 4.59.

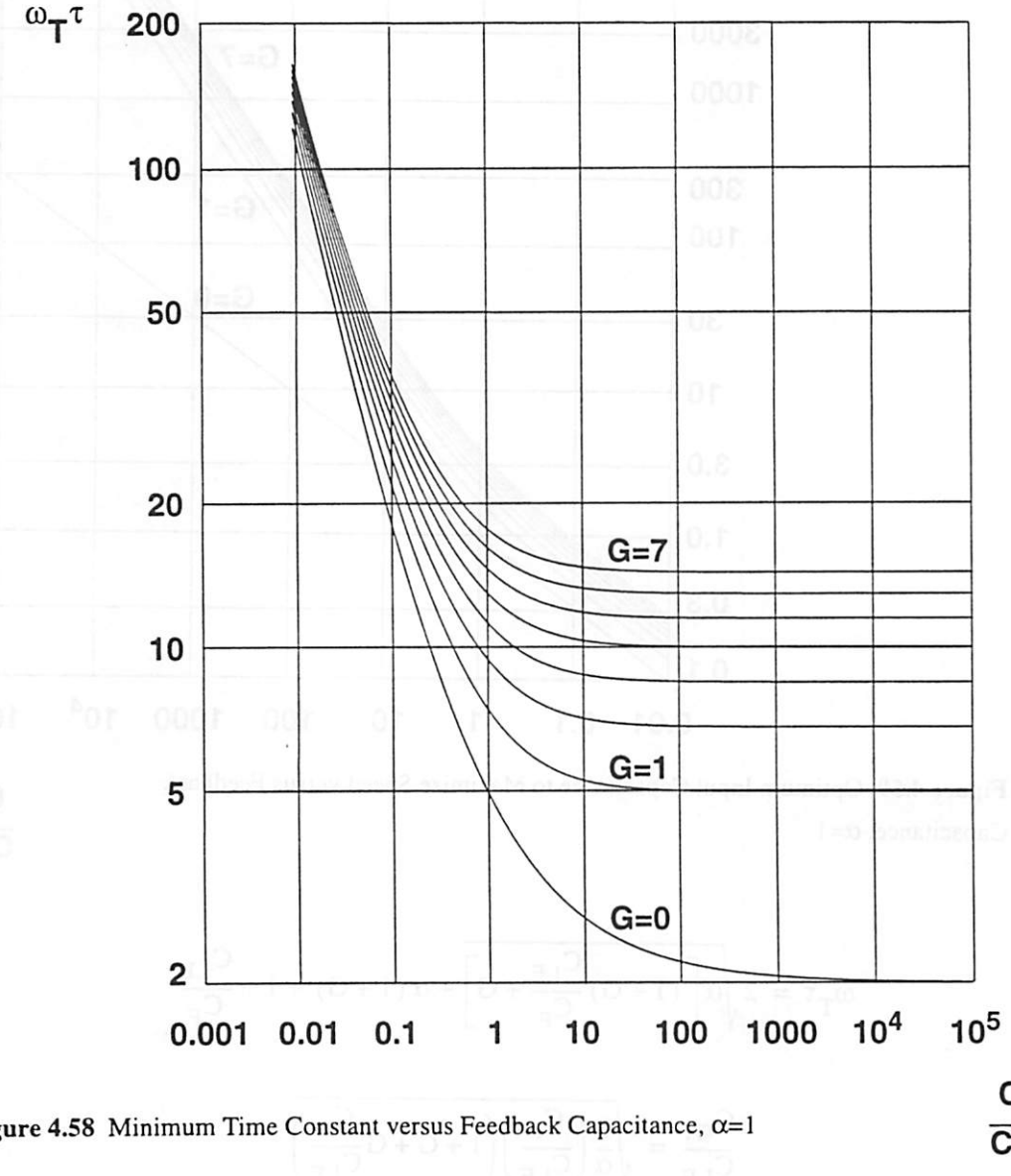


Figure 4.58 Minimum Time Constant versus Feedback Capacitance, α=1

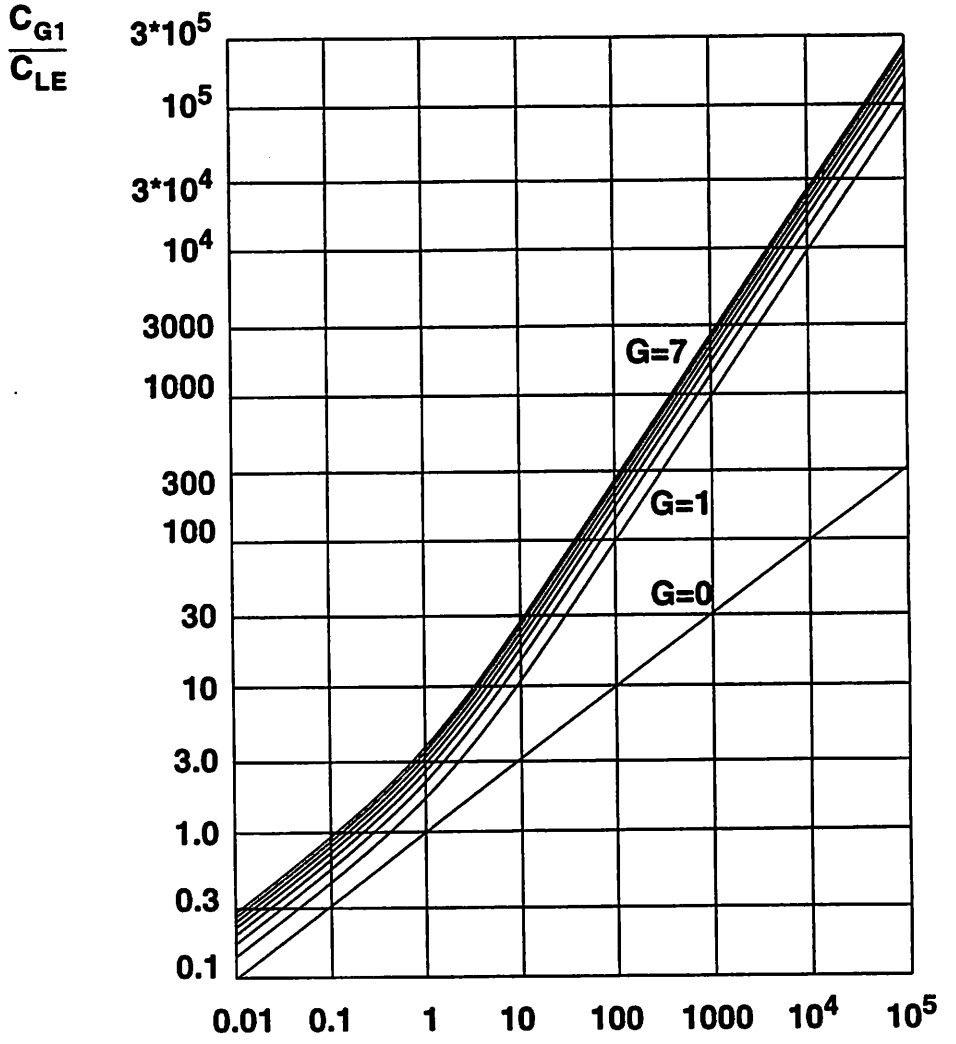


Figure 4.59 Optimum Input Capacitance to Maximize Speed versus Feedback Capacitance, α=1

$$\omega_T\tau = 2\sqrt{\alpha\left[(1+G)\frac{C_{LE}}{C_F}+G\right]}+\alpha(1+G)+1+\frac{C_{LE}}{C_F} \tag{4.116}$$

$$\frac{C_{G1}}{C_{LE}} = \sqrt{\frac{1}{\alpha}\left(\frac{C_F}{C_{LE}}\right)\left(1+G+G\frac{C_F}{C_{LE}}\right)} \tag{4.117}$$

4.A.1.3 Speed and Power Optimization with Variable Current Density-Long Channel Model

In this analysis, the current density is allowed to vary. The long channel model given by equation (4.3) is used to model the transconductance. This analysis has the advantage of allowing the optimum current density to be determined, but its applicability is limited to cases where the long channel model is valid.

Nodal analysis of figure 4.8 results in the following formula for the time constant.

$$\tau = \frac{C_{LE} + C_F(1-f) + \alpha C_{G1}}{f g_m} = \frac{C_{LE} + C_F(1-f) + \alpha C_{G1}}{f\sqrt{2\mu C_{ox}\frac{W}{L} I}} \tag{4.118}$$

The width W is eliminated from the above equation by assuming that $C_{G1}=C_{ox}WL$. This substitution leads to the result shown below.

$$\tau = \frac{C_{LE} + C_F(1-f) + \alpha C_{G1}}{f\sqrt{\frac{2\mu}{L^2} C_{G1} I}} \tag{4.119}$$

The formula for the feedback factor from equation (4.33) is then plugged into the above equation to obtain the following formula.

$$\tau\sqrt{\frac{I}{C_{LE}}} = \sqrt{\frac{L^2}{2\mu}}\left\{T\left(\frac{C_{G1}}{C_{LE}}\right)^{-1/2} + B\left(\frac{C_{G1}}{C_{LE}}\right)^{1/2} + \frac{\alpha C_{LE}}{C_F}\left(\frac{C_{G1}}{C_{LE}}\right)^{3/2}\right\} \tag{4.120}$$

In the above formula, B and T are given as follows.

$$B = \alpha(1+G) + 1 + \frac{C_{LE}}{C_F} \tag{4.121}$$

$$T = 1 + G + \frac{GC_F}{C_{LE}} \tag{4.122}$$

In order to determine an optimal design, equation (4.120) is differentiated with respect to C_{G1}, and the derivative of $\tau\sqrt{I}$ is set to zero. The result is shown below.

$$\frac{C_{G1opt}}{C_{LE}} = \frac{B}{6\alpha}\left(\frac{C_F}{C_{LE}}\right)\left(-1+\sqrt{1+\frac{12\alpha T}{B^2\left(\frac{C_F}{C_{LE}}\right)}}\right) \tag{4.123}$$

$$\left(\frac{I}{W}\right)_{opt} = \frac{C_{ox}L^3}{2\mu\tau^2}\left(\frac{4B^2}{9}\right)\left(2+\sqrt{1+\frac{12\alpha T}{B^2\left(\frac{C_F}{C_{LE}}\right)}}\right) \tag{4.124}$$

$$\frac{\tau^2 I}{C_{LE}} = \frac{2}{27}\left(\frac{L^2}{2\mu}\right)\frac{B^3}{\alpha}\left(\frac{C_F}{C_{LE}}\right)\left\{-1+\frac{36\alpha T}{B^2}\left(\frac{C_{LE}}{C_F}\right)+\left[1+\frac{12\alpha T}{B^2}\left(\frac{C_{LE}}{C_F}\right)\right]^{3/2}\right\} \tag{4.125}$$

Note that the optimal current density is proportional to the cube of the channel length. Figure 4.60 shows a plot of the optimum value of C_{G1} versus C_F. Figure 4.61 shows a graph of equation (4.125).

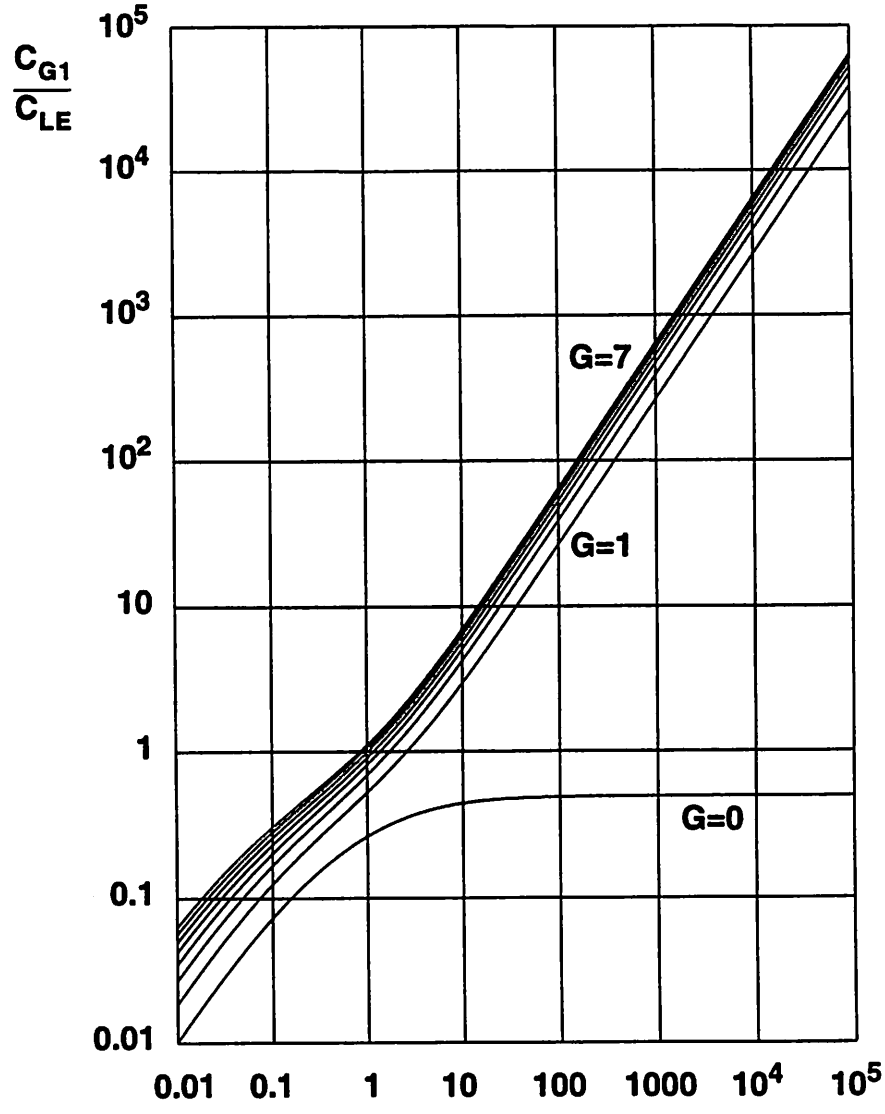


Figure 4.60 Optimal Input Capacitance to Minimize $I\tau^2$ Based on the Long Channel Model, α=1

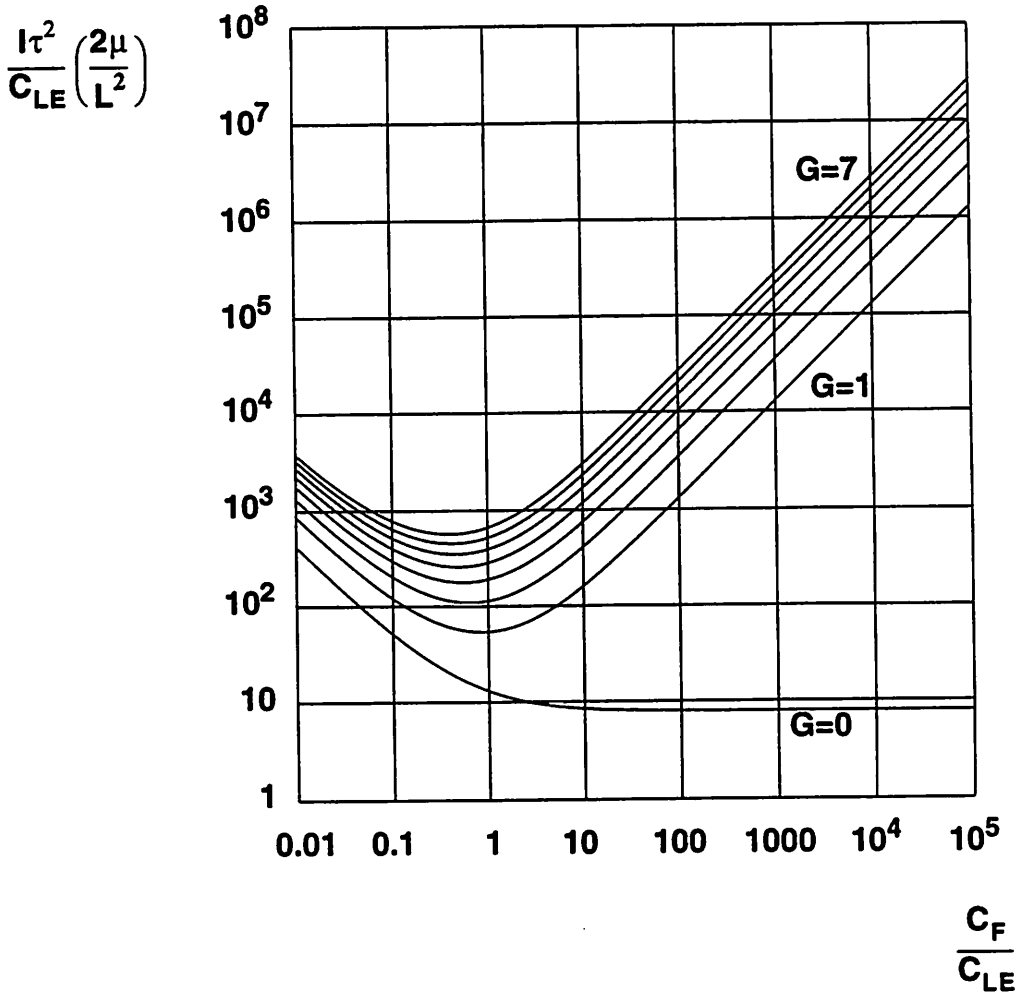


Figure 4.61 Minimum Current versus Feedback Capacitance, α=1

4.A.1.3.1 Speed and Power Optimization with Variable Feedback Capacitance

If the optimization on equation (4.125) is extended, an optimal value of feedback capacitance to minimize power or maximize speed can be determined. The resulting relations are shown below.

$$\frac{C_{Fopt}}{C_{LE}} = \sqrt{\frac{C_{G1}}{GC_{LE}}\left(\alpha\frac{C_{G1}}{C_{LE}}+1\right)} \tag{4.126}$$

$$\left(\frac{C_{G1opt}}{C_{LE}}\right)^3 + V\left(\frac{C_{G1opt}}{C_{LE}}\right)^2 + X\frac{C_{G1opt}}{C_{LE}} + Y = 0 \tag{4.127}$$

$$U = \alpha[\alpha(1+G)+1]^2 - 4\alpha^2 G \tag{4.128}$$

$$V = \frac{1-\alpha^2(1+G)^2}{U} \tag{4.129}$$

$$X = \frac{-(1+G)[2+\alpha(1+G)]}{U} \tag{4.130}$$

$$Y = \frac{(1+G)^2}{U} \tag{4.131}$$

The following variables are defined in order to help in solving the cubic equation given by equation (4.127).

$$C_A = \frac{(V^2-3X)}{3} \tag{4.132}$$

$$C_B = \frac{(2V^3-9VX+27Y)}{27} \tag{4.133}$$

$$M = \frac{C_A^3}{27} - \frac{C_B^2}{4} \tag{4.134}$$

$$\theta = \operatorname{atan}\left(\frac{2\sqrt{m}}{C_B}\right) \tag{4.135}$$

$$s_m = \left(\frac{C_B^2}{4}+M\right)^{1/6} \tag{4.136}$$

$$\frac{C_{G1}}{C_{LE}} = s_m\left[\cos\left(\frac{\theta}{3}\right) - \sqrt{3}\sin\left(\frac{\theta}{3}\right)\right] \tag{4.137}$$

Figure 4.62 shows a plot of the optimized input capacitance versus G. Figure 4.63 shows a plot of the optimized feedback capacitance versus G, and figure 4.64 shows a plot of the product of the current and the square of the time constant normalized to the technology.

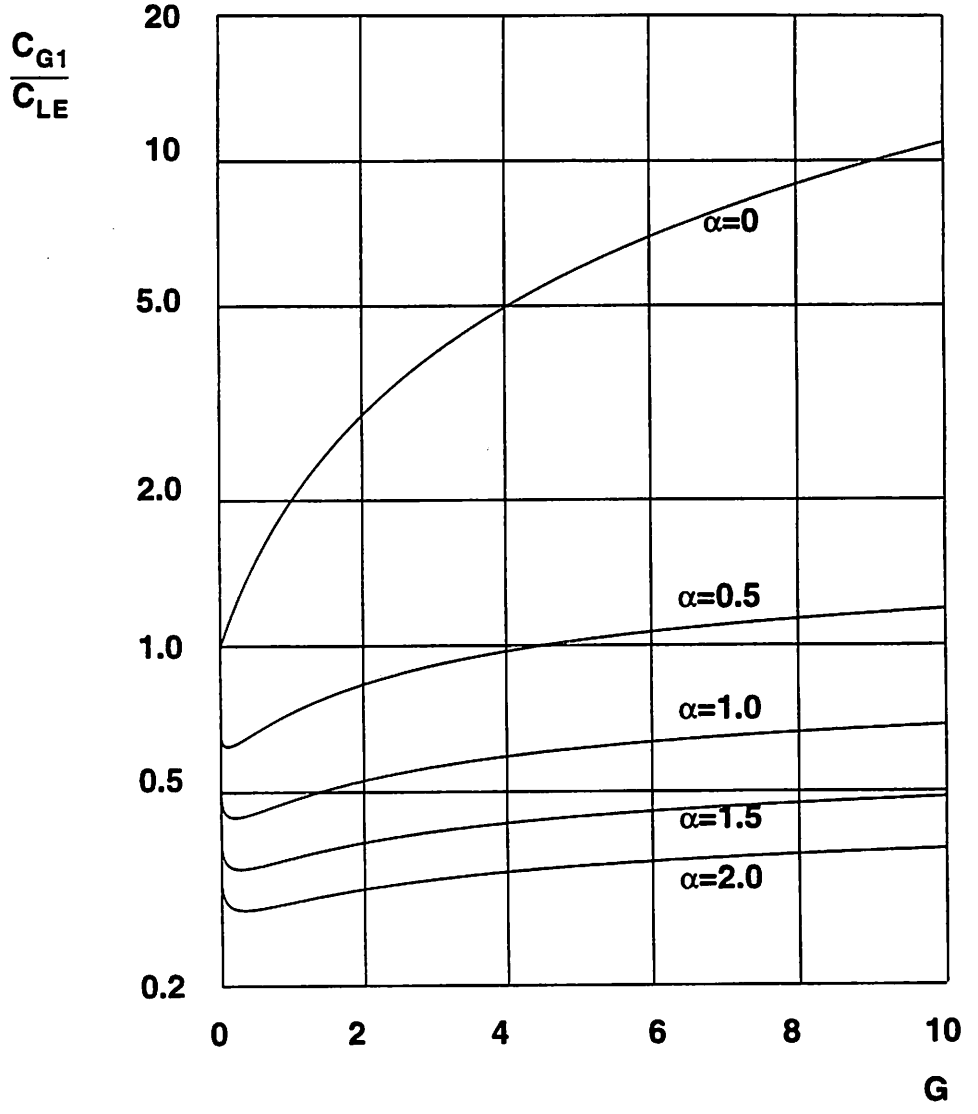


Figure 4.62 Optimized Input Capacitance versus G

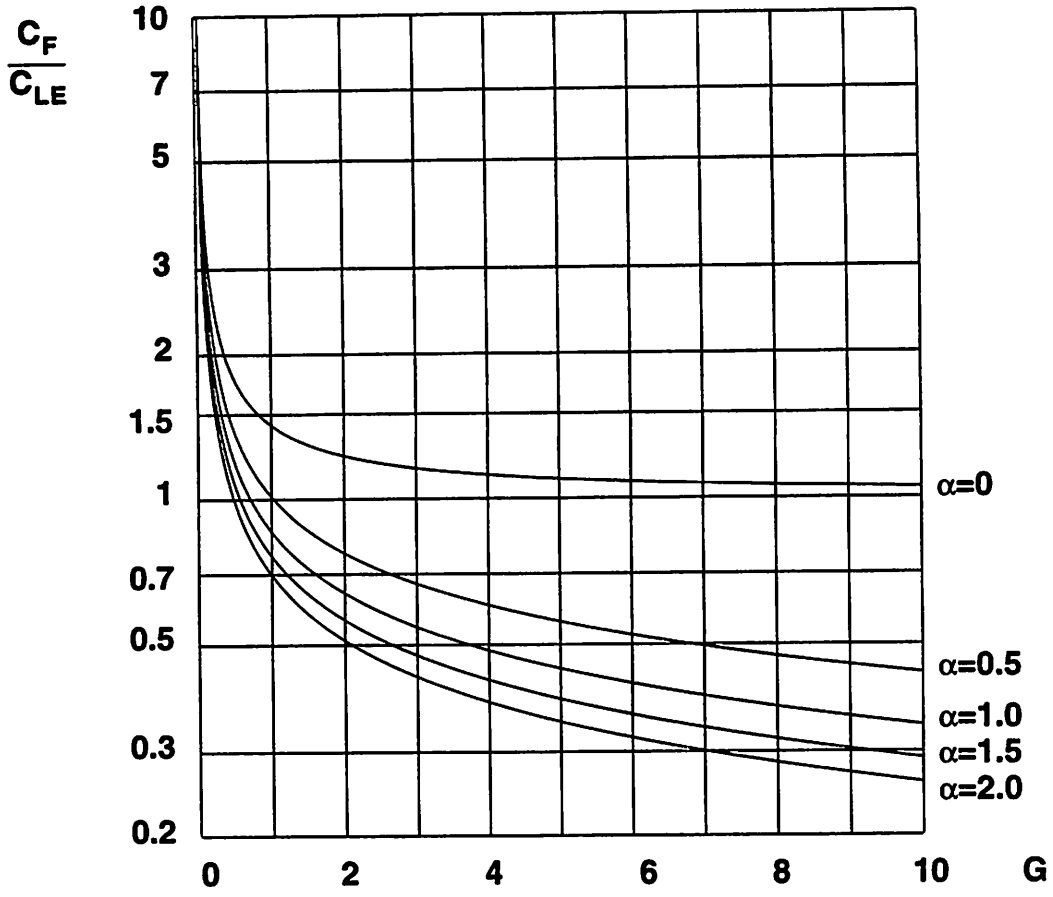


Figure 4.63 Optimal Feedback Capacitance versus G

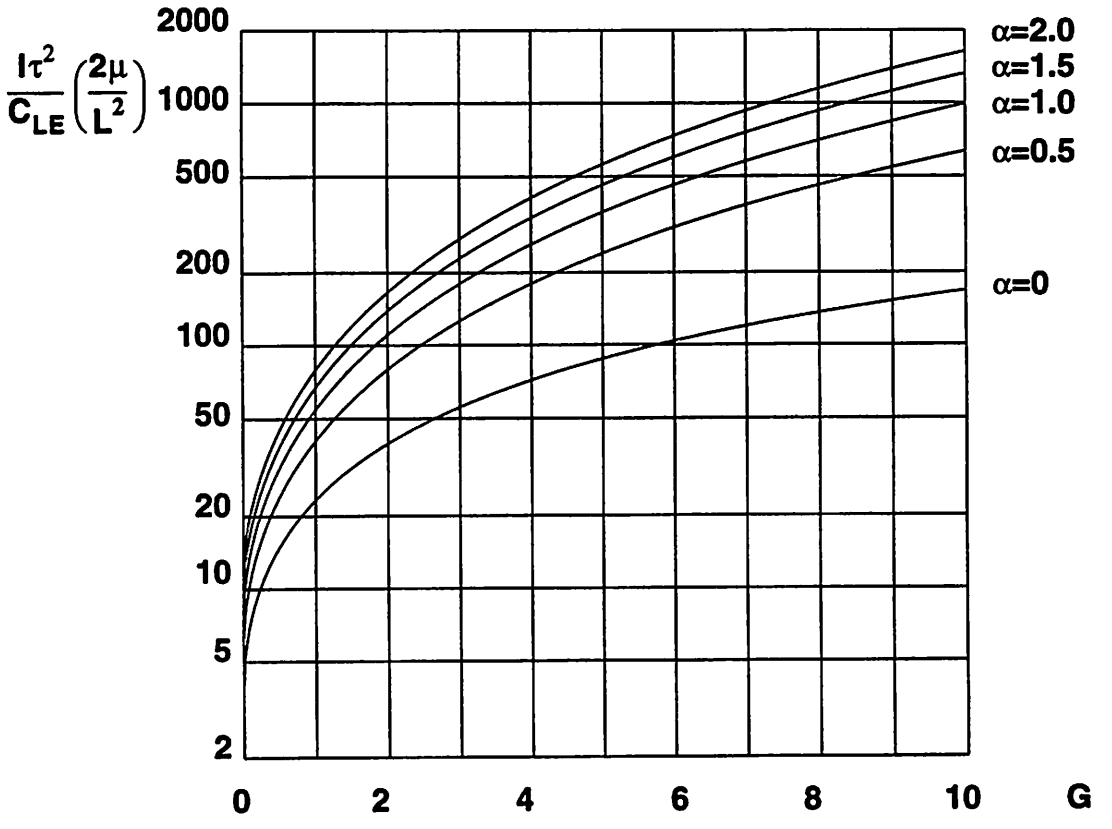


Figure 4.64 Minimum Current versus Gain

4.A.1.3.2 Special Case: No Output Parasitic Capacitance ($\alpha = 0$)

In order to provide some intuition to the above analysis, the case where the output parasitic capacitance is negligible is analyzed. Setting α to zero greatly simplifies the analysis. With $\alpha = 0$ equation (4.120) simplifies to the following formula.

$$\tau\sqrt{\frac{I}{C_{LE}}} = \sqrt{\frac{L^2}{2\mu}}\left[\left(1+G+\frac{GC_F}{C_{LE}}\right)\left(\frac{C_{G1}}{C_{LE}}\right)^{\frac{-1}{2}} + \left(1+\frac{C_{LE}}{C_F}\right)\left(\frac{C_{G1}}{C_{LE}}\right)^{1/2}\right] \tag{4.138}$$

When the input capacitance is optimized to minimize the time constant or the current density, the following relation is obtained.

$$\left(\frac{C_{G1}}{C_{LE}}\right)_{opt} = \frac{1+G+G\dfrac{C_F}{C_{LE}}}{1+\dfrac{C_{LE}}{C_F}} \tag{4.139}$$

The corresponding optimized time constant is given by the following formula.

$$\tau^2\frac{I}{C_{LE}} = 4\left(\frac{L^2}{2\mu}\right)\left(1+G+G\frac{C_F}{C_{LE}}\right)\left(1+\frac{C_{LE}}{C_F}\right) \tag{4.140}$$

As in the previous section, the optimization can be taken further by finding the feedback capacitance that minimizes the time constant. When this is done, the following relations are obtained.

$$\left(\frac{C_F}{C_{LE}}\right)_{opt} = \sqrt{1+\frac{1}{G}} \tag{4.141}$$

$$\left(\frac{C_{G1}}{C_{LE}}\right)_{opt} = 1+G \tag{4.142}$$

The corresponding optimized time constant is given by the following formula.

$$\tau^2\frac{I}{C_{LE}} = 4\left(\frac{L^2}{2\mu}\right)\left(\sqrt{1+G}+\sqrt{G}\right)^2 \tag{4.143}$$

Equation (4.142) states that the input capacitance of the amplifier should optimally be as large as, or in a high gain design much larger than, the load capacitance. However, the output capacitance has an important effect, and figure 4.62 shows that the optimum input capacitance may be much smaller than the load capacitance if the output capacitance is comparable to the input capacitance.

4.A.2 Telescopic Cascode Amplifier Optimizations

In this section, some optimizations of the telescopic cascode amplifier are treated. The telescopic cascode amplifier is not as simple as the single pole amplifier, but it is simple enough that some degree of closed form analysis is possible.

4.A.2.1 Optimization of Feedback Capacitance to Minimize Power - Fixed Current Density and Speed

Equations (4.56) through (4.65) may be combined to obtain the following equation for the time constant as a function of the feedback capacitance, the input capacitance, the feedback factor, the unity gain frequency, and r.

$$\tau = \frac{C_{LE} + C_F(1-f) + \alpha_3 r C_{G1}}{2\zeta^2 \omega_{T1} C_{G1} f} \tag{4.144}$$

Now it is assumed that the time constant and current densities are fixed parameters. This implies that r is also fixed. In that case, the power is minimized by minimizing the input capacitance C_{G1}. In this section, the optimal value of C_F to minimize the input capacitance is determined with the following results.

$$\frac{C_{G1}}{C_{LE}} = \frac{(1+G)E + 2G + \sqrt{(4G)\left[(1+G)^2\alpha_3 r + (1+G)E + G\right]}}{(E^2 - 4\alpha_3 r G)} \tag{4.145}$$

$$\frac{C_F}{C_{LE}} = \sqrt{\left(\frac{1}{G}\right)\frac{C_{G1}}{C_{LE}}\left(\alpha_3 r \frac{C_{G1}}{C_{LE}} + 1\right)} \tag{4.146}$$

$$E = 2\zeta^2 \omega_{T1}\tau - (1+G)\alpha_3 r - 1 \tag{4.147}$$

4.A.2.1.1 Results of Optimization Using the Long Channel Model

If the long channel model is assumed, g_{m2} can be related to g_{m1} by the following expression.

$$g_{m2} = g_{m1}\sqrt{r} \tag{4.148}$$

Then, substitution of (4.64) and (4.148) into (4.56) results in the following equation for $\omega_{T1}\tau$.

$$\omega_{T1}\tau = \frac{\alpha_1 + (1+\alpha_2)r}{\sqrt{r}} \tag{4.149}$$

Substituting this result into the above three equations makes them independent of the time constant and unity gain frequencies. However, the equations are still dependent on r. E can now be rewritten in terms of r as shown below.

$$E = \frac{4\zeta^2 [\alpha_1 + (1+\alpha_2) r]}{\sqrt{r}} - (1+G)\alpha_3 r - 1 \tag{4.150}$$

The ratio of current to input capacitance may be expressed in terms of the ratio r and the time constant by combining (4.4) and (4.149) as shown below.

$$\frac{I}{C_{G1}} = \frac{4 [\alpha_1 + (1+\alpha_2) r]^2}{\tau^2 r} \left(\frac{L^2}{2\mu}\right) \tag{4.151}$$

Now a figure of inefficiency, the product of current, time constant squared, and $2\mu/L^2$ divided by the load capacitance can be written using equations (4.145) and (4.151). The result is shown below. This result is the optimized figure of inefficiency.

$$\frac{2\mu}{L^2}(\tau^2)\frac{I}{C_{LE}} = \frac{4 [\alpha_1 + (1+\alpha_2) r]^2 \{ (1+G)E + 2G + \sqrt{(4G)[(1+G)^2\alpha_3 r + (1+G)E + G]} \}}{r(E^2 - 4\alpha_3 rG)}. \tag{4.152}$$

The results of this analysis are plotted in figures 4.65, 4.66, and 4.67.

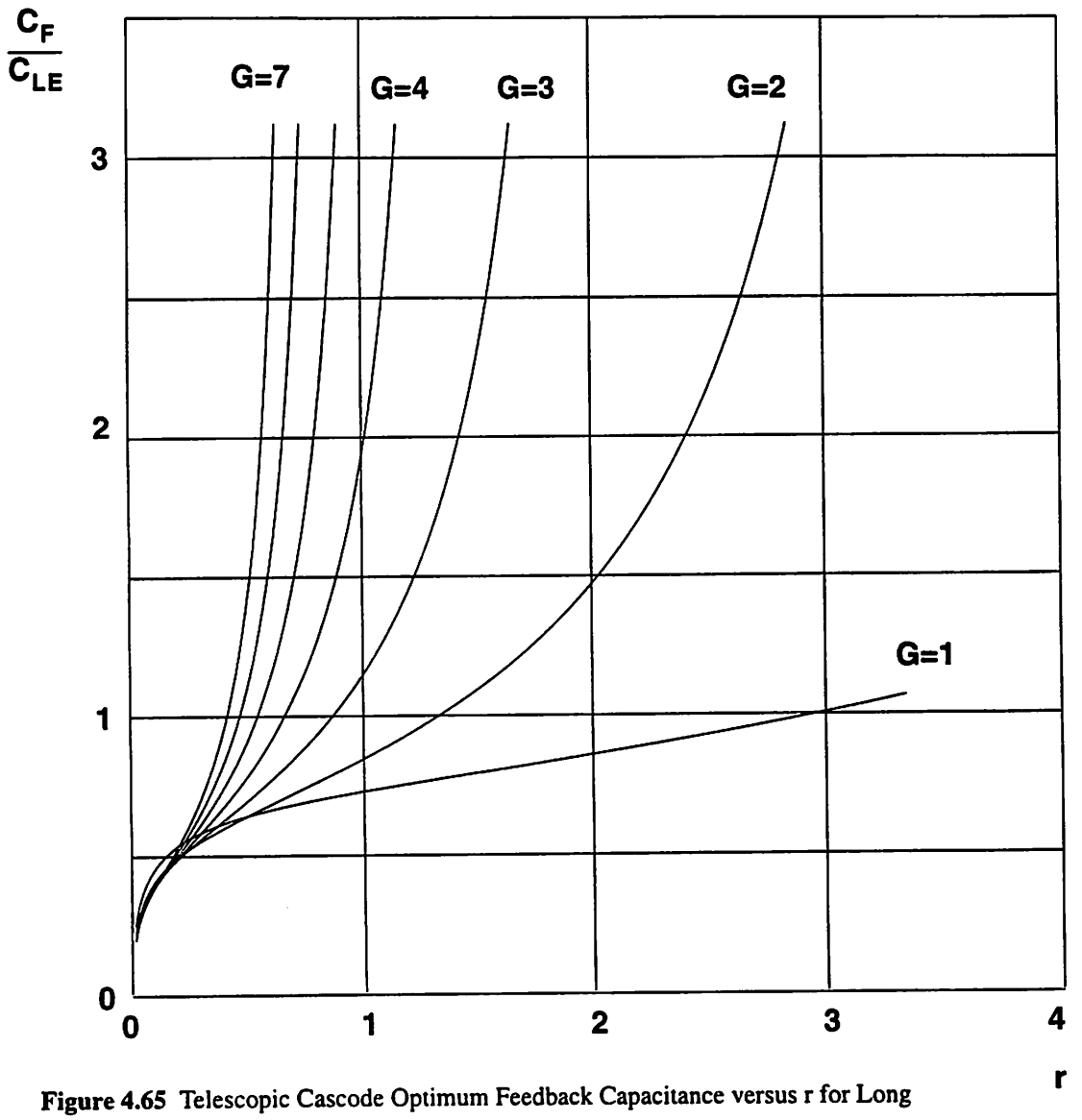


Figure 4.65 Telescopic Cascode Optimum Feedback Capacitance versus r for Long Channel Model, $\alpha_1=\alpha_2=\alpha_3=1$, $\zeta=1$

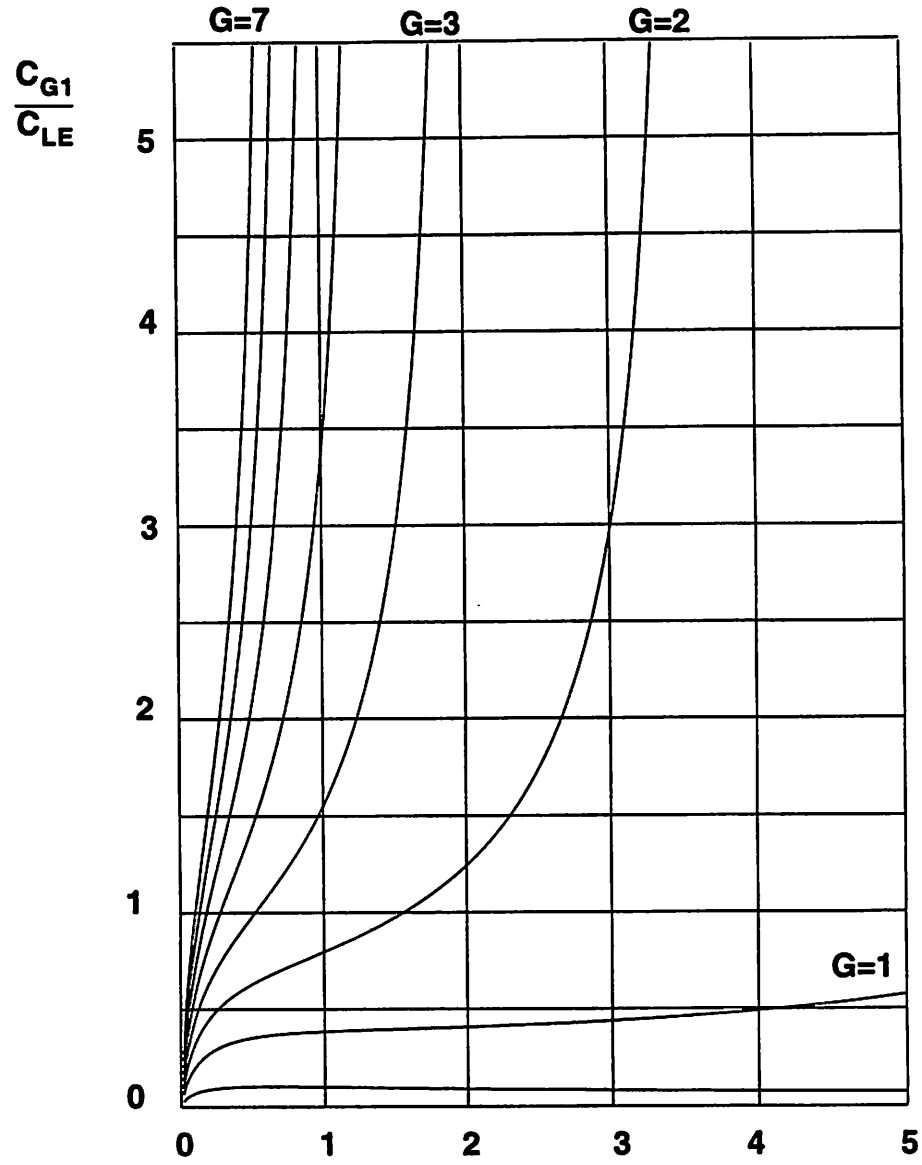


Figure 4.66 Telescopic Cascode Optimum Input Capacitance versus r for Long Channel Model, $\alpha_1=\alpha_2=\alpha_3=1$, $\zeta=1$

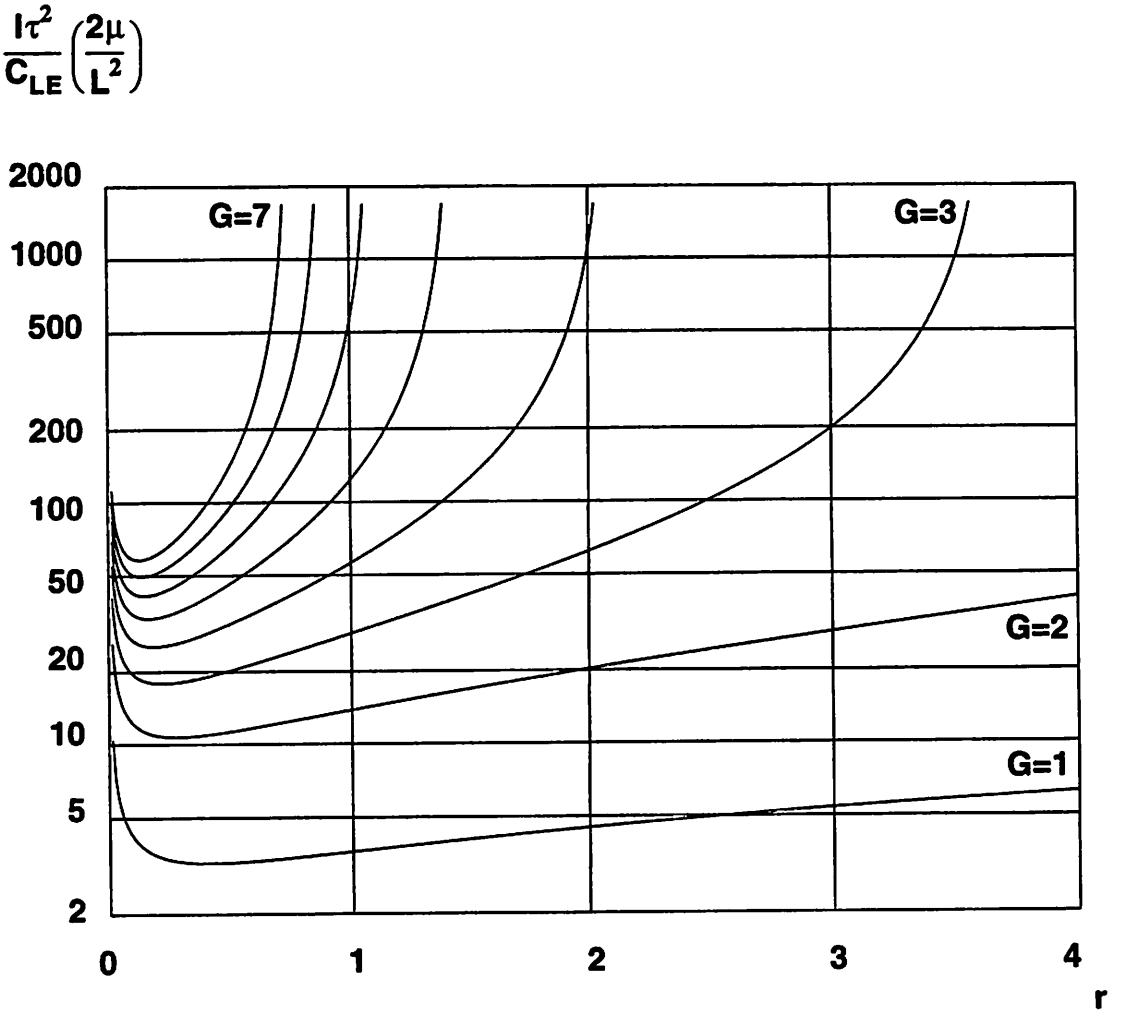


Figure 4.67 Telescopic Cascode Minimum Current versus r for Long Channel Model, $\alpha_1=\alpha_2=\alpha_3=1$, $\zeta=1$

References

[Ahuja83] Bhupendra K. Ahuja, "An Improved Frequency Compensation Technique for CMOS Operational Amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-18, no. 6, Dec. 1983, pp. 629-633.

[Cho95A] Thomas Byunghak Cho and Paul R. Gray, "A 10 b, 20 Msample/s, 35 mW Pipeline A/D Converter," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, March 1995, pp. 166-172.

[Cho95B] Thomas B. Cho, "Low-Power Low-Voltage Analog-to-Digital Conversion Techniques Using Pipelined Architectures," Memorandum No.UCB/ERL M95/23, Electronics Research Laboratory, U. C. Berkeley, April 1995, pp. 107-109.

[Conroy94] Cormac 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, pp.180-206.

[Eschauzier92] Ruud G. H. Eschauzier, Leo P. T. Kerklaan, Johan H. Huijsing, "A 100-MHz 100-dB Operational Amplifier with Multipath Nested Miller Compensation Structure," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1709-1717.

[Gray93] Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd. ed., New York: Wiley, 1993, pp. 607-623.

[Hu88] Chenming Hu, "Solid-State Devices," EECS 231 Class Notes, University of California at Berkeley, Spring 1988.

[Muller86] Richard S. Muller and Theodore I. Kamins, *Device Electronics for Integrated Circuits*, 2nd. ed., New York: Wiley, 1986, pp. 424-435.

[Ribner84] David B. Ribner and Miles A. Copeland, "Design Techniques for Cascoded CMOS Op Amps with Improved PSRR and Common-Mode Input Range," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 919-925.

[Tsividis87A] Yannis P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1987, pp. 141-148.

[Tsividis87B] Yannis P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1987, pp. 175-181.

[Tsividis87C] Yannis P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1987, pp. 305-306.

CHAPTER 5

OPTIMIZATION TECHNIQUES FOR PIPELINED ANALOG TO DIGITAL CONVERTERS

A number of issues are involved in the design of pipelined analog to digital converters. These issues require the designer to make a number of design decisions. For example, the designer must choose the supply voltage, the resolution of each stage, and the sizes of the capacitors in the sample and hold circuits. In this chapter, these issues are discussed in order to facilitate these design choices.

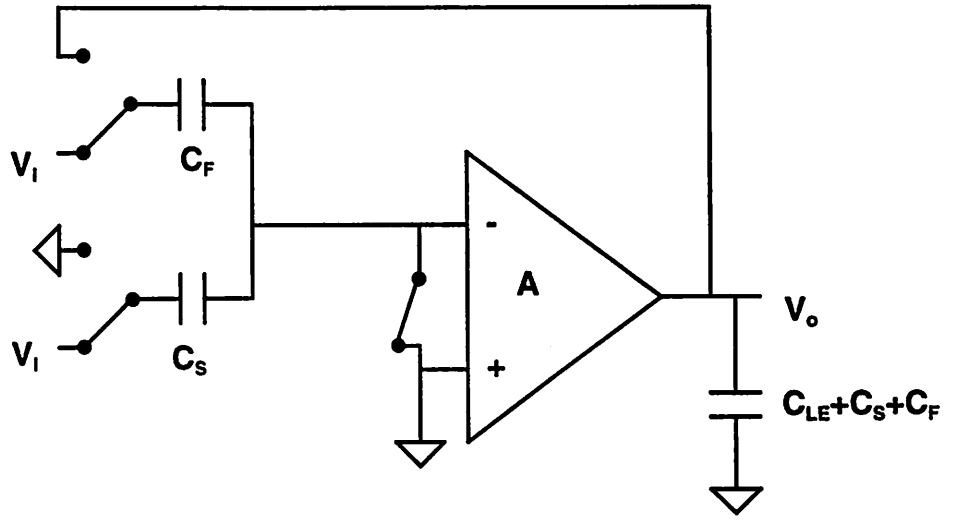
5.1 Pipelined Analog to Digital Converter Design in the Absence of Noise

If the resolution requirement is sufficiently low or if the power supply voltage is sufficiently high, thermal noise is not a serious issue in the design of pipelined analog to digital converters. In this case, the design is then likely to involve a simple trade-off between speed and power. In this section, such a situation is analyzed with the goal of meeting a speed requirement with minimum power. It is assumed here that the residue amplifiers in the pipeline use class A operational amplifiers and that these amplifiers are the only significant contributors to the power dissipation. The optimization will begin by considering an individual amplifier in the pipeline and sizing the sampling capacitor to meet a speed requirement with minimum power. In this analysis, each stage of the pipeline is assumed to be equal in size. Therefore, the capacitance loading each amplifier includes a component equal to the size of the sampling capacitance.

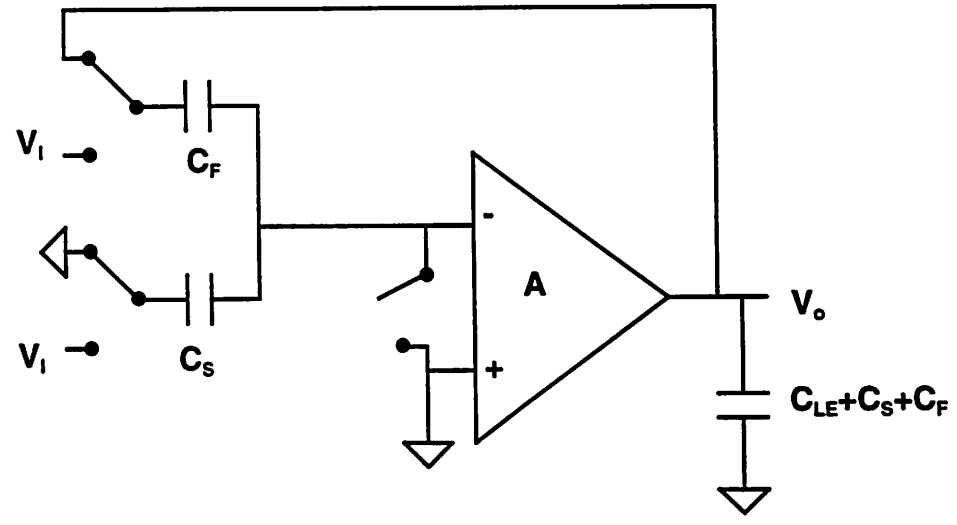
5.1.1 Optimum Sampling Capacitor Size to Minimize Pipeline Power

Here, the speed and power dissipation of a single residue amplifier in the pipeline are discussed. In this section, it is assumed that the sample and hold circuit operates as shown in figure 5.1. Note that it is assumed here that the feedback capacitor is used for both feedback and sam-

pling. The circuit shown in the figure is single ended. In practice, a fully differential configuration would usually be used, but in this model a single ended configuration is shown for simplicity.



(a) sampling phase



(b) hold phase

Figure 5.1 Switched Capacitor Sample and Hold Circuit Operation

The circuit is composed of an operational amplifier, capacitors, and switches. The circuit operates on two phases. During the sampling phase, the input signal is connected to the bottom plate of the sampling capacitor and feedback capacitor while the top plates are grounded. During the hold phase, the top plates of the capacitors float while the bottom plate of the sampling capacitor is connected to a DC value, and the bottom plate of the feedback capacitor is connected to the output. The amplifier is assumed to drive the next pipeline stage of the same size. Therefore, the load to the amplifier is composed of C_S, C_F, and C_{LE}. The capacitor C_{LE} includes the comparator input capacitance and parasitic capacitance due to interconnect. As in chapter 4, the symbol G is used to describe the ratio of the sampling capacitance to the feedback capacitance.

$$G = \frac{C_S}{C_F} \tag{5.1}$$

The above ratio has a strong influence on the feedback factor. The feedback factor is defined as the ratio of the input voltage v_i to the output voltage v_o in the closed loop circuit.

$$f = \frac{v_o}{v_i} = \frac{1}{1 + \frac{C_S}{C_F} + \frac{C_{G1}}{C_F}} = \frac{1}{1 + G + \frac{C_{G1}}{C_F}} \tag{5.2}$$

In order to simplify the analysis here, the amplifier is modeled as a simple one transistor amplifier as shown in figure 5.2, and slewing effects are ignored.

To analyze the step response of the circuit, the circuit is modeled as shown in figure 5.3.

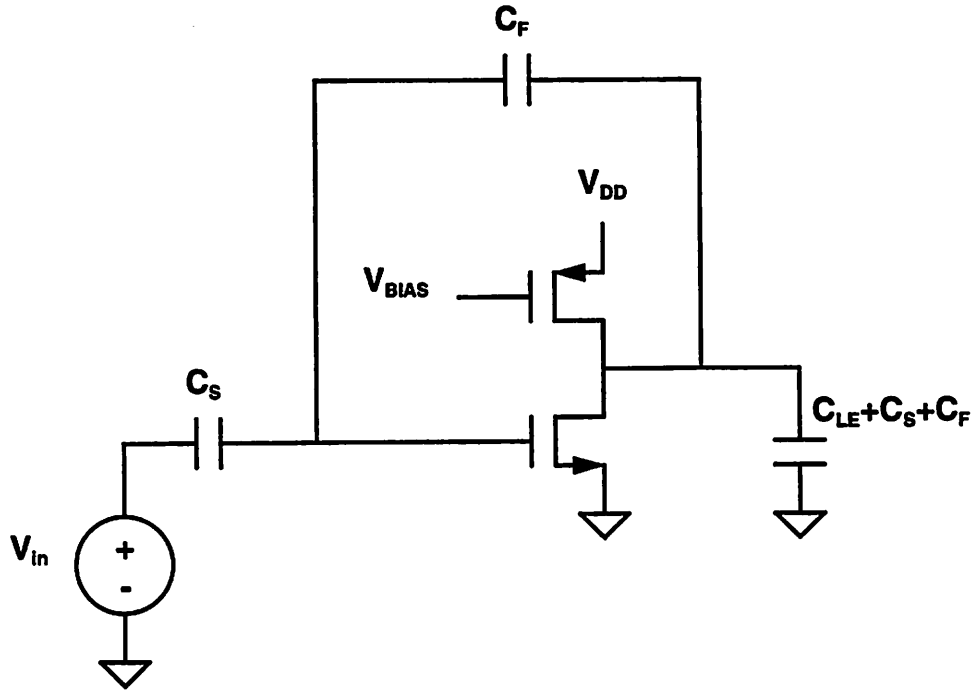


Figure 5.2 Circuit Model for the Hold Phase of a Switched Capacitor Gain Stage Using a Single Transistor Opamp

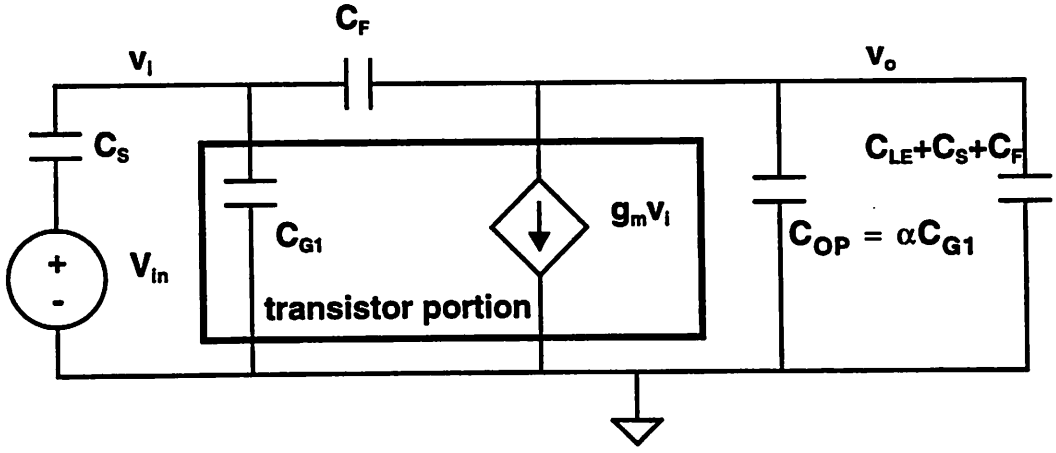


Figure 5.3 Small Signal Equivalent Circuit of the Hold Phase of a Switched Capacitor Gain Stage Using a Single Transistor Opamp

In this analysis, the current density is assumed to be fixed. This assumption implies that the unity gain frequency of the transistor in the critical path is also fixed. This assumption leads to an analysis that is technology independent. The analysis is general enough to apply to bipolar transistors as well as MOSFETs. For more information, this type of analysis is treated carefully by Conroy[Conroy94A].

To begin this analysis, the transconductance is related to the input capacitance by the unity gain frequency ω_T as shown below.

$$g_m = \omega_T C_{G1} \tag{5.3}$$

In the analysis here, the ratio G of sampling capacitance to feedback capacitance is assumed to be constant. A nodal analysis of figure 5.3 yields the following formula for the time constant τ.

$$\tau = \frac{C_{LE} + C_S + C_F + C_F(1-f) + \alpha C_{G1}}{f g_m} \tag{5.4}$$

Now equations (5.1), (5.2), (5.3), and (5.4) are combined to eliminate f, g_m, and C_S. The result is shown below.

$$\tau = \frac{1}{\omega_T C_{G1}}\left[\left(1 + G + \frac{C_{G1}}{C_F}\right)(C_{LE} + (G+2)C_F + \alpha C_{G1}) - C_F\right] \tag{5.5}$$

The goal of this analysis is to meet a given speed requirement with minimum power while the feedback capacitance is allowed to vary. A trade-off exists in choosing the feedback capacitance, because as the feedback capacitance is increased, the load capacitance and the feedback factor both tend to increase. Increased load capacitance tends to slow things down, but increased feedback factor tends to speed things up. Figure 5.4 shows a plot of C_{G1} versus C_F based on equation (5.5). Because the current density is fixed, power in this graph is proportional to C_{G1}. From this figure, it can be seen that there is an optimum feedback capacitance to minimize the power. The following example can give some context to the graph. For the case where the ratio of current to transistor width is 1μA/μm, C_{ox}=1.535 fF/(μm)^2, and channel length is 1.2μm, C_{G1}/C_{LE}=1 corresponds to a bias current of 543μA/μm.

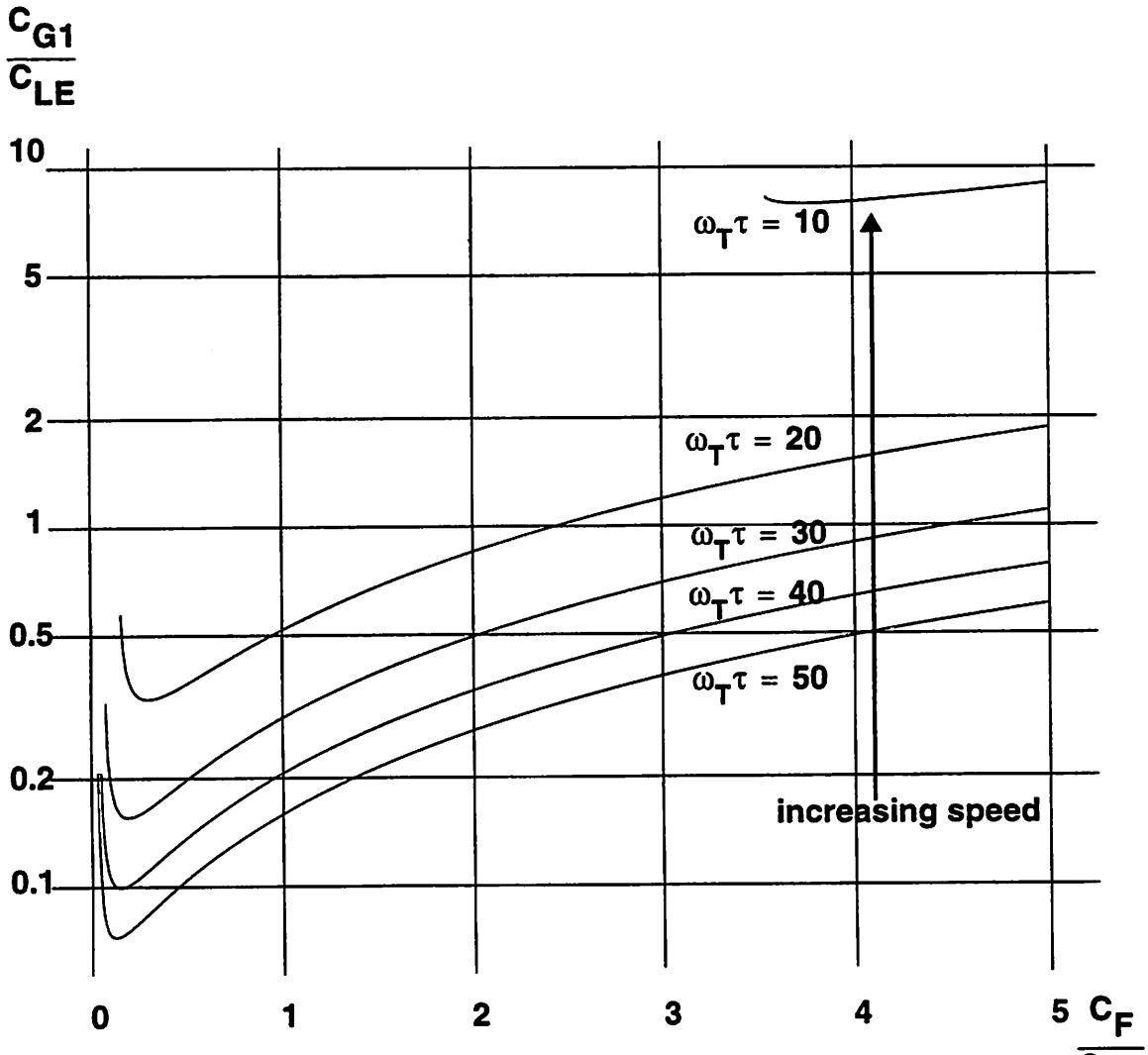


Figure 5.4 Input Capacitance of Residue Amplifier Versus Feedback Capacitance, G=1, α=1

Because the speed is fixed, the required time constant τ is fixed, and because the current density is fixed, the current and thus the power are proportional to the input capacitance C_{G1}. Therefore, minimizing the input capacitance minimizes the power dissipation. To minimize the power dissipation, equation (5.5) is differentiated with respect to C_F, and the derivative of C_{G1} with respect to C_F is set to zero as shown below.

$$\frac{dC_{G1}}{dC_F} = 0 \tag{5.6}$$

The above operation results in the following relation between C_{G1} and C_F.

$$C_{G1} = \frac{(G+1)C_{LE} + 2(1+3G+G^2)C_F}{\omega_T\tau - (G+2) - \alpha(G+1)} \tag{5.7}$$

When equations [5.5] and [5.7] are combined, the following relations are obtained.

$$\frac{C_F}{C_{LE}} = \frac{[\omega_T\tau - (G+2) - \alpha(G+1)]\left[1 + \sqrt{\frac{(G+1)\omega_T\tau - 1}{G^2+3G+1}}\right] + 2\alpha(G+1)}{[\omega_T\tau - (G+2) - \alpha(G+1)]^2 - 4\alpha(G^2+3G+1)} \tag{5.8}$$

$$\frac{C_{G1}}{C_{LE}} = \frac{\left[\sqrt{\omega_T\tau(G+1) - 1} + \sqrt{G^2+3G+1}\right]^2 - \alpha(G+1)^2}{[\omega_T\tau + (G+2) - \alpha(G+1)]^2 - 4\alpha(G^2+3G+1)} \tag{5.9}$$

From equation (5.9) it is possible to determine the maximum possible speed (minimum time constant τ_{min}) by setting the denominator to zero. This corresponds to the time constant for which the minimum power goes to infinity. This time constant corresponds to the maximum speed. The result is shown below.

$$\omega_T\tau_{min} = (G+2) + \alpha(G+1) + \sqrt{4\alpha(G^2+3G+1)} \tag{5.10}$$

The results of this optimization analysis are shown in figures 5.5 and 5.6.

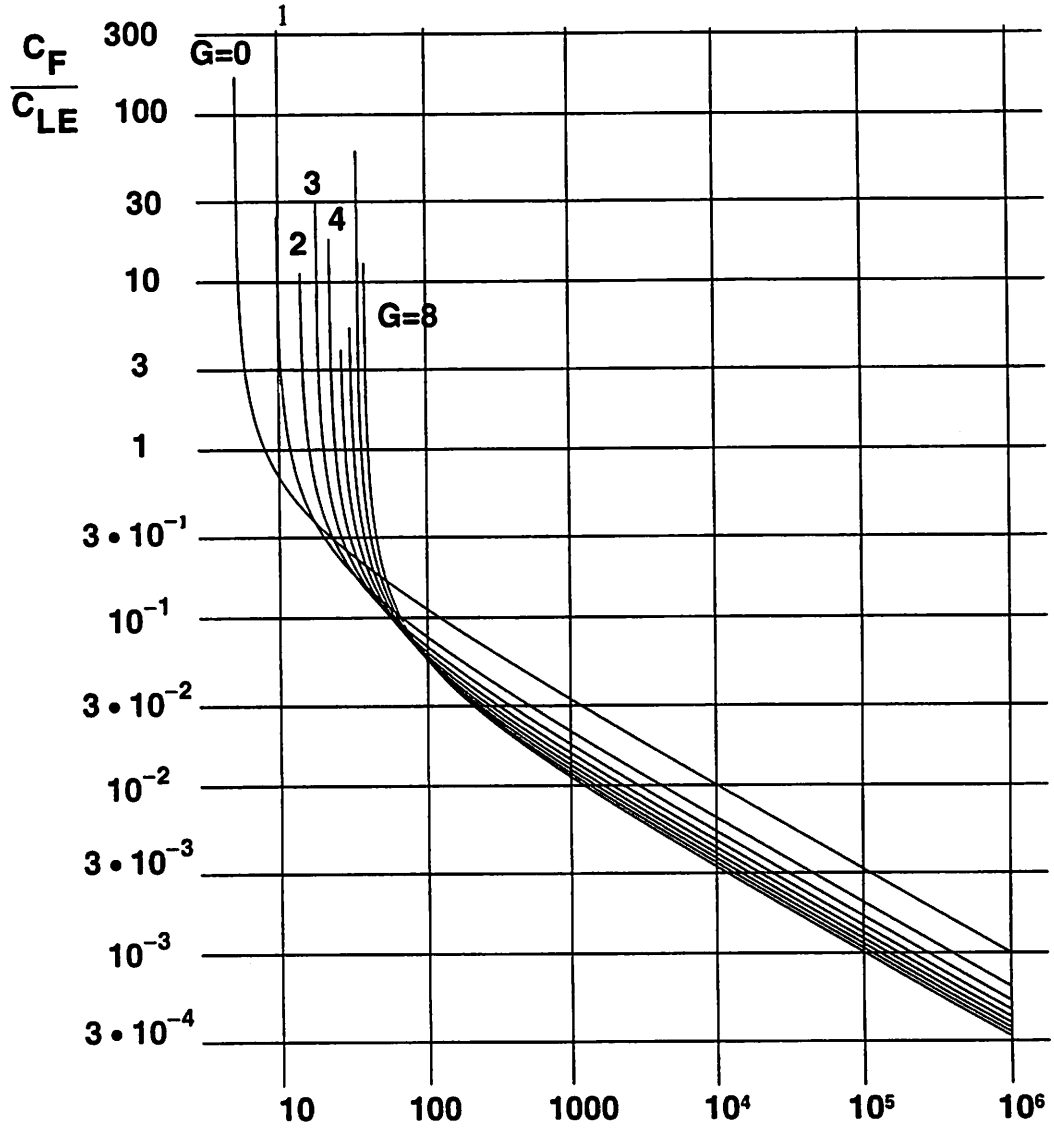


Figure 5.5 Feedback Capacitance Versus Normalized Time Constant, $\alpha = 1$

These graphs show that both the optimum feedback capacitance and optimum input capacitance tend to be reduced as speed is reduced if the current density is fixed. Figure 5.6 shows that power dissipation tends to increase as the gain is increased. Figure 5.5 shows that at low speed increasing the gain tends to result in a smaller value for the optimum feedback capacitance. However, increasing the gain tends to increase the optimum feedback capacitance as the speed limit of the technology is approached.

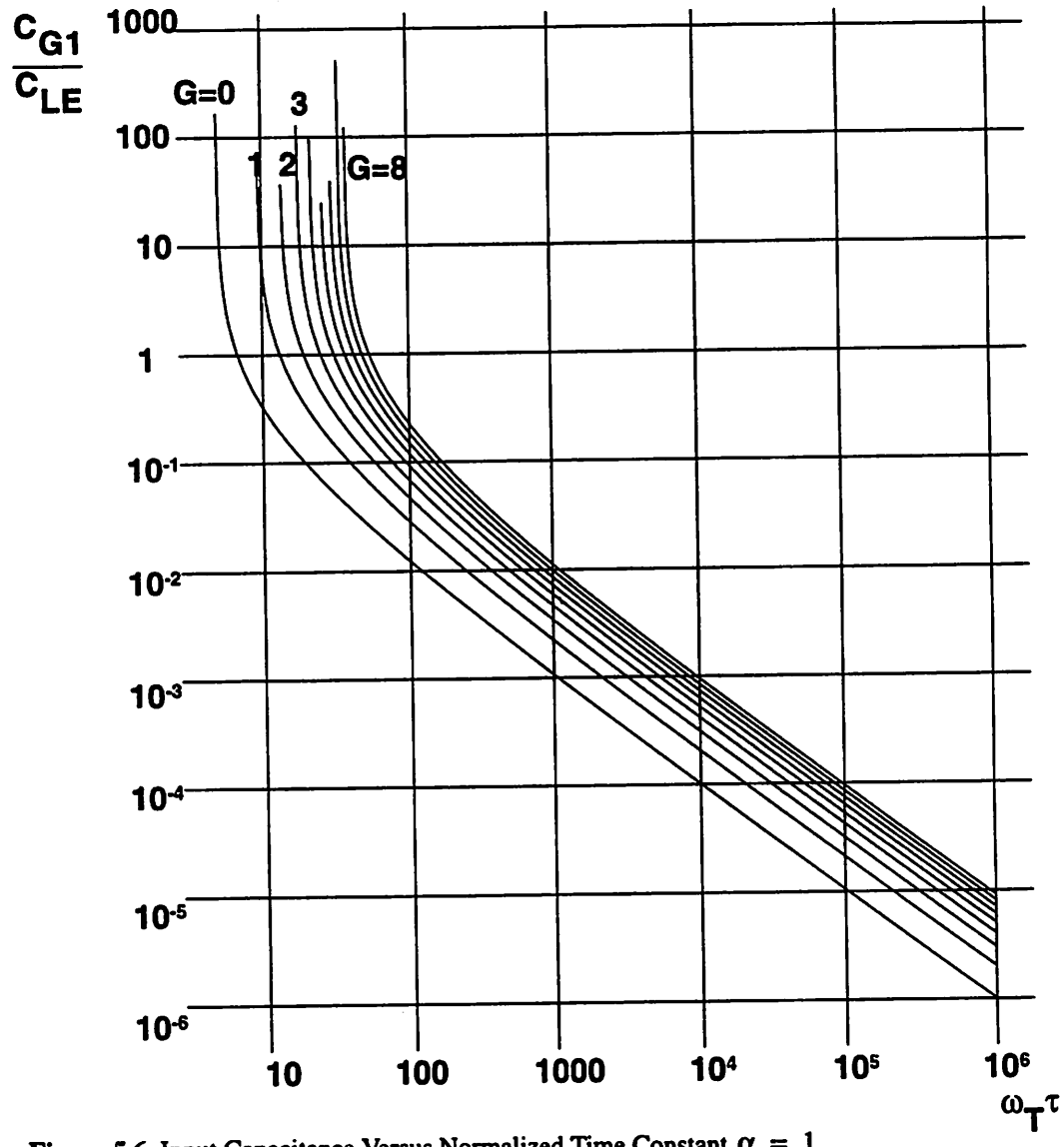


Figure 5.6 Input Capacitance Versus Normalized Time Constant $\alpha = 1$

5.1.2 Optimum Number of Bits Per Stage to Minimize Pipeline Power

In pipelined analog to digital converter design, the number of bits resolved by a single stage of the pipeline is a key design decision. Resolving a large number of bits in each stage requires the use of an operational amplifier with a large closed loop gain, and thus a low feedback factor. Resolving a large number of bits in each stage also implies the need for a large number of comparators in each stage. Because the amplifier must drive these comparators, increasing the number of comparators increases the load capacitance presented to the amplifier. Therefore, if the speed and

overall resolution requirements are fixed, increasing the per stage resolution increases the per stage power dissipation. On the other hand, increasing the per stage resolution reduces the number of stages required in the pipeline. This tends to save power. Therefore, a trade-off exists in choosing the per stage resolution of a pipelined analog to digital converter. In this section, the per stage resolution will be chosen to minimize the power dissipation.

An analysis is presented here in an effort to quantify the optimization of the per stage resolution. This problem has also been treated by Lewis[Lewis92]. The analysis here builds on the results of the previous section. The power dissipated per stage is proportional to the input capacitance of the amplifier as given by equation (5.9) and repeated here for reference.

$$C_{G1} = C_{LE}\left\{\frac{\left[\sqrt{\omega_T\tau\,(G+1)-1}+\sqrt{G^2+3G+1}\right]^2-\alpha\,(G+1)^2}{[\omega_T\tau+(G+2)-\alpha\,(G+1)]^2-4\alpha\,(G^2+3G+1)}\right\} \tag{5.11}$$

C_{LE} in the above expression is assumed here to be composed of two components: a parasitic capacitance C_{IN} due to interconnect, and the combined load presented by the comparators. Therefore, C_{LE} is given by the following expression where C_{comp} is the input capacitance of a comparator and n is the per stage resolution.

$$C_{LE} = 2\,(2^n-1)\,C_{comp}+C_{IN} \tag{5.12}$$

The above formula assumes that each stage of the pipeline has 1 bit of redundancy for error correction. This is common practice. For very small values of n, more redundancy than predicted by this formula is required because the number of comparators cannot be less than one. Therefore, this analysis is expected to underestimate the required power dissipation for pipelines with small values of n.

The total power of the pipelined analog to digital converter is proportional to the input capacitance given in equation (5.9) multiplied by the number of stages N_s in the pipeline. The number of stages in the pipeline is simply the overall resolution N_B divided by the per stage resolution.

$$N_s = \frac{N_B}{n} \tag{5.13}$$

Therefore, the combined input capacitance of all of the amplifiers in the pipeline is given by the following equation if all of the stages are assumed to be identical.

$$\frac{C_T}{N_B} = \frac{1}{n}\left[2\left(2^n-1\right)C_{comp}+C_{IN}\right]\left\{\frac{\left[\sqrt{\omega_T\tau\left(G+1\right)-1}+\sqrt{G^2+3G+1}\right]^2-\alpha\left(G+1\right)^2}{\left[\omega_T\tau+\left(G+2\right)-\alpha\left(G+1\right)\right]^2-4\alpha\left(G^2+3G+1\right)}\right\} \tag{5.14}$$

With the assumption that $C_{comp}=C_{IN}$ and $G=2^n-1$, this equation is plotted in figure 5.7. Note that the optimal per stage resolution is less than 1 regardless of the speed. The power dissipation for small values of n is actually higher than what is shown in this graph due to the fact that fractional numbers of comparators cannot be realized. However, the general trend that the optimal per stage resolution is less than 1 is still true even if the minimum number of comparators is set at 2.

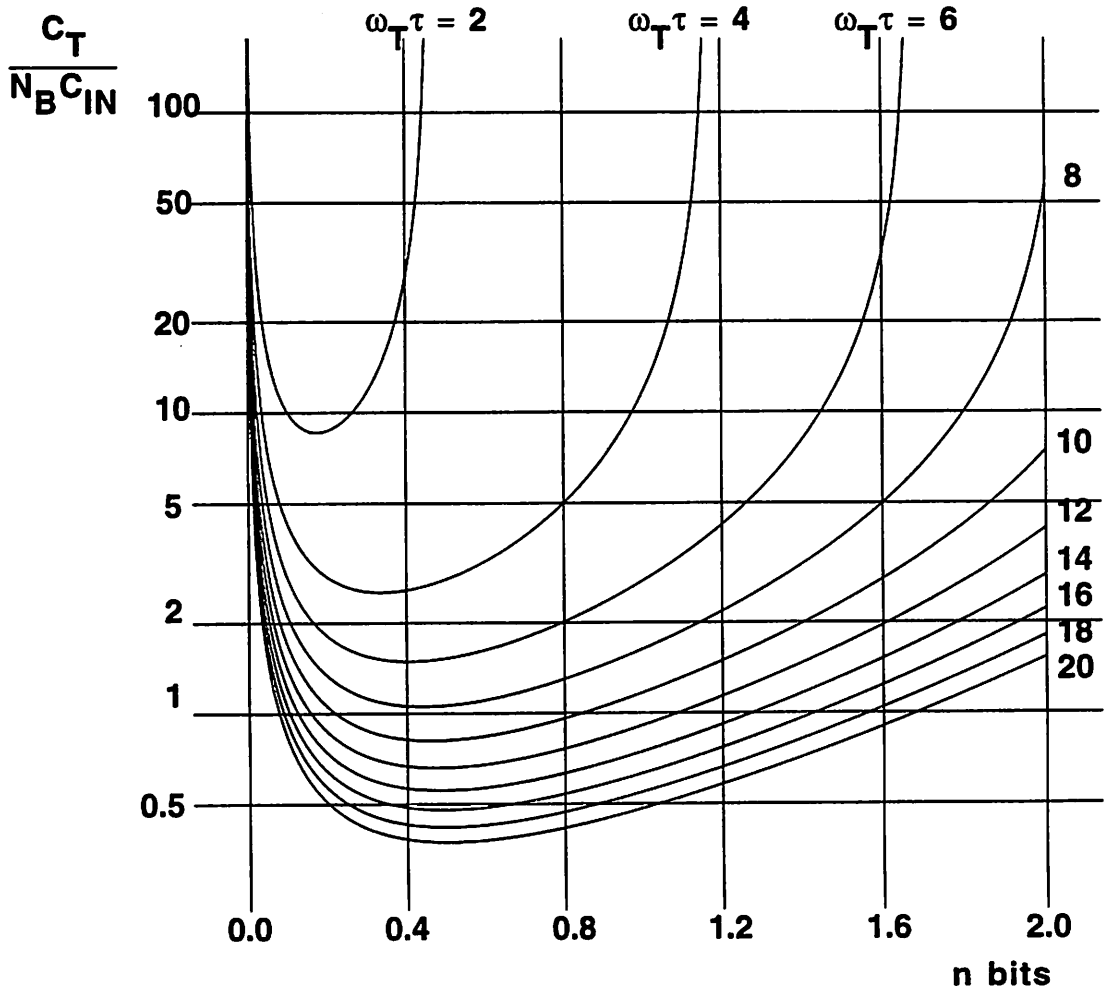


Figure 5.7 Total Amplifier Input Capacitance Versus the Per Stage Resolution

5.2 Pipelined Analog to Digital Converter Design in the Presence of Noise

If the required resolution is high, or if the available supply voltage is low, thermal noise becomes a significant issue. The thermal noise issue greatly affects the choice of the size of the sampling capacitors and the per stage resolution. In this section, techniques for determining the appropriate values for the sampling capacitance and per stage resolution are discussed. The effects of supply voltage on power dissipation are also discussed.

This section discusses three analysis techniques for optimization having differing levels of accuracy. The table below summarizes the assumptions of each of the optimizations presented.

Table 5.1

Section	Effects Included	Effects Neglected
5.2.2	Capacitive Loading by Feedback Network Loading by Sampling Capacitance of Following - Stage	Opamp Input Capacitance Parasitic Opamp Output Capacitance Comparator Input Capacitance Interconnect Capacitance
5.2.3	Capacitive Loading by Feedback Network Loading by Sampling Capacitance of Following - Stage Opamp Input Capacitance Parasitic Opamp Output Capacitance	Comparator Input Capacitance Interconnect Capacitance
5.2.4	Capacitive Loading by Feedback Network Loading by Sampling Capacitance of Following - Stage Opamp Input Capacitance Parasitic Opamp Output Capacitance Comparator Input Capacitance Interconnect Capacitance	

5.2.1 Thermal Noise in Switched Capacitor Gain Blocks

In a switched capacitor gain block such as the one shown in figure 5.1, two sources of thermal noise are important. The sampling switches used to connect the input signal to the sampling capacitor and to connect the sampling capacitor to ground are one important source of thermal noise. The noise resulting from this source is commonly called kT/C noise because it is proportional to k_BT/C_S where k_B is Boltzmann's constant, T is temperature, and C_S is the sampling capacitance. Thermal noise in the amplifier also makes a significant contribution to the input referred noise of the switched capacitor gain stage. This section attempts to quantify the resolution degradation caused by thermal noise from the sampling switches and the amplifier.

5.2.1.1 Thermal Noise Contribution of the Sampling Switches (kT/C Noise)

The kT/C noise contribution to the input referred noise in a switched capacitor gain block has been treated carefully by Conroy[Conroy94B] and Gregorian[Gregorian86]. A simple analysis is shown here to illustrate the basic concepts. The sampling of thermal noise in the circuit of figure 5.1 may be modeled by the circuit shown in figure 5.8. The noise in the sampling switch that connects the

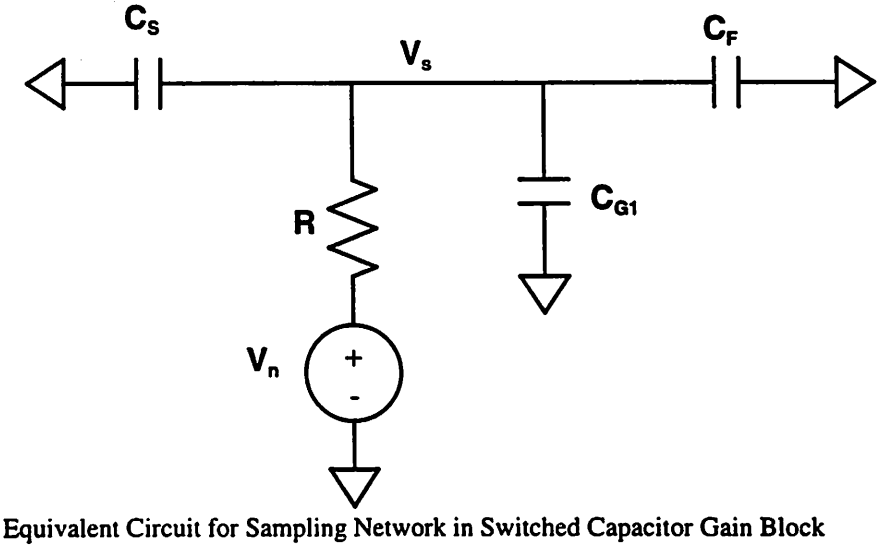


Figure 5.8 Equivalent Circuit for Sampling Network in Switched Capacitor Gain Block

input of the amplifier is represented by a Thevenin equivalent. R is the resistance of the switch, and

C_{G1} is the input capacitance of the amplifier. The above circuit acts as a low pass filter between the noise source V_n and V_s. The frequency domain value of V_s is given as shown below.

$$V_s = \frac{V_n}{1 + j\omega RC} \tag{5.15}$$

C in the above equation is the sum of the three capacitors in the circuit.

$$C = C_S + C_F + C_{G1} \tag{5.16}$$

The noise charge Q on the capacitor network is given as a function of frequency as shown below.

$$Q = CV_s = \frac{CV_n}{1 + j\omega RC} \tag{5.17}$$

In order to determine the total noise charge Q_s sampled onto the capacitor network, it is necessary to integrate the noise charge over all frequencies. Because the noise charge is a random quantity, the spectral density of the noise charge rather than the noise charge itself is integrated over all frequencies. The spectral density S_Q of the noise charge is found by squaring the magnitude of the noise charge as shown below.

$$S_Q = \frac{C^2\sigma_{V_n}^2}{1 + \omega^2 R^2 C^2} \tag{5.18}$$

The standard deviation σ_{Vn} of the thermal noise voltage V_n in the angular frequency domain is given as follows.

$$\sigma_{V_n}^2 = \frac{4k_B TR}{2\pi} \tag{5.19}$$

Using this value, the charge sampled onto the capacitor network is given as shown below.

$$Q_s^2 = \int_0^\infty S_Q d\omega = k_B TC \tag{5.20}$$

During the hold phase of operation, the charge is all moved to the feedback capacitor. Thus, the standard deviation σ_{Von} of the resulting noise voltage V_{on} at the amplifier output is given by the following formula.

$$\sigma_{Von} = \frac{Q_s}{C_F} = \frac{1}{C_F}\sqrt{k_B T(C_S + C_F + C_{G1})} \tag{5.21}$$

Referring this noise voltage at the output back to the input of the sample and hold amplifier gives the following result.

$$\sigma_{V_{in}} = \frac{\sigma_{V_{on}}}{A_C} = \sqrt{\frac{k_B T}{f A_C^2 C_F}} \tag{5.22}$$

The above result is true for the single ended circuit shown. For a fully differential circuit, the input referred noise voltage is larger by a factor of $\sqrt{2}$ as shown below.

$$\sigma_{V_{indifferential}} = \sqrt{\frac{2k_B T}{f A_C^2 C_F}} \tag{5.23}$$

If the feedback capacitor is also used for sampling as it is in figure 5.1, then $A_C=1+C_S/C_F$ and the RMS value of the input referred noise may be expressed as shown below.

$$\sigma_{V_{indifferential}} = \sqrt{\frac{2k_B T}{f A_C (C_F + C_S)}} \tag{5.24}$$

5.2.1.2 Thermal Noise Contribution of the Transconductance Amplifier

The resistive channel of the MOSFET devices in the operational transconductance amplifier also has thermal noise and contributes to the input referred noise of the switched capacitor gain block.

As an example, consider the telescopic cascode amplifier shown in figure 5.9. In this amplifier, transistors M_1, M_2, M_7, and M_8 contribute to the thermal noise seen at the output. It is useful to define a parameter β that is defined as the ratio of the sum of the transconductances of all the noise contributing devices to the transconductance of the input device.

$$\beta = \frac{\sum_{noisecontributors} g_m}{g_{minput}} \tag{5.25}$$

For the telescopic amplifier shown in figure 5.9, β has the following value.

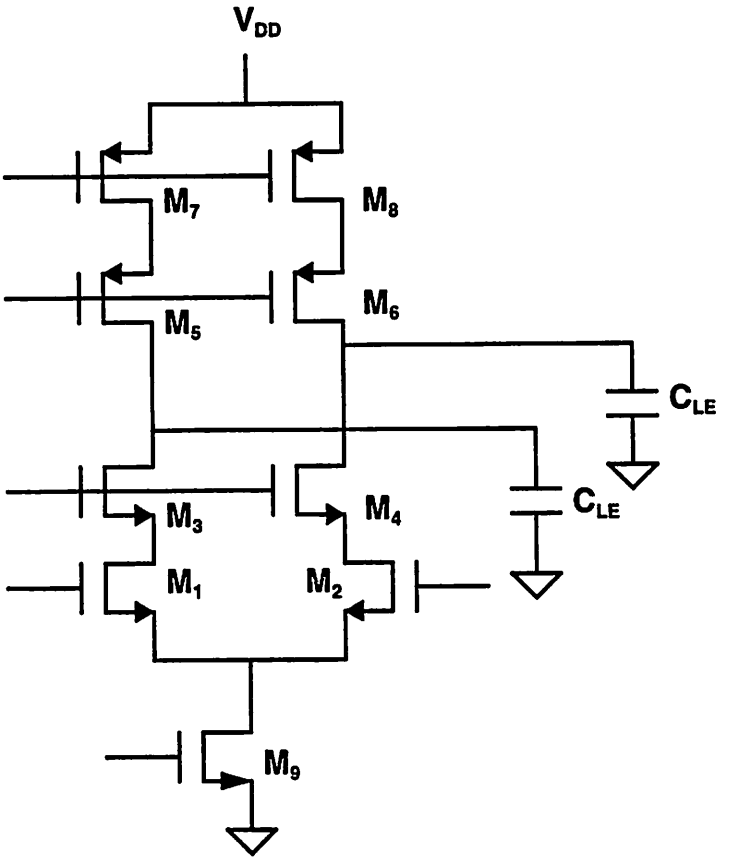


Figure 5.9 Telescopic Cascode Operational Amplifier

$$\beta = \frac{g_{m1} + g_{m7}}{g_{m1}} \tag{5.26}$$

This value will be used later to express the thermal noise contribution in terms of circuit element values. Next, the thermal noise contribution of several operational amplifier topologies will be analyzed.

5.2.1.2.1 Thermal Noise of a Single-Stage Amplifier

In order to determine the thermal noise contribution of a single stage amplifier, the small signal model of figure 5.10 is used. In the figure, the thermal noise in the channel of the noise contributing transistors is represented by the current source i_{dn}. A nodal analysis of the circuit yields the following expression for the frequency domain output voltage in terms of the noise current.

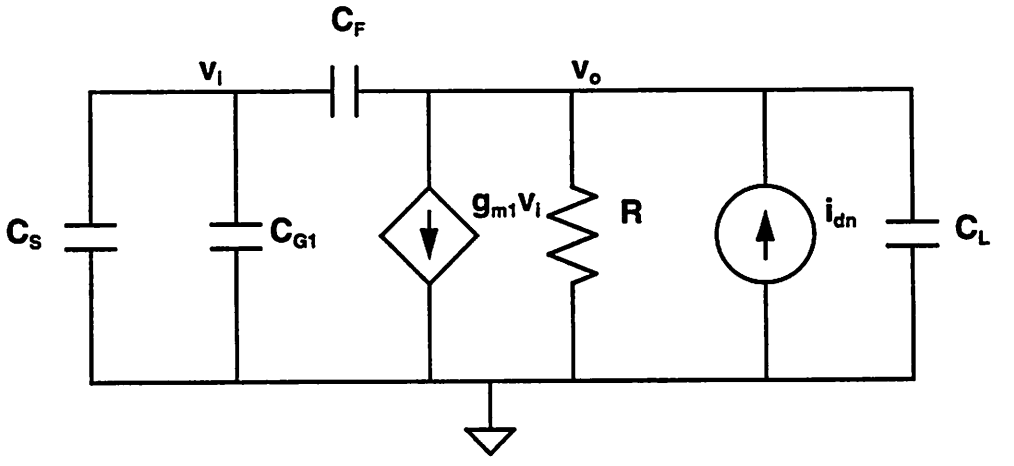


Figure 5.10 Small Signal Model for the Analysis of Thermal Noise in a Single Stage Amplifier During Hold Phase of a Switched Capacitor Circuit

$$v_o = \frac{\frac{i_{dn}}{C_2}}{j\omega + \frac{g_m f + \frac{1}{R}}{C_2}} \tag{5.27}$$

In the above formula, f is the feedback factor and C_2 represents the total capacitance loading the amplifier. This capacitance is given by the following expression.

$$C_2 = C_L + C_F(1 - f) \tag{5.28}$$

Now the noise voltage v_o at the output is integrated over all frequencies to get the total noise voltage v_{on} seen at the output. The result is shown below.

$$v_{on}^2 = \frac{\pi}{2}\left(\frac{i_{dn}^2}{\omega_c C_2^2}\right) \tag{5.29}$$

The cutoff frequency ω_c is given by the following expression.

$$\omega_c = \frac{1}{C_2}(g_m f + \frac{1}{R}) \tag{5.30}$$

The spectral density of the noise current i_{dni} of the input device is proportional to its transconductance and is given by the following formula.

$$i_{dni}^2 = \frac{4k_B T \frac{2}{3} g_{m1}}{2\pi} \tag{5.31}$$

Because the amplifier often has load devices that also contribute to the noise, the total spectral density of the noise current is found by multiplying the above formula by β as shown below.

$$i_{dn}^2 = \beta i_{dni}^2 = \frac{4\beta k_B T \frac{2}{3} g_{m1}}{2\pi} \tag{5.32}$$

Substituting equations (5.30) and (5.32) into equation (5.29) and assuming R approaches infinity yields the following formula for the noise voltage at the output.

$$v_{on}^2 = \frac{\frac{2}{3}\beta k_B T}{fC_2} \tag{5.33}$$

For a fully differential circuit, the variance of the noise is larger by a factor of 2.

$$v_{ondiff}^2 = \frac{\frac{4}{3}\beta k_B T}{fC_2} \tag{5.34}$$

Finally, the input referred fully differential noise v_{in} is given by dividing the output noise in the above equation by the square of the closed loop gain.

$$v_{in}^2 = \frac{4\beta k_B T}{3fA_C^2[C_L + C_F(1-f)]} \tag{5.35}$$

Note that the thermal noise contribution of the amplifier is inversely proportional to a capacitance, in this case the load capacitance. Thus, the combined thermal noise contribution of the sampling switches and the operational amplifier is reduced by increasing capacitor sizes.

5.2.1.2.2 Thermal Noise of a Critically Damped Telescopic Cascode Amplifier

A telescopic cascode amplifier such as the one in figure 5.9 is often modeled as a single stage, single pole amplifier. However, this amplifier does contain more than one pole. In some instances the presence of the nondominant poles may be significant. In this section, it is assumed that the closed loop system using the telescopic cascode amplifier has two real poles at the same frequency (critical damping). The analysis procedure is similar to the procedure for the simple single stage amplifier described above with the following exceptions. Because the nondominant pole in the amplifier is significant, the cascode transistors contribute a significant amount of noise to the output. This would not be true if the nondominant pole were at a much higher frequency than the dominant pole. The analysis for this topology is based on the small signal model shown in figure 5.10. A nodal analysis of the circuit yields the following s-domain expression for the output voltage.

$$v_o = \frac{\frac{1}{C_2}\left(\frac{g_{mn2}}{C_{N1}} i_{dn1} + s i_{dn2}\right)}{s^2 + \frac{g_{mn2}}{C_{N1}} s + \frac{f g_{mn1} g_{mn2}}{C_{N1} C_2}} + \frac{\frac{1}{C_2}\left(s + \frac{g_{mn2}}{C_{N1}}\right)\left(\frac{g_{mp2}}{C_{P1}} i_{dp1} + s i_{dp2}\right)}{\left(s + \frac{g_{mp2}}{C_{P1}}\right)\left(s^2 + \frac{g_{mn2}}{C_{N1}} + \frac{f g_{mn1} g_{mn2}}{C_{N1} C_2}\right)} \tag{5.36}$$

In the above circuit, $C_2 = C_L + C_F(1-f)$. In this analysis, the poles are all assumed to be real and equal. This assumption allows the above equation to be rewritten as shown below.

$$v_o = \frac{\frac{1}{C_2}\left(\frac{g_{mn2}}{C_{N1}} i_{dn1} + s i_{dn2}\right)}{(s+\omega_n)^2} + \frac{\frac{1}{C_2}(s+2\omega_n)\left(\frac{g_{mp2}}{C_{P1}} i_{dp1} + s i_{dp2}\right)}{(s+\omega_n)^3} \tag{5.37}$$

In the above equation ω_n is given by the following relation.

$$\omega_n = \frac{g_{mn2}}{2C_{N1}} \tag{5.38}$$

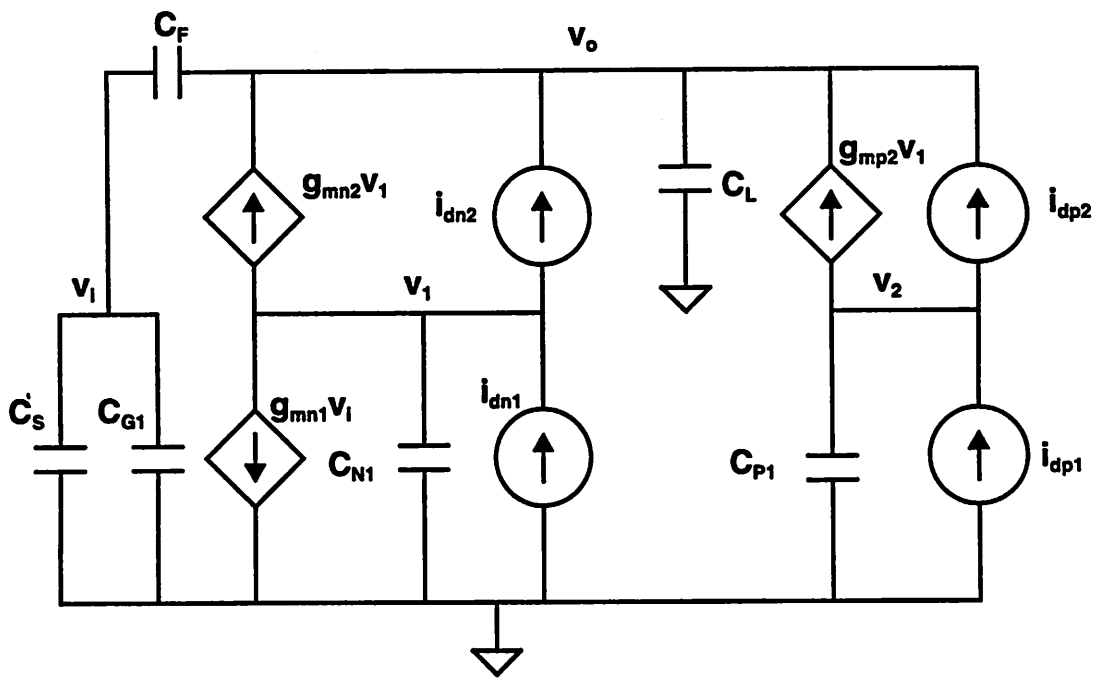


Figure 5.11 Small Signal Model for the Analysis of Thermal Noise in a Telescopic Cascode Amplifier

The spectral density of the noise at the output can then be found by taking the magnitude squared of equation (5.37) assuming that all four of the noise currents are independent. This spectral density is integrated over all frequencies to obtain the noise power at the output. This noise power is then referred back to the input of the closed loop system with the following result for the fully differential case.

$$v_{idiff}^2 = \frac{\frac{4}{3}k_B T}{f A_C^2 C_2}\left\{1 + \frac{1}{4}\frac{g_{mn2}}{g_{mn1}} + \frac{13}{16}\frac{g_{mp1}}{g_{mn1}} + \frac{7}{16}\frac{g_{mp2}}{g_{mn1}}\right\} \tag{5.39}$$

5.2.1.2.3 Thermal Noise of a Preamplifier Driving a Single Stage Amplifier

The analysis procedure to determine the thermal noise contribution of a single stage amplifier driven by a preamplifier is similar to the procedure for the simple single stage amplifier described earlier. Because the gain of the preamplifier is low, both stages of the amplifier contribute a signif-

icant amount of noise. The analysis for this topology is based on the small signal model shown in figure 5.10. A nodal analysis of the circuit yields the following s-domain expression for the output

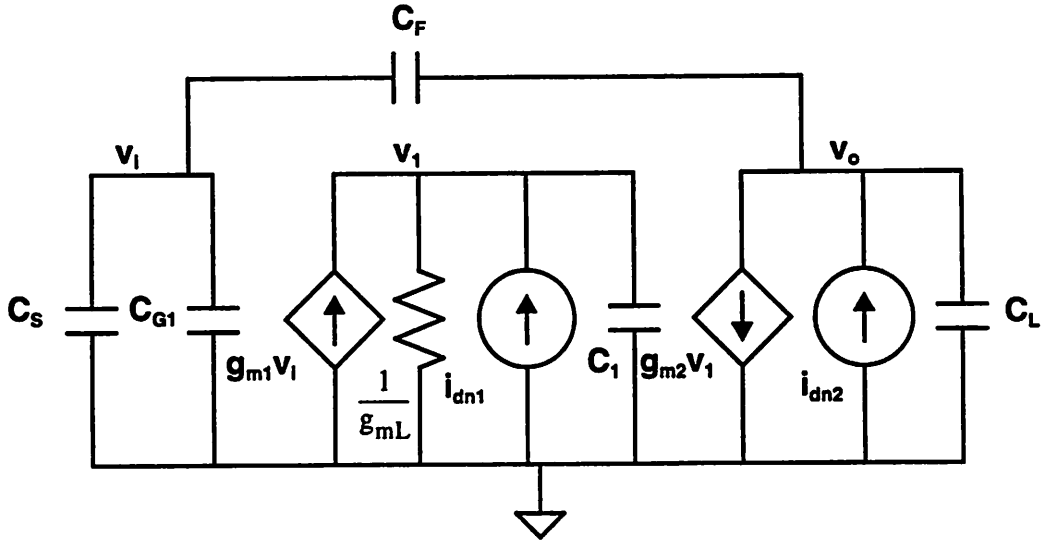


Figure 5.12 Small Signal Model for the Analysis of Thermal Noise in a Single Stage Amplifier Driven by a Wideband Preamplifier

voltage in terms of the noise currents.

$$v_o = \frac{\frac{-g_{m2}i_{dn1}}{C_1C_2} + \left(s + \frac{g_{mL}}{C_1}\right)\frac{i_{dn2}}{C_2}}{s^2 + \frac{g_{mL}}{C_1}s + \frac{g_{m1}g_{m2}f}{C_1C_2}} \tag{5.40}$$

C_2 in the above equation is given by the following equation.

$$C_2 = C_L + C_F(1-f) \tag{5.41}$$

In this analysis, the two poles are assumed to be real and equal. With this assumption the above equation may be rewritten as follows where ω_n is the magnitude of the pole position.

$$v_o = \frac{\frac{-g_{m2}i_{dn1}}{C_1C_2} + (s+2\omega_n)\frac{i_{dn2}}{C_2}}{(s+\omega_n)^2} \tag{5.42}$$

The total noise at the output is found by integrating the spectral density of v_o over all frequencies. The spectral densities S_{idn1} and S_{idn2} of the noise currents are given by the following equations.

$$S_{i_{dn1}} d\omega = \frac{4k_B T}{2\pi} (\frac{2}{3}) \beta_1 g_{m1} d\omega \tag{5.43}$$

$$S_{i_{dn2}} d\omega = \frac{4k_B T}{2\pi} (\frac{2}{3}) \beta_2 g_{m2} d\omega \tag{5.44}$$

$$\beta_1 = \frac{g_{m1} + g_{mL}}{g_{m1}} \tag{5.45}$$

Assuming that the two noise currents are independent, the total noise appearing at the output is given by the following equation.

$$v_{on}^2 = \frac{k_B T}{6 f A_{PRE}} \left(\frac{\beta_1}{f C_1} + \frac{5\beta_2}{C_2} \right) \tag{5.46}$$

The input referred noise voltage is then found by dividing the noise at the output by the closed loop gain A_C with the result shown below.

$$v_{inputreferred}^2 = \frac{k_B T}{6 f A_{PRE} A_C^2} \left(\frac{\beta_1}{f C_1} + \frac{5\beta_2}{C_2} \right) \tag{5.47}$$

For a fully differential circuit, this input referred noise is higher by a factor of 2 as shown below.

$$v_{fullydifferential}^2 = \frac{k_B T}{3 f A_{PRE} A_C^2} \left(\frac{\beta_1}{f C_1} + \frac{5\beta_2}{C_2} \right) \tag{5.48}$$

5.2.1.2.4 Thermal Noise of a Two-Stage Miller Compensated Amplifier

In this section, the thermal noise of a two stage amplifier with Miller compensation is analyzed. The analysis for this topology is based on the small signal model shown in figure 5.10. A

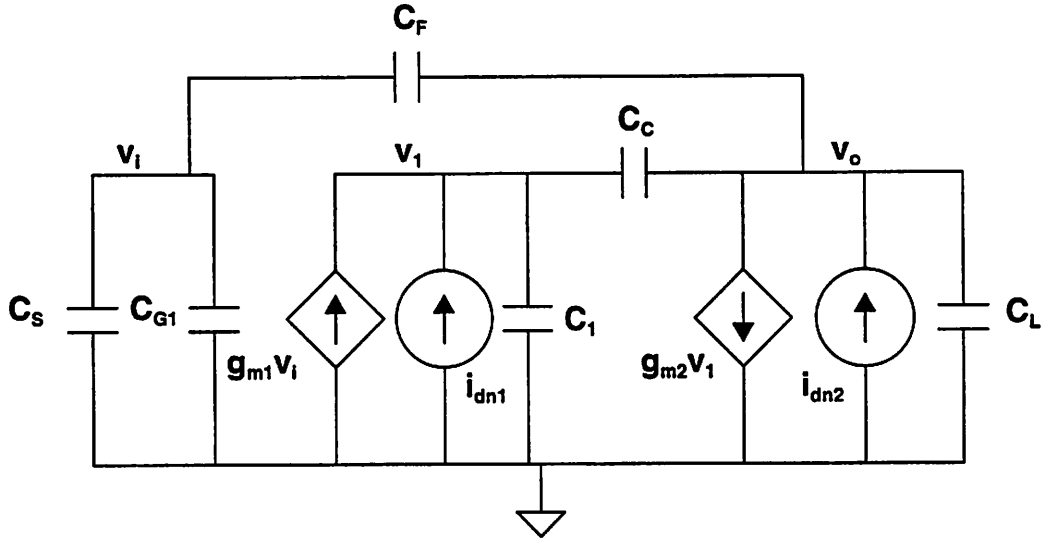


Figure 5.13 Small Signal Model for the Analysis of Thermal Noise in a Two Stage Amplifier with Miller Compensation

nodal analysis of the circuit yields the following s-domain expression for the noise voltage at the output.

$$v_o = \frac{\frac{(C_C s - g_{m2})\, i_{dn1}}{C_T^2} + \frac{(C_1 + C_C)\, s i_{dn2}}{C_T^2}}{s^2 + \frac{(g_{m2} - g_{m1} f)\, C_C}{C_T^2} s + \frac{f g_{m1} g_{m2}}{C_T^2}} \tag{5.49}$$

In the above equation, C_T is given by the following formula.

$$C_T^2 = C_1 C_2 + C_1 C_C + C_2 C_C \tag{5.50}$$

C_2 in the above equation is given by the following formula.

$$C_2 = C_L + C_F (1 - f) \tag{5.51}$$

In this analysis, the two poles are assumed to be real and equal. With this assumption the above equation may be rewritten as follows where ω_n is the magnitude of the pole position.

$$v_o = \frac{\frac{(C_C s - g_{m2})\, i_{dn1} + (C_1 + C_C)\, si_{dn2}}{C_T^2}}{(s+\omega_n)^2} \tag{5.52}$$

The total noise at the output is found by integrating the spectral density of v_o over all frequencies. The spectral densities S_{idn1} and S_{idn2} of the noise currents are given by the following equations.

$$S_{i_{dn1}} d\omega = \frac{4k_B T}{2\pi}\left(\frac{2}{3}\right)\beta_1 g_{m1} d\omega \tag{5.53}$$

$$S_{i_{dn2}} d\omega = \frac{4k_B T}{2\pi}\left(\frac{2}{3}\right)\beta_2 g_{m2} d\omega \tag{5.54}$$

Assuming that the two noise currents are independent, the total noise appearing at the output is given by the following equation.

$$v_o^2 = \frac{\frac{1}{3}k_B T}{C_C}\left\{\frac{\beta_1}{f}\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}+\frac{C_C^4}{C_T^4}}\right)+\beta_2\left(\frac{C_1+C_C}{C_T}\right)^2\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}}\right)\right\} \tag{5.55}$$

The input referred noise voltage is then found by dividing the noise at the output by the closed loop gain A_C with the result shown below.

$$v_o^2 = \frac{\frac{1}{3}k_B T}{A_C^2 C_C}\left\{\frac{\beta_1}{f}\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}+\frac{C_C^4}{C_T^4}}\right)+\beta_2\left(\frac{C_1+C_C}{C_T}\right)^2\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}}\right)\right\} \tag{5.56}$$

For a fully differential circuit, this input referred noise is higher by a factor of 2 as shown below.

$$v_o^2 = \frac{\frac{2}{3}k_B T}{A_C^2 C_C}\left\{\frac{\beta_1}{f}\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}+\frac{C_C^4}{C_T^4}}\right)+\beta_2\left(\frac{C_1+C_C}{C_T}\right)^2\left(1+\sqrt{1+\frac{C_C^2}{C_T^2}}\right)\right\} \tag{5.57}$$

5.2.1.2.5 Thermal Noise of an Amplifier with Ahuja Style Compensation

In this section, the thermal noise of an amplifier with Ahuja style compensation is analyzed. The analysis for this topology is based on the small signal model shown in figure 5.10. A nodal

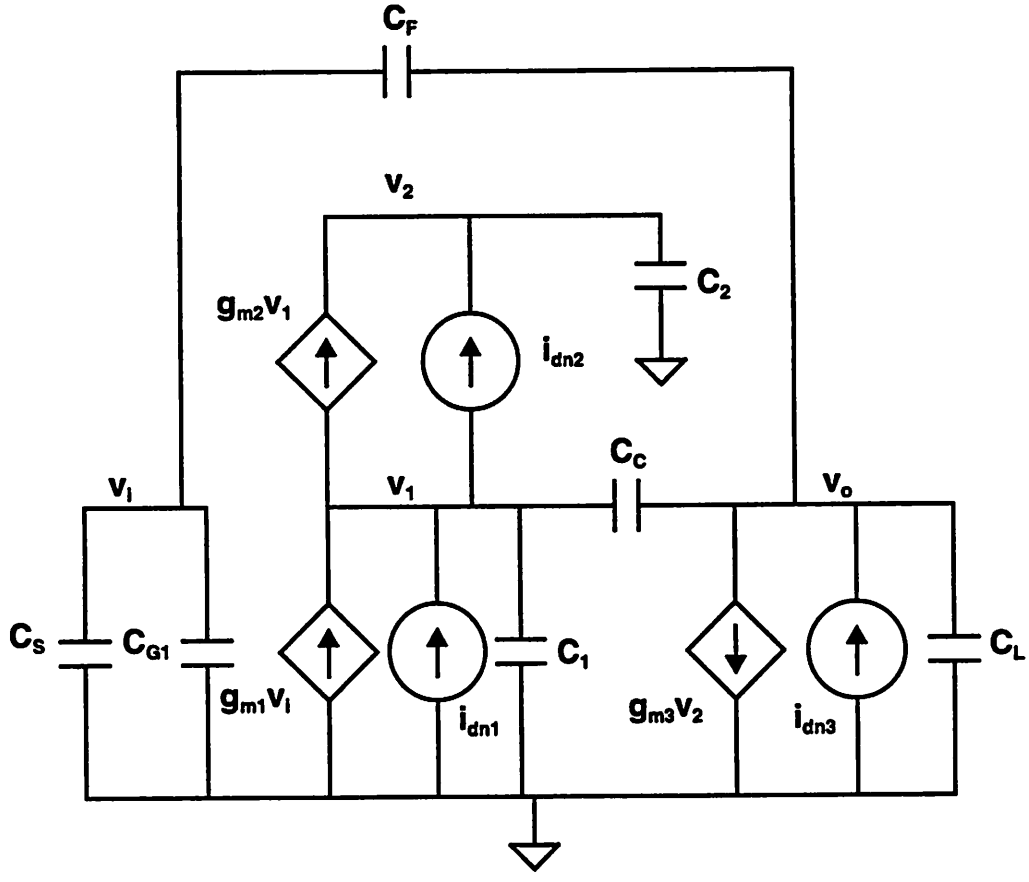


Figure 5.14 Small Signal Model for the Analysis of Thermal Noise in an Amplifier with Ahuja Compensation

analysis of the circuit yields the following s-domain expression for the noise voltage at the output.

$$v_o = \frac{\left(\frac{C_C}{C_T^2}\right)\left(s^2 - \frac{g_{m2}g_{m3}}{C_C C_2}\right)i_{dn1} - \left(\frac{C_C}{C_T^2}\right)s\left(s + \frac{g_{m3}(C_1+C_C)}{C_2 C_C}\right)i_{dn2} + \frac{(C_1+C_C)s}{C_T^2}\left(s + \frac{g_{m2}}{C_1+C_C}\right)i_{dn3}}{D} \quad (5.58)$$

In the above equation, C_T is given by the following formula.

$$C_T^2 = C_1C_3 + C_1C_C + C_3C_C \tag{5.59}$$

C_2 in the above equation is given by the following formula.

$$C_3 = C_L + C_F(1-f) \tag{5.60}$$

In this analysis, the three poles are assumed to be real and equal. Using techniques similar to that used in the previous sections, the fully differential input referred noise contributions, v_{diff1} for stage 1, v_{diff2} for stage 2, and v_{diff3} for stage 3, are given by the following equations.

$$v_{idiff1}^2 = \frac{\frac{3}{2}k_BT}{fA_C^2C_C}\left[1 + \frac{2}{9}\left(\frac{C_C}{C_T}\right)^2 + \frac{1}{9}\left(\frac{C_C}{C_T}\right)^4\right] \tag{5.61}$$

$$v_{idiff2}^2 = \frac{\frac{1}{18}k_BT}{fA_C^2C_C}\left(\frac{g_{m2}}{g_{m1}}\right)\left\{\frac{\left[1 + \left(\frac{C_C}{C_T}\right)^2\right]^2}{\left[1 + \frac{1}{9}\left(\frac{C_C}{C_T}\right)^2\right]^2} + 3\frac{C_C^4}{C_T^4}\right\} \tag{5.62}$$

$$v_{idiff3}^2 = \frac{\frac{1}{18}k_BT}{fA_C^2C_C}\left(\frac{g_{m3}}{g_{m1}}\right)\left[\frac{(C_1+C_C)C_C}{C_T^2}\right]^2\left\{3 + 9\frac{\left[1 + \frac{1}{9}\left(\frac{C_C}{C_T}\right)^2\right]^2}{\left[1 + \left(\frac{C_C}{C_T}\right)^2\right]^2}\right\} \tag{5.63}$$

5.2.2 Optimal Capacitor Sizing in High Resolution-Low Speed Pipelined Analog to Digital Converters

In a high resolution pipelined analog to digital converter where the maximum signal to noise ratio is limited by the thermal noise, the resolution can be improved by increasing the sizes of the sampling capacitors in the front end stages of the pipeline. Because the resolution requirements are relaxed in later stages of the pipeline, power can be saved by making the sampling capacitors in the later stages smaller than they are in the front end stage. This size reduction saves power in two ways. First, reducing the sizes of the sampling capacitors in the later stages of the pipeline reduces the load capacitances presented to the stages that drive them. Therefore, the power dissipated by the stages driving the sampling capacitors can be reduced. Secondly, if the closed loop gain is fixed, the size of the feedback capacitor is reduced as the size of the sampling capacitor is reduced. The feedback capacitor in some stage of the pipeline loads the amplifier in that same stage of the pipeline. Therefore, reducing the size of the feedback capacitor in a given stage can reduce the power dissipation of that stage. Thus, reducing the size of a sampling capacitor at some stage in the pipeline can potentially reduce the power dissipation of that stage and the stage before it.

With this in mind, it is worthwhile to investigate the problem of how much to reduce the size of the sampling capacitor from one stage to the next. The analysis that follows attempts to answer the question, "What is the optimum scaling factor for the capacitors in a high resolution pipeline limited by thermal noise?". The scaling factor s is defined here as the ratio of the sampling capacitance in one stage of the pipeline to the sampling capacitance in the following stage of the pipeline. In order to simplify the analysis, it is assumed here that the scaling factor is constant throughout the pipeline. It is also assumed that the parasitic capacitances and comparator input capacitances are negligible in comparison to the sampling capacitances. Furthermore, it is assumed that the required speed is low enough that the amplifier input capacitance and output capacitance can be neglected. Finally, the speed requirement is assumed to be fixed. In the following sections, some of the above mentioned assumptions are removed in an effort to obtain a more accurate optimization that is applicable over a wider range of design problems.

Figure 5.15 illustrates a pipeline with the features described here. In the figure, s is the scaling factor.

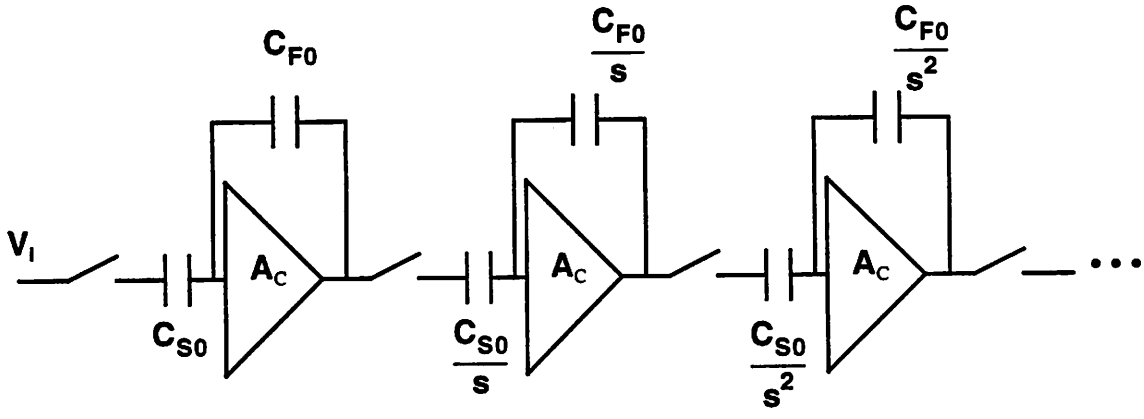


Figure 5.15 Pipelined Analog to Digital Converter with Capacitor Scaling to Save Power

Given the above assumptions, the goal is now to meet some fixed speed requirement and thermal noise requirement with minimal power. Using a large scaling factor can potentially save power because as the scaling factor goes up, the power dissipated by the later stages of the pipeline is reduced relative to the front end. The disadvantage of using a large scaling factor is that the noise contributed by the later stages in the pipeline is increased as the scaling factor goes up. In order to compensate for the increased noise from the later stages of the pipeline, the noise of the front end stage of the pipeline must be reduced by increasing the front end sampling capacitance. Thus, the choice of the optimum scaling factor is an exercise in optimally distributing the thermal noise budget among the pipeline stages.

For the analysis here, the thermal noise is assumed to come from two sources: the sampling switches and the amplifier. It is assumed here that single pole amplifiers are used. Therefore, the input referred noise of a single stage of the pipeline is given by combining equations (5.23) and (5.35). The following formula results.

$$v_{idiffk}^2 = \frac{2k_BT}{fA_C^2C_{Fk}} + \frac{\frac{4}{3}\beta k_BT}{fA_C^2[C_{Lk} + C_{Fk}(1-f)]} \tag{5.64}$$

The load capacitance C_L is assumed to be dominated by the sampling capacitance of the following stage. It is assumed that the sampling capacitor C_{Sk+1} of the following stage is related to the sampling capacitance of the current stage by the following relation.

$$C_{S(k+1)} = \frac{C_{Sk}}{s} = \frac{C_{Sk}}{A_C^y} \tag{5.65}$$

In the above equation, s is the scaling factor. The above equation also introduces a parameter y which will be referred to as the scaling exponent. The scaling factor is related to the gain and scaling exponent by the following formula.

$$s = A_C^y \tag{5.66}$$

Therefore, the load capacitance is given by the following formula.

$$C_L = A_C C_{Fk+1} = A_C^{1-y} C_{Fk} \tag{5.67}$$

It is then assumed that y is close enough to 1 that C_L is comparable to C_F. Assuming $f << 1$, equation (5.64) then may be rewritten as follows.

$$v_{idiff}^2 = \frac{2k_B T}{f A_C^2 C_{Fk}} (1 + \frac{1}{3}\beta) \tag{5.68}$$

The contribution of stage k is then scaled down by the gains of all the stages that precede it. Therefore,

$$v_{idiffk}^2 = \frac{2k_B T}{f A_C C_{Fk} A_C^{2k}} (1 + \frac{1}{3}\beta) \tag{5.69}$$

The total noise in the pipeline is then determined by summing the contribution of all of the stages as shown below. N_s is the number of stages in the pipeline.

$$v_{idiff}^2 = \sum_{k=0}^{N_s-1} v_{idffk}^2 \approx \sum_{k=0}^{\infty} v_{idiffk}^2 \tag{5.70}$$

Equations (5.69) and (5.70) are now combined with the following result.

$$v_{idiff}^2 = \sum_{k=0}^{\infty} \left(\frac{2k_BT}{fA_C^2 C_{F0} A_C^{(2-y)k}}\right)\left(1+\frac{1}{3}\beta\right) = \left(1+\frac{1}{3}\beta\right)\frac{2k_BT}{fA_C^2 C_{F0}}\left(\frac{1}{1-A^{y-2}}\right) \tag{5.71}$$

The above equation relates the noise to the front end feedback capacitance C_{F0} and to the scaling exponent y. The next step in the process of determining the optimum scaling factor is to relate the power dissipation to the front end feedback capacitance and scaling factor. To begin, single stage amplifiers are assumed. The time constant of the settling of the amplifier is related to the transconductance g_m of the input device, the feedback factor f, the load capacitance C_L, and the feedback capacitance C_F as shown below.

$$\tau = \frac{C_L + C_F(1-f)}{fg_m} \tag{5.72}$$

Now, equation (5.67) is substituted into the above equation, and it is solved for the transconductance with the result shown below.

$$g_{mk} = \frac{[C_{Lk} + C_{Fk}(1-f)]}{f\tau} = \frac{1}{f\tau}\left[A_C^{1-y}C_{Fk} + C_{Fk}(1-f)\right] \tag{5.73}$$

Now, the total transconductance required of all the pipeline stages to meet the noise and speed requirements is calculated by summing over k as shown below. It is useful to find the total transconductance required because it is assumed that the total power dissipation is proportional to the total transconductance. It is assumed that the feedback factor is the same for all of the stages in the pipeline.

$$g_{mtotal} = \frac{1}{f\tau}(A_C^{1-y} + 1 - f)\sum_{k=0}^{N_s-1} C_{Fk} = \frac{1}{f\tau}(A_C^{1-y} + 1 - f)C_{F0}\sum_{k=0}^{N_s-1}\frac{1}{A_C^{yk}} \tag{5.74}$$

$$g_{mtotal} = \frac{1}{f\tau}\left(A_C^{1-y} + 1 - f\right) C_{F0}\left(\frac{1}{1-\frac{1}{A_C^y}}\right) \tag{5.75}$$

Now equation (5.71) is used to solve for C_{F0}, and this is substituted into the above equation.

$$C_{F0} = \frac{2k_BT}{fA_C^2 v_{idiff}^2}\left(1+\frac{1}{3}\beta\right)\left(\frac{1}{1-A_C^{y-2}}\right) \tag{5.76}$$

$$g_{mtotal} = \frac{1}{f\tau}\left(\frac{2k_BT}{fA_C^2 v_{idiff}^2}\right)\left(1+\frac{1}{3}\beta\right)\left(A_C^{1-y}+1-f\right)\left(\frac{1}{1-\frac{1}{A_C^y}}\right)\left(\frac{1}{1-A_C^{y-2}}\right) \tag{5.77}$$

The above formula is then differentiated with respect to y. To find the optimum value, the derivative is set to 0. The result for the optimum value of the scaling factor is shown below.

$$s_{opt} = A_C^{y_{opt}} = \left(\frac{A_C}{1-f}\right)\left[-1+\sqrt{1+(1-f)^2+A_C(1-f)+\frac{1}{A_C}(1-f)}\right] \tag{5.78}$$

If it is assumed that $f=1/A_C$, the above equation simplifies to the following.

$$s_{opt} = \left(\frac{A_C^2}{A_C-1}\right)\left(-1+\sqrt{1+A_C-\frac{1}{A_C}}\right) \tag{5.79}$$

The above equation is tabulated below and graphed in figure 5.17. Also, the normalized power dissipation given by equation (5.77) is graphed in figure 5.16.

Table 5.2 Optimum Scaling Factor Versus Closed Loop Gain for High Resolution - Low Speed Pipelined Analog to Digital Converters

A_C	y_{opt}	s_{opt}
1	-	1
1.5	1.148	1.593
2	1.217	2.325
3	1.289	4.117
4	1.327	6.290
8	1.39	18.09
∞	1.5	∞

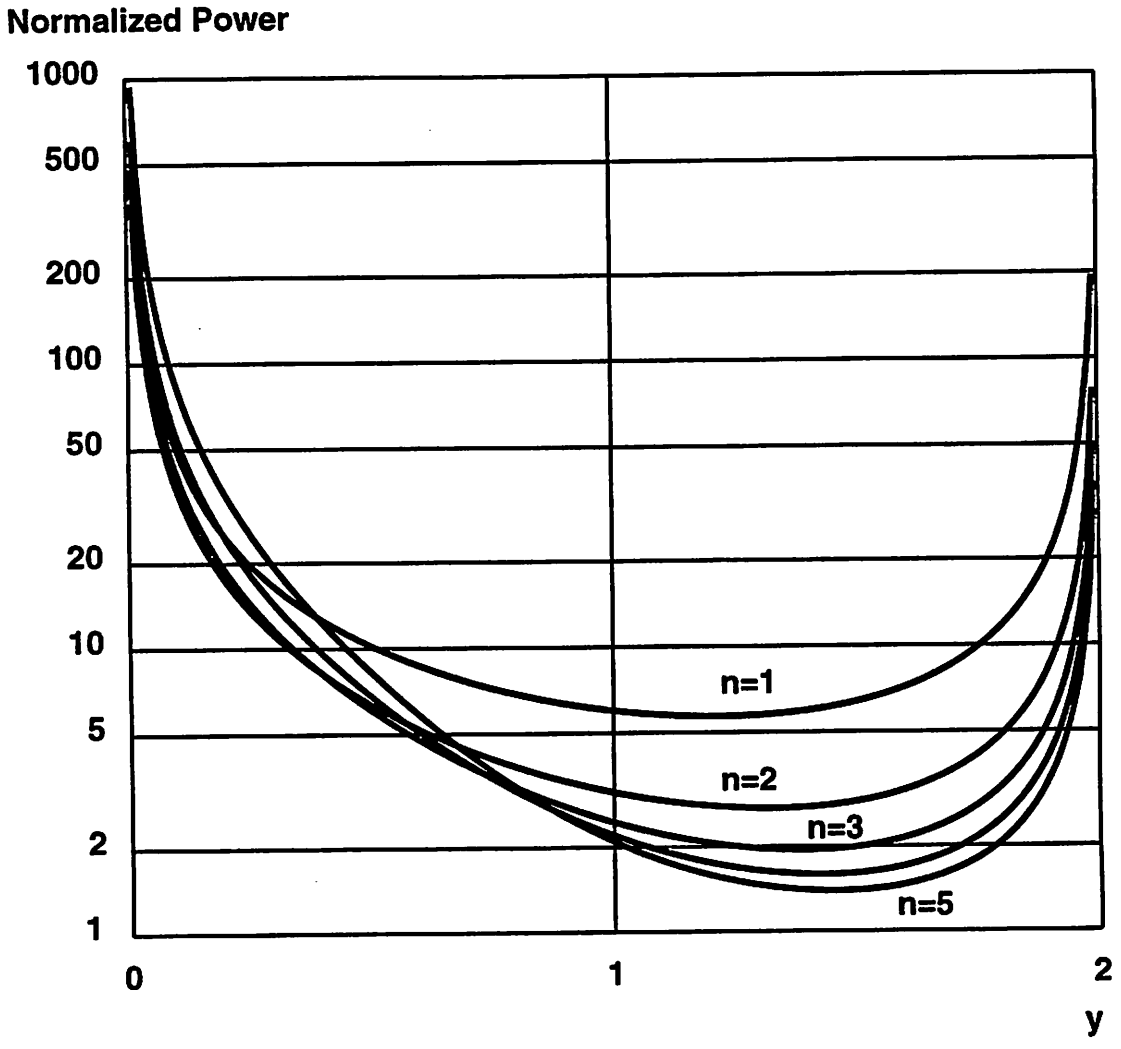


Figure 5.16 Normalized Power Versus Scaling Factor for First Order Scaling Analysis of a Pipelined ADC

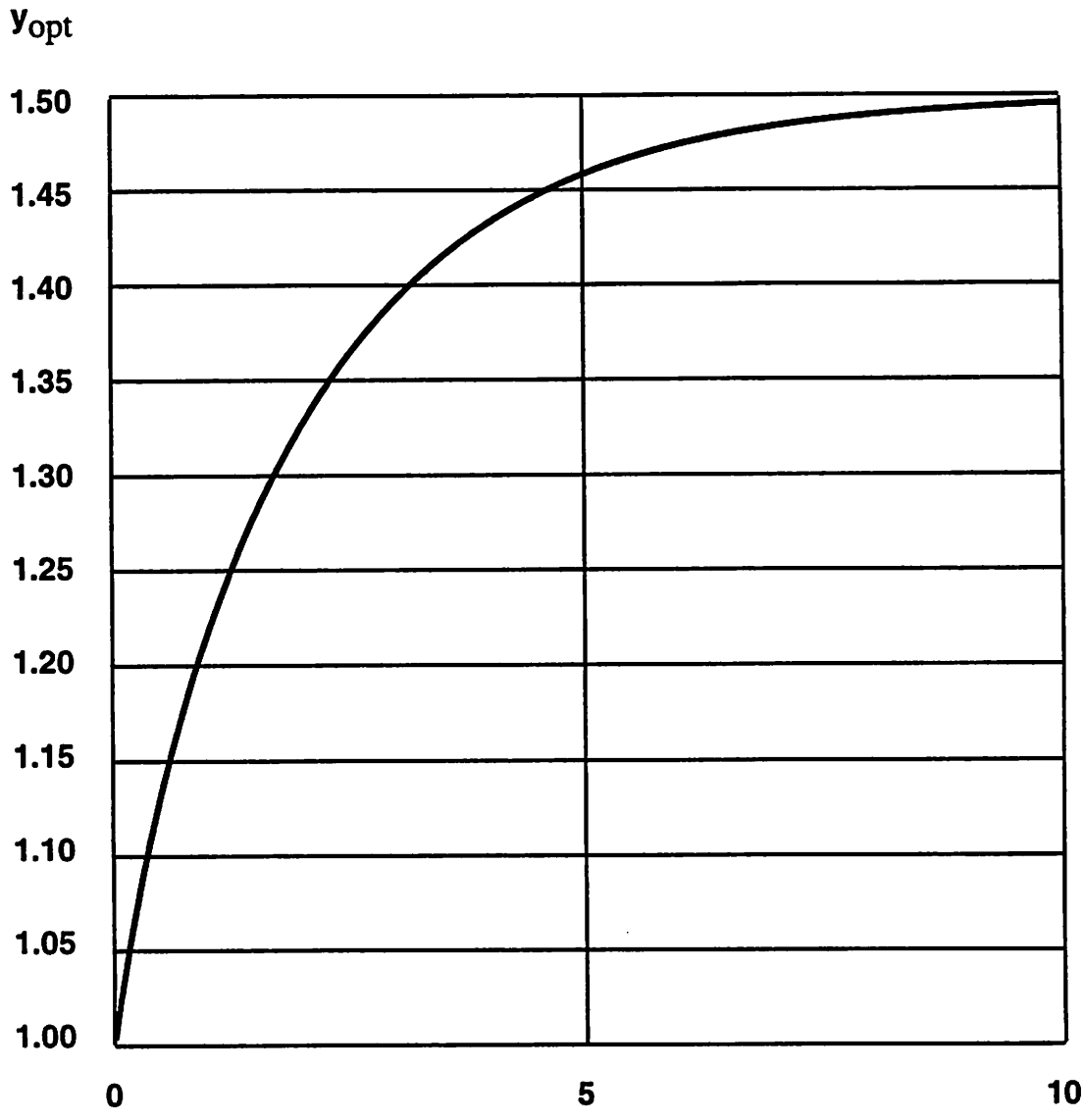


Figure 5.17 Optimum Scaling Exponent versus Per Stage Resolution

Note that the ratio of the optimum scaling factor to the closed loop gain goes up as the closed loop gain goes up. Also note that according to equations (5.77) and (5.79), increasing the inter-stage gain always reduces the power. In reality, this only holds true as long as the speed requirement is well below the limit of the technology. When the speed requirement approaches the speed limit of the technology, the optimum value for the gain is reduced because at high speeds it is harder to design high gain amplifiers than low gain amplifiers. Furthermore, at some speeds the

resolution and speed design requirements can only be met if low gain amplifiers are used. The optimum value for the scaling factor is also reduced when the speed limit of the technology is approached because using feedback capacitors that are too small reduces the feedback factor, and reducing the feedback factor reduces the speed.

The above analysis also assumes that the loading effects of parasitic capacitances are negligible in comparison to the sampling capacitors. This is true only for cases involving very high resolution or very small signal swings. In many practical cases, parasitic capacitances are very significant. These capacitances tend to reduce the optimum value for the interstage gain and the optimum value for the scaling factor.

5.2.3 Optimal Closed Loop Gain of Interstage Gain Amplifiers in High Resolution-High Speed Pipelined Analog to Digital Converters

In the previous section, the optimum scaling factor to minimize the power dissipation of a pipelined ADC with fixed noise and speed requirements was determined as a function of the closed loop gain. This analysis was performed for the case when the speed requirement is far below the technology limit. In this section, the same problem will be treated for the high speed case. High speed means that the amplifier is large enough that the input capacitance and output capacitance of the amplifier cannot be neglected. The other parasitic capacitances due to interconnect and comparators are still assumed to be negligible. The effect of including these capacitances will be addressed in t he analysis of the following section.

This analysis differs from that of the previous section in that a capacitance αC_{Gk} proportional to the size of the amplifier is assumed to exist at the output in parallel with the load capacitance. Therefore, equation [5.72] from the previous section is modified to include this capacitance as shown below.

$$\tau = \frac{C_{Lk} + C_{Fk}(1-f) + \alpha C_{Gk}}{f g_{mk}} \tag{5.80}$$

It is assumed here that the current density is fixed. Therefore, the transconductance g_{mk} is proportional to the input capacitance C_{Gk} of the amplifier as shown below.

$$g_{mk} = \omega_T C_{Gk} \tag{5.81}$$

Now the sampling capacitor is assumed to scale in the same way as described in the previous section. Therefore, C_{Lk} is given by equation (5.67). Substituting this and equation (5.81) into equation (5.72) yields the following equation for the time constant.

$$\tau = \frac{(A_C^{1-y} + 1 - f)\, C_{Fk} + \alpha C_{Gk}}{f \omega_T C_{Gk}} \tag{5.82}$$

It is assumed here that the feedback capacitor is used for sampling as well as feedback. Therefore, the feedback factor is related to the closed loop gain and the input capacitance by the following relation.

$$f = \frac{1}{A_C + \dfrac{C_{Gk}}{C_{Fk}}} \tag{5.83}$$

The above two equations can be used to solve for the feedback factor as a function of the closed loop gain and the speed with the result shown below. Note that the feedback factor is the same for all of the stages in the pipeline.

$$f = \frac{[\omega_T \tau + \alpha A_C - (A_C^{1-y} + 1)]}{2(\omega_T \tau A_C - 1)} \left\{ 1 + \sqrt{1 - \frac{4\alpha(\omega_T \tau A_C - 1)}{[\omega_T \tau + \alpha A_C - (A_C^{1-y} + 1)]^2}} \right\} \tag{5.84}$$

Equation (5.64) from the previous section is also modified by the output capacitance of the amplifier as shown below.

$$v_{idiffk}^2 = \frac{2k_B T}{f A_C^2 C_{Fk}} + \frac{\frac{4}{3}\beta k_B T}{f A_C^2 [C_{Lk} + C_{Fk}(1-f) + \alpha C_{Gk}]} \tag{5.85}$$

The total noise power in the pipeline can be determined by summing equation (5.64) over all the pipeline stages and substituting in equations (5.67) and (5.83). The contribution of a given

stage of the pipeline to the noise budget is reduced by the square of all the gains that precede that stage in the pipeline. The result is shown below.

$$v_{idiff}^2 = \frac{2k_BT}{fA_C^2}\left\{1+\frac{\frac{2}{3}\beta}{(A_C^{1-y}+1-f)\left(1+\frac{\alpha}{f\omega_T\tau-\alpha}\right)}\right\}\sum_{k=0}^{N_s-1}\frac{1}{A_C^{2k}C_{Fk}} \tag{5.86}$$

Note that the feedback factor can be taken outside the sum because it is the same for each stage in the pipeline.

It is assumed here that the feedback capacitance always scales by the same amount from one stage to the next. Therefore, the feedback capacitance C_{Fk} of stage k can be expressed as shown below.

$$C_{Fk} = \frac{C_{F0}}{A_C^{yk}} \tag{5.87}$$

This value for C_{Fk} is then substituted into equation (5.86). The sum can be simplified by letting it go to infinity. This approximation is very good. The resulting noise power is then given by the following equation.

$$v_{idiff}^2 = \frac{2k_BT}{fA_C^2C_{F0}}\left\{1+\frac{\frac{2}{3}\beta}{(A_C^{1-y}+1-f)\left(1+\frac{\alpha}{f\omega_T\tau-\alpha}\right)}\right\}\left\{\frac{1}{1-A_C^{y-2}}\right\} \tag{5.88}$$

This equation is now rearranged to express the feedback capacitance of the first stage as a function of the interstage gain, the scaling exponent, and the noise required.

$$C_{F0} = \frac{2k_BT}{fA_C^2v_{idiff}^2}\left\{1+\frac{\frac{2}{3}\beta}{(A_C^{1-y}+1-f)\left(1+\frac{\alpha}{f\omega_T\tau-\alpha}\right)}\right\}\left\{\frac{1}{1-A_C^{y-2}}\right\} \tag{5.89}$$

Because the current density is assumed to be fixed, the total power dissipated in the pipeline is proportional to the sum of the input capacitances C_{Gk} of all of the amplifiers. Therefore, it is this sum that needs to be optimized. The input capacitance can be related to the feedback factor, which has already been calculated, by rearranging equation (5.82) as shown below.

$$C_{Gk} = \frac{(A_C^{1-y} + 1 - f)\, C_{Fk}}{f\omega_T\tau - \alpha} \tag{5.90}$$

Using the above equation and equation (5.87), the total capacitance may be expressed as follows.

$$C_{Gtotal} = \sum_{k=0}^{N_s-1} C_{Gk} = C_{F0}\frac{(A_C^{1-y} + 1 - f)}{(f\omega_T\tau - \alpha)} \sum_{k=0}^{N_s-1} A_C^{-yk} \tag{5.91}$$

Again, the upper limit of the sum is allowed to go to infinity to simplify the result to the following expression.

$$C_{Gtotal} = \left(\frac{2k_BT}{fA_C^2v_{idiff}^2}\right)\left(\frac{1}{1-A_C^{y-2}}\right)\left(\frac{1}{1-A_C^{-y}}\right)\left\{\left(1+\frac{\frac{2}{3}\beta}{\omega_T\tau}\right)\frac{1}{f} - A_C\right\} \tag{5.92}$$

Thus, the total power has been expressed as a function of the noise, the closed loop gain, the scaling exponent, and the speed. The feedback factor also appears in the above equation, but that has already been calculated as a function of the closed loop gain, the scaling exponent, and the speed. It is worth noting that a dimensionless figure of merit can be obtained from the above expression. This figure of merit F, shown below, is the ratio of kT/C_{Gtotal} to the noise. This figure of merit indicates the efficiency with which a pipelined ADC design meets the noise and speed specifications.

$$F = \frac{\left(\frac{k_BT}{C_{Gtotal}}\right)}{v_{idiff}^2} = \frac{fA_C^2(1-A_C^{y-2})\,(1-A_C^{-y})}{2\left\{\left(1+\frac{\frac{2}{3}\beta}{\omega_T\tau}\right)\frac{1}{f} - A_C\right\}} \tag{5.93}$$

This expression is then used to find the optimum scaling exponent. The results are shown in figures 5.18 and 5.19. From these graphs, it can be seen that the optimum scaling exponent

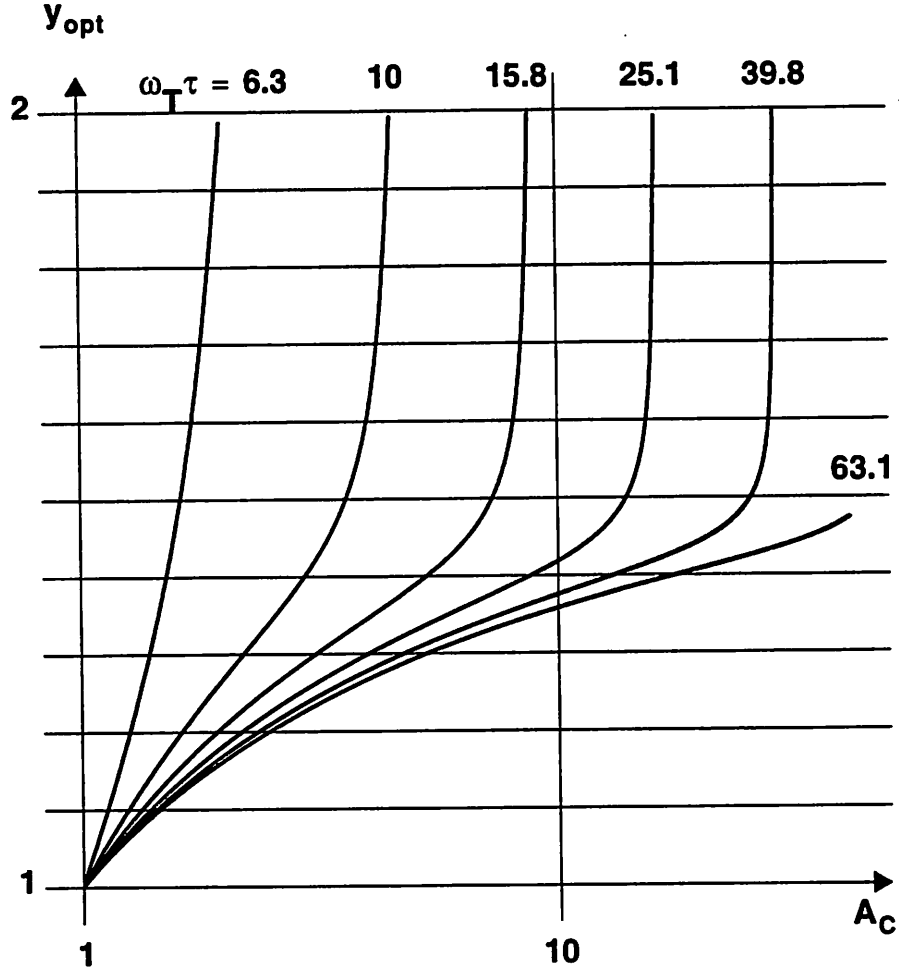


Figure 5.18 Optimum Scaling Exponent Versus Closed Loop Gain for a High Resolution Pipelined Analog to Digital Converter, $\beta = 1$

increases as the closed loop gain increases and as the speed increases. This trend can be understood in the following way. As the scaling exponent is increased, the ratio of the load capacitance to the feedback capacitance is reduced. This reduction increases the maximum speed attainable from the amplifier. As the speed requirement approaches the maximum allowable by the technology, reducing the ratio of load capacitance to the feedback capacitance may be the only way to

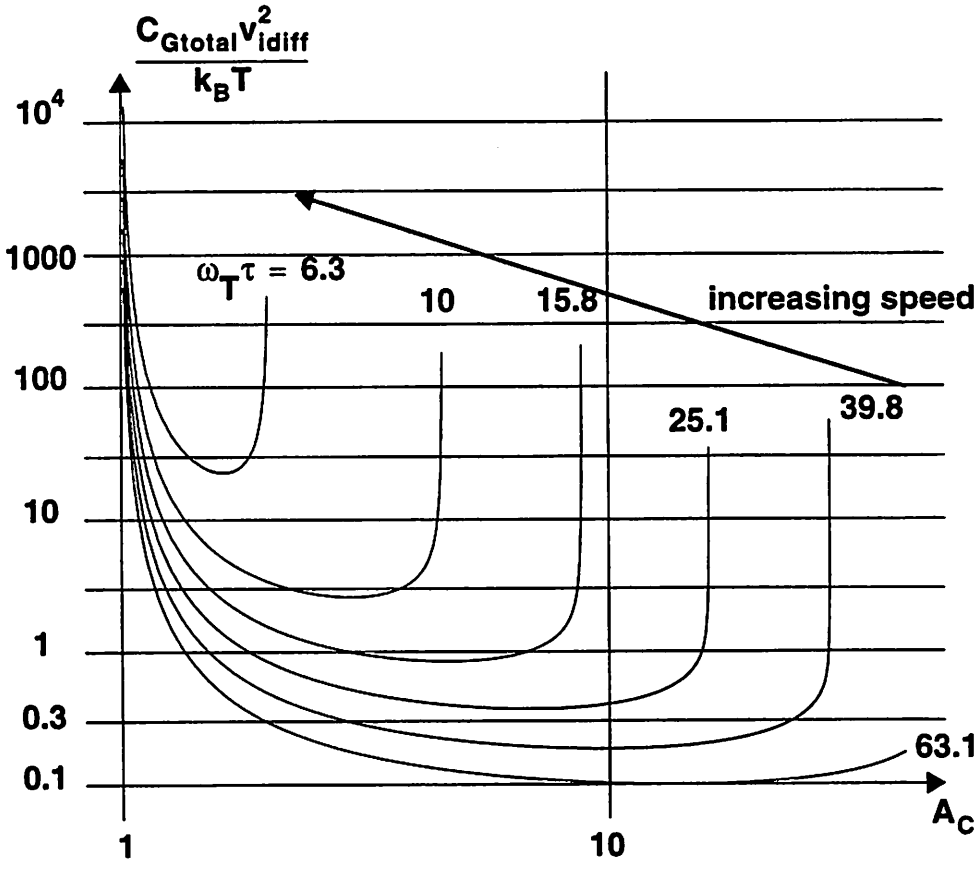


Figure 5.19 Noise versus Gain for a Pipelined Analog to Digital Converter, $\beta = 1$

meet the specification. For this reason, the optimum scaling factor increases with increasing closed loop gain and increasing speed.

The maximum scaling exponent of 2 corresponds to the case where all stages of the pipeline contribute equally to the noise. This case is highly nonoptimal unless the speed requirement pushes the technology close to the limit of its performance capability.

From figure 5.19 it can be seen that an optimum value exists for the closed loop gain. This optimum value for the closed loop gain is plotted in figure 5.20. The optimum valued for the

closed loop gain is seen to decrease as the speed increases. This makes sense because as the speed limit of a technology is approached, the only way to satisfy the speed requirement is to reduce the gain. The corresponding optimum scaling exponent is plotted in figure 5.21, and the corresponding total amplifier input capacitance is plotted in figure 5.22.

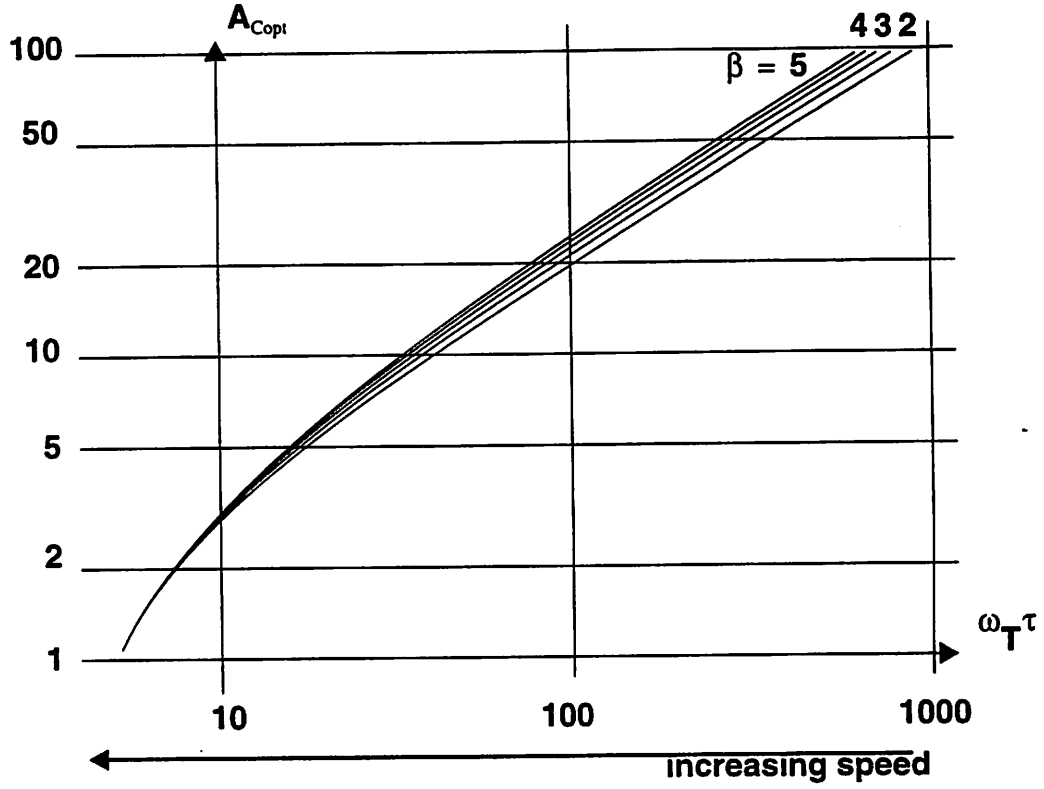


Figure 5.20 Optimum Closed Loop Gain of a High Resolution Pipeline Neglecting Parasitics due to Interconnect and Comparator Inputs

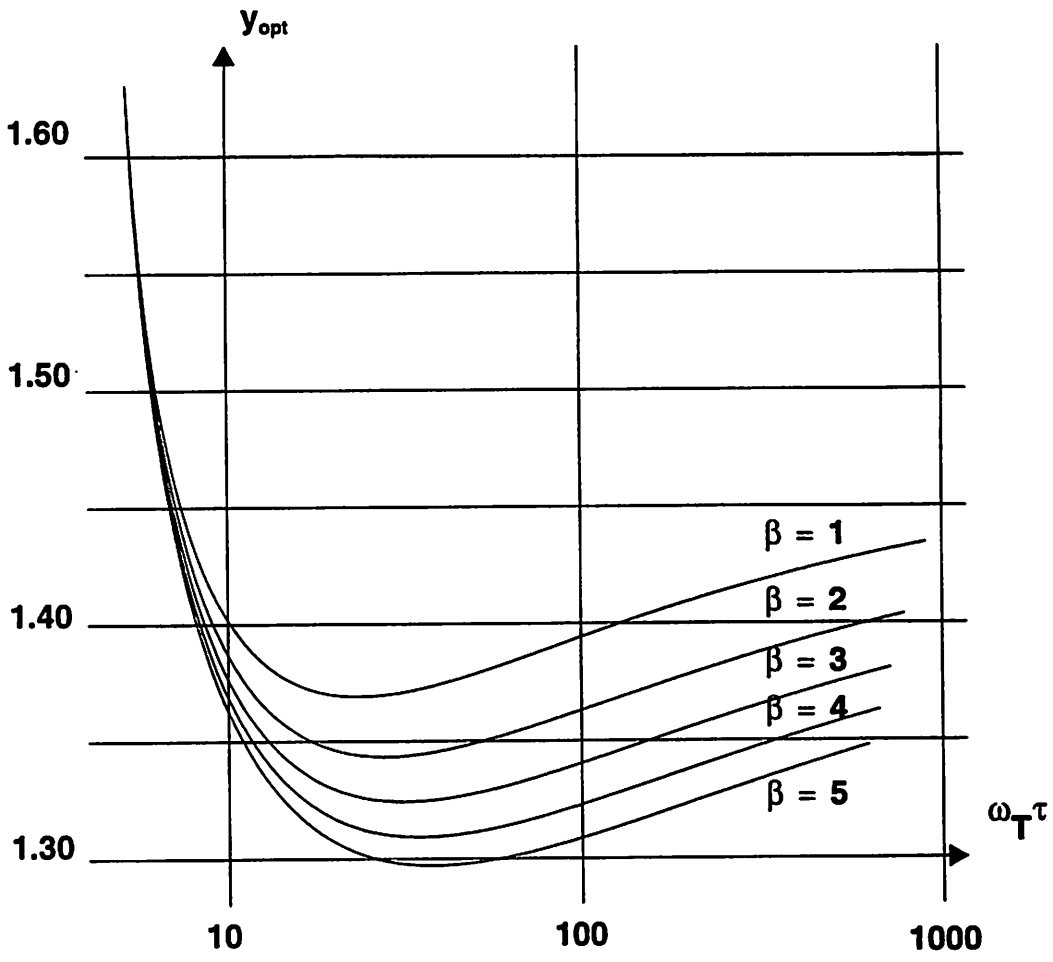


Figure 5.21 Optimum Scaling Exponent for a High Resolution Pipelined Analog to Digital Converter Neglecting Parasitics Due to Interconnect and Comparator Inputs

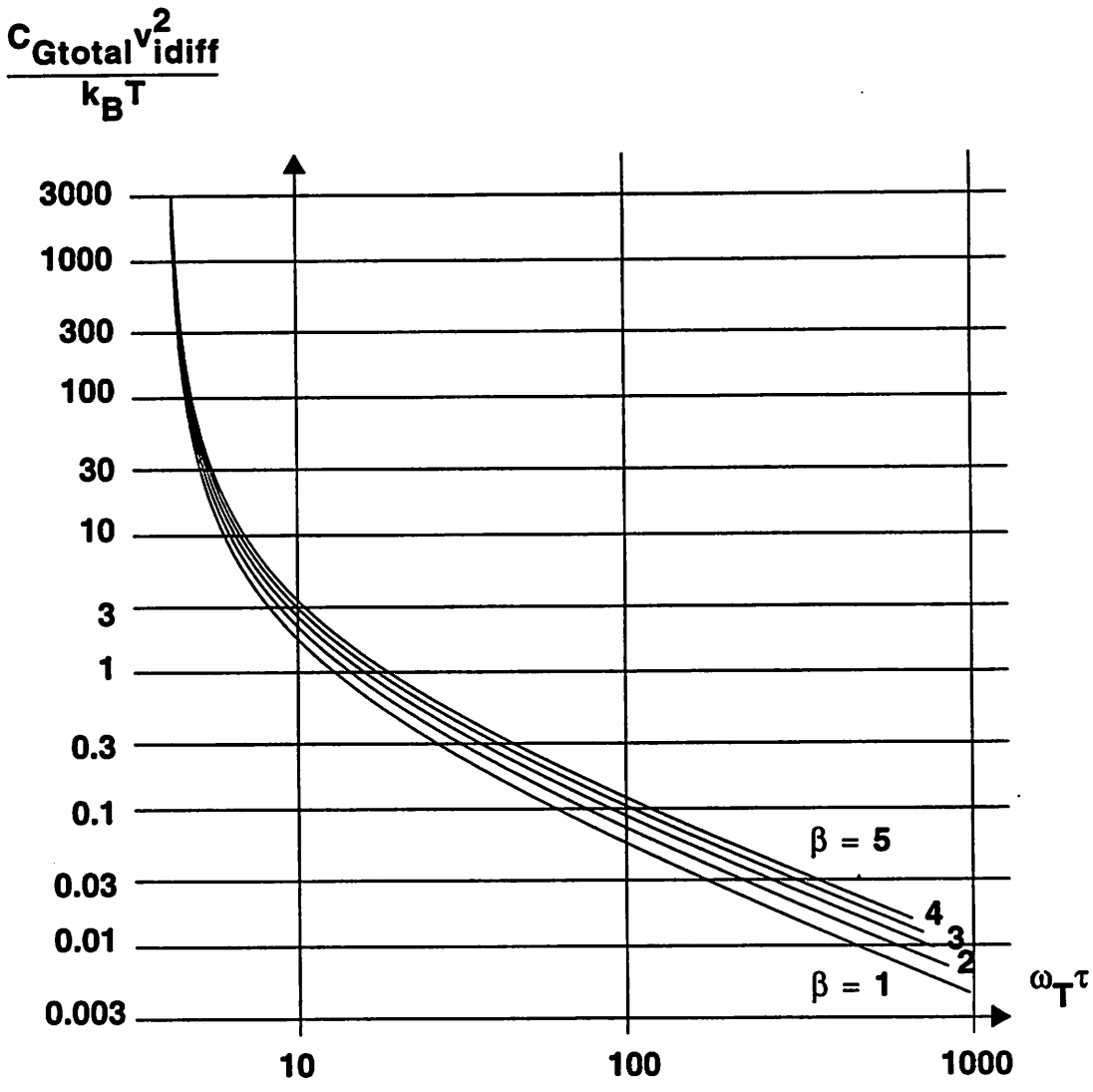


Figure 5.22 Total Capacitance of a High Resolution Pipelined Analog to Digital Converter Neglecting Parasitics Due to Interconnect and Comparator Inputs

From these graphs, it can be seen that the optimum interstage gain steadily increases as the speed requirement is reduced. This trend displays the trade-off between speed, which favors low interstage gain, and thermal noise, which favors high interstage gain.

5.2.4 Optimum Closed Loop Gain of Interstage Gain Amplifiers in a Pipelined ADC with Parasitics Included

In the previous sections, the analog to digital converter was assumed to be of a resolution high enough that the parasitic loading capacitances due to interconnect and comparator inputs could be neglected. In other words, all of the capacitors were assumed to be proportional to the size of the sampling capacitor. In reality, the amplifiers must drive a capacitive component that does not change from one stage to the next. This fixed parasitic capacitance sets the lower limit on the size of the sampling capacitors in the tail end of the pipeline. This fixed parasitic capacitance prevents the sampling capacitor from being scaled down as it would be if the parasitic capacitance were proportional to either the sampling capacitance or the amplifier size. As a result, the pipeline may be thought of to contain two sections; a noise limited section at the front end, and a parasitic limited section at the tail end. Therefore, it is optimal to have a nonconstant scaling factor that is large at the front end of the pipeline, and small at the tail end of the pipeline. In this section, the optimization of the scaling factor and the interstage gain given these constraints is considered.

The analysis in this section proceeds in much the same way as that in the previous section with the exception that equation [5.67] for C_{Lk} is modified to include the fixed parasitic capacitance as shown below.

$$C_{Lk} = A_C C_{Fk+1} + C_P \tag{5.94}$$

$$C_P = 2(A_C - 1)C_{comp} + C_{INT} \tag{5.95}$$

In the above equation the fixed parasitic capacitance C_P is assumed to be composed of interconnect capacitance $C_{INT}$and comparator input capacitances C_{comp}.

Equation (5.87) relating the feedback capacitance C_{Fk} of stage k to the feedback capacitance C_{F0} of stage 0 is also modified as shown below.

$$C_{Fk} = \frac{C_{F0} - C_{FF}}{A_C^{yk}} + C_{FF} \tag{5.96}$$

When C_{FF}=0, the above equation degenerates to equation (5.87). C_{FF} is the optimum feedback capacitance for the interstage amplifier when thermal noise is not an issue. This capacitance is

determined by the analysis of section Figure 5.1 and is given by equation (5.8). The above equation causes the feedback capacitance to scale down as A_C^y at the front end of the pipeline and to remain nearly constant at the tail end of the pipeline.

Given these assumptions, an analysis similar to that discussed in the previous section leads to the following relation for the feedback factor.

$$\frac{1}{f} = \frac{(\omega_T\tau + \alpha A_C - 1 - A_C R_1 - R_2)}{2\alpha}\left\{1 - \sqrt{1 - \frac{4\alpha(A_C\omega_T\tau - 1)}{(\omega_T\tau + \alpha A_C - 1 - A_C R_1 - R_2)^2}}\right\} \quad (5.97)$$

In the above equation, the quantities R_1 and R_2 have been introduced and are given by the following relations.

$$R_1 = \frac{C_{Fk+1}}{C_{Fk}} = \frac{A_C^{-y}\left(\frac{C_{F0}}{C_{FF}} - 1\right) + A_C^{yk}}{\frac{C_{F0}}{C_{FF}} - 1 + A_C^{yk}} \quad (5.98)$$

$$R_2 = \frac{C_P}{C_{Fk}} \quad (5.99)$$

Note that the feedback factor is now dependent on the value of the stage index k. In the previous analysis the feedback factor was independent of k. As a result of this variable feedback factor, the analysis cannot proceed to the relatively simple relation of the previous section. Therefore, the rest of this analysis was performed by computer. In the results shown below, the amplifiers are assumed to be fully differential. The peak voltage swing is assumed to be 5V, and β is assumed to be 1. The value for C_{GT} is the single ended total amplifier input capacitance.

the other hand, when thermal noise is not an important issue, a high per stage resolution is not optimal because of the cost of adding a large number of comparators and the cost of achieving amplifiers with high closed loop gain.

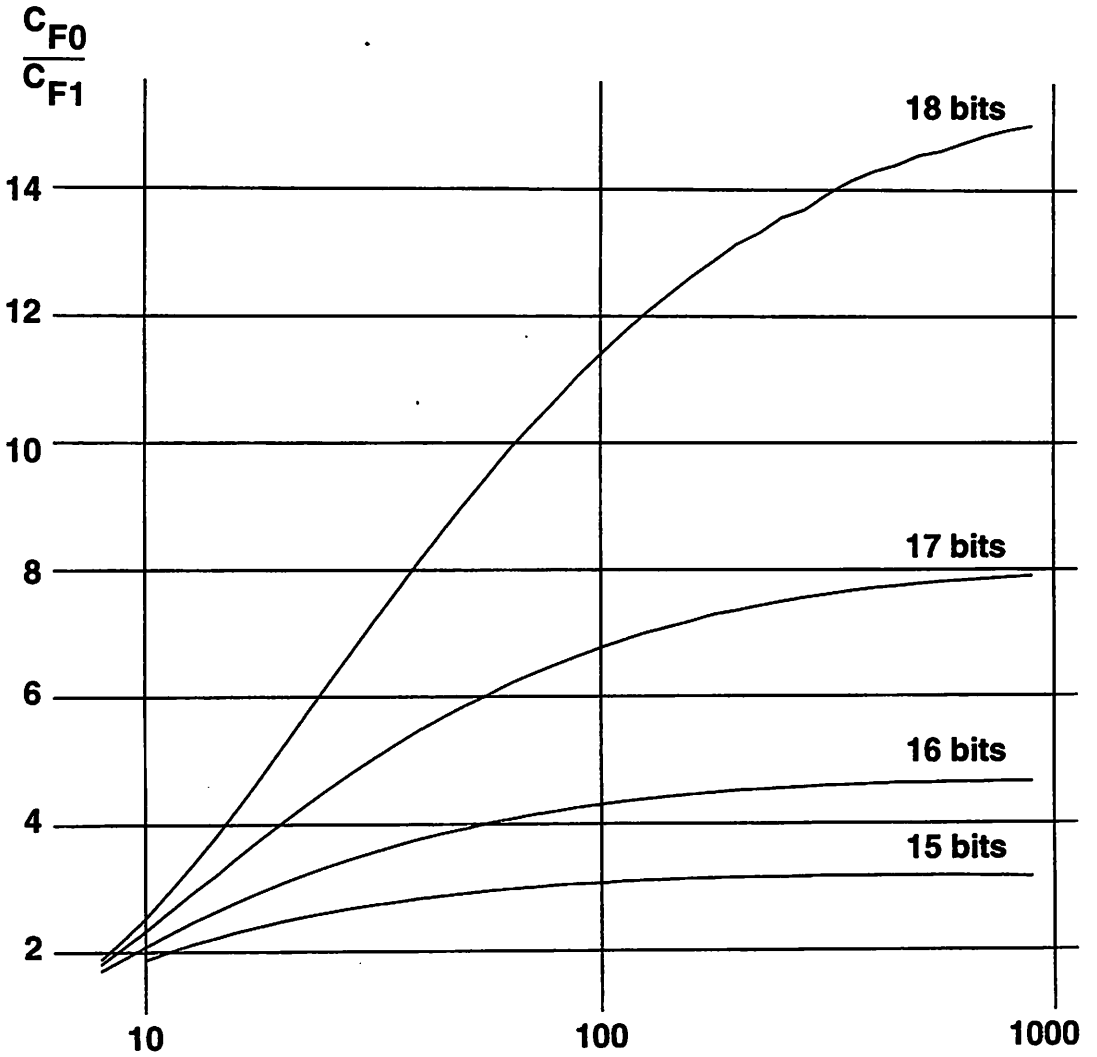


Figure 5.24 Front End Scaling Factor (Ratio of Feedback Capacitance of Stage 0 to Feedback Capacitance of Stage 1) Versus Time Constant

Figure 5.25 shows a plot of the optimum front end scaling exponent versus the time constant for a number of different resolution requirements. This graph does not show well defined trends. The reason for this fact is that there are two primary reasons for increasing the size of the sampling

capacitor. One reason to increase its size is to improve the feedback factor to help improve the speed. A second reason to use a large sampling capacitor is to reduce the amount of thermal noise sampled onto the capacitor. If the speed is high enough, thermal noise is not an issue because the sampling capacitor must be very large anyway in order to meet the speed requirement. In this case, the gain tends to be rather low, and the front end scaling exponent is high to help reduce the load capacitance. On the other hand, if the speed requirement is relaxed the sizes of the sampling capacitors are determined by thermal noise. In this case, rapid scaling is desirable, but this is achieved by increasing the interstage gain. Therefore, the scaling exponent does not need to be especially large.

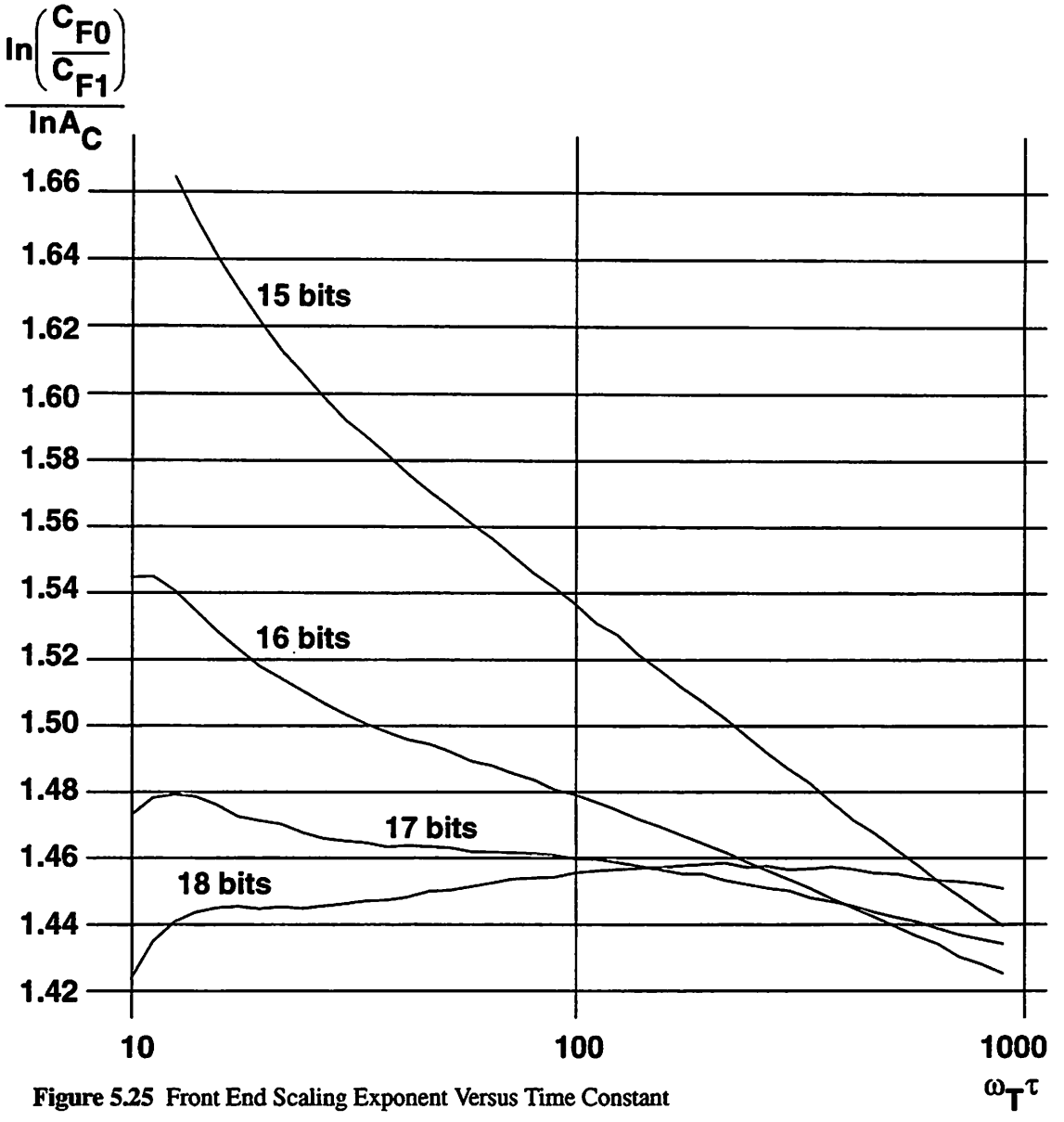


Figure 5.25 Front End Scaling Exponent Versus Time Constant

Figure 5.26 shows a plot of the front end feedback capacitance as a function of speed. As the speed is increased the optimum front end feedback capacitance also increases because a higher feedback factor is desirable to meet the speed requirement. In the high speed region of the graph where the size of the feedback capacitor is not constant, thermal noise is not an important factor in determining the size of the capacitor. However, the curves flatten out in the low speed portion of

the graph because thermal noise is more important than speed in determining the size of the feedback capacitor in the low speed case.

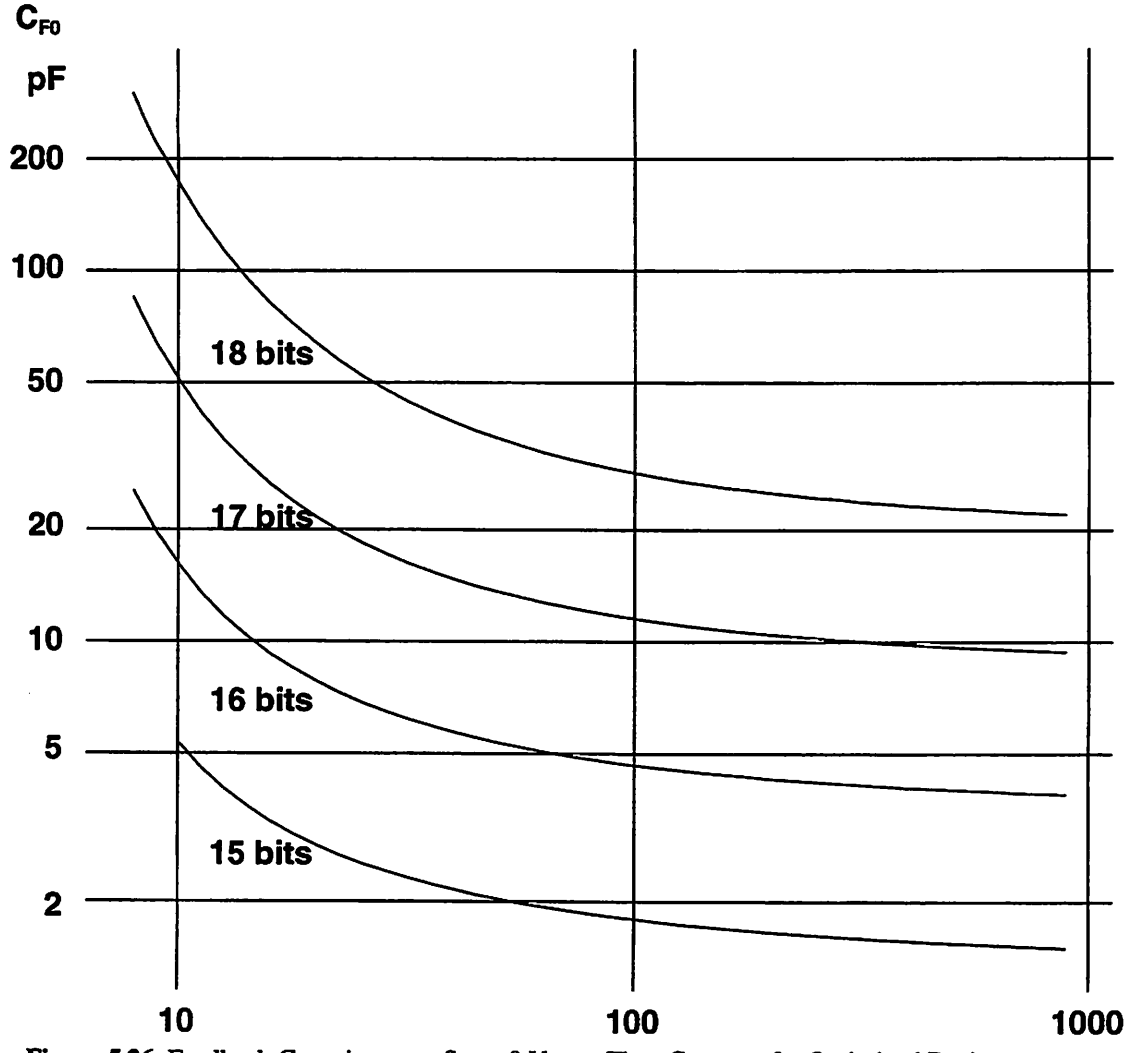


Figure 5.26 Feedback Capacitance at Stage 0 Versus Time Constant for Optimized Design

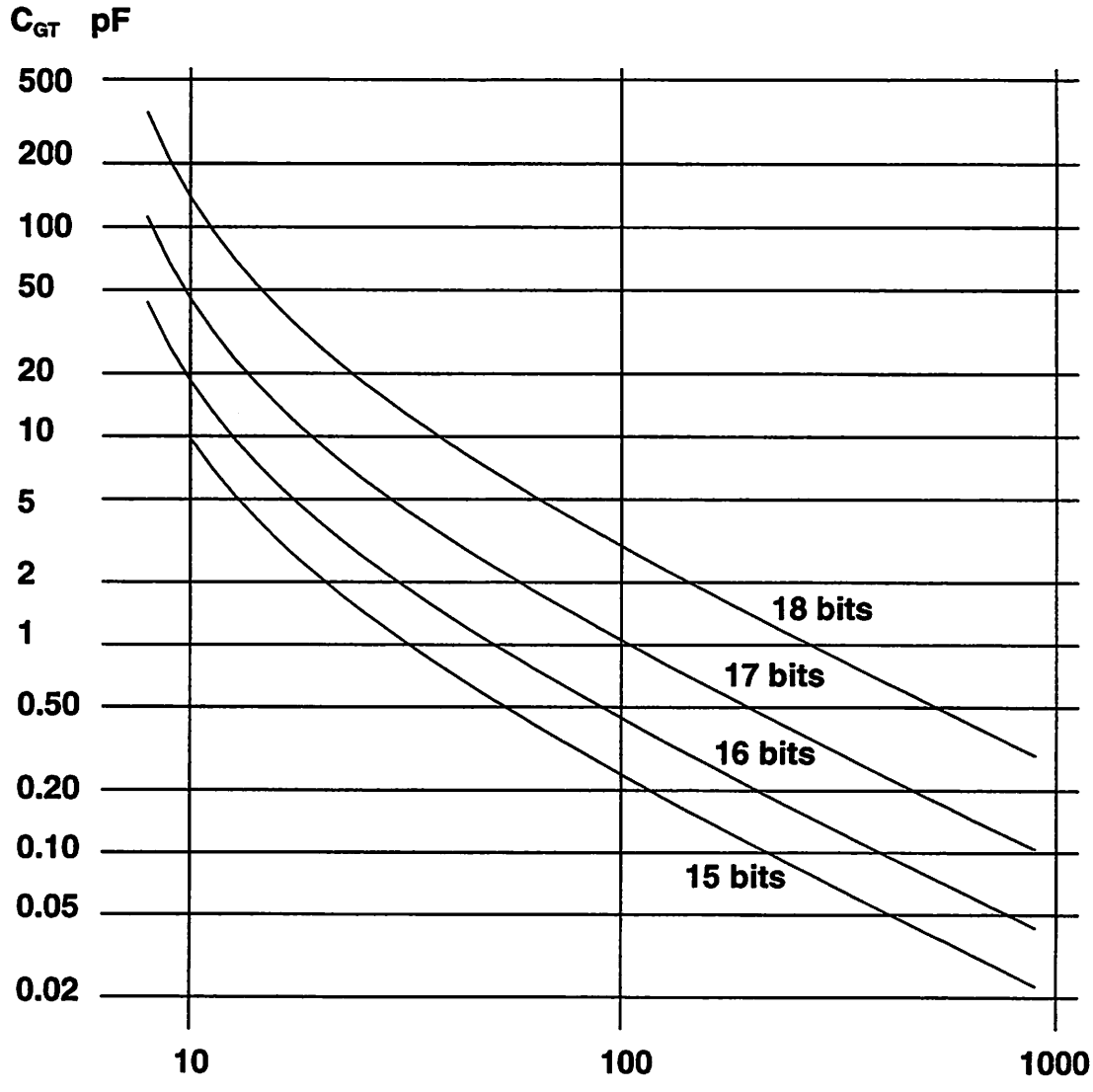


Figure 5.27 Sum Total of the Amplifier Single Ended Input Capacitances Versus Time Constant

Figure 5.27 shows a plot of the total input capacitance versus the time constant. Power dissipation is proportional to the total input capacitance because the current density is fixed. It is no surprise that both high speed and high resolution tend to drive up the required power dissipation.

5.2.5 Optimum Supply Voltage for Power Dissipation in a Pipelined ADC

The power dissipation required to meet resolution and speed requirements in a pipelined ADC is significantly dependent on the supply voltage used. If the design requirements allow some flexi-

bility in the choice of the supply voltage, the supply voltage can sometimes be adjusted to save power.

The next section of this document describes fundamental limits to power dissipation in a pipelined ADC and the relation of these fundamental limits to the supply voltage. The section following this discussion of fundamental limits to power dissipation presents the results of a computer optimization of a pipeline showing how supply voltage affects the power dissipation.

5.2.5.1 Thermal Noise Limits to Power Dissipation in a Pipelined ADC

As discussed earlier in this chapter, corruption of a signal by thermal noise is fundamental to sampling the signal onto a capacitor. For a simple sampling operation, the variance V_n^2 of the stored voltage on a capacitor C is given by the following equation.

$$V_n^2 = \frac{k_B T}{C} \tag{5.100}$$

As stated by Weste[Weste85], the process of sampling a signal by storing charge from the signal onto a capacitor also inherently must dissipate energy. The power P required to alternately charge a capacitor to a voltage V_{DD} and then discharge back to 0 at a frequency f_s is given by the following formula.

$$P = CV_{DD}^2 f_s \tag{5.101}$$

If the input signal is a sinusoid with a peak to peak amplitude of V_{DD}, the power is smaller as shown below.

$$P = \frac{1}{8} CV_{DD}^2 f_s \tag{5.102}$$

The voltage signal to noise ratio (SNR) is given by the following formula.

$$SNR = \frac{V_{signalrms}}{V_n} = \frac{V_{DD}}{V_n 2\sqrt{2}} \tag{5.103}$$

Using the above four equations, the power required can be expressed in terms of the signal to noise ratio and the frequency as shown below.

$$P = f_s k_B T (SNR)^2 \tag{5.104}$$

In the above case having no operational amplifiers, the power is proportional to the sampling rate, the thermal energy, and the required ratio of signal power and noise power. The power is independent of the sampling capacitance and the magnitude of the signal if the signal to noise ratio is held constant.

In actual implementations, the power dissipation is much higher than this thermal limit for a number of reasons. First of all, a number of extra parasitic capacitors must be charged and discharged along with the sampling capacitor. The extra parasitic capacitors include comparators, interconnect, and amplifier input and output capacitances, and they are usually significant. Furthermore, in a pipelined ADC, each stage performs a sampling operation and contributes to the power dissipation. Furthermore, operational amplifiers are often included as part of the sampling circuit in order to make the design independent of parasitic capacitances. A class A architecture is often used in order to reduce supply noise effects and improve settling speed. Because of the class A architecture, power is dissipated whether or not a signal is present, and this results in power dissipation that does not contribute to the signal power. Finally, the amplifier usually does not have unity feedback. This fact increases the noise and the power required to achieve a given speed. The end result is that the power dissipation of the operational amplifier usually is far higher than the fundamental limit described above.

For this case, the power dissipation P_A is fundamentally given by the supply voltage and the bias current I.

$$P_A = V_{DD}I \tag{5.105}$$

Assuming a very simple single transistor amplifier with ideal feedback, the time constant of the amplifier's response to a unit step is given by the following equation.

$$\tau = \frac{C}{g_m} \tag{5.106}$$

The transconductance g_m can be related to the current by the following formula assuming a long channel model holds.

$$\frac{g_m}{I} = \frac{2}{V_{GS} - V_T} \tag{5.107}$$

Combining equations (5.103), (5.105), (5.106), and (5.107) yields the following result for the power dissipation when an operational amplifier is present.

$$P_A = \frac{4k_B T (SNR)^2}{\tau}\left(\frac{V_{GS} - V_T}{V_{DD}}\right) \tag{5.108}$$

If it is assumed that n time constants are required to settle to the required accuracy and that the settling time is half the period of the sampling, then the following relation can be used to relate the time constant τ to the sampling rate.

$$\tau = \frac{1}{2nf_s} \tag{5.109}$$

Combining the above two equations results in the following formula for the power.

$$P_A = 8nf_s k_B T (SNR)^2\left(\frac{V_{GS} - V_T}{V_{DD}}\right) \tag{5.110}$$

The above equation is optimistic because the capacitance to be charged and discharged is probably significantly higher than the sampling capacitor. Note that in the above expression the power dissipation is inversely proportional to the supply voltage. This result contrasts the with the case discussed earlier without opamps. In the case without opamps, the power dissipation is independent of supply voltage.

5.2.5.2 Power Dissipation Trade-offs in Choosing the Supply Voltage

From the previous section and the above equation, it can be seen that in general, high resolution designs limited by thermal noise tend to favor the use of higher supply voltages because the larger supply allows larger signal swings. Thus, the noise requirement is relaxed. For example, a factor of 2 increase in signal swing allows a factor of 4 reduction in capacitor size. This can lead to a factor of 4 reduction in current. Thus, if P=VI, the power is reduced by a factor of 2.

On the other hand, lower resolution designs that are not limited by thermal noise, but instead by parasitic capacitances tend to favor lower supply voltages. This happens because the capacitor sizes in a low resolution design cannot be reduced by using larger supply voltages. As a result, the supply voltage goes up while the capacitor sizes and currents remain the same and an increase in power dissipation results.

Thus, there is a trade-off between thermal noise issues, which drive the design toward higher supply voltages, and parasitic capacitor issues, which drive the design toward lower supply voltages.

In this section, the analysis discussed in section 5.2.4 is performed again. However, this time the supply voltage is allowed to vary. As before, the analysis was performed using a computer, and the results are graphed here. Power is proportional to the total gate capacitance in this analysis. Therefore, the optimum supply voltage can be determined by reading the graphs of total gate

capacitance C_{Gtotal} versus supply voltage. As expected, the optimum supply voltage tends to

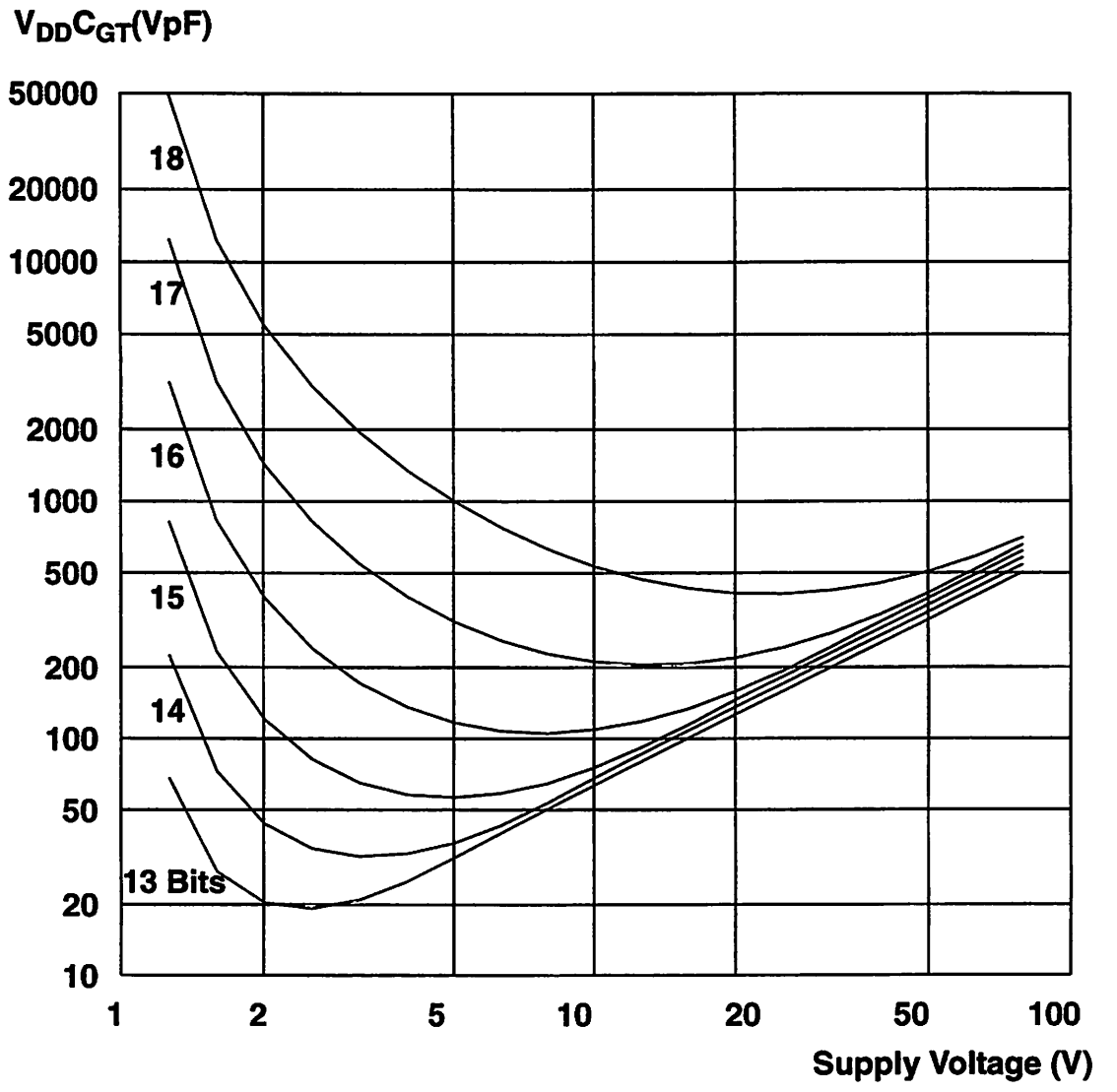


Figure 5.28 Total Power Dissipation of Optimized Pipeline versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

increase as the required resolution is increased. This effect occurs because increasing the supply voltage allows the noise constraint to be relaxed. With the noise constraint relaxed, smaller capacitors can be used, and this use of smaller capacitors reduces the bias current required to drive them. It turns out that the required bias current drops with the square of the supply voltage, so the power dissipation, which is the product of supply voltage and current, drops as supply voltage increases. In low resolution pipelines, increasing the supply voltage tends to increase power dissipation

because the capacitor size is set by parasitic capacitances rather than thermal noise. Since the parasitic capacitances do not generally get smaller as supply voltage is increased, the current is constant with respect to supply voltage. In this case increasing the supply voltage tends to increase the power dissipation.

Four graphs of total capacitance versus supply voltage are included here in order to indicate the minimum supply voltage for four different speeds. From these graphs it is seen that changing the speed requirement does not have a large effect on the optimum supply voltage. However, as expected, increasing the speed requirement creates a large increase in the power dissipation.

More graphs showing the optimal interstage gain, scaling factor, and front end feedback capacitance are shown in the appendix at the end of this chapter.

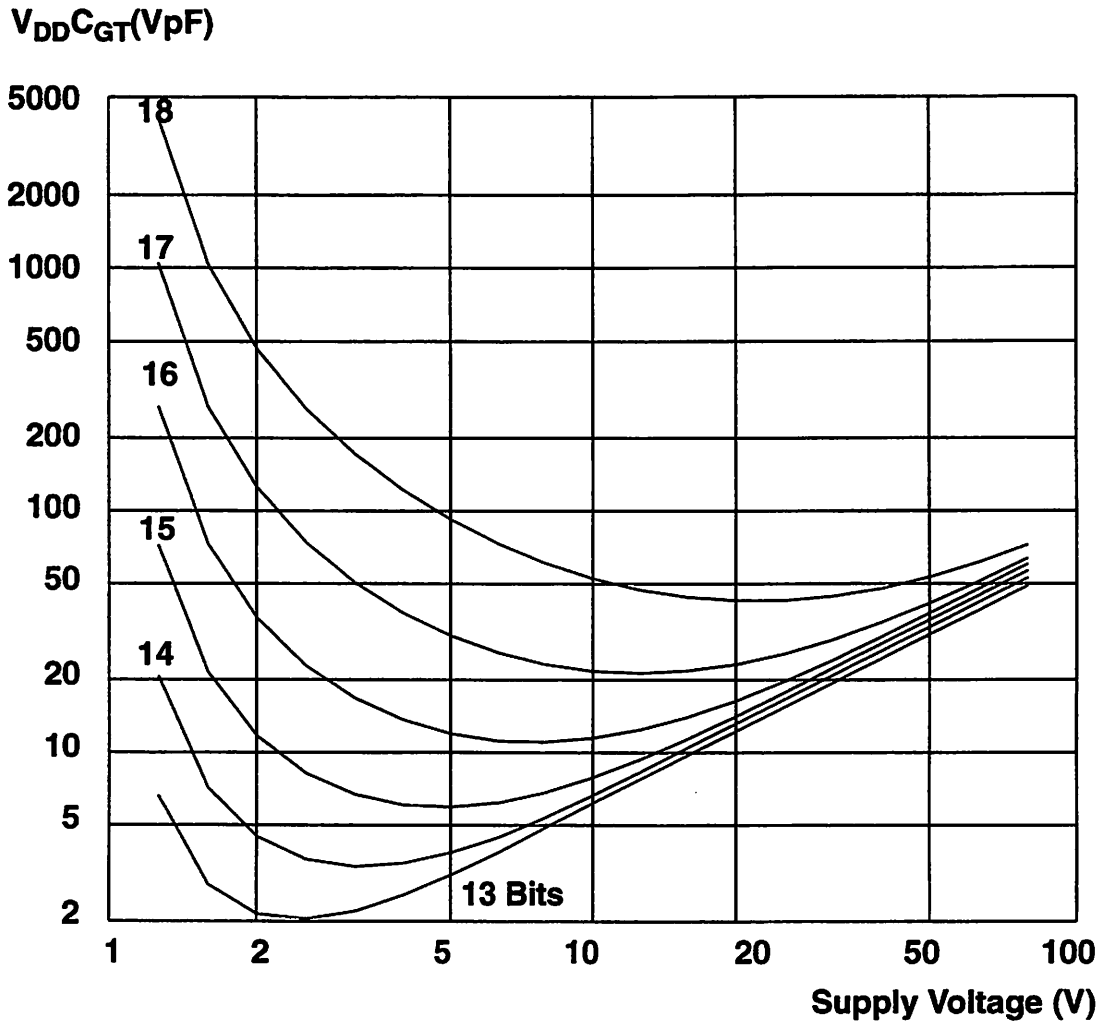


Figure 5.29 Total Power Dissipation of Optimized Pipeline versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

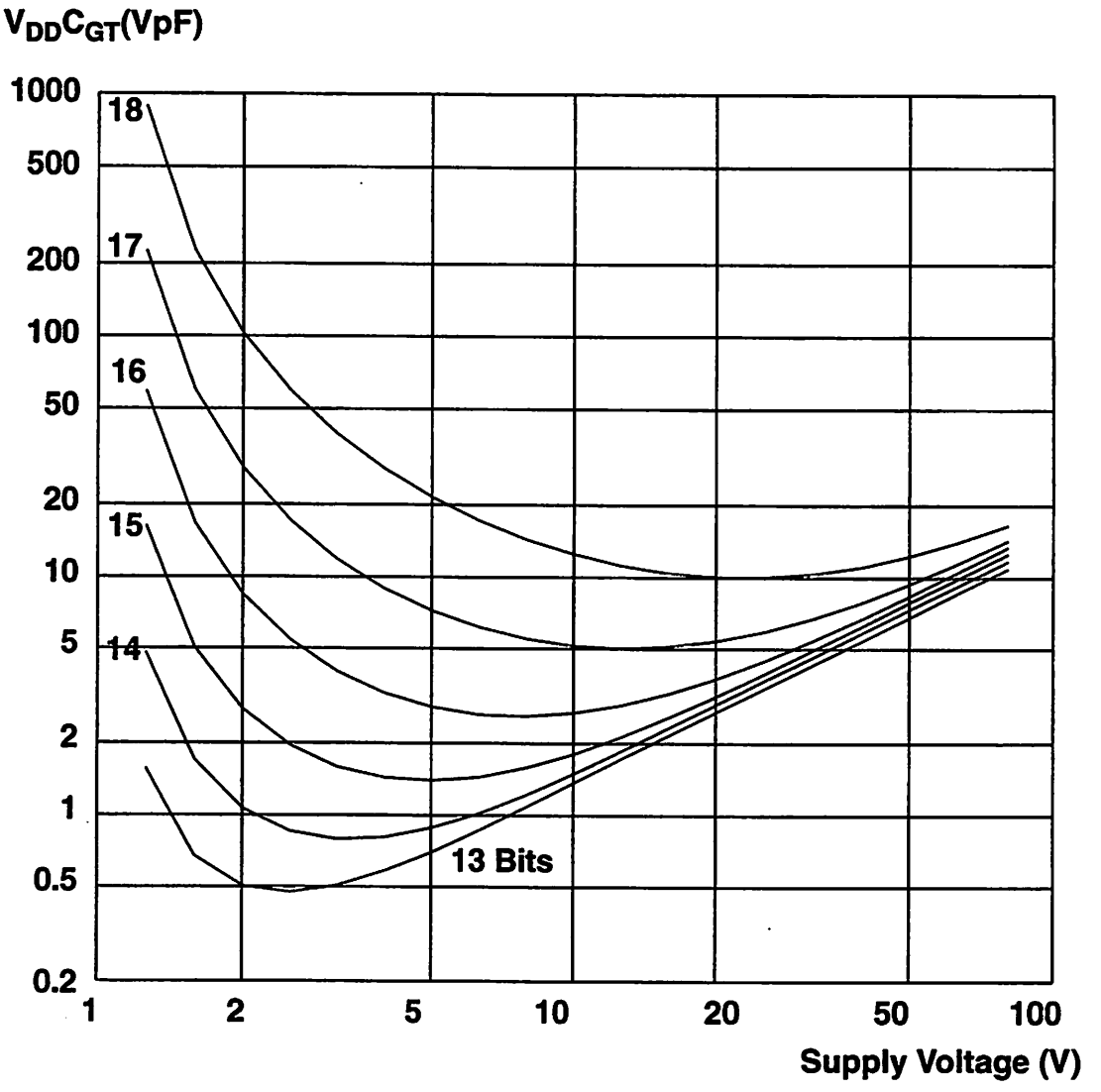


Figure 5.30 Total Power Dissipation of Optimized Pipeline versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

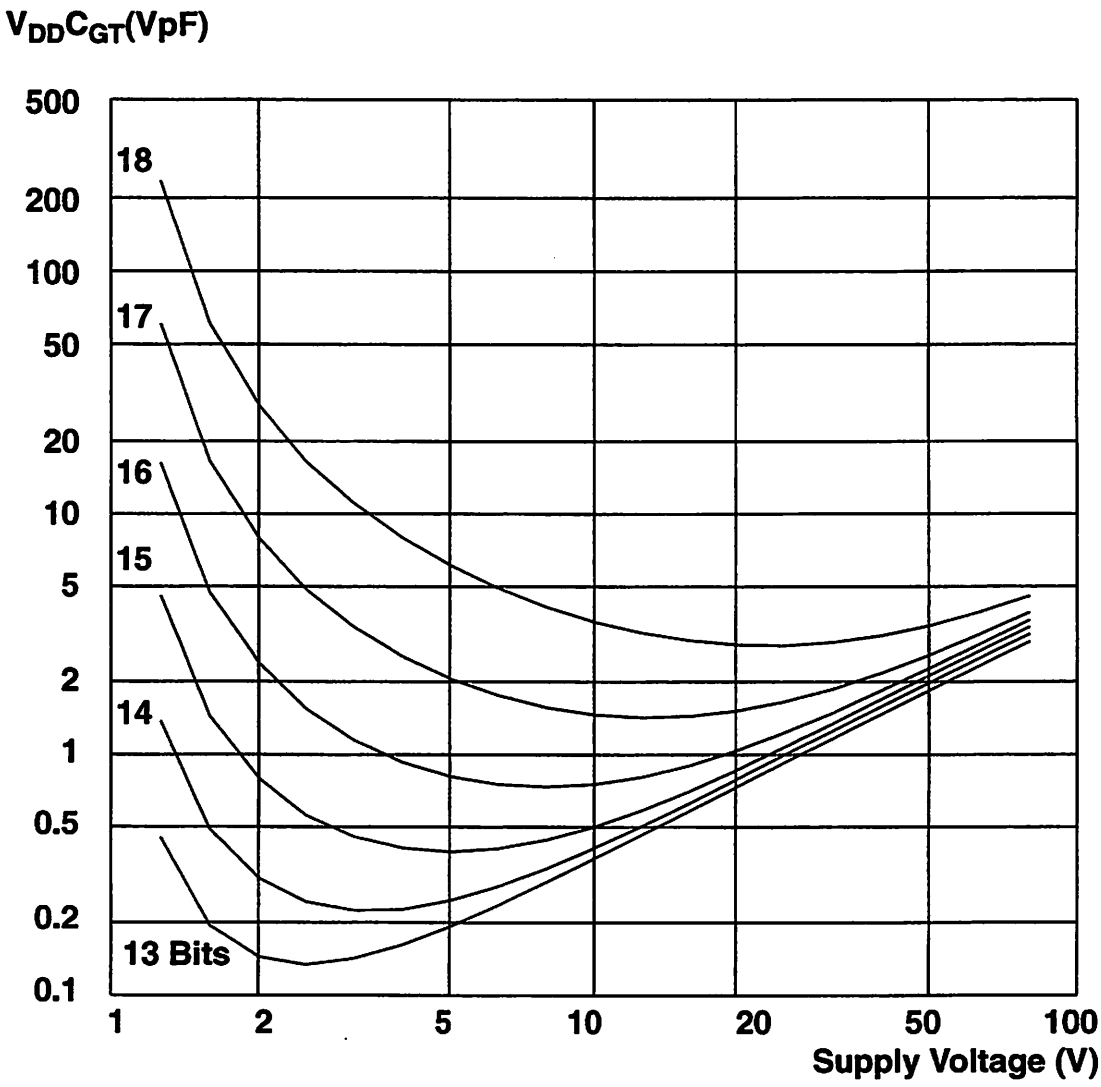


Figure 5.31 Total Power Dissipation of Optimized Pipeline versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

5.3 Scaling of Power and Speed in an Optimized Pipelined ADC

It is often of great interest to know how changing process technologies might affect the performance and power dissipation of a pipeline ADC. Therefore, in this section, a very approximate quantitative analysis of the scaling of power and performance with process technology is performed. The intent of this analysis is to predict how much performance or power dissipation might

improve if a pipelined ADC is redesigned in a new technology with transistors having shorter channel lengths.

In this analysis, the pipeline is assumed to consist of two sections: a front end section limited by thermal noise, and a tail end section limited by parasitic capacitances. It is assumed here that both the front end section and the tail end section contain amplifiers that must meet the same settling time specification. This settling time is assumed to be composed of two periods: a slew rate limited period, and an exponential settling period. In this analysis, the total settling time is assumed to be a simple sum of the slewing time and some number of time constants. This analysis also assumes that single pole class A amplifiers are used, and that the amplifiers swing to the rails. With these assumptions, the slewing period t_{slew} is assumed to be approximately equal to the product of the load capacitance C_{LOAD} and the supply voltage V_{DD} divided by the bias current I as shown below.

$$t_{slew} = \frac{C_{LOAD} V_{DD}}{I} \tag{5.111}$$

The exponential settling t_{exp} period is assumed to be equal to some number x_1 of time constants τ. It turns out that the number of time constants will become unimportant later in this analysis.

$$t_{exp} = x_1 \tau = \frac{x_1 C_{LOAD}}{f g_m} \tag{5.112}$$

In the above equation, f is the feedback factor and is approximately equal to $1/2^n$ where n is the per stage resolution.

In order to determine how the transconductance relates to the current, the following process technology scaling laws are used. First of all, the bias condition is assumed to be defined by the density Q of charge in the channel, and this density of charge is assumed to remain fixed from one technology to the next.

$$Q = C_{ox} (V_{GS} - V_T) \tag{5.113}$$

The gate oxide capacitance C_{ox} per unit area is assumed to be inversely proportional to the channel length L as shown below.

$$C_{ox} \sim \frac{1}{L} \tag{5.114}$$

With these two assumptions, the bias voltage V_{GS}-V_T is proportional to L.

$$V_{GS} - V_T \sim L \tag{5.115}$$

With this information, it can be determined how the bias current density, the ratio of current to channel width, scales with technology. The current is assumed to be given by the long channel model as shown below.

$$I = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS} - V_T)^2 \tag{5.116}$$

The result is that the current density is independent of channel length.

$$\frac{I}{W} = \text{constant} \tag{5.117}$$

With the above information, the scaling of the transconductance g_m with technology may be determined. With the long channel model, the transconductance is given by the following formula.

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I} \tag{5.118}$$

The result is that the transconductance is proportional to the ratio of bias current to channel length as shown below.

$$g_m \sim \frac{I}{L} \tag{5.119}$$

Also, the unity gain frequency ω_T is inversely proportional to the channel length as shown below.

$$\omega_T \sim \frac{g_m}{C_{ox}WL} \sim \frac{1}{L} \tag{5.120}$$

With this result, the exponential settling time given by (5.111) is given as shown below.

$$t_{exp} = \frac{2^n x_1 C_{LOAD} L}{I} \tag{5.121}$$

Therefore, the total settling time t_s is given by the sum of t_{slew} and t_{exp} as shown below.

$$t_s = t_{slew} + t_{exp} = \frac{C_{LOAD}V_{DD}}{I} + \frac{2^n x_1 C_{LOAD} L}{I} \tag{5.122}$$

Now the above equation is rearranged slightly to express the power settling time product in terms of four variables, supply voltage, load capacitance, per stage resolution, and channel length. The result is shown below.

$$Pt_s = V_{DD}It_s = C_{LOAD}V_{DD}\left(V_{DD} + 2^n x_1 L\right) \tag{5.123}$$

The above expression applies to both the front end of the pipeline and the tail end of the pipeline. The total pipeline power dissipation is assumed to be a weighted sum of the contribution from the front end and from the tail end as shown below.

$$P_{total}t_s = (P_{front} + P_{tail})\,t_s = (C_{LOADfront} + C_{LOADtail})\,V_{DD}\left(V_{DD} + 2^n x_1 L\right) \tag{5.124}$$

To proceed further, it is now necessary to determine the load capacitances for the front end and for the tail end. The front end load capacitance is determined by the signal to noise ratio specification, which is determined by the required resolution. It is assumed that the signal amplitude is proportional to the supply voltage, and the RMS value of the thermal noise is assumed to be inversely proportional to the square root of the front end sampling capacitance C_{S0}. Therefore, the voltage signal to noise ratio SNR is given by the following proportionality.

$$SNR \sim V_{DD}\sqrt{\frac{C_{S0}}{k_B T}} \tag{5.125}$$

It is assumed here that the sampling capacitors are scaled down by 2^n from one stage to the next. In the front end of the pipeline the load capacitance is assumed to be dominated by the load capacitance of the following stage ($C_{S0}/2^n$) and the front end feedback capacitance ($C_{S0}/2^n$). Therefore, the load capacitance on the front end is given approximately by the following formula.

$$C_{LOADfront} \approx \frac{2C_{S0}}{2^n} \sim \frac{2k_B T}{2^n}\left(\frac{SNR}{V_{DD}}\right)^2 \tag{5.126}$$

In the tail end of the pipeline, the load capacitance is assumed to be dominated by the input capacitances of the comparators. It is assumed here that the number of comparators is proportional to 2^n. Therefore, the tail end load capacitance is also proportional to 2^n.

$$C_{LOADtail} \sim 2^n C_{comp} \tag{5.127}$$

The comparator input capacitance C_{comp} is assumed to be proportional to the gate capacitance of a minimum sized transistor. It is assumed that the minimum width W_{min} of a minimum sized transistor is proportional to the channel length. Therefore, C_{comp} is proportional to L as shown below.

$$C_{comp} \sim C_{ox} W_{min} L \sim L \tag{5.128}$$

Equation [5.124] can now be rewritten as follows.

$$P_{total} t_s \sim \left(\frac{(SNR)^2}{2^n V_{DD}^2} + 2^n L x_2 \right) V_{DD} \left(V_{DD} + 2^n x_1 L \right) \tag{5.129}$$

In the above equation x_2 is a constant factor that indicates the weight of the tail end contribution relative to the front end contribution. Both x_1 and x_2 will drop out later in the analysis.

Now, the above equation is rewritten slightly to obtain the following relation.

$$P_{total} t_s \sim \left(\frac{(SNR)^2 L}{2^n L V_{DD}} + x_2 2^n L V_{DD} \right) \left(V_{DD} + x_1 2^n L \right) \tag{5.130}$$

The above expression relates power dissipation, settling time, signal to noise ratio, supply voltage, per stage resolution, and channel length. This equation actually possesses a great deal of symmetry, which is revealed by rewriting the equation in terms of a new variable $y = x_1 2^n L V_{DD}$.

$$P_{total} t_s \sim \frac{x_2}{x_1} \left(\frac{\frac{x_1^2 L}{x_2} (SNR)^2}{y} + y \right) \left(V_{DD} + \frac{y}{V_{DD}} \right) \tag{5.131}$$

Now the product of power dissipation and settling time is expressed in terms of channel length and signal to noise ratio. The equation also includes two variables to be optimized, y and V_{DD}. To perform this optimization, y is first assumed to be constant. For a constant y, the supply voltage V_{DD} is optimized when it is equal to $\sqrt{y}$. Then the equation may be rewritten as follows.

$$P_{total}t_s \sim \frac{2x_2}{x_1}\left(\frac{\frac{x_1^2 L}{x_2}(SNR)^2}{y} + y\right)\sqrt{y} \tag{5.132}$$

Now y is the only variable left to optimize. To find the optimum another change of variable is done. This time, a new variable z is defined by the following formula.

$$z = \frac{y}{SNR\sqrt{\frac{x_2}{x_1^2 L}}} \tag{5.133}$$

In terms of z, the power settling time product is given by the following formula.

$$P t_s = 2x_2^{1/4} x_1^{1/2} L^{3/4} (SNR)^{3/2} (z^{3/2} + z^{-1/2}) \tag{5.134}$$

The optimum value of z to minimize the above product is given by the following formula.

$$z = \sqrt{\frac{1}{3}} \tag{5.135}$$

Actually, the optimum value of z is not important here. However, it is important that the optimum value of z is a constant. This implies that y scales as follows.

$$y \sim SNR\sqrt{L} \tag{5.136}$$

Using this information, it can be found that the optimum supply voltage, per stage resolution, and power settling time product scale as follows.

$$V_{DD} \sim L^{1/4} SNR^{1/2} \tag{5.137}$$

$$2^n \sim \frac{SNR^{1/2}}{L^{3/4}} \tag{5.138}$$

$$Pt_s \sim L^{3/4} SNR^{3/2} \tag{5.139}$$

Thus, for a given resolution and power dissipation, the speed improvement obtained by reducing the channel length is expected to be less than linear in an optimized pipelined ADC.

APPENDIX

This appendix includes more graphs of the results of section 5.2.5.2. These plots show the value for the optimum interstage gain, optimum scaling factor, and optimum front end feedback capacitance versus the supply voltage for four different speed requirements.

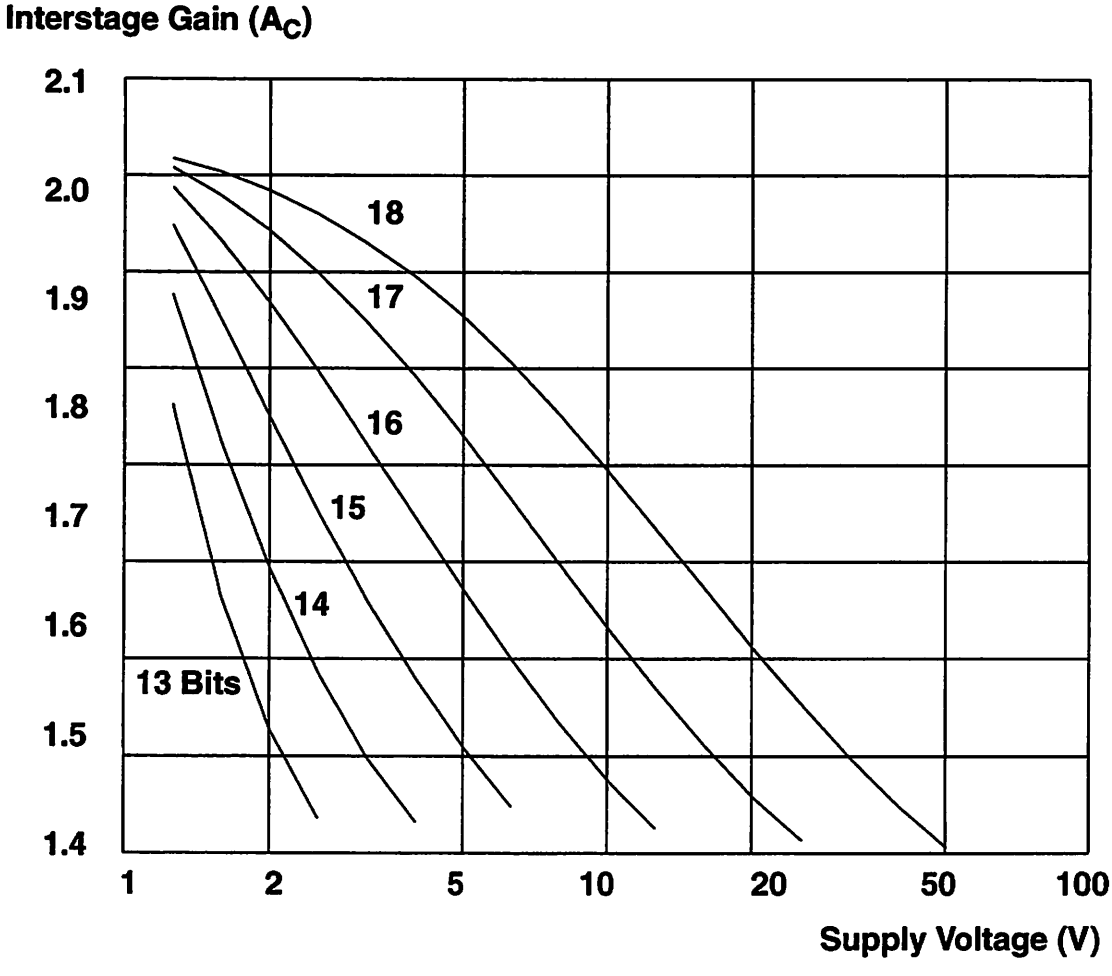


Figure 5.32 Optimum Interstage Gain versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

Figures 5.32 through 5.35 show graphs of the optimum interstage gain versus the supply voltage. It can be seen from these graphs that increasing the supply voltage tends to reduce the optimum value for the interstage gain, and thus also reduces the optimum per stage resolution. This behavior makes sense because increasing the supply voltage relaxes the thermal noise constraint. Increasing the interstage gain relaxes the thermal noise constraint in the tail end stages of the pipeline, but this relaxation of the thermal noise constraint is of little value if the thermal noise requirement is already relaxed to the point that it is not an issue in the design. It is also worth noting that the optimal interstage gain tends to increase as the speed requirement is relaxed. This effect occurs because relaxing the speed requirement increases the relative importance of the thermal noise issue, and thermal noise limited designs tend to be improved by increasing the interstage gain.

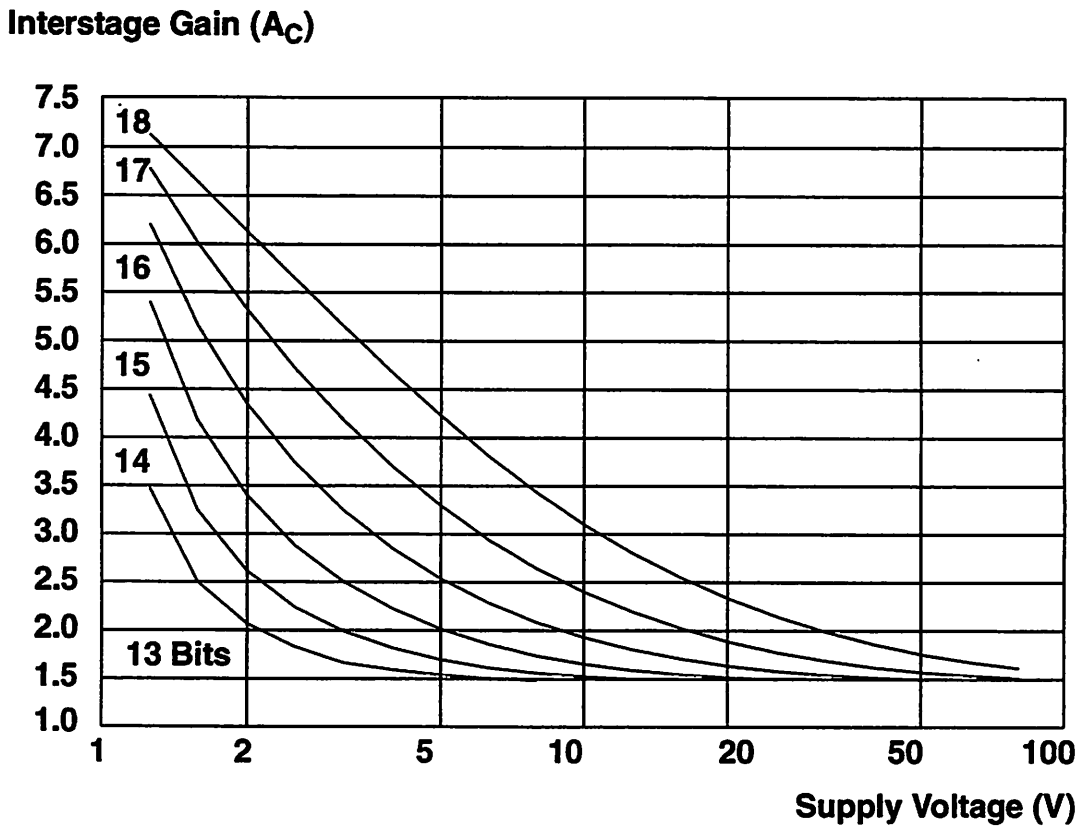


Figure 5.33 Optimum Interstage Gain versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

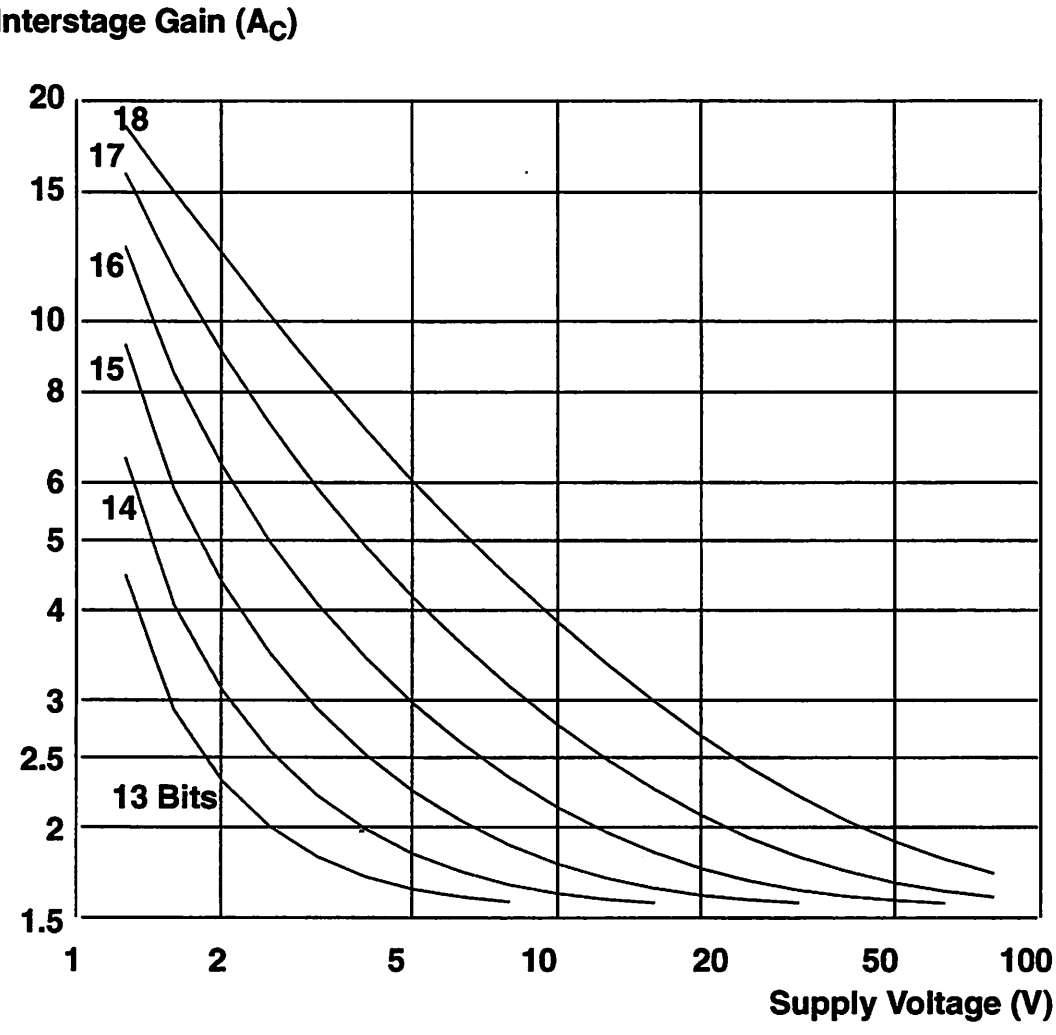


Figure 5.34 Optimum Interstage Gain versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

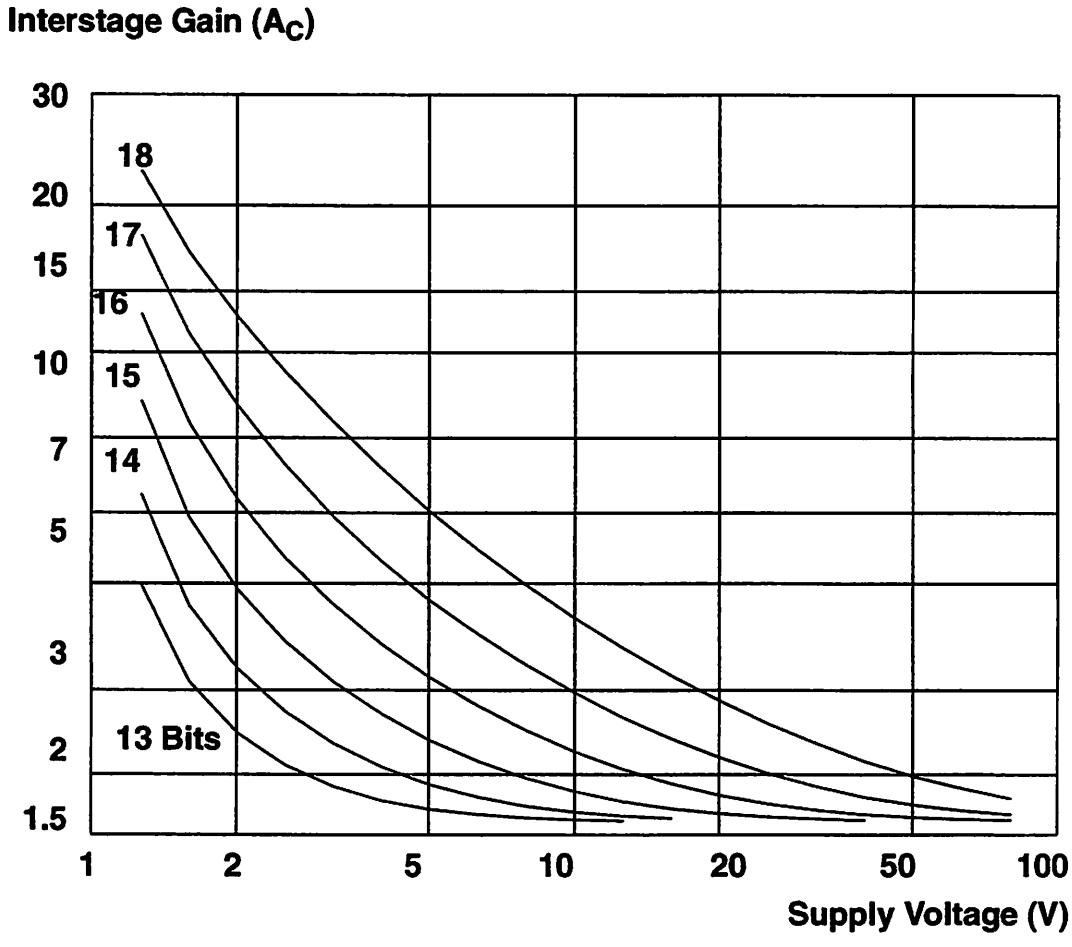


Figure 5.35 Optimum Interstage Gain versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

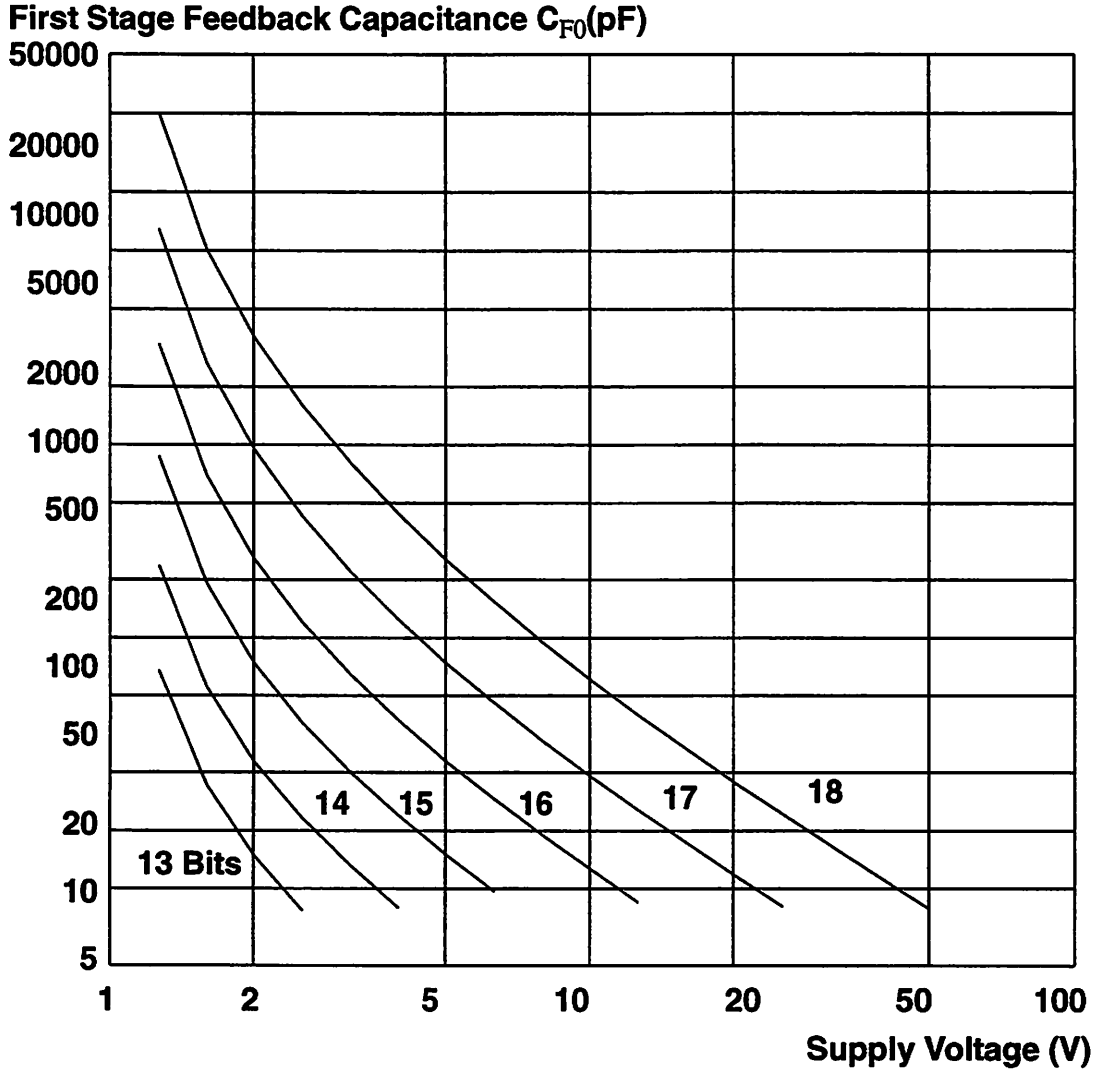


Figure 5.36 Optimum Feedback Capacitance versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

Figures 5.36 through 5.39 show graphs of optimum front end feedback capacitance versus supply voltage. It is no surprise that the optimum front end feedback capacitance is reduced as the supply voltage is increased. This occurs because increasing the supply voltage relaxes the thermal noise constraint and makes it possible to meet the specifications using smaller capacitors.

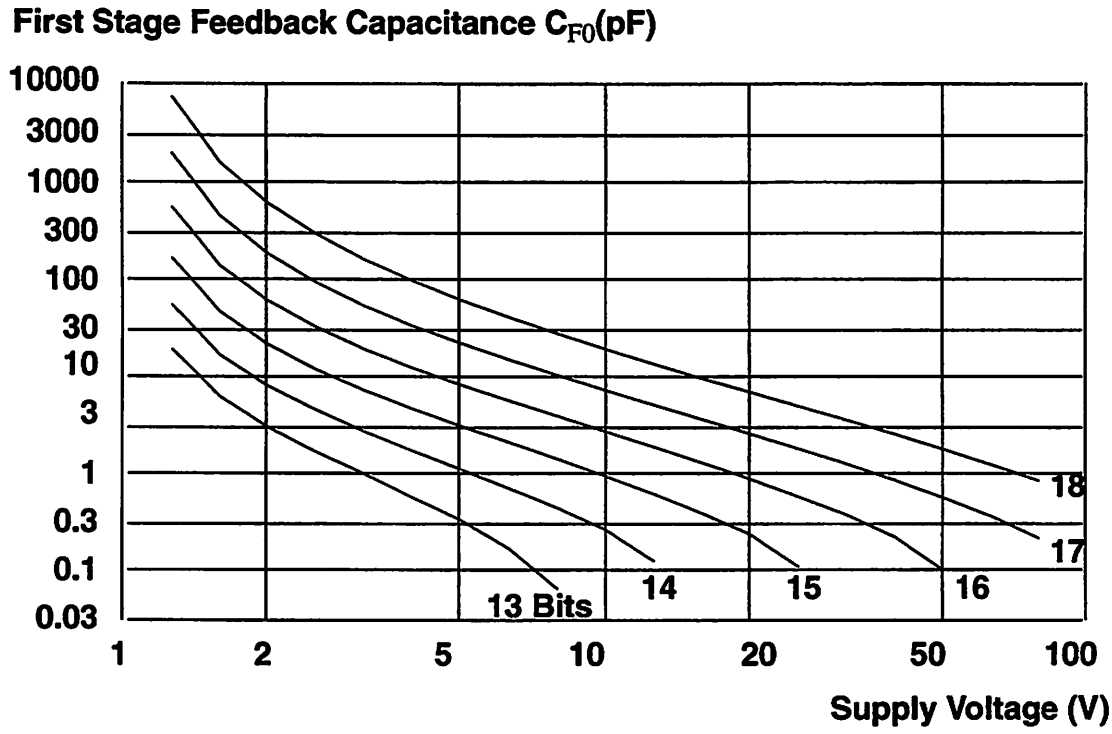


Figure 5.37 Optimum Feedback Capacitance versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

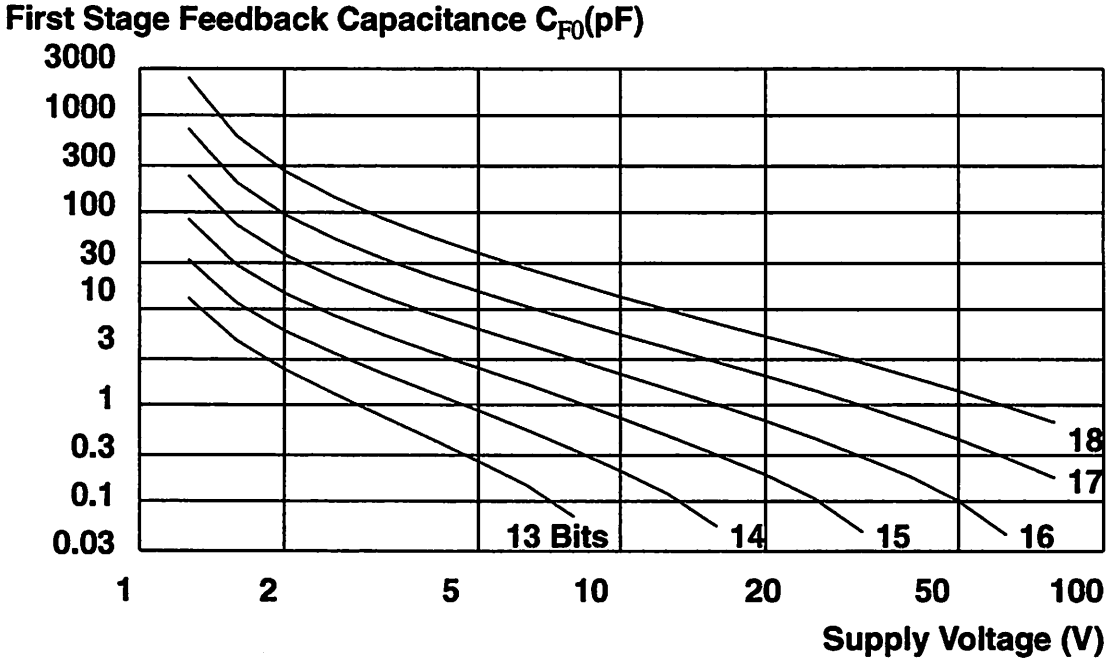


Figure 5.38 Optimum Feedback Capacitance versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

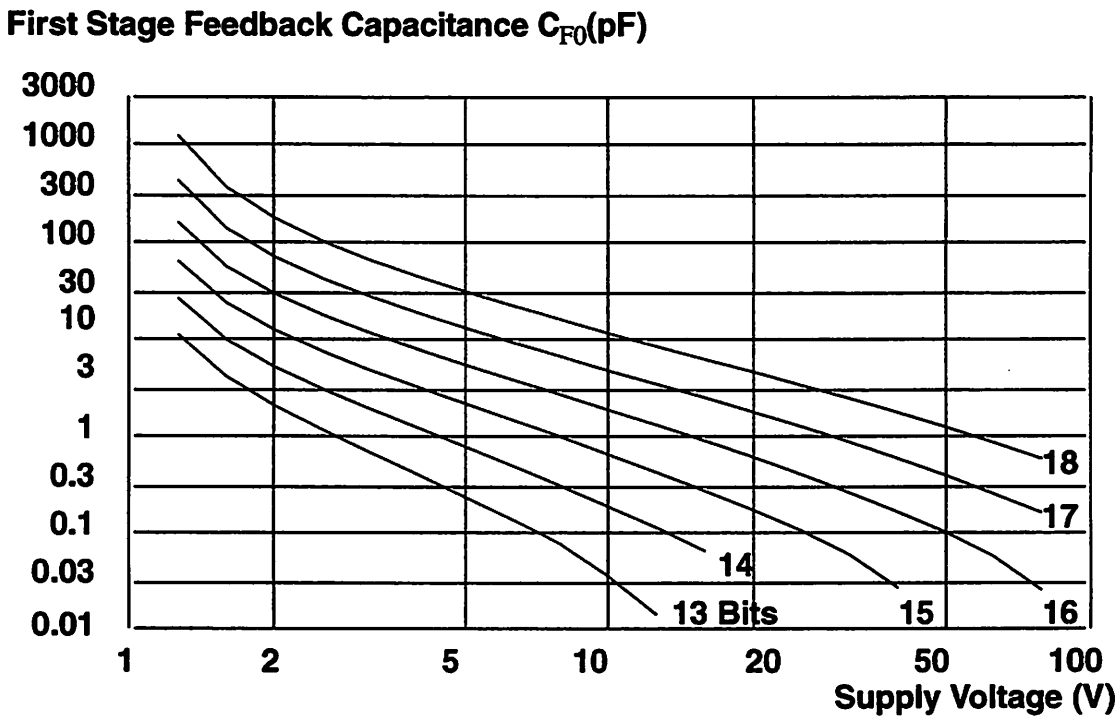


Figure 5.39 Optimum Feedback Capacitance versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

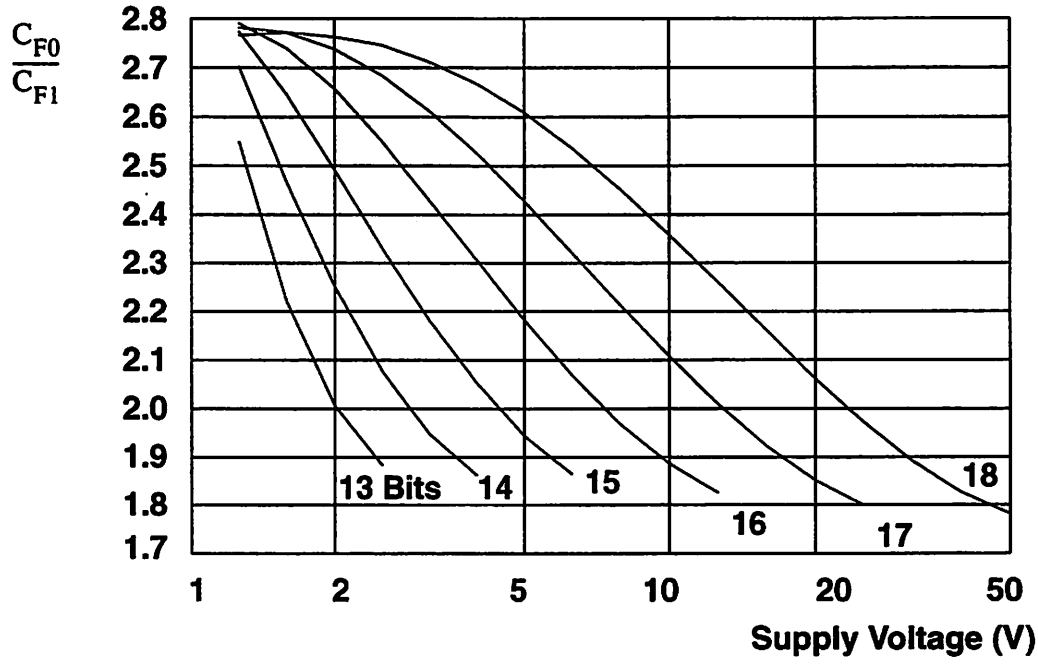


Figure 5.40 Optimum Front End Scaling Factor versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

Figures 5.40 through 5.43 show graphs of the optimum front end scaling factor versus the supply voltage for four different speed requirements. The optimum scaling factor is reduced as the supply voltage is increased because increasing the supply voltage relaxes the thermal noise specification. As a result, scaling of the capacitors from one stage to the next does not help because it will just further relax an already easy to meet noise specification.

Reducing the speed requirement has the tendency to increase the optimum scaling factor. This occurs because reducing the speed requirement increases the relative importance of the thermal noise constraint. Thermal noise limited pipelines tend to favor rapid capacitor scaling. Therefore, the optimum scaling factor is increased.

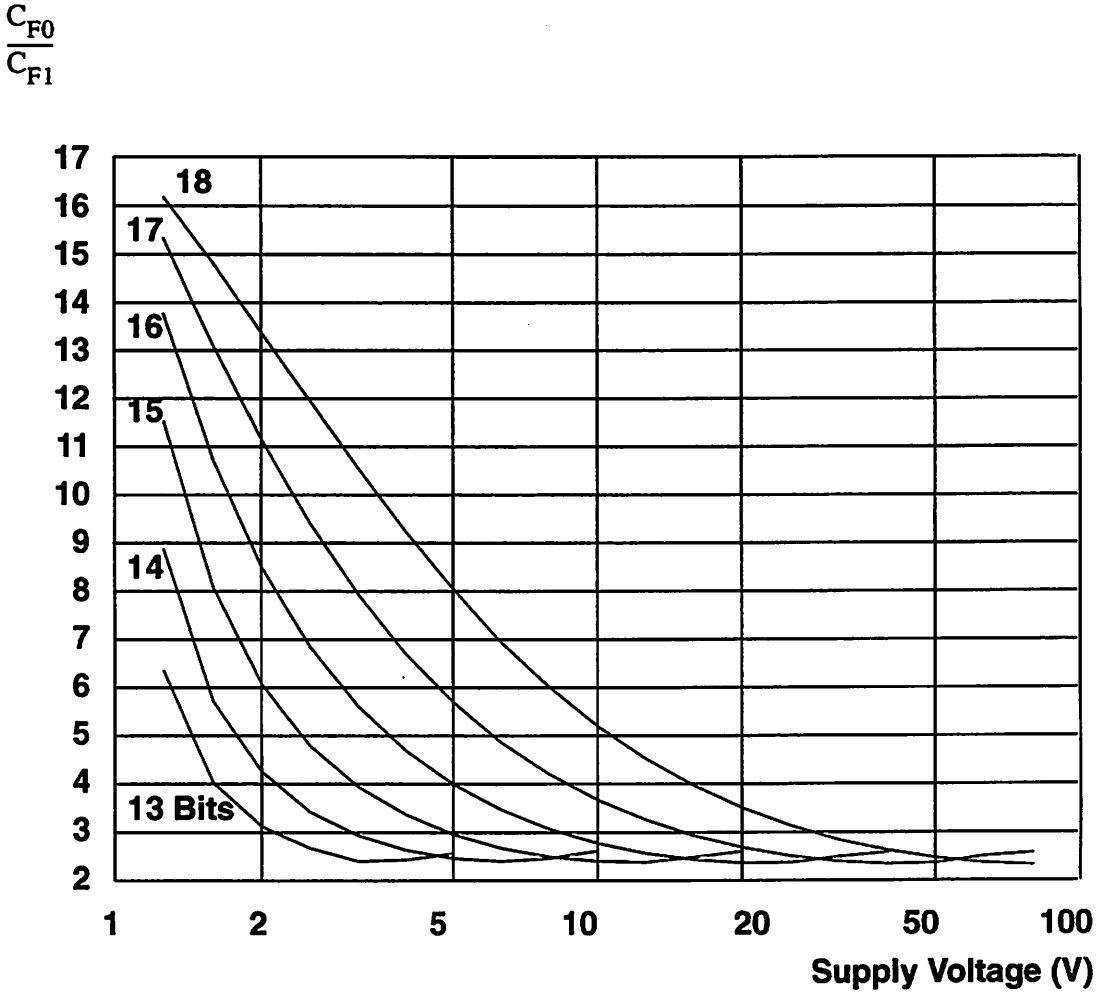


Figure 5.41 Optimum Front End Scaling Factor versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

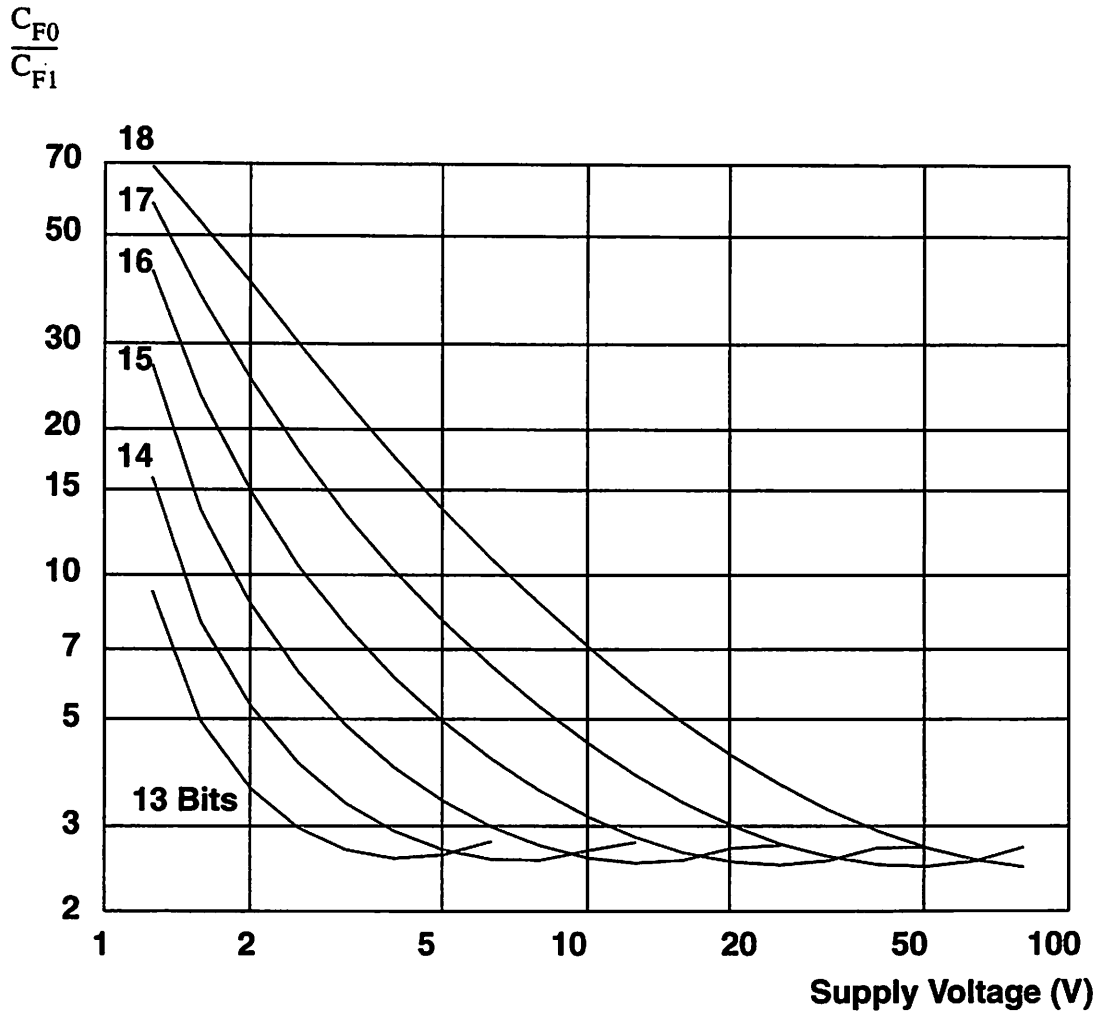


Figure 5.42 Optimum Front End Scaling Factor versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

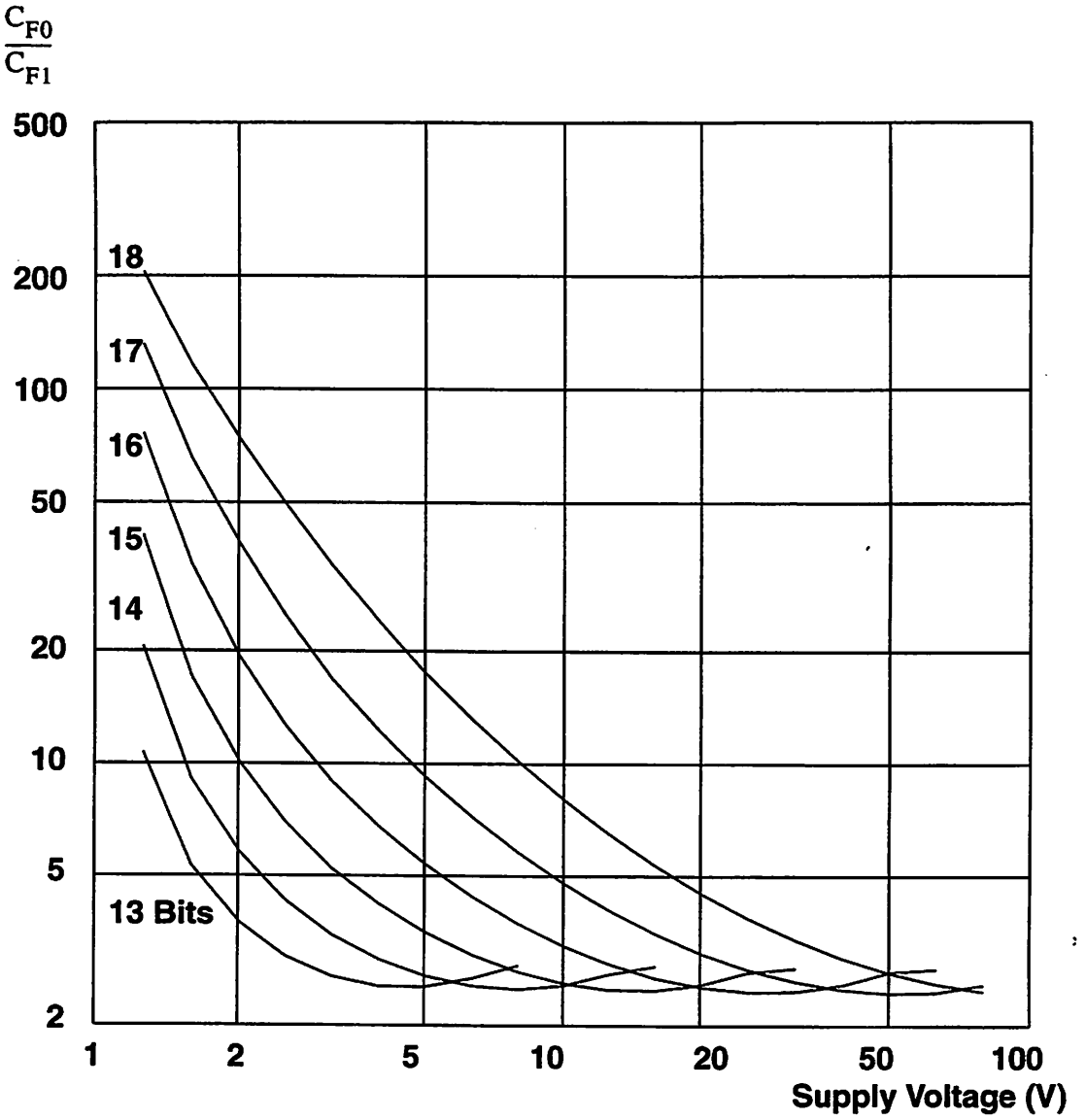


Figure 5.43 Optimum Front End Scaling Factor versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

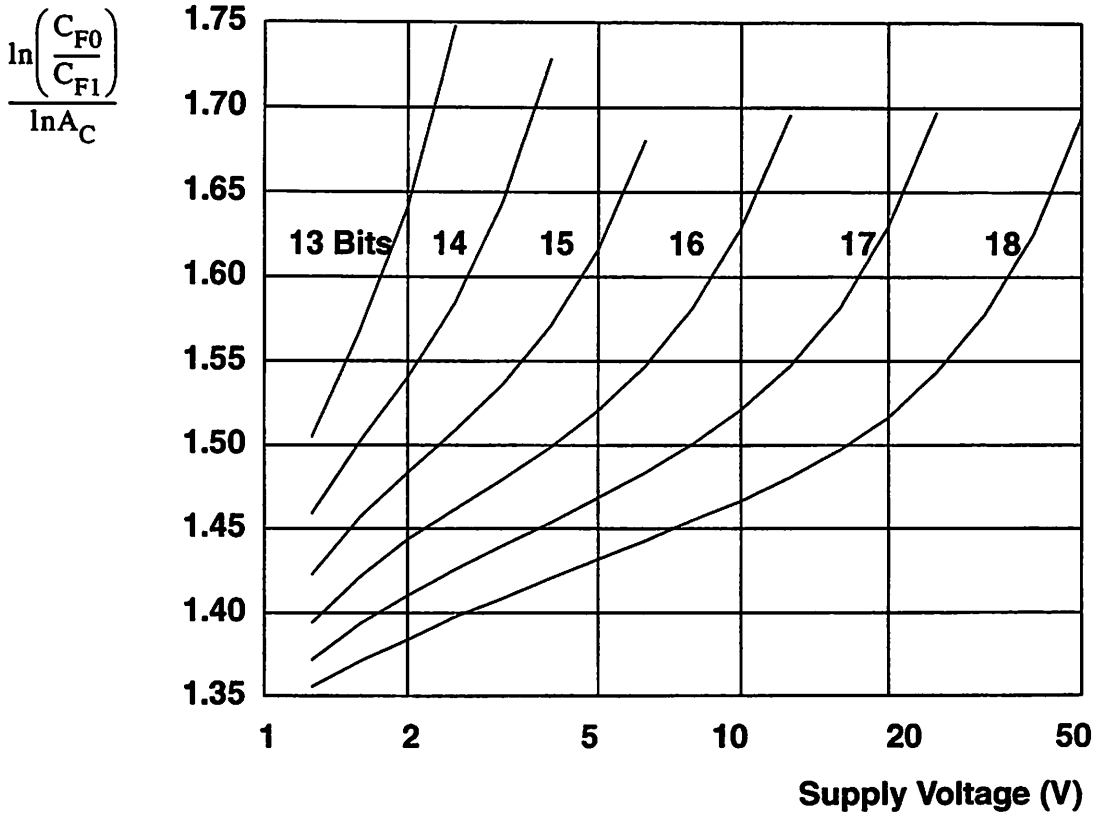


Figure 5.44 Optimum Front End Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

Figures 5.44 through 5.47 show graphs of the front end scaling exponent versus supply voltage. From these graphs it can be seen that increasing the supply voltage tends to increase the optimum front end scaling exponent. This effect occurs because increasing the supply voltage tends to reduce the optimum interstage gain as discussed earlier. In spite of the fact that increasing the supply voltage reduces the optimum front end scaling factor, the optimum interstage gain drops fast

enough that the front end scaling exponent, which relates the scaling factor to the interstage gain, is increased.

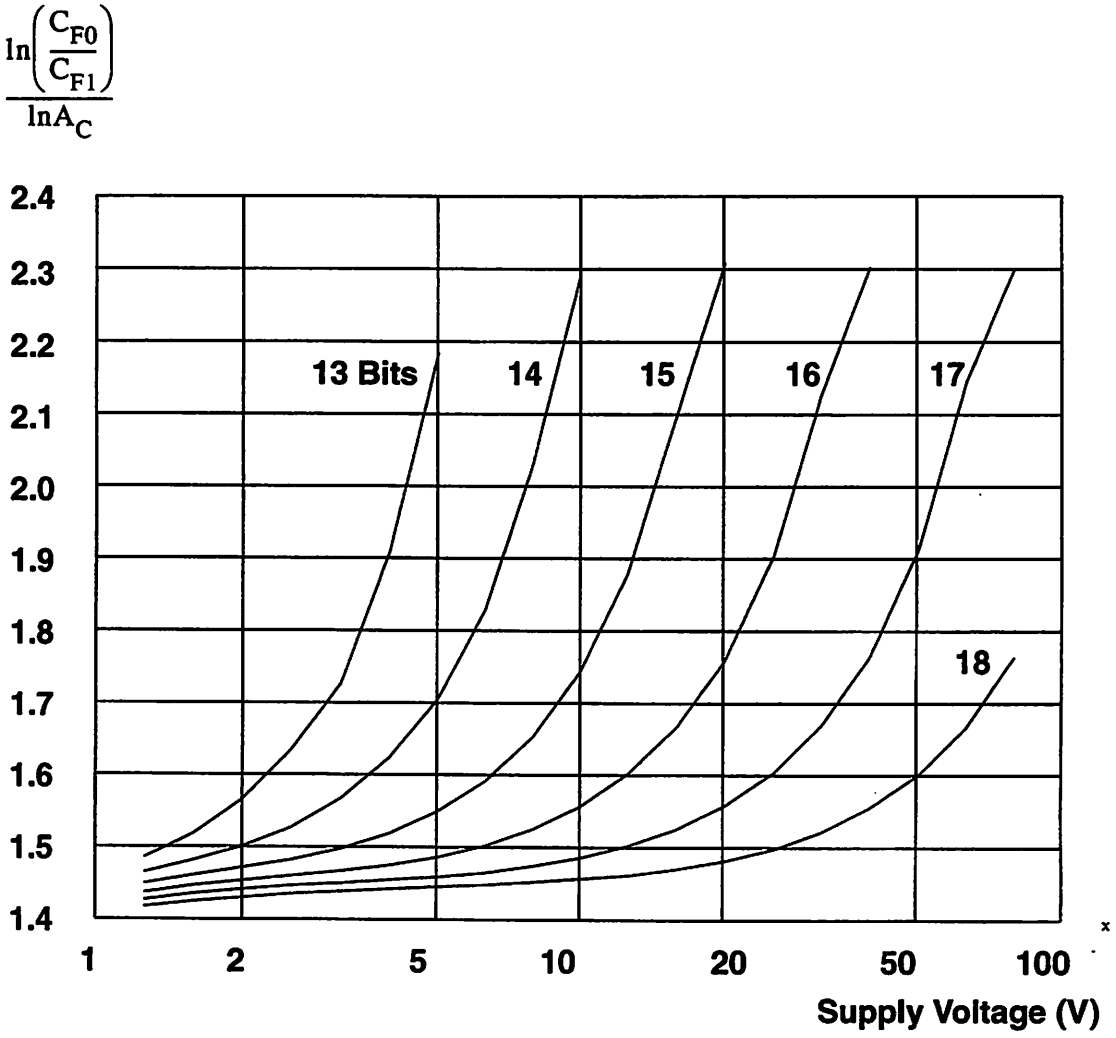


Figure 5.45 Optimum Front End Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

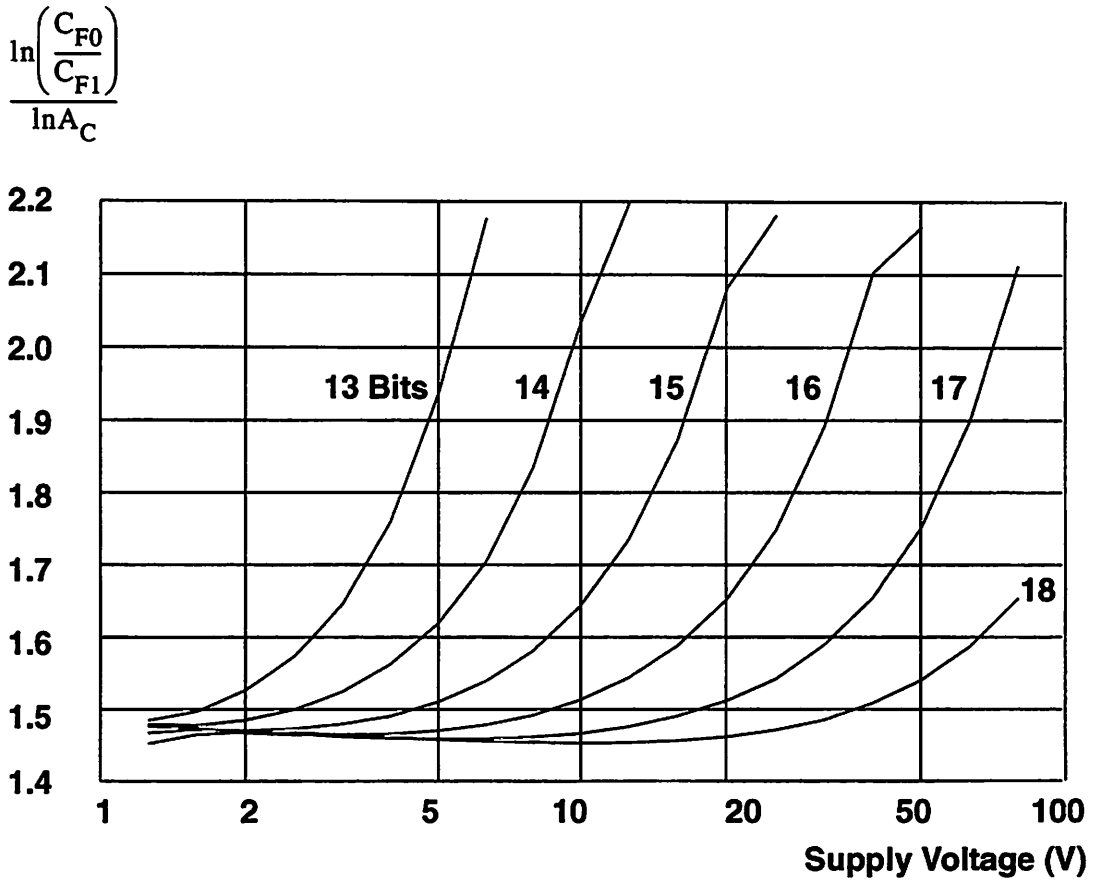


Figure 5.46 Optimum Front End Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

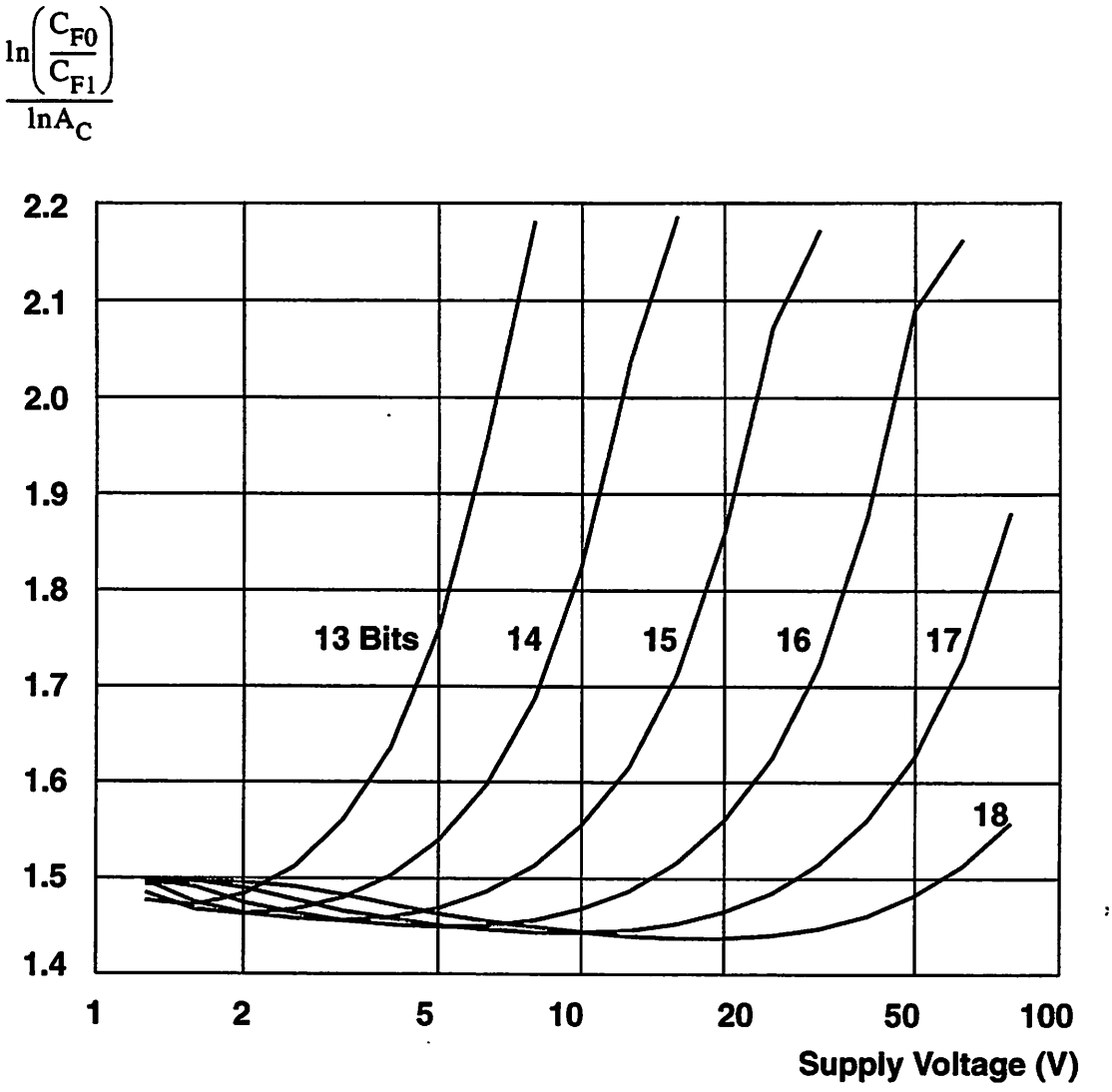


Figure 5.47 Optimum Front End Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

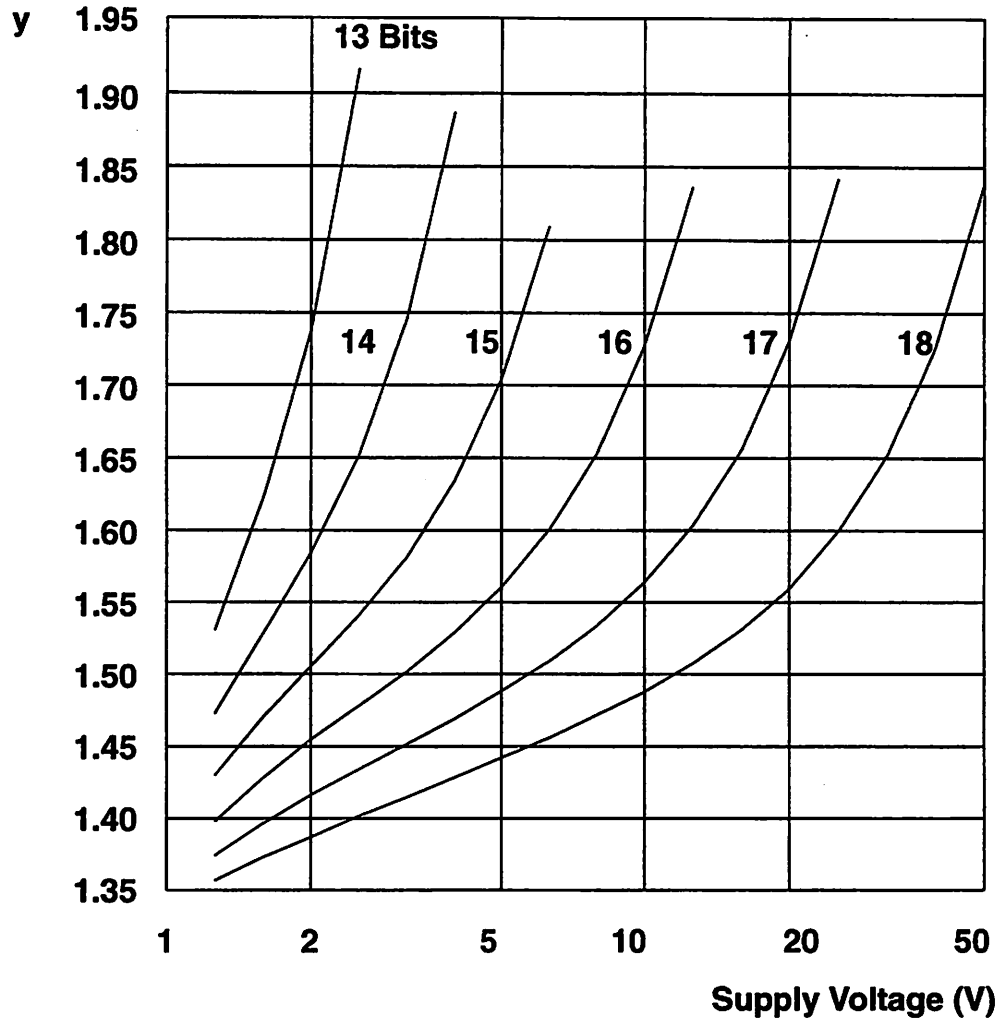


Figure 5.48 Optimum Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=10)

Figures 5.48 through 5.51 show graphs of the optimum value for the scaling exponent versus the supply voltage. In the optimization formulated here, the scaling exponent has less physical significance than the front end scaling exponent. However, these graphs are included here for completeness.

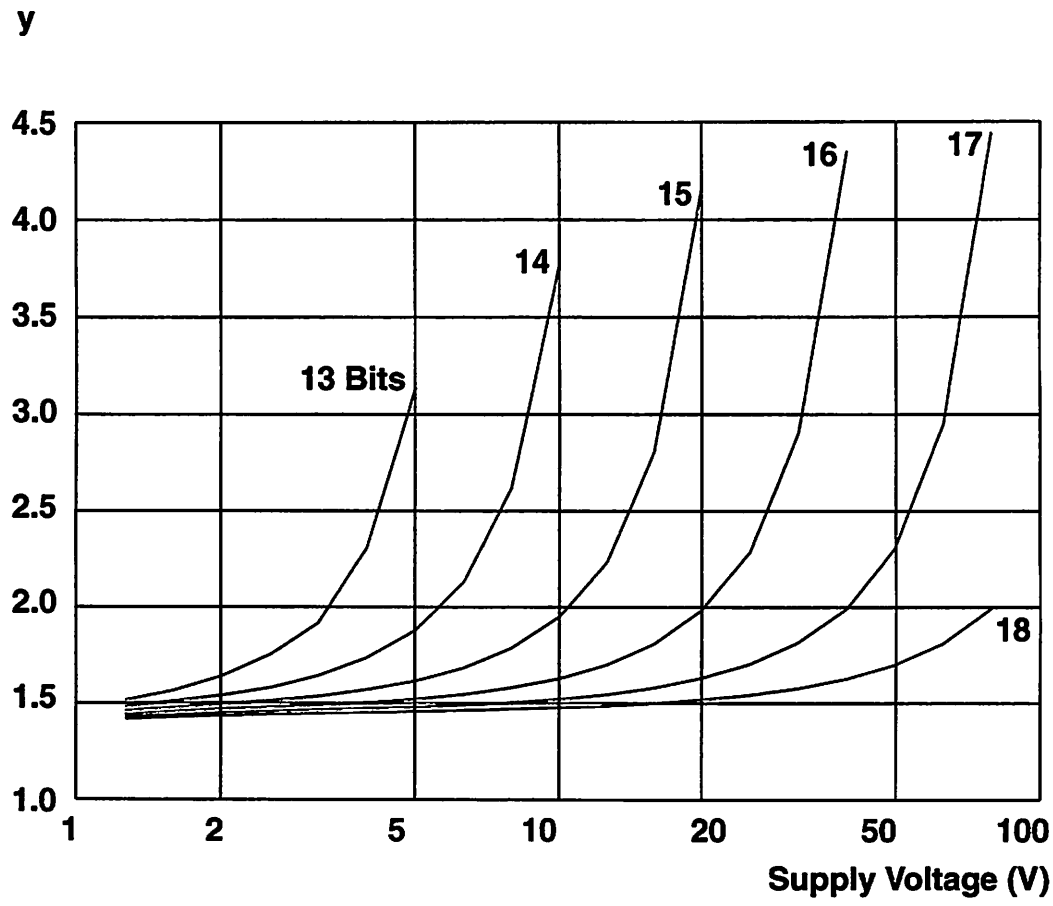


Figure 5.49 Optimum Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=32)

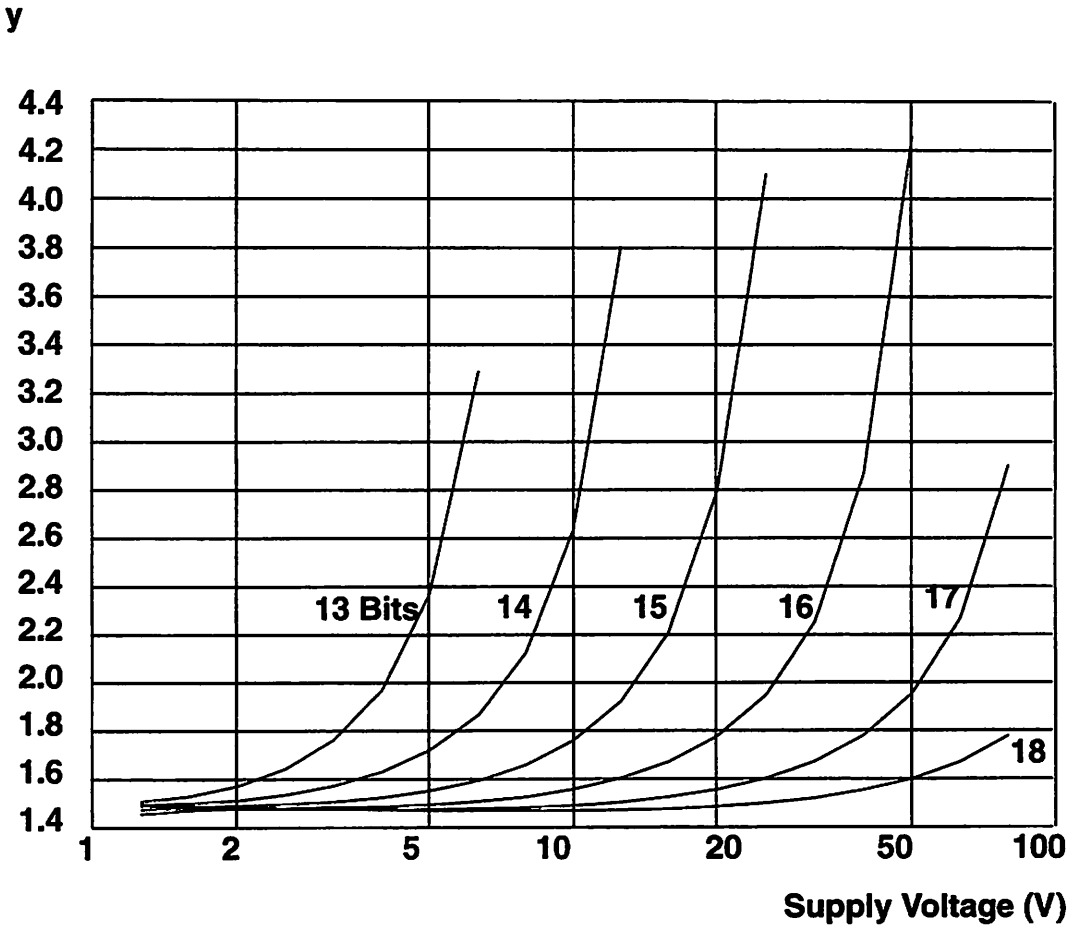


Figure 5.50 Optimum Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=100)

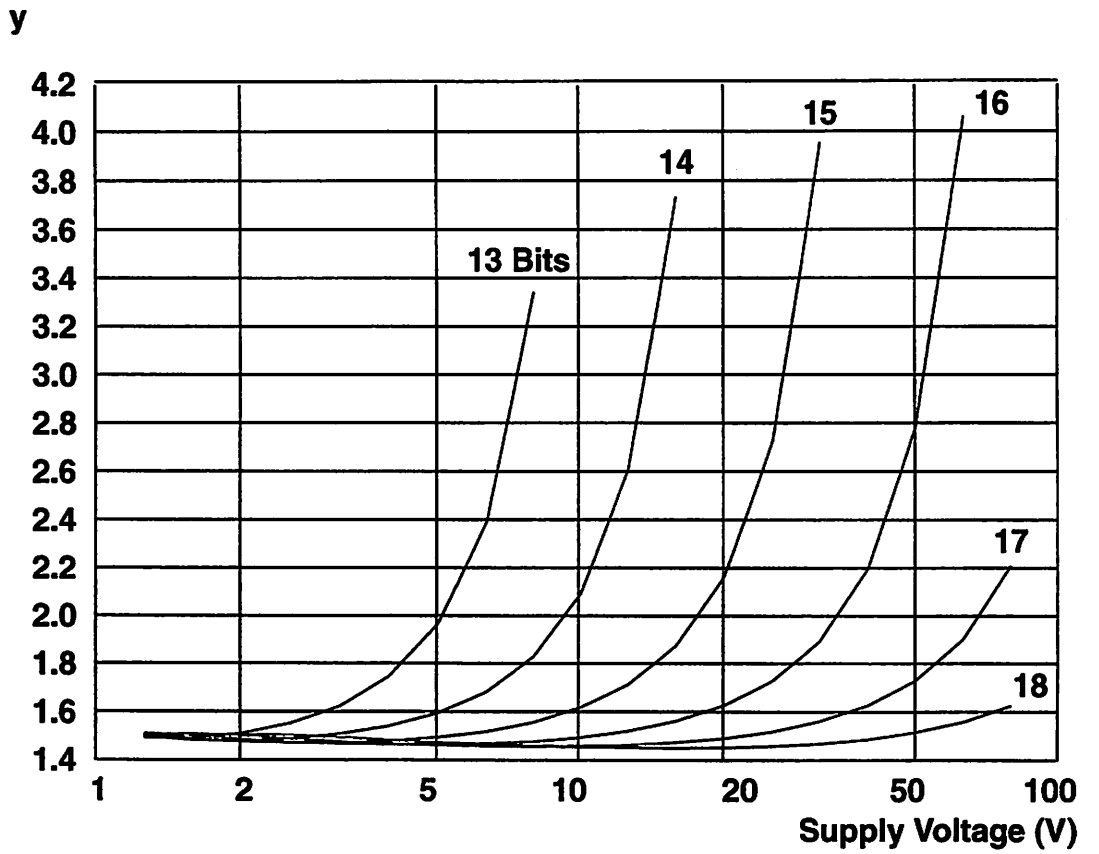


Figure 5.51 Optimum Scaling Exponent versus Supply Voltage for Various ADC Resolutions ($\omega_T\tau$=316)

References

[Conroy94A] Cormac 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, pp.180-206.

[Conroy94B] Cormac 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, pp.119-132.

[Gregorian86] Roubik Gregorian and Gabor C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, New York: Wiley, 1986, pp. 500-513.

[Lewis92] S. H. Lewis, "Optimizing the Stage Resolution in Pipelined, Multistage, Analog-to-Digital Converters for Video Rate Applications," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 39, no. 8, Aug. 1992, pp. 516-523.

[Weste85] Neil H. E. Weste and Kamram Eshraghian, *Principles of CMOS VLSI Design: A Systems Perspective*, Reading, Mass.: Addison-Wesley, 1985, pp. 147-149.

CHAPTER 6

PROTOTYPE DESIGN AND DESCRIPTION

A prototype analog to digital converter was designed, laid out, and fabricated based on the principles discussed in the previous chapters. The design is summarized in this chapter, and the results of testing are discussed in chapter 7.

6.1 Design Goal

The goal of this project was to demonstrate an analog to digital converter with 14 bit linearity and dynamic range at a sampling rate of 10 MHz with as little power dissipation as possible. This corresponds to a signal to noise + distortion ratio of 86dB.

6.2 Architecture

The pipelined ADC architecture was selected because of its ability to achieve high throughput rates with relatively low power and for its tolerance to comparator offsets. Each stage is composed of six comparators, a charge redistribution digital to analog converter (DAC), and a gain block with a gain of 4. Thus, each stage effectively resolves 2 bits. Digital calibration[Karanicolas93] was used to correct for capacitor mismatches in the digital to analog converters of the first six stages of the pipelined ADC. The pipeline is composed of a total of 9 stages for a nominal resolution of 18 bits. The two additional stages are used to improve the calibration accuracy of the previous stages. The post-processing required to generate the correct digital output was done off chip by computer. A block diagram of the entire system is shown in figure 6.1. Figure 6.2 shows the structure of a single stage of the pipeline.

6.2.1 Comparator Architecture

Because of digital error correction, a high comparator offset voltage can be tolerated. In the prototype design, the signal swing is +/-3.3V, the interstage gain is 4, and the number of compara-

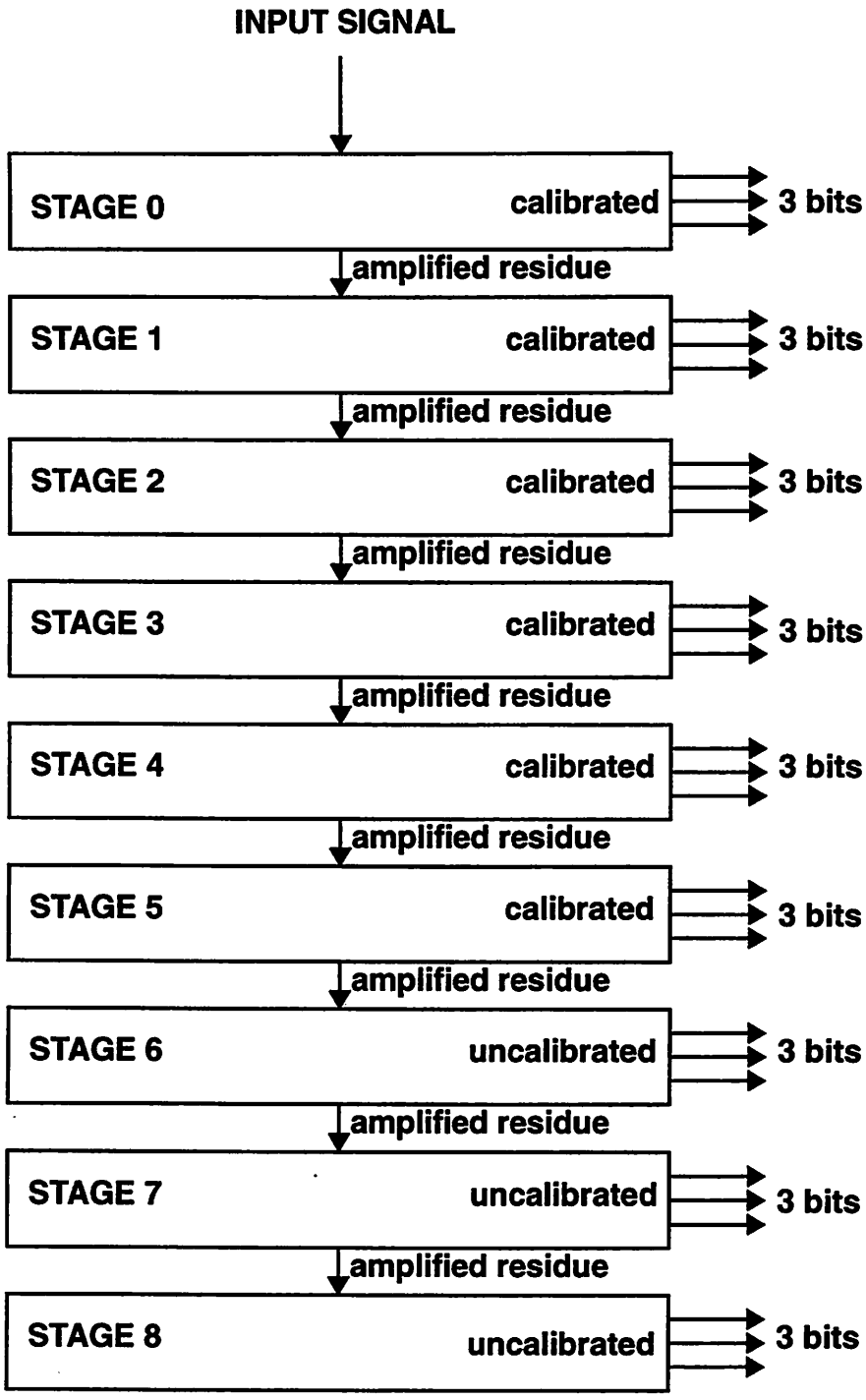


Figure 6.1 Overall Structure of the Prototype Pipelined ADC

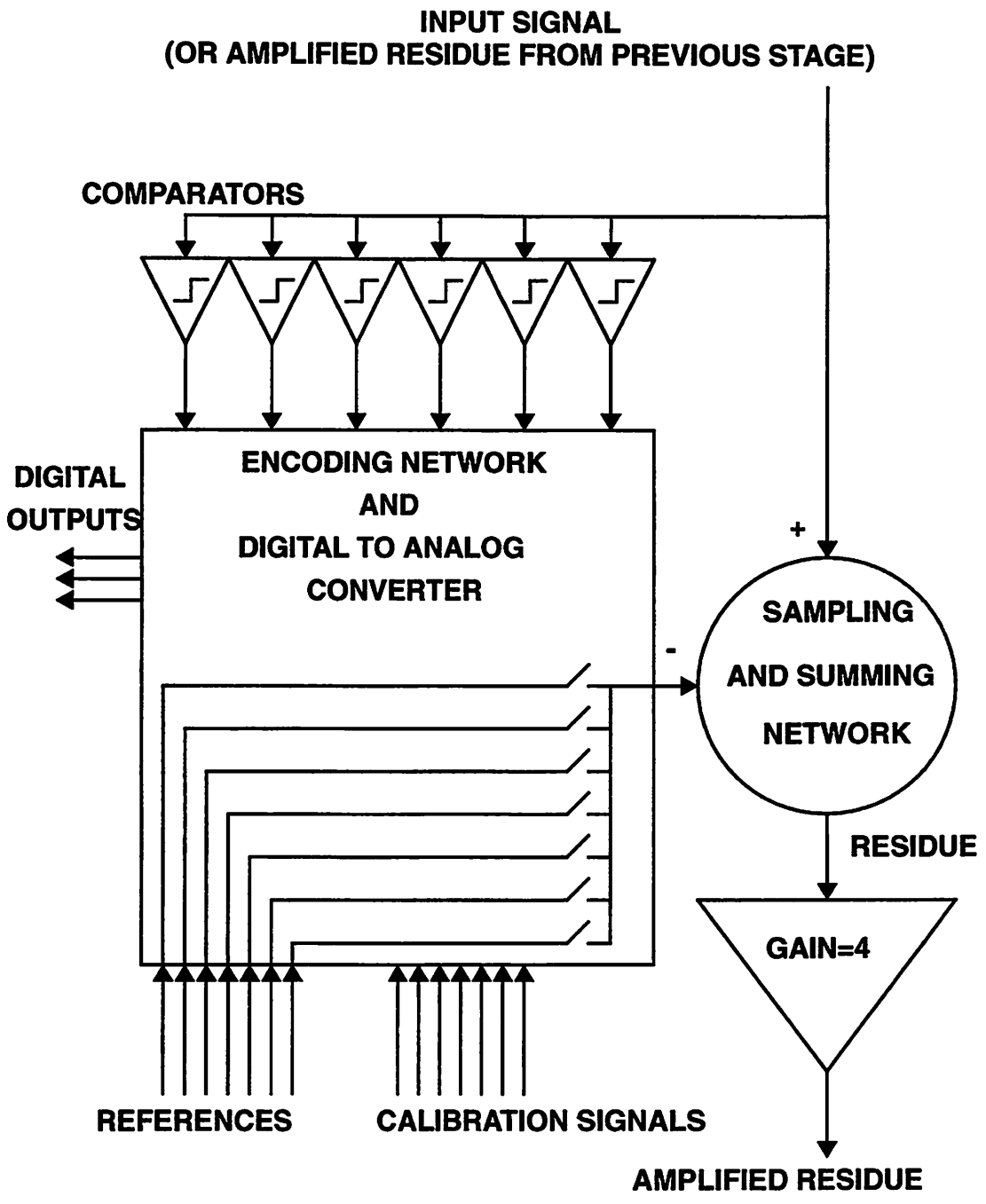


Figure 6.2 One Stage of the Pipelined ADC

tors is 6. Using the analysis of section 3.4.2, an offset voltage of 412.5mV can be tolerated. Because such a high offset voltage can be tolerated, it was possible to use a very simple dynamic comparator. A number of references exist for dynamic comparators[Yukawa85][Cho95][Song95]. The circuit diagram for the comparator used in this project is shown in figure 6.3. This comparator circuit consists of a latch (transistors ML), precharge transistors (MP), input transistors (MI), cascode transistors (MC), bias transistors (MBIAS), input sampling switches (MSI), and reference sampling switches (MSR).

The operation of this circuit is as follows. While clocks ϕ_{A1}, ϕ_{A2}, and ϕ_{A3} are high, the charges at nodes V_{S+} and V_{S-} are initialized to values determined by the reference voltages V_{R+} and V_{R-} and the bias voltage. While clocks ϕ_{S1}, ϕ_{S2}, and ϕ_{S3} are high, the output nodes V_{o+} and V_{o-} are precharged to V_{DD}. V_{C+} and V_{C-} are also precharged to V_{DD} at this time. V_{A+} and V_{A-} are precharged to ground at this time. During this time, V_{S+} and V_{S-} are floating and follow the input voltages V_{I+} and V_{I-} through the capacitive coupling provided by C_S and C_R. When ϕ_{S2} goes low, latching begins while the input remains connected. Both V_{o+} and V_{o-} begin to discharge through transistors MC and MI. The differential input voltage produces a differential voltage on nodes V_{S+} and V_{S-}. This differential voltage produces a mismatch in impedance that causes the nodes V_{o+} and V_{o-} to discharge at unequal rates. Eventually, one of the nodes reaches a voltage low enough to cut-off discharge of the other node. At this time, the slow node is charged back to V_{DD} while the fast node continues to discharge to ground. Thus a decision has been made by the circuit.

Figure 6.4 shows how six of these comparators are used together to form a small ADC. Each comparator has a corresponding input threshold voltage of 0V. For V_{S+}-V_{S-}<0, the output is high. Otherwise the output is low. In order to construct a useful ADC, a set of comparators with a range of input threshold voltages is required. This is accomplished by appropriately choosing the value for the ratio of C_R to C_G. This ratio determines the amount of charge initially stored at the summing nodes V_{S+} and V_{S-}.

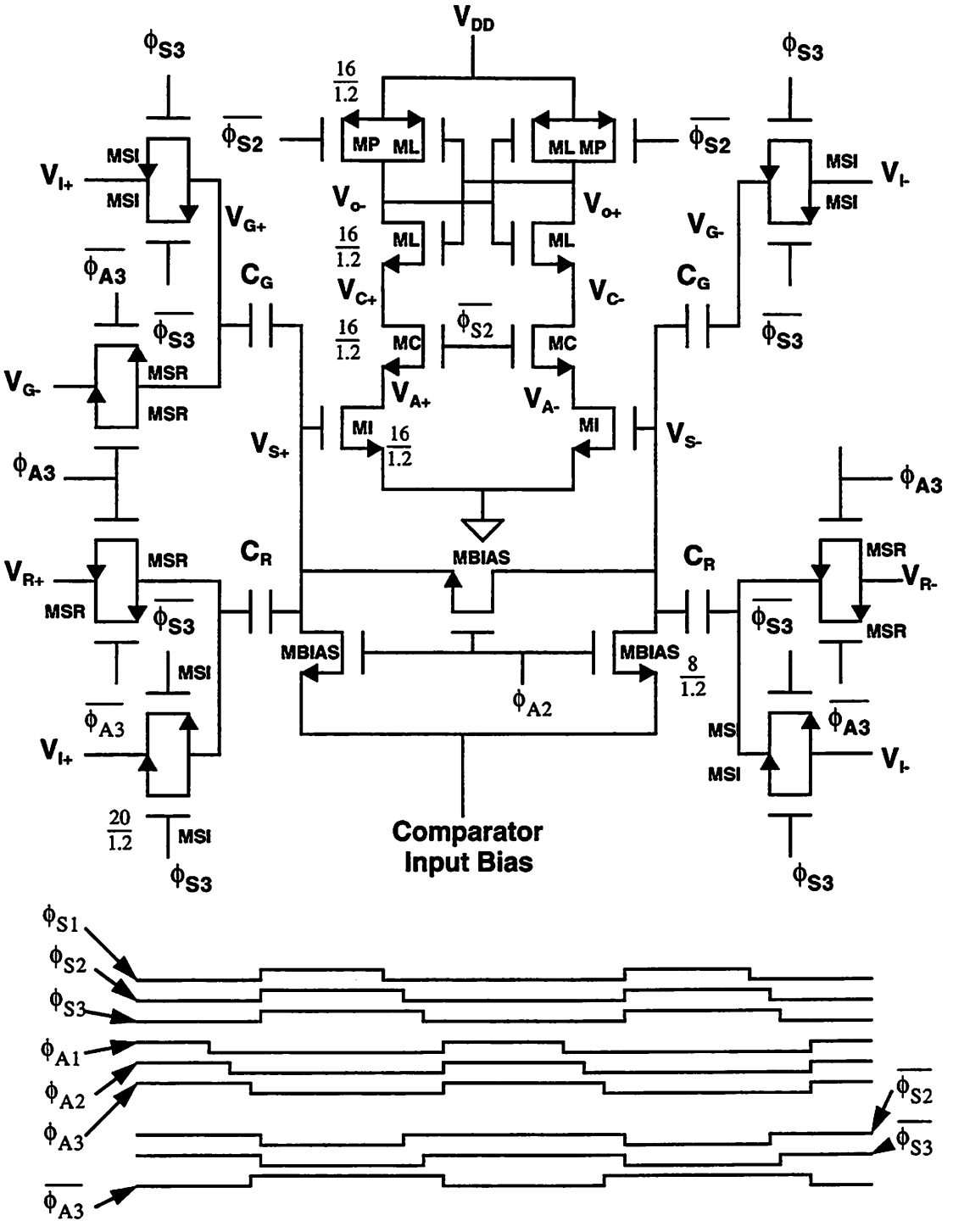


Figure 6.3 Comparator and Comparator Sampling Network Schematic

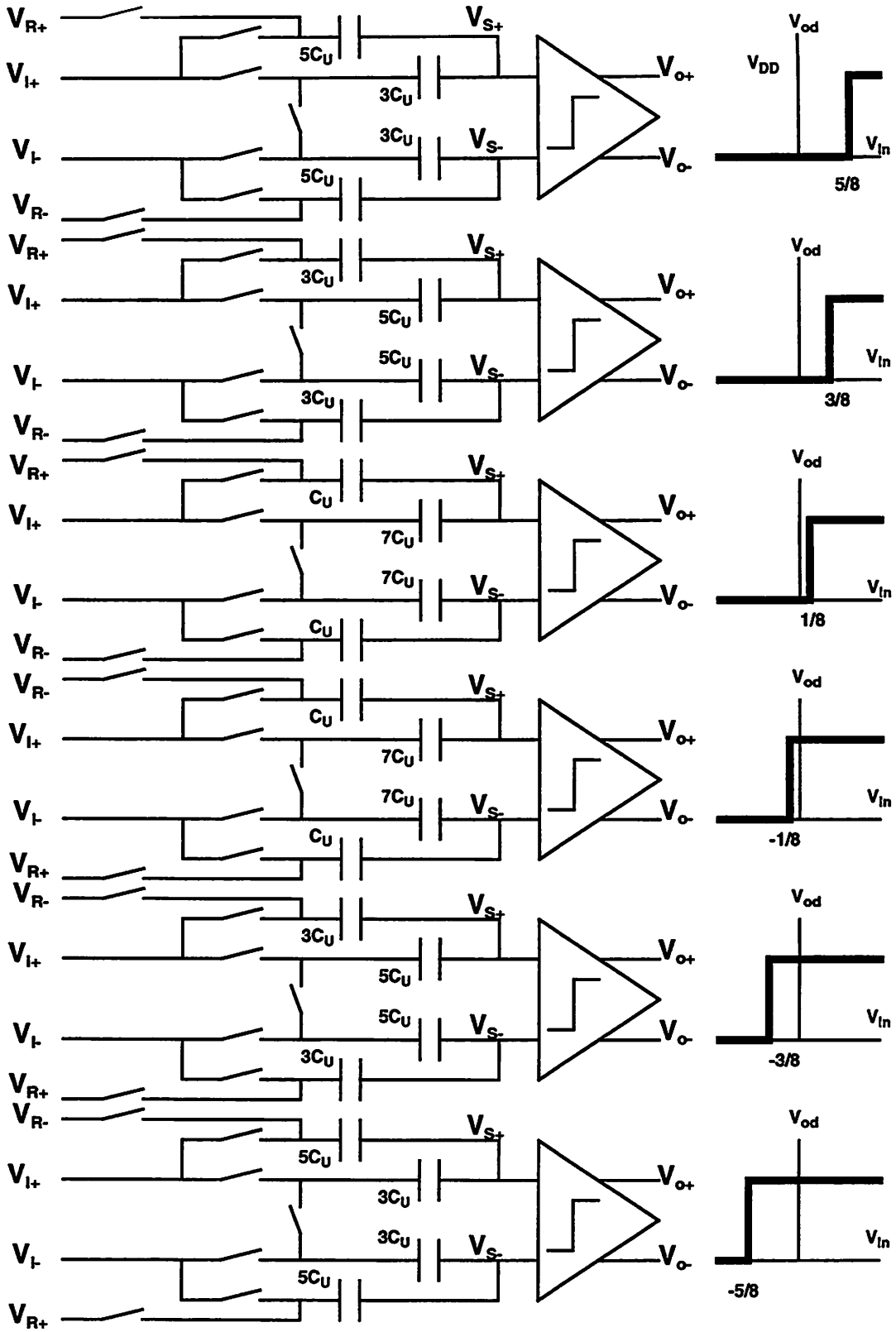


Figure 6.4 Comparator Bank Architecture

$$V_{od} = V_{o+} - V_{o-} \qquad V_{In} = \frac{V_{I+} - V_{I-}}{V_{R+} - V_{R-}}$$

6.2.2 Encoding Network

6.2.2.1 Voting Error Correction

The input offset voltage of the comparator described in the previous section can be quite large. In order to improve the tolerance to offsets in this comparator, extra digital circuitry was added following the comparator outputs to correct for gross errors[Mangelsdorf90][Flynn95]. This scheme has been referred to as voting error correction and addresses the case where the input offset voltage of a comparator is so large that the hierarchy of switching thresholds shown in figure 6.4 is altered. Figure 6.5 shows an example of the comparator outputs for such a case and the desired correction. Note that the large offset results in a thermometer code with a "bubble" at the comparator outputs.

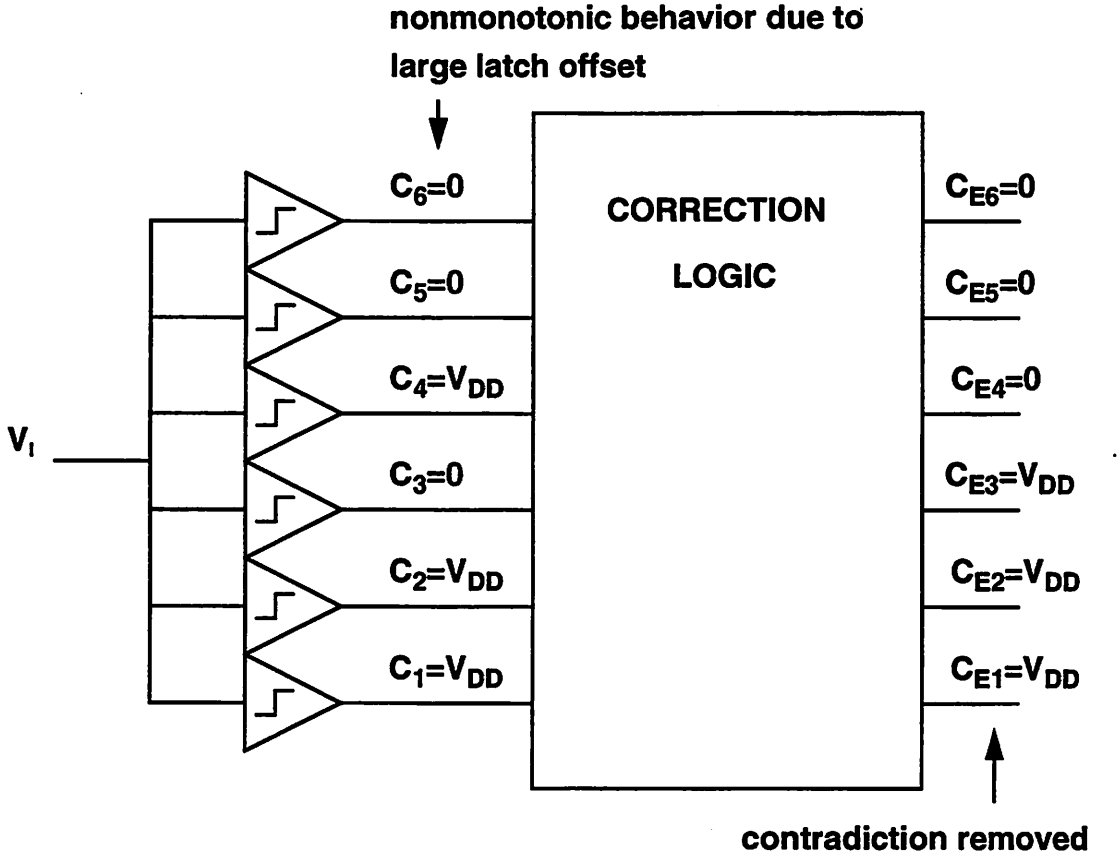


Figure 6.5 Example of Voting Error Correction in Action

To correct for this mistake, correction logic is added to make sure that the output C_{Ek+1} is always

less than or equal to the output C_{Ek}. A logic diagram showing the implementation of this scheme is shown in figure 6.6.

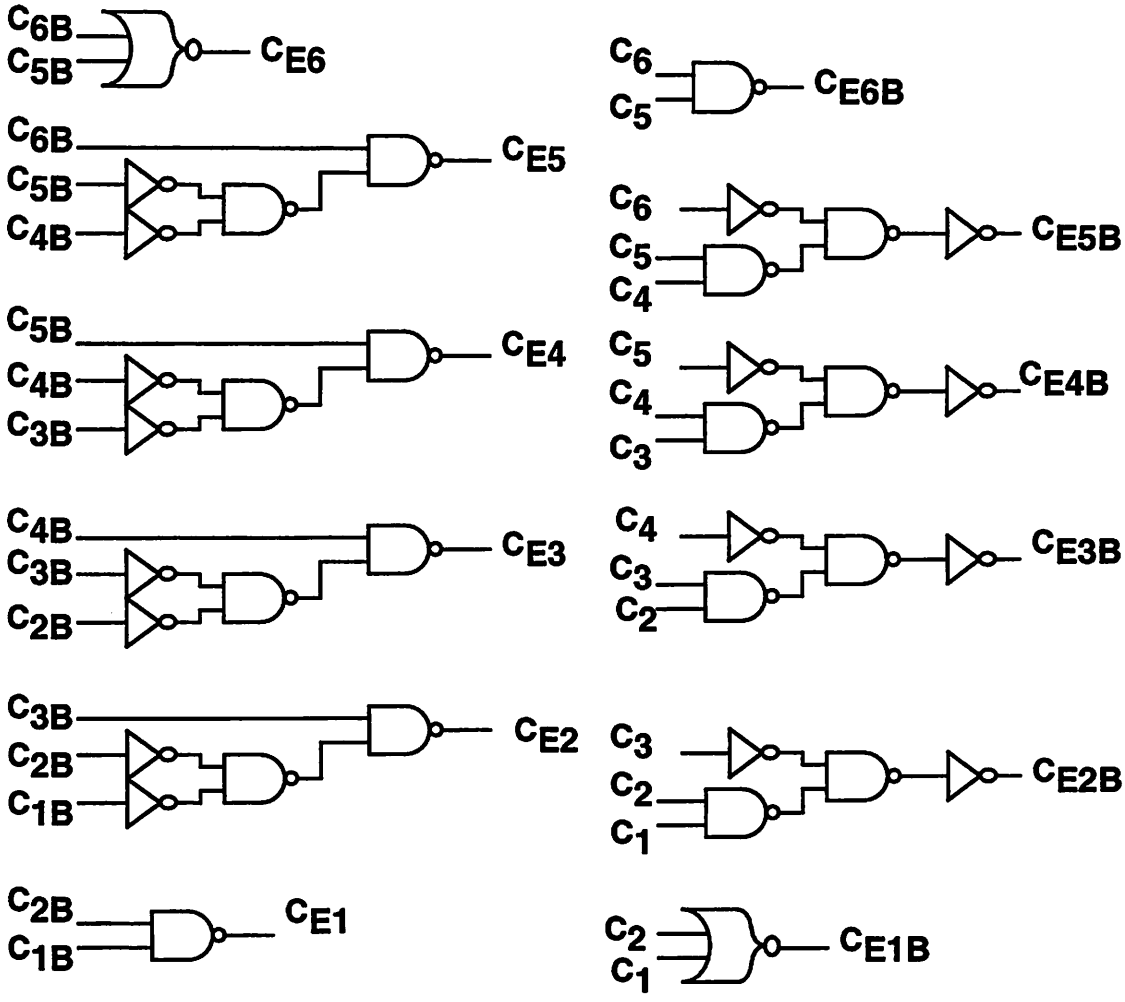


Figure 6.6 Logic Implementation of Voting Error Correction

6.2.2.2 Digital to Analog Converter

The digital to analog converter (DAC) block is an analog multiplexer (MUX). This block uses the comparator outputs as control signals that select a reference voltage to be subtracted from the input signal. The DAC used in the prototype uses a charge redistribution architecture[McCreary75][Lin91]. This operation of this type of DAC is intimately tied to the subtraction operation and the

operation of the sample and hold amplifier. Therefore, all three of these must be discussed together.

The subtraction operation is obtained by operating the sample and hold on two phases as discussed in section 3.1. During the sample phase, the signal to be sampled is connected to the sampling capacitor. During the hold phase, the signal to be subtracted is connected. Fractional values of the reference can be generated by dividing the sampling capacitor into smaller units and connecting some capacitors to the reference and some to ground. Therefore, the DAC operates by connecting the appropriate reference to the appropriate capacitor.

Figure 6.7 shows how the DAC is organized. This figure also shows the digital calibration scheme. The calibration is used to measure errors in the DAC. Therefore, its operation is closely associated with the DAC. The digital calibration scheme shown is similar to that discussed by Karanicolas[Karanicolas93]. For the case of the prototype chip, the calibration was not implemented on the chip, but was controlled manually off the chip. Therefore, the calibration circuits shown illustrate how the calibration might be implemented if it were included on the chip.

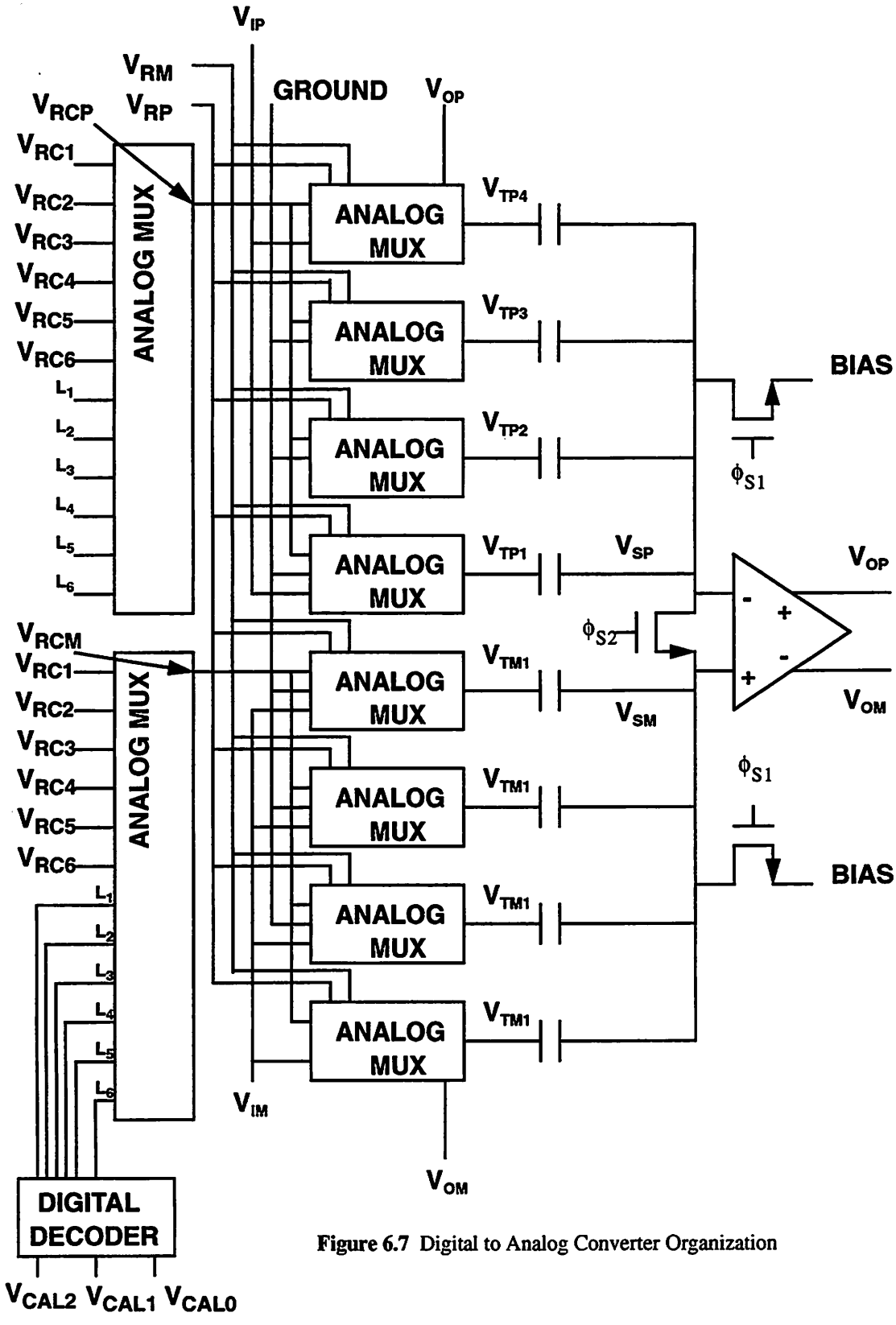
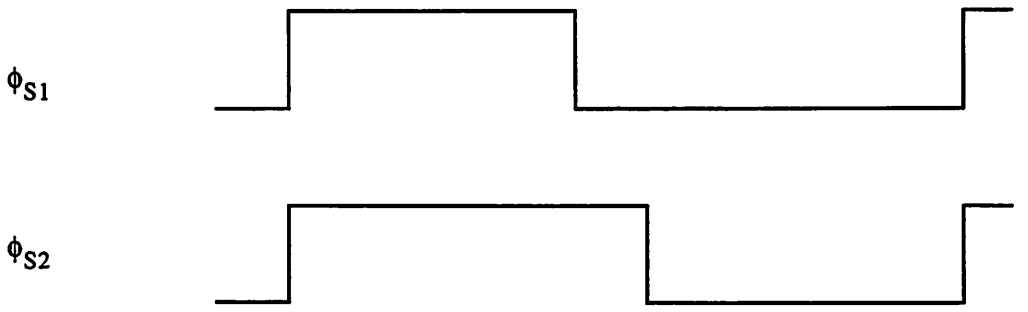


Figure 6.7 Digital to Analog Converter Organization

The timing diagram shown in figure 6.8 helps to illustrate the how the DAC works with the sample and hold. During the sampling phase of operation, the capacitors are connected either to



	SAMPLING PHASE	HOLD PHASE
possible sampling capacitor connections	ADC input signal previous stage output V_{RC}	V_{RP} V_{RM} 0

Figure 6.8 Possible Locations for Connecting Sampling Capacitor During the Sampling Phase and Hold Phase of Operation

the input to the ADC, the output of the previous stage of the ADC, or to V_{RC}, depending on the mode of operation. During calibration, the capacitors are connected to V_{RCP} and V_{RCM}. During normal operation, the capacitors are connected to the output of the previous stage. In the case of the first stage of the ADC, the capacitors are connected to the input signal.

Next, the switches biasing the summing nodes V_{SP} and V_{SM} are turned off. The switches are turned off in an ordered way to minimize the differential charge injection[Lewis87A][Lin91][Conroy93]. The switch controlled by ϕ_{S1} turns off first. The switch controlled by ϕ_{S2} is the next to turn off. These switches both turn off before any of the switches in the DAC turn off.

After the switches biasing V_{SP} and V_{SM} are turned off, V_{SP} and V_{SM} are floating. During the next phase of operation, the hold phase, capacitors 1 through 3 are connected either to the positive reference V_{RP}, the negative reference V_{RM}, or ground. Capacitors C_{SP4} and C_{SM4} are the feedback capacitors. Therefore, these capacitors are always connected to the amplifier output. During normal operation, the node to which each of the sampling capacitors is connected is controlled by the comparator outputs. During calibration mode, the node to which each of the sampling capacitors is connected is determined by the calibration control inputs V_{CAL0}, V_{CAL1}, and V_{CAL2}. These inputs also determine which resistor string voltage is connected to V_{RC}.

The resistor string shown in figure 6.9 is used to generate a set of voltages that are approximately equal to the switching threshold voltages of the comparators as discussed in section 3.4.5. It is needed only during calibration and may be turned off at other times.

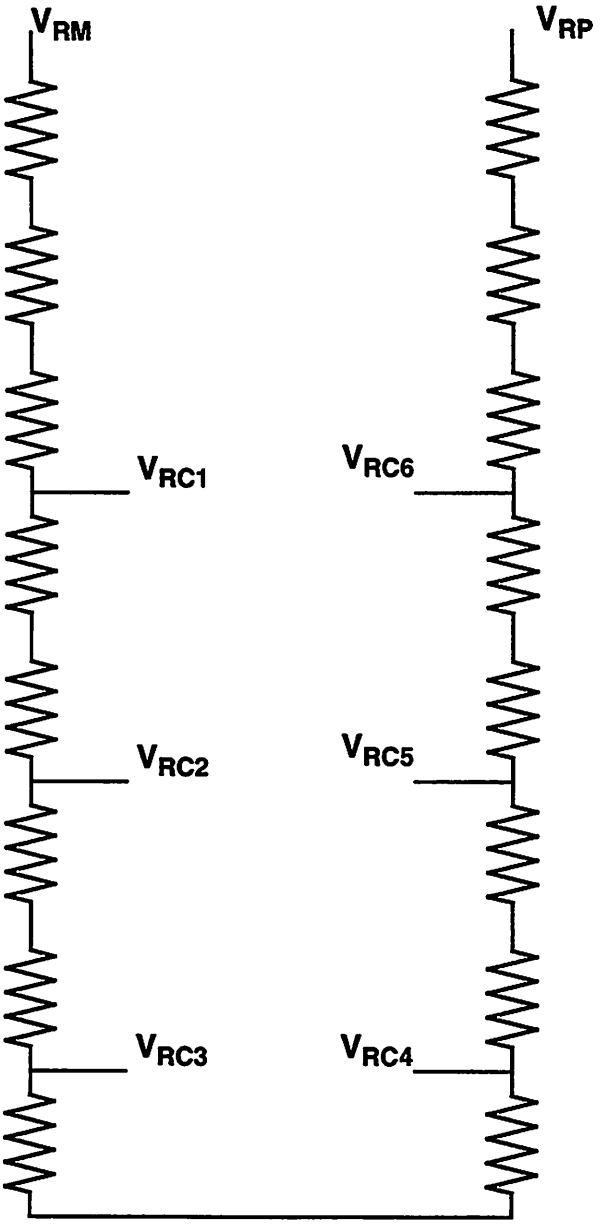


Figure 6.9 Resistor String for Generating Input Signal Near Comparator Switching Thresholds

The truth tables in tables 6.1 through 6.5 further illustrate the operation of the DAC circuits described above.

During calibration, the input signal is near a comparator switching threshold voltage. Therefore, it is admissable during this time to subtract either the closest reference below the switching threshold or the closest reference above the switching threshold. For both cases, the resulting residue will be within the range of values that can be handled by the following stages of the ADC. Subtracting the lower reference will result in a positive residue, and subtracting the higher reference will result in a negative residue. By subtracting these two residues, the size of the DAC segment can be determined as discussed in section 3.4.5. Since that is the purpose of performing the calibration, the calibration mode is divided into two smaller modes, one where the lower reference is subtracted and one where the higher reference is subtracted. The operation is summarized in tables 6.4 and 6.5.

Table 6.1 Sampling Capacitor Connections During Sampling Phase of Operation

	V_{TP1}	V_{TP2}	V_{TP3}	V_{TP4}	V_{TM1}	V_{TM2}	V_{TM3}	V_{TM4}
normal operation	V_{OPk}	V_{OPk}	V_{OPk}	V_{OPk}	V_{OMk}	V_{OMk}	V_{OMk}	V_{OMk}
calibration mode	V_{RCP}	V_{RCP}	V_{RCP}	V_{RCP}	V_{RCM}	V_{RCM}	V_{RCM}	V_{RCM}

Table 6.2 Sampling Capacitor Connection During Hold Phase of Normal Operation

L_6	L_5	L_4	L_3	L_2	L_1	V_{TP1}	V_{TP2}	V_{TP3}	V_{TP4}	V_{TM1}	V_{TM2}	V_{TM3}	V_{TM4}
0	0	0	0	0	0	V_{RM}	V_{RM}	V_{RM}	V_{oP}	V_{RP}	V_{RP}	V_{RP}	V_{om}
0	0	0	0	0	1	V_{RM}	V_{RM}	0	V_{oP}	V_{RP}	V_{RP}	0	V_{om}
0	0	0	0	1	1	V_{RM}	0	0	V_{oP}	V_{RP}	0	0	V_{om}
0	0	0	1	1	1	0	0	0	V_{oP}	0	0	0	V_{om}
0	0	1	1	1	1	V_{RP}	0	0	V_{oP}	V_{RM}	0	0	V_{om}
0	1	1	1	1	1	V_{RP}	V_{RP}	0	V_{oP}	V_{RM}	V_{RM}	0	V_{om}

Table 6.2 Sampling Capacitor Connection During Hold Phase of Normal Operation

L_6	L_5	L_4	L_3	L_2	L_1	V_{TP1}	V_{TP2}	V_{TP3}	V_{TP4}	V_{TM1}	V_{TM2}	V_{TM3}	V_{TM4}
1	1	1	1	1	1	V_{RP}	V_{RP}	V_{RP}	V_{oP}	V_{RM}	V_{RM}	V_{RM}	V_{om}

Table 6.3 Calibration Control Decoder Truth Table

V_{CAL2}	V_{CAL1}	V_{CAL0}	L_6	L_5	L_4	L_3	L_2	L_1	V_{RC}=?
0	0	0	0	0	0	0	0	V_{DD}	V_{RC1}
0	0	V_{DD}	0	0	0	0	V_{DD}	0	V_{RC2}
0	V_{DD}	0	0	0	0	V_{DD}	0	0	V_{RC3}
0	V_{DD}	V_{DD}	0	0	V_{DD}	0	0	0	V_{RC4}
V_{DD}	0	0	0	V_{DD}	0	0	0	0	V_{RC5}
V_{DD}	0	1	V_{DD}	0	0	0	0	0	V_{RC6}
V_{DD}	V_{DD}	0	0	0	0	0	0	0	not conn.
V_{DD}	V_{DD}	V_{DD}	0	0	0	0	0	0	not conn.

Table 6.4 Sampling Capacitor Connection During Hold Phase of Calibration Mode - Lower Reference Subtracted

L_6	L_5	L_4	L_3	L_2	L_1	V_{TP1}	V_{TP2}	V_{TP3}	V_{TP4}	V_{TM1}	V_{TM2}	V_{TM3}	V_{TM4}
0	0	0	0	0	1	V_{RM}	V_{RM}	V_{RM}	V_{oP}	V_{RP}	V_{RP}	V_{RP}	V_{om}
0	0	0	0	1	0	V_{RM}	V_{RM}	0	V_{oP}	V_{RP}	V_{RP}	0	V_{om}
0	0	0	1	0	0	V_{RM}	0	0	V_{oP}	V_{RP}	0	0	V_{om}
0	0	1	0	0	0	0	0	0	V_{oP}	0	0	0	V_{om}
0	1	0	0	0	0	V_{RP}	0	0	V_{oP}	V_{RM}	0	0	V_{om}
1	0	0	0	0	0	V_{RP}	V_{RP}	0	V_{oP}	V_{RM}	V_{RM}	0	V_{om}

Table 6.5 Sampling Capacitor Connection During Hold Phase of Calibration Mode - Higher Reference Subtracted

L_6	L_5	L_4	L_3	L_2	L_1	V_{TP1}	V_{TP2}	V_{TP3}	V_{TP4}	V_{TM1}	V_{TM2}	V_{TM3}	V_{TM4}
0	0	0	0	0	1	V_{RM}	V_{RM}	0	V_{oP}	V_{RP}	0	0	V_{om}
0	0	0	0	1	0	V_{RM}	0	0	V_{oP}	0	0	0	V_{om}
0	0	0	1	0	0	0	0	0	V_{oP}	0	0	0	V_{om}
0	0	1	0	0	0	V_{RP}	0	0	V_{oP}	V_{RM}	0	0	V_{om}
0	1	0	0	0	0	V_{RP}	V_{RP}	0	V_{oP}	V_{RM}	V_{RM}	0	V_{om}
1	0	0	0	0	0	V_{RP}	V_{RP}	V_{RP}	V_{oP}	V_{RM}	V_{RM}	V_{RM}	V_{om}

6.2.3 Sample and Hold Amplifier Architecture

Because the sample and hold, DAC, and subtraction operations are combined, the operation of this circuit was already largely discussed in the previous section. In this section, its operation is discussed again with a different emphasis.

The basic structure of the sample and hold circuit used in the prototype is shown in figure 6.10. This circuit is the standard switched capacitor sample and hold circuit used in charge redistribution analog to digital converters[McCreary75][Lewis87A][Lin91][Conroy93].

The operation of the sample and hold circuit is as follows. The circuit has two phases of operation: the sampling phase and the hold phase. During the sampling phase, ϕ_{S1}, ϕ_{S2}, and ϕ_{S3} are high. During this time the amplifier output nodes are tied to an output bias voltage. The amplifier input nodes V_{SP} and V_{SM} are also tied to a bias voltage at this time. The capacitor bottom plates V_{P1}, V_{P2}, V_{P3}, V_{P4}, V_{M1}, V_{M2}, V_{M3}, and V_{M4} are all tied to the appropriate input voltage, either V_{IP} or V_{IM}. The input voltage is either the input to the adc or the output from the previous stage during normal operation. During calibration, the input is one of the comparator switching thresholds. At the end of the sampling phase, the switches connecting the input to the bias voltage are turned off first. The charge injected by these switches is then equilibrated by the switch shorting the two input nodes together. This switch turns off next. The time at which this switch turns off defines the

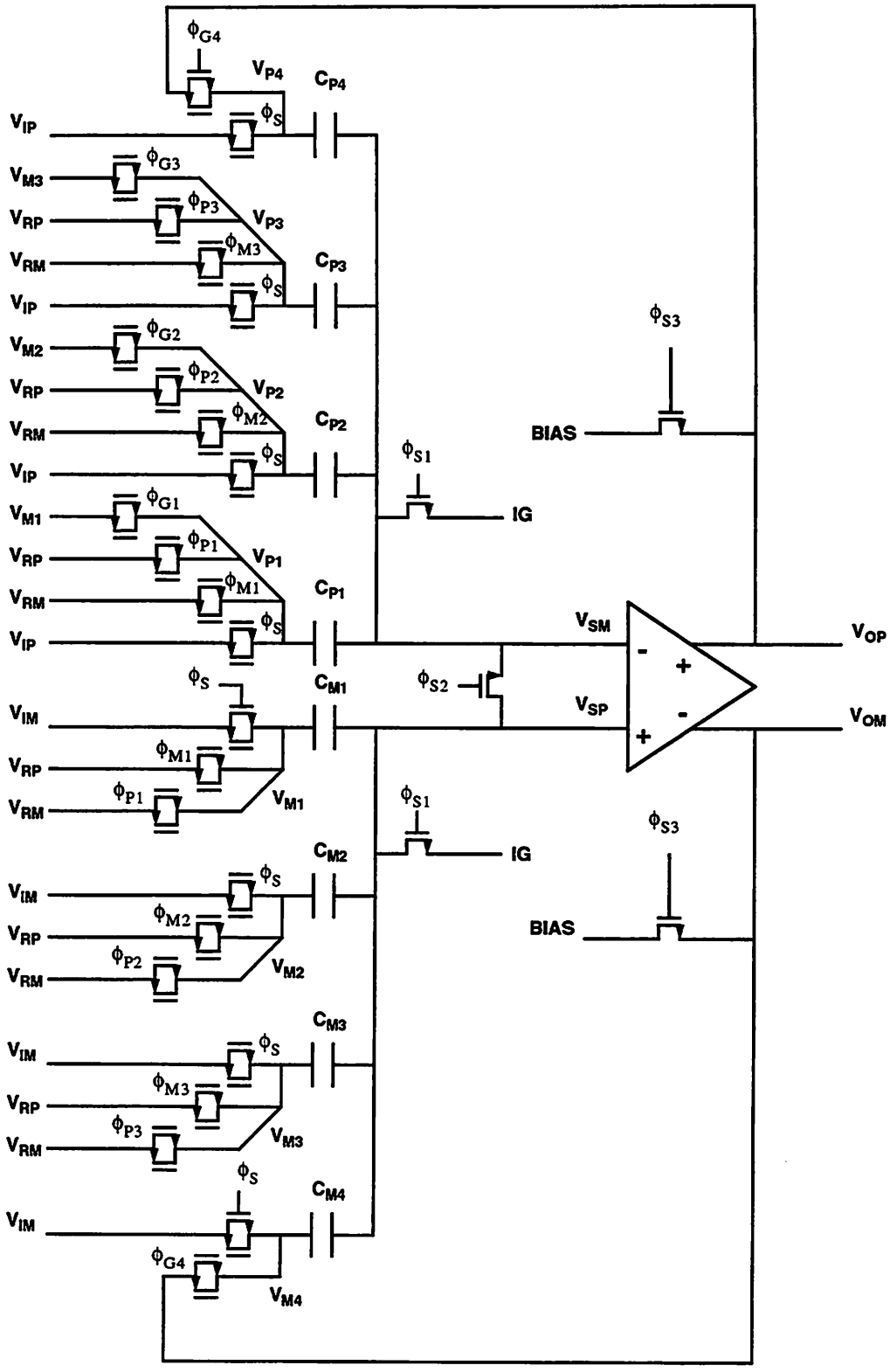


Figure 6.10 Sample and Hold Amplifier and Differencing Circuit

sampling instant. Lastly, the switches connecting the input signal to the capacitor bottom plates and the switches biasing the amplifier outputs are turned off.

During the hold phase of operation, the amplifier input nodes float. At this time, the capacitor bottom plates are connected to either a reference voltage or another capacitor bottom plate. The connection made is determined by the comparator outputs during normal operation and by the calibration inputs during calibration mode.

6.2.4 Operational Amplifier Architecture

The sample and hold amplifier discussed above includes an operational amplifier. This operational amplifier is a key block that limits the performance of the analog to digital converter. This block is the primary source of speed limitation and power dissipation. This block also contributes thermal noise limiting the ADC resolution.

Figure 6.11 shows the operational amplifier architecture used in the prototype design.

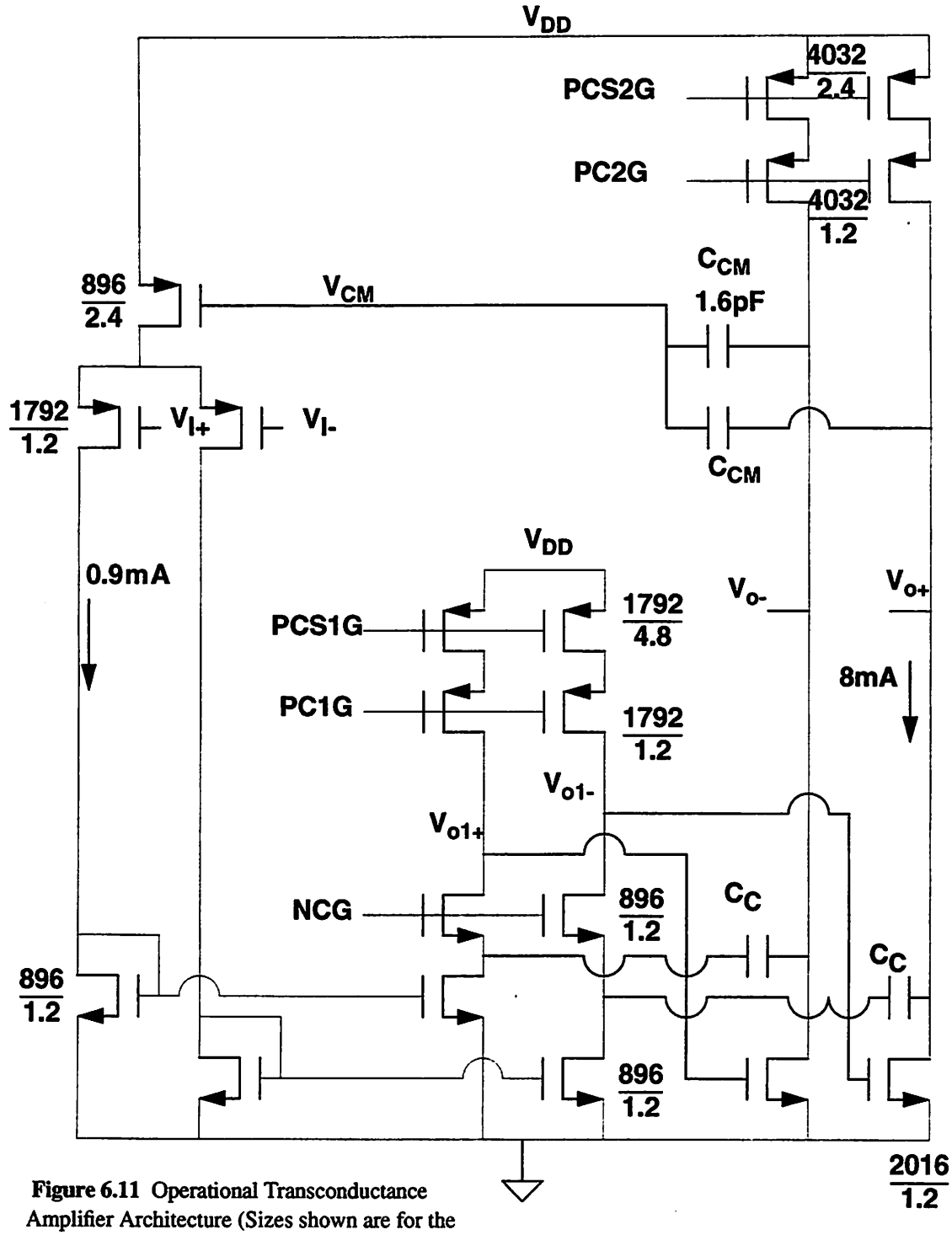


Figure 6.11 Operational Transconductance Amplifier Architecture (Sizes shown are for the front end.)

The transistor sizes given are for the first stage. The architecture is basically a two stage design with pole splitting Miller compensation modified according to the technique discussed by Ahuja[Ahuja83]. The settling time performance of this amplifier is also discussed in chapter 4, and the thermal noise is discussed in chapter 5. The amplifier is fully differential in order to reduce power supply coupling. Because of the fully differential architecture, common mode feedback is necessary. This feedback is provided by capacitors C_{CM}. If the common mode voltage at the output rises, the voltage at node V_{CM} also rises. This results in reduced current through the input transistors and reduced current in the NMOS transistors sinking current from nodes V_{o1+} and V_{o1-}. Therefore, the voltage at these two nodes goes up, and this results in more current being sunk from nodes V_{o+} and V_{o-}. Therefore, the common mode voltage at the output is pulled back down.

A one stage amplifier architecture is often favored because of its simplicity and potential for high speed. However, distortion specifications require the use of amplifiers with high open loop gain. To meet open loop gain requirements, a one stage amplifier requires the use of Cascode transistors that limit the swing available from the amplifier. For a fixed dynamic range specification, the reduced swing results in tighter limits on the noise. To meet the tighter requirements, large capacitors must be used. If the sampling rate specification is fixed, the larger capacitors force the use of higher bias currents. Therefore, the two stage architecture was chosen for this prototype. The output stage has a moderate gain of about 10 and has a wide voltage swing. Because of the gain of the second stage, the output of the first stage does not need to have a large swing. Therefore, it is possible to use Cascode transistors in this stage to achieve high gain. Thus, a high open loop gain amplifier is achieved without losing signal swing.

6.2.5 Bias Circuit for the Operational Amplifier

The operational amplifier bias circuit is largely based on the circuit in figure 6.12. This circuit is discussed by Lin[Lin90] and is used to bias the gates of the current sources and the cascode transistors in the operational amplifier. The basic principle of the circuit is as follows. The current I_B flowing through MB1 and MB2 forces MB1 into the triode region and MB2 into the saturation region. By appropriate choice of current and transistor sizes, the value of V_C can be adjusted to bias a cascode transistor above a current source.

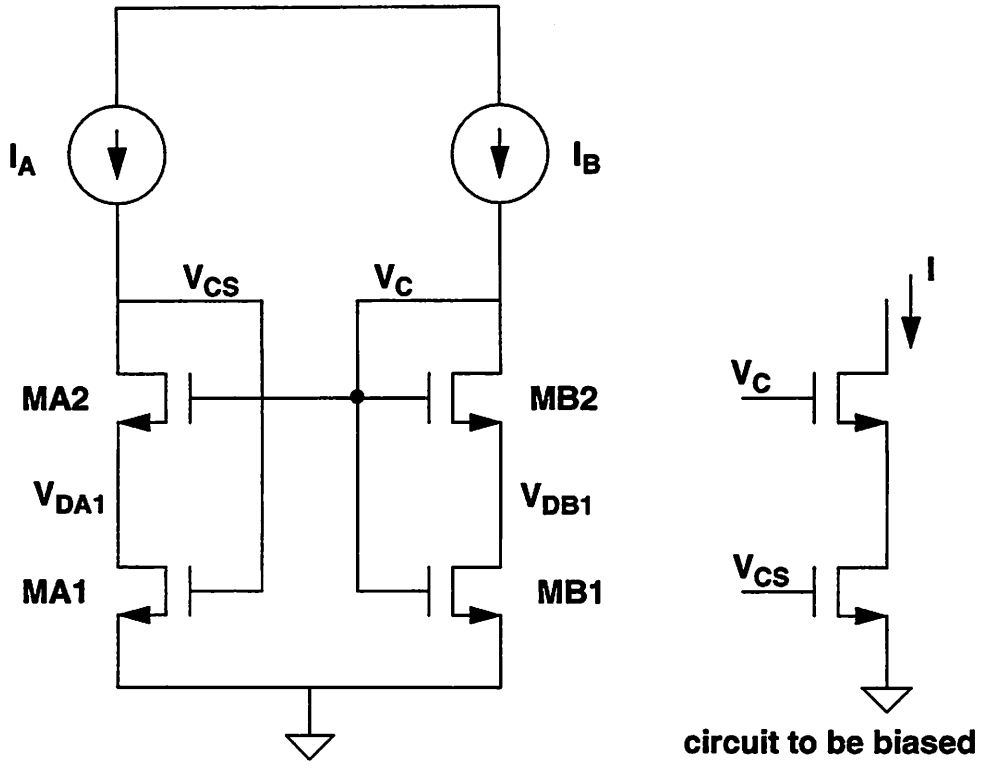


Figure 6.12 Current Source and Cascode Biasing Scheme

The circuit is also used to generate the gate voltage V_{CS} of the current source. This voltage is often generated by running a bias current through a diode connected transistor. The problem with that method is that the drain to source voltage on the biasing transistor often does not match the drain to source voltage of the transistor it is biasing. To alleviate this problem, the circuit is modified by adding cascode transistor MA2 to the bias circuit. This modification results in a more accurate bias voltage for the current source by forcing the drain voltage of the bias transistor to be similar to that of the current source.

The following analysis gives a first order expression for the bias voltage. This analysis assumes long channel models for the transistors. The current I_B is expressed as the following function of node voltages.

$$I_B = \frac{1}{2}\mu C_{ox}\frac{W_{B2}}{L_{B1}}(V_C - V_{DB1} - V_T)^2 = \mu C_{ox}\frac{W_{B1}}{L_{B1}}\{(V_C - V_T)V_{DB1} - \frac{1}{2}V_{DB1}^2\} \quad (6.1)$$

From the above equation, the drain voltage V_{DB1} is given by the following equation.

$$V_{DB1} = \sqrt{\frac{2I_B}{\mu C_{ox}\frac{W_{B2}}{L_{B2}}}}\left(-1+\sqrt{1+\left(\frac{W_{B2}}{W_{B1}}\right)\left(\frac{L_{B1}}{L_{B2}}\right)}\right) \tag{6.2}$$

Also, the bias voltage V_C is given by the following equation.

$$V_C = V_T + \sqrt{\frac{2I_B}{\mu C_{ox}}\left(\frac{L_{B2}}{W_{B2}}+\frac{L_{B1}}{W_{B1}}\right)} \tag{6.3}$$

Thus, by appropriately choosing the current and the device sizes, the desired bias voltage can be obtained.

Since good matching between V_{DA1} and V_{DB1} implies that the bias current I will be closely matched to the reference current I_A, it is worthwhile to calculate V_{DA1}-V_{DB1} in order to indicate how well the currents match. Both MA1 and MA2 are in the saturation region of operation. Therefore, V_{DA1} is given as follows.

$$V_{DA1} = \sqrt{\frac{2I_B}{\mu C_{ox}\frac{W_{B2}}{L_{B2}}}}\left(-\sqrt{\frac{I_A}{I_B}\left(\frac{W_{B2}}{W_{A2}}\right)\left(\frac{L_{A2}}{L_{B2}}\right)}+\sqrt{1+\left(\frac{W_{B2}}{W_{B1}}\right)\left(\frac{L_{B1}}{L_{B2}}\right)}\right) \tag{6.4}$$

The mismatch is then given by the following equation.

$$V_{DA1}-V_{DB1} = \sqrt{\frac{2I_B}{\mu C_{ox}\frac{W_{B2}}{L_{B2}}}}-\sqrt{\frac{2I_A}{\mu C_{ox}\frac{W_{A2}}{L_{A2}}}} \tag{6.5}$$

From the above equation, it can be seen that if the current densities in MA2 and MB2 are matched, the drain voltages will match. Thus, the bias current will closely track the reference current.

The output impedance of the bias circuit is also of interest. The bias circuit is designed to generate a DC voltage. However, switching events in the circuit being biased may result in transient deviations in the bias voltages. In order to achieve rapid recovery from these switching events, it is desirable that the output impedance be low. The output impedance can be determined by differen-

tiating the bias voltage with respect to the bias current. The result is that the output resistance R_C of the cascode bias circuit and R_{CS} of the current source bias circuit are given by the following equations.

$$R_C = \frac{V_C - V_T}{2I_B} \tag{6.6}$$

$$R_{CS} = \frac{V_{CS} - V_T}{2I_A} \tag{6.7}$$

Thus, for a given bias voltage, the output resistance can only be reduced by increasing the bias current. Therefore, there is a direct trade-off between output resistance and power dissipation of the bias circuit.

The complete circuit used to bias the operational amplifier is shown in figure 6.13.

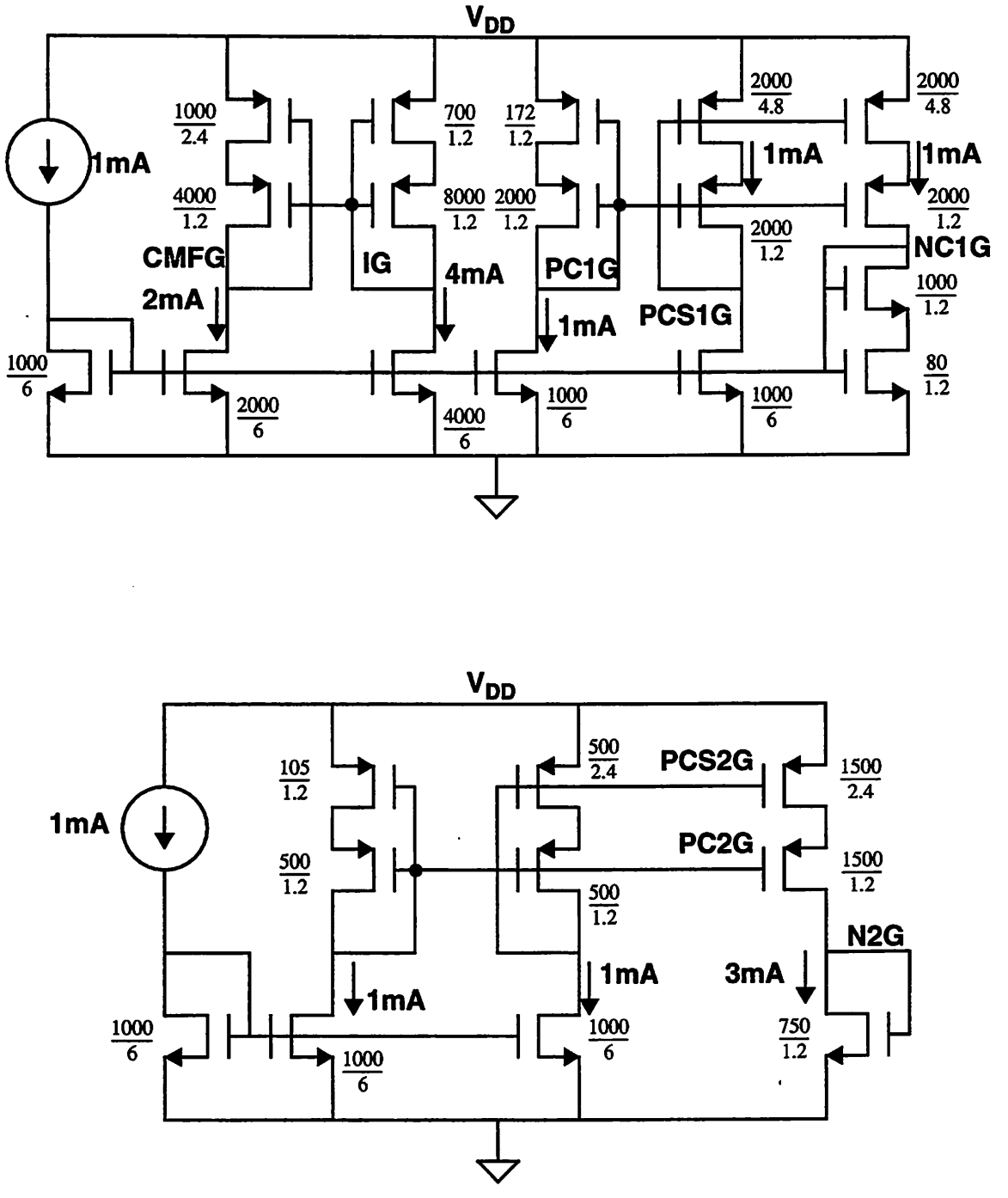


Figure 6.13 Bias Circuit for the Operational Amplifier in Stages 0 and 1 of the Pipelined ADC

6.2.6 Clock Generation

In order to control the timing of operation, it is necessary to generate a set of clock signals with two nonoverlapping phases. It is essential that the sampling phase and hold phase do not overlap. Overlapping of the phases would result in signal charge leaking away during the hold phase.

The circuit used for generating these nonoverlapping clocks in the prototype design is shown

in figure 6.14. A timing diagram showing how these clocks are generated is shown in figure 6.15.

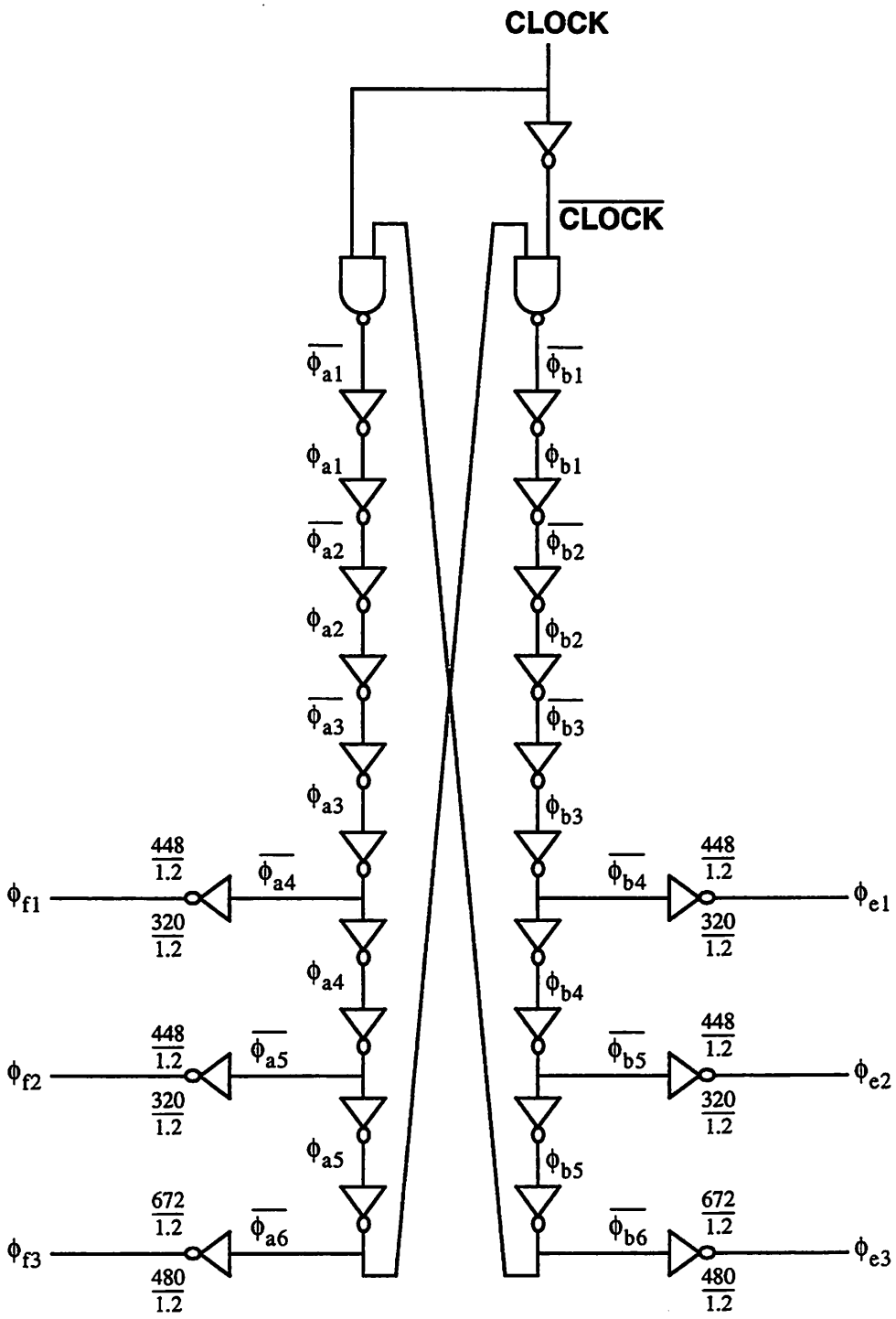


Figure 6.14 2 Phase Nonoverlapping Clock Generator

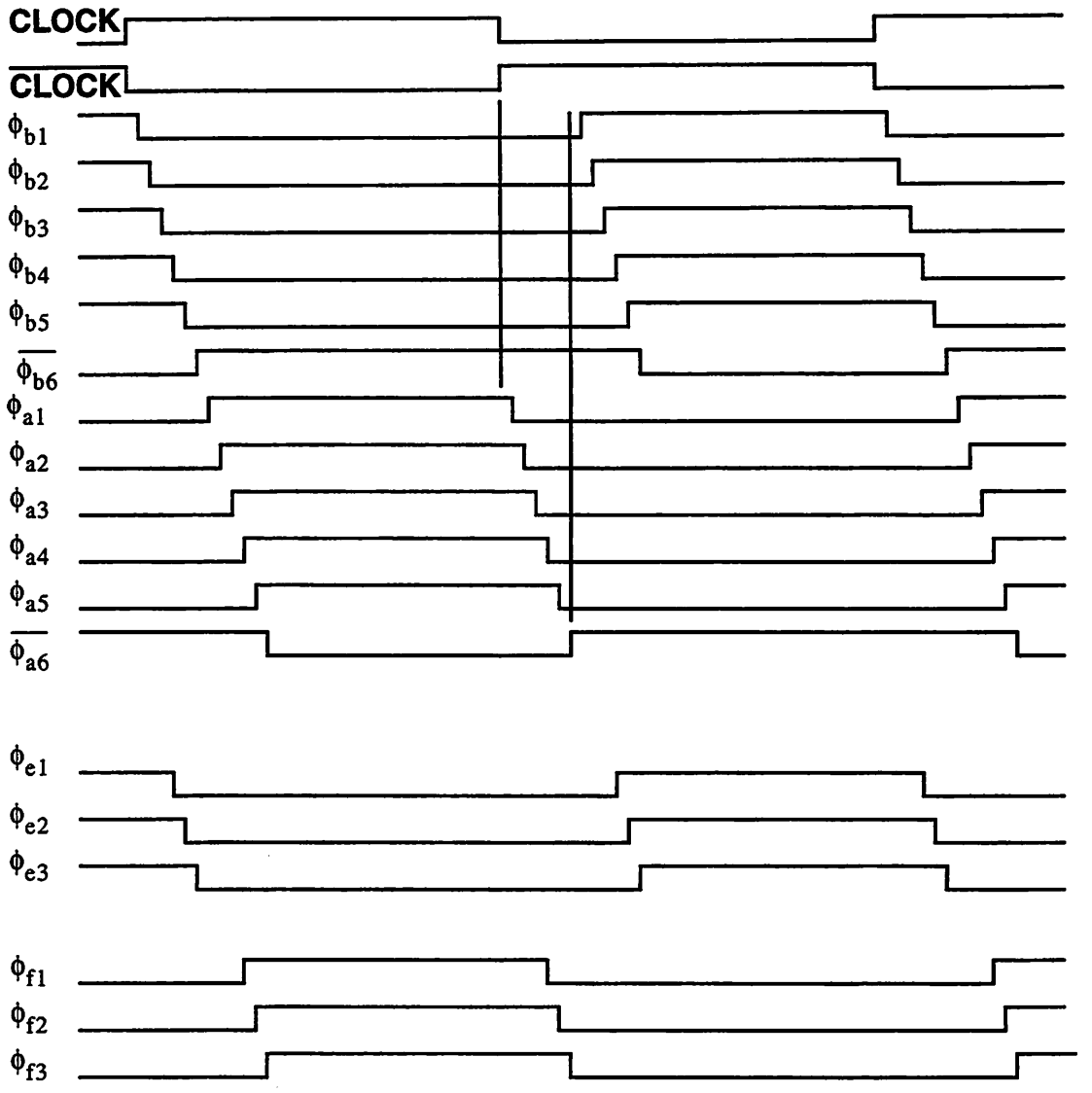


Figure 6.15 Timing Diagram for Clock Generation Circuit

This circuit is similar to the circuit used by Lewis[Lewis87B]. CLOCK is an external input signal to the integrated circuit. The clocks ϕ_{e1}, ϕ_{e2}, and ϕ_{e3} correspond to sampling clocks ϕ_{s1}, ϕ_{s2}, and ϕ_{s3} respectively for stages 0, 2, 4, 6, and 8. For stages 1, 3, 5, and 7 ϕ_{e1}, ϕ_{e2}, and ϕ_{e3} correspond to hold clocks ϕ_{h1}, ϕ_{h2}, and ϕ_{h3}. Likewise, clocks ϕ_{f1}, ϕ_{f2}, and ϕ_{f3} correspond to sampling clocks ϕ_{s1}, ϕ_{s2}, and ϕ_{s3} for stages 1, 3, 5, and 7 and hold clocks ϕ_{h1}, ϕ_{h2}, and ϕ_{h3} for stages 0, 2, 4, 6, and 8. The three skewed clocks are generated

for the sampling phase in order to minimize the effects of charge injection[Lewis87A][Lin91][Conroy93]. On the falling edge of ϕ_{s1} the switch connecting the operational amplifier inputs to the bias value are turned off. As these switches turn off, charge is injected from them onto the input nodes of the operational amplifier. However, another switch controlled by ϕ_{s2} equalizes the charge injection so that the differential charge injection is zero. Then this switch turns off on the falling edge of ϕ_{s2}. Charge is also injected from this switch, but ideally the charge injection should be equal for both nodes. Finally, on the falling edge of ϕ_{s3} the input signal is disconnected from the bottom plate of the sampling capacitor.

6.2.7 Layout Considerations

The resolution desired from the prototype analog to digital converter makes interference a significant concern. Therefore, care was taken in the layout to isolate sensitive nodes in the circuit from sources of interference. This was accomplished primarily by shielding and separate ground and supply voltages.

6.2.7.1 Sampling Capacitor Layout

The sample and hold amplifier input is the most critical node of the integrated circuit. Therefore, a great deal of care was taken to shield this node from sources of interference. The input to the ADC is also rather sensitive. Therefore, special attention was given to the layout of the sampling capacitors, which connect both of these nodes. The sampling capacitors were constructed by

overlapping metal and polysilicon as shown in the cross section of figure 6.16. Note the presence

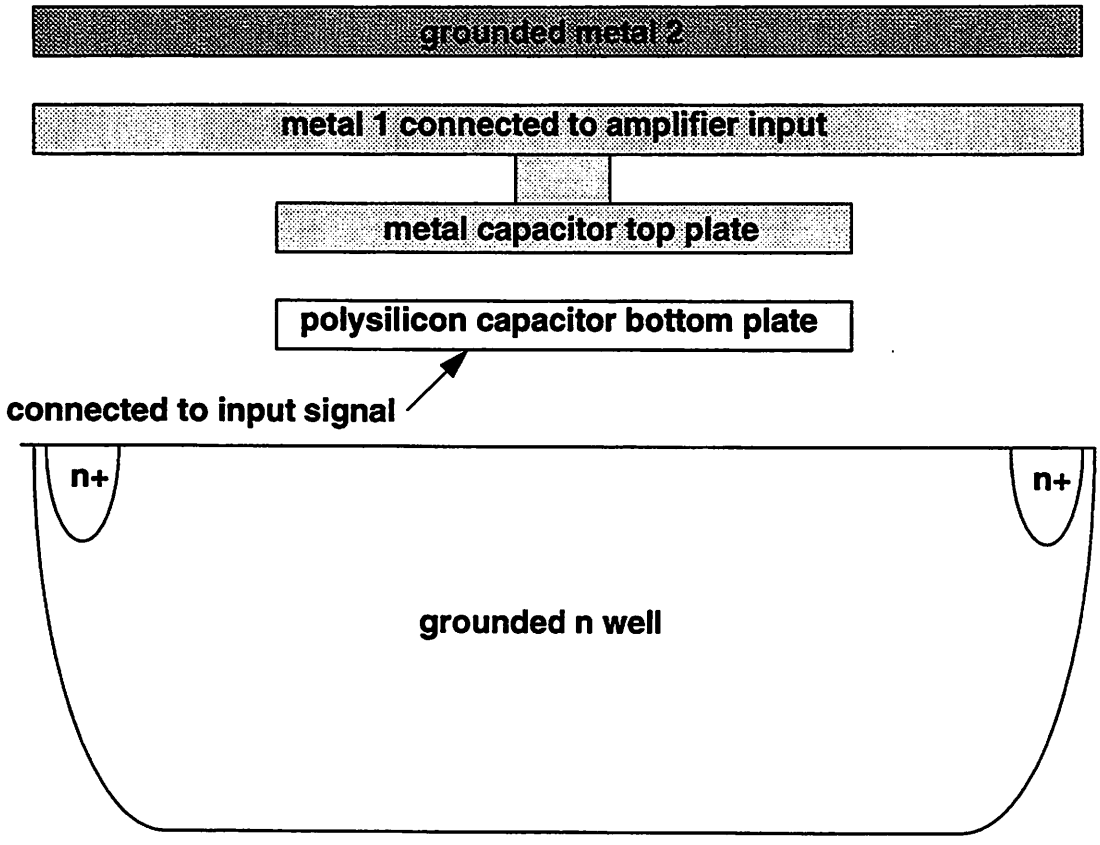


Figure 6.16 Cross Section of a Sampling Capacitor

of grounded shielding both above and below the capacitor. The plan view of the sampling capaci-

tors is shown in figure 6.16. Here again, the attempt to use grounded shielding to protect the ampli-

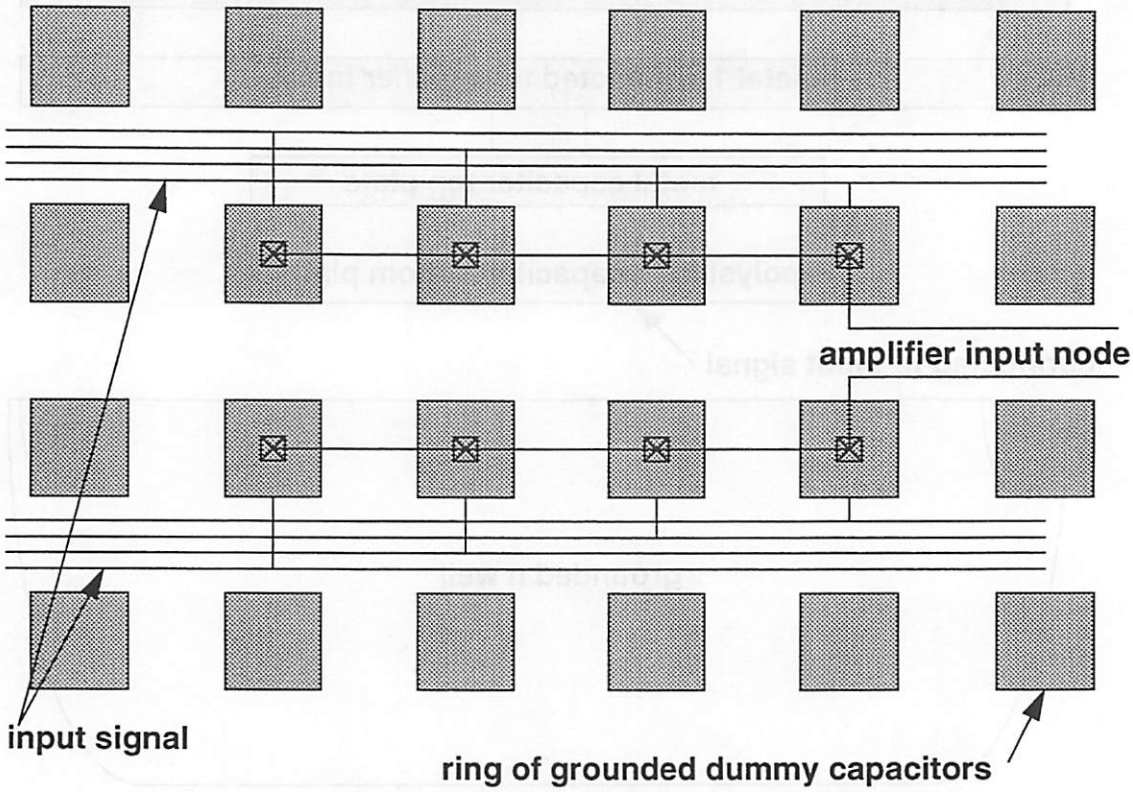


Figure 6.17 Plan View of the Sampling Capacitor Array

fier input node can be seen. The 8 sampling capacitors in the middle of the block are surrounded by a ring of grounded dummy capacitors. These dummy capacitors are used to shield the sampling capacitors from interference and to improve matching by making the edge effects similar[McNutt94][McCreary81][Shyu84].

6.2.7.2 Interconnect Layout

Because the amplifier outputs and inputs are sensitive nodes, the interconnect used to connect the output of one stage to the next stage was also shielded as shown by the cross section in figure

6.18 and the plan view shown in figure 6.19. Keeping the signal lines close together prevents other

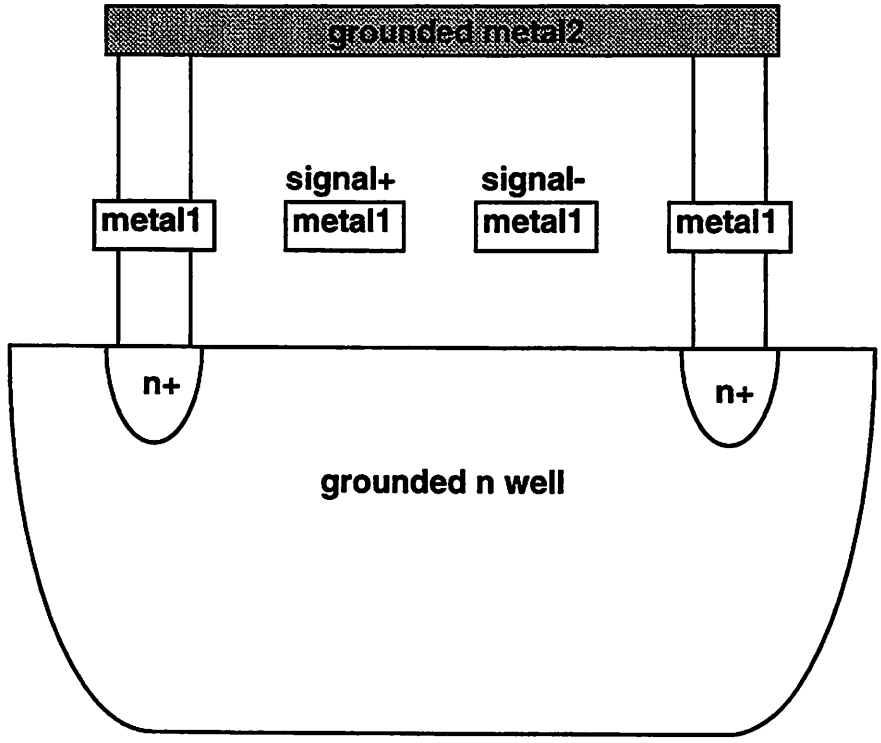


Figure 6.18 Cross Sectional View of the Interconnect Line Between Pipeline Stages

signals from coupling inductively to the signal. Also, the signal lines are surrounded on four sides by grounded conductors to minimize capacitive coupling of sources of interference onto the signal.

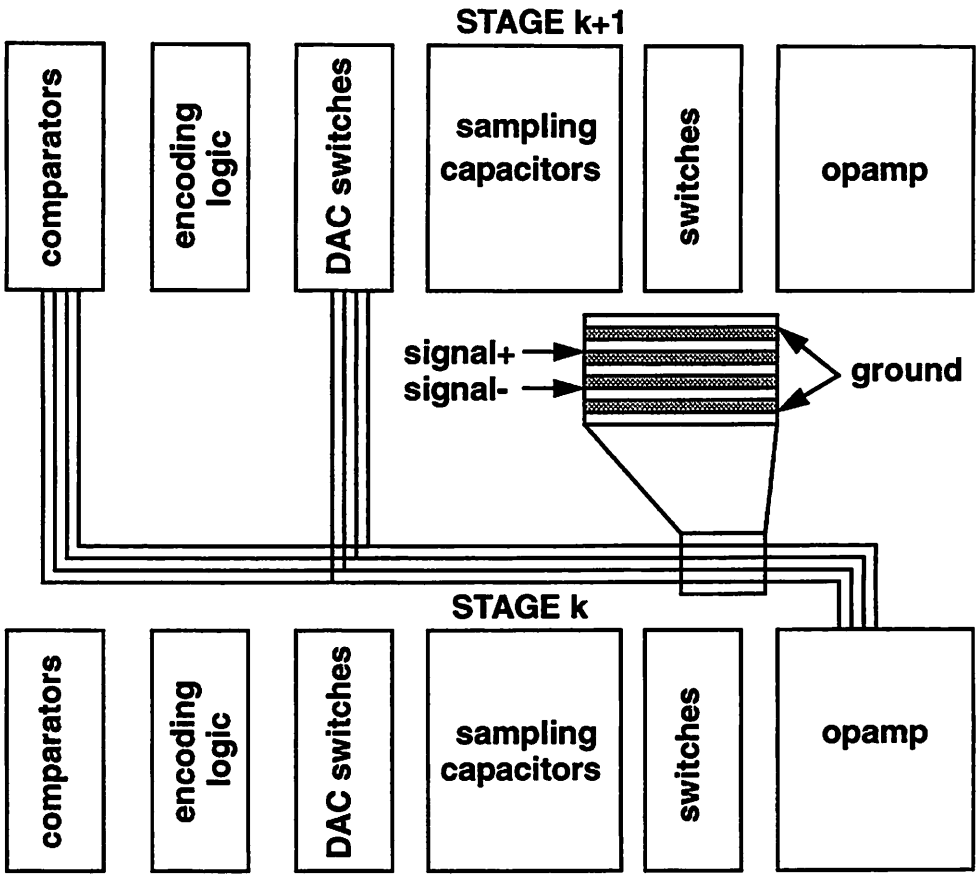


Figure 6.19 Plan View of the Interconnect Between Pipeline Stages

6.3 Power Supply Noise Issues

Supply noise is a significant problem that often limits the performance of high resolution high speed analog circuits. Supply noise occurs when variable currents are injected into the supply. In the ideal case, when the supply impedance is zero, these currents have no effect. However, a real supply voltage source has a finite impedance. As a result, these noise currents cause variations in the supply voltage. These supply voltage variations are then coupled into the signal.

To alleviate this problem, three techniques are used. One technique is to design the circuit in such a way that supply noise is not coupled into the signal. In other words, the circuit is designed such that the power supply rejection ratio (PSRR) is high. One very effective way of accomplish-

ing this is to use fully differential circuit architectures. A fully differential circuit architecture has both positive and negative input and output terminals. Therefore, the signal never has to be referenced to either the positive or negative supply. To first order, the supply noise couples equally onto both the positive and negative signal. When the difference is taken, this supply noise is cancelled out. The fully differential architecture also has the advantage that variations in current are reduced because increases in current in a positive branch of the circuit tend to be balanced by reductions in a negative branch. This technique is very effective, but it has the disadvantage that extra hardware and complexity is required in comparison to a single ended circuit.

Another way of minimizing degradation due to supply noise is to reduce the noise current on the supply. This can be accomplished by using multiple supplies and isolating the supplies requiring low noise from the digital supplies and other noisy supplies. Furthermore, the use of fully differential circuit architectures helps because changes in current in one branch of the circuit tend to be cancelled by the current change in the matching branch. Also, class A circuit designs are helpful because the quantity of current in such a design is relatively independent of the signal in comparison to a class B design. The disadvantage of these techniques is that multiple supplies may drive up the cost of the system. Also, class A designs have high standby power and are poor for handheld and battery operated systems.

Finally, supply noise can be reduced by reducing the impedance of the supply. By reducing the impedance, the quantity of supply noise for a given noise current can be reduced. This method also has limitations. These limitations will be treated further in the following discussion.

The power supply of an integrated circuit may be modeled simply by the circuit in figure 6.20. In this model, the external supply is assumed to be connected to the integrated circuit through inductive bond wires L_B. C represents bypass capacitors present on the chip, and R represents conductive paths between V_{DD} and ground. The current i represents noise currents injected onto the supply by circuits switching on the chip.

From this figure, three approaches to reducing the supply impedance can be seen. First, a reduction of the inductance between the chip and the external supply is quite useful. This can be accomplished by double bonding the supply pads on the IC and by using multiple supply pads. The

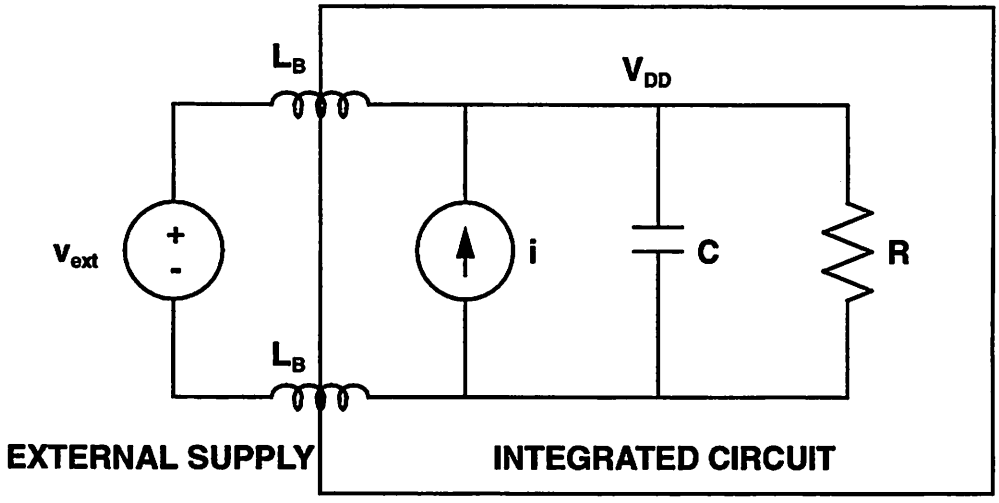


Figure 6.20 Circuit Model for Integrated Circuit Power by an External Supply

inductance per bond wire is generally somewhere in the vicinity of 10nH. Multiple supply pads reduce the supply inductance by putting these bond wires in parallel. Thus, to first order the inductance is given by the following formula.

$$L_{total} = \frac{\text{inductance per pad}}{\text{number of pads}} \tag{6.8}$$

This technique is effective, but extra pads can be expensive and could force the use of a larger package.

Increasing the on chip capacitance is another commonly used technique to reduce the impedance of the on chip supply. The problem here is that in order to be effective, very large values of capacitance are required. Often, the area required to realize such a capacitance exceeds the area available on a typical die, even if the entire die were covered with the capacitor.

Finally, the supply can be bypassed with a shunt resistor R. This can help by reducing the quality factor of the resonance in the supply circuit. The supply impedance is reduced as R is reduced. However, this shunt resistor dissipates power, and this power dissipation increases as the resistance is reduced. In a low power application, this side effect could be severe.

The effect of adding bypass capacitors to the on chip supply will now be studied further. First, the model in figure 6.20 is rearranged as shown in figure 6.20 to make the analysis more convenient. In this model, the two bond wire inductors are combined. The external voltage source is

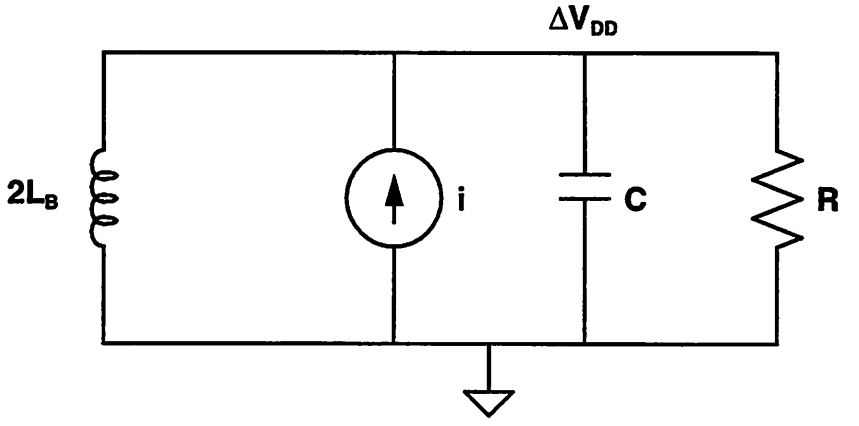


Figure 6.21 Modified Circuit Model for the Supply Circuit of an Integrated Circuit

removed because only AC behavior is of interest here. The s domain transfer function from noise current to V_{DD} is given by the following equation.

$$\frac{\Delta V_{DD}}{I} = \frac{\frac{1}{C}s}{s^2 + \frac{1}{RC}s + \frac{1}{2L_B C}} \tag{6.9}$$

Given this transfer function, the magnitude of the frequency domain response is given by the following formula.

$$\left|\frac{\Delta V_{DD}}{I}\right| = \frac{\frac{\omega}{C}}{\sqrt{\left(\omega_n^2 - \omega^2\right)^2 + (2\zeta\omega_n\omega)^2}} \tag{6.10}$$

The frequencies shown in the above equation are given as follows.

$$\omega_n = \sqrt{\frac{1}{2L_B C}} \tag{6.11}$$

$$\zeta = \frac{1}{R\sqrt{2}}\sqrt{\frac{L_B}{C}} \tag{6.12}$$

Assuming the inductance L_B is 10nH and R=100 ohms, the magnitude of the impedance given by the above equation is plotted in figure 6.22 for several capacitor values. Note that as C is

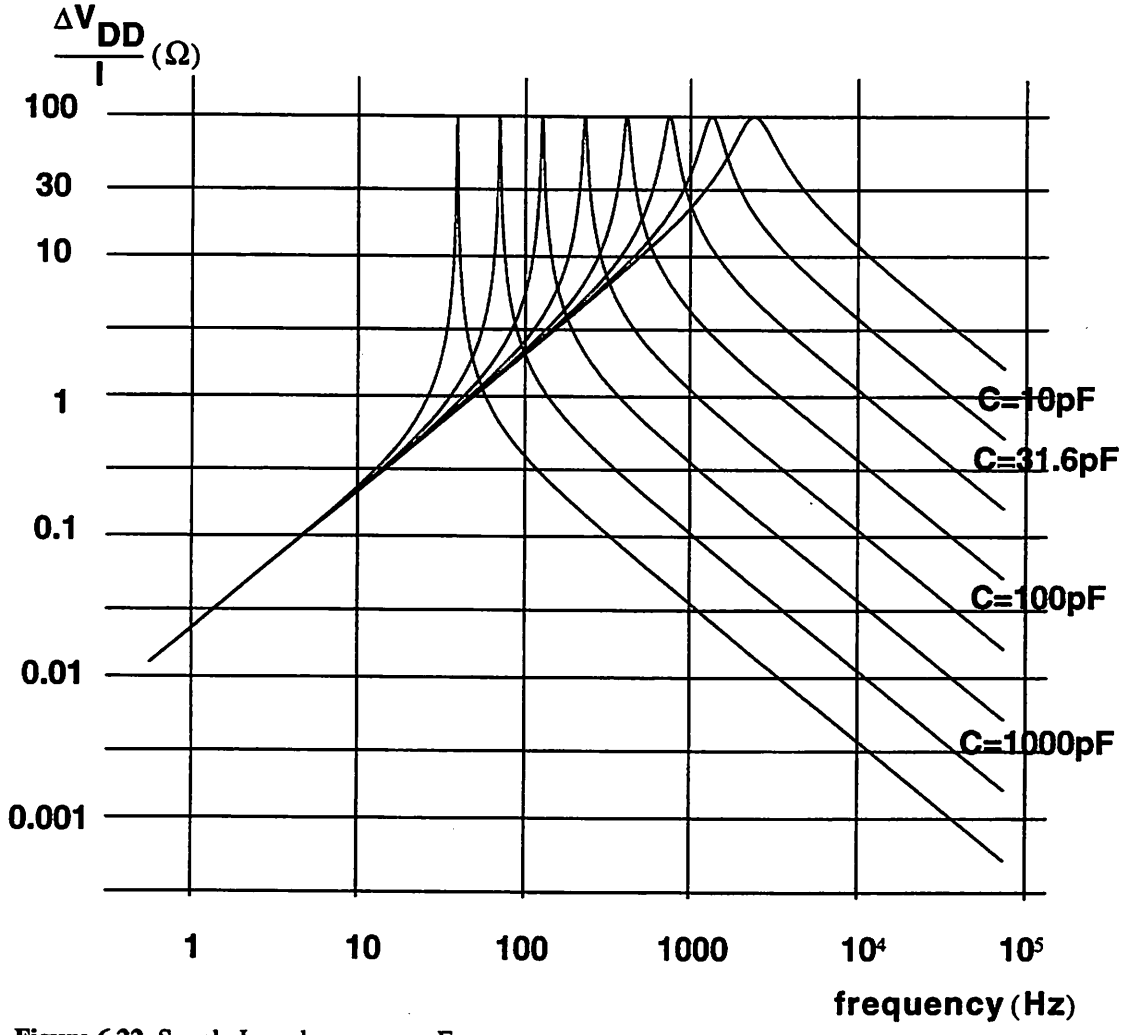


Figure 6.22 Supply Impedance versus Frequency

increased, the impedance is reduced for frequencies above the resonant frequency, but for low frequencies the impedance is not affected. Also note that for reasonable capacitor values (values of

capacitance that can be laid out on an integrated circuit, generally less than 10nF) the resonant frequency is above 100MHz. Thus, systems with a bandwidth less than 100MHz probably cannot be improved by adding bypass capacitors on chip. Furthermore, the additional capacitance might move the resonance into the bandwidth of interest and make things worse.

It is also interesting to look at the above problem from a time domain point of view. For this approach, the response of the supply voltage to an impulse $Q\delta(t)$ of current is analyzed. Q is the charge delivered by the impulse of current. This response is found by taking the inverse Laplace transform of the response given by equation (6.9). The result is given by the following equation.

$$\Delta V_{DD} = \frac{Qe^{-\zeta\omega_n t}}{C\sqrt{1-\zeta^2}}\cos(\omega_D t + \theta) \tag{6.13}$$

In the above equation, θ is given as follows.

$$\theta = \text{atan}\left(\frac{\zeta}{\sqrt{1-\zeta^2}}\right) \tag{6.14}$$

The results of this analysis are plotted in figure 6.23 as a function of the capacitance C for the case where R=100 ohms and L_B=10nH. Supply noise is often tolerable during part of the cycle if it is not present during the sampling instant. Therefore, the response is plotted at various times t. It can be seen from the graph that for long waiting times adding additional capacitance can make things worse rather than better. Therefore, capacitive supply bypassing on an integrated circuit must be used with caution. Because of the problems discussed here, on chip capacitive bypassing was not added to the prototype circuit.

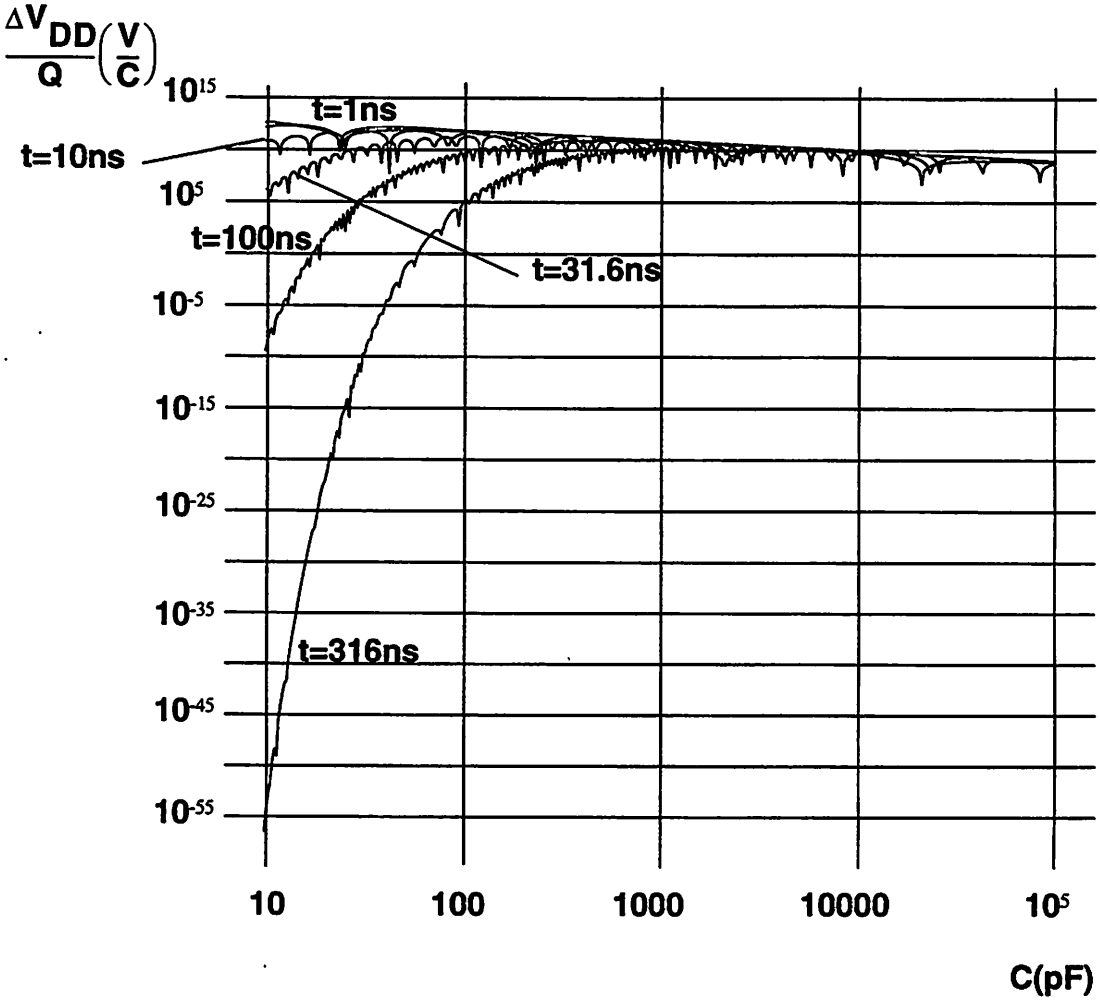


Figure 6.23 Response of Supply Versus Frequency

References

[Ahuja83] Bhupendra K. Ahuja, "An Improved Frequency Compensation Technique for CMOS Operational Amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-18, no. 6, Dec. 1983, pp. 629-633.

[Cho95] Thomas Byunghak Cho and Paul R. Gray, "A 10 b, 20 Msample/s, 35 mW Pipeline A/D Converter," *IEEE J. Solid-State Circuits*, vol. 30, no. 3, March 1995, pp. 166-172.

[Conroy93] Cormac S. G. Conroy, David W. Cline, and Paul R. Gray, "An 8-b 85-MS/s Parallel Pipeline A/D Converter in 1-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 447-454.

[Flynn95] Michael P. Flynn and David J. Allstot, "CMOS Folding ADCs with Current-Mode Interpolation," *ISSCC Dig. Tech. Papers*, Feb. 1995, pp. 274-275, 378.

[Karanicolas93] Andrew N. Karanicolas, Hae-Seung Lee, and Kantilal L. Bacrania, "A 15-b 1-Msample/s Digitally Self-Calibrated Pipeline ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1207-1215.

[Lewis87A] Stephen H. Lewis and Paul R. Gray, "A Pipelined 5-Msample/s 9-bit Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 954-961.

[Lewis87B] Stephen H. Lewis, "Video-Rate Analog-to-Digital Conversion Using Pipelined Architectures," Memorandum No. UCB/ERL M87/90, Electronics Research Laboratory, U. C. Berkeley, November 1987, pp. 96-98.

[Lin90] Yuh-Min Lin, "Performance Limitations on High-Resolution Video-Rate Analog-to-Digital Interfaces," Memorandum No.UCB/ERL M90/55, Electronics Research Laboratory, U. C. Berkeley, June 1990, pp. 76-80.

[Lin91] Yuh-Min Lin, Beomsup Kim, and Paul R. Gray, "A 13-b 2.5-MHz Self-Calibrated Pipelined A/D Converter in 3-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, April 1991, pp. 628-636.

[Mangelsdorf90] Christopher W. Mangelsdorf, "A 400MHz Input Flash Converter with Error Correction," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, Feb. 1990, pp. 184-191.

[McCreary75] James L. McCreary and Paul R. Gray, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques-Part I," *IEEE J. Solid-State Circuits*, vol. SC-10, Dec. 1975, pp. 371-379.

[McCreary81] James L. McCreary, "Matching Properties, and Voltage and Temperature Dependence of MOS Capacitors," *IEEE J. Solid-State Circuits*, vol. SC-16, no. 6, Dec. 1981, pp. 608-616.

[McNutt94] M. J. McNutt, S. LeMarquis, J. L. Dunkley, "Systematic Capacitor

Matching Errors and Corrective Layout Procedures," *IEEE J. Solid-State Circuits*, vol. 29, no. 5, May 1994, pp. 611-616.

[Shyu84] Jyn-Bang Shyu, Gabor C. Temes, and Francois Krummenacher, "Random Error Effects in Matched MOS Capacitors and Current Sources," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 948-955.

[Song95] Won-Chul Song, Hae-Wook Choi, Sung-Ung Kwak, and Bang-Sup Song, "A 10-b 20-Msample/s Low-Power CMOS ADC," *IEEE J. Solid-State Circuits*, vol. 30, no. 5, May 1995, pp. 514-521.

[Yukawa85] Akira Yukawa, "A CMOS 8-Bit High-Speed A/D Converter IC," *IEEE J. Solid-State Circuits*, vol. SC-20, no.3, June 1985, pp. 775-779.

CHAPTER 7

EXPERIMENTAL RESULTS

The prototype analog to digital converter was tested, and the results are summarized in this chapter. The tests performed include the code density test to measure integral and differential nonlinearity, a sine wave fit to measure signal to noise ratio and harmonic distortion, and an idle channel noise test with no input signal present.

7.1 Die Photograph

The die photograph for the prototype ADC is shown in figure 7.1. The pipelined ADC is visible in the right side of the photograph. The input stage of the pipeline is at the bottom. The structure on the left side of the photograph is a shift register memory circuit that was included to allow the digital output data to be saved without driving signals off chip during conversion. Total die area is 6.9mm by 6.9mm. Core die area of the pipelined ADC excluding pads and memory is 4.5mm by 6.2mm.

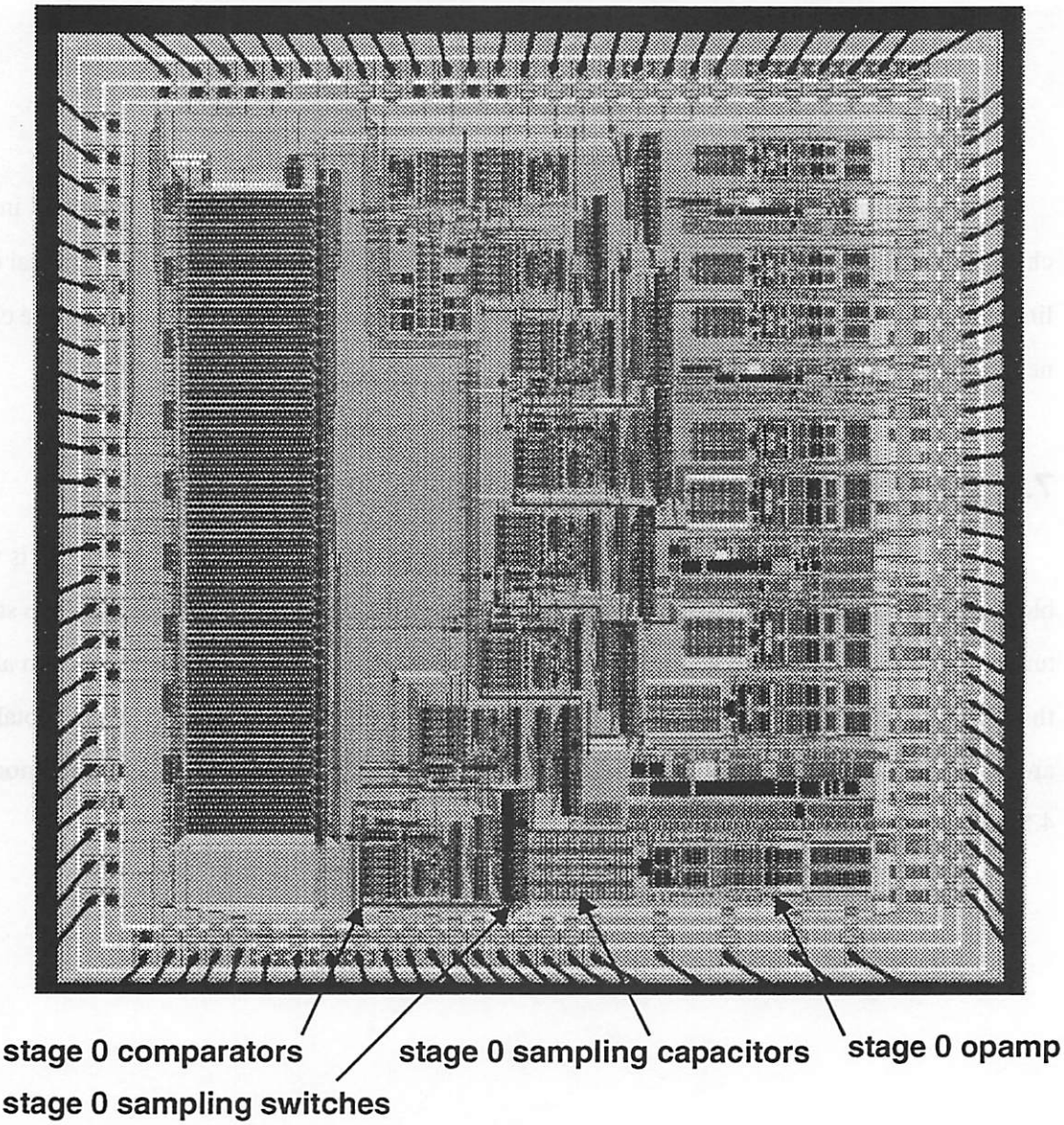


Figure 7.1 Die Photograph of the Prototype Integrated Circuit

7.2 Code Density Test

The transfer characteristic of an ADC is often measured using a code density test with a known waveform at the input[Doernberg84]. In a code density test, a large number of samples of an input signal are collected and converted to digital codes. Then the number of occurrences of each

digital code is plotted on a histogram of frequency of occurrence versus code. If enough samples are taken, the effects of noise are averaged out and all of the information about the transfer characteristic of the ADC can be obtained.

The procedure for using the code density test to determine the transfer characteristic of the ADC is illustrated in the following discussion. This analysis is based on the ADC with the transfer characteristic shown in figure 7.2. The transfer characteristic is defined by the values of the threshold voltages V_{Tk}. If these thresholds are known, the transfer characteristic is completely known, and therefore the DNL and INL are also known.

It is assumed here that a symmetrical input signal is used during the code density test. Symmetrical is taken here to mean that the positive going part of the signal has the same shape as the negative going part of the signal.

The application of the input signal as shown in figure 7.2 results in each output code lasting for a certain period of time as indicated in the bottom left corner of the figure. The frequency of each code is proportional to the length of time each code lasts. Therefore, the code density test results in a histogram output like that shown in figure 7.3. The number of occurrences N_k of the k-th code is given by the following formula.

$$\frac{N_k}{N_{total}} = \frac{t(V_{Tk+1}) - t(V_{Tk})}{t_{total}} \tag{7.1}$$

The value N_{total} in the above equation is the total number of samples taken during the code density test, and it can be expressed by the following formula.

$$N_{total} = \sum_{k=0}^{N_{code}-1} N_k \tag{7.2}$$

N_{code} in the above equation is equal to the number of possible codes output by the ADC. In the example here N_{code} is 9.

When a code density test is run, the values N_k are measured, and therefore these values are known. The values V_{Tk} of the threshold voltages are not known. These thresholds must be determined by rearranging equation (7.1). The result is shown below.

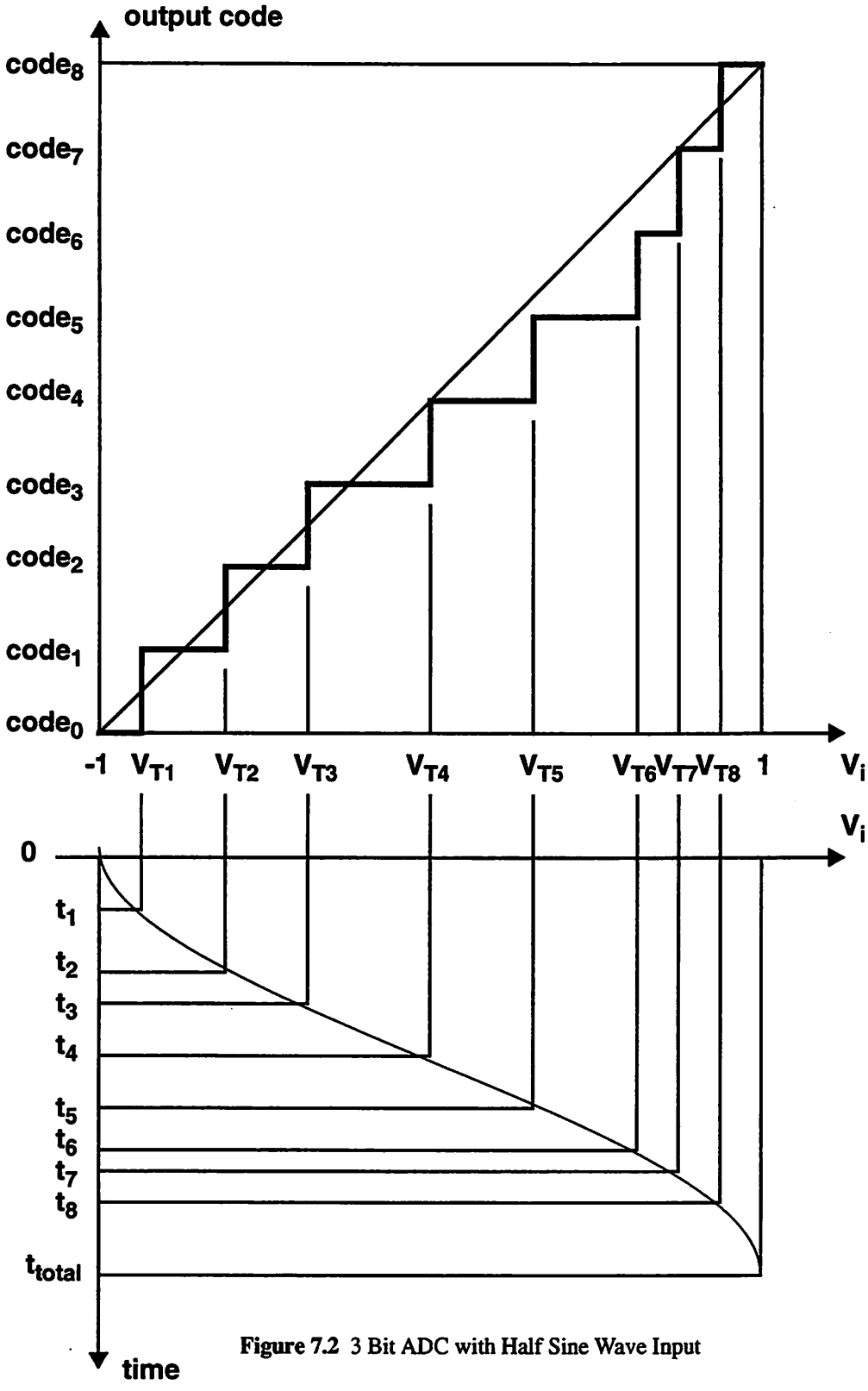


Figure 7.2 3 Bit ADC with Half Sine Wave Input

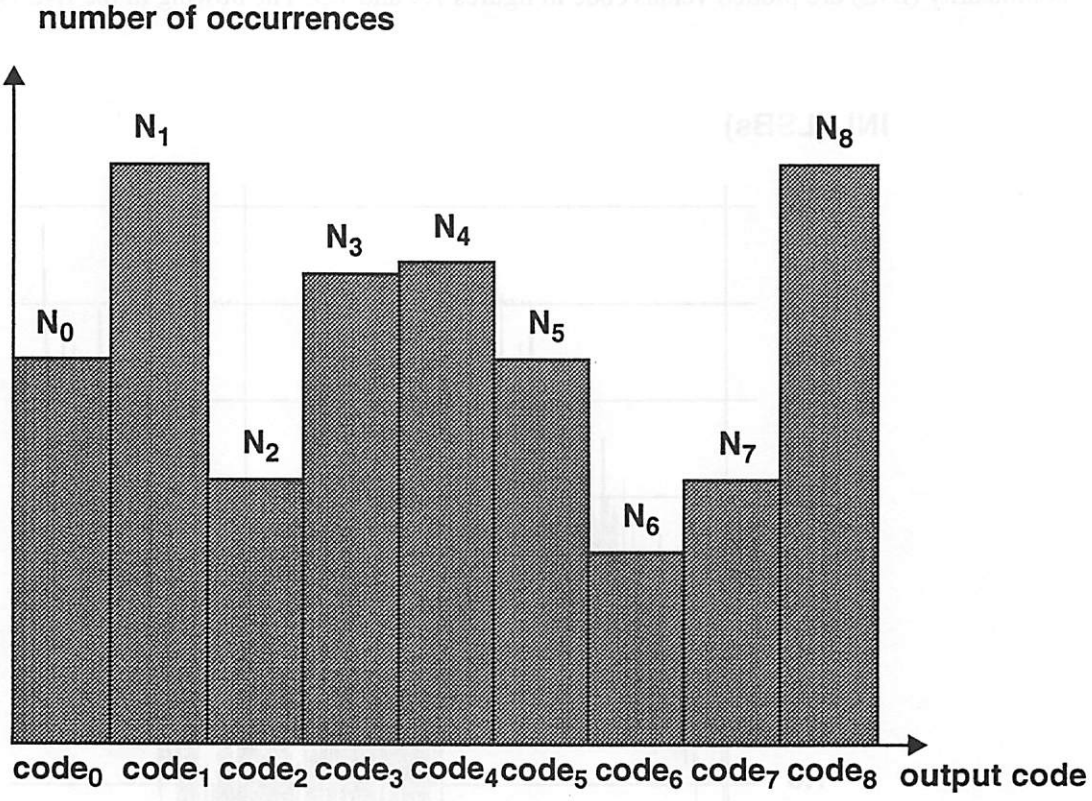


Figure 7.3 Histogram Output for 3 Bit ADC

$$t(V_{Tk+1}) = t(V_{Tk}) + \frac{N_{k+1}}{N_{total}} t_{total} \tag{7.3}$$

The above result holds for a general input waveform. The code density test is often performed with a sine wave input as indicated in figure 7.2. In this case, t_{total} is equal to half the period of the input. The above equation can then be rewritten as shown below.

$$V_{Tk+1} = -\cos\left[\operatorname{acos}(-V_{Tk}) + \pi\frac{N_{k+1}}{N_{total}}\right] \tag{7.4}$$

$$V_{Tk} = -\cos\left[\frac{\pi}{N_{total}}\left(\sum_{m=0}^{k} N_m\right)\right] \tag{7.5}$$

The code density test results presented here were based on 4194304 samples of a 100kHz sine wave input signal sampled at 5MHz. The resulting differential nonlinearity (DNL) and integral nonlinearity (INL) are plotted versus code in figures 7.4 and 7.5. The bowing in the INL is caused

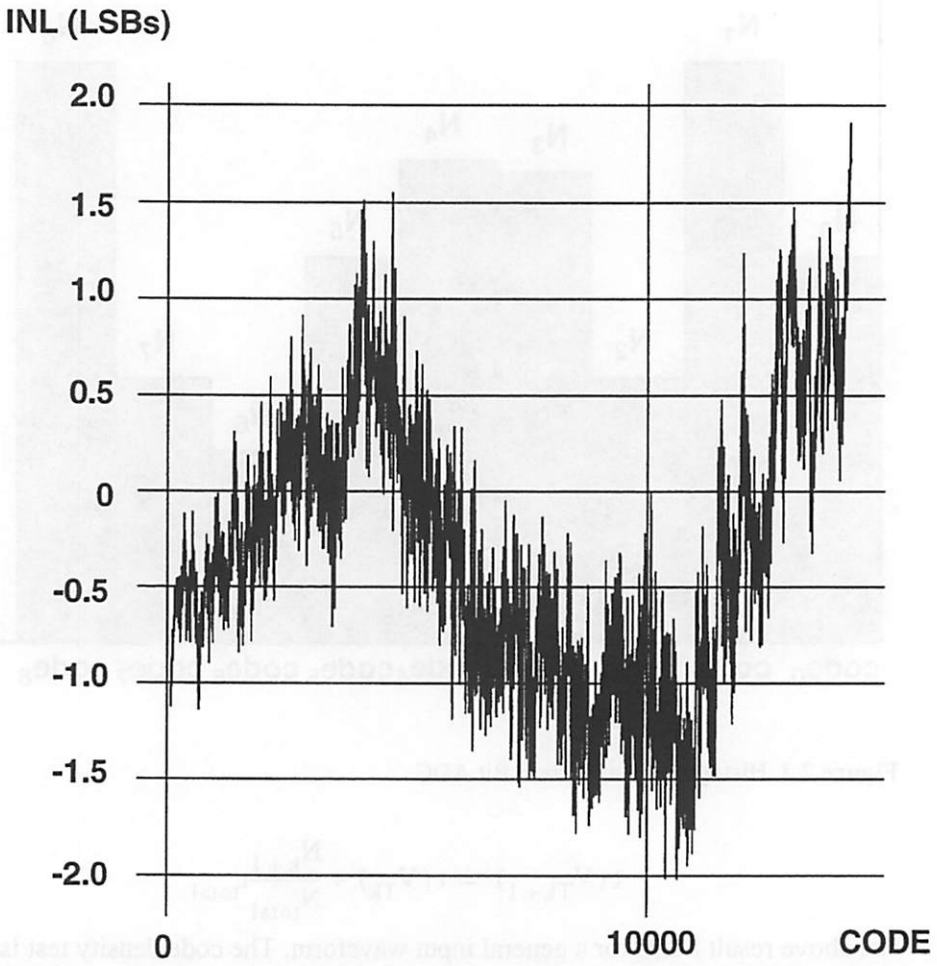


Figure 7.4 Integral Nonlinearity Versus Code in LSBs at 14 Bits, sampling rate=5MHz, input frequency=100kHz

by signal dependent charge injection by the center shorting switch of the sample and hold circuit[Shieh87][Lin90].

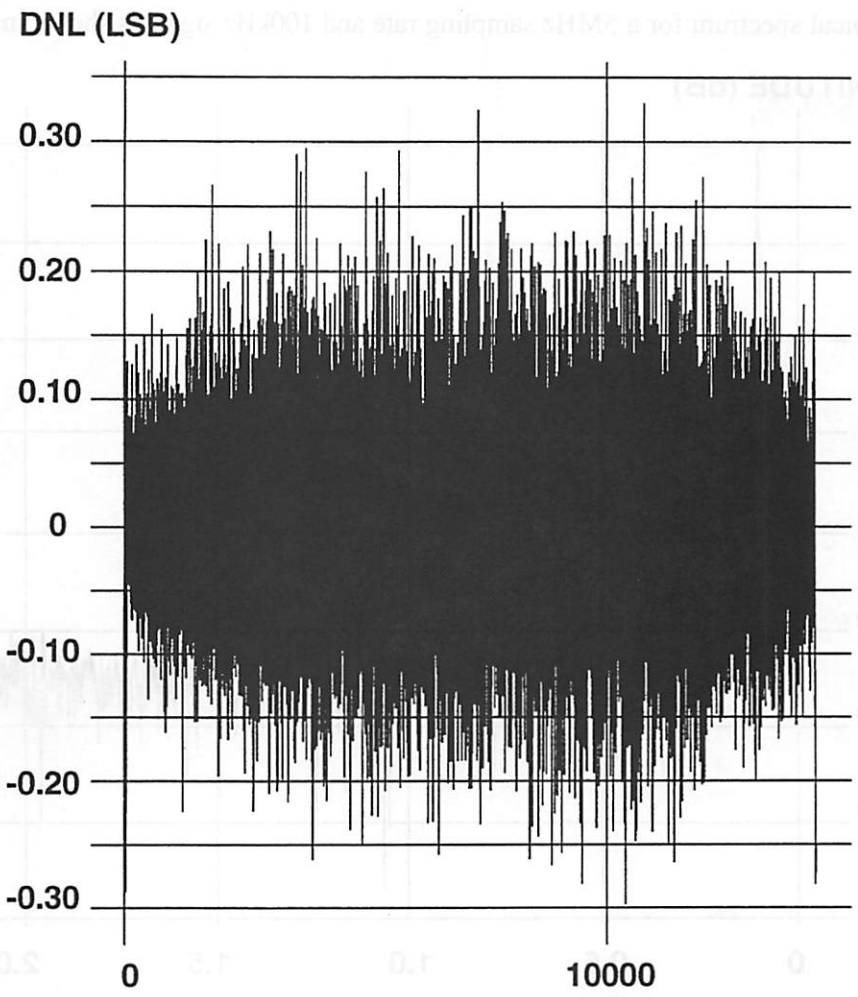


Figure 7.5 Differential Nonlinearity versus Code in LSBs at 14 Bits

7.3 Signal to Noise Ratio and Distortion

The signal to noise ratio was calculated by collecting 2048 samples of the input signal and performing a 2048 point fast Fourier transform. This test was performed at a sampling frequency of 5MHz and input frequency of 100kHz and 2MHz.

7.3.1 Results for Low Input Frequency (100kHz)

A typical spectrum for a 5MHz sampling rate and 100kHz signal is shown in figure 7.6. This

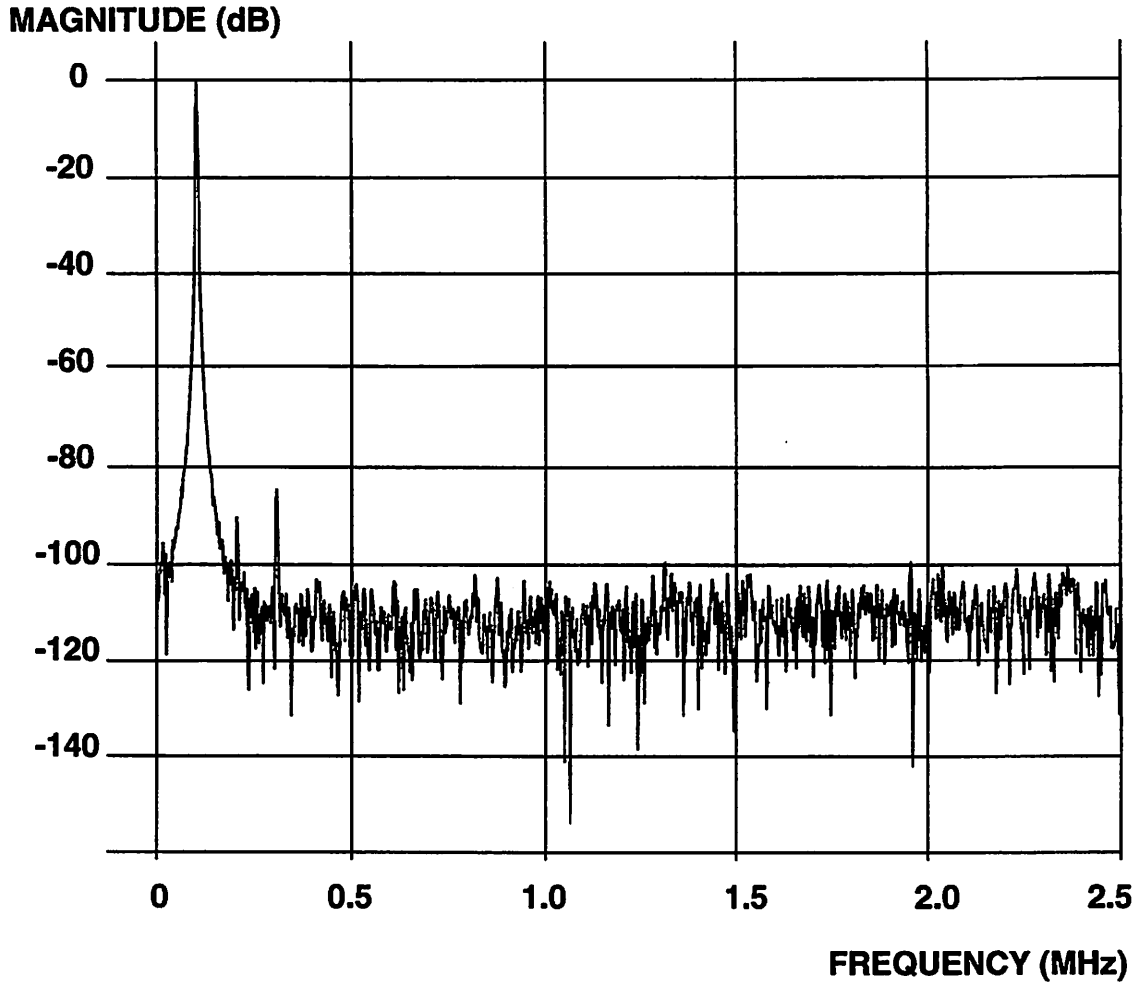


Figure 7.6 Typical 2048 Point FFT of a 100kHz Input Sinusoid at a 5 MHz Sampling Rate

input spectrum was plotted for an input signal that was -4.22dB below full scale. A third harmonic about 85dB below the fundamental is clearly observable in the spectrum. This distortion is caused by signal dependent charge injection from the center shorting switch. The charge injection is signal dependent because the impedance of the sampling switches varies with the signal.[Lin90]

7.3.2 Results for High Input Frequency (2MHz)

A typical spectrum for a 5MHz sampling rate and 2MHz input signal is shown in figure 7.7.

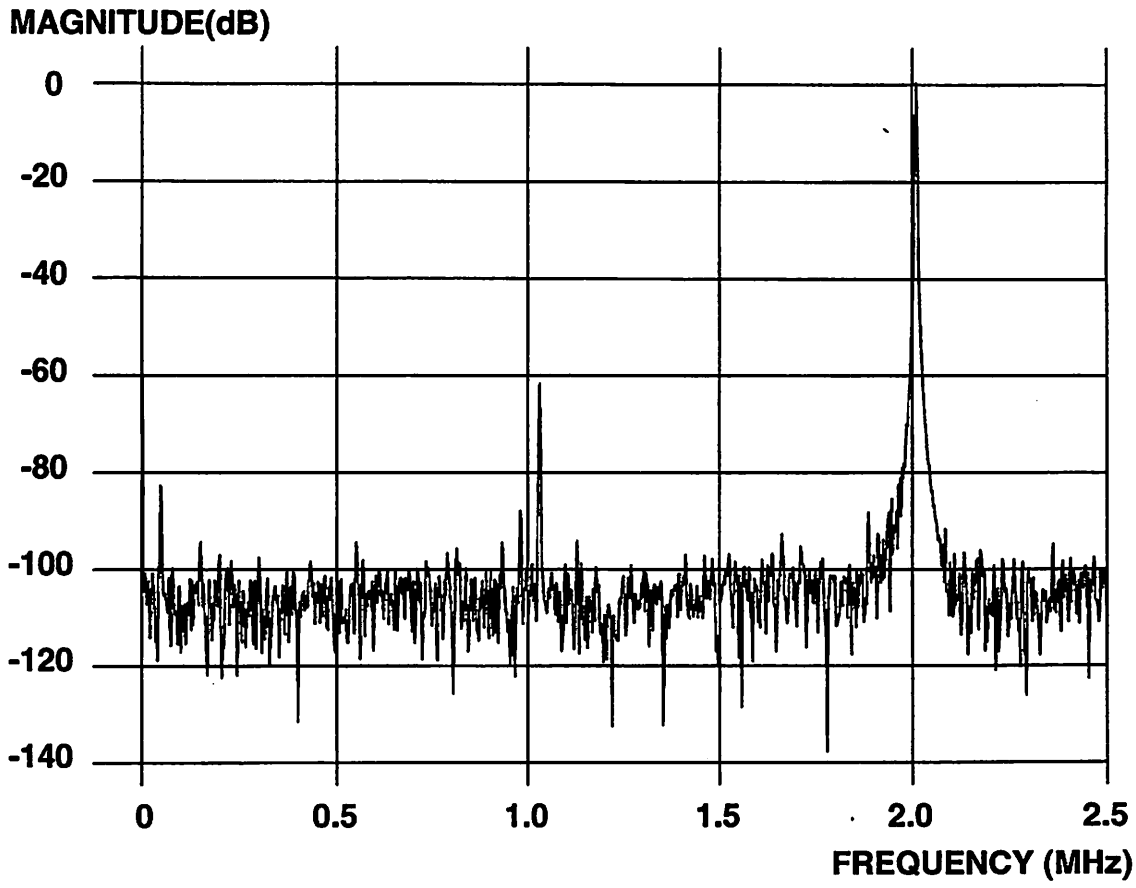


Figure 7.7 Typical 2048 Point FFT of a 2MHz Input Sinusoid at a 5MHz Sampling Rate

The spike at 1 MHz is a 6MHz third harmonic signal sampled at 5 MHz and aliased down to the baseband. This third harmonic is caused by the nonlinear resistance of the sampling switches in the first stage sample and hold circuit. This effect is modeled in chapter 3. Basically, the sample and hold network is a resistor in series with a capacitor that acts like a voltage divider. At low frequencies, the capacitor impedance looks infinite, and the nonlinearity of the resistance does not matter.

As the frequency is increased, the impedance of the capacitor drops. As this happens, the transfer function of the voltage divider becomes voltage dependent, and this results in distortion.

7.3.3 Signal to Noise Ratio Versus Input Amplitude

Figure 7.8 shows a plot of the signal to noise ratio (SNR) as a function of input amplitude for a

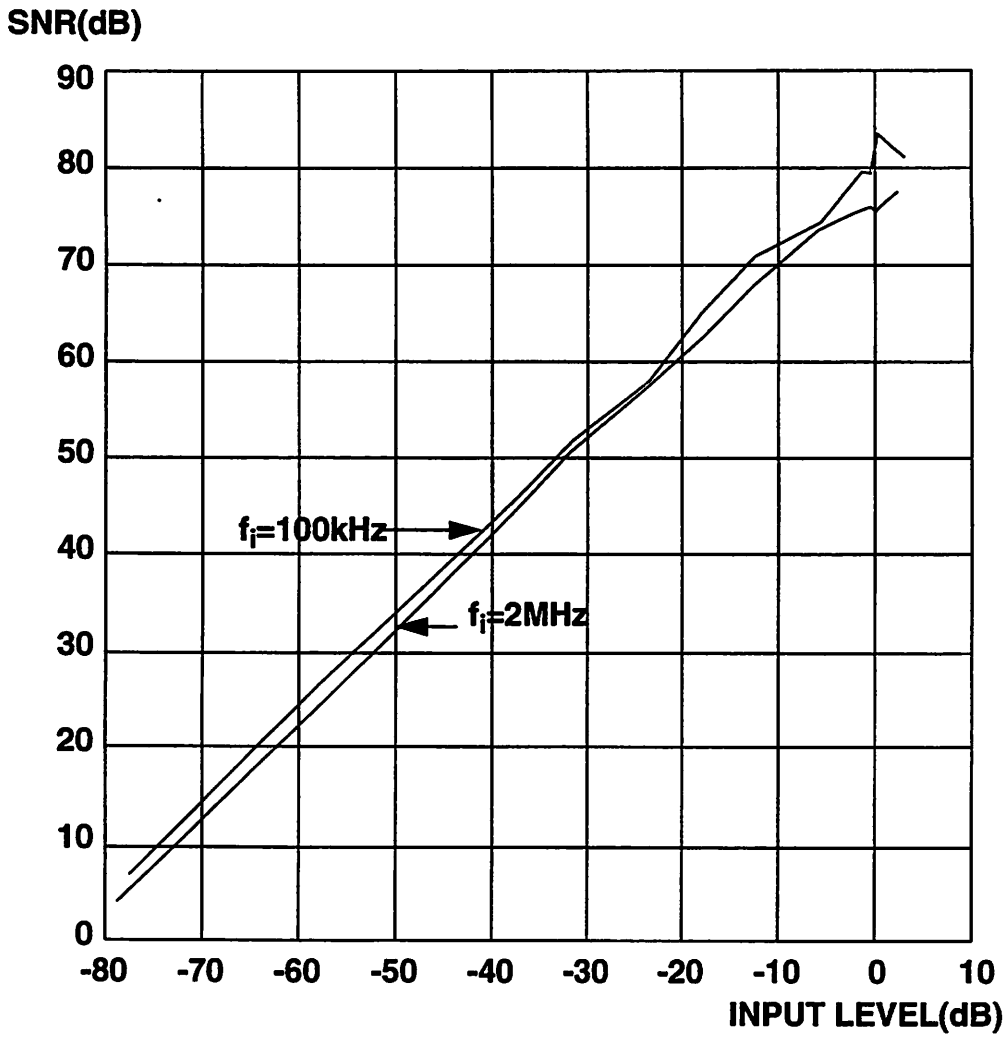


Figure 7.8 SNR versus Input Level at a Sampling Rate of 5MHz

100kHz input signal and a 2MHz input signal. The sampling rate is 5MHz in both cases. For both cases, the signal to noise ratio increases rather linearly with input amplitude.

Figure 7.9 shows a plot of the signal to noise+distortion ratio (SNDR) as a function of input amplitude. Note that the SNDR reaches a peak value at an input amplitude below full scale and

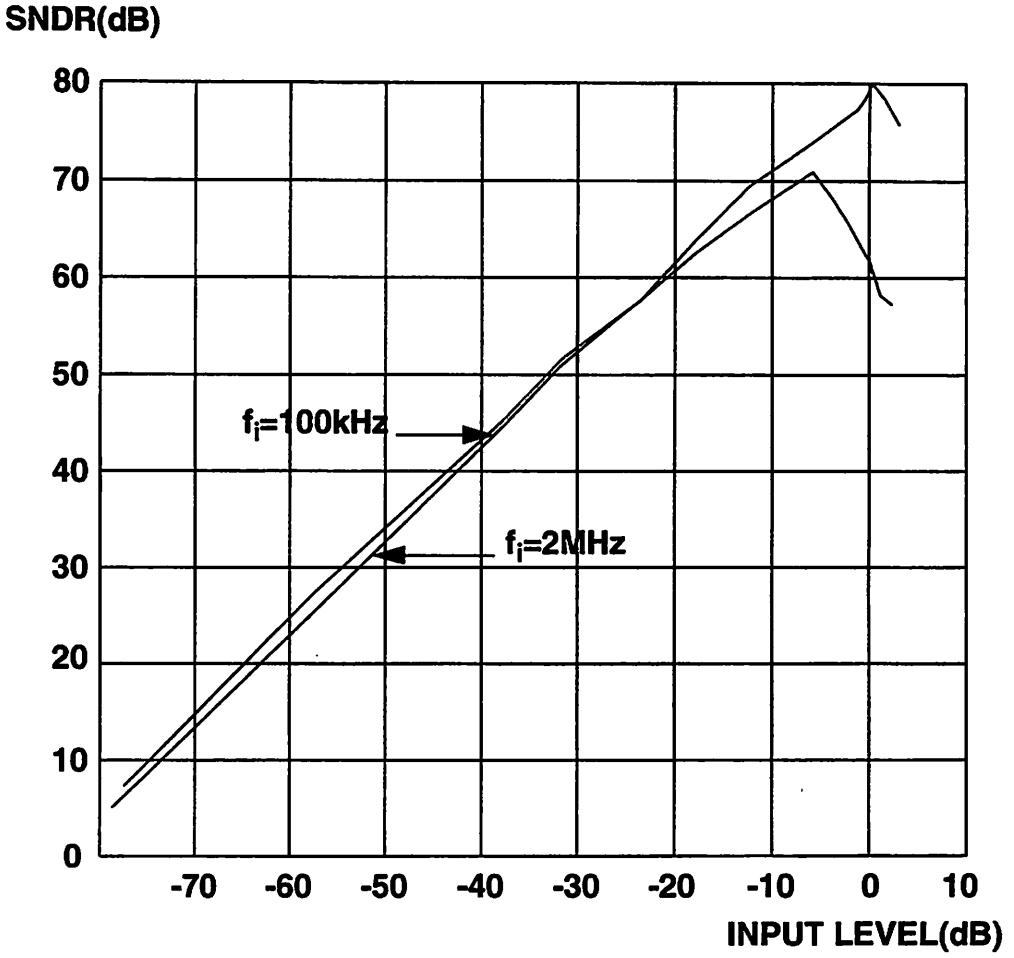


Figure 7.9 Signal to Noise+Distortion (SNDR) versus Input Level at a Sampling Rate of 5MHz

then rolls off. The peak amplitude is lower for a high frequency input signal than for the low frequency input. This occurs because the distortion produced by the nonlinearity in the sampling switches is proportional to the frequency.

7.3.4 Idle Channel Noise

The standard deviation of the output was measured with no input signal applied. The RMS value of the variation is 0.21LSB at 14 bits of resolution. The maximum signal amplitude is 6.6V peak to peak. Therefore, this idle channel noise corresponds to 84.6μV RMS. This noise indicates that the maximum signal to noise ratio for this circuit should be 88.8dB barring other factors.

7.4 Summary of Results

The following table summarizes the experimental results. The small signal input bandwidth was measured by applying a 20mV peak to peak sinusoidal input and measuring the frequency at which the SNDR was down by 3 dB relative to the maximum SNDR for an input signal of that amplitude.

Table 7.1 Summary of Results

parameter	value
resolution	14 bits
sampling rate	5MHz
INL	+/-2.0LSB
DNL	+/-0.35LSB
peak SNR for a 100kHz input	83dB
peak SNDR for a 100kHz input	80dB
peak SNR for a 2MHz input	77dB
peak SNDR for a 2MHz input	71dB
small signal input bandwidth	50 MHz
process technology	1.2μm double poly double metal CMOS
total die area	6.9mm x 6.9mm
core area (excluding pads and memory)	4.5mm x 6.2mm
full scale differential signal swing	6.6V peak to peak
supply voltage	0,5V

Table 7.1 Summary of Results

parameter	value
power dissipation	166mW

References

[Doernberg84] Joey Doernberg, Hae-Seung Lee , and David A. Hodges, "Full-Speed Testing of A/D Converters," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 820-827.

[Lin90] Yuh-Min Lin, "Performance Limitations on High-Resolution Video-Rate Analog-to-Digital Interfaces," Memorandum No.UCB/ERL M90/55, Electronics Research Laboratory, U. C. Berkeley, June 1990, pp. 8-40.

[Shieh 87] Je-Hurn Shieh, Mahesh Patil, Bing J. Sheu, "Measurement and Analysis of Charge Injection in MOS Analog Switches," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 2, April 1987, pp. 277-281.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

This research was concerned primarily with two topics: optimizing switched capacitor sample and hold circuits in pipelined ADCs to minimize power dissipation, and comparing the power dissipation and speed performance of operational amplifiers used in switched capacitor applications

8.1.1 Optimization of Pipelined ADCs

The following conclusions about the optimization of pipelined analog to digital converters were obtained. The discussion divides the pipelines into two classes; low resolution pipelines and high resolution pipelines. For the purpose of this discussion, a high resolution pipeline is a pipeline where the sampling capacitance required to meet the thermal noise requirement is significantly higher than the parasitic capacitances.

- Raising the supply voltage saves power in a noise limited pipelined ADC.

The supply voltage tends to determine the maximum signal amplitude that can be accommodated at an acceptable level of distortion. This maximum signal amplitude sets the noise constraint for a given resolution. A factor of 2 increase in supply voltage results in at least a factor of 2 increase in allowable signal amplitude. For every factor of 2 increase in allowable signal amplitude, the noise voltage can also be increased by a factor of 2. This allows a factor of 4 reduction in the size of the sampling capacitor. For a fixed speed this tends to allow a factor of 4 reduction in bias current. As a result, the power dissipation tends to be reduced by a factor of 2 in a class A opamp.

If the supply voltage is not constrained by something else, the optimum supply voltage is determined by a trade-off between the high resolution analog circuits, which favor a high supply voltage, and the low resolution analog circuits and digital circuits, which favor a lower supply voltage.

• Capacitor tapering saves power in a noise limited pipelined ADC.

In high resolution pipelines a significant reduction in power dissipation can be achieved by tapering the capacitor sizes in later stages of the pipeline relative to the front end. The optimum rate of tapering is dependent on the resolution requirements of the ADC and the available supply voltage. High resolution pipelines and pipelines operating on a low supply voltage are both likely to require front end sampling capacitors that are significantly larger than parasitic capacitances in order to meet noise constraints. As a result, these pipelines tend to favor rather rapid tapering of capacitors as the noise requirements trail off in the later stages of the pipeline. A good rule of thumb is to start with a taper factor equal to the interstage gain. Pipelines with a moderate resolution tend to favor a smaller taper factor.

• Some capacitor tapering saves power in a low resolution pipeline.

Power dissipation can also potentially be reduced by scaling down the tail end capacitors of a low resolution pipeline where thermal noise is not an issue. In this case, a relatively large front end sampling capacitor is beneficial because it can improve the feedback around the opamp assuming the required interstage gain is fixed. Increasing the front end sampling capacitance also increases the load capacitance on the front end opamp, but the effect tends to be relatively small since only a fraction of the sampling capacitance loads the output of the opamp. As a result, the power saved by improving the feedback can more than compensate for the extra load capacitance. The optimum rate of tapering for a low resolution pipeline tends to be lower than it is for a high resolution pipeline.

• High per stage resolution saves power in a noise limited pipeline.

Increasing the per stage resolution in a pipelined ADC increases the interstage gain. As a result, the feedback is degraded, and this tends to increase the power dissipation. However, the optimum capacitor taper factor is also increased. Because increasing the per stage resolution allows the capacitors to be scaled down more rapidly, the load capacitance to each opamp tends to be reduced. The power savings caused by this reduced load capacitance can more than compensate for the power penalty caused by lower feedback factor.

• High per stage resolution increases power dissipation in a low resolution pipeline.

As mentioned in the previous paragraph, increasing the per stage resolution leads to a degradation of the feedback around the opamp. Furthermore, increasing the per stage resolution increases the number of comparators required in each stage. This tends to increase the load capacitance of each opamp. In a low resolution pipeline, the load capacitance contributed by the comparators is significant. Because of these two effects, reduced feedback factor and increased load capacitance, increasing the per stage resolution results in higher power dissipation. Low resolution pipelines are optimized by using a per stage resolution of 1 or less.

• High resolution pipelines tend to favor the use of multistage opamps.

In addition to affecting optimum capacitor sizes and opamp architecture, noise considerations also influence the choice of opamp architecture. Multistage designs may be necessary in order to achieve adequate gain and output swing, especially in cases where the supply voltage is less than 5V. However, as mentioned later in this chapter, in some cases a single stage architecture might still be practical in a high resolution pipeline.

• Low resolution pipelines tend to favor the use of single stage opamps.

Because of the simplicity, smaller quantity of hardware, and reduced number of poles in a single stage opamp, this type of opamp tends to dissipate less power than a multistage opamp to settle within a fixed period of time. In a low resolution pipeline, the linearity requirements are relaxed enough to make a single stage opamp a practical choice.

8.1.2 Opamp Architectures for Switched Capacitor Applications

Many factors generally influence opamp design. This research attempted to compare a few architectures strictly on the basis of speed and power dissipation. It was no surprise that the single stage single pole opamp was the best for achieving high speed in situations where the required closed loop gain was low. However, at lower speeds that only modestly push the limits of technology it was found that a preamplifier driving a single stage amplifier could meet the speed specification with less power. For higher closed loop gains, the single stage opamp architecture was found to have poor speed performance in comparison to multistage architectures. However, the tele-

scopic cascode architecture, which is a simple modification to the single stage opamp was found to perform well at high speeds, even when the closed loop gain was very high.

8.2 Pipelined ADC with Low Swing Amplifiers and Extra Comparators

In this project the prototype ADC was built using two stage amplifiers in order to simultaneously achieve high gain and high swing. Using this two stage amplifier architecture is a big penalty in terms of speed and power dissipation. It would be desirable to use a telescopic cascode amplifier to improve the speed and save power. However, the limited open loop gain from this architecture would lead to linearity problems as discussed in chapter 3. Furthermore, the telescopic cascode amplifier would have less output swing than the two stage amplifier, and this would lead to a tightening of the noise constraints. Therefore, larger sampling capacitors would be required.

The key question is then, "How can a low power high resolution pipelined ADC be built using telescopic cascode amplifiers?" It turns out that both the linearity problem and the problem of limited swing can be tolerated if extra comparators are added to the first stage of the pipeline. By adding extra comparators without increasing the interstage gain, the opamp output swing can be reduced without reducing the input signal swing. Because the input swing has not been reduced, the noise constraints do not need to be tightened. Furthermore, the reduction in output swing allows the open loop gain of the opamp to be improved by using cascode transistors in the output stage of the opamp. The reduction in output swing also relaxes the linearity requirement, and thus the open loop gain requirement, of the opamp. As a result, it may be possible to achieve adequate open loop gain with a single stage telescopic cascode amplifier.

The penalty paid for these simplified opamps is the extra comparators added to the front end stage of the pipeline. However, if these comparators are simple dynamic latches, the power added by these comparators should be very small. Unfortunately, adding extra comparators to the front end of the pipeline tightens the maximum tolerable offset of these comparators. As a result, this technique is limited by the extent to which the comparator offset voltage can be reduced without increasing the power dissipation of the comparator. Some techniques for alleviating the problem of comparator offsets have been developed, and these techniques might be useful in the design of a

pipeline with extra comparators in the front end[Kusumoto93].

To provide some quantitative insight to relate the peak input voltage, peak output voltage, and maximum comparator offset voltage, a graph of the amplified output residue of the first pipeline stage versus the input voltage is plotted in figure 8.1. From this figure, the ratio of the peak output

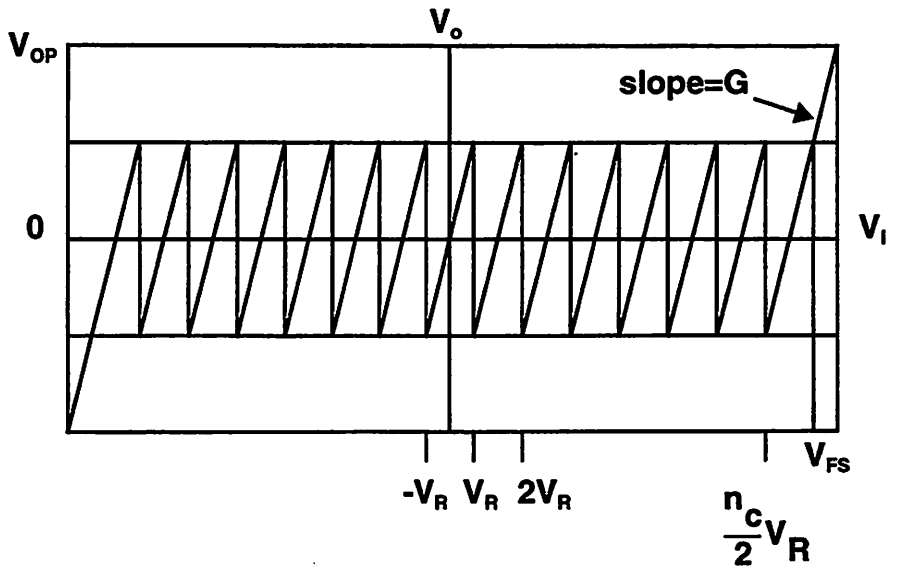


Figure 8.1 Amplified Output Residue Versus Input Voltage for a Pipelined Stage with Extra Comparators

voltage V_P to the full scale input voltage V_{FS} can be related to the gain G and the number of comparators n_C. The result is given by the following equation.

$$\frac{V_P}{V_{FS}} = \left(\frac{G}{n_C + 1}\right) \tag{8.1}$$

The maximum comparator offset can then be calculated using the above graph and the above equation. The result shown below relates the maximum comparator offset voltage V_{Coff} to the maximum available output swing V_{OP} from the amplifier, the closed loop gain G, and the number of comparators n_C. The procedure used here to calculate the maximum comparator offset is similar to the technique used in chapter 3.

$$V_{Coff} = V_{CT} - V_{CTN} = \frac{V_{OP}}{G} - \frac{V_{FS}}{(n_C + 1)} \tag{8.2}$$

To illustrate one way a pipeline with extra comparators might be implemented, the following example is given.

Table 8.1 Implementation Example of a Pipeline with Extra Comparators

Resolution	14 bits
Supply Voltage V_{DD}	5V
Full Scale Input Swing V_{FS}	4.6875V peak
Opamp Architecture	Telescopic Triple Cascode
Maximum Opamp Output Swing V_{OP}	1.5V peak
Nominal Maximum Opamp Output	+/-0.75V
Closed Loop Gain G	4
Number of Comparators n_C in First Stage of Pipeline	24
Maximum Comparator Offset Voltage V_{Coff}	312.5 mV
Number of Comparators per Stage After First Stage	6
Number of Pipeline Stages	7

References

[Kusumoto93] Keiichi Kusumoto, Akira Matsuzawa, and Kenji Murata, "A 10-b 20-MHz 30-mW Pipelined Interpolating CMOS ADC," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, Dec. 1993, pp. 1200-1206.