1

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

THÈSES CANADIENNES

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

CHARACTERIZATION AND MODELING OF MISMATCH
IN MOS DEVICES AND APPLICATION TO
PRECISION ANALOG DESIGN

by

K. R. Lakshmi Kumar, M.E.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfilment of the requirement
for the degree of
Doctor of Philosophy

Faculty of Engineering
Department of Electronics
Carleton University

September 1985

© copyright

1985, K.R. Lakshmi Kumar

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-29870-7

The undersigned hereby recommends to the Faculty of
Graduate Studies and Research, acceptance of the thesis,

"CHARACTERIZATION AND MODELING OF MISMATCH IN MOS
DEVICES AND APPLICATION TO PRECISION ANALOG DESIGN"

submitted by K.R. Lakshmi Kumar, in partial fulfilment of the
requirements for the degree of Doctor of Philosophy.

MA Copeland

Thesis Supervisor

R Hadaway

Thesis Supervisor

J. B. Hagelboom

External Examiner

A. Boonin

Chairman,
Department of Electronics

Carleton University

October 1985

ABSTRACT

This thesis is concerned with the design of precision MOS analog circuits. The work addresses two major aspects of the design, viz, the characterization and modeling of the mismatch in MOS devices, particularly transistors and resistors, and application of this knowledge to circuit design. Statistical methods are used to develop models for the random mismatch in MOS transistors and resistors. The parameters that a circuit designer will have freedom to choose are the dimensions of the devices. Therefore, analytical models have been developed that relate the mismatch to the device dimensions. Extensive measurements on specially designed test structures have been carried out to validate the theoretical analysis.

The second part of the thesis demonstrates the application of the matching characterization studies for precision analog design. A video bandwidth 8-bit Digital-to-Analog Converter (DAC) is taken as a specific example. In order that this test vehicle will also be a useful library cell, considerations like glitch and thermal drift were taken into account in the design. Salient features of the DAC are 80 ns conversion time, 500 pV-s glitch energy, and 23 ppm/°C worst case drift. These values are comparable to those achieved in hybrid DACs. Finally, the circuit yield of the DAC was made to exceed 97% as a result of the knowledge generated in this thesis regarding the matching characteristics of the devices.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Professor Miles Copeland, Department of Electronics, Carleton University, and Mr. Robert Hadaway, Northern Telecom Electronics Ltd., whose ideas led to this research effort. I am also indebted to them for their continued guidance and encouragement throughout the course of this work.

I would like to pay special thanks to Dr. Martine Simard-Normandin for her help with device characterization. Thanks are also due to Mr. Mike King for many suggestions and discussions. I place on record my deep sense of appreciation to Mr. Narasimha Prasad for his assistance concerning statistical analyses.

Northern Telecom Electronics Ltd., Ottawa, Canada provided design, fabrication and test facilities, and also offered financial assistance, for which I am thankful.

Financial support from the Canadian Commonwealth Scholarship and Fellowship Committee, and Carleton University is gratefully acknowledged. My thanks are also due to Bharat Electronics Ltd., Bangalore, India for granting me study leave to pursue this work.

Finally, I wish to thank my wife Vathsala and daughter Pavala for their co-operation.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF FIGURES AND TABLES	viii
LIST OF SYMBOLS AND ABBREVIATIONS	x
 CHAPTER 1 -- INTRODUCTION	 1
1.1 Historical Overview	2
1.2 Thesis Objectives	3
1.3 Thesis Format	5
 CHAPTER 2 -- TRANSISTOR MATCHING STUDIES	 8
2.1 Design of Test Structures	9
2.2 Measurement Setup	10
2.3 Measurement Philosophy	13
2.3.1 Choice of the Model	15
2.3.2 Description of the Model	16
2.4 Factors Causing Mismatch	21
2.4.1 Mismatch in V_T	22
2.4.2 Mismatch in K	37
2.4.3 Correlation Between V_T and K	48
2.4.4 Mismatch in Saturation Region	56
2.5 Range of Applicability	57
2.6 Tracking With Temperature	62
2.7 Summary	63
 CHAPTER 3 -- RESISTOR MATCHING STUDIES	 64
3.1 Resistor Types	64
3.2 Mismatch in Resistance	66
3.3 Summary	72

CHAPTER 4 -- HIGH SPEED D/A CONVERTER	74
4.1 Current Steering Configurations	74
4.2 Multiple Current Source Approach	79
4.2.1 Interstage Divider Multiple Current Source DAC	82
4.3 Error Analysis	85
4.3.1 General Considerations	86
4.3.2 Mathematical Analysis	87
4.3.3 Accuracy of Divide By 16 Network	94
4.4 Circuit Blocks	95
4.4.1 Unit Current Source	95
4.4.2 Bit Cell	102
4.5 Video Bandwidth DAC	108
4.6 Summary	115
CHAPTER 5 -- CONTROL AMPLIFIER	116
5.1 Causes of Offset and Drift	116
5.2 Autozeroing Technique	123
5.3 Chopper Amplifier	123
5.4 Chopper-Stabilized Amplifier	125
5.5 Basic Operational Amplifier Block	136
5.5.1 Open Loop Gain	140
5.5.2 Bandwidth	145
5.5.3 Noise	148
5.5.4 Biasing	155
5.6 Null Inputs to the Amplifier	157
5.6.1 Configuration 1	157
5.6.2 Configuration 2	157
5.6.3 Configuration 3	158
5.6.4 Configuration 4	158
5.7 The Complete Amplifier	158
5.8 Summary	161
CHAPTER 6 -- RESULTS	163
6.1 Chopper Stabilized Amplifier	163

6.2	8-Bit DAC	165
6.3	Summary	177
CHAPTER 7 -- CONCLUSIONS AND FUTURE RESEARCH		178
7.1	Conclusions	178
7.2	Future Research	179
REFERENCES		183
APPENDIX A BASIC STATISTICAL RELATIONSHIPS		187
APPENDIX B DERIVATION OF EQUATION (2.19)		192

LIST OF FIGURES AND TABLES

FIGURES

Fig.2.1	Matched transistor lay out	11
Fig.2.2	Automated measurement set up	12
Fig.2.3	Current-voltage relationship of a MOSFET in triode region	19
Fig.2.4	Gate oxide thickness variation	26
Fig.2.5	Threshold voltage mismatch versus dimension	35
Fig.2.6(a)	Conductance constant mismatch versus dimension for N-channel transistors	45
Fig.2.6(b)	Conductance constant mismatch versus dimension for P-channel transistors	47
Fig.2.7(a)	Drain current mismatch versus dimension for N-channel transistors	58
Fig.2.7(b)	Drain current mismatch versus dimension for P-channel transistors	59
Fig.3.1	Cross section of capacitor implant resistor	67
Fig.3.2	Monolithic resistor	68
Fig.3.3	Capacitor implant resistor mismatch versus dimension	73
Fig.4.1	Drain current versus gate width of narrow channel transistors	77
Fig.4.2	Channel narrowing effects	78
Fig.4.3	Simple multiple current source approach	80
Fig.4.4	Interstage divider multiple current source DAC	84
Fig.4.5	Divide by 16 network	84
Fig.4.6	DAC yield versus current source matching	93
Fig.4.7	Unit current source	96
Fig.4.8	Current steering bit cell	107
Fig.4.9	Switching waveform	107
Fig.4.10	Bit cell driver	107
Fig.4.11	Schematic diagram of the video DAC	109
Fig.4.12	Equivalent circuit of the DAC	112
Fig.5.1	Two-stage amplifier illustrating interstage coupling constraints	118
Fig.5.2(a)	A simple MOS differential stage	120
Fig.5.2(b)	Equivalent circuit	120
Fig.5.3	Principle of autozeroing	124
	(a) Autozero mode; (b) Amplification mode	
Fig.5.4	Chopper amplifier	126
Fig.5.5	Chopper stabilized amplifier	126
Fig.5.6	Modified chopper stabilized amplifier	128

Fig.5.7	(a) Null amplifier block diagram for clock phase 1	131
	(b) System block diagram during clock phase 2	131
Fig.5.8	Frequency response of the amplifiers	
	(a) Main amplifier; (b) Null amplifier with R2C2 filter network	134
Fig.5.8(c)	Frequency response of the overall amplifier	135
Fig.5.9	Basic operational amplifier block	139
Fig.5.10	Equivalent circuit of the amplifier	141
Fig.5.11	Equivalent circuits for calculating the output impedances of the (a) input and N-channel cascode stage; (b) differential to single ended converter	144
Fig.5.12	Equivalent circuit to determine the output impedance (a) at the source node of M5; (b) at the source node of M9	147
Fig.5.13	Individual noise sources in the op amp	150
Fig.5.14	(a) Noise equivalent of the cascode stage (b) & (c) equivalent circuits when the two noise sources are independently considered	151
Fig.5.15	Simplified circuit for calculating the equivalent input noise of the amplifier	153
Fig.5.16	(a) Cascode stage biasing arrangement (b) I-V characteristics	156
Fig.5.17	Various configurations for providing offset null inputs	159
Fig.5.18	Schematic of the complete amplifier block	160
Fig.5.19	Reference voltage for the non-inverting null input of the null amplifier	162
Fig.6.1	Circuit to measure op amp gain and offset	164
Fig.6.2	Automated DAC static test set up	167
Fig.6.3	Typical DAC error plot	168
Fig.6.4	Full scale settling time of the DAC (10X attenuation due to picoprobe)	171
Fig.6.5	Linearity display at 10 MHz clock frequency	173
Fig.6.6	Worst case glitch of the DAC	174
Fig.6.7	Photomicrograph of the DAC chip	175

TABLES

Table 2.1	Comparison of Theoretical and Experimental Values of Correlation Coefficient for N-Channel Transistors	53
Table 2.2	Comparison of Theoretical and Experimental Values of Correlation Coefficient for P-Channel Transistors	55
Table 4.1	Drain Current Matching With And Without A Source Feedback Resistor	101
Table 6.1	Chopper Stabilized Amplifier Performance	166
Table 6.2	DAC Yield	169
Table 6.3	DAC Performance	176

LIST OF SYMBOLS AND ABBREVIATIONS

A	Amplifier d.c. open loop gain
A/D	Analog-to-Digital
C	Gate oxide capacitance per unit area
CMOS	Complementary Metal Oxide Semiconductor
D/A	Digital-to-Analog
DAC	Digital-to-Analog Converter
D_I	Threshold adjust implant dose
$E(x)$	Expected value of x
g_m	Transconductance
g_o	Output conductance
I_D	Drain current
K	Conductance Constant
L	Channel Length
MOS	Metal Oxide Semiconductor
MOSFET	MOS Field Effect Transistor
N_A	Substrate doping
n_i	Intrinsic carrier concentration
q	Electronic charge
Q_B	Depletion charge density per unit area
Q_f	Fixed oxide charge density per unit area
Q_m	Mobile ionic charge density per unit area
Q_{ot}	Oxide trapped charge per unit area
r	Correlation coefficient
t	Gate oxide thickness

VLSI	Very Large Scale Integration
V_{DS}	Drain to source voltage
$V_{DS,Sat}$	Drain to source Saturation voltage
V_{FB}	Flat band voltage
V_{GS}	Gate to source voltage
V_{SB}	Bulk to source reverse bias
V_T	Threshold voltage
W	Channel width
W_d	Depletion layer width
x_j	Junction depth
$\bar{x}$	Mean value of x
ϵ_{ox}	Permittivity of silicon dioxide
ϵ_s	Permittivity of silicon
μ	Channel mobility
σ_x	Standard deviation of x
ϕ_B	Fermi potential in the bulk
ϕ_{ms}	Gate-semiconductor work function difference

CHAPTER 1

INTRODUCTION

This thesis is concerned with the design of precision analog circuits in MOS VLSI. The performance of many of these circuits is critically dependent on the tight matching of components. Statistical methods are generally used to tackle problems like yield analysis and design centering in such circuits. The results of any statistical analysis are only as accurate as the characterization of the device electrical parameters. This work addresses two major aspects of precision analog design. First, characterization and modeling of the mismatch in MOS devices is carried out. The parameters that a circuit designer has freedom to choose are the dimensions of the devices. Therefore, analytical expressions have been derived that relate the mismatch to the device dimensions. Statistical methods are used to develop these models. Extensive measurements on specially designed test structures have been carried out to validate the analysis and extract the model parameters. The next objective is to demonstrate the application of the matching characterization in designing precision analog circuits. The design of a high performance 8-bit D/A converter is chosen as an example. The knowledge of matching of devices has made it possible to achieve a high speed circuit with a yield of greater than 97% without using any trimming, and yet occupying a small chip area.

1.1 Historical Overview

The traditional technology used for the implementation of precision analog functions has been bipolar. Both active and passive components have been observed to match well in this technology. However, MOS technology is the most generally used digital technology. This resulted in hybrid circuits consisting of one or more bipolar analog chips and an MOS chip to economically perform the digital functions. Further, the rapidly decreasing cost of digital integrated circuits made it possible for digital technology to displace analog circuitry in a number of application areas. The primary importance of analog circuits in VLSI systems, therefore, is in the area of interface circuits such as filters and data converters. These circuits connect the real world of analog, continuous time signals with the digital world of sampled encoded signals.

For increased reliability and cost effectiveness, it is desirable to have these analog functions and the digital circuits on the same chip. As MOS technology has the most desirable characteristics for realizing complex digital circuits, conventional bipolar integrated circuit and hybrid circuit approaches to realize analog functions were found unsuitable. Instead, circuit techniques which employ only MOS technology were sought. Perhaps the first reported work in this direction is the development of an all-MOS charge redistribution A/D conversion technique in 1975 [1]. MOS capacitors are used for data conversion here. Soon, this technique dominated the MOS data conversion

field. The important factor that prompted the wide spread use of this approach is the accuracy and stability possible with MOS capacitors when proper layout rules and process control are employed, capacitor ratio matching accuracy better than 0.1 percent can be achieved with virtually unmeasurable voltage and temperature effects [2].

In 1977, the switched capacitor filtering technique for MOS LSI was invented [3,4]. Here, the resistor found in RC active filter topologies is simulated by a switched capacitor. The compatibility of this technique with the charge redistribution data conversion approach made it possible to realize such complex functions as filter-codecs for telephony. The matching and other properties of MOS capacitors was published in 1981 [5]. To further assist the MOS analog designer, analytical models for the random mismatch in MOS capacitors was reported in 1982 [6].

1.2 Thesis Objectives

The first part of the thesis is devoted to characterization and modeling of the mismatch in MOS transistors and resistors. This would complement the knowledge that already exists for MOS capacitors. The need for such a study stems from the fact that precision analog circuits cannot be designed using capacitors alone. For certain applications such as high speed data conversion, capacitive techniques tend to be too slow. Further, a digital VLSI process may not offer

linear capacitors. The interest in such a study is evidenced by recent publications [7-8]. In [7] experimental results of matching of MOS current mirrors is discussed. However, this paper does not give a physical insight into the causes of mismatch. The work reported in [8] attempts to break down the causes of mismatch in MOS current sources, but it does not consider all the variables involved. Also the experimental results are limited to large area N-channel devices only. The work reported here is aimed at providing a complete understanding of the causes of mismatch in both P- and N-channel MOS devices. Experimental data is obtained from a 5 volt, 3 μ m, p-well CMOS process with n^+ poly silicon gates, that is in use at Northern Telecom Electronics Limited, Ottawa, Canada.

The characterization of the mismatch in MOS transistors is more complex than in the case of capacitors. The drain current matching not only depends on the device dimensions but also on the operating point. A characterization methodology is developed that accurately determines the mismatch in drain current over a wide operating region, using a minimum set of measurement data. The characterization is valid in the case of small geometry devices also.

The overall objective is not only to develop models for the mismatch behavior, but also to establish a comprehensive design approach for MOS precision analog circuits. The design of a high-speed, 8-bit current steering digital-to-analog converter (DAC) is chosen as an example. An important specification of the DAC is to

obtain at least 95% of circuits with less than $\pm 1/2$ LSB integral nonlinearity without any trimming. Based on statistical error analysis and the knowledge of matching behavior, the dimensions of the conversion elements are chosen to provide the required yield. It was also intended that this test vehicle be a useful library cell for custom/semi-custom IC design. Therefore, aspects such as glitch and drift were also considered. A new current steering bit cell is designed to provide low-glitch operation. Thermal drift of the DAC is minimized by designing a chopper stabilized control amplifier. Video bandwidth conversion speed has been achieved which was not possible before with capacitors. Salient features of the DAC are 80 ns conversion time, 500 pV-s glitch energy, and 23 ppm/ $^{\circ}$ C worst case thermal drift. These values are comparable to those achieved in hybrid DACs. As a result of this work, the circuit yield of the DAC exceeds 97% [9].

1.3. Thesis Format

The thesis is organized into seven chapters.

In Chapter 1, following the introductory remarks, the objectives of the work and the thesis format are stated.

An exhaustive study of transistor matching characteristics is provided in Chapter 2. The factors causing random mismatch are discussed in detail and analytical models are developed to relate the

mismatch to the device dimensions. Extensive measurement results are also reported. It is shown that the mismatch in the drain current can be determined accurately over a wide operating range with a minimum set of measured data. Finally, it is demonstrated that the characterization methodology is valid even for small geometry devices and that temperature variations do not affect the matching significantly.

The mismatch in integrated monolithic resistors is discussed in Chapter 3. First, a brief discussion of the different types of resistors available in a CMOS process is given. Then an analysis of the mismatch is provided followed by experimental data.

Chapter 4 is devoted to the design of a high-speed current steering DAC which serves as a test vehicle to demonstrate the application of the matching characterization studies. Based on a study of several configurations, an interstage divider multiple current source approach is chosen. Error analysis is presented to relate the circuit yield to the component matching accuracy. Design of a new current steering bit cell is presented followed by a discussion of a video-bandwidth DAC. This chapter concludes with the requirement of a low drift operational control amplifier necessary for the DAC to have a low drift.

The design of the control amplifier subsystem is presented in Chapter 5. The causes of offset and drift in MOS operational

amplifiers are presented followed by techniques to overcome them. A chopper stabilized control amplifier configuration is discussed. The design of a transconductance amplifier that serves as the basic amplifier for the chopper stabilized amplifier subsystem is also discussed.

The results of measurements on fabricated samples of the video DAC are presented in Chapter 6. Important parameters of the chopper stabilized control amplifier are discussed first. An automatic test set up for static testing and yield determination of the DAC is then presented. Finally, dynamic test results are reported.

Chapter 7 presents conclusions and suggestions for further research.

CHAPTER 2

TRANSISTOR MATCHING STUDIES

In this chapter we discuss the matching characteristics of both P- and N-channel transistors, and develop models to relate the random mismatch to the device dimensions. In general, there are two variations to consider in an integrated circuit process.

- (i) Total variation in a component value over a wafer or a batch, known as global variation.
- (ii) Variations in the value of a component on a die, with reference to an adjacent component on the same chip, known as local variation or mismatch.

From the earliest days of integrated circuit design; it is well known that component ratios are held within closer limits than absolute values of the components, and all precision designs are based on this fact. Therefore, we have concentrated on item (ii) above. Before we proceed any further we would like to elaborate on the phrase "modeling the mismatch". Usually, the term modeling is used to describe new physical mechanisms concerning the operation of a device and/or the analytical expressions for use in computer aided design. In this thesis we are not providing any new theories about the device behavior as such. On the other hand, based on the already existing knowledge of the device and the related models, we explore the physical mechanisms which cause random mismatch in MOS devices. We also develop

analytical expressions to relate the mismatch to the device dimensions. We name these quantitative relationships as mismatch models. This is the sense of the term modeling as used in this thesis.

We begin this chapter with a discussion of the design of test structures to extract the random mismatch data, followed by a description of the measurement set up. Then we discuss the measurement philosophy and show how the mismatch in the drain current can be determined accurately over a wide operating range with a minimum set of measured data. The factors causing the mismatch in threshold voltage (V_T) and conductance constant (K) are discussed next. Analytical models are derived and compared with experimental data. The correlation between the mismatches in V_T and K are discussed and related to the drain current mismatch in the saturation region. The next section deals with the range of applicability of the models. Although the threshold voltage of small geometry devices has a dimensional dependence, it is shown that the models developed are accurate down to and even beyond the feature size permitted by the technology. Finally, it is shown that temperature variation does not affect matching significantly.

2.1 Design of Test Structures

The test structures required to measure the mismatch must be so designed that all systematic error components are minimized so that random mismatch errors can be measured accurately. Figure 2.1 shows the arrangement used for the case of transistors. Transistors marked

1 & 2 are laid out in such a way that their channels run parallel to the wafer flat and we designate them as the horizontal pair. The vertical pair is formed by the devices 3 & 4, in which the channels run perpendicular to the wafer flat. Both the horizontal and vertical pairs of devices are enclosed by "dummy" devices at the ends as shown, so as to provide a uniform etching environment and thus minimize loading variations during dry etching [10]. This would minimize any systematic errors in edge definitions. With this arrangement we can study the effects of orientation of the device with respect to the wafer flat, if any. Also the mismatch in orthogonally laid out devices such as devices 2 and 3 can be attempted. In order to verify the analytical expressions for the mismatch such structures with transistor W/L ratios 6/3, 12/3, 12/6, 24/6, 24/12, 48/12, $\mu\text{m}/\mu\text{m}$ are designed for both P- and N-channel types.

2.2 Measurement Setup

The large amount of data required for statistical analysis necessitated an automatic measurement system. The system block diagram is shown in Figure 2.2. An electroglass 1034XA automatic prober is used to step the wafer. The many devices on a single die are efficiently characterized using an HP3695A Multiplexer Scanner. The Scanner switches analog input signals between the device under test (DUT) and an appropriate test/measuring instrument. The voltage inputs needed for the measurement are provided by a HP4140B pA Meter/DC voltage source.

Fig.2.1 Matched transistor lay out

Fig.2.2 Automated measurement set up

Although this instrument is capable of measuring currents in the pA range, its limited resolution of 3 1/2 digits was found to be inadequate. Therefore, a 5 1/2 digit Keithley 619 Electrometer/Multimeter was used for current measurement.

2.3 Measurement Philosophy

The measurement approach suitable for ratioed type precision analog work can be divided into two categories [11]:

- (i) Terminal approach -- measure two-port parameter data at the required operating point.
- (ii) Equivalent circuit approach -- measure the various parameters in a circuit model.

The advantages and disadvantages of these approaches for precision analog design are discussed below:

- (i) Terminal Approach: In this approach, a transistor can be viewed as a two terminal circuit element. The measured data can be obtained extremely precisely and used directly to investigate circuit performance. There is no loss of accuracy in the representation of the specific devices in question. Also the method is quick for a given application. An example of this approach is in the design of a current steering D/A converter [9]. The matching accuracy of a transistor of certain size, operating as a constant current source at a

certain current level is measured precisely and this data is used to predict the yield of the DAC. There are several disadvantages to this approach. Data at different temperatures and at several bias points imposes a measurement problem as well as a storage problem. Statistical data is expensive and therefore it is necessary to depend on a minimum set of data points and maximize their use. This approach requires the right kind of devices be available on which measurements can be made and hence new designs requiring new devices cannot be handled easily. Further, the available statistical data, in general, will not precisely indicate paths to take to improve any given situation.

(ii) Equivalent Circuit Approach: In this approach the mismatch in various parameters in a circuit model are measured. This eliminates the difficulties cited above and has the following advantages: The number of parameters is dependent only on the complexity of the model. The mismatch in these parameters can be related to the operating point and dimensions of the devices. Thus the mismatch in these parameters for different sized devices and/or at different operating points can be predicted without requiring new measurements. The main problem with the equivalent circuit approach is the range of applicability of the model particularly for small geometry devices.

2.3.1 Choice of the Model

The model chosen should be such that it gives an adequate description of the electrical behavior of the device, and at the same time should have readily measurable parameters that are amenable to statistical description. From the measurement standpoint the most important consideration is that of effectively decoupling the model parameters. This is done such that each measurement, or set of measurements, uniquely defines specific parameters. The ability to do this depends very much on the complexity of the model structure and may be difficult to achieve in more detailed circuit models. Also, an elaborate model may greatly exceed the accuracy of measurable data or may hamper the extraction of statistically significant model parameters. Further, the model should be such that the measured parameters may be applied over a wide range of operating conditions. Finally, it should be noted that we are only looking for local variations and are not so much concerned about the estimation of the absolute value of the parameters. In other words, any small model error would cancel out while estimating the mismatch so that a simple model would suffice for our application.

Based on the above criteria, we chose the square-law model which is a simplified form of the bulk charge model [12]. The statistically significant parameters of this model are the threshold voltage V_T , the conductance constant K , and the correlation between the mismatches in V_T and K . The mismatch in V_T accounts for the variations in the different charge quantities and in the gate oxide

capacitance. The variations in the dimensions, channel mobility and gate oxide capacitance are measured as the mismatch in K . As both V_T and K are dependent on the gate oxide capacitance, we need to measure the correlation between the mismatches in V_T and K . For small geometry devices, the threshold voltage is also dependent on the dimensions. The dimensional variations in a well controlled process are so small that these deviations can be accounted for entirely by the mismatch in K and do not contribute significantly to the mismatch in V_T . We elaborate on this further in Section 2.5 dealing with the range of applicability of the mismatch models we will be developing.

2.3.2 Description of the Model

The square-law model describes the current voltage relationship in the triode region as follows:

$$I_D = \frac{CW}{L} \left[V_{GS} - V_{FB} - 2\phi_B + \frac{\sqrt{2\epsilon_S q N_A (2\phi_B + V_{SB})}}{C} - \frac{V_{DS}}{2} \right] V_{DS} \quad (2.1)$$

$$= \frac{CW}{L} \left(V_{GS} - V_T - \frac{V_{DS}}{2} \right) V_{DS} \quad (2.2)$$

where

$$V_T = V_{FB} + 2\phi_B + \frac{\sqrt{2\epsilon_S q N_A (2\phi_B + V_{SB})}}{C} \quad (2.3)$$

is the threshold voltage or the turn on voltage of the device, i.e., the voltage required on the gate to induce a conducting channel.

There are several assumptions made while deriving the one dimensional bulk charge model. Since some of these assumptions are not strictly applicable, a discussion to justify the use of this model is in order.

The gradual channel approximation and the assumption that the substrate is uniformly doped do not hold good for small geometry devices. An accurate analysis calls for a two dimensional solution of Poisson's equation. However, in order to develop analytical expressions for the mismatch behavior, we use a one dimensional circuit model and apply appropriate corrections to the model such as, the effect of dimensional dependence on threshold voltage and nonuniform doping of the substrate. Also we have assumed a simplified picture of the oxide-silicon interface, that oxide fixed charge and interface trap charges are considered to be smeared-out uniform charge sheets. In fact the oxide fixed charge and charged interface traps are localized, and not sheets [14]. The accurate calculation of the surface potential would then be a two dimensional problem. We are simplifying the problem by estimating the aggregate of the localized nonuniformity over the entire area of the channel by using the simplified one dimensional circuit model. In fact it is this information that is required for circuit design. With these approximations, we develop analytical expressions for the mismatch behavior that compare remarkably well with experimental data.

The square-law model overestimates the drain current for any given bias point. The error is more for larger drain-to-source voltages. Also the saturation voltage predicted by this model is more than the actual one. The net effect is that the absolute value of the current predicted by this model will be different from the actual one both in triode and saturation regions. However, since we are interested in the mismatch in the otherwise identical devices, we assume that the model error component cancels out.

A typical plot of the drain current versus the gate-to-source voltage for a small value of drain-to-source voltage is shown in Figure 2.3. The curve shows three regions of operation. Initially, as the gate-to-source voltage is increased from zero, the current increases very slowly until the onset of strong inversion. Then, the current increases linearly with V_{GS} . Finally, the slope decreases due to mobility degradation at high transverse electric field. The weak inversion region and the high transverse electric field region are not described by the Grove model. However, the maximum slope of the curve is an accurate representation of the transconductance of the device as obtained by differentiating equation (2.2) with respect to V_{GS} . The intercept of the maximum slope on the V_{GS} axis is equal to $V_T + V_{DS}/2$. Therefore, measurement of the current-voltage characteristics in the triode region on matched pairs of transistors simultaneously provides the mismatches in the conductance constant (ΔK) and the threshold voltage (ΔV_T).

Fig.2.3. Current-Voltage relationship of a MOSFET in triode region

Generally, MOS transistors will be operating in the saturation region in analog circuits. Therefore, we should relate the measured mismatch in K and V_T to the saturation region. This is done as follows: The saturation voltage according to the square-law model is

$$V_{DS, \text{Sat}} = V_{GS} - V_T \quad (2.4)$$

Therefore, the drain current in saturation is given by

$$I_D = \frac{K}{2} (V_{GS} - V_T)^2 \quad (2.5)$$

where $K = \mu C W/L$

The addition of a channel length modulation term to this simple expression to correct for the increase in drain current with increase in drain-to-source voltage is really not necessary. This is because, while designing precision analog circuits, matched devices have identical bias conditions. Further, the degradation of mobility due to transverse electric field and velocity saturation are not taken into account in equation (2.5). This causes the conductance constant K , in the saturation region to be different from the one measured in the triode region. Again, we treat this as a constant error not affecting the mismatch measurements and proceed on to derive an expression for the mismatch in the drain current in the saturation region in terms of the measured mismatches in V_T and K , and the correlation between ΔV_T and ΔK , all measured in the triode region. Strictly speaking, the term correlation should be used only when there is a linear dependence between two quantities. The threshold voltage

is not linearly dependent on the oxide capacitance per unit area.

Yet, we use the term correlation to indicate the degree of dependence between V_T and K brought about by the variation in oxide capacitance.

Let σ_I , σ_{VT} , and σ_K be the standard deviations in the drain current, threshold voltage and conductance constant. Let r be the correlation coefficient between V_T and K . Using equation (A.13) σ_I can be expressed in terms of σ_K , σ_{VT} and r as

$$\frac{\sigma_I^2}{I^2} = \frac{\sigma_K^2}{K^2} + 4 \cdot \frac{\sigma_{VT}^2}{(V_{GS} - \bar{V}_T)^2} - 4r \frac{\sigma_{VT}}{(V_{GS} - \bar{V}_T)} \cdot \frac{\sigma_K}{\bar{K}} \quad (2.6)$$

where $\bar{K}$ is the expected value of K , and so on. Using equation (2.6) the standard deviation in drain current at any operating point can be determined if σ_K , σ_{VT} and r are known. In the next section we will analyze the causes for the mismatch in V_T and K and develop analytical models. We also discuss the correlation coefficient r .

2.4 Factors Causing Mismatch

The factors that are responsible for the mismatch are discussed in this section. Analytical expressions for σ_{VT} , σ_K and r are derived. This will provide an insight into the mismatch mechanisms and enable the circuit designer to optimally select device types, dimensions, etc., for any given situation. Experimental results are compared with theory.

2.4.1 Mismatch in V_T

The threshold voltage as given in equation (2.3) is

$$V_T = V_{FB} + 2\phi_B + \frac{Q_B}{C} \quad (2.7)$$

This expression assumes the ideal situation where there are no oxide charges and where the substrate is uniformly doped. Equation (2.7) must be modified to account for the oxide charges and the addition of a threshold adjust implant, if any.

Oxide charges include the oxide fixed charge Q_f , the oxide trapped charge Q_{ot} , and the mobile ionic charge Q_m . In the presence of these charges, the flat-band voltage is no longer equal to the gate-semiconductor work function difference ϕ_{ms} , but is given by

$$V_{FB} = \phi_{ms} - \frac{Q_o}{C} \quad (2.8)$$

$$\text{where } Q_o = Q_f + Q_{ot} + Q_m \quad (2.9)$$

For a well controlled process $Q_o \approx Q_f$.

In modern MOS processes, a shallow implant is done at the Si - SiO₂ interface for threshold voltage adjustment. A simple approximation to the effect of this implant is to treat the implanted ions as a delta function of charge at the Si - SiO₂ interface. This is equivalent to the modification of the fixed oxide charge by an amount $\pm q D_I$, where D_I is the implant dose [13]. This results in a

threshold shift

$$\Delta V_T = V_T(D_I) - V_T(D_I=0) = \pm q D_I / C \quad (2.10)$$

with the +ve sign when the dopant is ionized acceptors and -ve sign for donors. Incorporating equations (2.8) and (2.10) into (2.7)

$$V_T = \phi_{ms} + 2\phi_B + \frac{Q_B}{C} - \frac{Q_f}{C} \pm \frac{q D_I}{C} \quad (2.11)$$

$$\text{where } Q_B = \sqrt{2\epsilon_s q N_A (2\phi_B + V_{SB}) / C} \quad (2.12)$$

The standard deviation in V_T may be computed if we can find the standard deviations of the various terms on the right hand side of equation (2.11). ϕ_B has a logarithmic dependence on the doping in the substrate, and ϕ_{ms} has a similar dependence on the substrate doping and the doping of the polysilicon gate. Therefore any small variations in these doping concentrations over adjacent devices will have negligible effect on the threshold voltage through ϕ_{ms} and ϕ_B . Hence these terms may be regarded as constants.

Next we consider the oxide fixed charge. It has been reported that this has a Poisson distribution [14, pp.242]. If $\bar{Q}_f$ is the mean oxide fixed charge per unit area, then its variance is given by

$$\sigma_{Qf}^2 = \frac{q \bar{Q}_f}{L W} \quad (2.13)$$

where $\bar{L}$ and $\bar{W}$ are the mean values of the length and width of the channel.

The depletion charge per unit area Q_B is also a random variable dependent on the distribution of the dopant atoms. This is an important difference between the treatment given here and the one reported in [8], where Q_B is treated as a constant. In fact, it is reported in [14, pp.237] that the dopant ions are nonuniformly distributed in MOS devices. No theoretical treatment of fluctuations in dopant ion density is available. However, we shall show that the physical conditions in the substrate favour a Poisson distribution [15]. The number of atoms per unit volume in silicon is $5 \times 10^{22} \text{ cm}^{-3}$. Only a very small fraction of these sites are occupied by the dopant atoms. The number of dopant atoms in nonoverlapping volumes are independent. Further, for domains of sufficiently small volume, the probability of finding exactly one dopant atom in a domain is proportional to the volume, and the probability of finding more than one atom is negligible. Hence the dopant ions may be considered to have a Poisson distribution. Then if $\bar{W}_d$ is the expected value of the depletion width, the variance in the dopant ions per unit volume is

$$\sigma_{NA}^2 = \frac{\bar{N}_A}{L \bar{W} \bar{W}_d} \quad (2.14)$$

The depletion charge per unit area is given by

$$Q_B = \frac{\sqrt{2\epsilon_s q N_A (2\phi_B + V_{SB})}}{C}$$

Hence

$$\frac{\sigma_{QB}^2}{\bar{Q}_B^2} = \frac{1}{4 L \bar{W} \bar{W}_d \bar{N}_A} \quad (2.15)$$

Now we consider the variability due to the threshold adjust implant. We will again assume the implant ions to follow a Poisson distribution. If $\bar{D}_I$ is the mean implant dose per unit area, then the variance of D_I is given by

$$\sigma_{DI}^2 = \frac{\bar{D}_I}{\bar{L} \bar{W}} \quad (2.16)$$

The only other random variable to consider is C , the oxide capacitance per unit area. It depends on the oxide thickness and the permittivity of the oxide. First we shall consider the nonuniformity of the oxide thickness. The gate oxide layer may be divided into a number of small regions over which the thickness is uniform. As shown in Figure 2.4, let the oxide thickness in an infinitesimal area $dx dy$ at the point (x,y) be $t(x,y)$. Then the equivalent oxide thickness required to compute the capacitance is the average thickness of such elemental regions. If $\bar{L}$ and $\bar{W}$ are the expected values of the dimensions of the device, then the equivalent thickness is

$$\tau = \frac{1}{\bar{L} \bar{W}} \int_0^{\bar{L}} \int_0^{\bar{W}} t(x,y) dx dy \quad (2.17)$$

Here we are fixing the dimensions of the device and considering only the variability in the thickness. Our task is to find the variance in τ . The evaluation of the integral in equation (2.17) can be simplified by regarding $t(x,y)$ as a stationary random process, meaning that the two processes $t(x,y)$ and $t(x+h, y+k)$ have the same statistics for any h and k [16, pp.300]. This assumption is not strictly true due to the formation of birds-beak which causes the oxide to be thicker near

Fig.2.4 Gate oxide thickness variation

the edges of the channel. Since we are dealing with the mismatch in two otherwise identical devices, we can regard the birds-beak to be a systematic component and not contributing to the random mismatch, and thus can proceed with the assumption that $t(x,y)$ is stationary.

From equation (2.17), the expected value of τ is given by

$$\begin{aligned} E(\tau) &= \frac{1}{L W} \int_0^L \int_0^W E(t(x,y)) dx dy \\ &= \frac{1}{L W} \int_0^L \int_0^W \bar{t} dx dy \\ &= \bar{t} \end{aligned} \tag{2.18}$$

where $\bar{t}$ is the mean of $t(x,y)$.

The variance of τ is evaluated in a similar manner as in [8]. The mathematical derivation is given in Appendix B.

$$\sigma_\tau^2 = \frac{1}{L W} \sigma_t^2 d_1^2 \tag{2.19}$$

where d_1 is the region over which the autocovariance function of the oxide thickness has any significant value.

Thus the standard deviation in the equivalent gate oxide thickness τ is inversely proportional to the square root of the gate area.

The variability in the permittivity of the oxide may be treated in a similar way as the oxide thickness by defining an equivalent permittivity as

$$\bar{\epsilon} = \frac{1}{L \bar{W}} \int_0^L \int_0^{\bar{W}} \epsilon(x,y) dx dy$$

$$E(\bar{\epsilon}) = \bar{\epsilon} \quad (2.20)$$

and the variance of ϵ will be

$$\sigma_{\bar{\epsilon}}^2 = \frac{1}{L \bar{W}} \sigma_{\epsilon}^2 d_2^2 \quad (2.21)$$

The oxide capacitance per unit area is

$$C = \bar{\epsilon} / \tau \quad (2.22)$$

The permittivity of the oxide is independent of the thickness.

Therefore, from equation (A.13)

$$\frac{\sigma_C^2}{\bar{C}^2} = \frac{\sigma_{\bar{\epsilon}}^2}{\bar{\epsilon}^2} + \frac{\sigma_{\tau}^2}{\tau^2} \quad (2.23)$$

Using equations (2.18), (2.19), (2.20) and (2.21)

$$\begin{aligned} \frac{\sigma_C^2}{\bar{C}^2} &= \frac{1}{L \bar{W}} \left(\frac{\sigma_t^2 d_1^2}{\tau^2} + \frac{\sigma_{\epsilon}^2 d_2^2}{\bar{\epsilon}^2} \right) \\ &= \frac{1}{L \bar{W}} A_{ox} \end{aligned} \quad (2.24)$$

where A_{ox} is the quantity in parenthesis and has the units of cm^2 .

This parameter will be extracted from measurements.

The random variables Q_f , Q_B , D_I and C are all independent.
Hence the variance in V_T may be written using equation (2.11) as

$$\sigma_{VT}^2 = \left(\frac{1}{C}\right)^2 \sigma_{QB}^2 + \left(-\frac{\bar{Q}_B}{C^2}\right)^2 \sigma_C^2 + \left(-\frac{\bar{Q}_f}{C^2}\right)^2 \sigma_C^2 + \left(\frac{1}{C}\right)^2 \sigma_{Qf}^2$$

$$+ \left(\frac{q}{C}\right)^2 \sigma_{DI}^2 + \left(-\frac{q \bar{D}_I}{C^2}\right)^2 \sigma_C^2$$

i.e.

$$\sigma_{VT}^2 = \frac{1}{C^2} (\sigma_{QB}^2 + \sigma_{Qf}^2 + q^2 \sigma_{DI}^2)$$

$$+ \frac{\sigma_C^2}{C^2} \left(\frac{\bar{Q}_B^2}{C^2} + \frac{\bar{Q}_f^2}{C^2} + \frac{q^2 \bar{D}_I^2}{C^2} \right) \quad (2.25)$$

Substituting equations (2.15), (2.16) and (2.24) into (2.25) we have

$$\sigma_{VT}^2 = \frac{1}{L W C^2} [q(\bar{Q}_B + \bar{Q}_f + q \bar{D}_I) + A_{ox}(\bar{Q}_B^2 + \bar{Q}_f^2 + q^2 \bar{D}_I^2)] \quad (2.26)$$

Now let us examine the importance of the various terms on the right hand side of equation (2.26), for both P- and N-channel devices.

Consider N-channel devices first. In our process, the threshold adjust implant is carried out for P-channel transistors only.

Therefore, $q D_I = 0$ for N-channel devices. The P-well doping concentration is typically of the order of $10^{16}/\text{cm}^3$. For a typical gate oxide thickness of 500 Å, the oxide capacitance per unit area is $6.9 \times 10^{-8} \text{ F/cm}^2$. The quantities required to estimate σ_{VT} are calculated below.

Assuming uniform doping in the substrate, the Fermi potential is given by

$$\phi_B = \frac{KT}{q} \ln \frac{N_A}{n_i} = 0.3723 \text{ V}$$

and the depletion region width is

$$W_d = \left[\frac{2 \epsilon_s (2\phi_B)}{q N_A} \right]^{1/2} = 2.005 \times 10^{-5} \text{ cm}$$

Therefore, the depletion charge density per unit area is given by

$$\begin{aligned} Q_B &= q N_A W_d \\ &= 7.7 \times 10^{-8} \text{ Coulombs/cm}^2 \end{aligned} \quad (2.27)$$

In a well controlled process the number of fixed oxide charges can be reduced to about $2 \times 10^{10} / \text{cm}^2$. Then

$$Q_f = 1.6 \times 10^{-19} \times 2 \times 10^{10} = 3.2 \times 10^{-9} \text{ Coulombs/cm}^2 \quad (2.28)$$

Comparing equations (2.27) and (2.28), it is seen that the fixed oxide charge is an order of magnitude smaller than the depletion charge and hence its contribution to the mismatch in threshold voltage (equation (2.26)) may be neglected.

The measured relative standard deviation of the threshold voltage (σ_{V_T}/V_T) is plotted as a function of the channel area in Figure 2.5. For collecting statistics, 128 device pairs of each type and size were measured on every wafer. The vertical error bars reflect the spread in measured values over four different wafers. The measured data fit well with the theoretical straight line relationship given by equation (2.26). This equation may be approximated for

N-channel devices as,

$$\sigma_{VT}^2 = \frac{1}{[\bar{L} \bar{W} \bar{C}^2]} [q \bar{Q}_B + A_{ox} \bar{Q}_B^2] \quad (2.29)$$

Substituting the values of $\bar{Q}_B$ and $\bar{C}$ we have

$$\frac{\sigma_{VT}}{V_T} = \left(\frac{1}{[\bar{L} \bar{W}]} \right)^{1/2} (2.5875 \times 10^{-12} + A_{ox} 1.2421)^{1/2} / V_T \quad (2.30)$$

where the effective dimensions $\bar{L}$ and $\bar{W}$ are in units of cm. Now we compare the measured slope with the theoretical slope. The slope factor using equation (2.30) is

$$(2.5875 \times 10^{-12} + A_{ox} 1.2421)^{1/2} / 0.7$$

where we have assumed an average threshold voltage of 0.7 volts. The measured slope is 2.33×10^{-6} cm. Comparing this with the theoretical value, we have

$$[2.5875 \times 10^{-12} + 1.2421 A_{ox}]^{1/2} = 0.7 \times 2.33 \times 10^{-6}$$

$$\text{Therefore, } A_{ox} = 6.4631 \times 10^{-14} \text{ cm}^2 \quad (2.31)$$

Substituting this value of A_{ox} in equation (2.24), we can estimate the variation in the oxide capacitance.

$$\frac{\sigma_C}{\bar{C}} = \left(\frac{1}{[\bar{L} \bar{W}]} \right)^{1/2} 2.5423 \times 10^{-7} \quad (2.32)$$

For example, for a $24 \times 6 \mu\text{m}^2$ gate, $\sigma_C / \bar{C}$ is 0.02%. This low value agrees with the extremely uniform nature of the gate oxide observed in other measurements [17].

Now, we consider P-channel devices. The threshold-adjust boron implant is done through the gate oxide. As it is a very shallow implant, a considerable portion of the implanted ions is retained in the oxide. Although this results in the formation of charge states in the oxide, they are readily annealed during subsequent processing [18]. However, due to the random distribution of the impurity atoms in the gate oxide, we may expect a slightly poorer capacitance matching than in the case of N-channel devices. We now estimate σ_{VT} using the following typical process parameters.

$$\text{N-substrate concentration} = 10^{16} \text{ cm}^{-3}$$

$$\text{Implant Dose} = 5 \times 10^{11} / \text{cm}^2$$

$$\phi_B = \frac{KT}{q} \ln \frac{N_A}{n_i} = 0.35 \text{ V}$$

$$W_d = \left[\frac{2\epsilon_s (2\phi_B)}{q N_A} \right]^{1/2}$$

$$= 3.012 \times 10^{-5} \text{ cm}$$

$$Q_B = q N_A W_d$$

$$= 4.819 \times 10^{-8} \text{ Coulomb/cm}^2$$

$$q D_I = 8.0 \times 10^{-8} \text{ Coulomb/cm}^2$$

$$Q_f = 3.2 \times 10^{-9} \text{ Coulomb/cm}^2$$

Again, as Q_f is an order of magnitude smaller than Q_B , equation (2.26) may be approximated as

$$\sigma_{VT}^2 = \frac{1}{L W C^2} [q (\bar{Q}_B + q \bar{D}_I) + A_{ox} (\bar{Q}_B^2 + q^2 \bar{D}_I^2)]$$

Substituting the values of $\bar{Q}_B$, $q \bar{D}_I$ and C , we have,

$$\frac{\sigma_{VT}}{V_T} = \left(\frac{1}{L W} \right)^{1/2} (4.2945 \times 10^{-12} + A_{ox} 1.8463)^{1/2} / V_T \quad (2.33)$$

The coefficients in equation (2.33) are larger than the corresponding ones in equation (2.30) indicating a larger mismatch in P-channel devices owing to the additional threshold adjust implant. This may be physically interpreted as a larger variation in the surface concentration due to the differential doping occurring at the surface.

The mismatch in threshold voltage for P-channel devices is plotted on the same set of axes as N-channel devices to provide a ready comparison. It is shown in Figure 2.5. Again we will compare the slope of the fitted line with the theoretical value. Using equation (2.33) and for an average threshold voltage of 0.8 V, the theoretical slope is

$$\frac{1}{0.8} [4.2945 \times 10^{-12} + A_{ox} (1.8463)]^{1/2}$$

Comparing this with the measured value of 3.9333×10^{-6} cm, we find that

$$A_{ox} = 3.0369 \times 10^{-12} \text{ cm}^2 \quad (2.34)$$

and

$$\frac{\sigma_C}{C} = \left(\frac{1}{L W} \right)^{1/2} 1.7427 \times 10^{-6} \quad (2.35)$$

Comparison of equations (2.32) and (2.35) suggest that the gate oxide capacitance matching is poorer for P-channel devices than for N-channel ones, as we anticipated. Such a behavior is also noticed while comparing the conductance constant mismatch in the P- and N-channel devices in Section 2.4.2.

In Figure 2.5 the mismatch is plotted as a function of the effective area. The effective channel length is less than the drawn channel length due to the lateral diffusion of the source drain regions. The encroachment of the field oxide region and the channel stop implant causes the effective width to be smaller than the drawn width. The effective length was measured by using an array of devices with constant width but differing lengths. Similarly, the effective width was measured with an array of devices having constant length but differing widths. The measurements were done on several different locations on each of the wafers used to measure the mismatch. The spread in these values appear as the horizontal error bars in Figure 2.5. As this effect is felt most for the smaller devices, the measured spread in mismatch is also more for the smaller devices. At this point, it may be noted that the effective channel length is independent of the gate bias while the effective width is not. The encroachment of the field oxide region and the channel stop implant causes a larger proportion of the channel region to be inverted as the gate bias is increased. The appropriate gate bias to choose for the effective width measurement is that voltage where the V_T and K were measured to determine the mismatch. It may also be noted that in the

Fig.2.5 Threshold voltage mismatch versus dimension

case of P-channel devices, there is no channel stop implant, and hence the effective width is slightly larger than that in the case of N-channel devices. It was also observed that the effective length for the P-channel devices is larger than the corresponding quantity for N-channel transistors. The net result of these effects is that there is a shift in the x-direction in Figure 2.5 between the P- and N-channel devices for the same drawn W/L. Finally, it may be seen that for both P- and N-channel devices, the measured mismatch for the larger devices is more than the theoretical value predicted by the straight line relationship. This is because the measurement errors become more prominent when the actual mismatch is very small.

In our analysis, we have assumed that the P-channel device is a surface channel transistor. The threshold adjust implant can be so adjusted to form a normally off buried channel device [13]. Then the analytical expression for the threshold voltage mismatch may be worked out starting from the buried channel threshold model discussed in [19].

We will now summarize the threshold voltage mismatch behaviour. These findings are particular to this thesis.

- (a) The standard deviation of mismatch is inversely proportional to the square root of the effective channel area.
- (b) In a well controlled process, the fixed oxide charge density is about an order of magnitude smaller than the depletion charge density. Hence the nonuniform distribution of the

fixed oxide charges has negligible effect on threshold voltage mismatch.

- (c) The nonuniform distribution of the dopant atoms in the bulk is a major contributor to the threshold voltage mismatch. The assumption that these atoms follow a Poisson distribution has resulted in excellent agreement with measurements.
- (d) Devices which use a compensating threshold adjust implant have a higher mismatch in threshold voltage due to the differential doping occurring at the surface. This is the major reason for the significantly larger mismatch noticed in P-channel devices as compared to N-channel transistors.
- (e) The variation in the gate oxide capacitance is quite uniform and hence has little influence on the threshold voltage mismatch. However, between N- and P-channel devices, the gate oxide capacitance of the latter has slightly poorer matching characteristic. This could be due to the nonuniform distribution of the threshold adjust implant atoms in the gate oxide.

2.4.2 Mismatch in K

The conductance constant is given by

$$K = \mu C W/L \quad (2.36)$$

We can express the variance in K in terms of the variances in μ , C , W and L . Let us first consider the length of the device.

The channel length is determined by the etching of the polysilicon gate. Electrically, the channel length is the average distance between the source and the drain diffusions. Any raggedness in the definition of the poly may not be exactly reproduced in the source and drain diffusion edges. Further, we are not so much concerned about this raggedness; rather we are interested in the difference in the electrical length from one device to the next. Such a mismatch in length may not be due to the nonuniformity of the edge alone. In the absence of a complete knowledge of the causes of variation in length, we will simply call the variance of the length as σ_L^2 and make no attempt to derive any expression for it. In [8], the nonuniformity of the edges is considered as the only cause for the mismatch and an expression for σ_L^2 is derived which is inversely proportional to the width of the device. This would mean that the mismatch in length would tend to zero for very wide devices. We have observed results that contradict this. In fact we have noticed that σ_L is more or less independent of the device width.

The width of the device may be treated similarly and let the variance in W be σ_W^2 . The definition of the length and width occur during different stages of processing. Also the layers that are etched, the etchants, and other process conditions are all different. Hence L and W may be treated as independent random variables.

The oxide capacitance per unit area is already considered in connection with the threshold voltage mismatch. The variance of C

is given by equation (2.24).

Next we consider the channel mobility. We briefly reproduce the treatment given in [20] and use the empirical relations developed there to determine the standard deviation of mismatch in electron mobility.

In general, at least three different scattering mechanisms have been proposed to account for the mobility behavior in the region where the gate voltage is above threshold.

- (1) Phonon scattering due to the various modes of lattice vibration including surface acoustic wave phonons and optical phonons. Phonon scattering is important at room temperature and can be ignored at very low temperatures.
- (2) Coulomb scattering due to charged centers, including fixed oxide charge, interface-state charge, and localized charge due to ionized impurities. The effects of Coulomb scattering are important for lightly inverted surfaces. High surface charge densities or substrate doping concentrations imply increased Coulomb scattering. This type of scattering becomes less effective for a heavily inverted surface because of carrier screening.
- (3) Surface roughness scattering is caused by the deviation of the interface from the ideal plane. This is important under strong inversion conditions because the strength of the interaction is governed by the distance of the carriers from the surface; the closer the carriers are to the surface, the stronger the scattering due to surface roughness will be.

The relative importance of these scattering mechanisms depends on the operating temperature and the strength of the electric field. At low temperatures, the mobility is determined by the combined effects of Coulomb scattering, which dominates the low-gate-field region, and by surface-roughness scattering which dominates the high field region. At room temperature, the mobility is governed by Coulomb scattering due to charged centers and phonon scattering in the low field region. It is dominated by surface roughness and phonon scattering under strong inversion.

The mobility near the threshold voltage at room temperature is mainly governed by scattering due to interface charge centers and phonons, except for more heavily doped samples where surface-roughness scattering may also be important. The effect of impurity concentration on the maximum mobility μ is mainly through the gate field required to produce inversion. The mobility is primarily a function of effective gate field rather than of the carrier density; more lightly doped samples have lower fields at the onset of inversion and correspondingly higher μ . Experimentally, it is observed that μ at room temperature decreases hyperbolically with increasing fixed oxide charge density (Q_f) for a given doping level and is independent of the surface orientation. This hyperbolic behavior with Q_f does not change with the doping concentration, even though as expected, the mobility does decrease as concentration increases, due to the higher electric field required to produce inversion. An empirical relationship for μ as a function of Q_f and N_A is

$$\mu = \frac{\mu_0(N_A)}{1 + \alpha(N_A)N_f} \quad (2.37)$$

where N_A is the average surface impurity concentration in cm^{-3} , N_f is the fixed oxide charge carrier density in units of 10^{11} cm^{-2} , and

$$\mu_0(N_A) = 3490 - 164 \log N_A \quad (2.38)$$

$$\alpha(N_A) = -1.04 \times 10^{-1} + 1.93 \times 10^{-2} \log N_A \quad (2.39)$$

Equations (2.38) and (2.39) indicate that both μ_0 and α are very slowly varying functions of the dopant concentration. Therefore, the mismatch in μ may be approximated to be entirely due to the mismatch in Q_f . Then the variance in the mobility is

$$\begin{aligned} \sigma_\mu^2 &= \frac{\mu_0^2 \alpha^2}{(1 + \alpha Q_f)^4} \cdot \sigma_{Nf}^2 \\ &= \frac{\mu_0^2 \alpha^2}{(1 + \alpha Q_f)^4} \frac{N_f}{L W} \end{aligned} \quad (2.40)$$

as the fixed oxide charge has a Poisson distribution. Thus the standard deviation of the maximum mobility is inversely proportional to the square root of the effective area of the channel. The order of magnitude of this deviation may be estimated as follows.

The expected value of μ may be expressed as

$$\begin{aligned}\mu &= \frac{\mu_0}{1 + \alpha \bar{N}_f} + \frac{\mu_0 \alpha^2}{(1 + \alpha \bar{N}_f)^3} \cdot \sigma_{Nf}^2 \\ &= \frac{\mu_0}{(1 + \alpha \bar{N}_f)} \left[1 + \frac{\alpha^2}{(1 + \alpha \bar{N}_f)^2} \cdot \frac{\bar{N}_f}{L \bar{W}} \right]\end{aligned}\quad (2.41)$$

When N_f is expressed in units of cm^{-2} , α is given by

$$\alpha = (-1.04 \times 10^{-1} + 1.93 \times 10^{-2} \log N_A) \times 10^{-11}$$

and for a typical example of $N_f = 2 \times 10^{11}/\text{cm}^2$ and $N_A = 10^{16}/\text{cm}^3$, the second term in the square brackets of (2.41) is very small compared to 1 even for very small devices. Then

$$\mu = \frac{\mu_0}{1 + \alpha \bar{N}_f} \quad (2.42)$$

Substituting equation (2.42) into (2.40),

$$\frac{\sigma_\mu^2}{\mu^2} = \frac{\alpha^2}{(1 + \alpha \bar{N}_f)^2} \cdot \frac{\bar{N}_f}{L \bar{W}}$$

Using the values of α , $\bar{N}_f$, and N_A given above,

$$\frac{\sigma_\mu}{\mu} = \frac{1}{(L \bar{W})^{1/2}} 4.95 \times 10^{-7} \quad (2.43)$$

where L and $\bar{W}$ have the units of cm. Equation (2.43) indicates that the mismatch in the maximum mobility is small. For example, for a $24 \times 6 \mu\text{m}^2$ device, the mismatch in mobility is 0.04%. In comparison,

the average standard deviation of mismatch in K is 0.25%.

The discussion given above for the mobility mismatch is for electrons only. Unfortunately, we are not aware of any model that relates the mobility of holes to the doping concentration in the bulk and the fixed oxide charge density. The situation in the case of P-channel devices is further complicated by the threshold adjust implant. This could cause some damage in the substrate which may not be completely annealed, resulting in a poorer mobility matching than in the case of N-channel devices. In spite of these uncertainties, it is still reasonable to assume that the mobility mismatch has a similar dependence on channel dimensions as given by equation (2.43). This is due to the Poisson type distribution of the various non-uniformities. Therefore, we may write, in general,

$$\frac{\sigma_{\mu}}{\mu} = \left(\frac{A}{L W} \right)^{1/2} \quad (2.44)$$

where $\sqrt{A} = 4.95 \times 10^{-7}$ cm for N-channel devices and is not known for P-channel transistors.

Having determined the mismatch in the various factors contributing to the conductance, we may write the mismatch in K from equation (2.36). The factors on the right hand side of this equation are all independent. Thus,

$$\begin{aligned} \left(\frac{\sigma_K}{K}\right)^2 &= \frac{\sigma_\mu^2}{\mu^2} + \frac{\sigma_C^2}{C^2} + \frac{\sigma_L^2}{L^2} + \frac{\sigma_W^2}{W^2} \\ &= \frac{1}{L W} (A_\mu + A_{ox}) + \frac{\sigma_L^2}{L^2} + \frac{\sigma_W^2}{W^2} \end{aligned} \quad (2.45)$$

after substituting for the capacitance and mobility variations from equations (2.24) and (2.44) respectively. For N-channel devices, $A_\mu = 2.45 \times 10^{-13} \text{ cm}^2$ and the estimated value of A_{ox} from equation (2.31) is $6.4631 \times 10^{-14} \text{ cm}^2$. Then equation (2.45) may be written as

$$\left(\frac{\sigma_K}{K}\right)^2 = \frac{1}{L W} (2.45 \times 10^{-13} + 0.6463 \times 10^{-13}) + \frac{\sigma_L^2}{L^2} + \frac{\sigma_W^2}{W^2} \quad (2.46)$$

Using the measured values of (σ_K/K) for different sized devices, equation (2.46) can be solved for σ_L and σ_W . It is found that both σ_L and σ_W are approximately the same and are in the range 0.01 to 0.03 μm . To provide a feel for the relative importance of the factors contributing to the mismatch in K , we may substitute $\sigma_L = \sigma_W = 0.02 \mu\text{m}$ in equation (2.46). Then,

$$\left(\frac{\sigma_K}{K}\right)^2 = (2.45 \times 10^{-13} + 0.646 \times 10^{-13}) \frac{1}{L W} + 4 \times 10^{-12} \left(\frac{1}{L^2} + \frac{1}{W^2}\right) \quad (2.47)$$

where the effective dimensions L and W have the units of cm. The coefficient of the $(1/L^2 + 1/W^2)$ term is larger than that of $1/L W$. Therefore, we have plotted (σ_K/K) versus $(1/L^2 + 1/W^2)^{1/2}$ in Figure 2.6(a).

Fig.2.6(a) Conductance constant mismatch versus dimension for N-channel transistors

For the case of P-channel devices, as we have already stated, the mobility mismatch is not known. The oxide variation as given by equation (2.34) is $A_{ox} = 3.0369 \times 10^{-12} \text{ cm}^2$. Assuming that the edge variations are of the same order as for N-channel devices, i.e., $\sigma_L = \sigma_W = 0.02 \text{ } \mu\text{m}$, we may write equation (2.46) as

$$\left(\frac{\sigma_K}{\bar{K}}\right)^2 = (A_{\mu} + 3.0369 \times 10^{-12}) \frac{1}{\bar{L} \bar{W}} + 4 \times 10^{-12} \left(\frac{1}{\bar{L}^2} + \frac{1}{\bar{W}^2}\right) \quad (2.48)$$

The measured $(\sigma_K/\bar{K})$ values are plotted against $(1/\bar{L}^2 + 1/\bar{W}^2)^{1/2}$ in Figure 2.6(b). Comparing this with the results of N-channel devices shown in Figure 2.6(a), it is seen that the P-channel devices have a larger mismatch in conductance constant. One reason for this behavior is the poorer gate oxide capacitance matching as has already been pointed out in Section 2.4.1. Another factor could be a larger mobility variation.

Mismatch in conductance constant is not a linear function of $(1/\bar{L}^2 + 1/\bar{W}^2)^{1/2}$ as seen in Figures 2.6(a) and 2.6(b). The curvature is due to the $(1/\bar{L} \bar{W})$ term in equations (2.47) and (2.48). The curvature increases as the device dimensions decrease owing to the increasing contribution of the $1/\bar{L} \bar{W}$ term. The ratio of $(1/\bar{L} \bar{W})^{1/2}$ to $(1/\bar{L}^2 + 1/\bar{W}^2)^{1/2}$ is 0.48 for 48/12 devices and increases of 0.63 for the 6/3 devices.

We will now summarize the mismatch behavior in conductance constant. These are particular findings of this thesis.

Fig.2.6(b), Conductance constant mismatch versus dimension for P-channel transistors

- (a) The mismatch in K due to edge variations is proportional to $(1/L^2 + 1/W^2)^{1/2}$. The standard deviation of mismatch in length and width is in the range 0.01 to 0.03 μm . For N-channel devices, this is the dominant source of mismatch in K .
- (b) The larger gate oxide capacitance variation observed in P-channel devices in connection with V_T mismatch agrees with the larger mismatch in K .
- (c) For N-channel devices, the variation in mobility has the little effect on the mismatch in K . The corresponding quantity for P-channel transistors, however, could be larger.

2.4.3 Correlation Between V_T and K

The threshold voltage as well as the conductance constant are functions of the gate oxide capacitance. Therefore we can expect a dependence between V_T and K . Generally, the correlation coefficient is used as a measure of dependence between two quantities. In the strict mathematical sense, the term correlation is used only when there is a linear relationship between two quantities. In a case such as this we are only allowed to express the covariance. Unfortunately, the covariance does not really give a good feel for the dependence. On the other hand, the correlation coefficient is a unit-less quantity in the range -1 to 1, and gives a better insight into the nature of

dependence. The correlation between the absolute value of V_T and K will have a negative sign for N-channel devices and positive sign for P-channel devices, as shown below.

For a P-well CMOS process, with n^+ poly silicon gate and no threshold adjust implant, the threshold voltage of N-channel devices is given by

$$\begin{aligned} V_T &= \phi_{ms} + 2\phi_B + \frac{Q_B}{C} \\ &= -0.36 + \frac{Q_B}{C} \end{aligned} \quad (2.49)$$

where typical values of ϕ_{ms} and ϕ_B are used. The second term in equation (2.49) has a larger magnitude than the first one. Therefore, an increase in the oxide capacitance value decreases the threshold voltage. For the case of P-channel devices, however, a threshold adjust implant is carried out and the threshold voltage is approximated by

$$\begin{aligned} V_T &= \phi_{ms} + 2\phi_B - \frac{Q_B}{C} + \frac{q D_I}{C} \\ &= -1.1 - \frac{Q_B}{C} + \frac{q D_I}{C} \end{aligned} \quad (2.50)$$

The third term in equation (2.50) has a larger magnitude than the second one and hence the absolute value of the threshold voltage increases with increase in capacitance. Thus the correlation between

magnitude of V_T and K will have opposite signs in case of P-channel devices as compared to N-channel transistors. This causes some difficulty in deriving a generalized expression for the dependence of V_T on K . This problem can be overcome if we can determine the correlation between the mismatches in V_T and K . After all, we are more interested in the differences in V_T and in K , and not so much concerned about their absolute values. The threshold voltage may be written in general as

$$V_T = A + Q/C \quad (2.51)$$

where A is a constant given by

$$A = \phi_{ms} + 2\phi_B$$

and $\frac{Q}{C} = \frac{Q_B}{C}$ for N-channel

$$= \frac{q D_I}{C} - \frac{Q_B}{C} \text{ for P-channel}$$

For both N- and P-channel devices, A is negative and Q/C is positive. The difference in V_T can be expressed in terms of the partial derivatives as

$$\Delta V_T = \frac{\bar{Q}}{\bar{C}} \left(\frac{\Delta Q}{\bar{Q}} - \frac{\Delta C}{\bar{C}} \right) \quad (2.52)$$

Similarly, the difference in K is given by

$$\Delta K = K \left[\frac{\Delta \mu}{\bar{\mu}} + \frac{\Delta C}{\bar{C}} + \frac{\Delta W}{\bar{W}} - \frac{\Delta L}{\bar{L}} \right] \quad (2.53)$$

The correlation coefficient is defined as

$$r = \frac{\mu_{11}}{\sigma_{\Delta VT} \sigma_{\Delta K}} \quad (2.54)$$

where $\sigma_{\Delta VT} = \sigma_{VT}$

and $\sigma_{\Delta K} = \sigma_K$

and the covariance μ_{11} is given by

$$\mu_{11} = E(\Delta V_T \Delta K) - E(\Delta V_T) E(\Delta K) \quad (2.55)$$

Since the mean values of the differences ΔV_T and ΔK are zero, and since the quantities $\Delta\mu$, ΔC , ΔQ , ΔW , ΔL are all independent, equation (2.55) may be simplified as

$$\begin{aligned} \mu_{11} &= \bar{K} \cdot \frac{\bar{Q}}{\bar{C}} E\left(-\frac{\Delta C^2}{\bar{C}^2}\right) \\ &= -\frac{\bar{Q}}{\bar{C}} \cdot \bar{K} \cdot \frac{\sigma_C^2}{\bar{C}^2} \end{aligned}$$

Then, equation (2.54) reduces to

$$r = - \frac{(\bar{Q}/\bar{C}) \sigma_C^2 / \bar{C}^2}{\sigma_{VT} \cdot \frac{\sigma_K}{\bar{K}}} \quad (2.56)$$

Using equations (2.25) and (2.45) in (2.56)

$$r = \frac{-\frac{\sigma_C^2}{\bar{C}^2} \cdot \frac{\bar{Q}}{\bar{C}}}{\left[\frac{1}{\bar{C}^2} (\sigma_{QB}^2 + \sigma_{Qf}^2 + q^2 \sigma_{DI}^2) + \frac{\sigma_C^2}{\bar{C}^2} \left(\frac{\bar{Q}_B^2}{\bar{C}^2} + \frac{\bar{Q}_f^2}{\bar{C}^2} + \frac{q^2 \bar{Q}_I^2}{\bar{C}^2} \right) \right]^{1/2}} \times \frac{1}{\left[\frac{\sigma_\mu^2}{\bar{\mu}^2} + \frac{\sigma_C^2}{\bar{C}^2} + \frac{\sigma_L^2}{\bar{L}^2} + \frac{\sigma_W^2}{\bar{W}^2} \right]^{1/2}} \quad (2.57)$$

If the capacitance variation were to be the only source of mismatch, and for N-channel devices $\bar{Q}/\bar{C} = \bar{Q}_B/\bar{C}$ and since Q_f is an order of magnitude smaller than Q_B , the correlation coefficient would be -1. But as there are other variations also, we can expect the correlation coefficient to be between -1 and 0.

Now we can calculate r and compare it with measured values. Q_B/C may be estimated as shown below. For N-channel devices,

$$V_T = \phi_{ms} + 2\phi_B + \frac{Q_B}{C}$$

$$0.7 = -1.1 + 2 \times 0.37 + Q_B/C$$

$$\text{Thus, } Q_B/C = 1.06 \text{ volts}$$

$$A_{ox} = 6.4631 \times 10^{-14} \text{ cm}^2$$

Using equation (2.56) we can compute r . The computed values are compared with experimental ones in Table 2.1 which shows a good

TABLE 2.1

Comparison of Theoretical and Experimental Values
Of Correlation Coefficient for N-Channel Transistors

W/L	Theoretical r	Experimental r (Range)
6/3	- 0.16	- 0.01 to - 0.14
12/3	- 0.12	+ 0.03 to - 0.11
12/6	- 0.22	+ 0.08 to - 0.31
24/6	- 0.15	+ 0.08 to - 0.05
24/12	- 0.10	- 0.15 to - 0.42
48/12	- 0.10	- 0.02 to - 0.29

agreement. Also it is seen that the correlation coefficient is a very small value indicating that the gate oxide capacitance variation has very little effect on the mismatch in N-channel devices. We have noticed this fact in the threshold mismatch and conductance mismatch measurements.

For the case of P-channel devices, the threshold adjust implant should also be considered in evaluating Q . It is done as follows:

$$V_T = \phi_{ms} + 2\phi_B - \frac{Q_B}{C} + \frac{q D_I}{C}$$

$$- 0.8 = -0.4 - 2 \times 0.35 - \frac{Q_B}{C} + \frac{q D_I}{C}$$

$$\frac{Q_B}{C} - \frac{q D_I}{C} = 0.3 \text{ Volts}$$

From threshold voltage mismatch measurements

$$A_{ox} = 3.0369 \times 10^{-12} \text{ cm}^2$$

Again, using equation (2.56), the correlation coefficient is calculated for devices with different sizes and compared with measured values, in Table 2.2.

The agreement between the theoretical value and the measurements is not good for the P-channel devices. Also there is a large spread in the measured values over different wafers. However, the effect of this disagreement on the estimation of the mismatch in drain current in saturation is about 10% in the worst case.

TABLE 2.2

Comparison of Theoretical and Experimental Values
Of Correlation Coefficient for P-Channel Transistors

W/L	Theoretical r	Experimental r (Range)
6/3	- 0.10	+ 0.06 to - 0.28
12/3	- 0.09	+ 0.12 to - 0.29
12/6	- 0.11	- 0.16 to - 0.3
24/6	- 0.10	- 0.06 to - 0.36
24/12	- 0.12	- 0.23 to - 0.5
48/12	- 0.07	- 0.27 to - 0.74

2.4.4 Mismatch in Saturation Region

The drain current mismatch in saturation may be estimated using the measured values of the standard deviations of the threshold voltage and conductance constant, and the correlation between them. As shown in equation (2.6) the standard deviation of the drain current may be written as

$$\frac{\sigma_I^2}{\bar{I}^2} = \frac{\sigma_K^2}{\bar{K}^2} + 4 \cdot \frac{\sigma_{VT}^2}{(\bar{V}_{GS} - \bar{V}_T)^2} - 4r \frac{\sigma_K}{\bar{K}} \cdot \frac{\sigma_{VT}}{\bar{V}_{GS} - \bar{V}_T} \quad (2.57)$$

At low values of gate to source voltage the dominant factor causing the mismatch in the drain current is due to the threshold voltage. For higher bias voltages approaching the mid-rail levels, the conductance constant and threshold voltage mismatches have almost equal contributions in the drain current mismatch. For N-channel devices, the last term in equation (2.57) has negligible contribution due to the low values of the correlation coefficient. Now we may actually measure the mismatch in drain current at different gate biases and compare it with the estimated values. This would serve as a verification to the characterization methodology we have adopted and also to the measurements. The results of such an exercise is shown in Figure 2.7(a) for N-channel devices. The standard deviation of the drain current is plotted against the square root of the reciprocal of the channel area for two different gate bias voltages. The estimated value determined from the measured average values of

σ_{VT} , σ_K and r is indicated by cross marks on the same graph. As seen the estimated values lie within the measured spread in most cases validating our theory and measurement.

A similar exercise is carried out for P-channel devices also. The results are shown in Figure 2.7(b). Due to the larger mismatch observed in both the threshold voltage and the conductance constant, we may expect the standard deviation in the drain current to be larger than in the case of N-channel devices. The estimated values again indicated by crosses are calculated using the theoretically estimated values of the correlation coefficient, because the measured correlation coefficients have a wide spread over different wafers. However, it was observed that the difference in the estimated value of the drain current mismatch obtained using the theoretically estimated correlation coefficient and the measured values of the correlation coefficient is about 10%. Even in the presence of this discrepancy, the agreement between the estimated value and the measured values is good indicating the characterization approach is valid for the P-channel devices also.

2.5 Range of Applicability

In the foregoing analysis we have assumed that the dimensional variations are accounted entirely by the mismatch in the conductance constant and has no influence on the threshold voltage. For small geometry devices, however, it is not true. Reduction in channel length reduces the threshold voltage. An increase in channel width has a similar effect. In this section, we will estimate the component of the mismatch in threshold voltage brought

Fig.2.7(a) Drain current mismatch versus dimension for N-channel transistors

Fig.2.7(b) Drain current mismatch versus dimension

about by the mismatch in dimensions and show that it is negligible as compared with the total threshold voltage mismatch.

The estimated standard deviation of the device dimensions in our process is in the range 0.01 to 0.03 μm . For this order of dimensional variation, the shift in the threshold voltage for a device with drawn dimensions 3 μm x 3 μm will be first estimated. In the worst case, the effective dimensions of such a device will be 2 μm x 2 μm . First we determine the shift in threshold voltage between two devices with gate lengths 2 μm and 2.03 μm using the trapezoidal approximation [13, pp.475]. The shift in the threshold voltage for two devices with lengths L_1 and L_2 is given by

$$\Delta V_T = \frac{q N_A W_d r_j}{2 C} \left[\left(\left(1 + \frac{2y_s}{r_j} \right)^{1/2} - 1 \right) + \left(\left(1 + \frac{2y_d}{r_j} \right)^{1/2} - 1 \right) \right] \left(\frac{1}{L_1} - \frac{1}{L_2} \right) \quad (2.58)$$

where

$$y_s = \left(\frac{2\epsilon_s}{qN_A} (V_{bi} - 2\phi_B) \right)^{1/2}$$

$$y_d = \left(\frac{2\epsilon_s}{qN_A} (V_{bi} - 2\phi_B + V_{DS}) \right)^{1/2}$$

$$W_d = \left(\frac{2\epsilon_s}{qN_A} (2\phi_B + V_{BS}) \right)^{1/2}$$

The horizontal depletion layer widths y_s and y_d are smaller than the vertical depletion width W_d . In equation (2.58) r_j is the junction depth. For $L_1 = 2\mu\text{m}$ and $L_2 = 2.03\mu\text{m}$ and for our process parameters, $\Delta V_T \approx 0.7\text{ mV}$ which is a negligible quantity as compared to the total

threshold voltage mismatch. Therefore, the mismatch in length is not reflected as a mismatch in threshold voltage.

Next, we consider the effect of the mismatch in width on the threshold voltage. This shift is related to the depletion region spreading laterally in the substrate along the channel width [13, pp.475]. The threshold voltage difference for two devices with widths W_1 and W_2 is given by,

$$\Delta V_T = \frac{q N_A W_d}{C} \cdot \frac{\pi}{2} \cdot W_d \left(\frac{1}{W_1} - \frac{1}{W_2} \right) \quad (2.59)$$

For this process, and for $W_1 = 2 \mu\text{m}$ and $W_2 = 2.03 \mu\text{m}$, ΔV_T is 2.6 mV which is equivalent to a relative shift of 0.34%. The estimated mismatch in threshold voltage for a $2 \mu\text{m} \times 2 \mu\text{m}$ device obtained by extrapolating the curve for N-channel devices in Figure 2.5 is about 3%. Therefore, the contribution to the threshold mismatch due to dimensional variations is about an order of magnitude smaller than that due to charge and oxide variations. Thus, we may attribute the dimensional variations entirely to the mismatch in the conductance constant and not to threshold voltage down to the feature size of the technology.

We also verified this fact to even smaller device dimensions. The process parameters used for the calculations are the ones given in [21] where a generalized guide for MOSFET miniaturization is provided. Even for a $1 \mu\text{m} \times 1 \mu\text{m}$ device the mismatch in dimensions does not affect the mismatch in threshold voltage; despite the assumption that

the dimensional mismatch in a 1 μm process is the same as that in a 3 μm process. In fact, we may expect a better dimensional mismatch in a smaller geometry process due to better process control. Therefore, we may conclude that the characterization methodology we have described here is valid over a wide range of device dimensions.

2.6 Tracking With Temperature

The threshold voltage as well as the conductance constant vary with temperature. Let us now see how the matching in these quantities are affected by temperature.

The threshold voltage is given by

$$V_T = \phi_{ms} + 2\phi_B + \frac{\sqrt{2\epsilon_s q N_A (2\phi_B + V_{SB})}}{C} - \frac{Q_f}{C}$$

The only terms that are dependent on temperature are ϕ_{ms} and ϕ_B . We have seen that the contribution of these terms to the threshold voltage mismatch is negligible. Therefore, we may expect the matching behavior of the threshold voltage to be almost independent of temperature.

The only factor through which the conductance constant can be affected is the temperature dependence of mobility. However, we have seen that the mismatch in conductance constant is largely due to the photolithographic edge variations, and mobility variations have the least effect. Thus the temperature variations have very little effect on the conductance constant. Since the mismatch in drain current is

due to the mismatches in the threshold voltage and conductance constant, we can expect the current mismatch to be unaffected over a wide temperature range.

2.7 Summary

We have discussed the matching characteristics of P-channel and N-channel transistors. The factors causing mismatch are discussed in detail and analytical models are developed to relate the random mismatch to the device dimensions. Extensive measurements are carried out. We have shown that it is possible to accurately determine the mismatch in drain current over a wide operating range by simply measuring the mismatches in the threshold voltage and the conductance constant. It is also shown that the characterization is valid for small geometry devices, and that temperature variations do not affect matching significantly.

The P-channel devices have a poorer matching than N-channel ones. This is due to the threshold adjust implant used for the P-channel transistors. Due to the differential doping occurring at the surface, there is a larger threshold voltage mismatch. Further, there is a degradation of the gate oxide capacitance matching and possibly the hole mobility matching. This results in a larger mismatch in conductance constant also.

CHAPTER 3

RESISTOR MATCHING STUDIES

Resistors are also used as precision components in analog circuits. In this chapter we study the matching characteristics of resistors available in a p-well CMOS process with particular reference to Northern Telecom's 3 μm technology. First we provide a brief discussion of the resistor types and then discuss their matching.

3.1 Resistor Types

Resistors can be implemented in several different ways in a CMOS process. The p-well region in a bulk CMOS process inherently has a high sheet resistivity in the range of 1000 to 5000 ohms per square. Unfortunately, such lightly doped diffusions suffer from a large voltage coefficient of resistance which make their use in a circuit difficult.

The source-drain implantations in a VLSI process, in general, provide higher values of sheet resistance than a conventional process due to the shallower junctions. Typical source-drain areas exhibit a sheet resistance in the range 50 to 100 ohms per square. Hence they may be used to realize moderately high values of resistance.

Polycrystalline silicon resistors, available in a silicon gate process, are totally isolated from the rest of the circuit by thick field oxide. Therefore parasitic capacitance is very low and, if desired, the resistor may operate outside the range of the power supplies. There is no problem of voltage coefficient of resistance. Also poly resistors could be packed to a higher density because the minimum pitch for poly is usually smaller than that for diffusion. Finally resistor trimming is possible at wafer stage by electrically blowing narrow fuse links formed in the poly. It is not suitable for realizing moderately high valued resistors, however, since the gate poly has a low sheet resistivity.

Thin film resistors deposited on the wafer via additional fabrication steps are perhaps the most obvious way to obtain resistors with characteristics relatively independent of standard MOS processes. This has been done successfully for high resolution A/D and D/A converters using deposited silicon-chromium resistors. Laser trimming of resistor values is used for highest accuracy. The costs of additional processing steps and, laser trimming are quite significant. They can be justified only if the desired specifications cannot be met by simpler methods.

Northern Telecom's 3 μ m process, has an additional p-type implant intended to form the bottom plate of poly to diffusion capacitors. This has a moderately high sheet resistance of the order

of 150 ohms per square. The cross sectional view of a capacitor implant resistor is shown in Figure 3.1. The implanted region is covered by the poly silicon gate to prevent the source-drain diffusions from getting in. This results in an additional parasitic capacitance which is a drawback in certain applications such as output load resistors.

3.2 Mismatch in Resistance

The resistance of a bar of uniform material of length L and area of cross-section A is given by

$$R = \rho L/A \quad (3.1)$$

where the resistivity ρ is the reciprocal of the conductivity given by

$$\sigma = \frac{1}{\rho} = q\mu n \quad (3.2)$$

In equation (3.2) q is the electronic charge, n is the number of carriers per unit volume and μ is the mobility of the carriers. For the case of monolithic resistors which have a non-uniform charge concentration, the conductance may be evaluated as follows:

Fig.3.1 Cross section of capacitor implant resistor

Fig.3.2 Monolithic resistor

The monolithic resistor is shown in Figure 3.2. W is the width of the resistor and x_j is the junction depth. The total conductance G is calculated as a sum of the conductances of a number of infinitesimally thin layers of uniform charge concentration,

$$G = \frac{W}{L} \int_0^{x_j} q \mu n(x) \cdot dx \quad (3.3)$$

The distribution of the carriers $n(x)$ is approximately the same as the distribution of the dopant atoms. The mobility μ is a function of the carrier concentration, and, therefore, a function of the distance from the surface. Thus, the evaluation of the integral in equation (3.3) is in general very complicated. Hence, to derive an exact expression for the standard deviation in G is very difficult. However, a feel for the causes of the variation in G can be obtained by approximating the diffused layer to have effective uniform dopant concentration N_A and mobility μ_A over a depth x_j . Then

$$G = q N_A \mu_A \frac{W}{L} x_j \quad (3.4)$$

or

$$R = \frac{1}{q N_A \mu_A} \frac{L}{W x_j} \quad (3.5)$$

The standard deviation in R may be expressed in terms of the standard deviations in L , W , x_j , N_A and μ_A . The treatment of L and W is similar to that in the case of MOS transistors. L and W are independent random variables and let their variances be σ_L^2 and σ_W^2 . The effective dopant concentration N_A may be approximated to have a Poisson distribution for the same reasons as given for MOS transistors.

Then

$$\sigma_{NA} = \left(\frac{\bar{N}_A}{L W \bar{x}_j} \right)^{1/2} \quad (3.6)$$

In order to determine the standard deviation of the mobility, it is necessary to know the factors that affect mobility. Mobility is dependent on the frequency of collision of the charged carriers due to various scattering mechanisms. Impurity scattering and lattice scattering are the two most important scattering mechanisms [12, pp.108]. Impurity scattering is due to the fact that when a carrier travels past a fixed charge particle, e.g. an ionized acceptor or donor, its path will be deflected by the charge on that fixed particle. The mobility dominated by impurity scattering is shown to be proportional to $T^{3/2}/C_T$, where T is the absolute temperature and C_T is the total concentration of ionized impurities present in the

crystal, i.e. the sum of the concentrations of negatively and positively charged ions. Lattice scattering, on the other hand, is due to the thermal vibration of the atoms of the crystal lattice which disrupts the periodicity of the lattice and thereby impedes the motion of electrons. It has been shown that the mobility dominated by lattice scattering is proportional $1/T^{3/2}$. Therefore, at a given temperature, the mobility may be written as

$$\mu_A \propto \frac{1}{C_T} \quad (3.7)$$

Hence

$$\frac{\sigma_{\mu_A}^2}{\bar{\mu}_A^2} = \frac{\sigma_{C_T}^2}{\bar{C}_T^2} \quad (3.8)$$

Again, we may assume C_T to have Poisson distribution. Then

$$\sigma_{C_T} = \left[\frac{\bar{C}_T}{L \bar{W} \bar{x}_j} \right]^{1/2} \quad (3.9)$$

Thus,

$$\frac{\sigma_{\mu}^2}{\bar{\mu}^2} = \frac{1}{\bar{C}_T [L \bar{W} \bar{x}_j]} \quad (3.10)$$

Now, we shall calculate the standard deviation in resistance.

In equation (3.5) L, W, x_j are all independent random variables.

Also they are independent of the doping. However, μ_A and N_A are not independent due to the dependence of μ_A on the total dopant concentration. Therefore,

$$\frac{\sigma_R^2}{R^2} = \frac{\sigma_L^2}{L^2} + \frac{\sigma_W^2}{W^2} + \frac{\sigma_{xj}^2}{\bar{x}_j^2} + \frac{\sigma_{NA}^2}{N_A^2} + \frac{\sigma_A^2}{A^2} + 2 A_1 \frac{\sigma_{NA}}{N_A} \cdot \frac{\sigma_{\mu A}}{\mu A} \quad (3.11)$$

where the last term is due to the dependence of μ_A on N_A . A_1 is a constant in the range -1 to 0. Substituting equations (3.6) and (3.10) in (3.11) we have

$$\frac{\sigma_R^2}{R^2} = \frac{\sigma_L^2}{L^2} + \frac{\sigma_W^2}{W^2} + \frac{\sigma_{xj}^2}{\bar{x}_j^2} + \frac{1}{N_A L W \bar{x}_j} + \frac{1}{C_T L W \bar{x}_j} + \frac{2A_1}{\sqrt{N_A C_T} L W \bar{x}_j} \quad (3.12)$$

We may assume σ_L and σ_W to be of the same order. Further, resistors are usually designed such that $L \gg W$, then equation (3.12) may be written as

$$\frac{\sigma_R^2}{R^2} = \frac{\sigma_W^2}{W^2} + \frac{\sigma_{xj}^2}{\bar{x}_j^2} + \frac{1}{L W \bar{x}_j} \left[\frac{1}{N_A} + \frac{1}{C_T} + \frac{2A_1}{\sqrt{N_A C_T}} \right] \quad (3.13)$$

The structure used to measure the mismatch in resistors is similar to the transistors shown in Figure 2.1. Such test structures were designed for poly, p^+ source-drain, N^+ source-drain, and P-capacitor implant resistors. In every case resistors patterns with widths 6 μm , 9 μm and 12 μm was laid out. Characterization

results show that except in the case of p-capacitor implant resistors, there is a large systematic component of mismatch. A possible explanation for this is metal thinning when it passes over a device well.

We have plotted the standard deviation of the p-type capacitor implant resistors as a function of resistor width in Figure 3.3. The dots in the measurement spreads correspond to the average value over four wafers. Even the 6 μm wide resistors match better than 0.15 percent. By extrapolating these results it is expected that 3 μm wide resistors will have a matching better than 0.25%. Hence large valued resistors can be realized with very good matching accuracy in a small area.

3.3 Summary

In this chapter we have briefly discussed the various types of resistors available in a CMOS process. An analytical expression for the mismatch in resistors is developed. Experimental data about the capacitor implant resistors are presented. This is the type of resistor used in the D/A converter discussed in Chapter 4.

Fig.3.3 Capacitor implant resistor mismatch versus dimension

CHAPTER 4

HIGH-SPEED D/A CONVERTER

We begin this chapter with a brief introduction to current-steering DAC configurations. The disadvantages of most of the configurations is summarized. Then a multiple current source conversion approach is discussed which overcomes most of these disadvantages. A modification to the approach is also presented which is suitable for realizing high resolution converters. Then we discuss error analysis to relate the circuit yield of the DAC to the basic component matching accuracy of the conversion elements. Finally, the design of the different circuit blocks of the DAC are presented, followed by a discussion of the complete video bandwidth DAC.

4.1 Current Steering Configurations

In a current steering D/A converter, for each input bit, a current source proportional to a reference signal and the weight of the code bit is established. The weighted current may be achieved by the current source itself or by the current source in combination with an output weighting network. Each code bit acts to steer its related current source into either the output directly (or via the aforementioned weighting network), or into an alternate output such as ground. The current source is established independently of the switching. Currents are

always drawn from the source whether or not that current is steered to the output, i.e. the current drain from the current source is constant and independent of the code. Thus, components which establish the current source are always in a steady-state condition, and therefore switching and thermal transients effects are minimized. The output of the D/A converter may be either the sum of the weighted currents themselves, available for direct use over some range of output compliance voltage determined by the particular circuitry, or may be applied typically to an output circuit such as null point inverting operational amplifier.

MOS current steering D/A converters are rarely reported in the literature [22] - [24]. In [22] the current division is achieved by a thin film R-2R network. Both [23] and [24] use resistorless active current sources. In bipolar technology, though, the current steering technique is a popular one and there have been several configurations reported. A number of bipolar designs [25] - [33] were examined to assess the suitability of the configurations for implementation in a modern MOS VLSI process in general, and the Northern Telecom Electronics CMOS3 in particular. The CMOS3 process has a feature size of 3 μm . The devices exhibit low threshold voltage, high gain and high speed. The process has provision for a moderately high-sheet-resistivity implant. Therefore, it is possible to use resistors in combination with MOS transistors to design the current sources required in a D/A converter.

Most of the configurations studied suffer from one or more of the following disadvantages:

1. Binary weighting of the currents requires the gate to source voltage of all the current source devices to be the same. With long and wide devices, this can be achieved by binary scaling of the width to length (W/L) ratio of the devices. Unfortunately, for small geometry devices, the current does not linearly scale with the aspect ratio. The drain current of an array of transistors studied here having constant length and differing widths is shown in Figure 4.1. The departure from the ideal linear plot is characteristic of many small geometry MOS processes. There are two reasons for this departure. First, there is a spreading of the depletion region beyond the edges of the channel along the width. This amounts to an increase in threshold voltage which becomes more pronounced as the channel width decreases [13, pp.475]. Secondly, the effective channel width is less than the drawn channel width due to the encroachment of the channel stop implant and the field oxide region into the channel region, in a LOCOS process. These effects are shown in Figure 4.2 and lead to reduction in device gain, increase in body effect, and increase in threshold voltage [34]. The net result is that as the channel width decreases, the drain current decreases faster than the ideal case. In principle, one could trim this variation out by careful design over some iterations of the circuit. However, this is a time consuming process and is not suited to modern VLSI design methods. Thus, it is not practical to scale the aspect ratio of small geometry devices to achieve precise current ratios.

Fig.4.1 Drain current versus gate width of narrow channel transistors.

Fig.4.2 Channel narrowing effects

2. The voltage coefficient of monolithic resistors, caused by the depletion layer penetration into the diffused resistor due to the reverse bias voltage on the p-n junction, introduce errors. A binary weighted voltage drop is generated across the shunt resistors of a R-2R ladder network. This introduces nonlinearity in the D/A converter output.

3. Requires large power supply voltage.

4. Slow conversion speed.

5. High output glitch energy. Glitch is an undesirable output occurring during transitions of the digital inputs. The worst case occurs when there is a major carry transition, i.e. when the input word is changing between 011...1 and 100...0. Data skew and unequal switching times of the various bits can cause a momentary erroneous code of 000...0 or 111...1, resulting in a large output spike. Low glitch is an important requirement for many high speed applications.

6. Complex circuitry and layout.

4.2 Multiple Current Source Approach

A conversion approach which does not suffer from the above disadvantages is the multiple current source approach. A conceptual 3-bit D/A converter following such an approach is shown in Figure 4.3 to illustrate the technique. Each MOS transistor has the same size and has the same value of resistance in the source so that they carry equal currents. The LSB controls one "unit" current source, bit 2 controls two unit current sources, and the MSB controls four unit

Fig.4.3 Simple multiple current source approach

current sources. In general, for an n -bit converter, the MSB has to control 2^{n-1} unit current sources and, the total number of such current sources in the entire converter will be $2^n - 1$. Each of these current sources will have its own current steering bit cell to maximize speed and minimize glitch. A method to overcome this exponential growth in circuit size is discussed in Section 4.2.1.

The multiple current source approach brings with it several advantages. Apart from eliminating the problem of voltage coefficient of resistance and the need to scale transistor geometries, the scheme offers an improvement in accuracy. If the unit current source has a mean value I , and standard deviation σ , connecting ' n ' such sources in parallel would produce an equivalent current source with a mean value nI and standard deviation $\sqrt{n}\sigma$. If we are using one unit current source for the LSB and n of them in parallel for the MSB, then the accuracy of the LSB is $\sigma/I \times 100\%$ while that of the MSB is $(\sqrt{n}\sigma/nI) \times 100\%$. Therefore, there is an improvement in accuracy in the MSB by a factor $\sqrt{n}$. Thus, the bit current sources will have a tapered error distribution with the accuracy improving towards the MSB. It will be shown in Section 4.3 (Error Analysis) that the MSB current source is the most critical, then comes the bit 2, and so on. Thus, overall accuracy improvement is possible in this approach.

It is possible to achieve an improvement in speed also. This is due to the fact that all the bits are switching equal current values; only the number of such parallel sources vary depending on the significance

of the bit. On the other hand, if we had scaled currents obtained by some other means, the LSB would take a longer time to charge up the parasitic nodal capacitance and settle down, than the MSB. Of course, the speed also depends on the design of the bit cell, whose function is to steer the current to the analog output node or to ground. If the currents in all the bit cells are identical as in this approach, then the bit cell can be designed for optimum performance.

By the same argument given above, a multiple current source DAC should produce less glitch energy at the output as compared to conventional techniques.

Another advantage of this approach is that the complete DAC can be constructed by using identical building blocks. Therefore, the layout should be relatively easy. But this can also turn out to be a disadvantage when the number of such building blocks increases. For example, for an 8-bit DAC the MSB has to have 128 such blocks in parallel, the layout could get quite involved. This would also take up a lot of area. Now we will present a technique which solves this problem.

4.2.1 Interstage Divider Multiple Current Source DAC

This configuration solves the problem of large ratio of current sources in the MSB to that in the LSB by using an interstage attenuator. As shown in Figure 4.4 it consists of two four bit DACs with a divide by 16 network interconnected between the two sections. Both sections

contain eight, four, two and one unit current sources. The interstage current divider, as shown in Figure 4.5, is a simple resistive network which routes 15/16 of the least significant bit current to ground and 1/16 to the output bus. The division ratio is independent of the load resistance as shown below [32].

With the notation of Figure 4.5, we can write three simultaneous equations:

$$I_{OUT} R_L = I_1 R + I_2 nR \quad (4.1)$$

$$I_1 = I_2 + I_L \quad (4.2)$$

and $I_2 = I_M - I_{OUT} \quad (4.3)$

from (4.2) and (4.3)

$$I_1 = I_M + I_L - I_{OUT} \quad (4.4)$$

Substituting in (4.1)

$$I_{OUT} R_L = (I_M + I_L - I_{OUT}) R + (I_M - I_{OUT}) nR \quad (4.5)$$

Therefore,

$$I_{OUT} = [I_M (n+1) + I_L] \frac{R}{[R_L + (n+1)R]} \quad (4.6)$$

Equation (4.6) indicates that the contribution to I_{OUT} of I_M and I_L is in the relation $(n+1):1$ and is independent of R and R_L . Hence, if the interstage divider has a ratio of 15:1, I_L contributes proportionately 1/16 of I_{OUT} , while I_M contributes 15/16. The

Fig.4.4 Interstage divider multiple current source DAC

Fig.4.5 Divide by 16 network

2

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1910a
(ANSI and ISO TEST CHART No. 2)

relationship holds for any R and R_L with the limitation that $R_L I_{OUT}$ must not exceed the voltage compliance limits.

The attenuation provided by the resistor network is independent of process and temperature variations. As will be shown in Section 4.3, the resistor ratio need only be accurate to within 3.2%, if this were the only source of error. We have seen that 6 μm wide capacitor implant resistors have a standard deviation of matching better than 0.2 percent. Therefore, the interstage current divider will not be a cause for device failures.

A slight drawback of this configuration as compared to the straight multiple current source configuration is that the supply current is about twice the full scale current of the converter. But the saving in area is considerable bringing down the number of unit currents sources and the associated bit cells from 255 to only 30.

4.3 Error Analysis

The circuit yield of a DAC is a strong function of the errors associated with it. The criterion for a device to be considered good is that the total error should be less than a pre-assigned value. In general, the errors generated by a DAC consist of linearity, offset and gain errors. Usually, a DAC may be calibrated for zero gain and offset errors. Linearity errors, however, occur due to the random

mismatch in the conversion elements. In this section we relate the nonlinearity to the basic component matching accuracy. First we discuss a few general considerations and then provide a mathematical analysis.

4.3.1 General Considerations

The accuracy of an A/D or D/A conversion system is a function of the algebraic summation of the actual errors of the individual elements of the system. In addition to the errors in conversion caused by the component imperfections, there is the error due to the finite resolution of the word, known as the quantization error. The quantization error is normally not considered as part of the D/A conversion system error, because the DAC can only convert to the resolution of the input digital word.

In making an error allocation or analysis one of two general methods can be used. Given the overall system requirements, the errors can be allocated to the various elements so that the algebraic sum of their maximum possible errors in the same direction is equal to or less than the allowed overall errors. Such an approach is too pessimistic since the possibility of this worst case ever occurring is extremely remote. A more reasonable procedure is a statistical error allocation and computation. This is a less costly design approach even though it produces a small number of rejects in testing. This is the approach taken in this chapter.

In our analysis, we assume that the output of the DAC is a current. A voltage output may be obtained by passing this current through a precision low valued resistance or by using an op-amp.

Error can be computed either as a percentage of the full-scale analog output current or as a percentage of the particular output current (reading) that should be present [35]. When the value of the digital word is small, the error based on the percentage of reading will reach a maximum value. Consider, for example a n-bit DAC which has all its bit currents equal to their ideal values except for the LSB which is off, say, by 30%. When only the LSB is ON, and all the other bits are OFF, the error when referred to the actual output is 30%. Actually, the error is within 1/2 LSB, which is acceptable in most cases. Therefore, referring the error to the actual output could be quite misleading.

In most cases, the error is computed on the basis of full scale analog output. It is generally expressed in terms of fractions of LSB or as a percentage. This is the definition we have followed in our analysis.

4.3.2 Mathematical Analysis

Let x be the output of the DAC for a given input word and y be the analog complement of the output. The full scale output of the converter is the sum $x+y$. To express the error as a fraction of the

full scale, first we normalize the output to the full scale as follows

$$z(x,y) = \frac{x}{x+y} \quad (4.7)$$

z is the normalized output, and x and y are dependent on the input digital word and the DAC configuration. For the 8-bit interstage divider DAC shown in Figure 4.4,

$$x = [8D_1 + 4D_2 + 2D_3 + D_4 + \frac{1}{16} (8D_5 + 4D_6 + 2D_7 + D_8)] * I_{\text{unit}} \quad (4.8)$$

and

$$y = [8(1-D_1) + 4(1-D_2) + 2(1-D_3) + (1-D_4) + \frac{1}{16} \{8(1-D_5) + 4(1-D_6) + 2(1-D_7) + (1-D_8)\}] * I_{\text{unit}} \quad (4.9)$$

where $D_i = 0$ or 1 $i = 1, 2, \dots, 8$

D_1 is the MSB

D_8 is the LSB

and I_{unit} is the unit current source with a mean value $\bar{I}$ and standard deviation of matching σ . As the unit current sources are random and uncorrelated, we may treat x and y to be independent random variables with standard deviations σ_x and σ_y respectively. Now we may determine the standard deviation of z in terms of σ_x and σ_y using equation (A.13)

$$\sigma_z^2 = \frac{\bar{y}^2 \sigma_x^2 + \bar{x}^2 \sigma_y^2}{(\bar{x} + \bar{y})^4} \quad (4.10)$$

σ_x and σ_y are evaluated as follows

$$\bar{x} = [8D_1 + 4D_2 + 2D_3 + D_4 + \frac{1}{16} (8D_5 + 4D_6 + 2D_7 + D_8)] \bar{I}$$

and

$$\begin{aligned} \sigma_x^2 &= [8D_1 + 4D_2 + 2D_3 + D_4 + \frac{1}{16} (8D_5 + 4D_6 + 2D_7 + D_8)] \sigma^2 \\ &= \frac{\bar{x}}{\bar{I}} \cdot \sigma^2 \end{aligned} \quad (4.11)$$

Similarly,

$$\sigma_y^2 = \frac{\bar{y}}{\bar{I}} \cdot \sigma^2 \quad (4.12)$$

Substituting equations (4.11) and (4.12) into equation (4.10)

$$\sigma_z^2 = \frac{1}{\bar{I}} \cdot \frac{\bar{x} \cdot \bar{y}}{(\bar{x} + \bar{y})^3} \cdot \sigma^2 \quad (4.13)$$

The expected value of z may be determined using equation (A.12) as

$$\bar{z} = \frac{\bar{x}}{\bar{x} + \bar{y}} + \frac{1}{2} \left[\frac{-2 \bar{y} \sigma_x^2}{(\bar{x} + \bar{y})^3} + \frac{2 \bar{x} \sigma_y^2}{(\bar{x} + \bar{y})^3} \right] \quad (4.14)$$

Substituting equations (4.11) and (4.12) into (4.14), the term in the brackets vanishes and

$$\bar{z} = \frac{\bar{x}}{\bar{x} + \bar{y}} \quad (4.15)$$

Using equation (4.15) in (4.13), we have

$$\sigma_z^2 = \frac{\bar{z} (1 - \bar{z})}{\bar{I} (\bar{x} + \bar{y})} \cdot \sigma^2 \quad (4.16)$$

The above result may also be derived by determining the joint probability density function of z as shown in [36]. Equation (4.16) expresses the

variances of the D/A outputs for different digital words. The function $\bar{z}(1-\bar{z})$ will have a maximum value when $z = 1/2$, i.e. when the output is half way through the full scale, and falls off towards zero for minimum and maximum input word combinations. This observation suggests that the MSB current could be the most critical and should have the highest accuracy and error contributions of the bits taper off towards the LSB. In order to prove this, we assume the contrary. In other words, ignore for the moment that the bit currents are realized using the multiple current source approach and assume that the relative errors of all the bit current sources are identical, i.e. if ϵ is the error in current source I , $\epsilon/2$ is the error in $I/2$, and so on. Consider an n -bit converter with bits current, $I, I/2, \dots, I/2^{n-1}$. Let the input word be $100\dots 0$ for which the maximum error referred to the full scale occurs. Now, let the current in the third M.S.B., $I/4$, be high in value by an amount equal to $\epsilon/4$. Bit 1, bit 2, and all other bit currents are assumed to be in exact ratios. The full scale current is

$$I_{FS} = I + \frac{I}{2} + \frac{I}{4} + \frac{\epsilon}{4} + \frac{I}{8} + \frac{I}{16} + \dots = 2I + \epsilon/4 \quad (4.17)$$

Therefore, the output for the word $10\dots 0$ should be

$$I_{10\dots 0} = I + \epsilon/8 \quad (4.18)$$

But the output current is $I_{OUT} = I$. Therefore, error referred to full scale = $\frac{\epsilon/8}{2I + \epsilon/4}$

$$= \frac{1}{16} \cdot \frac{\epsilon}{I} \left(1 - \frac{1}{8} \cdot \frac{\epsilon}{I}\right) \approx \frac{1}{16} \cdot \left(\frac{\epsilon}{I}\right) \quad (4.19)$$

Thus, for any error in this bit current is reduced to $1/16$ of its value when referred to the full scale analog output current. Similarly, it could be demonstrated that the error in the second MSB is attenuated by a factor of $1/8$ at the analog output and the fourth MSB will have its error reduced by $1/32$, and so on. Then if n for the MSB is 1, and for the NMSB is 2 etc., the current error attenuation factor for the n th bit is $(1/2)^{n+1}$.

This suggests that the relative error contributions of all the bit current sources need not be the same. The MSB is the most critical and the accuracy of the following bits could taper off towards the LSB. Such an error distribution is indeed a natural consequence in the multiple current source approach where the relative accuracy of the bits improve towards the MSB as we have already shown.

The analysis so far has shown that maximum error occurs half way through the full scale and the error contributions of the individual bits taper off towards the LSB. Now we proceed further to relate the circuit yield to the standard deviation of the unit current sources. We have chosen $\pm 1/2$ LSB as the maximum tolerable integral nonlinearity. The yield is then obtained by multiplying the probabilities of each of the 256 DAC outputs with less than $1/2$ LSB error. For normal distribution with variances given by equation (4.16), the yield is

$$G = \prod_{i=2}^{255} \frac{1}{\sqrt{2\pi}\sigma_z} \int_{z-\frac{1}{512}}^{z+\frac{1}{512}} \exp\left\{-\frac{(z-\bar{z})^2}{2\sigma_z^2}\right\} \cdot dz$$

$$= \prod_{i=2}^{255} \text{erf}(Q/\sqrt{2}) \quad (4.20)$$

where $1/512$ is the normalized $1/2$ LSB value and

$$Q = \frac{1}{512 \left(\frac{\bar{z}(1-\bar{z})}{15 + 15/16}\right)^{1/2}} \cdot \frac{\sigma}{I} \quad (4.21)$$

The yield is plotted as a function of (σ/I) in Figure 4.6. Because of the error function nature of the relation between yield and current source matching, there is an almost flat region close to the 100% yield level, followed by a very steep region and finally the yield asymptotically approaches zero. To avoid the possibility of any process variations from batch to batch affecting the yield adversely, the design should be such that we don't go into the region where the yield is very sensitive to the matching accuracy. On the otherhand, a very conservative matching tolerance is also not desirable because the improvement in yield is marginal with improvement in matching accuracy beyond a certain point. Therefore we chose 95% yield level as an optimum value that would not appreciably shift due to process variations from batch to batch. The standard deviation of matching of the current sources should be about 0.45% to achieve this yield level. It may be noted that for an 8-bit DAC, an integral nonlinearity of $\pm 1/2$ LSB is equivalent to 0.2% of full scale. Thus, with this configuration, it

Fig.4.6 DAC yield versus current source matching

is possible to provide good integral linearity associated with a high circuit yield without requiring an equivalent degree of component matching.

The method we have indicated above is quite general. It can be easily extended to different resolutions and configurations. The analysis provides the circuit designer a very systematic design approach for data converters. Next we discuss the accuracy requirement of the divide by 16 network.

4.3.3 Accuracy of Divide By 16 Network

To visualize the accuracy requirement of the divide by 16 network, assume that all the bit currents are in exact ratio and the divider network alone is the source of error. Instead of dividing by 16, let it divide by A . Obviously, this network will not introduce any errors when the currents in the LSB array are zero and introduce maximum error when all the LSB current sources are on. The maximum nonlinearity would be felt between the code words 00010000 and 00001111. If the value of the MSB current source is I , then

$$I_{00010000} = I/8$$

and

$$I_{00001111} = \frac{1}{A} \left(I + \frac{I}{2} + \frac{I}{4} + \frac{I}{8} \right) = \frac{15}{8A}$$

The difference between the two currents should ideally be equal to

1 LSB which is $I/128$, but in no case be more than $\pm 1/2$ LSB from the ideal value if a differential nonlinearity of $1/2$ LSB is to be achieved.

Thus

$$\frac{I}{8} - \frac{15}{8A} = \frac{I}{128} \pm \frac{I}{256}$$

which means $15.48 \leq A \leq 16.55$ and therefore the error in the division ratio can be as high as 3.2%. We have seen that the resistors match a lot better than this value and hence the divide by 16 network is not a source of yield loss.

4.4 Circuit Blocks

In this section the circuit blocks of the 8-bit interstage divider multiple current source DAC shown in Figure 4.4 are presented. The design of the divide by 16 network is already presented in Section 4.2.1. Therefore the remaining circuit blocks are the unit current sources and the current steering bit cells. We discuss these first and then the complete DAC.

4.4.1 Unit Current Source

MOS current steering DACs in the past [21] - [23] have not used the combination of MOS transistors and resistors to set up the bit current values. The inclusion of a resistance in series with the source as shown in Figure 4.7, provides a local negative feedback and improves the matching accuracy of the current source. Topographically,

Fig.4.7 Unit current source

this is analogous to connecting a resistor in series with the emitter of a bipolar transistor. Such an arrangement is used in almost all bipolar current steering DACs. The effect of this resistor, however, is not the same in MOS and bipolar cases. In the case of bipolar technology, due to the exponential relationship between the current and voltage there will be a large negative feedback. As a result the matching of the current sources is almost solely determined by the matching of the resistors. In the case of MOS transistors, owing to the square-law dependence of the current on voltage, the feedback will not be as much as it is in bipolar circuits. Further, in MOS transistors, in addition to the photolithographic variations, the threshold voltage mismatch also contributes to the drain current mismatch. An analogous situation does not exist in case of bipolar transistors. The net result of all these is that the improvement in current matching brought about by the resistor is not as much as it is in case of bipolar transistors. Yet, under certain conditions the presence of a source feedback resistor can bring about a significant improvement in matching accuracy. Also this arrangement offers an easy way of biasing the current sources and provides a good temperature tracking with the output on-chip load resistor. These aspects will be discussed in greater detail in Section 4.5. Now we will demonstrate the improvement in matching accuracy.

The current-voltage relationship of the current source shown in Figure 4.7 when the transistor is operating in the saturation region is given by

$$I = \frac{K}{2} (V_G - IR - V_T)^2 \quad (4.22)$$

It is difficult to analytically derive an expression for the standard deviation of the current in terms of the standard deviations in K and V_T from equation (4.22). However, in order to get a feel of the influence of various terms on the current matching, we evaluate the deviation in current by finding partial derivatives. We can show from equation (4.22)

$$\frac{\partial I}{I} = \frac{1}{1 + R\sqrt{2IK}} \cdot \frac{\partial K}{K}$$

$$\frac{\partial I}{I} = - \frac{\sqrt{2K/I}}{1 + R\sqrt{2IK}} \cdot \partial V_T$$

and, the total change in current may be written as

$$\frac{\Delta I}{I} = \frac{1}{1 + R\sqrt{2IK}} \left(\frac{\Delta K}{K} - 2\sqrt{K/I} \cdot \Delta V_T \right) \quad (4.23)$$

The quantity in the parentheses of equation (4.23) is the deviation in current if there is no resistance in series with the source, and designating this as $\Delta I'/I$, we have

$$\frac{\Delta I}{I} = \frac{1}{1 + R\sqrt{2IK}} \cdot \frac{\Delta I'}{I} \quad (4.24)$$

Equation (4.24) indicates that there is an improvement in the current source matching by a factor $(1 + R\sqrt{2IK})$. For a typical case when

$$I = 125 \mu A$$

$$R = 4.5 K \text{ ohm}$$

$$W = 48 \mu m$$

$$L = 6 \mu m$$

$$K = 36 W/L \mu A/V^2$$

the improvement is about 55%.

So far, we have assumed that the resistor is ideal. The effect of mismatch in resistors may be accounted for as follows:

$$\frac{\partial I}{I} = - \frac{R \sqrt{2IK}}{1 + R\sqrt{2IK}} \cdot \frac{\partial R}{R} \quad (4.25)$$

The total deviation in current is obtained by combining equations (4.24) and (4.25).

$$\frac{\Delta I}{I} = \frac{1}{1 + R\sqrt{2IK}} \left(\frac{\Delta I'}{I} - R\sqrt{2IK} \cdot \frac{\Delta R}{R} \right) \quad (4.26)$$

If σ_R is the standard deviation of the resistors and $\sigma_{I'}$ is the standard deviation of the current sources measured without including the resistors, then the net standard deviation in current I is, from equation (4.26)

$$\frac{\sigma_I^2}{I^2} = \frac{1}{(1+x)^2} \left(\frac{\sigma_{I'}^2}{I^2} + x^2 \frac{\sigma_R^2}{R^2} \right)$$

where $x = R\sqrt{2IK}$.

If $\sigma_R/R = M \cdot \sigma_I/I$, then for a typical value of $X=1.2$ as shown above, the total current variation for different values of M are listed below:

$$\text{For } M=0, \quad \frac{\sigma_I}{I} = 0.45 \frac{\sigma_I}{I} \quad (4.27a)$$

$$\text{For } M=0.5, \quad \frac{\sigma_I}{I} = 0.53 \frac{\sigma_I}{I} \quad (4.27b)$$

$$\text{For } M=1.0, \quad \frac{\sigma_I}{I} = 0.71 \frac{\sigma_I}{I} \quad (4.27c)$$

Equations (4.27) indicate that the matching of the resistors is not very critical. Even in the case when the standard deviation of matching is the same as that of the current source, as shown by equation (4.27c) there is still an improvement by about 30%. We have seen that 6 μm wide capacitor implant resistors have a mismatch of only 0.15%. Hence we can expect an improvement in current matching of about 50%. We have verified the above analysis by Monte-Carlo simulations using the measured standard deviations of V_T , K and R in equation (4.22). An experimental investigation has also been done by connecting a discrete 4.7 K ohm resistor to the source of a 24 μm wide and 6 μm long transistor. The results are summarized in Table 4.1 which shows an improvement in matching by about 50% when the resistor is present.

Now we would like to state that this method of introducing a resistor to improve matching is area intensive unless a high sheet resistance implant and/or narrow width resistors are used. The same degree of improvement could have been obtained under identical conditions

TABLE 4.1

DRAIN CURRENT MATCHING WITH AND WITHOUT A
SOURCE FEEDBACK RESISTOR

Batch No.	No. of Pairs	Standard Deviation Of Matching (%)	Value of Resistance In Source (Ohm)
1	156	0.85	0
		0.41	4.7 k
2	90	1.20	0
		0.58	4.7 k

of gate-source bias and drain current by using a device with a larger channel area. When the DAC was originally laid out complete information about the dependence of transistor matching on dimensions was not available. With the knowledge of matching generated in this thesis, a better circuit may now be designed. However, the present configuration allows to have an on-chip load resistor without introducing any errors due to temperature co-efficient of resistance. We will elaborate on this in Section 4.5.

4.4.2 Bit Cell

The function of the bit cell in a current-steering DAC is to steer the bit current either to the current summing node or to the digital ground. The design of the bit cell has a considerable impact on the final accuracy and speed of a DAC. The MOS current steering DACs reported earlier in [23] and [24] do not make reference to the bit cell design.

The bit cell should be designed in such a way that logic input swings are isolated from the output current line (switch signal feed-through). Also, the voltage swing across the current source during switching should be held as low as possible. If this voltage swing is large, the parasitic capacitance at the drain of the current source transistor would take longer to charge up. This voltage swing could also be coupled on to the common gate bus of the current sources which is driven by the control amplifier. This would require the settling

time of the amplifier to be small. Another problem with the large voltage swings across the current source is the disturbance of the steady state value of the current sources due to channel length modulation.

A differential pair may be used for current steering as shown in Figure 4.8. Devices Q_{11} and R make up the bit current source. By suitably controlling the voltages on the gates of the transistors Q_9 and Q_{10} , the bit current can be steered through one or the other transistor. If the change in voltage at the common source of Q_9 and Q_{10} is held low during switching, then a high switching speed can be achieved. If a symmetrical push-pull voltage is applied to the gates of Q_9 and Q_{10} , then there would be a large change in voltage at the drain of Q_{11} during switching due to the square law relationship of the drain current to gate voltage. This is shown below.

Let at some time when the voltage at the gate of Q_9 be 5V, and that at the gate of Q_{10} be 0V. Then,

$$V_{GS9} = \sqrt{\frac{2I}{K}} + V_{TH} \quad (4.28)$$

where V_{TH} is the threshold voltage including body effect. The voltage at the drain of Q_{11} is

$$V_{D11} = 5 - V_{GS9} \quad (4.29)$$

If the drive signal is symmetric, then at some instant of time the gate

voltages of Q_9 and Q_{10} are equal to 2.5V, and the current through each of them is $I/2$. Then the gate source voltage of Q_9 is

$$V_{GS9} = \sqrt{\frac{I}{K}} + V_{TH} \quad (4.30)$$

Here it is assumed that the body effect does not alter the threshold voltage during switching. The voltage at the drain of Q_{11} is now,

$$V_{D11} = 2.5 - V_{GS9} \quad (4.31)$$

Therefore, the change in voltage at the drain of Q_{11} is

$$\begin{aligned} V_{D11} - V_{D11}' &= 2.5 - (V_{GS9} - V_{GS9}') \\ &= 2.5 - \left(\sqrt{\frac{2I}{K}} - \sqrt{\frac{I}{K}} \right) \end{aligned} \quad (4.32)$$

The term in brackets is a small positive quantity. Therefore, there is a large swing in voltage during switching. This would even push the current source transistor Q_{11} into the triode region, thereby altering the current value by a large amount.

The bit cell design also plays an important role in determining the glitch energy of the DAC. Even when external sources of glitch such as data skew are minimized, the unequal switching times of the bit currents could result in a large glitch output. The symmetric push-pull swing applied to the gates of the differential pair discussed above also results in an asymmetric current output in the differential pair. Suppose the gate voltage Q_{10} is increasing from 0 to 5 V. Until

the gate voltage exceeds the threshold voltage of Q_{10} , there won't be any conduction. On the other hand, the current through Q_9 will start decreasing the moment the voltage on the gate of Q_9 decreases from 5 V towards 0, because of the change in value of the current source brought about by channel length modulation. Thus there is an asymmetric current output in the bit cell resulting in a large glitch.

The problems of speed and glitch can be simultaneously solved by providing an asymmetric push-pull drive to the gates of Q_9 and Q_{10} . Consider the situation when the gates of Q_9 and Q_{10} have the full power supply voltage simultaneously, for a brief interval as shown in the switching waveform of Figure 4.9. At $t = t_0$,

$$V_{GS9} = \sqrt{\frac{2I}{K}} + V_{TH} \quad (4.33)$$

$$V_{D11} = 5 - V_{GS9} \quad (4.34)$$

At $t = t_1$, the bit current is equally shared by both Q_9 and Q_{10} and hence

$$V_{GS9} = \sqrt{\frac{I}{K}} + V_{TH} \quad (4.35)$$

$$V_{D11} = 5 - V_{GS9} \quad (4.36)$$

From (4.34) and (4.36) the change in voltage during switching is given by

$$V_{D11} - V'_{D11} = - \left(\sqrt{\frac{2I}{K}} - \sqrt{\frac{I}{K}} \right) \quad (4.37)$$

which is very small. Hence the switching is fast, and also the output current change is symmetric as the current source value is not

altered during switching.

A circuit that can provide the asymmetric push-pull voltage of the type we need is shown in Figure 4.10. Transistors $Q_1 - Q_4$ form two inverters which provide the complementary drive to the cross-coupled latch comprising of devices $Q_5 - Q_8$. Even when the p-channel and n-channel devices in the latch are all electrically identical, the nature of the cross-coupling provides a faster low to high transition than a high to low transition. The dimensions of the devices have been further optimized with the help of simulations to provide a high switching speed and low glitch energy.

It is interesting to note that in steady state, the current source transistor and the ON transistor in the bit cell form a cascode connection. This enables a load to be connected directly to the drain of the bit cell transistors, as the change in voltage across the load does not affect the value of the current source. The load, however, should be such that the bit cell transistor is operating in the saturation region so that the full advantage of the cascode combination can be realized. Thus the drain of the bit cell transistor cannot be more than one threshold below the gate. By placing the differential pair in the same p-well as the current source, the threshold voltage of the bit cell devices can be increased to about 1 volt due to body effect. Hence a full scale output voltage of 1 volt can be achieved with this configuration.

Fig.4.8 Current steering bit cell

Fig.4.9 Switching waveform

Fig.4.10 Bit cell driver

4.5 Video Bandwidth DAC

Video converters generally operate in the 1 to 20 MHz range. The composite video signal has an amplitude of approximately 1 volt. This signal amplitude can be handled with a total supply voltage of 5 volts without much difficulty. However, the low signal amplitude poses other concerns such as noise, and drift due to time and temperature, which become comparable to the quantization error of the converter. These problems are not peculiar to video converters alone. They are common to present day analog circuits in VLSI. The scaling down of device dimensions accompanied by scaling the total power supply voltage brings down the signal amplitude. Therefore, the selection of the video DAC as a test vehicle also demonstrates a design approach in analog VLSI.

The schematic diagram of the video DAC is shown in Figure 4.11. The white level in the video signal is usually taken as zero and the signal increases in the negative direction towards black levels. Therefore, we have used a -5 V supply as shown. The unit current sources are biased by mirroring the current in the reference transistor Q_{REF} . Placing Q_{REF} in a closed loop with the control amplifier compensates for systematic variations in threshold voltage of the current source transistors due to time and temperature. Eight such reference current sources are interdigitized throughout the current source array. The reference current is given by

Fig.4.11 Schematic diagram of the video DAC

$$I_{REF} = \frac{V_{SS} - \frac{R_1}{R_1 + R_2} \cdot V_{REF}}{R_S} \quad (4.38)$$

where R_1 and R_2 are two external variable resistors used to set the full scale voltage of the DAC. V_{REF} could be an on-chip reference voltage. The resistors in the divide by 16 network and the load resistor are chosen such that

$$R_3 = R_L = R$$

and

$$R_4 = 15R$$

so that the output current (equation (4.6)) may be written as

$$I_{OUT} = (16 I_M + I_L) \frac{R}{R + 16R} \quad (4.39)$$

where I_M and I_L are the currents from the MSB array and LSB array respectively. The nominal values of the resistors are chosen such that a full scale voltage of 1 volt is developed across the load resistor. The resistor network is laid out such that R_3 and R_L have five unit resistors in parallel while R_4 has three unit resistors in series. This eliminates the error introduced by the contact resistance to a first order.

If the nominal LSB current is chosen to be $8 \mu A$, the full scale current is

$$I_{FS} = 2048 \cdot 8 = 2040 \mu A$$

when $I_{OUT} = I_{FS}$, let $I_M = I_L$, then from equation (4.39)

$$I = 2040 \mu A$$

and hence the unit current = $2040/15 = 136 \mu A$.

The temperature dependance of the output voltage may be determined from the equivalent circuit shown in Figure 4.12. Here I_1 represents the total reference current of all the reference sources, I_2 is the total current from the MSB array and I_3 is that from the LSB array. R_{11} , R_{12} and R_{13} are the corresponding equivalent resistances. The voltage at the inverting input of the operational amplifier is given by

$$V_1 = \frac{R_1}{R_1 + R_2} V_{REF} + V_{OS}$$

where V_{OS} is the input offset voltage of the amplifier. The reference current is therefore,

$$\begin{aligned} I_1 &= \frac{V_{SS} - \frac{R_1}{R_1 + R_2} V_{REF} + V_{OS}}{R_{11}} \\ &= \frac{V_2 + V_{OS}}{R_{11}} \end{aligned} \quad (4.40)$$

where

$$V_2 = V_{SS} - \frac{R_1}{R_1 + R_2} V_{REF} \quad (4.41)$$

$$I_1 R_{11} = I_2 R_{12} = I_3 R_{13} \quad (4.42)$$

$$\text{Thus, } I_2 = \frac{R_{11}}{R_{12}} I_1 \quad (4.43)$$

and

$$I_3 = \frac{R_{11}}{R_{13}} I_1 \quad (4.44)$$

Fig.4.12 Equivalent circuit of the DAC

From equation (4.39), the output current is

$$I_{OUT} = [16 I_2 + I_3] \frac{1}{17} \quad (4.45)$$

Substituting equations (4.43), (4.44) and (4.40) in (4.45)

$$I_{OUT} = \left[\frac{16 R_{13} + R_{12}}{R_{12} R_{13}} \right] \frac{R_{11}}{17} \frac{V_2 + V_{OS}}{R_{11}} \quad (4.46)$$

The output analog voltage is given by

$$\begin{aligned} V_{OUT} &= I_{OUT} R \\ &= R \left[\frac{16 R_{13} + R_{12}}{R_{12} R_{13}} \right] \frac{V_2 + V_{OS}}{17} \end{aligned} \quad (4.47)$$

In this equation the resistors R , R_{12} , R_{13} are all of the same type. Thus, absolute variations in sheet resistance do not affect the analog output voltage. Also, the offset voltage of the control amplifier, V_{OS} , does not affect the output of the converter. The voltage V_2 (equation (4.41)) can be adjusted by means of the external resistors R_1 and R_2 to take care of the offset. However, the drift of the offset could cause an error as shown below.

The resistors R , $15R$, R_{11} , R_{12} , and R_{13} in Figure 4.12 being of the same type can be made to track well over temperature with a good layout. The quantity V_2 (equation (4.41)) is independent of temperature because V_{REF} is a stable reference voltage and, R_1 and R_2 are precision external resistors. The error voltage, caused due to temperature variations may be written as

$$\Delta = \left[\frac{16 R_{13} + R_{12}}{R_{12} R_{13}} \right] \frac{R}{17} V_{OS}(T) \quad (4.48)$$

This error is maximum when the equivalent resistors R_{12} and R_{13} are minimum i.e., when the currents from the MSB array and LSB array are maximum.

$$\text{Minimum } R_{12}, R_{13} = 4500/15 = 300 \text{ ohms}$$

$$R = 500 \text{ ohms}$$

$$\text{Thus, } \Delta_{\max} \approx 1.67 V_{OS}(T).$$

Assuming that all the other components of the D/A converter are ideal and the only error component is due to the offset drift of the control amplifier, the maximum bound on the drift may be obtained by equalizing the maximum error to 1/2 LSB.

$$\text{i.e., } 1.67 V_{OS}(T) \leq 2\text{mV}$$

$$\text{or } V_{OS}(T) \leq 1.2\text{mV}$$

Thus, the offset of the control amplifier should not change by more than 1.2mV over temperature. The same holds true with time also. If the D/A converter is to operate in an industrial environment (-25°C to +85°C), the offset drift of the control amplifier should be as low as 11μV/°C. The design of such a low drift amplifier is discussed at length in Chapter 5.

4.6 Summary

In this chapter we have discussed the design of a multiple current source 8-bit video speed DAC. We have briefly outlined the disadvantages of most current steering configurations and proposed the interstage multiple current source DAC configuration as an optimum one. A statistical error analysis has been performed on this configuration to relate the circuit yield to the matching accuracy of the current sources. Although the analysis is specific to an 8-bit converter, the method can be easily extended to different resolutions, and even to other configurations. We have also discussed the design of the unit current source. We have used a transistor and resistor in combination to improve the current matching of the transistor by about 50%. The characterization and modelling presented in Chapter 2 indicates that this improvement could be obtained by increasing the channel area by a factor of four and then we can do away with the source feedback resistor. The choice between these two approaches should be based on area considerations and the circuit configuration.

The speed and glitch performance of the DAC is very much dependent on the bit cell design. A new bit cell configuration has been described. The complete video bandwidth DAC has also been presented. Finally, the chapter concludes with a need for a low drift control amplifier.

CHAPTER 5

CONTROL AMPLIFIER

It was shown in Chapter 4 that the offset voltage drift of the current control amplifier should be very small if the overall accuracy of the D/A converter be ensured. This chapter is devoted to the design of a low drift operational amplifier subsystem. First a general discussion of the causes of offset and drift is presented. Then a low drift technique suitable for control amplifier application is discussed. This is followed by the design of a basic amplifier block to be used in the low drift amplifier subsystem.

5.1 Causes Of Offset and Drift

The input offset voltage of an amplifier is composed of two components, the systematic offset and the random offset [37]. The former results from the design of the circuit and is present even when all of the matched devices in the circuit are indeed identical. Consider, for example, a two stage amplifier consisting of a differential input stage and a common source gain stage as shown in Figure 5.1. Assuming perfectly matched devices, for zero input the quiescent voltage at the drain of M_4 is equal to that at the drain of M_3 . However, the value of the gate voltage of M_6 which is required to force the output voltage to zero may be different from the quiescent output voltage of the

first stage. For a first stage gain of 50, for example, each 50 mV difference in these voltages results in a 1mV of input-referred systematic offset. This means that the (W/L) ratios of M_3 , M_4 and M_6 must be chosen so that the current density in these devices are equal. This implies

$$\frac{(W/L)_3}{(W/L)_6} = \frac{(W/L)_4}{(W/L)_6} = \frac{1}{2} \frac{(W/L)_5}{(W/L)_7} \quad (5.1)$$

In order that this ratio be maintained over process induced variations in channel length, M_3 , M_4 and M_6 must have the same channel length, and the ratios be provided by properly choosing the channel widths. However, for low input noise the input devices should have long channel lengths, and for stable frequency response under capacitive loading M_6 should have high transconductance necessitating small channel length. In addition, the systematic offset voltage could be dependent on operating conditions such as the common mode bias voltage. And since the gain per stage of the MOS amplifiers is generally lower than their bipolar counterparts, the systematic offset voltage of MOS amplifiers is generally higher.

Random input offset voltage results from mismatch in supposedly identical pairs of devices. MOS differential amplifiers display higher input offset voltage than bipolar differential pairs. Consider the circuit shown in Figure 5.2. Let there be a mismatch in the load resistors by δR . In order the output voltage be zero, the amount of

Fig.5.1 Two-stage amplifier illustrating interstage coupling constraints

mismatch in the current δI to be introduced is given by

$$I(R + \delta R) - (I + \delta I) R = 0$$

or $\delta I = \frac{\delta R}{R} \cdot I$ (5.2)

where $I = I_{\text{bias}}/2$

This change in current is produced by a d.c. input difference voltage given by

$$\begin{aligned} \delta V_{gs} &= \frac{\delta I}{g_m} \\ &= \frac{I}{g_m} \cdot \frac{\delta R}{R} \end{aligned} \quad (5.3)$$

Thus, the input offset is proportional to the I/g_m ratio of the input devices. For bipolar devices, the I/g_m ratio is equal to kT/q or 26mV at room temperature. For MOS devices, the ratio is $(V_{gs} - V_T)/2$, a bias dependent quantity which is normally in the 100 - 500mV range. While the offset voltage can be substantially improved by operating the devices at low values of V_{gs} , the result is typically a somewhat larger offset voltage than in the bipolar case.

An additional contributor to the random offset of MOS operational amplifiers which is not present in bipolar ones is the mismatch in the threshold voltage itself. This component is bias current independent.

(a)

(b)

Fig.5.2 (a) simple MOS differential stage
(b) Equivalent circuit

The above discussion serves to point out that the input offset voltage of MOS operational amplifiers cannot be made negligibly small even with good care in design, layout and processing. However, it was highlighted in Chapter 4 that the absolute value of the offset is not a problem, but it is the drift that has to be minimized. It will now be shown that the offset drift is unfortunately directly proportional to the initial offset voltage.

Again, consider the circuit shown in Figure 5.2(a). The overall offset caused due to the input devices, load devices and any subsequent stages is the amount of differential input voltage V_{OS} to be applied at the input to produce zero output voltage.

$$\text{i.e., } V_{OS} = V_{gs1} - V_{gs2} \quad (5.4)$$

Thermal offset drift is the derivative of the offset voltage with respect to temperature.

$$\frac{dV_{OS}}{dT} = \frac{dV_{gs1}}{dT} - \frac{dV_{gs2}}{dT} \quad (5.5)$$

Therefore, the drift is caused as a result of inaccurate thermal tracking of gate-to-source voltages. It is experimentally observed that both n- and p-channel MOS transistors can be operated such that they exhibit either positive, negative or zero temperature coefficients depending on the drain current. A good approximation to the behaviour in the saturation region is given by [38].

$$\frac{d|V_{gs}|}{dT} = n - 2.5 \times |V_{gs} - V_T| \text{ mV/}^\circ\text{C} \quad (5.6)$$

where V_T is the measured threshold voltage obtained from the straight line extrapolation of an I_D/g_m plot on the V_{gs} axis, and n is a constant whose value is somewhat dependent on the substrate doping but normally lies between 2 and 6.

From equations (5.5) and (5.6) the offset drift may be written as

$$\begin{aligned} \frac{dV_{OS}}{dT} &= -2.5 [V_{gs1} - V_{gs2} - V_{T1} + V_{T2}] \\ &= -2.5 V_{OS} + 2.5 (V_{T1} - V_{T2}) \text{ mV/}^\circ\text{C} \end{aligned} \quad (5.7)$$

Thus, the offset drift with temperature is directly proportional to the offset voltage.

An additional problem with MOS amplifiers is the surface traps and oxide charge. A redistribution of this charge shifts the threshold voltage of the devices bringing about a drift component with time.

From the foregoing discussion, it is clear that it is not possible to produce MOS operational amplifiers with negligible offset drift. Special circuit techniques which overcome this problem are discussed below.

5.2 Autozeroing Technique

Specific advantages of MOS devices can be used to reduce offset voltage and drift. MOS transistors can be readily connected as high quality switches and, the high input impedance of MOS gates allow long time storage of offset voltage on a capacitor. In sampled data systems, the offset voltage can be sampled during a clock phase when the input signal is not present, by shorting the output of the amplifier to its input (Figure 5.3(a)). This is known as the autozero phase. Subsequently, the offset is subtracted from the combined signal plus offset (Figure 5.3(b)). This technique is also known as correlated double sampling. For continuous time systems, however, chopper stabilized techniques are required to reduce the offset and drift.

5.3 Chopper Amplifier

The offset or drift of an amplifier is a d.c. or low frequency signal. This component may therefore be removed by making the amplifier an a.c. coupled one. Figure 5.4 shows a classical chopper amplifier. The idea is to chop the input signal by means of a switch S_1 , producing a square wave whose amplitude is proportional to the input. The chopped signal is then amplified by an a.c. coupled amplifier in which the d.c. drifts are not passed to the output. The output of the amplifier is then demodulated by a second chopper switch S_2 and smoothed by the $R_3 C_3$ filter to give the output signal. Since the

Fig.5.3 Principle of autozeroing

(a) Autozero mode (b) Amplification mode

smoothing time constant must be much larger than the period of the switching frequency to give satisfactory ripple reduction, the bandwidth of the system will be less than the chopping frequency.

Increasing the chopping frequency, however, leads to problems of capacitive feed through of the switch-drive into the amplifier input. A modification of this approach which uses smaller capacitors is reported in [39]. It uses a differential configuration to compensate for the switching parameters, and, the leakage currents of switches and capacitors. Even then the bandwidth of the system is only a few kilohertz.

5.4 Chopper-Stabilized Amplifier

A chopper amplifier can, however, be combined with a wide band amplifier to result in a configuration that has low drift of a chopper amplifier and yet have a wide bandwidth. Such a combination is the chopper-stabilized amplifier. The operation of a chopper-stabilized amplifier may be explained with reference to Figure 5.5. Here A_2 is the main amplifier and A_1 is an auxiliary amplifier connected as a chopper amplifier. The d.c. and low frequency components of V_{in} are chopped by the modulator M , amplified by A_1 , demodulated by D , smoothed by $R_3 C_3$ and, amplified by A_2 . Thus the net gain of these components is $A_1 A_2$. Higher frequency components of V_{in} pass directly to A_2 via $C_4 R_4$ so that the gain of these is only A_2 . The low drift of the overall system can be explained in two ways. As described above, the chopper

Fig.5.4 Chopper amplifier

Fig.5.5 Chopper stabilized amplifier

amplifier can amplify dc signals and contributes negligible drift. The drift at the input of A_2 due to A_2 must now be compared with A_1 times the input signal V_{in} , rather than just V_{in} , and so is A_1 times less significant. This view holds without the necessity of feedback components R_i and R_f . If the feedback is connected, then an alternative description, that gives the system its name, is more appropriate. Drift at the input of A_2 will be amplified and fed back to node V_G . This signal will be amplified and inverted by A_1 and fed to A_2 in a sense to counteract the original drift. Thus, the chopper amplifier stabilizes the d.c. operating point of A_2 , and hence the name [40].

A major difficulty in implementing the circuit of Figure 5.5 in monolithic form is the presence of coupling capacitors. A modification of the basic approach which does not require coupling capacitors but only requires two external capacitors for filtering and storing the offset is reported in [41]. A slight variation of this technique is used to realize the control amplifier as described below.

Figure 5.6 shows the block diagram of the chopper stabilized amplifier. There are two amplifiers. The main amplifier and the null amplifier. Both have differential offset nulling inputs in addition to the usual differential signal inputs. The null amplifier is not an a.c. coupled amplifier. On the other hand, the null amplifier is connected in such a way that it alternately nulls itself and the main amplifier under the control of the chopping clock storing correction

Fig.5.6 Modified chopper stabilized amplifier

voltages on two external capacitors. These voltages drive the offset null adjustment nodes of the amplifiers. The main amplifier is never switched and hence no output spikes are produced.

Like the conventional chopper stabilized amplifier, the main amplifier and the null amplifier need not be identical. A low bandwidth amplifier can be used as the null amplifier. Also, as shown below, the gain for the signal inputs and the null inputs in each amplifier need not be identical.

The response of the system at low frequencies (below the chopping frequency) can be seen by considering each clock phase separately (Figure 5.7). Let the gain for the signal inputs be A_N and that for the null inputs be $K_N A_N$. Referring to Figure 5.6, during ϕ_1 , the nulling amplifier signal inputs are shorted and a voltage almost equal to the offset of the amplifier is stored on C_1 . This situation is shown in Figure 5.7(a), where V_{OSN} is the offset of the null amplifier and V_N is the offset adjust voltage. V_N may be written as

$$\begin{aligned} V_N &= A_N V_{OSN} - K_N A_N V_N \\ \therefore V_N &= \frac{A_N V_{OSN}}{1 + K_N A_N} \end{aligned} \quad (5.8)$$

The effective offset voltage seen at the signal inputs may then be written as

$$\begin{aligned}
 V_{OSN(eff)} &= \frac{A_{II} \cdot V_{OSN} - K_N A_{II} \cdot V_N}{A_{II}} \\
 &= \frac{V_{OSN}}{1 + K_N A_{II}} \quad (5.9)
 \end{aligned}$$

Thus, the offset voltage is reduced by a factor almost equal to the null input gain.

When the switches are reversed during ϕ_2 , the arrangement is shown in Figure 5.7(b). The signal input is amplified by the null amplifier and then applied to the non-inverting null input of the main amplifier. A deviation from zero corresponds to an input offset in the main amplifier which after amplification by the nulling amplifier is stored on C_2 . If A_M is the signal gain and $K_M A_M$ is the offset gain of the main amplifier, the overall output V_{OUT} may be written as

$$\begin{aligned}
 V_{OUT} &= A_M [V_+ - V_- + V_{OSM}] \\
 &\quad + K_M A_M [A_N (V_+ - V_- + V_{OSN(eff)}) - V_N] \quad (5.10)
 \end{aligned}$$

Substituting for $V_{OSN(eff)}$ and V_N from equations (5.8) and (5.9), we have

$$V_{OUT} = A_M A_N [K_N (V_+ - V_-) + \frac{V_{OSM}}{A_N}] \quad (5.11)$$

Therefore, the resulting offset of the system is the offset of the main amplifier divided by the open loop gain of the null amplifier. By using a high gain null amplifier, the offset voltage can be made

(a)

(b)

Fig.5.7 (a) Null amplifier block diagram for clock phase 1
(b) System block diagram during clock phase 2

negligibly small. As seen above, this technique does not require the main amplifier and the null amplifier to be identical. In addition, the signal gain and the offset null gain need not be the same thus avoiding any tight matching requirements.

A subtle observation can be made at this point about the non-inverting input of the null amplifier. This node may be connected to an appropriate d.c. potential which need not be precise nor be stable. Also, high frequency noise on this input will not affect the output due to the large time constant of the filter network at the output of the null amplifier. Therefore, a simple on chip voltage divider type reference may be used.

It is interesting to determine the overall frequency response of the system. At low frequencies, the dominant signal path is through the cascade of both amplifiers giving overall system open loop gain equal to the product of the individual amplifier open loop gains. At high frequencies, however, the presence of the filters at the output of the null amplifier decreases the gain through this path and thus the feed-forward path from the inputs directly to the main amplifier dominates. Thus high frequency components of the chopping clock are not a problem.

The time constants of the filter networks should be selected in such a way that the overall system be stable in the unity gain configuration.

Let both the amplifiers have single pole responses. The frequency response of the main amplifier be as shown in Figure 5.8(a). The time constant $R_2 C_2$ must be so chosen that the overall frequency response of the null amplifier and the filter reduces to unity or 0dB at the break frequency of the main amplifier as shown in Figure 5.8(b). Then the response of the overall system is approximately a single pole response as shown in Figure 5.8(c).

The chopper stabilized technique described here can be used to reduce the noise of an amplifier. For the system, any noise below the chopping frequency appears as an offset error and is therefore reduced. In voice band applications of switched capacitor filters, the dominant noise source is often the $1/f$ noise component of the operational amplifier. One circuit technique for reducing the $1/f$ noise at low frequency is the correlated double sampling (see Section 5.2). This operation requires that the output node of the amplifier slew back and forth between the signal level and the initialized level each clock period. This puts a severe requirement on the operational amplifier settling time. Thus, the chopper stabilized technique described here should be a better approach to reduce the low frequency noise.

Another advantage brought about by the chopper stabilized technique is the improvement in common mode rejection ratio (CMRR) and power supply rejection ratio (PSRR). The nulling operation is independent of input common mode voltage or power supply voltages

Fig.5.8 Frequency response of the amplifiers

(a) Main amplifier

(b) Null amplifier with R2C2 filter network

Fig. 5.8(c) Frequency response of the overall amplifier

and hence the system provides exceptional CMRR and PSRR at frequencies below the switching frequency, where the two amplifiers are effectively cascaded.

An inherent disadvantage with the chopper stabilization technique is the long over load recovery time. When the output limits, the nulling circuit will attempt to correct the non zero input to the main amplifier and will limit as well. After the overload is removed the long time constant of the nulling loop determined by the RC networks will prolong recovery of the entire system. This should pose no threat to its use as a control amplifier in the D/A converter.

Having discussed the drift reduction technique for the control amplifier, we will now describe the basic operational amplifier block that is part of the chopper stabilized system.

5.5 Basic Operational Amplifier Block

The performance objectives for operational amplifiers to be used within a monolithic analog subsystem are often quite different from those of traditional stand alone bipolar amplifiers. Perhaps the most important difference is the fact that the load of the amplifier is well defined and, is often purely capacitive with values of a few picofarads.

The control amplifier of the D/A converter is one of this type. The load offered to it is the common gate of all the current source transistors. This permits the output stage of the amplifier to have a rather high impedance in contrast to the low output impedance of "ideal" operational amplifiers. The dominant pole required to provide a stable frequency response may then be placed at the output. This is in contrast to the pole splitting type compensation of conventional two stage amplifiers whose dominant pole is at the output of the first stage. In other words, the load capacitance of a high output impedance amplifier itself acts as the compensating capacitor, and there is no need for an additional compensating capacitor.

The design of an operational amplifier which has its dominant pole at the output is simpler than a conventional two stage amplifier. Besides, such a configuration offers savings in power consumption and die area. In conventional pole splitting type frequency compensation, the overall stability depends on the output load capacitance. As this value increases the parasitic pole frequency decreases and instability could result. To place the parasitic pole due to the load capacitance beyond the unity gain frequency of a wide band amplifier, a very low output impedance stage has to be provided, even though the amplifier may not have to drive large resistive loads. This requires an increase in bias current level. An order of magnitude increase in bandwidth dictates an order of magnitude increase in current and power consumption. Also to provide low voltage drops across the devices calls for their

size to increase. These two factors are not in keeping with the VLSI context. These considerations prompted the use of an operational transconductance amplifier to serve as the basic amplifier block for the control amplifier sub system.

Two configurations, both using cascode stages for the output were investigated [37], [42]. Both were originally intended for switched capacitor filter applications. The amplifier reported in [42] is actually a two stage one with current mirrors to translate the signal from the differential input stage to the output stage. It was found that even at moderate values of the current levels the stability of the amplifier degraded due to the parasitic pole at the output of the first stage. It was decided that the current in the output stage of the basic amplifier block to be a few tens of micro amperes. This is to allow the external filter capacitors in the chopper stabilized system to recover quickly from any minor disturbances. Therefore, the single stage configuration of [37] was chosen to be the basic amplifier block.

The circuit schematic of the basic amplifier block is shown in Figure 5.9. It consists of a p-channel differential stage comprising of transistors $M_1 - M_3$. The output of this stage is connected to the source terminals of the common gate devices M_4 and M_5 . Transistors $M_6 - M_9$ form the cascoded differential to single ended converter. The common gate devices M_4 and M_5 are biased by the constant current sources M_{10} and M_{11} respectively. The bias current levels in various devices

is shown in Figure 5.9. The reason for using p-channel devices for the input is that it would permit an input common mode voltage almost down to the V_{SS} rail. The common mode voltage of the control amplifier is closer to V_{SS} than V_{DD} .

We will now determine the various performance parameters of this amplifier stage.

5.5.1 Open Loop Gain

The amplifier output, due to its very high impedance (resulting from the cascode stage, as shown below) may be modelled as a current source. The amplifier may therefore be represented by the equivalent circuit shown in Figure 5.10. Thus, the gain of the amplifier is given by the product of the transconductance of the input transistors and the impedance at the output node. The output impedance may be computed in two parts, the one due to the input along with the n-channel cascode transistors and, the other resulting from the differential to single ended converter.

The equivalent circuit of the input along with the n-channel cascode devices is shown in Figure 5.11(a). As shown, a voltage source is connected to the output and the resultant current is calculated with the input voltage V_{in} set to zero. Transistors M_3 and M_{11} are simply represented by their output conductances. The common gate device M_5 is replaced by a current source equal to the source conductance of M_5 .

Fig.5.16 Equivalent circuit of the amplifier

times the signal at the source V_{S1} in parallel with its conductance.

$$g_{OUT1} = \frac{I_{OUT1}}{V_{OUT1}} = \frac{V_{S1} (g_{03} + g_{011})}{V_{OUT1}} \quad (5.12)$$

The voltage V_{S1} may be found by superimposing the effects of the sources g_{S5} , V_{S1} and V_{OUT1} .

$$V_{S1} = \frac{-g_{S5} \cdot V_{S1}}{g_{03} + g_{011} + g_{05}} + \frac{g_{05} \cdot V_{OUT1}}{g_{03} + g_{011} + g_{05}}$$

$$\text{i.e., } \frac{V_{S1}}{V_{OUT1}} = \frac{g_{05} (g_{03} + g_{011})}{g_{03} + g_{011} + g_{05} + g_{S5}} \quad (5.13)$$

Substituting equation (5.13) in (5.12)

$$g_{OUT1} = \frac{g_{05} (g_{03} + g_{011})}{g_{03} + g_{011} + g_{05} + g_{S5}} \quad (5.14)$$

where the source transconductance g_S is given by

$$g_S = g_m + g_{mb}$$

$$g_m \triangleq \frac{\partial I_D}{\partial V_{GS}}$$

$$g_{mb} \triangleq \frac{\partial I_D}{\partial V_{BS}} = n g_m$$

$$n \triangleq \frac{\gamma}{2\sqrt{(-V_{BS} + 2\phi_f)}}$$

$$\text{and } \gamma = \frac{\sqrt{2 \epsilon_s q N_A}}{C}$$

ϵ_s is the permittivity of silicon, q is the electronic charge, N_A is the substrate impurity concentration and, C is the gate oxide capacitance per unit area. The value of the source conductance g_s is much greater than the output conductance g_o . Hence equation (5.14) may be approximated as

$$g_{OUT1} = \frac{g_{05} (g_{03} + g_{011})}{(1 + n) g_{m5}} \quad (5.15)$$

Thus, there is an effective increase in the output impedance by a factor equal to the gain (g_{m5}/g_{05}) of the common gate device M_5 .

The output impedance of the differential to single ended converter may be calculated in a similar way. Again, setting the input to zero and connecting a source at the output, the equivalent circuit may be written as shown in Figure 5.11(b).

$$g_{OUT2} = \frac{I_0}{V_{OUT2}} = \frac{V_{S2} g_{m7}}{V_{OUT2}} \quad (5.16)$$

Again, V_{S2} is found by superimposition

$$V_{S2} = \frac{-g_{S9} \cdot V_{S2}}{g_{m7} + g_{09}} + \frac{g_{09} \cdot V_{OUT2}}{g_{m7} + g_{09}}$$

$$\frac{V_{S2}}{V_{OUT}} = \frac{g_{09}}{g_{m7} + g_{09} + g_{S9}} \quad (5.17)$$

From equations (5.16) and (5.17)

Fig.5.11 Equivalent circuits for calculating the output impedances of the (a) input and N-channel cascode stage (b) differential to single ended converter

$$\begin{aligned}
 g_{OUT2} &= \frac{g_{m7} \cdot g_{09}}{g_{m7} + g_{09} + g_{s9}} \\
 &= \frac{g_{m7} \cdot g_{09}}{g_{m7} + (1+n) g_{m9}}
 \end{aligned} \tag{5.18}$$

The total output conductance is the sum of the two conductances given by equations (5.15) and (5.18). Hence the voltage gain is given by

$$A_V = g_{m3} / \left[\frac{g_{05} (g_{03} + g_{011})}{(1+n) g_{m5}} + \frac{g_{m7} \cdot g_{09}}{g_{m7} + (1+n) g_{m9}} \right] \tag{5.19}$$

5.5.2 Bandwidth

The 3-dB bandwidth of the amplifier is given by

$$f_{3dB} = \frac{1}{2\pi R_{OUT} C_L} = \frac{g_{OUT}}{2\pi C_L} \tag{5.20}$$

and the unity gain bandwidth product is

$$GBW = g_{m3} / 2\pi C_L \tag{5.21}$$

As the transconductance varies directly as the square root of the current, the unity gain frequency increases as the square root of the bias current.

Two parasitic poles arise due to the capacitances at the source nodes of M_5 and M_9 . The impedance at the source of M_5 may be determined from Figure 5.12(a). The total output current due to an applied output

voltage is given by

$$g_{OUT} = V_{OUT} (g_{03} + g_{05} + g_{011}) + g_{S5} \cdot V_{OUT}$$

$$\therefore g_{OUT} = g_{03} + g_{05} + g_{011} + (1+n) g_{m5}$$

$$\approx (1+n) g_{m5} \quad (5.22)$$

Also, from Figure 5.12(a), the capacitance at this node is the sum of the gate to source capacitance of M_5 , drain to gate capacitance of M_{11} and any parasitic capacitance due to the source and drain diffusions of M_5 and M_{11} . Therefore, the parasitic pole occurs at a frequency given by

$$f_{p1} = \frac{(1+n) g_{m5}}{C_{gS5} + C_{gd11} + C_{p1}} \quad (5.23)$$

where C_{p1} is the total stray capacitance at this node.

Similarly, the impedance at the source node of M_9 may be determined from Figure 5.12(b). The output current due to the external voltage is given by

$$i_{OUT} = V_{OUT} (g_{m7} + g_{09}) + g_{S9} \cdot V_{OUT}$$

$$\therefore g_{OUT} = g_{09} + g_{m7} + (1+n) g_{m9}$$

$$\approx g_{m7} + (1+n) g_{m9} \quad (5.24)$$

(a)

(b)

Fig.5.12 Equivalent circuit to determine the output impedance (a) at the source node of M5 (b) at the source node of M9

The total capacitance at this node is due to the gate-source capacitance of M_9 and, any stray capacitance C_{p2} due to the source diffusion of M_9 and drain diffusion of M_7 . Thus the other non dominant pole occurs at

$$f_{p2} = \frac{g_{m7} + (1+n) g_{m9}}{C_{gs9} + C_{p2}} \quad (5.25)$$

Comparing equations (5.23) and (5.25) we see that the first non-dominant pole is at f_{p1} and the second one is at f_{p2} .

5.5.3 Noise

The noise of MOS amplifiers working in the low frequency range is high due to the dominant contribution of flicker noise, or $1/f$ noise. The most widely accepted model for $1/f$ noise is that for a given device, the gate referred mean-squared voltage noise is approximately independent of bias conditions, in saturation, and is inversely proportional to the gate area of the device.

$$e_n = \sqrt{\frac{a_n}{WL}} \quad (5.26)$$

where W and L are, respectively, the width and length of the channel and a_n is a parameter dependent on technology. It will be shown below that increasing the WL product of the input transistors does not decrease the input noise, as all the devices in the circuit contribute to the input noise [43].

The circuit diagram of the operational amplifier including the input noise sources is shown in Figure 5.13. The analysis may be simplified by considering two separate half sections. Further, the contribution of the cascode devices may be determined by referring to Figure 5.14. Considering V_{n11} alone, (Figure 5.14(b))

$$V_{S5} = - \frac{g_{m11} V_{n11}}{g_{011}} \quad (5.27)$$

$$\begin{aligned} V_{OUT} &= V_{S5} + \frac{(1+n) g_{m5} \cdot V_{S5}}{g_{05}} \\ &= - \frac{g_{m11} V_{n11}}{g_{011}} [1 + (1+n) \frac{g_{m5}}{g_{05}}] \\ \therefore A_{n11} &= \frac{V_{OUT}}{V_{n11}} = - \frac{g_{m11}}{g_{011}} \frac{(1+n) g_{m5}}{g_{05}} \end{aligned} \quad (5.28)$$

Now consider V_{n5} alone. The equivalent circuit is shown in Figure 5.14(c).

$$V_{S5} = \frac{g_{m5} V_{n5} - (1+n) g_{m5} V_{S5}}{g_{011}} \quad (5.29)$$

$$\begin{aligned} V_{OUT} &= V_{S5} - [g_{m5} V_{n5} - (1+n) g_{m5} V_{S5}] \cdot g_{05} \\ &= V_{S5} [1 + (1+n) g_{m5} g_{05}] - g_{m5} g_{05} V_{n5} \end{aligned}$$

Substituting for V_{S5} from equation (5.29) and noting that $g_{05} \ll (1+n) g_{m5}$, we may write the gain for the noise input V_{n5} as

Fig.5.13 Individual noise sources in the op amp

Fig.5.14 (a) Noise equivalent of the cascode stage
(b)&(c) equivalent circuits when the two noise sources are independently considered

$$A_{n5} = \frac{V_{CUT}}{V_{n5}} \approx \frac{1}{1+\eta} \leq 1 \quad (5.30)$$

Since the two noise sources are uncorrelated, the equivalent input noise at the gate of M_{11} is given by

$$\begin{aligned} V_n (IN) &= \frac{1}{A_{n11}} [(A_{n5} \cdot e_{n5})^2 + (A_{n11} \cdot V_{n11})^2]^{1/2} \\ &= [(\frac{A_{n5}}{A_{n11}} \cdot e_{n5})^2 + V_{n11}^2]^{1/2} \end{aligned}$$

From equations (5.28) and (5.30)

$$\frac{A_{n5}}{A_{n11}} \ll 1$$

$$\therefore V_n (IN) \approx V_{n11} \quad (5.31)$$

Thus the effective input noise is due to M_{11} alone and the contribution of M_5 may be ignored. Similarly, it can be shown for the p-channel differential to single ended converter, that the major contributor to output noise is M_7 and M_9 may be ignored. Then the operational amplifier circuit for noise calculations alone may be written as shown in Figure 5.15. Again, since the various noise sources are uncorrelated, the equivalent noise at the input of M_3 is given by

$$\begin{aligned} V_N (IN) &= \frac{1}{A_{n3}} [(A_{n3} \cdot V_{n3})^2 + (A_{n5} \cdot V_{n5})^2 + (A_{n11} \cdot V_{n11})^2]^{1/2} \\ &= [(V_{n3})^2 + (\frac{A_{n5}}{A_{n3}})^2 \cdot V_{n5}^2 + (\frac{A_{n11}}{A_{n3}})^2 \cdot V_{n11}^2]^{1/2} \quad (5.32) \end{aligned}$$

Fig.5.15 Simplified circuit for calculating the equivalent input noise of the amplifier

The various gains are given by

$$A_{n3} = g_{m3} / g_{OUT} \propto \sqrt{(I K_p W_3 / L_3)} \quad (5.33a)$$

$$A_{n5} = g_{m5} / g_{OUT} \propto \sqrt{(I K_p W_5 / L_5)} \quad (5.33b)$$

$$\text{and } A_{n11} = g_{m11} / g_{OUT} \propto \sqrt{(2I K_N W_{11} / L_{11})} \quad (5.33c)$$

From equation (5.26) the input noise voltages may be written as

$$V_{n3} = \sqrt{\frac{a_{np}}{W_3 L_3}} \quad (5.34a)$$

$$V_{n5} = \sqrt{\frac{a_{np}}{W_5 L_5}} \quad (5.34b)$$

$$V_{n11} = \sqrt{\frac{a_{nn}}{W_{11} L_{11}}} \quad (5.34c)$$

where the second subscript in a_{nn} and a_{np} refer to the n- and p-channel devices respectively. Substituting equations (5.33) and (5.34) in (5.32) we get

$$V_{in} (IN) = \left[\frac{a_{np}}{W_3 L_3} \left(1 + \left(\frac{L_3}{L_5} \right)^2 + \left(\frac{2K_N}{K_p} \cdot \frac{a_{nn}}{a_{np}} \cdot \frac{L_3}{L_{11}} \right)^2 \right) \right]^{1/2} \quad (5.35)$$

To determine the total input noise for the differential inputs, the other half section should also be considered which would result in a similar expression as equation (5.35). Therefore, the total equivalent input noise is obtained by multiplying equation (5.35) by $\sqrt{2}$.

$$V_N (IN) = \left[\frac{2a_{np}}{W_3 L_3} \left(1 + \left(\frac{L_3}{L_5} \right)^2 + \left(\frac{2K_n}{K_p} \cdot \frac{a_{nn}}{a_{np}} \cdot \frac{L_3}{L_{11}} \right)^2 \right) \right]^{1/2} \quad (5.30)$$

Thus, the noise decreased by increasing the width of the input transistors and the lengths of the common source transistors in the output cascode stage.

5.5.4 Biasing

A disadvantage of providing cascode devices at the output is the reduction of the output voltage swing due to the d.c. voltage drop across 4 devices as against that of 2 devices in conventional output stages. The devices can be made to operate on the verge of saturation to maximize the swing. The cascode differential to single ended converter is self biased and requires $(2V_{gs} - V_T)$ across it before clipping sets in. However, on the negative side, the voltage can swing down to $2(V_{gs} - V_T)$ of the negative supply rail, by using the arrangement shown in Figure 5.16 [44]. The operation of the circuit is self explanatory. Because VLSI processes need high substrate doping and hence show large body effect, the threshold voltages of all the devices are not nearly the same. This necessitates the source and bulk of transistors M_4 and M_5 to be shorted as shown.

(a)

(b)

Fig.5.16 (a) Cascode stage biasing arrangement
(b) I-V characteristics

5.6 Null Inputs To The Amplifier

The basic amplifier stage we have discussed so far has only two differential inputs for the signal. The chopper stabilized configuration requires two more inputs for nulling. We will now consider various possibilities for providing the additional inputs.

5.6.1 Configuration 1

The signal input and the null input may be combined as shown in Figure 5.17(a) and the output may be applied to the input of the differential stage. The disadvantage of this arrangement is that the input common mode range for the signal is severely limited. The only way to overcome this problem is to make the top device a depletion mode one [41]. This requires additional process steps and is therefore not suitable.

5.6.2 Configuration 2

The back gate or substrate of a MOS transistor can also be used as a high impedance input if care is taken to see that the source substrate junction is always reverse biased. Hence the signal input and offset input can be connected to the same device as shown in Figure 5.17(b) [45]. Such an arrangement cannot be used here because the input devices are p-channel ones in a common n-substrate.

5.6.3 Configuration 3

The third possibility is shown in Figure 5.17(c). The drawback with this arrangement is that the common mode levels for the signal and offset greatly influence their gain and could even shut one section completely off.

5.6.4 Configuration 4

The bias current sharing problem of the previous configuration can be overcome by having separate current sources for the signal and offset sections as shown in Figure 5.17(d). The currents from the two sections are appropriately summed by tying the corresponding drains together and, then they are fed into the output cascode stage. This configuration enables widely differing common mode levels for the signal and the offset. This is the one used for the control amplifier.

5.7 The Complete Amplifier

The complete amplifier with signal inputs, null inputs and the biasing circuitry is shown in Figure 5.18. The same configuration is used for both the main amplifier and the null amplifier to simplify layout. The non-inverting null input of the nulling amplifier is to be connected to a reference voltage within the input common mode range of the amplifier.

(a)

(b)

(c)

(d)

Fig.5.17 Various configurations for providing offset null inputs

Fig.5.18 Schematic of the complete amplifier block

As this voltage need not be stable (see Section 5.4), the reference voltage generator is as shown in Figure 5.19. The switches and switch drivers are also laid out. The clock is external to the circuit although internal clock circuitry could be added without much difficulty. The active area of the complete chopper stabilized amplifier is approximately 0.3mm^2 .

5.8 Summary

In this chapter we have discussed the design of the current control amplifier in detail. The causes of offset and drift in MOS amplifiers and methods to overcome them are discussed. The design of an operational amplifier with the dominant pole due to the output load capacitance is discussed. Methods to include the offset null inputs to the amplifier are also shown.

Fig.5.19 Reference voltage for the non-inverting null input of the null amplifier

CHAPTER 6

RESULTS

The complete 8-bit DAC including the chopper stabilized control amplifier was fabricated at Northern Telecom Electronics using their 3 μm CMOS process. In this chapter, we present measurement results of fabricated samples. As the chopper stabilized amplifier is a part of the DAC, we will discuss its performance first.

6.1 Chopper Stabilized Amplifier

The most important parameter of interest to us is the offset drift of the amplifier over temperature. In addition to this, we also measured the D.C. gain of the individual amplifiers and the overall unity gain bandwidth of the complete subsystem. The circuit used to measure the gain and offset is shown in Figure 6.1. The relationship between the output and input is

$$V_{\text{out}} = A (V_{\text{in}} - V_{\text{out}}) + A \cdot V_{\text{os}} \quad (6.1)$$

where, V_{os} is the input offset voltage of the amplifier and A is the d.c. gain. A variable d.c. source is used for V_{in} . The voltages V_{out} and $V_{\text{in}} - V_{\text{out}}$ are measured using high-resolution multimeters. The slope and intercept of the straight line given by equation (6.1) give the d.c. gain and the offset. The offset of the overall amplifier was also measured using the same method. Large systematic offset was

Fig.6.1 Circuit to measure op amp gain and offset

deliberately introduced to the individual amplifiers to check the performance of the chopper stabilization technique. It was found that the composite amplifier had a very low offset in every case. We also measured the offset drift of the individual amplifiers and the complete subsystem over 0° to 80°C. The gain bandwidth product was also measured. The results are shown in Table 6.1.

6.2 8-Bit DAC

All functionally good D/A converters are to be tested for static performance to evaluate the yield. The automated measurement set up is shown in Figure 6.2. A Tektronix DAS 9100 Series Digital Analysis System is used as a pattern generator to provide the 256 digital word inputs to the DAC. This is controlled by a HP9845B desk top computer. An automatic prober and a 5 1/2 digit multimeter are also connected to the computer through the IEEE-488 bus. The computer was programmed to first determine the functionality of the DAC. If a catastrophic failure is sensed, the automatic prober would step to the next location. On a functionally good DAC, the analog output voltage is measured for each of the 256 word inputs. If for any input word, the output exceeds by more than $\pm 1/2$ LSB from the ideal linear relationship, the DAC is a reject. A typical error plot for a good device is shown in Figure 6.3. The yield statistics are shown in Table 6.2. It is seen that on most of the wafers, the circuit yield is 100% surpassing our goal, demonstrating the accuracy of the device characterization and the DAC design.

TABLE 6.1

Chopper Stabilized Amplifier

Performance

Main Amplifier:

DC Gain 61 dB

Offset Drift (0°-80°C) 4 mV

Overall Amplifier:

Offset Drift (0°-80°C) 2 μ V

GBW 1.4 MHz

Fig.6.2 Automated DAC static test set up

Fig.6.3 Typical DAC error plot

TABLE 6.2

DAC Yield

Wafer ID	Devices Tested	Functional Devices	Good Devices	Circuit Yield %
W04-101-DJ2	32	32	31	97
W04-101-DJ3	55	55	55	100
W04-101-DJ4	64	60	60	100
W04-101-DJ5	64	52	52	100
W04-101-DJ7	64	60	60	100

To perform dynamic testing of the DAC, a good high-frequency printed circuit board was designed. The conversion time is normally defined as the time required for the output to settle to within 1/2 LSB for the full scale transition, i.e. from an all zero to all one transition. For an 8-bit converter, the output should settle to within 0.2%. Because it is difficult to measure this low level on an oscilloscope, a commonly used method to determine the conversion time is to extrapolate the conversion time from rise time measurements [46]. In a single pole system, the D/A output at any instant may be obtained by the exponential relationship

$$V = V_0 (1 - e^{-t/\tau}) \quad (6.2)$$

where τ is the time constant, and V_0 is the final voltage of the DAC. The rise time t_r is related to the time constant by

$$t_r = 2.2\tau \quad (6.3)$$

From equation (6.2), the conversion time of an 8-bit DAC may be shown to be approximately $2.8 t_r$. The oscillogram of the full scale transition is shown in Figure 6.4 and displays a settling time of about 80 ns. This value agrees well with simulations when a 20 pF capacitor is connected across the 500 ohm load resistor. But for the large parasitic capacitance associated with the divide by 16 network and the load resistance, the conversion time would have been smaller. Yet, the value we have achieved is comparable to the state of the art for MOS D/A converters [48-49].

Fig.6.4 Full scale settling time of the DAC
(10X attenuation due to picoprobe)

The linearity measured at 10 MHz is shown in Figure 6.5. The transfer function does not exhibit any spikes.

The worst case glitch which occurs for the 1000 0000 to 0111 1111 transition is shown in Figure 6.6. The glitch energy is about 500 pV-sec. This is a low value and compares well with some of the commercially available video DACs [47]. The multiple current source approach and the new bit cell design are responsible for this good glitch performance.

The photomicrograph of the DAC is shown in Figure 6.7. The active area of the chip is 1.8 mm^2 including the chopper stabilized control amplifier. The source feedback resistors of the current sources occupy a large area. We have been very conservative and used $12 \text{ }\mu\text{m}$ wide resistors. Resistor matching studies show that $3 \text{ }\mu\text{m}$ wide resistors should be adequate.

The performance of the DAC is summarized in Table 6.3. The full-scale current of the DAC is 4.1 mA. Approximately half of this current is wasted to ground due to the divide by 16 network. The control amplifier draws very little current because the component amplifiers are output transconductance types with the dominant pole at the output. Finally, the worst case drift of the DAC is a very low value of $23 \text{ ppm}/^\circ\text{C}$. Again, this compares very favourably with $17 \text{ ppm}/^\circ\text{C}$ achieved in hybrid video DACs [47].

Fig. 6.5 Linearity display at 10MHz clock frequency

DAC output

1000 0000
to
0111 1111
transition

Fig.6.6 Worst case glitch of the DAC

Fig.6.7 Photomicrograph of the DAC chip

TABLE 6.3

DAC Performance

Nonlinearity - differential	$< \pm 1/2 \text{ LSB}$
- integral	$< \pm 1/2 \text{ LSB}$
Settling time to $\pm 1/2 \text{ LSB}$	80 ns
Worst-case glitch energy	500 pV-s
Supply voltage	5 V
Full-scale analog output	1 V
Supply current	4.1 mA
Chip area	1.8 mm^2
Worst-case drift	23 ppm/ $^{\circ}\text{C}$

The drift of the amplifier has been taken care of by chopper stabilization. This low overall drift for the complete DAC, therefore, demonstrates the good matching of transistors and resistors over temperature. This verifies our predictions about tracking of the transistor matching over temperature discussed in Section 2.6.

6.3 Summary

In this chapter we have discussed results of measurements on fabricated samples of the video DAC. Important parameters of the chopper stabilized control amplifier are discussed first. An automatic test set up for static testing and yield determination of the DAC is then presented. Finally, dynamic test results of the DAC are reported, and its performance is shown to be comparable to commercially available hybrid DACs.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

In this chapter we offer some concluding remarks and suggestions for further research.

7.1 Conclusions

The design of precision analog circuits for very large scale integration presents challenges in both device matching characterization and circuit design. Novel methodologies relevant to both aspects of the design have been presented.

The models developed for the mismatch in transistors and resistors complement the ones that already exist for MOS capacitors. The characterization of the mismatch in MOS transistors is more complex than in the case of capacitors. The matching not only depends on the dimensions of the device but also on the operating point. A characterization methodology is developed that accurately determines the mismatch in the drain current over a wide operating region, using a minimum set of measurement data. The overall objective of this work has been not only to provide a clear understanding of the random mismatch, but also to develop a comprehensive design approach for precision analog circuits. The parameters a circuit designer will have freedom to choose are the dimensions of the devices. Therefore, analytical models have been

developed that relate the mismatch to the device dimensions. Statistical methods are used to develop these models. Extensive measurements on specially designed test structures have been carried out to validate the theoretical analysis.

The second part of the thesis demonstrates the application of the matching characterization in circuit design. As the importance of analog circuits for VLSI is in the area of interface circuits, an 8-bit DAC is chosen as a test vehicle. In order that this test vehicle will also be a useful library cell, considerations like speed, glitch and drift were taken into account in the design. A new current steering bit cell was designed to provide high speed and low glitch energy. A chopper stabilized control amplifier has been designed to minimize the thermal drift of the DAC. The full scale conversion time of the DAC is 80 ns. The video speed achieved here is not possible with the older capacitive conversion techniques. Other important features of the DAC are 500 pV-s glitch energy, and 23 ppm/°C worst case drift. These values are comparable to those achieved in hybrid DACs. Finally, the circuit yield of the DAC was made to exceed 97% as a result of the studies of this thesis.

7.2 Future Research

There is scope for future research both in the area of device characterization and circuit design. The models we have developed have been tested for one process only. Process iterations throw more

light on the matching characteristics. For example, the P-channel devices can be fabricated without the threshold adjust implant. Then, comparison of the results with the standard process would pinpoint the effects of the threshold adjust implant. Even more drastic process changes such as an n-well process or a process having transistor gates with different work functions would be very useful. The device sizes in the new test structures may be so chosen as to provide information about the different factors causing mismatch by solving a set of simultaneous model equations.

Considering resistors, the test structures may be redesigned to eliminate systematic error components. Also there is a need for more resistor sizes. Implantations with different energies and doses may also be studied.

The device mismatch models we have developed may be incorporated into the model data base of a circuit analysis program which has provision for Monte Carlo simulations. Perhaps, this is the ultimate need of a circuit designer. Then circuits can be designed more efficiently.

In our analysis, we have only considered the intra-die or local variations. The knowledge about inter-die or global variations is also important for certain applications. Work may be pursued in this direction.

We have a few specific suggestions about improving the performance of the DAC. Although the DAC in its present form has successfully served as a test vehicle and meets state of the art specifications, the following modifications should enhance its performance.

In its present form, the speed of the DAC is limited by the on-chip load resistor. The divide by 16 network and the load resistor have a large parasitic capacitance. The use of an off-chip precision resistor as a load would greatly improve the speed.

When the unit current source for the DAC was designed, we did not have a clear understanding of the matching behavior as a function of the device dimensions. Therefore, to improve the matching of the current sources, a source feedback resistor was introduced. Now we find that the same degree of improvement can be obtained by increasing the channel area of the transistors by a factor of four, for the same gate-to-source bias. The resistors are then not required. This would reduce the chip area considerably.

The removal of the resistor would also reduce the voltage drop across the current source. This would enable a larger full scale analog output voltage. For example, with $\pm 2.5V$ power supplies and operating the digital section from 0 to $-2.5V$, the analog output can be ground referenced and a full scale voltage of zero to $2.5V$ is possible.

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

A drawback of the chopper stabilized control amplifier is the need to use two external capacitors. Alternative schemes which do not require these capacitors may be investigated.

The divide by 16 network is very inefficient in current handling. A large fraction of the supply current is wasted to ground. New configurations may be tried out.

Based on the results of matching characterization and the confidence we have achieved in the 8-bit DAC design, we may conclude that a 10-bit converter operating at video speeds and having good yields is feasible.

REFERENCES

- (1) J.L. McCreary and P.R. Gray, "All MOS Charge Redistribution Analog-to-Digital Conversion Techniques - Part I", IEEE J. Solid-State Circuits, Vol.SC-10, pp.371-379, December 1975.
- (2) D.A. Hodges, P.R. Gray, and R.W. Brodersep, "Potential of MOS Technologies for Analog Integrated Circuits", IEEE J. Solid-State Circuits, Vol.SC-13, pp.285-294, June 1978.
- (3) J.T. Caves, M.A. Copeland, C.F. Rahim, and S.D. Rosenbaum, "Sampled Analog Filtering Using Switched Capacitors As Resistor Equivalents", IEEE J. Solid-State Circuits, Vol.SC-12, pp.592-599, December 1977.
- (4) B.J. Hosticka, R.W. Brodersen, and P.R. Gray, "MOS Sampled Data Recursive Filters Using Switched Capacitor Integrators", IEEE J. Solid-State Circuits, Vol.SC-12, pp.600-608, December 1977.
- (5) J.L. McCreary, "Matching Properties, and Voltage and Temperature Dependence of MOS Capacitors", IEEE J. Solid-State Circuits, Vol.SC-16, pp.608-616, December 1981.
- (6) J.B. Shyu, G.C. Temes, and K. Yao, "Random Errors in MOS Capacitors", IEEE J. Solid-State Circuits, Vol.SC-17, pp.1070-1076, December 1982.
- (7) M. Akyia and S. Nakashima, "High-Precision MOS Current Mirror", IEE Proceedings, Vol.131, Pt.1, pp.170-175, October 1984.
- (8) J.B. Shyu, G.C. Temes, and F. Krummenacher, "Random Error Effects in Matched MOS Capacitors and Current Sources", IEEE J. Solid-State Circuits, Vol.SC-19, pp.948-955, December 1984.
- (9) K.R. Lakshmi Kumar, R.A. Hadaway, M.A. Copeland, and M.I.H. King, "A High-Speed 8-Bit Current Steering CMOS DAC", Proceedings, IEEE 1985 Custom Integrated Circuits Conference, pp.156-159.
- (10) S.M. Sze, "VLSI Technology", New York, McGraw-Hill, 1983, pp.328.
- (11) J. Logan, "Characterization and Modelling for Statistical Design", The Bell System Technical Journal, Vol.50, pp.1105-1147, April 1971.
- (12) A.S. Grove, "Physics and Technology of Semiconductor Devices", New York, Wiley, 1967.
- (13) S.M. Sze, "Physics of Semiconductor Devices", 2nd Edition, New York, Wiley, 1981.

- (14) E.H. Nicollian and J.R. Brews, "MOS Physics and Technology", New York, Wiley, 1982.
- (15) W. Feller, "An Introduction to Probability Theory and Its Applications, Vol. I", New York, Wiley, 1957, pp.146.
- (16) A. Papoulis, "Probability, Random Variables and Stochastic Processes", Tokyo, McGraw-Hill, Kogakusha, 1965.
- (17) M. Simard-Normandin, Private Communication.
- (18) S.K. Ghandhi, "VLSI Fabrication Principles Silicon and Gallium Arsenide", New York, Wiley, 1983, pp.353.
- (19) J.S.T. Huang, J.W. Schrankler, and J.S. Kueng, "Short-Channel Threshold Model for Buried-Channel MOSFET's", IEEE Trans. Electron Devices, Vol.ED-31, December 1984, pp.1889-1895.
- (20) S.C. Sun and J.D. Plummer, "Electron Mobility in Inversion and Accumulation Layers on Thermally Oxidized Silicon Surfaces", IEEE J. Solid-State Circuits, Vol.SC-15, pp.562-573, August 1980.
- (21) J.R. Brews, W. Fitchner, E.H. Nicollian and S.M. Sze, "Generalized Guide for MOSFET Miniaturization", IEEE Electron Dev. Letters, Vol.ED-1, January 1980, pp.2-4.
- (22) J.B. Cecil, "A CMOS 10-Bit D/A Converter", ISSCC Dig. Tech. Papers, pp.196-197, February 1974.
- (23) H.U. Post and K. Waldschmidt, "A High Speed NMOS A/D Converter With A Current Source Array", IEEE J. Solid-State Circuits, Vol.SC-15, pp.295-300, June 1980.
- (24) S. Kelley and D. Ulmer, "A Single Chip CMOS PCM Codec", IEEE J. Solid-State Circuits, Vol.SC-14, pp.54-59, February 1979.
- (25) D.J. Dooley, "A Complete Monolithic 10-Bit D/A Converter", IEEE J. Solid-State Circuits, Vol.SC-8, pp.404-408, December 1973.
- (26) G. Kelson, et al., "A Monolithic 10-Bit Digital to Analog Converter Using Ion Implantation", IEEE J. Solid-State Circuits, Vol.SC-8, pp.396-403, December 1973.
- (27) J. Taub and D. Schilling, "Digital Integrated Electronics", McGraw-Hill, 1977, pp.506.
- (28) T.S. Guy, et al., "A Sixteen-Bit Monolithic Bipolar DAC", IEEE J. Solid-State Circuits, Vol.SC-17, pp.1127-1132, December 1982.
- (29) J.A. Schoeff, "An Inherently Monotonic 12-Bit DAC", IEEE J. Solid-State Circuits, Vol.SC-14, pp.904-911, December 1979.

- (30) R.J. van de Plassche, "Dynamic Element Matching for High Accuracy Monolithic D/A Converters", IEEE J. Solid-State Circuits, Vol.SC-11, pp.795-800, December 1976.
- (31) R.J. van de Plassche, "A Monolithic 14-Bit D/A Converter", IEEE J. Solid-State Circuits, Vol.SC-17, pp.112-117, December 1982.
- (32) P.H. Saul, et al., "An 8-Bit, 5ns Monolithic D/A Converter Subsystem", IEEE J. Solid-State Circuits, Vol.SC-15, pp.1033-1039, December 1980.
- (33) P.H. Saul, et al., "Successive Approximation Analog-to-Digital Conversion at Video Rates", IEEE J. Solid-State Circuits, Vol.SC-16, pp.147-151, June 1981.
- (34) P. Yang and P.K. Chatterjee, "SPICE Modelling for Small Geometry MOSFET Circuits", IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, Vol.CAD-1, pp.169-182, October 1982.
- (35) D.F. Hoeschele, Jr., "Analog-to-Digital/Digital-to-Analog Converter Techniques", John Wiley, New York, 1968.
- (36) S. Kuboki, K. Kato, N. Miyakawa, and K. Matsubara, "Nonlinearity Analysis of Resistor String A/D Converters", IEEE Trans. Circuits and Systems, Vol.CAS-29, pp.383-390, June 1982.
- (37) P.R. Gray and R.G. Meyer, "MOS Operational Amplifier Design - A Tutorial Overview", IEEE J. Solid-State Circuits, Vol.SC-17, pp.969-982, December 1982.
- (38) R.S.C. Cobbold, "Theory and Applications of Field Effect Transistors", pp.261-264, Wiley-Interscience, 1970.
- (39) R. Poujois and J. Borel, "A Low Drift Fully Integrated MOSFET Operational Amplifier", IEEE J. Solid-State Circuits, Vol.SC-13, pp.499-503, August 1978.
- (40) T.D.S. Hamilton, "Handbook of Linear Integrated Electronics for Research", McGraw-Hill Book Company (U.K.) Limited, 1977.
- (41) M.C.W. Coln, "Chopper Stabilization of MOS Operational Amplifiers Using Feed-Forward Techniques", IEEE J. Solid-State Circuits, Vol.SC-16, pp.745-748, December 1981.
- (42) F. Krummenacher, "High Voltage Gain CMOS OTA for Micropower SC Filters", Electron. Lett., Vol.17, pp.160-162, 19th February 1981.
- (43) J.C. Bertails, "Low-Frequency Noise Considerations for MOS Amplifier Design", IEEE J. Solid-State Circuits, Vol.SC-14, pp.773-776, August 1979.

- (44) T.C. Choi, et al., "High-Frequency CMOS Switched Capacitor Filters for Communication Applications", ISSCC Dig. Tech. Papers, pp.246-247, February 25, 1983.
- (45) ICL 7650, Chopper Stabilized Operational Amplifier, Intersil, Inc., 1981.
- (46) W.A. Keşter, "Test Setups Judge Speed of Ultrafast 8-Bit DACs", Electronic Design, pp.207-210, May 14, 1981.
- (47) Analog Devices, Data Acquisition Databook, Vol.1 & 2, 1982.
- (48) V.W.K. Shen and D.A. Hodges, "A 60 ns Glitch-Free NMOS DAC", ISSCC Dig. Tech. Papers, pp.188-189, February 24, 1983.
- (49) P.H. Saul, et al., "An 8b CMOS Video DAC", ISSCC Dig. Tech. Papers, pp.32-33, February 13, 1985.

APPENDIX A

BASIC STATISTICAL RELATIONSHIPS

This appendix is provided as a quick reference for some of the statistical definitions and relationships used in the thesis [16].

The Gaussian or normal probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\eta)^2/2\sigma^2} \quad (A.1)$$

where η is the mean and σ is the standard deviation of the random variable x . The mean or the expected value of x is given by the integral

$$\eta = E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx \quad (A.2)$$

If x is of discrete type, taking values x_n with probability p_n , then

$$E(x) = \sum_n x_n P(x=x_n) = \sum_n x_n p_n \quad (A.3)$$

The mean η of a random variable places the center of gravity of $f(x)$. Another important parameter is its variance σ^2 , defined by

$$\sigma^2 = E((x-\eta)^2) = \int_{-\infty}^{\infty} (x-\eta)^2 f(x) \cdot dx \quad (A.4)$$

This quantity equals the moment of inertia of the probability masses and gives some notion of their spread about η . Its positive square root σ is called the standard deviation. If x is of discrete type, then,

$$\sigma^2 = \sum_n (x_n - n)^2 P(x=x_n) \quad (\text{A.5})$$

From equation (A.4)

$$\begin{aligned} \sigma^2 &= E(x^2 - 2xn + n^2) \\ &= E(x^2) - 2nE(x) + n^2 \\ &= E(x^2) - n^2 \end{aligned}$$

and thus obtain the important relationship

$$\sigma^2 = E(x^2) - E^2(x) \quad (\text{A.6})$$

It is often required to determine the mean and standard deviation of a function of a random variable. For example, with

$$y = g(x)$$

we need to determine $E(y)$ and σ_y . By definition

$$E(y) = \int_{-\infty}^{\infty} y f(y) dy$$

Thus, to determine the expected value of the random variable y from the above, we need to know the density function of y . The following approximate method permits the evaluation of the mean and variance of $g(x)$ in terms of the mean and variance of x , and without having to determine the density function of $g(x)$. It can be shown that

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx \quad (\text{A.7})$$

The proof of this theorem may be found in [16, pp.142]. Now we evaluate the integral in equation (A.7) in terms of the mean and standard deviation of x . First we write the Taylor's series expansion of $g(x)$ about n

$$g(x) = g(n) + g'(n)(x-n) + g''(n) \frac{(x-n)^2}{2!} + g'''(n) \frac{(x-n)^3}{3!} + \dots \quad (\text{A.8})$$

$f(x)$ takes significant values in the interval $n-\sigma$ to $n+\sigma$. If in this interval $g(x)$ is "smooth", then we can neglect derivatives of order 3 and above from equation (A.8). Then

$$g(x) = g(n) + g'(n)(x-n) + g''(n) \frac{(x-n)^2}{2}$$

Applying equation (A.7) we have

$$E(g(x)) = g(n) + g''(n) \frac{\sigma^2}{2} \quad (\text{A.9})$$

We shall now use the above equation to evaluate $E(g^2(x))$. With

$$h(x) = g^2(x)$$

we have

$$h'(x) = 2g'(x)g(x)$$

Using the estimate (A.9) for $E(h(x))$, we obtain

$$E(g^2(x)) = g^2 + [(g')^2 + gg''] \sigma^2 \quad (\text{A.10})$$

where all quantities are evaluated at $x=n$. Taking the square of equation (A.9) and subtracting from (A.10) we have

$$\sigma_g^2 = g^2 + ((g')^2 + gg'') \sigma^2 - (g + g'' \frac{\sigma^2}{2})^2 = [g'(n)]^2 \sigma^2 \quad (A.11)$$

neglecting σ^4 terms. Thus, equations (A.9) and (A.11) give the mean and variance of $g(x)$ in terms of the mean and variance of x .

We can derive similar relationships when g is a function of two random variables x and y . Let $n_x, n_y, \sigma_x^2, \sigma_y^2$ be the means and variances of x and y respectively. Expanding $g(x,y)$ in Taylor's series about (n_x, n_y) we have

$$g(x,y) = g(n_x, n_y) + (x-n_x) \frac{\partial g}{\partial x} + (y-n_y) \frac{\partial g}{\partial y} + \dots$$

If the probability masses $f(x,y)$ are concentrated near their center of gravity (n_x, n_y) and $g(x,y)$ is smooth in the vicinity of this point, then, retaining only up to second order terms, we obtain

$$E(g(x,y)) = g(n_x, n_y) + \frac{1}{2} \left(\frac{\partial^2 g}{\partial x^2} \sigma_x^2 + 2r \frac{\partial^2 g}{\partial x \partial y} \sigma_x \sigma_y + \frac{\partial^2 g}{\partial y^2} \sigma_y^2 \right) \quad (A.12)$$

Using the above estimate for the function $g^2(x,y)$, we have

$$E(g^2(x,y)) = g^2(n_x, n_y) + \left[\left(\frac{\partial g}{\partial x} \right)^2 + g \left(\frac{\partial^2 g}{\partial x^2} \right) \right] \sigma_x^2 + \left[\left(\frac{\partial g}{\partial y} \right)^2 + g \left(\frac{\partial^2 g}{\partial y^2} \right) \right] \sigma_y^2 + 2r \left(\frac{\partial g}{\partial x} \cdot \frac{\partial g}{\partial y} + g \frac{\partial^2 g}{\partial x \partial y} \right) \sigma_x \sigma_y$$

And within second order moments

$$\sigma_{g(x,y)}^2 = \left(\frac{\partial g}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial g}{\partial y} \right)^2 \sigma_y^2 + 2r \frac{\partial g}{\partial x} \frac{\partial g}{\partial y} \sigma_x \sigma_y \quad (A.13)$$

In the above equations, r is the correlation coefficient of x and y defined as

$$r = \frac{E((x - \bar{x})(y - \bar{y}))}{\sigma_x \sigma_y} \quad (A.14)$$

r is a unit less quantity in the range -1 to 1 . If the random variables x and y are independent then $r=0$.

APPENDIX B

DERIVATION OF EQUATION (2.19)

The variance of τ given by equation (2.19) is evaluated as follows [16, pp.324]:

$$\tau = \frac{1}{L \bar{W}} \int_0^L \int_0^{\bar{W}} t(x,y) dx dy \quad (B.1)$$

The square of τ may be written as

$$\begin{aligned} \tau^2 &= \frac{1}{L^2 \bar{W}^2} \int_0^L \int_0^{\bar{W}} t(x_1, y_1) dx_1 dy_1 \int_0^L \int_0^{\bar{W}} t(x_2, y_2) dx_2 dy_2 \\ &= \frac{1}{L^2 \bar{W}^2} \int_0^L \int_0^{\bar{W}} \int_0^L \int_0^{\bar{W}} t(x_1, y_1) t(x_2, y_2) dx_1 dy_1 dx_2 dy_2 \end{aligned}$$

$$\begin{aligned} E(\tau^2) &= \frac{1}{L^2 \bar{W}^2} \int_0^L \int_0^{\bar{W}} \int_0^L \int_0^{\bar{W}} E(t(x_1, y_1) t(x_2, y_2)) \cdot dx_1 dy_1 dx_2 dy_2 \\ &= \frac{1}{L^2 \bar{W}^2} \int_0^L \int_0^{\bar{W}} \int_0^L \int_0^{\bar{W}} R(x_1, x_2; y_1, y_2) dx_1 dy_1 dx_2 dy_2 \end{aligned}$$

where $R(\)$ is the autocorrelation function of $t(\)$. For a stationary process, the autocorrelation depends only on the differences $x_1 - x_2, y_1 - y_2$ and not on the absolute values of x_1, x_2, y_1, y_2 .

Then,

$$E(\tau^2) = \frac{1}{L^2 W^2} \int_0^L \int_0^W \int_0^L \int_0^W R(x_1 - x_2, y_1 - y_2) dx_1 dy_1 dx_2 dy_2 \quad (B.2)$$

The integral in equation (B.2) can be simplified by making suitable transformations. We will illustrate the method for one variable x only just to simplify the writing, and then generalize it to equation (B.2).

Let

$$A = \int_0^L \int_0^L R(x_1 - x_2) dx_1 dx_2 \quad (B.3)$$

Make the transformations

$$x_1 - x_2 = u$$

and

$$x_2 = w$$

The Jacobian of transformation is

$$\begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial w} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial w} \end{vmatrix} = 1$$

and hence $dx_1 dx_2 = du dw$. The limits for u and w should be carefully evaluated. Since $0 \leq x_1 \leq L$ and $0 \leq x_2 \leq L$, we have $-L \leq u \leq L$. The limits for w should be evaluated consistent with $-L \leq u \leq L$. Two cases arise for $0 \leq u \leq L$, $0 \leq w \leq L - u$ and for $-L \leq u \leq 0$, $0 \leq w \leq L + u$. Then, equation (B.3) is

$$\begin{aligned}
A &= \int_{u=-L}^0 \int_{w=0}^{L+u} R(u) \cdot dw \, du + \int_{u=0}^L \int_{w=0}^{L-u} R(u) \cdot dw \, du \\
&= \int_{u=-L}^L \int_{w=0}^{L-|u|} R(u) \cdot dw \cdot du \\
&= \int_{-L}^L (L-|u|) R(u) \cdot du
\end{aligned}$$

(B.4)

Thus the double integral in equation (B.3) is reduced to a single integral. It can be further simplified by approximating $R(u)$

$$R(u) = \text{COV}(t(x), t(x+u)) + \bar{t}^2$$

where $\text{COV}(\)$ is the autocovariance function of $t(x)$. For a system in which the $t(x)$ are weakly correlated;

$$\begin{aligned}
\text{COV}(t(x), t(x+u)) &= \sigma_t^2 \quad \text{for } u \rightarrow 0 \\
&= 0 \quad \text{otherwise.}
\end{aligned}$$

Then equation (B.4) may be written as

$$\begin{aligned}
A &= \int_{-d}^d (L-|u|) \cdot \sigma_t^2 \cdot du + L^2 \bar{t}^2 \\
&= 2L d_1 \sigma_t^2 + L^2 \bar{t}^2
\end{aligned}$$

where d_1 is the region over which the autocovariance function has any significant value. Carrying out a similar procedure for two variables equation (B.2) is simplified to

$$E(\tau^2) = \frac{1}{L^2 \bar{W}^2} (2L \bar{W} \cdot d_1^2 \sigma_t^2 + L^2 \bar{W}^2 \bar{t}^2)$$

$$= 2 \frac{\sigma_t^2}{L \bar{W}} \cdot d_1^2 + \bar{t}^2$$

Therefore

$$\sigma_\tau^2 = \frac{1}{L \bar{W}} \sigma_t^2 d_1^2$$

(B.5)

END

FIN

2 04 87