THE MEASUREMENT OF ELECTRICAL PARAMETERS AND TRACE IMPURITY EFFECTS IN MOS CAPACITORS

Thesis submitted by

IAN GRANT McGILLIVRAY

for the degree of

Doctor of Philosophy

Edinburgh Microfabrication Facility
Department of Electrical Engineering
University of Edinburgh

1987

DECLARATION

I declare that all the work done in this thesis is entirely my own except where otherwise indicated.

This thesis is a result of a CASE project with Hughs Microelectronics (Glenrothes) Ltd.

ABSTRACT

A wide ranging review and development of MOS capacitor electrical measurements is presented. Advances are reported in the high frequency CV method, pulsed CV measurements, the determination of interface state densities and calculation of the minority carrier generation lifetime.

A software package called EDUCATES (Edinburgh University Capacitor Test Software) has been written. This provides a comprehensive analysis of the MOS capacitor using accurate measurements implemented in a totally automated manner.

EDUCATES was used to investigate the electrical effects of trace levels of metal ions present in hydrofluoric acid which is used for pre-gate oxidation cleaning. It was found that there were no significant effects for concentrations of up to 5 p.p.m. Calcium, Chromium, Cobalt, Lead and Nickel although increases in the fixed and mobile charge densities were observed in iron contaminated oxide.

ACKNOWLEDGEMENTS

I would like to thank Professor J.M. Robertson for devising and supervising this CASE PhD project, and I am also indebted to Dr. A.J. Walton for his supervision.

I am grateful to the SERC for providing financial assistance.

I would like to acknowledge the support of Hughes Microelectronics (Glenrothes) Ltd. who collaborated in this project and provided me with additional financial support. I am particularly grateful for the assistance that I have received from Dr A. Hawrilyw my industrial supervisor.

I am grateful to Riedel-de-Hahn (Gmbh.) for the supply of chemicals. I am particularly indebted to Dr. W. Sievert of Riedel.

It would have been impossible to produce devices and test equipment for this project without the help of the staff at the Edinburgh Microfabrication Facility and at Hughes Microelectronics. I am sure the following list is not complete but I am grateful to Messrs A.M. Gundlach, R. Cambell, Z. Chen, M. Fallon, J. Fraser, W. Gammie, G. Gibbs, J. Gow, D. Hope, J. Laurie, J. S. Meek, B. Nielson, A. Ruthven, G. Saunders, J.T.M. Stevenson, G. White, and all the women in Hughes' clean room.

I have greatly appreciated the help in a variety of ways that I have had from my fellow PhD students, namely Messrs A. Buttar, E. Cameron, A. Gribben, G. Maxwell, and J. Serack.

Last, but certainly not least, I would like to acknowledge the assistance of Mrs E. Paterson, Miss J. Blanch and Miss J.M. Anderson in preparing this and other documents.

CONTENTS

CHAPTER ONE: INTRODUCTION 1

CHAPTER TWO: INTEGRATED CIRCUIT MANUFACTURING: A REVIEW

2.1 Introduction 7
2.2 Integrated Circuits 8
2.2.1 The Origin of the Integrated Circuit 8
2.2.2 NMOS Processing 8
2.2.2.1 EMF NMOS Process 9
2.2.2.2 Fabrication Technology 11
2.2.3 VLSI Processing 13
2.3 Thermal Oxidation and Cleanliness 15
2.3.1 Thermal Oxidation of Silicon: Kinetics and Technology 15
2.3.2 Wafer and Furnace Cleaning 23
2.4 Conclusions 28

CHAPTER THREE: THE METAL OXIDE SEMICONDUCTOR CAPACITOR AND ITS HIGH FREQUENCY CV CHARACTERISTIC

3.1 Introduction 29
3.2 Definition of Potentials and Bias Regimes of the MOSC 29
3.3 Ideal MOSC CV Characteristic at High Frequencies 34
3.4 Non-Ideal MOSC 41
3.5 Fabrication and Measurement of MOSC Samples 43
3.5.1 Sample Preparation 43
3.5.1.1 Gate Material 46
3.5.1.2 Gate Definition 46
3.5.1.3 Ohmic Back Contact 47
3.5.2 Measuring Equipment 48
3.5.2.1 Probe Needles, Wafer Chuck and Probing Box 48
3.5.2.2 Capacitance Meters 49
3.5.3 Automation of CV Measurements 50
3.6 Parameter Extraction from the High Frequency CV Characteristic 52
3.6.1 Calculation of Oxide Thickness 52
3.6.2 Determination of N_{sub} 56
3.6.3 Flatband Capacitance and Flatband Voltage 62
3.6.4 Barrier Heights, Work Functions and Work Function Differences 64
3.6.5 Oxide Charges 67
3.6.6 Threshold Voltage 72
3.7 Analysis of the CV Characteristic of a Non-Uniformly Doped MOSC 73
3.8 Control of the MOS Integrated Circuit Fabrication Process 74
3.8.1 Process Control Example 75
3.9 Measurement of Mobile Ion Density 79
3.10 CV Characteristics of MOSCs with Two Capacitive Contacts 82
3.10.1 Comparison of Capacitive Contact Measurements with Theory 84
3.10.2 Conclusions on Capacitive Contacts Measurements 86
3.11 Determination of C_{min} from the High Frequency CV Characteristic of MOSCs with High Carrier Lifetimes 86

3.11.1 Conventional DC Bias Sweep for High Frequency CV Measurements 87
3.11.2 Measurement of Equilibrium C_{min} for high Lifetime Samples 87
3.11.3 Alternative Methods of Determining C_{min}, N_{sub} and V_{fb} 92
3.11.4 Conclusions on the Measurement of Parameters of High Lifetime MOSCs 93
3.12 Conclusions 94

CHAPTER FOUR: INTERFACE TRAPS

4.1 Introduction 95
4.2 The Effect of Interface Traps on MOSC Characteristics 97
4.3 Determination of G_p from the Measured Conductance 102
104
4.4 Methods of Determining $D_{it}(E)$ and ψ_s-V_G from Admittance Measurements 104
4.4.1 The Low Frequency CV Method 104
4.4.2 The High Frequency CV Method 112
4.4.3 The Combined High and Low Frequency CV Method 115
4.4.4 The Conductance Methods 117
4.5 The Accuracy of MOSC Interface Trap Measurements 123
4.6 Post Metalisation Annealing of AlSi Gate MOSCS 128
4.7 Conclusions 132

CHAPTER FIVE: DOPING PROFILES IN SILICON

5.1 Introduction 133
5.2 Accurate Measurement of the Complete Doping Profile 134
5.2.1 Data Manipulation 134
5.2.2 Quantitative Error Assessment 138
5.2.3 Other Factors Effecting Doping Profile Measurements 142
5.3 Ion Implantation Dose Measurement 149
5.4 Conclusions 151

CHAPTER SIX: MINORITY CARRIER LIFETIMES IN SILICON

6.1 Introduction 152
6.2 Leakage Currents in the Deep-Depleted MOSC 153
6.3 Techniques for Measuring Minority Carrier Generation Lifetime 155
6.3.1 Accurate Implementation of the Zerbst Method 158
6.4 Generation Lifetime in Non-Uniformly Doped Silicon 168

CHAPTER SEVEN: EDUCATES (EDINBURGH UNIVERSITY CAPACITOR TEST SOFTWARE)

7.1 Introduction 172
7.2 EDUCATES Software 173

7.2.1 Program Access and Structure 173
7.2.2 Measurement Programs 177
7.2.3 Automatic Mode 184
7.3 EDUCATES System Hardware 186
7.4 Discussion 188
7.4.1 Current Software Status 188
7.4.2 Future Development 188
7.5 Conclusions 189

CHAPTER EIGHT: THE ELECTRICAL EFFECTS ON MOS CAPACITORS OF TRACE LEVELS OF DISSOLVED IMPURITIES IN HYDROFLUORIC ACID

8.1 Introduction 190
8.2 Measuring Impurity Effects 194
8.3 Preparations for Contamination Experiments 195
8.4 Contamination Experiments 198
8.4.1 Preparation of Contaminated 5%HF and MOSCs 198
8.4.2 Measurements 199
8.5 Results of Contamination Experiments 201
8.5.1 Calcium 205
8.5.2 Iron 205
8.5.3 Nickel 206
8.5.4 Chromium 207
8.5.5 Cobalt 208
8.5.6 Lead 208
8.6 Discussion 208
8.7 Conclusions 209

CHAPTER NINE: CONCLUSIONS 211

REFERENCES 215

CHAPTER ONE

INTRODUCTION

Metal Oxide Semiconductor, (MOS), Integrated Circuits, (ICs), are currently the most commercially important and sophisticated microelectronic components[1]. This situation is likely to endure for the forseeable future. The manufacturing technology for producing MOS devices has been extensively improved since Grove, Deal and Snow[2,3] first developed a commercially viable fabrication process in 1963.

Chapter 2 of this thesis reviews the current state-of-the-art in device processing with particular emphasis given to those processes which are of particular relevance to the subject matter of later chapters. Grove et al's first process[2,3] was made possible when it was discovered that trace levels of contamination, (later identified as sodium), were responsible for the instabilities in the MOS Field Effect Transistors (MOSFETS), which were being fabricated at that time. The instabilities were monitored and reduced by measuring the electrical properties of MOS Capacitors, (MOSCS), because they were far simpler to fabricate and analyse than the MOSFET, (which was the active circuit element), whilst they still allowed determination of all the critical factors.

From the early 60's to the present day many new fabrication processes have been developed, and also new devices have been introduced which have additional critical electrical parameters, for example minority carrier lifetime is critical in dynamic memories[4]. New measurements have been developed to measure such parameters, however the basic framework is still the same as that used by Grove et al., i.e. detection of an effect with a MOSC measurement, followed by the elimination or control of

that effect using MOSC experiments.

It was the aim of the work reported here, to investigate certain effects of trace levels of heavy metal ions, and to develop MOSC measurements into a more accurate, reliable and usable form of analysis. It is hoped that the latter effort will help alter the poor reputation that this author has found† such measurements have with MOS process and device engineers.

Personal experience† has shown that one major reason for this is that MOSC measurements are considerably more difficult to implement in practice, than would appear to be the case, when they are discussed in textbooks or scientific papers. The apparent simplicity leads to a situation where engineers perform experiments which produce patently incorrect results, thereby undermining their confidence in such techniques.

Another problem with MOSC measurements is the bewildering variety of techniques which have been proposed in the literature for measuring certain parameters. For example a recent review of Schroder and Kang[5] on the measurement of minority carrier lifetime found over 20 different methods. In their review and for other measurements which they do not cover the accuracy varies from excellent to so poor as to be useless. Unfortunately when new methods are introduced, special assumptions, samples or selected results make techniques appear more viable than they ever are for general use.

The introduction of automated MOSC measuring systems‡ in the late 1970's[6, 7, 8], ought to have brought about a substantial improvement in the

† This relates to views expressed by process and device engineers at a number of companies in particular Hughes (my sponsors), Motorola Ltd. of East Kilbride, Analogue Devices of Limerick, and Thorn EMI central research labs. I have also been able to obtain a broad view of affairs from engineers and academics I have met when they were visiting the EMF, and from my own recent visit to the USA.

‡ These are colloquially referred to as "CV systems", i.e. Capacitance Voltage systems, even although the range of their capability is often much greater than this, including measurements as a function of frequency or time, and also measurement of other quantities, such as ac conductance (G), and dc current (I).

unsatisfactory state of affairs, however this didn't happen to any appreciable extent, because they had their own substantial failings. These included a requirement to have a skilled operator initiate and supervise measurements, since important parts were not fully automated, and also the systems produced results even when there a fundamental error in the measuring technique, such as measuring inversion layer properties in the presence of light. Of course automation does nothing to improve the situation where there is a range of measuring techniques available.

In this thesis a range of established measurements has been developed to a point where they can be incorporated into a fully automated and accurate CV system. This system, which consists of a set of linked computer programs that drive commercially produced hardware[6] is described in chapter 7. In chapters 3-6 the range of measurements which this covers are critically examined and developed.

Since the standard texts of Nicollian and Brews[9], and Sze[10], more than adequately cover most of the device physics underpinning the work in these chapters, only the important results are quoted, because compared to these works this thesis aims to take a radically different perspective. They concentrate on detailed descriptions, (running to approximately 600 pages in the case of Nicollian and Brews), of device physics, largely dealing with concerns and equipment that were of interest in the late 60's and early 70's. In contrast this current work concentrates on the problems and applications of MOSC measurements in the latter half of the 80's, using the best measuring equipment currently available.

In chapter 3 the widely known high frequency MOS capacitor CV analysis is reviewed. The theory and basic shape of ideal MOS capacitor CV curves is described, to provide a basis for the study of curves measured for real capacitors. Following this a wide range of problems with the measurement and analysis of real capacitors is discussed. These problems include choosing the correct equivalent circuit for

measurements, sample preparation, limitations of work function 'constants', and the measuring environment. After this, examples of the application of high frequency CV measurements to a manufacturing process for the control of both mobile and immobile oxide charges is discussed. These types of examples are serious omissions from many treatments (eg.[9, 10, 11]), of the CV analysis, because the engineer in production is best helped by examples. In the penultimate section the measurement of the high frequency CV characteristic with high minority carrier lifetime is examined, and two methods proposed for parameter determination in a time-efficient manner. Finally it is shown that the dual top contact electrode, (normally formed with two mercury electrodes in a ring dot configuration), is a fundamentally unsound method with which to measure high frequency CV characteristics.

In chapter 4, the basis of capacitance and conductance determinations of interface trap densities is reviewed. The conductance method has been very rarely used in the past because it has been considered too time-consuming[9]. For example when this author visited Bell Laboratories† where the conductance method originated and has been developed, it was found that only the low frequency method for most interface trap density evaluations was used there, since there was no system with which to implement an automated conductance analysis. It is shown in the present work that with a simple modification of a method recently proposed by Brews[12] for analysing conductance characteristics, that it is possible to determine densities at seven bandgap energies per sample in less than 20 minutes. This new rapid conductance method is then shown to have superior range and accuracy, compared to all three of the capacitance methods which have traditionally been used for trap density determinations. To end the chapter, the power of the new conductance method is shown, with an investigation of a subtle difference between the annealing of aluminium gate and aluminium-silicon

† Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey, U.S.A.

gate MOSCS.

Many MOSC measurements are not as widely used as they might be because they have been incorrectly condemned as inaccurate in the past. An example is statements like,

"this technique (pulsed CV doping profile determination
of implant dose) could not be used for process monitoring."

This sounds very authoritative since it originates from IBM central research labs,‡ however it will be shown in chapter 5 that by using a voltage sweep which allows for the error in modern digital equipment, noise-free profiles can be determined that are suitable for either dose monitoring or for evaluating process simulation.

Of all the MOSC parameters, the most difficult to measure accurately and the one for which there is the greatest variety of measurement techniques is the minority carrier generation lifetime. In chapter 6 it is argued that where accuracy is paramount, a modification[13] of the Zerbst analysis[14] of C-t transients is the best technique to use. The drawbacks attributed to this method in a recent review[5], no longer occur when a fully automated measuring system is used. In this work full automation is achieved by accessing profile information determined using the methods in chapters 3 and 5, and also the equilibrium minimum capacitance using the method described in chapter 3, combined with newly developed methods of implementing the data acquisition in a time-efficient and noise-free manner. This new system is used for measuring lifetimes in chapter 8, and also to measure the carrier lifetime in ion implanted samples. This latter measurement was previously consider too difficult[5] to perform; however this can now be done routinely. This is of importance because in commercially produced

‡ IBM, Thomas J. Watson Research Centre, Yorktown Heights, New York, U.S.A.

circuits where lifetime is a critical parameter, ion implantation is always used during device processing.

Chapter 7 describes a new software system which provides a significant improvement in the automation and accuracy of MOSC measurements. This system was originally developed with one specific application in mind, however it has not been customised to this application, and for it only a subset of the systems capabilities is used. The application is the determination of the influence of trace impurities present in MOS grade chemicals on MOSC electrical parameters (chapter 8). The current specifications for chemicals are the tightest of any mass produced chemical. These specifications have arisen due to a fear of all unknown impurity effects rather than being based on experimental evidence. This fear of unknown impurities has its roots in the early 60's work of Grove, Deal and Snow[3] that has already been discussed. Over the intervening period to the present it has been common practice to blame all defects and instabilities which do not have a ready explanation on impurities. The present work presents the first known systematic experimental investigation into chemical specifications and their implications. This is valuable because the trend is currently, (and always has been), for MOS circuit manufacturers to demand ever-tighter specifications. Meeting this demand will certainly produce increased chemical costs, whilst possibly being a completely unnecessary move. From this experimental investigation, a new basis for assessing current and future specifications has been developed. Extension of the analysis applied here to other elements or other chemicals is straightforward and could be a means by which a team of workers could produce meaningful specifications.

CHAPTER 2

INTEGRATED CIRCUIT MANUFACTURING: A REVIEW

2.1. Introduction

The development of measurements, and the experiments in this thesis, are motivated, as they have been for other workers in the past, by a desire to produce as an end product, integrated circuits of greater complexity at a lower cost than has been previously possible. In this chapter a review of integrated circuit manufacturing technology is presented to provide the setting for subsequent chapters. The review begins by describing the origins of integrated modern integrated circuit process and a chronology of the major developments which have taken place. Following this, there is a description of EMF integrated circuit fabrication process. This process is typical of the fabrication used for LSI, (Large Scale Integration) products. It illustrates the magnitude of the complexity of even a 5μm process, thereby showing the usefulness of techniques described later for evaluating and controlling individual steps, each one of which must be executed to within a specification if any working devices are to be produced. After discussing LSI processing the additional technology requirements for VLSI, (Very Large Scale Integration) devices, which are the current state-of-the-art, are outlined. In section 2.3 those topics which are of particular relevance to later chapters, namely thermal oxidation and cleaning are looked at in greater detail.

2.2. The Integrated Circuit

2.2.1. The Origins of the Integrated Circuit

In March of 1959 at the Institute of Radio Engineers meeting in New York, when he was announcing the recent 'invention' of the Integrated Circuit (IC) by Jack Kily, Mark Shepard of Texas instruments said, "I consider this to be the most significant development ..... since the commercial availability of the transistor"[15]. The period since then has more than justified his bold statement, for it is with the integrated circuit that systems of previously undreamt complexity have been produced in volume, at low cost and having only a minute physical size. In 1958 a team at Fairchild semiconductors led by Jean Hoerni introduced the PLANAR process[16] for producing silicon bipolar transistors using photolithography and silicon dioxide diffusion barriers. This process represents the starting point from which modern integrated circuit technologies have developed. It is a process ideally suited to low cost, high volume production.

2.2.2. NMOS Processing

NMOS[1, 17] is an integrated circuit technology where the basic active circuit element is the MOSFET. After the problems of passivating the MOSFET had been solved in the mid-1960's, MOS based technology quickly moved to become the dominant commercial integrated circuit because it was better suited to digital applications where the demand for parts and the range of new applications was booming. The first commercial MOS technology was p-channel MOS (PMOS) which could be fabricated without any radically different technology from the then established bipolar processes. In Scotland PMOS was first produced by Hughes Microelectronics and Elliot Automation in 1966[18]. Later with the introduction of ion implantation, (at around 1972 in Scotland), NMOS could be manufactured in volume. This new process allowed very

small and precise changes to be made in the transistor turn-on (threshold) voltage. The higher mobility of the electrons in the n-channel which results in circuits with higher switching speeds, meant that it quickly overtook PMOS in importance. At the present time, complementary-MOS (CMOS) is becoming of increasing importance because of its lower static power requirement than NMOS. However NMOS still has the largest share of the IC market and will continue to be of importance for the forseeable future [see for example Semiconductor International, November, 1985].

2.2.2.1. EMF NMOS Process

Figure 2.1 is an outline of the major steps involved in the EMF 5-micron NMOS process[1]. (5 microns is the minimum dimension of any horizontal feature). The full process runsheet has many other steps and a wealth of detail on each individual step. As with any integrated circuit fabrication process the sequence achieves three things.

(1) Isolation of each active device from all the other active devices.

(2) Active circuit element definition.

(3) Interconnection of active elements to make a circuit.

In the NMOS processes, isolation is achieved with regions of thick field oxide on top of heavily boron-doped silicon. The field regions are defined with the LOCOS process[19] and implantation of boron (figure 2.2(a)).

The active element, (the MOSFET) is fabricated from a thin thermal oxide with heavily phosphorus-doped gate, source and drain regions (figure 2.2(b)). The silicon gate process self-aligns the gate[20,21], source and drain electrodes allowing smaller, faster circuits to be made than is possible with metal gates. Threshold voltages of enhancement (normally off), and depletion (normally on) transistors, are set by low dose boron and arsenic ion implantation respectively.

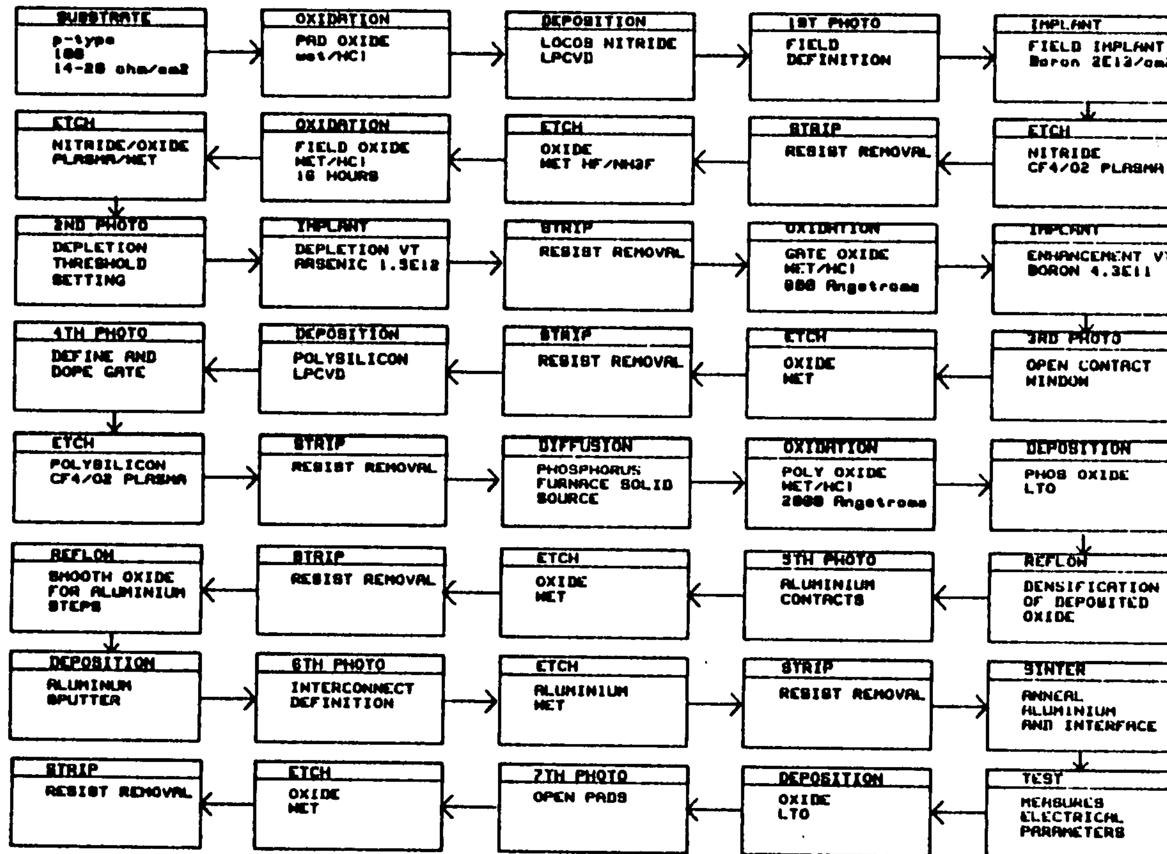


Figure 2.1 Simplified sequence of the major steps in the EMF's 5-micron NMOS integrated circuit fabrication process.

The interconnect which defines the circuit function is placed on top of a thick phosphorus-doped oxide. Contact to the active regions is made via holes etched in this layer and finally the entire circuit is protected by another phosphorus-doped oxide (figure 2.2(c)).

2.2.2.2. Fabrication Technology

The fabrication starts with 3" diameter, <100> orientation, p-type Czochralski silicon wafers[22, 23]. In the EMF these are processed in batches of up to 25 at one time. In industry, batch sizes can be up to 200 wafers of diameter between 3 and 6 inches. This batch production with several hundred IC's per wafer is the major reason for the cost effectiveness of the process.

The process schedule basically consists of repeated deposition or doping of layers, followed by photolithographic transfer patterns and then the etching of that design.

The photolithographic process (steps labeled 'photo' and strip in figure 2.1) has been described by McGillis and King[24, 25]. Some of the other important fabrication steps are:

(1) *Thermal Oxidation[9, 26, 27]*. Thermal oxidation of silicon in O_2 (dry) or H_2O (wet) ambients at temperatures in the range 800-1200^oC creates a stable, passivated oxide which can be used for both gate dielectric and field isolation. HCl is often added to the oxidising species to give a 'cleaner' oxide[28]. (Section 2.3 discusses oxidation in more detail).

(2) *Deposition[29, 30]*. Silicon-dioxide, Silicon-nitride and polysilicon can be deposited in thin uniform layers by Low Pressure Chemical Vapour Deposition (LPCVD). Doped oxides are formed by adding arsine, diborane or phosphene to the reactant gasses. Aluminium is deposited by sputtering an aluminium target with argon ions or by electron beam[31, 32] evaporation. The sputtered

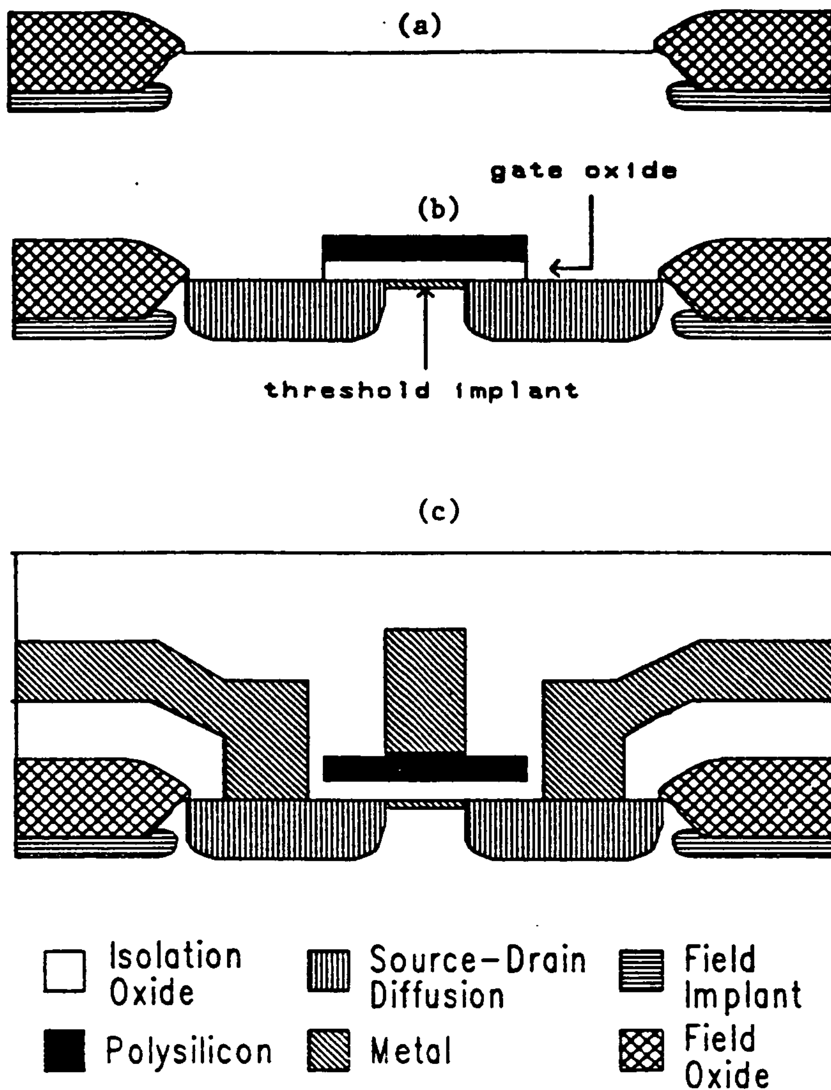


Figure 2.2 The three stages of integrated circuit fabrication: (a) the definition of isolation regions, (b) the fabrication of active devices, (in this case a MOSFET), and (c) connection of devices, (and passivation), to define a circuit function.

aluminium deposits in thin uniform layers on wafers in the sputter chamber.

(3) *Implantation/Diffusion[33, 34, 35, 36].* Ion implantation was the major advance in fabrication technology which allowed the commercial fabrication of NMOS circuits. It involves accelerating dopant ions into the silicon crystal at energies in the range 10-200 keV. Ions are typically stopped in the first two microns of the crystal. Implantation allows very small, well controlled adjustments to be made to the doping density near the oxide-semiconductor interface and hence well-defined enhancement and depletion threshold voltages can be set. Diffusion[36] introduces dopant ions into silicon from the gas phase in a furnace at 900-1200°C. The process is not nearly as controllable as implantation and is used only where heavy doping is required.

(4) *Etching[37, 38, 39, 40].* Etching is the selective removal of areas of a thin film to form geometric patterns. Areas left exposed by photoresist can be etched with wet chemicals, plasmas or reactive ions. The choice of etching method can greatly effect the resulting film edge geometry. Wet etching is isotropic, whilst reactive ions etch anisotropically perpendicular to the wafer surface. (Plasma etching represents a compromise between the two).

(5) High temperature inert gas treatments are used for a variety of purposes such as annealing implantation damage[33, 41], annealing of the oxide-semiconductor interface[9], or for smoothing(reflowing) and densifying deposited glass.

2.2.3. VLSI Processing

A process such as the standard EMF NMOS process can be scaled without any fundamental changes to have a gate length of 3 microns. If the gate and other dimensions are reduced further, then many new technologies and techniques are required[42]. These are;

(1) Lithography with contact or projection printing is not adequate and must be replaced by 'wafer stepper' lithography which can produce better resolved lines and more accurate layer to layer alignment[24, 25].

(2) The standard LOCOS field oxidation process results in a thickening of the gate oxide at the gate edge. This is called the bird's beak effect. A process somewhat similar to LOCOS which uses polysilicon in place of silicon nitride has been shown to give almost beak free field oxide[43]. Other VLSI isolation techniques include SWAMI[44], and trench (moat) isolation[45].

(3) The gate oxide thickness must be reduced to 200-300 angstroms. This creates problems with gate integrity and thickness uniformity. Lowering the oxidation temperature, and added emphasis on wafer cleaning, computer controlled furnace loading and other measures are all involved in improving gate reliability and quality[46, 47, 48]. Recently interest in thin thermally nitrided gates[49] has boomed because of the improved integrity and higher dielectric constant. The higher dielectric constant helps reduce the short channel effect[49].

(4) In order to prevent source and drain electrodes from becoming physically connected due to lateral diffusion during high temperature steps, three process changes need to be implemented. Source and drain are implanted rather than diffused for greater dose and depth control, and slower diffusing arsenic is used as the dopant in preference to phosphorus. In addition long high temperature steps can be replaced by rapid thermal processing[41, 50, 51, 52].

(5) An additional implant is required to raise the punchthrough voltage to well above the operating voltage[53]. This is a medium dose implant placed behind the threshold adjust implant.

(6) The resistance of the gate, source and drain electrodes will limit the speed of a small geometry circuit. This effect can be overcome by using low resistance silicides[54] in place of polysilicon.

Such a process is currently under development in the EMF. The channel length is 1.5 microns, there is a 250 angstrom thermally nitrided gate, source and drain are implanted arsenic, there is a punchthrough suppression implant, lithography is performed with a 10:1 reduction wafer stepper and etching is primarily performed by reactive ion etching. In the near future the addition of rapid thermal annealing and silicides to this process will realise the optimum circuit performance from this process.

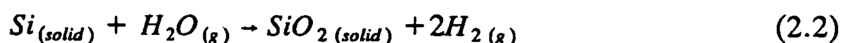
2.3. Oxidation and Cleaning

2.3.1. Thermal Oxidation of Silicon: Kinetics and Technology

The development of clean thermally grown SiO_2 films† described in chapter 1 was the main thrust for the commercialisation of MOS devices. Since that advance, many other insulator and semiconductor combinations have been assessed for MOS IC applications but none has been found to match the excellent dielectric and interface properties of the Si-SiO_2 system[55].

Thermal oxidation of silicon proceeds by one of two chemical reactions[9, 26, 56, 57],

$$Si_{(solid)} + O_{2\,(g)} \rightarrow SiO_{2\,(solid)} \qquad (2.1)$$



$$Si_{(solid)} + H_2O_{(g)} \rightarrow SiO_{2\,(solid)} + 2H_{2\,(g)} \qquad (2.2)$$

† Hereafter these are referred to as "oxide"

which are known as dry and wet oxidation respectively. In integrated circuit manufacturing the oxidation of silicon is performed in so called diffusion furnaces. (This is because this type of furnace was originally used for diffusion of dopants). Microprocessor control of furnace heating, gas flows and wafer loading, all enable the reproducible growth of high quality oxides[46]. It also allows the safe use of pyrogenic water vapour which is produced by combustion of hydrogen in oxygen in the diffusion furnace. Pyrogenic production of steam is much cleaner than the method which was originally used to produce steam for the oxidation of silicon, which involved the bubbling of oxygen through water at 95°C.

The kinetics of these reactions at temperatures between 700 and 1300°C are described by Deal and Grove's well known linear-parabolic relationship[56] for all thicknesses of wet oxide up to 2 microns and for dry oxides of thickness in the range 300 angstroms to 2 microns. The linear-parabolic relationship is,

$$D_{ox} = \frac{A}{2}\left[\left(1+\frac{t+\tau}{A^2/4B}\right)^{½}-1\right], \tag{2.3}$$

where D_{ox} is the oxide thickness, t is the elapsed time for oxide growth, τ is a shift in time to allow for an initial oxide thickness and where A and B are rate-dependent terms. There are two limits of equation 2.3 which are of interest. The first occurs for short oxidation times such that $(t + \tau) \ll A^2/4B$. In this case 2.3 reduces to

$$D_{ox} = \frac{B}{A}(t + \tau). \tag{2.4}$$

This is known as the linear rate law and B/A is the linear rate constant. The other limit occurs at longer times when $(t + \tau) \gg A^2/4B$ for which 2.3 becomes

$$D_{ox} = (Bt)^{1/2} \qquad (2.5)$$

where B is the parabolic rate constant. The detailed expressions for the linear and parabolic rate constants[56] show that different mechanisms control the reaction rate in these two regimes. In the linear regime there is an abundant supply of oxygen available for reaction to take place and the rate of oxide formation is controlled only by the concentration of silicon available for reaction and the reaction rate constant. Here oxide formation is said to be reaction limited. In the parabolic regime the reaction rate exceeds the supply of fresh oxidant at the reaction interface. In this case the reaction rate is controlled by the diffusion of oxidant to the interface and is said to be diffusion limited.

Plotting D_{ox} against $(t + \tau)/D_{ox}$ thickness yields the parabolic rate constant (B) from the slope and -A from the y-axis intercept. For wet oxides τ can be taken as zero. For dry oxides an effective initial oxide thickness D^{i}_{ox} is determined from the intercept at t=0 of the extrapolation of the linear portion of an oxide thickness versus time plot. The linear portion of this plot has a slope, k (=B/A), from which τ can be determined using $\tau = D^{i}_{ox}/k$.

Many of the details of the chemical reactions which lead to SiO_2 formation are unknown, however the effects of a number of experimental factors on reaction rates are well known and these have provided evidence for modelling certain parts of the oxidation process. The major factors affecting oxidation rates on non-degenerate silicon are now reviewed. The commercially important topics of oxidation of degenerate silicon and silicon dopant ion segregation during oxidation have been covered by Katz[26], Nicollian and Brews[9] and Plummer[57].

Table 2.1 lists the linear and table 2.2 the parabolic rate constants for wet and dry oxidation as determined by Deal and Grove[56], and they clearly show that

Temperature (^oC)	Wet B/A ($\mu/hour$)	Dry B/A ($\mu/hour$)	Wet/Dry
920	0.406	0.0208	20
1000	1.27	0.071	18
1100	4.64	0.30	15
1200	14.40	1.12	13

Table 2.1 Linear rate constants for wet and dry oxidation as determined by Deal and Grove.

Temperature (^oC)	Wet B ($\mu^2/hour$)	Dry B ($\mu^2/hour$)	Wet/Dry
920	0.203	0.0049	41
1000	0.287	0.0117	25
1100	0.510	0.027	19
1200	0.720	0.045	16

Table 2.2 Parabolic rate constants for wet and dry oxidation as determined by Deal and Grove.

oxidation in wet ambients is substantially faster in both regimes. This is primarily because the concentration of water vapour in oxide is a factor of 1000 higher than it is for oxygen in oxide. The fact that the diffusivity of water vapour is a factor of 10 less than that of oxygen is of only minor importance compared with the concentration effect in the parabolic (diffusion-limited) regime.

Increased oxidising ambient pressure will increase the concentration of oxidising agent in the oxide and speed its passage to the reaction interface and so both linear and parabolic rate constants are increased at higher pressures[58, 59]. The faster rate of oxide growth may be of commercial importance[60], but only if the problems of temperature uniformity[27], and particles[61] in high pressure furnaces are solved.

The addition of trace quantities of water vapour at concentrations as low as 0.1 ppm can measurably enhance dry oxidation rates and therefore the concentration of water vapour should be either minimised or strictly controlled[62, 63].

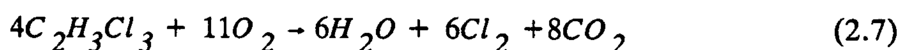
A careful consideration of kinetic data by Irene[64] indicates that both atomic and molecular oxygen are involved in dry oxidation, whilst Spitzer and Ligenza[65] showed that the linear pressure dependence of wet oxidation indicated that water vapour adsorbed from the gas phase was transported interstitially to the reaction interface. There is still considerable controversy as to the charge on the diffusing species though current evidence[9, 26] would favour an uncharged oxidant.

The silicon crystal orientation affects the oxidation rate in the linear growth regime. Ligenza[66] successfully modelled the orientation dependence by showing that the rate was dependent upon the density of silicon atoms in the reacting plane. The model predicted, and measurements were in close agreement with, relative rates of 1.00:1.707:1.227 for the orientations <110>,<311> and <111>. (The model predicts a value of 0.707 for the relative rate with <100> orientation). The experimental results of Irene[62] and Deal and Grove[56] showed the orientation effect to be

independent of the reacting species at temperatures in excess of $700^{o}C$, so steric effects are not important.

The presence of chlorine-containing species such as HCl[26], trichloroethene (TCE)[26], and trichloroethane (TCA)[67], in the oxidising ambient enhances the linear and parabolic rate constants in dry oxidation[68] but reduces them in wet oxidation[69]. A simple mechanism for the reduced rate for wet oxides[69] suggests that the chlorine containing species simply reduces the water vapour pressure. The mechanism of enhancement for dry oxidation is not understood. It has been proposed[68] that water vapour produced by the reaction of HCl, or TCA and O_2, by one of the following reactions,

$$4HCl + O_2 \rightarrow 2H_2O + 2Cl_2 \qquad (2.6)$$



$$4C_2H_3Cl_3 + 11O_2 \rightarrow 6H_2O + 6Cl_2 + 8CO_2 \qquad (2.7)$$

is responsible for the enhancement, however oxidation using pure Cl_2[70] as the source of chlorine also shows enhancement of the rate constants in which case the pres nce of water vapour can be discounted. Another suggestion which may have limited validity is that the chorine etches the silicon crystal thereby reducing the activation energy for the oxidation of silicon which should be observed as an increased linear rate constant whilst not affecting the parabolic constant. The observation of an orientation dependence of chlorine rate enhancement would support this model. However enhancement of the linear rate constant would only appear to be substantial for HCl additions up to 2% whereas the parabolic rate constant increases monatonically with HCl concentration. No complete model of chlorinated oxidation exists and as a result the kinetics have only been empirically modelled[71] as rate enhancement factors η and ϵ, in which case equations 2.4 and 2.5 become,

$$D_{ox} = \eta \frac{B}{A}(t + \tau). \tag{2.8}$$

$$D_{ox} = (\epsilon B t)^{1/2} \tag{2.9}$$

where the enhancement factors must be determined experimentally as a function of HCl concentration and temperature.

The final factor affecting oxidation rate is the presence of other impurities during oxidation. The only identified culprit to date is sodium[63] which enhances both linear and parabolic rate constants at high concentrations. It has been suggested that high sodium concentration will mean that the presence of Na-Si-O in the oxide will increase the concentration of dissolved oxidant sufficiently for it to affect the reaction rate[63].

McGillivray et al.[72], Matlock[73] and Schwettmann et al.[74] have all reported different aspects of the effect of surface cleanliness prior to thermal oxidation on oxidation rate. All three report growth enhancement due to unidentified species. The effects of cleaning on growth rate are reviewed in more detail in chapter 8.

Thin oxides show an enhanced oxidation growth rate for thicknesses of up to 300 angstroms[9, 56, 75], and this covers the entire growth period of an oxide for a VLSI gate application. The thickness over which the effect occurs is independent of temperature. Massoud et al.[76] have recently made a critical examination of a wide range of physical mechanisms which have been proposed to explain enhanced growth and have concluded that only a rather vaguely defined surface layer model can adequately account for all the experimentally observed behavior[77]. Several attempts have been made to empirically model thin oxide growth kinetics[78, 79, 80, 81]. The following model due to Massoud[78], is most likely to be widely used. It appends two factors to the Deal-Grove model and is now incorporated in the major process model program

SUPREM III[71]. Equation 2.7 becomes

$$D_{ox} = \eta \frac{B}{A} \left(1 + C_1 \exp(-D_{ox}/L_1) + C_2 \exp(-D_{ox}/L_2) \right) t. \tag{2.10}$$

L_1 and L_2 are characteristic lengths of 10 and 70 angstroms respectively and the constants C_1 and C_2 are constants dependent upon the activation energy for the fast oxidation process. The two pre-exponential terms are of the form[71, 77],

$$C = C_0 \exp(-E_a/kT) \tag{2.11}$$

where E_a is the activation energy of each of the two enhancements and C_0 is an orientation-dependent constant. Values of E_a, C_{10} and C_{20} have been determined experimentally for a range of orientations[77]. The first relatively short range super-enhanced growth regime can be neglected in the model with a resulting error of approximately 5%[71, 76].

The VLSI gate oxides described in section 2.1 exploit many of those experimental factors described here which reduce oxide growth rate thereby giving improved oxide thickness control.

In addition to controlling oxide thickness, it is important to control the level of oxide charges which remain after growth. Chapter 15 of Nicollian and Brews' book[9] discusses the effects of oxidation ambient, crystal orientation, oxidation temperature, radiation and annealing upon the level of oxide charges. Lai[82] has reviewed some of the electrical properties of oxides grown with oxidation recipes which are typical of modern manufacturing processes. The measurement of oxide charges and a variety of experimental factors affecting their level is the subject of much of the remainder of this thesis.

2.3.2. Wafer and Furnace Cleaning

In 1967 according to Burger and Donovan[83] "attempts to manufacture devices from high quality silicon ..... are beset by perplexing mysteries which provide fertile ground for the rites of the rainmaker, witch doctor and others of that ilk." The exasperation which is comically expressed in this statement reflects the extreme sensitivity of semiconductor devices to trace and even ultra-trace impurities. The DRAM is the most impurity-sensitive MOS circuit, where the basic active element is a storage capacitor which can show unacceptably high levels of leakage with gold concentrations in the silicon surface region of approximately one part in 10^{11}† The problems of contamination have been greatly reduced since 1967 with advances in passivation[9], clean room technology[85], impurity gettering[86], and wafer[87] and furnace cleaning[67]. The methods used to clean wafers and furnaces are now reviewed. In chapter 8 the electrical effects of the chemical purity of the cleaning fluids are examined.

Wafer cleaning[87,88,89,90,91,92,93] is used to remove organic films, inorganic ions and inorganic atoms before high temperature furnace processes where the contaminants can diffuse into device regions or interfere with layer growth, thereby causing functional failure of a device. The wafer cleaning sequence should leave the wafer with less particulate contamination than at its start. Monkowski[94] has reviewed the yield limiting effects of particles which have become of increased importance as device geometries have shrunk.

The detrimental effects of only a relatively few impurities are known, and thus cleaning procedures currently in use aim to reduce the level of all surface contamination to as near zero as possible. The vast majority of cleaning experiments which have been publicly reported[87,88,95,96,97,98,99] rely solely upon a chemical assessment

† one part in 10^{11}. This figure is calculated from data of Richou et al.[84] who related gold concentration to minority carrier lifetimes and the assumption of a lower limit of 100 microseconds for this DRAM parameter.

of removal, using very gross contaminations of wafers and therefore results may not be directly correlated with any real contamination/cleaning requirement. In chapter 8, trace contamination is assessed with electrical measurements, thereby highlighting only those impurity effects which are electrically active at concentrations which could conceivably occur in a fabrication plant. The proof of the effectiveness of the cleans which are discussed next, relies more on the fact that they are regularly used in the manufacture of working electrical devices, than on any published chemical removal data.

Burkmann[91] has stated that the optimum sequence of cleaning steps is:

(1) Gross organic film removal (normally photoresist). *

(2) Residual organic material removal.

(3) Native oxide removal. *

(4) Metal atom and ion removal.

* = optional

This is the procedure adopted in the widely used "RCA" clean described by Kern[87] in an update of the original cleaning process developed by Kern and Poutinen[88]. Table 2.3 details the steps involved in the RCA method. In IC production oxygen plasma stripping of resist is currently favoured[100] for this step because it is cleaner, cheaper and safer than sulphuric-peroxide mix. A commonly used alternative for photoresist removal is fuming nitric acid. This has the advantages over sulphuric-peroxide, that it can be used when underlying metal is exposed and also that it does not deteriorate nearly as rapidly.

Problems with the complete removal of resist can occur, especially after ion-implantation, where heating of the wafer by the ion beam causes the resist to bind more tightly to the underlying layer. In the EMF this 'baked-on' resist is removed by

1. Remove photoresist or other gross organic film in an oxygen plasma stripper or in Caro's acid. Çaro's acid is a mix of H_2SO_4 and H_2O_2 in the ratio 2:1 volumes. It should be used at 140^oC.

2. Remove residual resist in $NH_4OH/H_2O_2/H_2O$ mix in the volume ratio 1:1:5. The wafers should be immersed in the solution at room temperature and then the temperature raised to 80^oC and held there for 10 minutes.

3. Remove the native oxide film with dilute HF. Concentrations from 2-10% are commonly used. Immerse for 15-60 seconds.

4. Remove metallic and ionic contaminants in $HCL/H_2O_2/H_2O$ mix with a volume ratio (1:1:5). Heat to 80^oC then add the wafers for 10 minutes.

5. Rinse and Spin dry then transfer immediately to furnace loader.

Table 2.3 The RCA wafer cleaning recipe.

plasma stripping followed by fuming nitric acid immersion. The "RCA" clean uses $NH_4OH/H_2O_2/H_2O$ for residual organic removal. This is effective but expensive since it has to be used at approximately 80^oC and at that temperature the solution can only be used once because of rapid decomposition of the peroxide[88].

The HF dip to remove native oxide should be omitted in situations where the acid will attack sensitive areas of devices. This modification of the original RCA method gives improved metal ion removal by directly exposing the native silicon surface[87]. Dilute HF is used for a short time to minimise metal plating. Kern[88] recommends a very brief overflow immersion rinse, (30 seconds), in DI water to prevent reoxidation. This author prefers to use a 5 x 45 second cold spray rinse which should remove more fluoride ions[90], whilst not allowing significant reoxidation of the surface.

The extent of rinsing and the method of rinsing (spray, immersion, scrubbing, or spray/immersion) can have a significant effect on the overall performance of any clean[74, 99 101, 102,]. The mechanisms of rinsing effects are not understood and experimental results show marked inconsistencies.

The final RCA clean is used to remove metal atoms and ions. The only other metal clean commonly in use is hot (80^oC) nitric acid, for which experiments performed by this author[103] in an industrial environment showed good results for MOS capacitor electrical parameters. Both of these cleans grow a thin (<20 angstroms) chemical oxide. In some fabrication sequences this is removed with an additional HF dip. This practice is inadvisable since HF causes a higher degree of surface particulate contamination[96] than other cleans and the native silicon surface is more easily recontaminated than the chemically oxidised surface[99].

All cleaning sequences end with a thorough rinse in DI water followed by spin drying in a warm nitrogen ambient.

Throughout the cleaning sequence the quality of DI water[104] and the cleanliness of all utensils[87] is of the utmost importance. Teflon tweezers, unlike metal tweezers, can be cleaned using the same sequence as the wafers and should be used when loading wafers onto a furnace boat after cleaning[87, 105]. In volume production clean batch transfer systems[106] have recently been introduced.

There are three principal methods of implementing cleaning procedures[87]: immersion in baths of liquid, the RCA megasonic system and the FSI spray system. All three methods will give adequate comparable cleaning performance with the major differences between them being cost and safety[87]. In industrial plants the FSI spray system appears to be the dominant method of cleaning in new plants. The specification of commercial cleaning systems would appear rather poor (eg. guaranteeing only a 0.1V mobile ion shift), however this author believes that this is not because the systems are poor, but rather that the cleaning system manufactures do not understand how to assess electrical measurements properly! (They are not alone in this respect).

The work of Pearce and Schmidt[105], Hamsaki[28], and Kriegler[49] implicitly show that wafer cleaning alone is no guarantee that impurities will not find their way into wafers, because the diffusion furnace itself can be the principal source of contamination. Unwanted impurities can enter furnaces in process gasses or by diffusion through the furnace lining whereupon they can enter into wafers by the same mechanism which is used for the deliberate boron, arsenic and phosphorus doping of wafers.

Pearce and Schmidt clearly showed that diffusion through furnace liners is the most serious problem. Several alternative furnace liners to the quartz which is commonly used show better diffusion properties[105, 107]. However they are expensive and do not provide a complete solution. Double walled furnaces work well in inhibiting diffusion[105], but give poorer control over furnace temperature[108]. Since temperature control is of vital importance, particularly for thin oxide growth they are not

widely used. Gassing the furnace with HCl[28, 49, 105] or TCA† is the most widely applied solution to in-situ furnace contamination because it actively removes impurities from the tube. These gasses can also be used during oxidation cycles where chlorine also becomes incorporated into the growing oxide with beneficial effects.

The cleaning procedure for furnaces[67, 101, 109] normally consists of a dummy oxidation run at a temperature at least 100°C higher than the hottest operating temperature of that furnace. The furnace boat and paddle should be in the furnace during cleaning.

2.4. Conclusions

This review of MOS device processing has shown just how complicated their manufacture can be. This helps to explain the requirement for the measurements and analysis presented in later chapters which can be used to assess individual process steps, particularly those which might involve oxidation or contamination, both of which have been looked at in detail in this chapter.

† TCE is not used because it is a carcinogen.

CHAPTER THREE

Metal-Oxide-Semiconductor Capacitor (MOSC)

3.1. Introduction

It is widely accepted that the MOS Capacitor is the single most useful device with which to study the basic electrical properties of MOS system[8, 9, 10, 11, 110, 111, 112, 113, 114]. It can be seen from figure 3.1 that the MOSC forms an important functional part of more complex device structures which are some of the building blocks of monolithic integrated circuits.

The wide utility of the MOSC lies in its simplicity of fabrication and that nearly all the important surface, oxide and silicon parameters can be investigated using relatively simple two terminal measurements.

The useful information which can be derived from this test structure is the subject of this and later chapters. As was explained in chapter 1 the emphasis herein will be on how to measure and interpret the various characteristics with only the most important equations being quoted. The reader should refer to the standard texts for exhaustive proofs or minor detail.

3.2. Definition of Potentials and Bias Regimes of the MOSC*

The energy band diagram for the MOS capacitor is shown in figure 3.2. [115]. The diagram is presented here to define the potentials which will be used in later sections.

*The standard texts[9, 10] use slightly different nomenclatures therefore for clarity these quantities are redefined here.

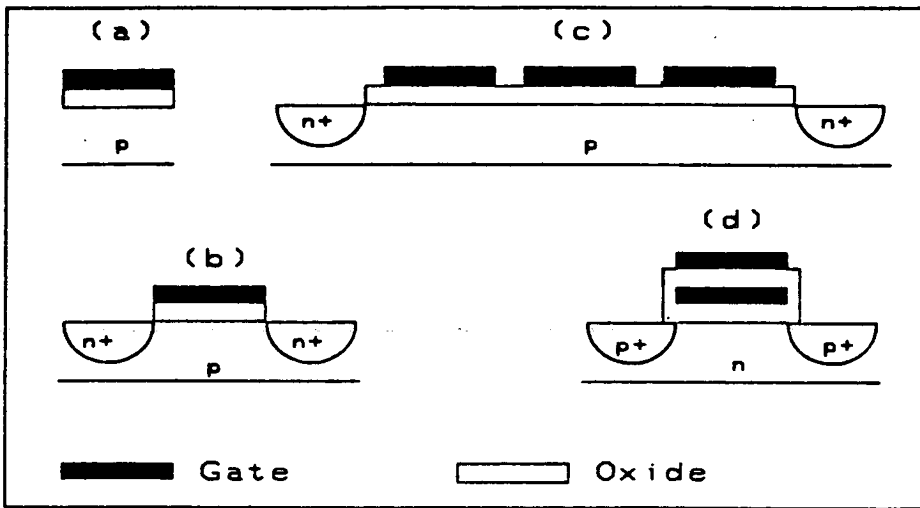


Figure 3.1. The MOS Capacitor (a) and tne devices some of whicn it is a functional component, (b) MOSFET, (c) 3-phase CCD, (d) SAMOS non-volatile memory cell.

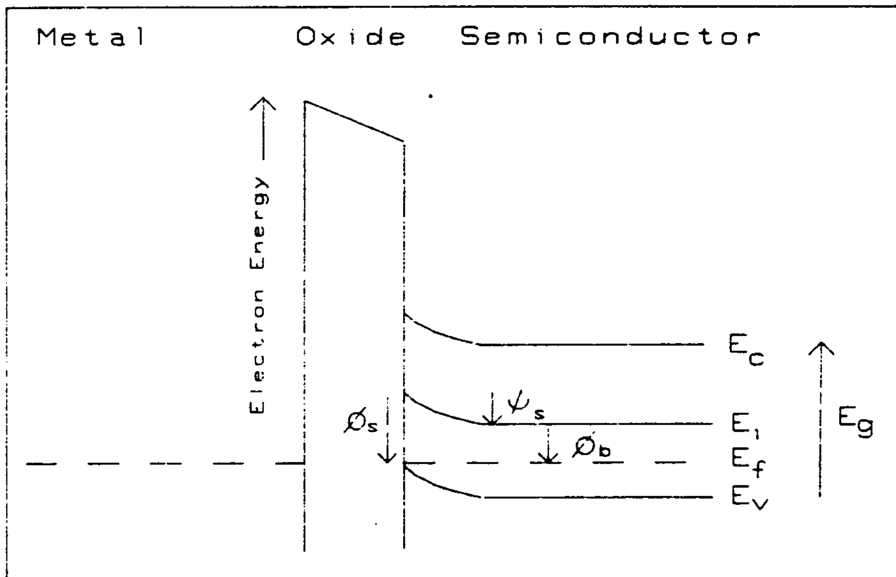


Figure 3.2. MOS Capacitor energy bands.

The bulk potential ϕ_B is defined first by the equation

$$q\phi_B = E_f - E_i \tag{3.1}$$

where E_f is the fermi level and E_i is the intrinsic level energy deep in the bulk of the semiconductor. When a bias is applied across the capacitor, the bands in the silicon bend in order to restore equilibrium. When the bands are bent, another potential, ψ_s the surface potential, can be defined such that

$$\psi_s = \phi_s - \phi_B . \tag{3.2}$$

The operation of the MOS capacitor can be most easily described by scanning surface potential over a range of negative and positive values. Only equations for a MOSC with a p-type semiconductor will be discussed here since the changes in arguments required for an n-type semiconductor relate only to the type of charge and the signs of potentials and do not involve any differences in fundamental concepts.

There are five† basic surface conditions of the uniformly doped MOSC, accumulation($\psi_s<0$), flatbands($\psi_s=0$), depletion ($0<\psi_s<\phi_B$), weak inversion ($\phi_B<\psi_s<2\phi_B$) and finally strong inversion where $\psi_s>2\phi_B$. These conditions along with the associated charge distributions in the semiconductor are illustrated in figure 3.3.

Accumulation occurs when the silicon bands bend upward causing the valence band to move closer to the fermi level. The hole density at the surface at any given

†There is a sixth non-equilibrium state viz. deep depletion.

surface potential is given by the equation

$$p_s = N_A \exp(-\beta\psi_s) \tag{3.3}$$

and the number of minority carriers is found from the pn product relationship

$$pn = n_i^2 \tag{3.4}$$

which holds for any value of p at equilibrium therefore

$$n_s = \frac{n_i^2}{(N_A \exp(-\beta\psi_s))} \tag{3.5}$$

Therefore in accumulation($\psi_s < 0$) there is a very large density of holes at the semiconductor surface and a density of electrons so small that it can effectively be ignored. The flatband condition, (hereafter referred to as "flatbands"), occurs when $\psi_s = 0$, at this point there is no nett potential across the semiconductor, and hence the nett charge density vanishes throughout. As the name flatbands implies, the energy bands in the silicon are flat all the way from the surface to deep in the bulk.

When the bands begin to bend downward the surface is initially in the depletion, ($0 < \psi_s < \phi_B$), regime where the density of holes in the near-surface region is depleted below the density of holes in the bulk of the semiconductor. The hole density vanishes so rapidly when surface potential exceeds a few eV that effectively the only charges present in the near-surface region are fixed ionised acceptor atoms. As surface

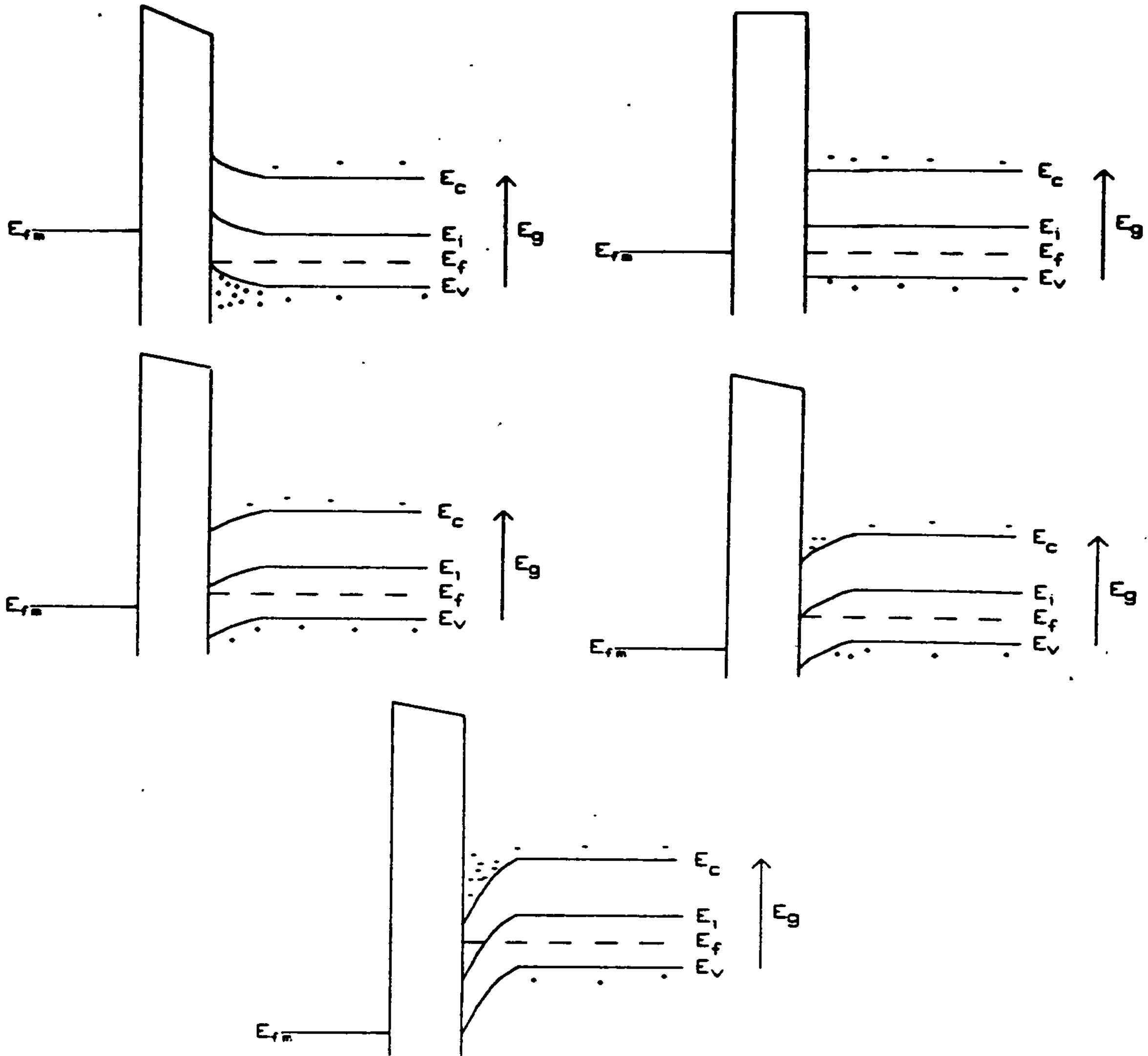


Figure 3.3. The surface conditions of the MOS Capacitor, (a) accumulation, (b) flatbands, (c) depletion, (d) weak inversion, (e) strong inversion.

potential is increased towards ϕ_B, the width of the depletion region increases in order to balance the charge on the gate.

When the surface potential reaches ϕ_B, a small number of electrons drift to the semiconductor surface creating an inversion layer, i.e. a layer of mobile charges forms near the surface which is of the opposite charge to the carriers for that substrate type. The region where ψ_s has a value of between ϕ_B and $2\phi_B$ is called weak inversion because the depletion layer charge exceeds that of the inversion layer.

The final regime is strong inversion ($\psi_s > 2\phi_B$). Here the bands have bent downward sufficiently that the inversion layer charge increases exponentially with increase in ψ_s so that the inversion layer charge becomes far greater than the depletion layer charge.

3.3. MOSC CV Characteristic at high Frequencies†

The high frequency CV (HFCV) characteristic of the MOSC is the most important electrical measurement used in process testing of silicon wafers in *all* MOS IC manufacturing plants and is equally important in research and development environments[9, 11, 113].

The MOSC was first proposed and analysed by a number of workers[116, 117, 118, 119] during the years 1959-1962. At the end of this period, Terman[120], applied the measurement of capacitance as a function of voltage and frequency to study surface states. However, the method did not give reproducible results at that time. In 1965 the classic paper of Grove et. al.[3] was the first to describe MOSC CV characteristics in a format similar to that which is used today. (This work represented a true milestone in the development of MOS processes for it included both

† Frequency in this context refers to the ac bias applied by a capacitance bridge in order to measure differential capacitance.

the practical application of CV measurements and the fabrication of a large quantity of stable MOS capacitors for the first time). An updated version of the classic description of MOSC CV curves is now presented.

In order to understand the CV characteristic of a real MOSC it is first necessary to consider the ideal MOSC which has the following properties, (1) when a dc bias is applied across the MOSC there is no dc current flow through the oxide i.e. the insulator has infinite resistivity, (2) there are no locally unbalanced charges in the oxide or at the oxide-semiconductor interface at zero bias, (3) The metal semiconductor work function is zero.

When these conditions hold, the structure behaves as a classical parallel plate capacitor whose capacitance is given by

$$C_{ox} = \frac{\epsilon_{ox}}{D_{ox}} . \tag{3.6}$$

In addition to the oxide capacitance the charge distributions which result from changes in surface potential result in there being a voltage-variable silicon capacitance. The measured MOS capacitance is a series combination of these two capacitors (figure 3.4). To determine the ideal Capacitance-Voltage (CV) characteristics, the silicon capacitance must be evaluated as a function of gate voltage*. It is simpler to determine silicon capacitance as a function of surface potential and then later use Gauss's law to determine the gate voltage for a given value of ψ_s. The general expression for the silicon capacitance as a function of ψ_s is [9],

* In order to simplify the discussion only the important results will be quoted.

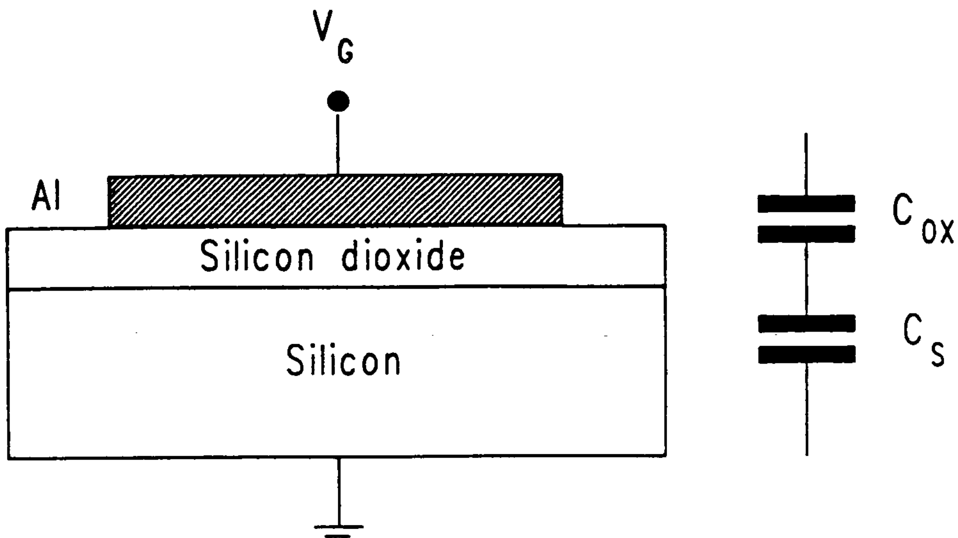


Figure 3.4. MOS Capacitor structure and equivalent circuit.

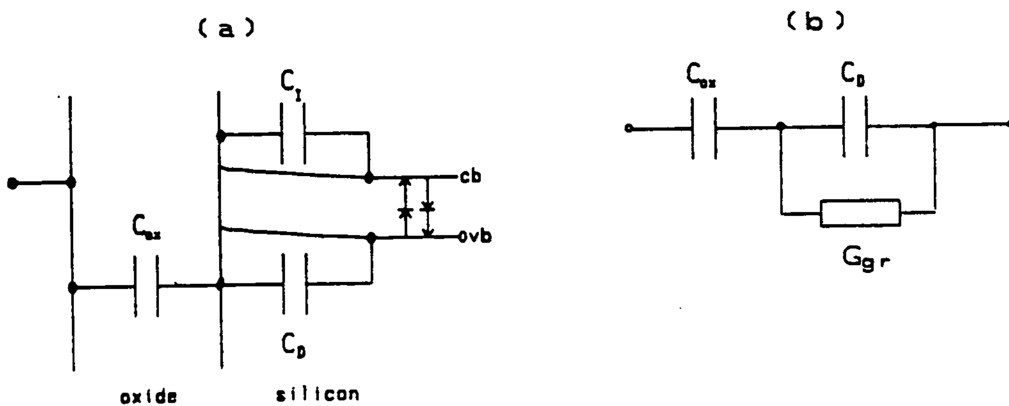


Figure 3.5. (a) Equivalent circuit including generation-recombination centers. (b) Simplification of (a) at high frequency.

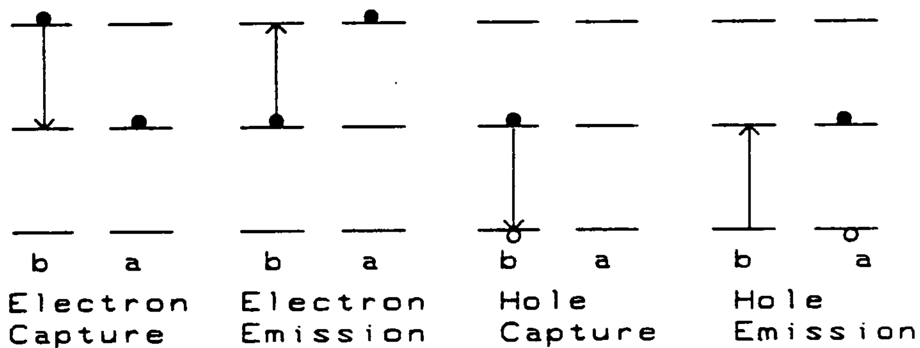


Figure 3.6. The capture and emission processes for charge generation and recombination.

$$C_s = \frac{C_{sfb}}{2^{1/2}} \frac{1-\exp(-\beta\psi_s)+(n_i/N_A)^2\exp(\beta\psi_s))}{[(\exp(\beta\psi_s)-1)+\exp(-\beta\psi_s)+(n_i/N_A)^2\exp(-\beta\psi_s)]^{1/2}} \tag{3.7}$$

where C_{sfb} is the silicon capacitance at flatbands which is defined by

$$C_{sfb} = \frac{\epsilon_{si}}{\lambda} \tag{3.8}$$

where λ is the Debye length and is given by

$$\lambda = \left(\frac{\epsilon_{si} kT}{q^2 N_A} \right)^{1/2}. \tag{3.9}$$

Finally, the measured capacitance C_m of the MOSC is a series combination of C_{ox} and C_s so that

$$\frac{1}{C_m} = \frac{1}{C_{ox}} + \frac{1}{C_s} \tag{3.10}$$

This however is not the capacitance which is measured over a complete range of ψ_s using conventional high frequency capacitance measuring equipment*. To show why this occurs the method by which capacitance is measured and the motion of minority carriers under non-equilibrium conditions are next discussed.

* The characteristic which is generated by equation 3.7 is discussed in chapter 4.

In a capacitance bridge such as the HP4275A LCR, the capacitance is measured by applying a high frequency ac voltage of small amplitude (<50mV) on top of the dc bias across the capacitor. This bridge measures differential capacitance i.e.

$$C_s = \frac{dQ_s}{d\psi_s} \tag{3.11}$$

Since C_s is strongly non-linear with respect to ψ_s in certain bias regimes, the amplitude of the ac variation should be kept as small as reasonably possible. In practice a 15mV amplitude suffices for nearly all applications.

The measured silicon capacitance differs from that predicted by equation 3.7 at high ac frequencies because minority carriers cannot redistribute quickly enough to follow the change in ac bias. Silicon capacitance can be measured in depletion and weak inversion because majority carriers respond instantaneously to ac bias frequencies up to at least 10MHz which is the upper frequency limit of practical MOSC measurements[3, 9].

Minority carriers on the other hand must redistribute by generation-recombination (GR) processes, which takes place at room temperature within the depletion layer[9]. They cannot be supplied by the external circuit where the contact which is ohmic for the majority carriers is blocking for minority carriers. Figure 3.5 illustrates the equivalent circuit of the MOSC including generation recombination centres in the depletion layer. These are the states in the silicon bandgap through which charge must pass to change the relative occupancy of the charge carrier energy bands. Figure 3.6 shows the capture and emission processes which are part of generation and recombination[114]. The role of these mechanisms and their effects on CV characteristics will now be described.

When an ac signal is suplied to the MOSC in inversion in the positive half of the cycle there will be too few electrons in the inversion layer. The capacitor is no longer in equilibrium and so electrons in the conduction band near the interface diffuse toward the surface. This transfers the disturbance from equilibrium away from the interface to a point at which there is a supply of fresh electrons from the valence band. This occurs at room temperature at a point where the band bending is equal but opposite to the bulk potential because generation-recombination only occurs efficiently with generation-recombination centers near midgap[9]. At this point a hole is emitted followed by an electron being emitted by the trap thus promoting an electron into the conduction band. The emitted hole diffuses toward the back contact. In the negative half of the cycle there are too many electrons in the inversion layer and so the recombination process which is similar to the generation process just described must reduce the inversion layer charge.

The GR processes take a finite time to occur. GR centers are impurity atoms with energies near midgap[9], therefore as will be shown later the purer the semiconductor the slower the GR process is at restoring equilibrium. For moderately pure silicon, GR cannot follow ac signals of more than a few tens of Herz and at typical measurement frequencies of 1MHz, it can be assumed that minority carriers do not follow the ac signal at all. This has the effect of making the measured MOS capacitance different from that predicted by 3.7 when the capacitor is in weak and strong inversion. In depletion and weak inversion, the high frequency silicon capacitance (C_s^{HF}) is[9],

$$C_s^{HF} = \frac{C_{sfb}}{2^{1/2}} \frac{[1-\exp(-\beta\psi_s)]}{[\beta\psi_s - 1 + \exp(-\beta\psi_s)]^{1/2}} \tag{3.12}$$

This equation is arrived at by neglecting contributions to the capacitance from minority

carriers.

In strong inversion the depletion layer width ceases to widen as it did in depletion and weak inversion. At high frequency this causes the measured MOS capacitance to settle out to a minimum value. The exact expression for the capacitance in strong inversion is rather complex[9]. However, a simple expression first developed by Linder[119], and then improved on by Brews[121], gives inversion layer capacitance to better than 1.5% accuracy which in practice is perfectly adequate. The Linder-Brews method involves estimating the value of band bending at which strong inversion occurs and then simply setting the capacitance at all other values of ψ_s in strong inversion equal to that at the match point where the match point is found from

$$\psi_{sm} = 2.10\phi_B + 1.33\frac{kT}{q}. \tag{3.13}$$

Thus using equation 3.7 in accumulation, equation 3.12 in depletion and weak inversion and equation 3.13 to find the capacitance in strong inversion the high frequency CV characteristic can be calculated over a full range of ψ_s. The only remaining problem is to relate the surface potential to the applied gate bias.

From Gauss's law, the charge on both the oxide and semiconductor capacitors is equal. Thus if Q_s can be calculated for each value of ψ_s, then Q_{ox} and C_{ox} are known and using V=Q/C the voltage across the oxide can be calculated. Now Q_s can be calculated from

$$Q_s = Sgn(-\beta\psi_s)\frac{\epsilon_s}{\lambda}\left(\frac{kT}{q}\right)F(\beta\psi_s) \tag{3.14}$$

where F is the electric field and is given by

$$F = \left(\frac{N_A}{n_i}\right)^{1/2} [(\beta\psi_s - 1) + \exp(-\beta\psi_s) + (n_i/N_A)^2 \exp(\beta\psi_s)]^{1/2} . \qquad (3.15)$$

Therefore gate voltage, V_g is

$$V_g = \frac{-Q_s}{C_{ox}} + \beta\psi_s . \qquad (3.16)$$

Figure 3.7 shows ideal high frequency CV characteristics for p-type substrates of various uniform doping densities and oxide thicknesses. Real MOS capacitors if carefully fabricated can approach this ideal model. Figure 3.8 shows a a comparison between an ideal curve and an overlayed measured curve.

3.4. Non-Ideal MOS Capacitor

Despite the fact that figure 3.8 showed that MOS capacitors with near ideal behavior can be made, this is only achieved with very careful oxide growth and annealing. Real MOS capacitors can support locally unbalanced charges in the oxide. Insulator charge has the effect of either shifting or distorting the high frequency CV characteristic and is of practical importance in IC manufacturing because it effects transistor parameters such as threshold voltage (V_T) and transconductance (g_m).

In 1980 a joint committee of the IEEE and Electrochemical Society[122] under the chairmanship of B. Deal produced an unambiguous oxide charge classification. This set up a working model of the MOS system which is now the common basis for

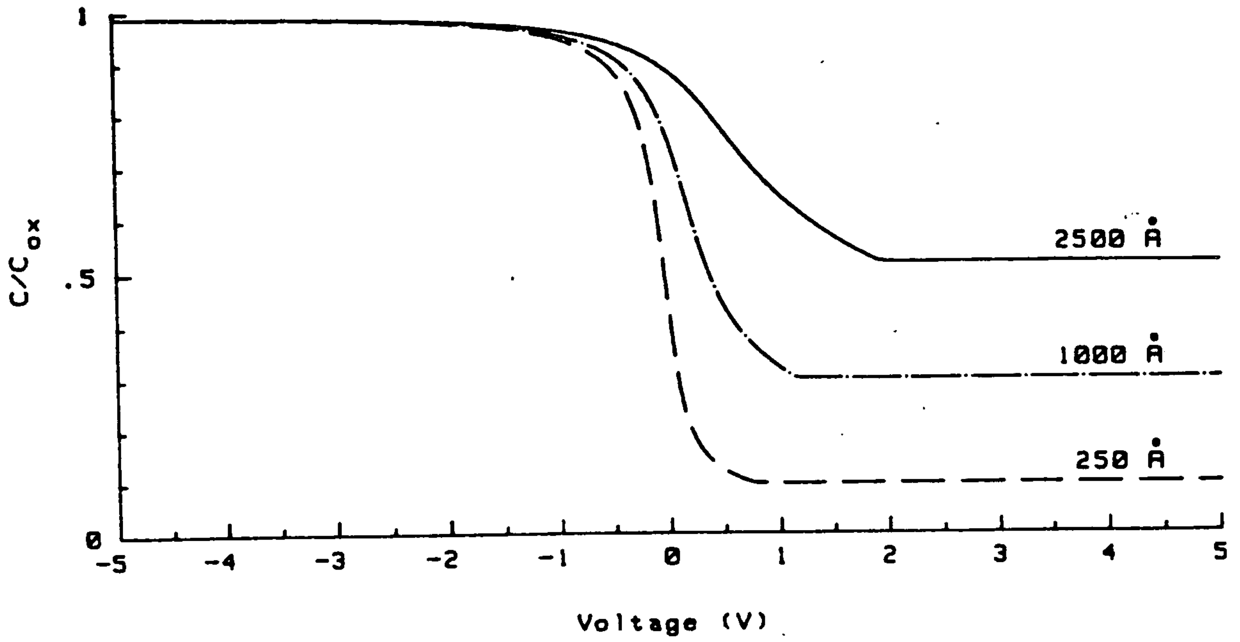


Figure 3.7 (a) Ideal high frequency CV characteristics for a range of oxide thicknesses.

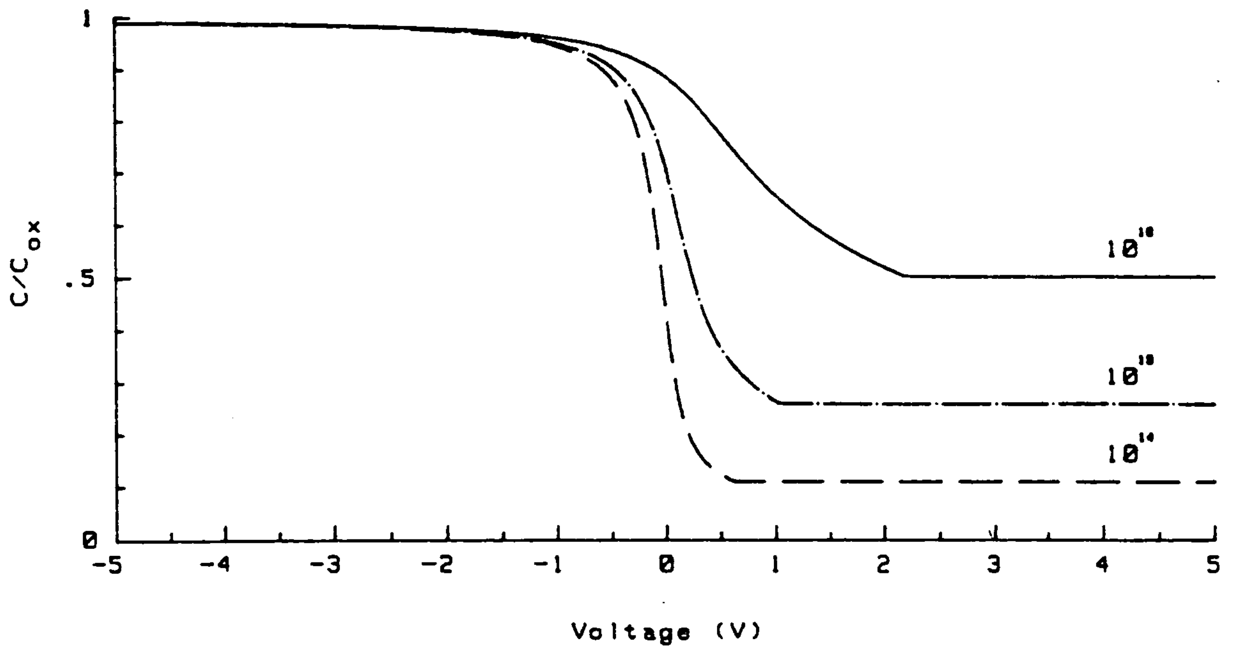


Figure 3.7 (b) Ideal high frequency CV characteristics for a range of substrate doping densities.

discussions of insulator charges. The Deal terminology will be followed in this text.

The four types of charges identified by the committee were,

Q_f	Fixed Oxide Charge
Q_m	Mobile Charge
Q_{it}	Interface Trapped Charge
Q_{ot}	Oxide Trapped Charge

Figure 3.9 shows the spatial location of these charges. In figure 3.10 the effect of a range of fixed oxide charge densities on the CV characteristic are shown. This is an idealised situation because in practice one type of charge seldom occurs on its own. In addition there are other effects which distort CV curves, (1) non- uniform doping profiles, (2) work functions, (3) measurement errors.

3.5. Preparation of MOSC samples and measuring equipment for CV experiments.

Many process engineers in IC fabrication plants are sceptical of the ability of MOSC measurements to provide reliable results. In most cases this is probably because of inadequate sample preparation, measuring equipment, or analysis. In this section the avoidance of the first two problems is discussed.

3.5.1. Fabrication Of MOSC Samples

The following chapters assume that samples are prepared in class 100 cleanrooms[123, 124] or better using modern 'clean' processing equipment using high quality single crystal silicon wafers which have been thermally oxidised at temperatures between 800 and 1200°C.

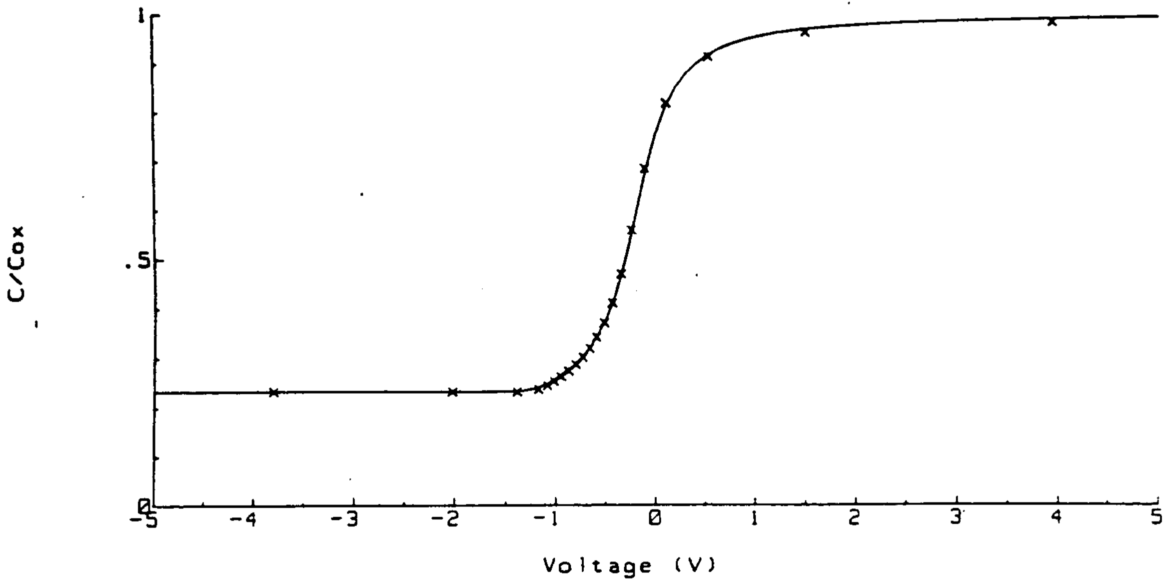


Figure 3.8. A comparison of a measured high frequency CV curve, (solid line), with a corresponding ideal CV characteristic (crosses).

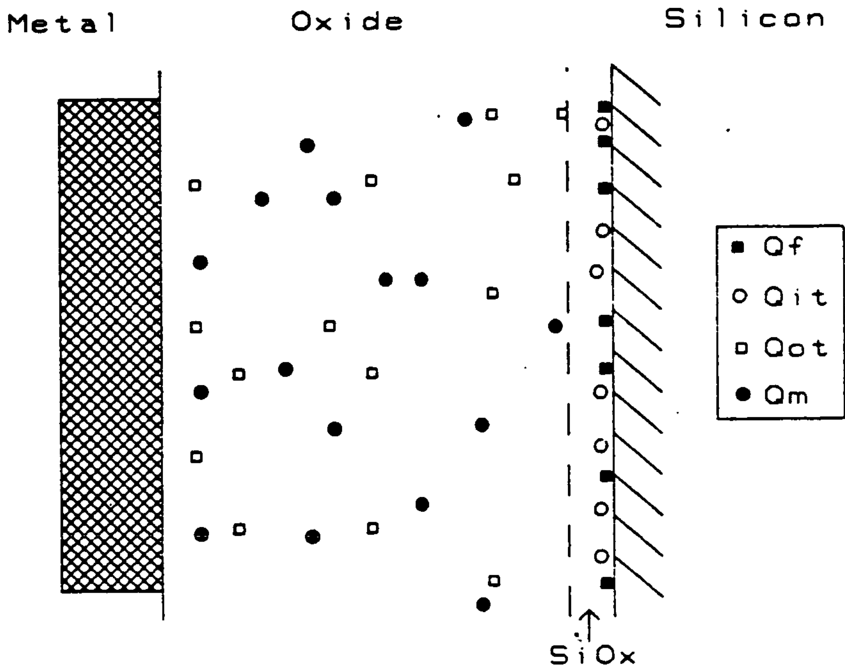


Figure 3.9. Spatial location of oxide charges as classified by Deal.

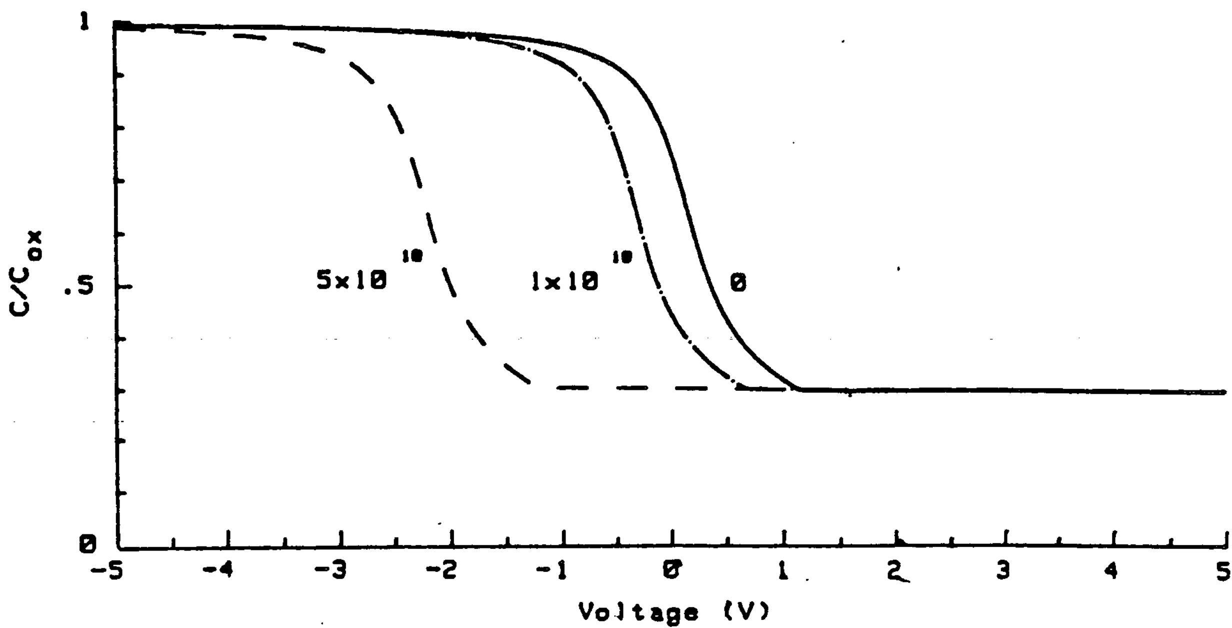


Figure 3.10. High frequency CV characteristics shifted by a range of fixed charge densities.

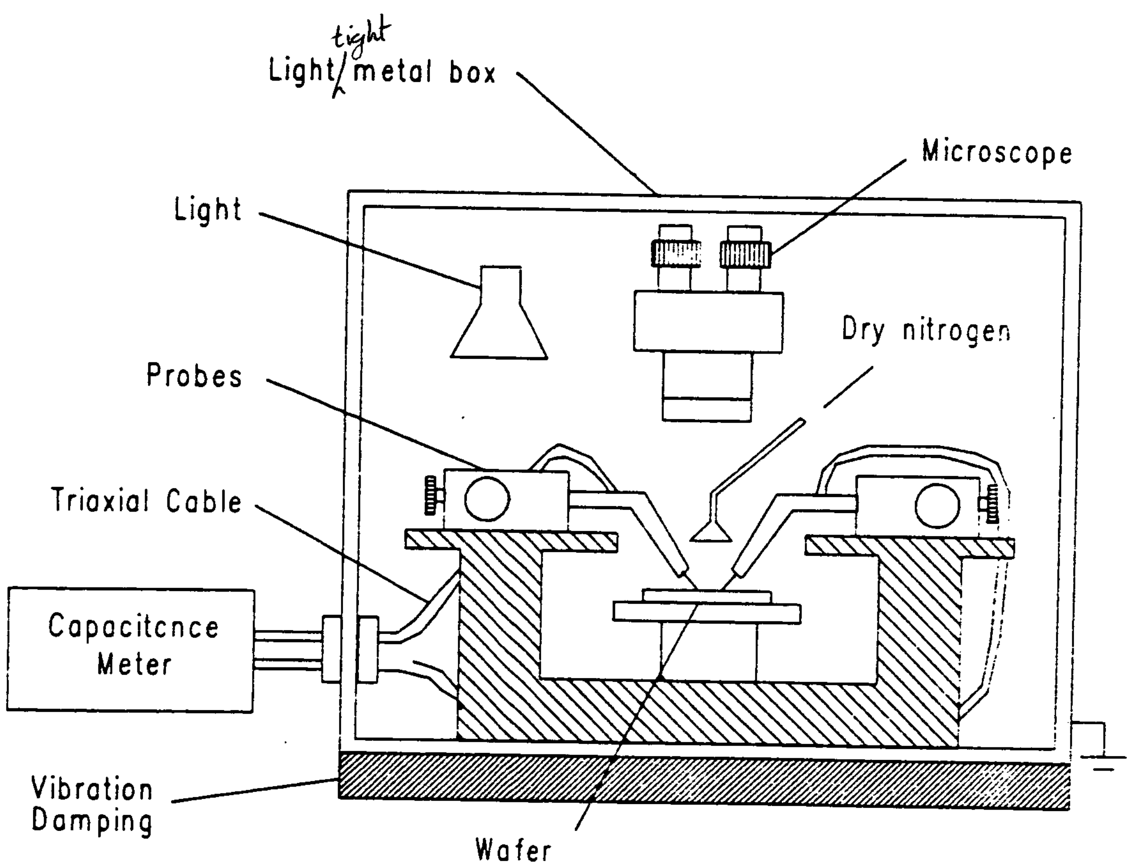


Figure 3.11. Probe station set-up for CV measurements.

3.5.1.1. Gate Material

MOS IC manufacturers use aluminium, polysilicon and silicides as gate material[31, 125]. Many other metals such as gold, magnesium and chromium are suitable as gate electrodes but their use is generally restricted to R&D. Different choices of gate material will result in different MOSC behavior[126, 125], therefore it is important to choose the gate metal carefully in order that experiments give meaningful results.

3.5.1.2. Gate Definition

A photolithographic definition, (chapter 2), is the best method for accurate work since this provides a gate with well defined area and uniform thickness. This is the only practical way to define polysilicon or silicide gates.

Shadow masking[127] is a simpler and quicker alternative where a thin flat plate with drilled holes is firmly attached to the front of the wafer. The plate and wafer are then placed in a metal evaporator/sputterer where the plate masks the majority of the wafer from metal. The exposed areas defined by the holes become the capacitor dots. The area definition is not as good as that obtained lithography since the edge definition is poor because metal will be deposited at points near the hole edge where the wafer and mask are not firmly in contact. In addition metal thickness can vary a great deal across a dot.

The final alternative is to use a liquid mercury column[128, 129]. Again the area is not well defined. The contact is generally maintained by placing the wafer face down onto a vacuum chuck and then a column of mercury within the chuck is formed upwards under pressure making contact with the wafer. This author has found the mercury probe difficult to use. Derived results compare very badly with those measured with lithography or shadow masks. Section 3.10 will show that the two top

contact mercury probe is not suitable for quantitative analysis.

When the area is not well defined, optical thickness measurements can be used to calibrate area, using equation 3.6. This usually still leaves some error particularly if the oxide is thin.

3.5.1.3. Ohmic Back Contact

An ohmic back contact of low resistance† is essential for high frequency CV analysis[3, 9]. Figure 3.10 shows the frequency dependence of the CV characteristics for a sample with a non-ohmic contact. Unfortunately each curve still has the right appearance if viewed in isolation.

One of two tests can be performed to ascertain whether a contact is ohmic or not. The first involves comparing the oxide thickness derived from the accumulation capacitance which has been corrected for series resistance with that derived by optical measurements. If the former value of the thickness is significantly less than the latter then the contact is non-ohmic. The second involves measuring accumulation capacitance corrected for series resistance over a range of frequency (e.g. 10kHz to 100kHz) and if the capacitance is frequency-dependent, the contact is non-ohmic. Care should be taken in the second measurement to ensure that the frequency dependence is not inherent in the measuring equipment. (See section 3.6 for more details).

The first step in making an ohmic contact is to remove any layers present on the back of the wafer to leave a bare silicon back[3]. The stripping of the back is best performed using wet chemical (rather than plasma or reactive ion) etchants since this is unlikely to effect the MOSC properties. A swab can be used to apply etchant to only the back of the wafer.

---

† In this and other sections ohmic is defined as having low, (i.e. $<100\Omega$), series resistance.

Aluminium is probably the most convenient metal to use in a fabrication facility. It provides a good contact particularly if after deposition it is annealed(sintered) at 400 to 500° C. It is also possible to perform additional processing on the MOSC with an Al gate after it has been probe tested since unlike other metals there is little chance of the sample contaminating the processing equipment.

Gold can be used for the back contact but adhesion is poor unless Chromium or Nichrome is evaporated first. Indium/Gallium paste[9] produces an ohmic contact which requires no processing above 300° C. This author has tried using silver-doped epoxy resin, (used to package ICs) as an ohmic contact which requires only room temperature processing but it proved to be unsuitable being messy and readily flaking off.

3.5.2. Measuring Equipment

3.5.2.1. Probe Needles, Wafer Chuck and Probing Box

Probe needles should be sharp tungsten for Al gates in order to pierce the native oxide. For soft metal gates with no native oxide a gold probe wire is adequate. Probes should be mounted on a micromanipulator preferably with x,y and z motions. Connections from the capacitance meter to the probe needles and chuck should be coaxial as far as is possible. Care should be taken not to scratch the gate once its area has been calibrated[130].

The wafer chuck should be gold-coated brass or aluminium mounted upon an insulating material in order to isolate it from ground potential. When the samples are large diameter wafers, mounting the chuck on an XY stage makes probing easier. If high temperature measurements are to be made then the design of the chuck is more complex. DC heating will cause less interference with measurements. All types of chuck should have a vacuum to hold the sample in intimate contact otherwise the

measurement will be affected in a similar way to that for a non-ohmic contact.

Coaxial probes, tungsten probe needles, manipulators, chucks hot chucks, black boxes and complete CV probe stations are available commercially[131]. The electrical properties of the MOSC structure are affected by light and MOSC electrical measurements are sensitive to electrical interference. To remove these problems, measurements should be carried out in a light-tight metal box. Leakage across the surface of the wafer is minimised by blowing a gentle stream of nitrogen across the wafer.

In order to aid the probing of small gate MOSC structures it is helpful to include a projection microscope in the probe station. This can be placed outside or more commonly inside the probe box. Figure 3.11 is a diagramatic view of the EMF CV probe station.

3.5.2.2. Capacitance Meters

There are a number of features required of any capacitance meter which would make it ideally suitable for high frequency CV measurements:

(1) Measurement over a wide absolute range of capacitance to a high degree of accuracy. (Four or five significant figure measurements).

(2) Measurement of ac conductance (G) or alternatively impedance measurement in order that series resistance effects can be calibrated.

(3) Cable capacitance and inductance correction.

(4) Auto-ranging (particularly for thin oxide measurements where C varies markedly over a voltage sweep).

(5) A computer interface preferably of the IEEE-488 type.

(6) Measurements over a range of frequencies.

(7) Control of the amplitude of the ac measuring signal

The HP4275A has all of these desireable qualities and represents the optimum choice of meter for most applications. When high speed capacitance measurement is required or where funds are limited the HP4280A and many meters manufactured by Boonton may be better suited.

3.5.3. Automation of CV Measurement and Analysis

Since the major use of high frequency CV measurements is for routine process monitoring, it is obviously useful to automate the acquisition and analysis of CV data to as great an extent as possible. In 1965 Zaininger[132] discussed the cumbersome nature of a manual point by point determination of the CV characteristic and described apparatus suitable for automating the measurement to give a rapid output of the entire characteristic on an XY plotter or curve tracer. From then until relatively recently, CV plotters were the standard way of obtaining data with the analysis being done manually or sometimes with the aid of charts[10, 110, 133]. This type of system has several disadvantages[7, 8, 134];

(1) The machine must be set up prior to each measurement with a number of dials and switches. This wastes time and can introduce operator error.

(2) The machine operates over fixed voltage and capacitance ranges. For very thick or very thin oxides this can cause difficulties.

(3) Digitisation of data is slow, particularly if it is to be performed accurately, unless an A to D converter is added to the CV plotter output.

A computer-controlled system can overcome all of these problems. In addition CV characteristics can be readily analysed, data acquisition and analysis times are reduced and derived parameter values are more accurate. The main problems of a computer-controlled CV system are:

(1) Software must be written to acquire data and analyse curves. The software must be 'user friendly' so that process operators can perform measurements and must be rugged enough to cope with bad input, operator errors and operator interference.

(2) The numerical analysis of curves must be appropriate to the sample measured and the algorithms used must be error tolerant. e.g. implanted samples should not be analysed as if they were uniformly doped.

Since the late 1970's and the introduction of relatively inexpensive microcomputers a number of computer controlled CV measurement and analysis systems have become available commercially. These range from hybrid analogue-digital systems such as that described by Gordon[7] to the fully digital systems sold by Hewlett-Packard. Shirley[8] has discussed the advantages of the HPIB (GPIB) bus system which is used in the HP4061A test system located in the EMF and the modified version of this system which has been assembled for Hughes Microelectronics[134].

The utility of any computer-controlled system is largely determined by the system software. The subject of automation is returned to in chapter 7, when a sophisticated software package capabable of fully automated measurement of several measurements is described.

UNIV. EDINBURGH

3.6. Parameter Extraction from the High Frequency CV Characteristic

The high frequency CV characteristic has been the major production control measurement of the integrated circuit fabrication industry since the late 1960's. It is an excellent measurement for identifying 1st order non-idealities in the MOS system[3, 10, 11, 110, 135], such as oxide thickness variations, sodium contamination, flatband voltage shifts and variation of threshold voltage.

For more detailed analysis it is also a powerful tool when used in conjunction with other measurements. The more advanced applications will be discussed in chapters 4, 5, 6 and 7. The remainder of this chapter critically examines the high frequency CV method as it is widely applied. There are many pitfalls for the unwary!

3.6.1. Calculation of Oxide Thickness

Oxide thickness can be derived from the oxide capacitance using equation 3.6. The oxide capacitance can be measured at biases well into accumulation where there is no depletion or inversion layer capacitance.

This seems straightforward. However, there are two major causes of error: series resistance and a variation in dielectric constant through the dielectric layer.

Variations of dielectric constant can occur when the oxide is thin or when the gate dielectric is a composite material such as thermally nitrided oxide in these circumstances. When an oxide thickness is calculated from the oxide capacitance the result is an effective oxide thickness $D_{ox}^{'}$, i.e. the thickness of oxide with fixed dielectric constant that would give the same accumulation (oxide) capacitance.

Series resistance[9, 136] (R_S) is the most commonly neglected part of proper sample preparation and is probably the most frequent cause of error in CV measurements. Because of its importance and widespread neglect, this effect is now examined in some detail.

The equivalent circuit changes from figure 3.12(a) to 3.12(b) when a series resistance is present. The primary effect of the extra circuit element is to reduce the measurement oxide capacitance although there is a marginal effect in depletion.

The effect of R_s becomes more important as frequency increases. This can be seen by looking at how C_{ox} is extracted from the measured admittance Y, such that

$$Y = 1/Z \tag{3.17}$$

where Z is the impedance and

$$Z = R_s + \frac{1}{j\omega C_{ox}} \tag{3.18}$$

therefore

$$Y = \frac{1}{R_s + \frac{1}{j\omega C_{ox}}} \tag{3.19}$$

from which it can be shown that

$$C_{ma} = \frac{C_{ox}}{1 + \omega^2 C_{ox}^2 R_s^2} \tag{3.20}$$

where C_{ma} is the measured accumulation capacitance.

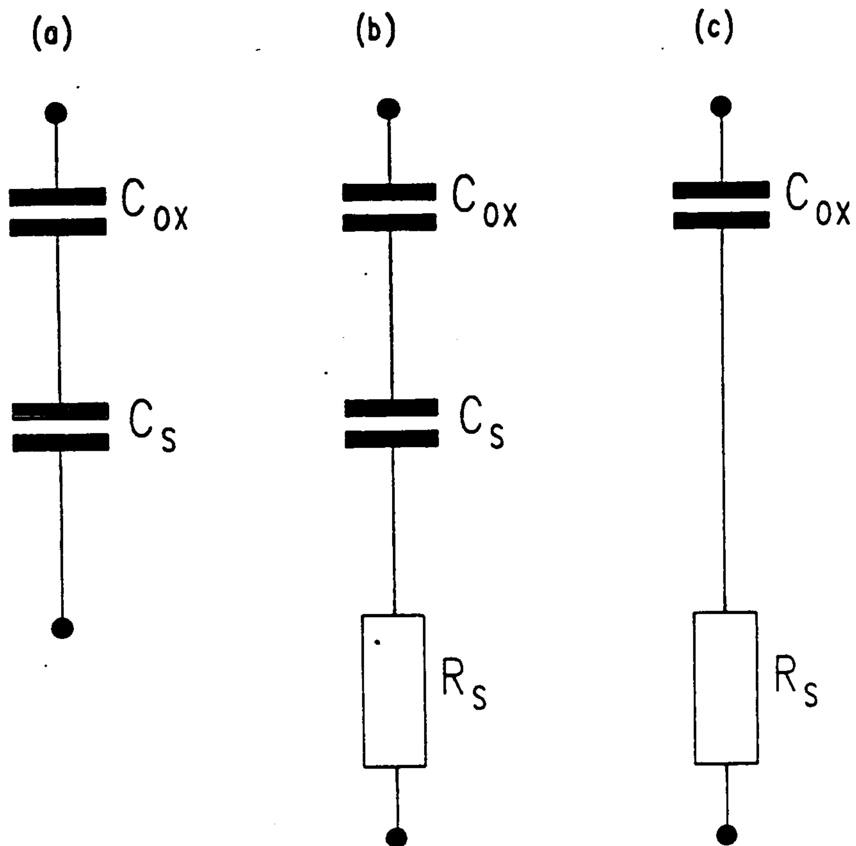


Figure 3.12. Equivalent circuits for, (a) an ideal MOS capacitor, (b) an MOS capacitor with substrate series resistance, (c) circuit (b) biased in accumulation.

It can readily be seen that as R_s increases C_{ma} decreases, and that the effect will be more significant at higher frequencies and when C_{ox} is large.

Figure 3.13 shows the effect of a series resistance of 250 Ω on a measured CV characteristic. The effect of R_s can be reduced in two ways, (1) the sample preparation can be modified and (2) R_s can be measured and corrected for.

R_s is significantly reduced by properly forming an ohmic back contact as described in section 3.5.1. The following factors also reduce R_s; larger capacitor area; thinner substrate and a heavily doped substrate. Many early research workers studied lightly doped silicon using epitaxial layers of lightly doped material on top of heavily doped substrate. The vast majority of substrates currently in use for production control are boron or phosphorus doped at a concentration of approximately 1 x 10^{15} cm^{-3}. With reasonable care about making an ohmic contact, measurement of such wafers should yield reasonable results for this type of substrate at frequencies up to 100 kHz. Special measures are necessary for higher frequencies.

A more sound approach for the accurate measurement of characteristics is to always take account of any resistance that is present. This can be done in two ways. The first involves using an Impedance (Z) rather than Admittance (Y) Bridge. With this measurement configuration, capacitance is measured independently of resistance. The second approach is to use a conventional admittance bridge to measure both capacitance and ac conductance (G). By measuring the conductance, R_s can be calculated and corrected for since G and R_s are related by

$$R_s = \frac{G_{ma}}{G_{ma}^2 + \omega^2 C_{ma}^2} \tag{3.21}$$

where the subscript ma denotes the conductance measured in strong accumulation. The entire CV curve can be corrected using the expression

$$C_c = \frac{(G_m^2 + \omega C_m^2) C_m}{(G_m - (G_m^2 + \omega^2 C_m^2) R_s)^2 + \omega^2 C_m^2} \tag{3.22}$$

where C_c is the corrected capacitance and C_m, G_m are the measured capacitance and conductance.

A practical example of this correction is shown in figure 3.13. The correction works best when the magnitude of the correction is small. This is because large values of R_s such as that for the sample in figure 3.10 are usually accompanied by an additional series capacitance which is not readily accounted for. Shirley[8] proposed a scheme for correcting out the effect of additional capacitance which involved the measurement of two capacitors of differing area, however his own plots show that this correction is inadequate since they still show a frequency dependent oxide capacitance.

Two optical methods[137], elipsometry and interferometry are currently the most popular for determining oxide and other thin film thicknesses since they do not require the fabrication of a MOSC. Values of D_{ox} from these measurements can be used to calibrate gate area if C_{ox} is known or alternatively C_{ox} if area is known.

3.6.2. Determination Of N_{sub} [9, 10, 110]

At the onset of strong inversion, the depletion layer reaches its maximum width, w_{max}. From the parallel plate capacitor equation this width can be defined by

$$w_{max} = \frac{\epsilon_s}{C_{smin}} \tag{3.23}$$

where C_{smin} is the minimum silicon capacitance. Rewriting this equation in terms of measurable quantities we have

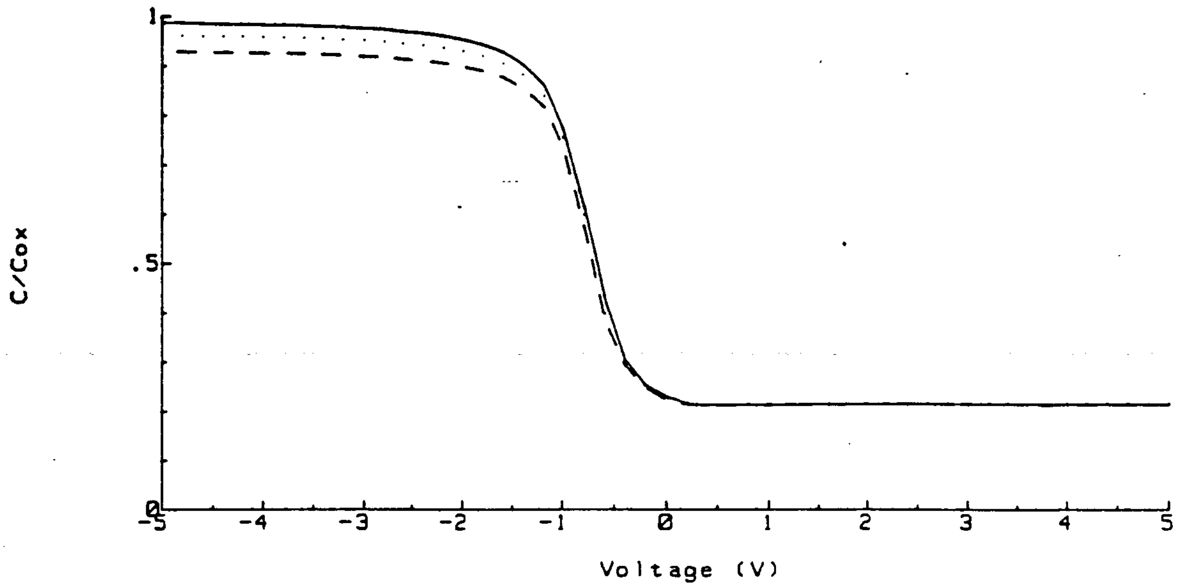


Figure 3.13 (a). CV characteristics at 10 kHz. (solid line), 40 kHz. (dotted line),and 100 kHz. (broken line). Series resistance was 250 ohms.

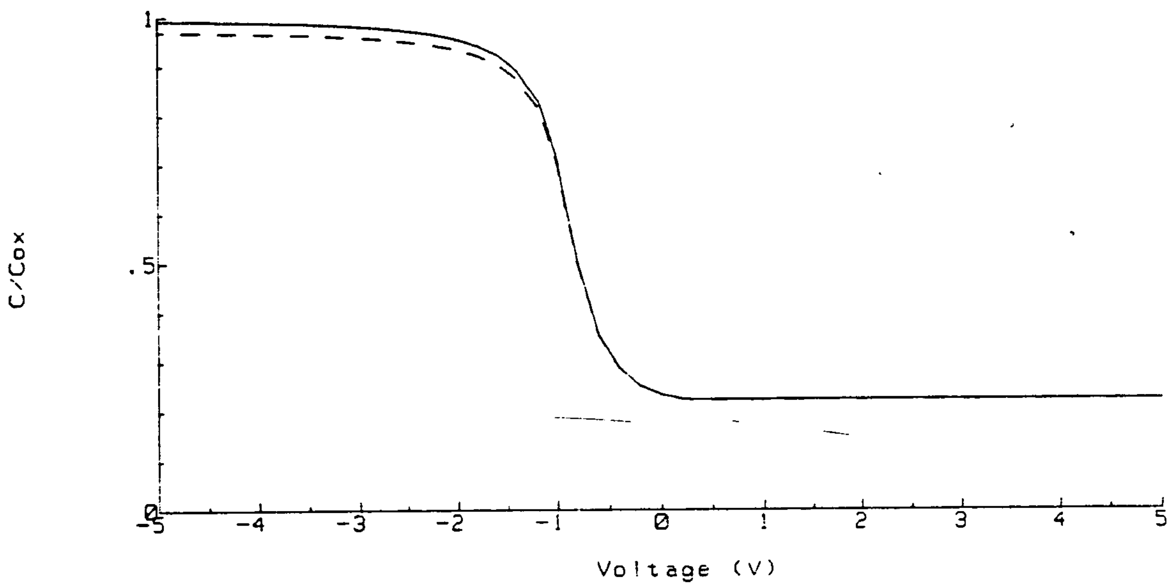


Figure 3.13 (b). CV characteristic at 40 kHz, as measured (broken line), and corrected for series resistance (solid line).

$$w_{max} = -\epsilon_s \left(\frac{1}{C_{ox}} - \frac{1}{C_{min}} \right) \qquad (3.24)$$

where C_{min} is the equilibrium minimum capacitance. The Linder-Brews expression shows that the surface potential at which the onset of strong inversion occurs is dependent upon substrate doping density. Re-writing equation 3.13 in terms of potential

$$\phi_m = 2\phi_b + \frac{kT}{q}\ln((2q\,\phi_b/kT)-1) \qquad (3.25)$$

Finally, in order to derive an expression for Nsub in terms of measurable quantities, a result from the solution of the 1-D Poisson equation for uniform doping is required[9]. This relates w_{max} to the maximum surface potential as follows

$$w_{max} = \left(\frac{2\epsilon_s \phi_m}{qN_{sub}} \right)^{1/2} \qquad (3.26)$$

Now by equating the right hand sides of equations 3.24 and 3.25 and rearranging terms we have

$$\frac{N_{sub}}{\ln(N_{sub}/n_i) + \frac{1}{2}\ln\left[2(N_{sub}/n_i)-1\right]} = \frac{4kT}{\epsilon_s q^2} \left(\frac{1}{C_{ox}} - \frac{1}{C_{min}} \right)^{-2} \qquad (3.27)$$

An iterative solution of 3.27 can be found using

$$N_{sub}^{j+1} = K\left(\ln[N_{sub}^{j}/n_i] + \tfrac{1}{2}\ln(2[N_{sub}^{j}/n_i]-1)\right) \quad (3.28)$$

where K is the right hand side of 3.27 and a suitable choice of $N_{sub}^{j=0}$ is 1×10^{20} cm^{-3}. Convergence of the iterative process usually occurs in approximately 6 cycles.

In cases where the substrate doping is non-uniform the value of N_{sub} that is determined is the mean doping density between the interface and the depletion layer edge.

The primary sources of error in N_{sub} are erroneous values of C_{ox} and C_{min}. The correct determination of the equilibrium C_{min} is inhibited by quite different factors from those affecting C_{ox}, (C_{min} is not affected by R_s as can be seen from figure 3.10). Two of the causes of error in C_{min} are:†

(a) Light[3, 114, 138] causes an increase in the minimum capacitance. Figure 3.14 shows the effect dim room lights and microscope lights have on the measured minimum capacitance. Often the effect can be much more severe than this and in some cases high frequency CV characteristics appear qualitatively similar to low frequency characteristics. The effect of light will be more severe for samples with small area. This is because of the increased importance of the periphery of the gate where the light generates carriers. Problems also occur if the gate electrode is transparent or semi-transparent, or if the sample contains interface, or bulk trapped charges which are sensitive to the applied illumination. The effect of light is difficult to model or control and so CV measurements ought to be made in a light-tight metal box such as that described in section 3.5.1.

(b) Inversion layer beyond the gate (ILBG)[9, 11, 139] is the least well understood and most difficult problem to overcome when measuring C_{min}. As the name

† A third cause, which has only become important recently is discussed in depth in section 3.11.

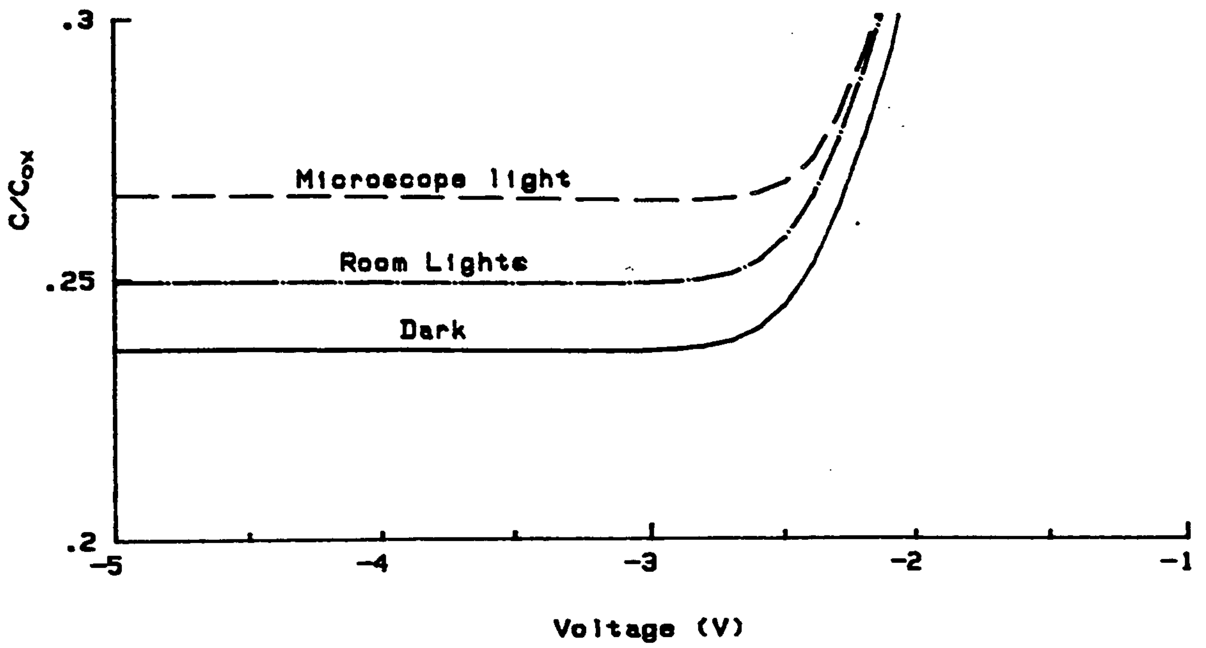


Figure 3.14. The effect of light on the measured minimum capacitance.

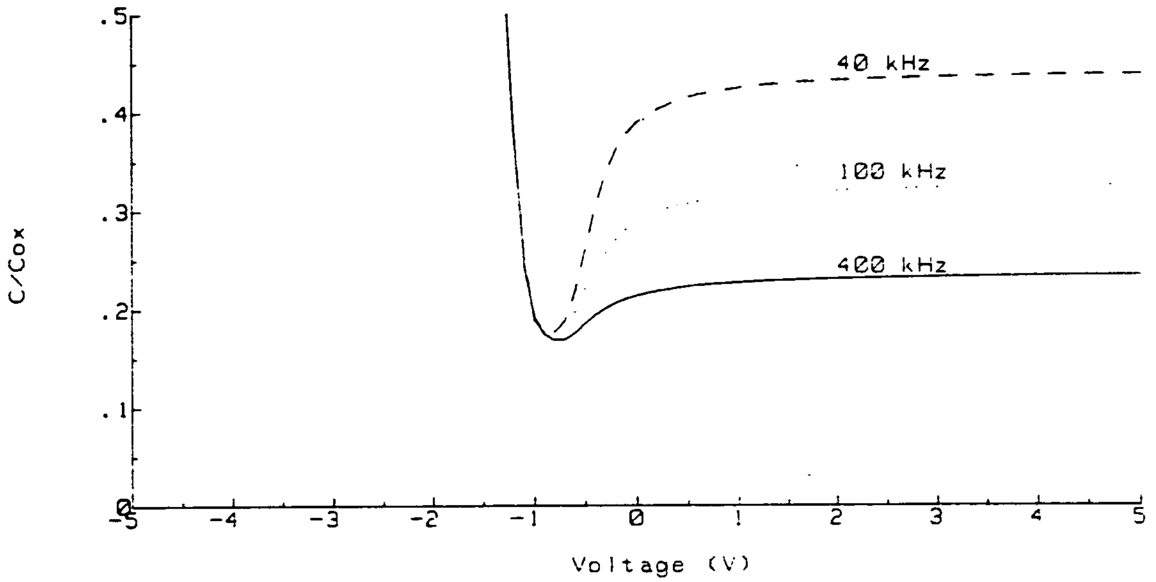


Figure 3.15. Measured capacitance as a function of frequency when there was an ILBG present.

suggests an inversion layer is present beyond the edge of the gate. This can arise in two ways. The first only occurs in lightly doped p-type substrates. For this type of substrate only a very small quantity of oxide charges can be sufficient to invert the silicon surface at zero bias and so all areas beyond the gate where there is no applied bias will be inverted. The second cause of ILBG occurs for both n and p-type substrates. In this case the gate area is enlarged by the gate charge inducing moisture present at the oxide-air interface around the gate to ionise.

The presence of ILBG changes the response of the inversion layer to the ac signal used to measure high frequency capacitance because instead of relying on generation-recombination to change its occupancy, carriers are rapidly fed in and out of the external inversion layer. Figure 3.15 shows the effect of ILBG on the CV characteristic over a range of frequencies. The effect is reduced but not eliminated at higher frequencies. ILBG will be most likely to occur with p-type samples of small area particularly if there is water adsorbed onto the oxide surface. CV characteristics of MOSFETS will always appear to suffer from ILBG because the source and drain regions readily supply minority carriers to the gate region.

The chances of water causing ILBG can be greatly reduced by ensuring that the sample does not come into contact with water after the final high temperature processing step. It is my experience that ILBG most frequently occurs because of DI water washing after back-oxide removal. This can be eliminated by sintering after backside metal deposition. If contact with water is unavoidable and only if ILBG shows up in measurements, it can sometimes be eliminated by a one hour bake at approximately $200^{o}C$ and cooling the sample in an inert gas ambient[11]. In addition a gentle stream of nitrogen blown over the wafer during measurements helps to prevent the adsorption of water onto the oxide surface.

Modification of the gate definition to include a guard ring, field oxide or heavily doping the region beyond the gate will greatly reduce the possibility of ILBG. Figure 3.16 shows a schematic drawing of these more complex structures.

Guard rings are biased so as to accumulate the region beyond the gate. The closer the guard ring is to the gate the more effective it will be. For a process using only wet etching the closest practical spacing for a guard ring is about 3 microns.

The use of a heavily doped region of the same type as the dopant under the gate stops ILBG by greatly increasing the inversion voltage of the surface. Typical source/drain diffusions or implants of the same dopant type as the substrate will completely suppress surface inversion. Field oxide around the gate periphery has a similar effect. The disadvantage of these test structures is that they require significantly more processing than a simple MOS capacitor or capacitor with guard ring. (A guard ring only requires a slightly more elaborate gate photomask). The heavy dose channel stop is only normally used where the additional processing would have to be done anyway e.g. on a process control chip.

3.6.3. Flatband Capacitance (C_{fb}) and Flatband Voltage (V_{fb})

For a uniformly doped substrate the surface potential (ψ_s) will be zero when the silicon energy bands are flat all the way from the oxide-semiconductor interface to deep in the bulk of the semiconductor. The voltage at which $\psi_s = 0$ is called the flatband voltage (V_{fb}). At the flatband voltage there is an associated silicon capacitance C_{sfb} and the measured flatband capacitance (C_{fb}) will be a series combination of C_{ox} and C_{sfb}. If N_{sub} is known then C_{sfb} can be calculated[10] using,

$$C_{sfb} = \frac{\epsilon_s}{\left(\frac{\epsilon_s kT}{q^2 N_{sub}}\right)} \qquad (3.29)$$

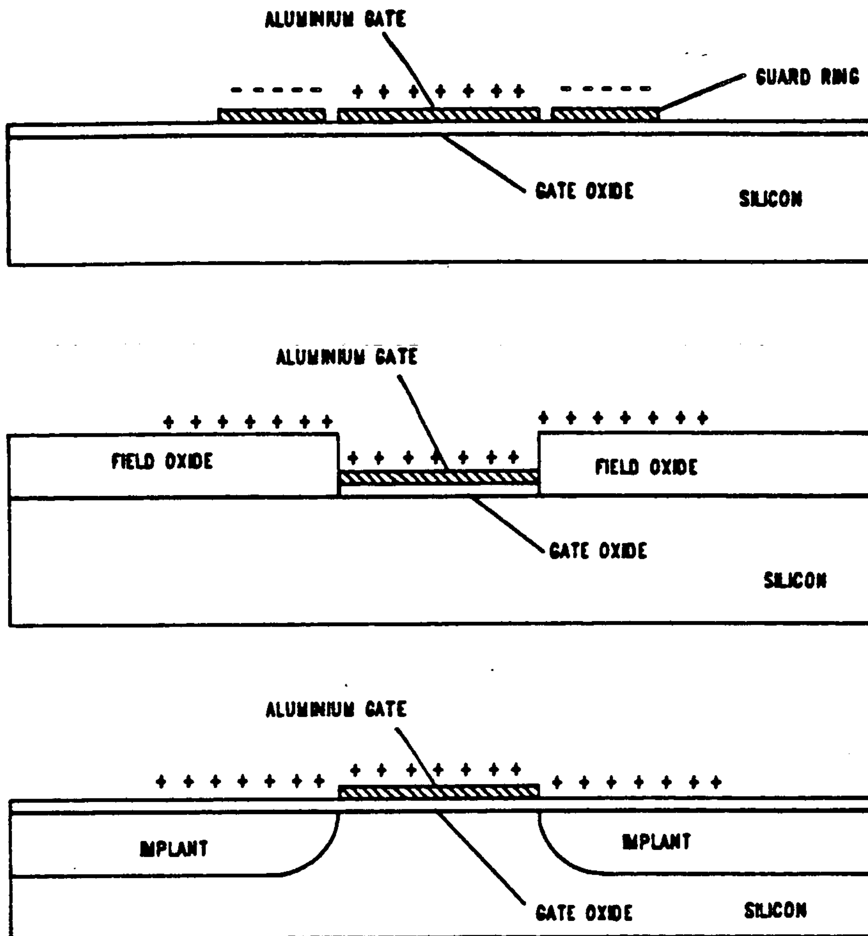


Figure 3.16. Structures to inhibit ILBG formation.

and

$$C_{fb} = \frac{C_{ox} C_{sfb}}{C_{ox} + C_{sfb}} \ . \tag{3.30}$$

In the ideal case V_{fb} ought to be zero, however real MOS capacitors have non-zero metal-semiconductor work function differences and oxide charges which give rise to a finite flatband voltage. V_{fb} is determined experimentally by finding the voltage at which the measured capacitance equals C_{fb}.

Errors in C_{fb} primarily occur due to C_{ox} and C_{min} (N_{sub}) errors. Careful choice of voltage step and the measurement of capacitance to as many significant figures as possible will prevent an an increase in error for V_{fb} compared to C_{fb}.

3.6.4. Barrier Heights, Work Function and Work Function Differences[9, 130, 140]

Figure 3.17 shows that there is a difference between the work function of aluminium and p-type silicon. More generally there will always be a work function difference between any metal or polysilicon gate and a silicon substrate. The work function for the metal is defined by the equation

$$W_m = \frac{E_{mo} + E_{vl}}{q} \tag{3.31}$$

where E_{mo} is the metal barrier height and E_{vl} is the energy gap between the oxide conduction band and the free electron level. The work function for silicon is defined by

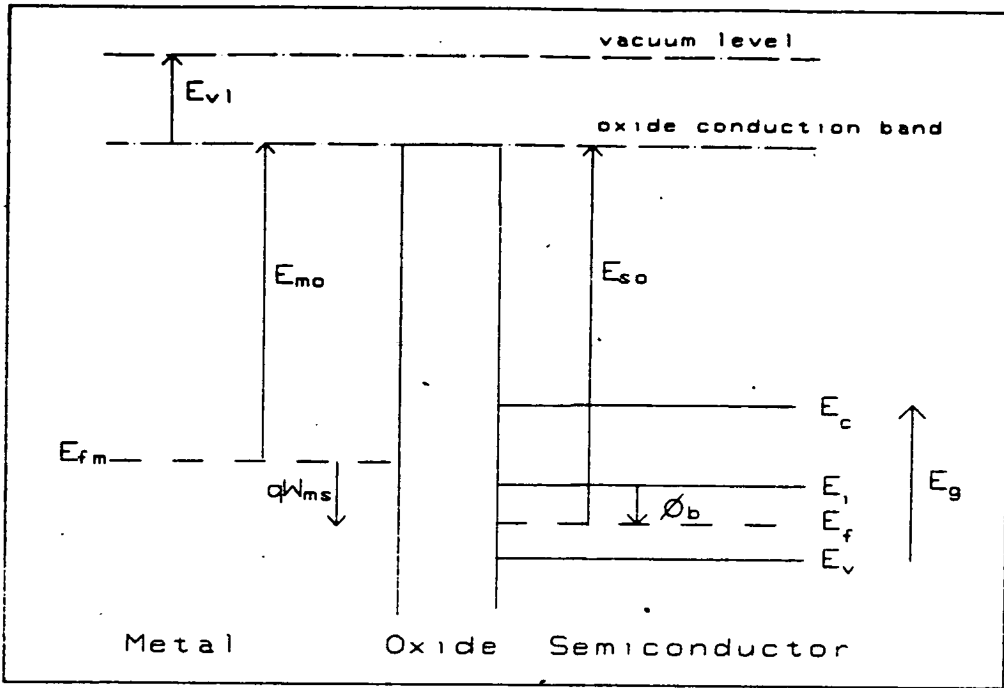


Figure 3.17. Energy band diagram defining work functions.

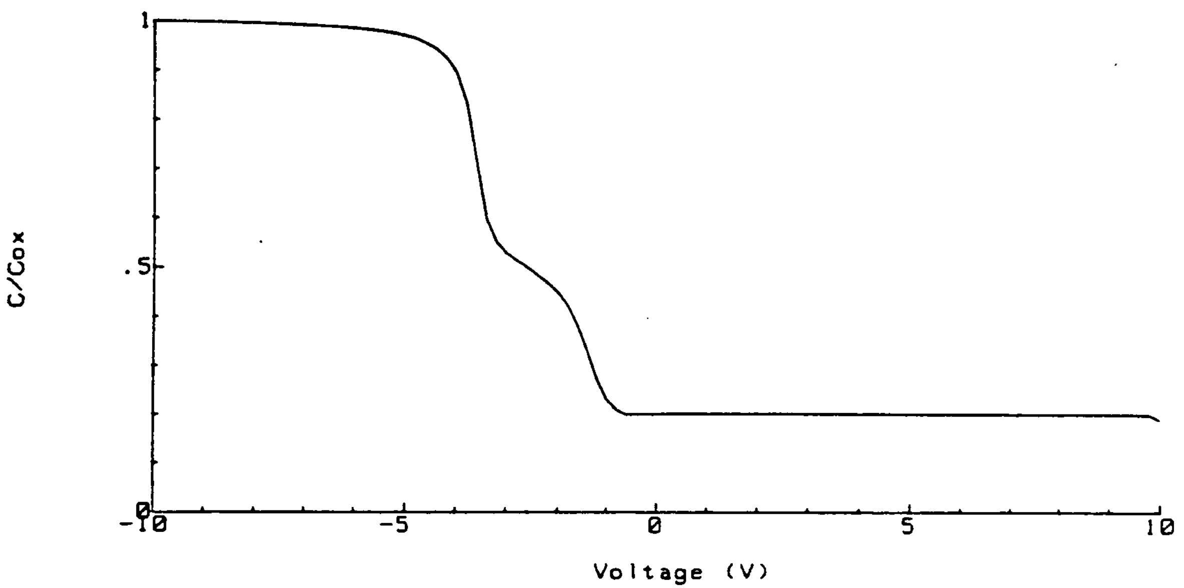


Figure 3.18. CV characteristic distorted by interface trap distribution of highly localised energy.

$$W_s = \frac{E_{so} + E_{vl} + E_g/2 + q\,\phi_b}{q} \tag{3.32}$$

where E_{so} is the silicon barrier height. Therefore the metal semiconductor work function difference (W_{ms}) is defined by†

$$W_{ms} = \frac{E_{mo} - E_{so} - E_g/2 - q\,\phi_b}{q} \,. \tag{3.33}$$

Table 3.1 lists values of E_{mo} and the intrinsic work function $W_{ms}^{\;i}$ which is defined as

$$W_{ms}^{i} = \frac{E_{mo} - E_{so} - E_g/2}{q} \tag{3.34}$$

i.e. the metal semiconductor work function difference for intrinsically doped silicon. Then for real samples with finite uniform doping

$$W_{ms} = W_{ms}^{i} + \phi_b \,. \tag{3.35}$$

Additional terms are required for non-uniformly doped substrates[140].

W_{ms} creates a built in potential across the oxide at zero bias. This means that even at zero applied bias the silicon energy bands are bent in the region close to the

† In certain text books the silicon work function is calculated using the electron affinity (χ) which is 1.12 eV less than the barrier height quoted here. When χ is used equation 3.33 becomes $qW_{ms} = E_{mo} - \chi + E_g/2 + q\,\phi_b$.

oxide-semiconductor interface. In the absence of oxide charges the application of a voltage equal to but opposite in polarity to W_{ms} will restore the semiconductor to the flatband condition.

The barrier heights quoted in table 3.1 have been determined by internal photoemission, the etch off CV method and the multi-metal CV method[9].

In principle the results from photoemission experiments ought to provide the most accurate values of W_{ms}, for although the theory and experimental procedure of the method are considerably more complex than both the CV methods, the technique is very sensitive to barrier height. The major problems of the CV methods are the assumptions that the metalisation has no influence on oxide properties and that even if this is the case that the oxide under the various gates exhibits a high degree of uniformity for all charge levels. (The validity of these two assumptions is addressed in more detail later in this section).

The Deal values for barrier heights are those most frequently quoted in the literature. It often appears that these values are given the status of fundamental constants. This is unfortunate because the degree of uncertainty in these values was not readily quantified at the time they were measured, and the purity of materials, quality of fabrication equipment and accuracy of test equipment were all very inferior to that which is currently available. Some consequences of the possible errors in W_{ms} values will be discussed in the next section.

3.6.5. Oxide Charges

In addition to the shift caused by metal-semiconductor work function differences, real CV curves are also shifted and also distorted by oxide charges. The simplest assessment of oxide charge density involves the determination of the total contribution (Q_{ss})[141] that all types of charge combine to contribute to the flatband voltage. Q_{ss}

metal	E_{mo}	W^{i}_{ms}
magnesium	2.4	-1.39
aluminium	3.2	-0.59
nickel	3.6	-0.19
copper	3.8	+0.01
silver	4.2	+0.41
gold	4.1	+0.31
p+ polysilicon	4.35	+0.56
n+ polysilicon	3.23	-0.56

Table 3.1 Metal-oxide barrier energies and metal-semiconductor work fuction differences.

in terms of the different types of oxide charge is defined by

$$\frac{Q_{ss}}{C_{ox}} = \frac{Q_f}{C_{ox}} + \frac{x_m Q_m}{\epsilon_{ox}} + \frac{x_{ot} Q_{ot}}{\epsilon_{ox}} + \frac{Q_{it}'}{C_{ox}} \tag{3.36}$$

The contribution from fixed oxide charge is accounted for since by definition it is all located at the oxide-semiconductor interface and therefore each oxide fixed charge must be balanced by an equal but opposite charge in the semiconductor. In general mobile ions and oxide trapped charge are distributed throughout the oxide. It can be shown[9] from a solution of Poisson's equation for a generalised distribution that the voltage shift due to the distribution is given by

$$\Delta V = \frac{\bar{x}Q}{\epsilon_{ox}} \tag{3.37}$$

where $\bar{x}$ is the centroid (first moment) of the distribution. In general $\bar{x}$ is not known and for this (and other) reasons the single high frequency CV curve cannot distinguish between the contributions from different charge types. Since the charge located in interface traps is bias dependent, only the nett occupancy of interface traps at the flat-band voltage (Q_{it}') contributes to flatband voltage shift.

Q_{SS} is determined from the measured flatband voltage using

$$V_{fb} - W_{ms} = -\frac{Q_{ss}}{C_{ox}} \tag{3.38}$$

It was widely assumed for many years that Q_{ss} was inherently positive, however measurement of negative values for Q_{ss} is now commonplace. This change suggests that either (1) As oxide quality has improved with enhanced growth methods, improved cleanliness and optimized annealing procedures means that negative rather than positive charges dominate Q_{ss} or (2) Metal-Semiconductor work functions are incorrect.

The following points are helpful in assessing which of these possibilities is more likely to be true.

(a) Deficiences in early processing methods meant that it was inevitable that measured values of Q_{ss} would be positive. A major factor in this category was a lack of control over mobile ion density. Since Q_m is known to be positive then when this is the dominant factor, Q_{ss} must also be positive.

(b) Prior to the introduction of implantation the vast majority of MOS ICs were p-channel, (n substrate) devices fabricated on <111> orientated wafers. Fixed oxide charge densities are known to be higher for <111> than for <100> crystal orientation[141, 142], with the nett effect invariably being a more positive shift for <111>.

(c) Segregation of semiconductor dopant[143] into the oxide will cause a positive shift in the CV characteristic compared to the CV characteristic of a uniformly doped substrate.

(d) The work functions commonly quoted today were measured in 1964-65. At that time metal gates were commonly evaporated from a pot heated by a tungsten filament. According to Gundlach[18] such evaporation systems were invariably contaminated with sodium and although low sodium filaments became available in 1968, it was not until the early 1970s and the introduction of electron beam evaporation of aluminium that significant improvements in the cleanliness of the metalisation process were achieved.

(e) It has been argued by Hamsaki[28] that negative values of Q_{ss} arise only for thin oxides where a negative charge present for all thicknesses at the metal-SiO_2 interface begins to be of importance. A later paper of Hamsaki[144] showed that interface trap density increases very rapidly for thin oxides below about 400 Å. If the interface traps are negatively charged then this could also explain negative Q_{ss}. This author has found that negative values of Q_{ss} can occur in even relatively thick oxide of approximately 800 angstroms with very low interface trap densities which is not to be expected from the arguments forwarded by Hamsaki.

(f) Values of Q_{it}' of sufficient accuracy to help decide whether the contribution from interface traps is sufficiently negative to cause negative Q_{ss} are impossible to obtain because interface trap measurements can only determine trap density to a reasonable level of accuracy over a very limited portion of the silicon bandgap[9].

The above points demonstrate that it is very difficult on presently available evidence to decide which is the correct explanation of negative Q_{ss}. Fortunately for the MOS manufacturer the most important parameter is V_{fb} as will be shown in the next section and so the dilemma discussed here is more of academic importance than practical importance. The lack of a full set of accurate measurement methods for all types of charge will mean that proper modelling of the charge contributions to V_{fb} will not be feasible for the forseeable future.†

† After this text was written in March 1986, This author had the opportunity to talk to Edward Nicollian about this dilemma. It was his opinion that negative Q_{ss} arose due to innaccurate work function values.

3.6.6. Threshold Voltage (V_T)

The simple first order transistor model of Sah[145] (which describes the IV characteristics of MOSFETS) defines a parameter, threshold voltage (V_T) which is the first voltage at which there is a non-zero current between source and drain electrodes. The simple model assumes that the threshold occurs when the silicon surface becomes strongly inverted. The MOS capacitor can be used to estimate threshold voltage for a uniformly doped substrate using

$$V_T = V_{fb} + 2\phi_b - \frac{Q_b}{C_{ox}} \tag{3.39}$$

where Q_b is the charge on the silicon capacitor at the onset of strong inversion and is given by

$$Q_b = -qN_{sub}\left(\frac{\epsilon_{ox}\epsilon_{si}}{C_{smin}}\right). \tag{3.40}$$

An early use of this equation was to estimate the threshold voltage of unimplanted metal-gate PMOS devices which were adequately modelled by the simple Sah model.

Equation 3.39 will only yield an accurate value of V_T if the only effect of oxide charges is to rigidly shift the CV curve along the voltage curve without distortion. Figure 3.10 shows an example of where this is true for a theoretical sample in which only a range of fixed oxide charge is present. When there is an appreciable density of interface traps, the CV curve will appear stretched out along the voltage axis. This stretchout can take a wide variety of forms from a relatively constant

stretchout to a highly localised stretchout as that shown in figure 3.18. In cases such as this, a better estimate of the threshold voltage can be obtained from the first point at which the capacitance reaches the strong inversion capacitance.

3.7. Analysis Of The CV Characteristics Of Non-Uniformly Doped MOS Capacitors

The analysis of CV curves described in section 3.6 cannot be used when the semiconductor doping profile is non-uniform. Only with a more sophisticated model can C_{sfb}† be calculated, and hence C_{fb} determined, and then V_{fb} found from the CV plot. The calculation of C_{sfb} requires that the doping profile be known accurately from the interface until it falls off to the substrate doping density, consequently the solution of Poisson's equation requires the use of a numerical method such as that of Jaeger et al[146].

Buttar[147] has developed software to determine C_{sfb}, C_{fb} and complete CV curves using profiles generated by SUPREM[71], combined with routines to implement Jaeger's solution of Poisson's equation. Using the theoretical value of C_{fb}, V_{fb} is determined from the measured CV plot. W_{ms} can be calculated using the formula proposed by Marashak[140] and hence using equation 3.38 a value for Q_{ss} can be calculated.

In practice this approach to determining V_{fb} of a non-uniformly doped sample is unlikely to produce reliable values of V_{fb} and Q_{ss} because uncertainty in the doping profile will lead to errors in both C_{fb} and W_{ms}. In addition to the lack of accuracy, implementation of this approach is not practical other than for a very few samples, because of the need to devote considerable computer resources and time for each calculation.

---

† The idea of flatband capacitance for non-uniform profiles is something of a misnomer since in general the bands never will be flat at any gate bias. In the following discussion it is assumed that the voltage at which the surface potential is zero will be the flatband voltage.

Despite the difficulty in determining V_{fb} there is still useful information in the CV curve. Oxide capacitance (equation 3.6), maximum deletion layer width (equation 3.24), mean doping between the interface and depletion layer edge (equation 3.27) and threshold voltage can all be determined. In addition, qualitative and semi-quantitative measurements of oxide charges can be made.

3.8. Control of MOS Integrated Circuit Fabrication Process

Chapter 2 showed that the manufacture of integrated circuits is a complex process involving a very large number of individual processing steps. In order to obtain a yield of working devices at the end of such a process it is necessary to ensure during the fabrication process that each step is being executed to within the required specification. High frequency CV measurements are a very important means of monitoring and controlling oxide charge levels[8, 11]. More recently, pulsed high frequency CV measurements are being used to monitor ion implantation doses[148].

A fabrication plant will normally check the proper performance of all oxidation furnaces and other critical high temperature furnaces as well as the cleanliness of metalisation at least once per day using a scheme similar to that in table 3.2. Table 3.3 lists some commonly occurring fabrication problems and the effect these have on high frequency CV characteristics.

Additional information can be extracted from CV data by plotting trend charts for various parameters. For example if oxide thickness plotted as a function of time shows a steady decrease then the cause is more likely to be a gradual drift downwards of temperature or one of the gas flows, rather than a gate area or ohmic contact problem.

3.8.1. Process Control Example†

The problem addressed in this example[149] is whether an oxide gate dielectric which is susceptible to mobile ion contamination is a better choice of gate dielectric for a metal gate NMOS process than a composite gate of silicon nitride atop silicon dioxide which is known to eliminate mobile ion contamination after oxidation but involves more complex processing.

To decide which is best, the unstressed CV parameters and mobile ion density of both types of samples were measured in a fabrication line over a period of two weeks. The measurement over a period of time simulates the variability in the processes which is important when a high yield must be obtained in a commercial operation.

Figure 3.19 shows a time plot of Q_{SS} for both oxide and oxide-nitride dielectric samples. As expected, the composite gate samples generally had immeasurably low levels of mobile ion contamination, whilst the oxide only samples showed levels up to 5×10^{10} cm^{-2}.

Although the results presented here for oxide are not excellent, such an oxide would give a reasonable yielding process provided the other fabrication was performed properly. The oxide-nitride results show that whilst they confirm that the nitride effectively excludes mobile ions, the price paid is an unacceptably high and variable level of oxide charge density. Such variability would result in MOSFET threshold variations of several volts which makes commercial fabrication impractical. (This example probably overstates the difference between these processes because of a poor nitridation procedure).

† This is not a true example of process control. However the daily measurements presented here clearly show a process in and out of control.

1. Fabricate MOSCS (1 per critical process step)
2. Measure HFCV characteristic (room temperature)
3. Inspect for gross abnormalities.
4. Determine D_{ox}, V_{fb}, Q_{ss}.
5. Bias temperature stress *.
6. Inspect for gross failure under stress.
7. Determine Q_m
8. Pass process, or Fail process.

Table 3.2 Process control recipe.

* see table 3.4 for details

Table 3.3 Fault Diagnosis from HFCV Characteristics		
Parameter	Fault	Possible Causes
D_{ox},(C_{ox})	low(high)	-furnace temperature low -incorrect gas flow -wrong oxidation procedure
	high(low)	-furnace temperature high -wrong oxidation procedure -poor cleaning/rinsing -poor MOSC fabrication (R_s)
V_{fb}, (Q_{ss})	more positive (more negative)	-wrong oxidation procedure -wrong gate metal -poor MOSC contact -wrong substrate doping
	more negative (more positive)	-as for positive and -poor cleaning -wrong crystal orientation -mobile ions -poor/omitted annealing
Q_m	high	-poor cleaning/rinsing -furnace contaminated -handling contamination

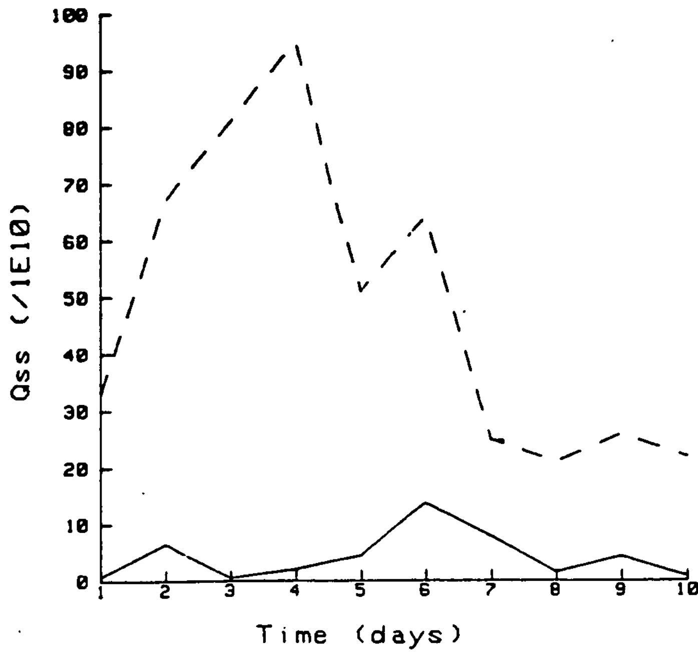


Figure 3.19. Q_{ss} as a function of time for oxide dielectric (solid line), and oxide-nitride dielectric (broken line).

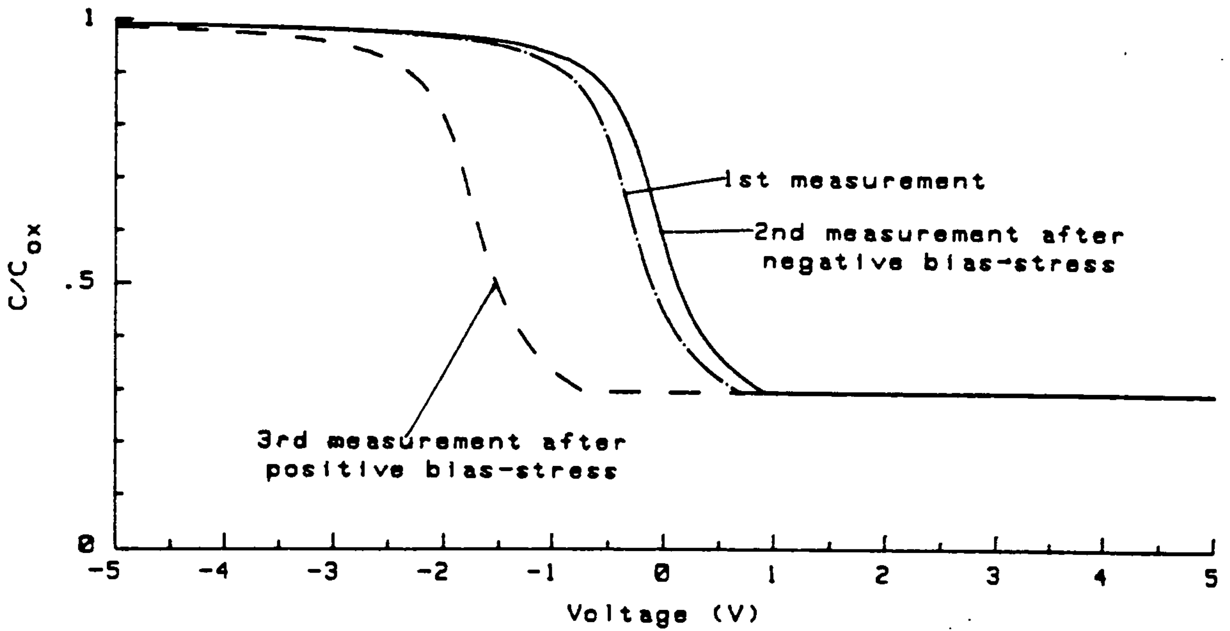


Figure 3.20. CV curves determined before and after bias temperature stresses.

3.9. Measurement of Mobile Ion Density

Mobile ion density is most commonly determined using the Bias-Temperature (BT) high frequency CV method[111, 135]. Q_m is inferred from the shifts in flatband voltage which occur after stressing. The main alternative measurement of mobile ion density is the Triangular Voltage Sweep (TVS) method which was developed by Kuhn and Silversmith[150] and Chou[151] from a procedure proposed by Yamin[152]. The latter method is considerably more difficult to implement but is favoured in fundamental studies because it provides individual ion densities and kinetic information independently of interface state creation during stressing. The CV method finds wide application in more routine monitoring applications.

A summary of the BT method is presented in table 3.4. There is no definitive set of values for the voltages, temperature or duration of stress which should be used, however the values quoted in table 3.4 ought to be adequate in the majority of situations.

Q_m is determined by applying a result from equation 3.34 where the shift ΔV_m due to mobile ions alone is

$$\Delta V_m = Q_m \frac{\bar{x}}{\epsilon_{ox}} \tag{3.41}$$

where $\bar{x}$ is in general unknown. The BT method determines Q_m by minimising $\bar{x}$ with a negative stress which accumulates the mobile charge at the metal-silicon dioxide interface. Thus $\bar{x}$ is effectively zero. The positive stress accumulates the mobile charge at the silicon dioxide- silicon interface thus maximising $\bar{x}$ such that $\bar{x} \approx D_{ox}$. Then if V_{fb}^- and V_{fb}^+ are the flatband voltage shifts after negative and positive stressing respectively then

1. Fabricate MOSC.
2. Measure high frequency CV characteristic at room temperature.
3. Determine V_{fb}.
4. Heat sample to 200°C.
5. Hold at high temperature with gate biased to -10V for 10 minutes.
6. Cool to room temperature whilst maintaining the gate at -10V.
7. Inspect for gross stress failure.
8. Remeasure high frequency CV characteristic and determine V_{fb}^{-}.
9. Repeat steps 4-8 for +10V bias.
10. Determine V_{fb}^{+}.
11. Calculate Q_m.

Table 3.4 Mobile ion test recipe.

$$Q_m = C_{ox}(V_{fb}^{+} - V_{fb}^{-}) \qquad (3.42)$$

Figure 3.20 shows an exaggerated example of the shifts that occur before and after stressing using the BT m.thód. Often in production environments only the voltage shift is of interest, (because they produce to a voltage tolerance), and in this case there is the advantage that the gate area and C_{ox} need not be known.

The BT method fails to yield accurate values for voltage shifts and hence Q_m when the applied stress induces some other oxide degradation mechanism such as interface state generation[153], or even dielectric breakdown[154]. The presence of interface state generation during stressing can be checked for by measuring the shift of the CV curve at several points in depletion. If there is no trap generation, all three shifts will be equal. ILBG is a problem which commonly occurs during stressing due to the condensation of water vapour onto the wafer surface during the cooling cycle. This however should not inhibit the determination of Q_m since C_{fb} which determines V_{fb} can be calculated from the unstressed plot.

In industry it is very common simply to measure only the shift after positive stress. This shortcut has been adopted because of the observation that the shift after positive stress is generally far greater than the shift after negative stress and that negative stressing often induces poòrly understood oxide degradations which would not appear to relate to any failure mode of real devices.

3.10. CV Characteristics of MOS Capacitors with Two Top Capacitive Contacts

Capacitive contacts, when used for CV measurements, have the advantage that the sample requires less preparation than for an ohmic contact to the back of the wafer. Two top capacitive contacts are also frequently used for mercury probe measurements[128, 129] and for insulating substrates such as SOS. The errors involved in this measurement are usually calculated assuming two capacitors in series giving a measured capacitance (C_m) in accumulation of

$$C_m = \frac{C_{ox1} C_{ox2}}{C_{ox1} + C_{ox2}} \tag{3.43}$$

where C_{ox1}, C_{ox2} are the oxide capacitances of the small and large capacitors. If $C_{ox2} \gg C_{ox1}$ then $C_m \simeq C_{ox1}$ and for a ratio of 100:1 the error is about 1%. This approximation is often inadequate because when the smaller capacitor is accumulated the second capacitor is depleted. This depletion may contribute a significant additional capacitance. An improvement on the simple model is shown in figure 3.21.

When the smaller capacitor is in accumulation its capacitance is C_{ox1} while the large capacitance is not simply C_{ox2} but is given by

$$C_2 = \frac{C_{ox2} C_{s2}}{C_{ox2} + C_{s2}} . \tag{3.44}$$

and therefore the assumption made in equation 3.43 is in general not valid.

The problem is further compounded by the fact that the effects of series resistance from the contacts and silicon is difficult to account for which can lead to significant

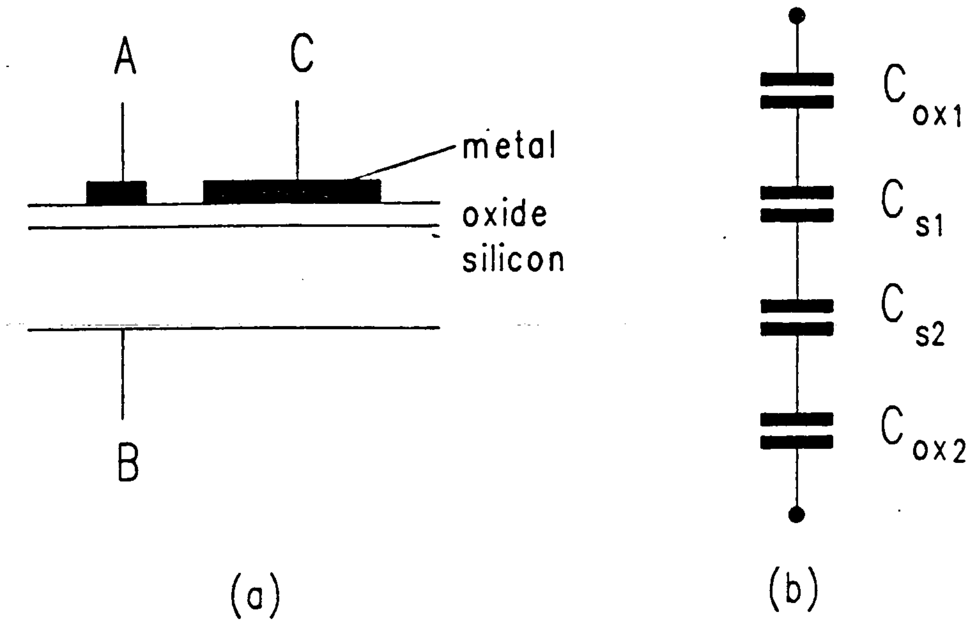


Figure 3.21. Electrode configuration and equivalent circuit for an MOS capacitor with two top contacts.

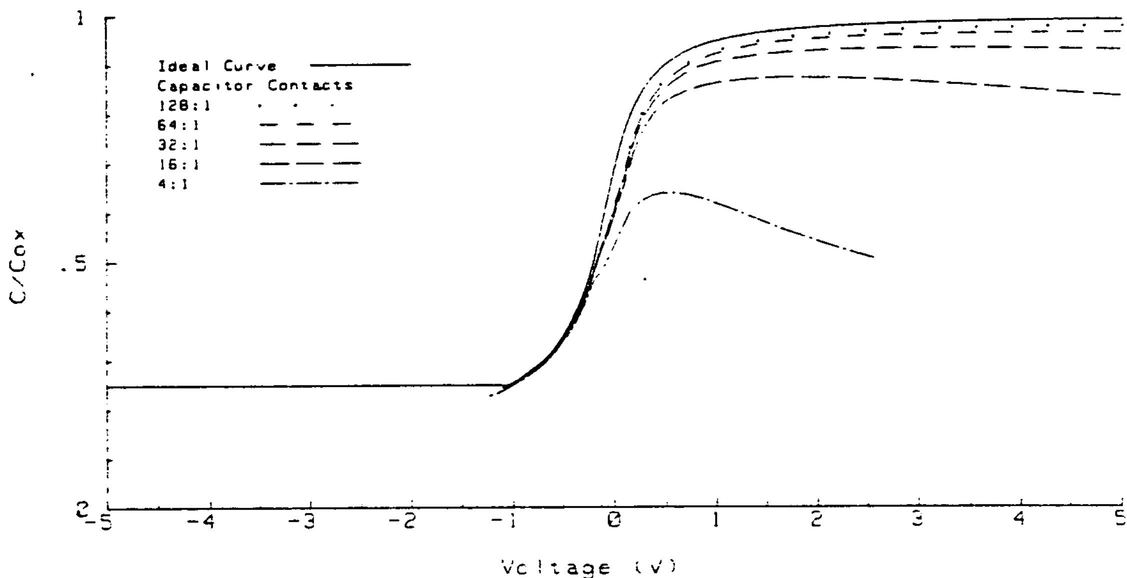


Figure 3.22. Theoretical CV characteristics of capacitors with different ratios of two top contacts.

errors

3.10.1. Comparison of Capacitive Contact Measurements with Theory

To illustrate these errors, capacitor dots for a number of ratios were fabricated using wafer stepper lithography to ensure good dimensional control. The substrate material was n-type <100> silicon with a dopant density of 1.5 x 10^{15} atoms cm^{-3}. The wafers were oxidised in dry oxygen with 5% HCl at 950°C. An ohmic back contact was made to the wafer using aluminium to allow direct comparison with the capacitive contact measurements. The final step was a 20 minute post-metallisation anneal at 435°C in forming gas. Figure 3.22 shows some typical CV curves measured for a number of different top contact ratios compared with the ohmic back contact measurement. It can be observed that as the ratio increases the accuracy improves but even at 1:128 the maximum capacitance is still in error. This is of great importance if parameters such as substrate doping density are are to be calculated because they depend upon the values of C_{ox} and the minimum capacitance measured in strong inversion. The repeatability of measurements across wafers was good for the larger ratios (128:1, 64:1) but was significantly poorer for the smaller ratios (16:1, 4:1).

Figure 3.23 shows the theoretical calculation assuming the model shown in figure 3.21. These curves were computer calculated (program "RINGDOT") by first calculating an ideal CV curve for both capacitors. An iterative process was then used to obtain the voltage drops across each capacitor assuming the charge on all the capacitor plates was equal. A comparison of figures 3.22 and 3.23 verifies the model of figure 3.21.

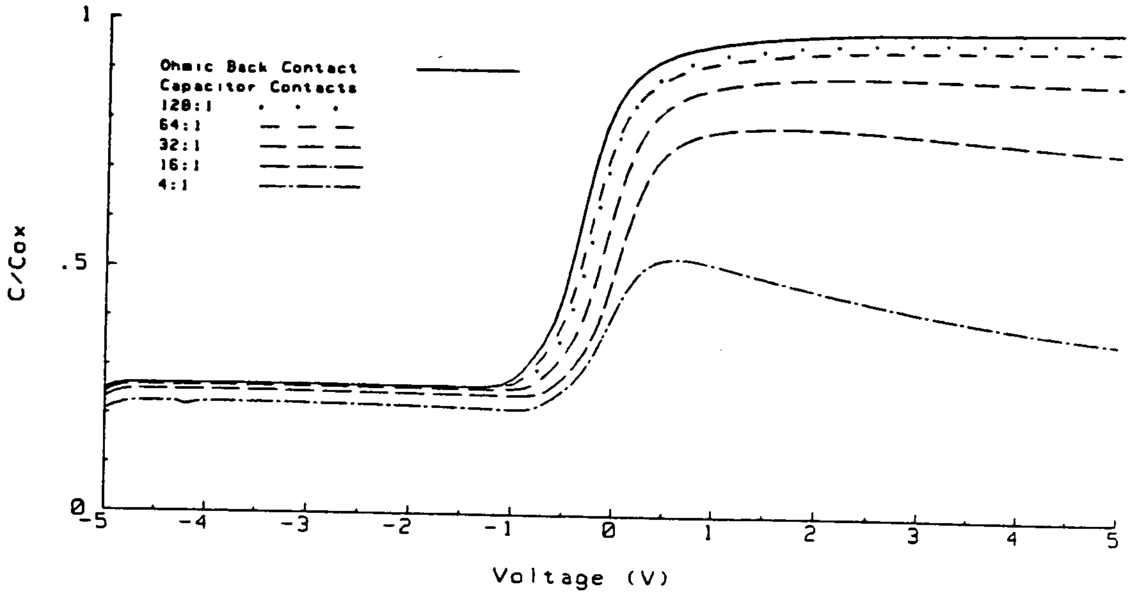


Figure 3.23. Measured CV characteristics of capacitors with different ratios of two top contacts.

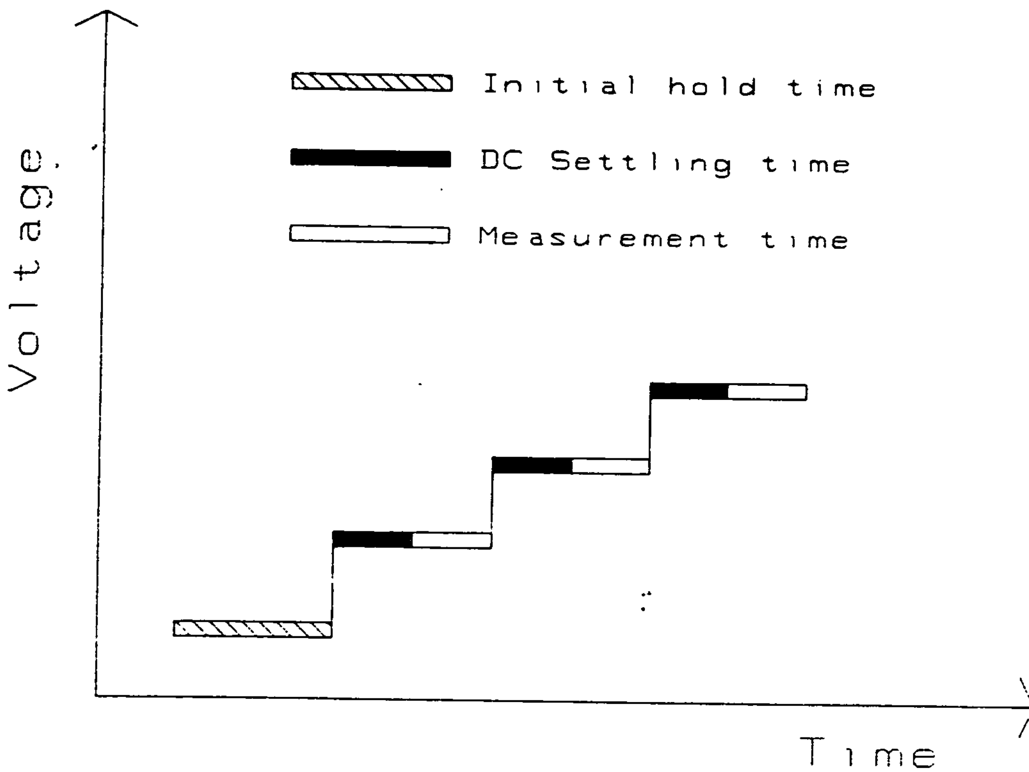


Figure 3.24. Conventional DC bias sweep for high frequency CV measurements.

3.10.2. Conclusions from Two Top Contact CV Measurements

The top capacitive contact is not a good method to use for CV measurements even if ratios are large as 128:1 are used. Double column mercury probes whose ratios are typically 50:1 are not accurate enough to be used for MOS production control where small changes in characteristics may be of importance. This problem is further compounded by the fact that the effect of series resistance is not easy to account for in the capacitive top constant measurement. It is therefore difficult to have confidence in any calculations which use parameters derived from curves using this structure.

3.11. Determination of C_{min} From the High Frequency CV Characteristic of High Lifetime MOS Capacitors

The conventional method of determining N_{sub} from the high frequency MOS capacitance was described in section 3.6.2. There the importance of correctly determining C_{ox} and C_{min} was emphasised. In this section a new phenomenon which has recently begun to affect the accuracy with which C_{min} can be determined is examined.

This phenomenon is a stagnant inversion layer to DC bias ramps. In the past this effect was only considered a problem at liquid nitrogen temperatures[9], but it is now being observed at room temperature because of recent improvements in MOS processing. In particular the introduction of HCl oxidations and gettering techniques have led to substantial increases in silicon minority carrier lifetime. During the late 1970's typical lifetimes were only a few tens of microseconds[155], whereas values of one millisecond or greater are now commonplace. This improvement in quality which gives rise to stagnant inversion layer response is the cause of the new difficulty with the equilibrium high frequency CV characteristic. Alternative approaches to CV parameter determination which overcome this problem are proposed.

3.11.1. Conventional DC Bias Sweep for High Frequency CV Measurement

Figure 3.24 shows the conventional DC voltage sweep used for the high frequency CV measurement with the bias being normally swept to take the capacitor from inversion to accumulation[9]. Equilibrium is more rapidly restored when sweeping in this direction because recombination is a considerably faster process than the generation mechanism which operates when bias is swept toward strong inversion. If voltage is swept too quickly for an inversion layer to form the capacitor deep depletes. In this case the depletion layer widens beyond its maximum width in equilibrium to balance the charge on the gate rather than the inversion layer balancing the gate charge. Normally the DC bias is initially set to put the MOS capacitor into inversion and held in that mode until the capacitance reading becomes constant. This allows the inversion layer to form before the measurement is started. The DC bias is then stepped incrementally and the capacitance allowed to settle before each measurement. Normally no check is made to ascertain whether equilibrium has been reached , instead the wait time is set to a value which is hoped to be adequate for settling to occur. The time is typically around 0.1 seconds.

3.11.2. Measurement Of Equilibrium C_{min} For High Lifetime Samples

Most computer programs for measuring C_{min} ignore the first few points which overcomes a problem for low lifetime samples if the initial hold time is not sufficient. During this period the capacitance is lower than the equilibrium value, and figure 3.25 illustrates this for a CV curve where the sweep was started before the inversion layer had time to form. Figure 3.26 show a qualitatively similar curve for a high lifetime sample, but in this case the mechanism is quite different from that in figure 3.25. It should be emphasised that this plot was obtained after allowing the inversion layer to completely equilibrate. The linear portion of figure 3.25 is the true minimum

capacitance whereas in figure 3.26 there is a spurious dependence of C_{min} on DC bias settling time. Figure 3.27 shows CV curves in the strong inversion region for one sample measured with different DC bias settling times. If the linear portion of these curves is assumed to be C_{min} then derived parameters will show a sweep rate dependence as is illustrated in table 3.5.

For high lifetime samples it is prohibitively slow to measure the entire equilibrium CV characteristic using a conventional sweep. For example it was determined experimentally for the sample in figure 3.26 that the initial hold time should be around two hours and the dc bias settling time about 2 minutes per step. The total sweep time for 100 points would then be about five hours which is obviously totally impractical.

These difficulties in measuring high frequency CV characteristics were noted in 1967 by Goetzberger[156], for samples held at liquid nitrogen temperatures. He reduced the total sweep time by briefly illuminating the sample at the starting bias. This fills the inversion layer above its equilibrium occupancy in the dark and the excess is removed by recombination. Equilibrium is reached approximately ten times faster than without illumination.

Figure 3.28 shows the inversion region of the high frequency curve using a sweep with an initial illumination step. Capacitance was measured immediately after each bias step and then again after 30 seconds. The sample clearly shows similar behaviour to that observed by Goetzberger. Initially after each bias step the inversion layer occupancy is greater than the equilibrium value. To balance the charge on the gate, the depletion region narrows since the inversion layer response is very slow. This results in the high capacitance values. The behaviour here is the opposite of that observed when a sample is deep depleted. After settling recombination reduces the inversion layer occupancy, which allows the depletion layer to widen resulting in a lowering of the measured capacitance.

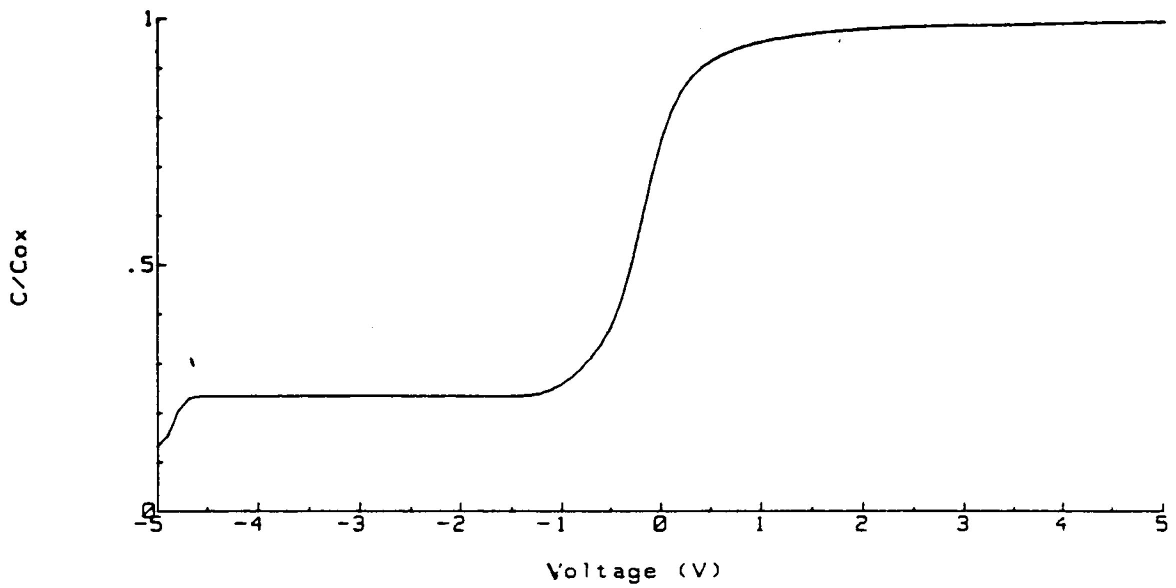


Figure 3.25. CV characteristic of a low lifetime sample with insufficient initial holding time. N.B. The sample is not equilibrated at the start of the measurement in this case.

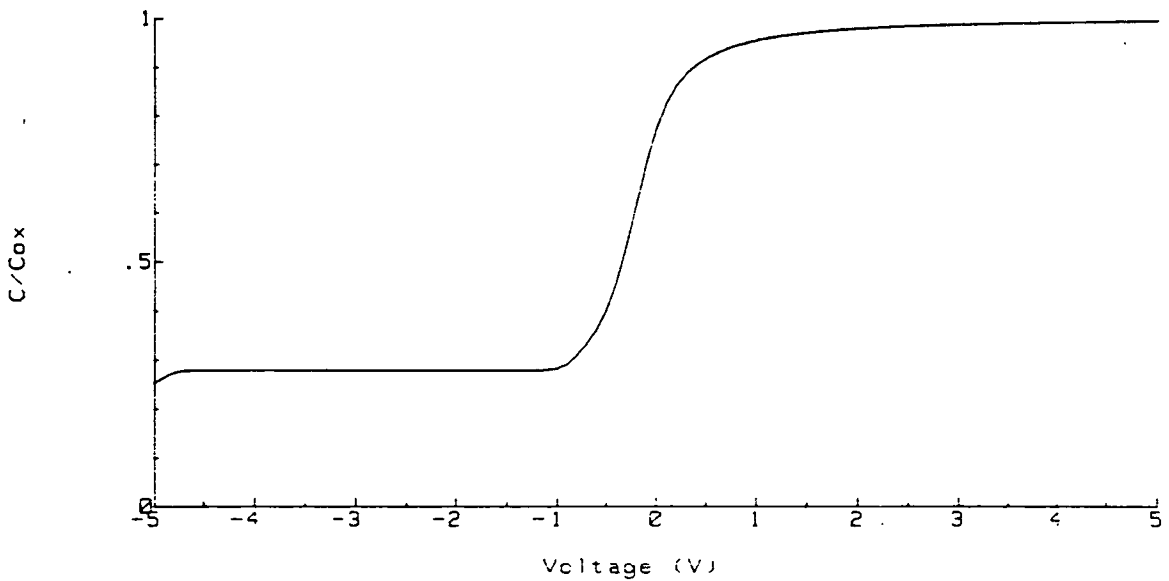


Figure 3.26. CV characteristic of a high lifetime sample. N.B. The sample was equilibrated at the start of the measurement.

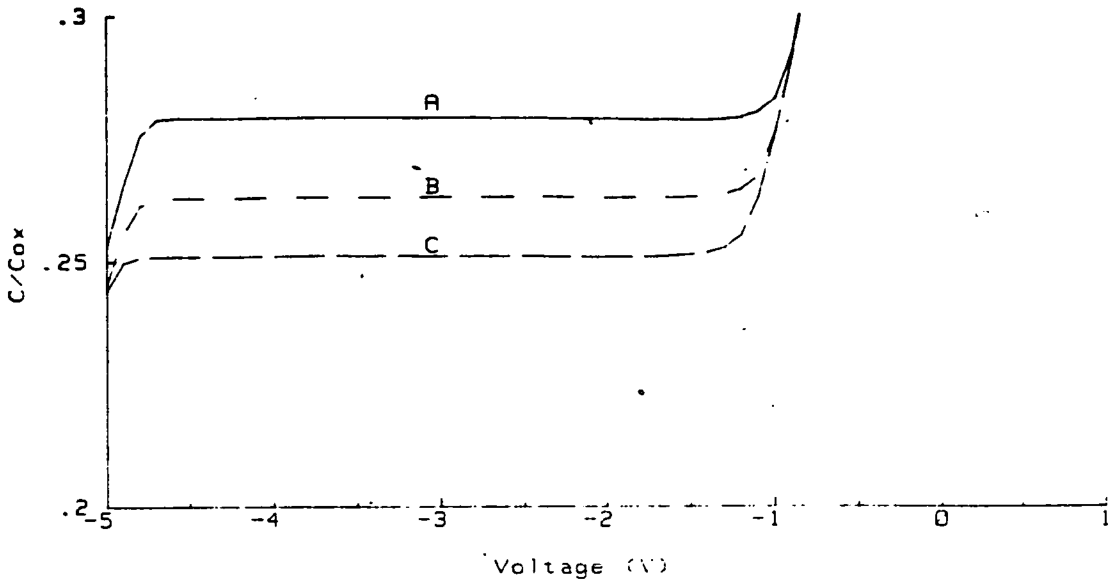


Figure 3.27. CV cnaracteristics for samples measured witn tnree different DC bias settling times, (a) no settling, (b) 1 second, (c) 10 seconds.

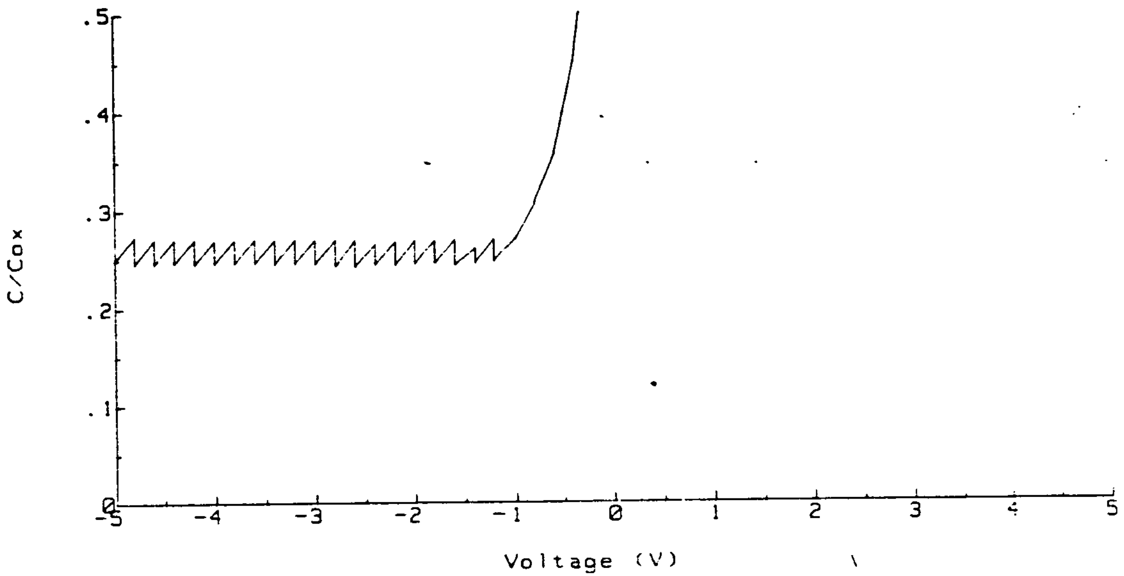


Figure 3.28. Settling of inversion layer capacitance after eacn bias step.

Settling Time (s)	C_{min} ($pF\ cm^{-2}$)	N_{sub} ($x\ 10^{15}\ cm^{-3}$)	V_{fb} (V)
0	11775	2.25	-0.05
0.5	11267	1.97	-0.07
0.75	11177	1.92	-0.07
1.0	11103	1.89	-0.08
5.0	10738	1.71	-0.09
10.0	10597	1.65	-0.10

Table 3.5 The effect of DC bias settling time on measured parameters.

Acceptance Level $(dC/dt)/C$	Hold Time for Acceptance (s)	C_{min} ($pF\ cm^{-2}$)	N_{sub} ($x\ 10^{15}\ cm^{-3}$)
0.001	7.4	10754	1.74
0.0005	43.3	10434	1.58
0.0001	114.4	10317	1.53

Table 3.6 Effect of choosing different criteria for settling.

3.11.3. Alternative methods of determining C_{min}, N_{sub} and V_{fb}

In order to measure the CV characteristic in a reasonable time period it is necessary to adopt one of two alternative sweep procedures. The first method relies on a single point determination of C_{min}. In a similar way to that described by Goetzberger, the sample is briefly illuminated at the initial bias. The true equilibrium capacitance is then determined by waiting until dC/dt reaches zero. The rest of the sweep is performed with a fairly rapid settling time of 0.1 seconds. The strong inversion part of the characteristic with this sweep procedure is not in equilibrium except for the first point. This is not a problem since the single point is adequate to determine C_{min} from which the other parameters are derived.

The correct determination of C_{min} with this method still requires care. Table 5.6 shows that there can be a significant error in C_{min} if very stringent criteria for the establishment of the equilibrium are not applied.

Using a completely different approach, the parameters N_{sub}, V_{fb} and oxide charge levels can be determined without the requirement of measuring C_{min}. (C_{min} itself is not a physically important parameter, since it is only used in an intermediate step in the determination of other parameters). This alternative method involves measuring N_{sub} directly, which can be performed very rapidly with the pulsed capacitance technique (see chapter 5). This approach can be readily inserted into existing CV measurement programs and as an example this author has modified a program to include two pulses from accumulation to deep depletion. With this procedure, which takes less than two seconds, a substrate doping density of 1.49×10^{15} cm^{-3} was measured, which agrees closely with best value achieved in table 3.6.

There are additional advantages with this method since there is no requirement to illuminate the sample and also the sweep can be performed from accumulation to inversion at a fast rate since the equilibrium capacitance in inversion is not required

for any parameter determination. The disadvantage is that it cannot be used with high accuracy for low lifetime samples because, in this case capacitance quickly rises from the deep depletion value to the equilibrium value. This makes accurate determination of the deep depletion capacitance difficult unless a very fast capacitance meter is available. Unfortunately, there is a trade off between measurement speed and accuracy with these instuments.

3.11.4. Conclusions on the Measurement of Parameters for High Lifetime MOSCs

It has been shown that the measurement of an equilibrium high frequency CV characteristic for high lifetime samples is problematic and alternative methods are required if accurate results are to be achieved in a reasonable time. One alternative approach to parameter extraction from the CV curve is to measure N_{sub} directly. It is simple to implement and also is considerably faster than the modification of the conventional approach. Unfortunately this new method is not well suited to low lifetime samples, unless a fast capacitance meter is available. With slow capacitance meters and low lifetime samples the traditional N_{sub} determination is better. Ideally, CV systems should be capable of both types of measurement. For automatic implementation this involves including an additional routine to decide on the basis of the initial transient response which method was most appropriate for the sample. If dC/dt is large after a pulse into deep depletion, then the conventional CV sweep method is appropriate, whereas when dC/dt is small then N_{sub} is better determined using a pulsed approach.

3.12. Conclusions on High Frequency CV Methods

This chapter has extensively reviewed the high frequency CV measurement of MOS capacitors, and shown that although the method does provide useful information there are a large number of factors which can adversely effect the result.

CHAPTER FOUR

INTERFACE TRAPS

4.1. Introduction

Interface traps[9] are defects located at the Si-SiO_2 interface which capture and emit carriers when the surface potential (ψ_s) changes. These defects are of importance because:

(1) Charge held in traps reduces the number of carriers in the surface region at any given gate bias. For example if there are N_t electrons in traps then by the law of mass action (equation 3.1) there will be n-N_t electrons in the conduction band.* This affects MOSFET parameters such as threshold voltage (V_T), transconductance (g), and subthreshold swing (S).

(2) In the DRAM storage capacitor or CCD, interface traps increase surface leakage currents[13]. (See chapter 6 for more details).

There has been speculation that interface traps can be caused by trace contamination of capacitors with alkali or heavy metals, and in chapter 8 of this thesis the effect of surface contamination on interface trap densities is investigated. In this chapter several methods of determining trap densities are assessed experimentally† in order to allow an unambiguous assessment of impurity effects to be made later. The experimental investigation used here is of more practical value than theoretical treatments[9], because the nature of the errors is such that a rigorous quantitative assessment of errors

* Although interface traps are considered to be oxide defects they can exchange charge with the silicon energy bands, and can therefore affect the surface concentration of carriers in the silicon.

† Table 4.1 details the codes and critical processing for samples used in this investigation.

Sample	Process Effecting Trap Density
TR1	Control sample†
TR2	1200°C post oxidation anneal.
TR3	Dual dielectric insulator (nitride-oxide).
TR4	Omission of any PMA.
Al1	Anneal control.†
AlSi1	AlSi gate metal instead of pure Al.
AlSi2	As above but without HCl during oxidation.

Table 4.1 Samples for interface trap measurements.

† Process details are given in table 3 of chapter 8.

is simply not possible even in principle. In addition the experimentally determined plots presented in this chapter provide a more readily understandable demonstration of the practical difficulties which can be encountered than the theoretical expressions for the magnitude of errors.

The final section of this chapter reports on an effect that a subtle change in the composition of an MOS capacitor gate has upon interface trap density. Analysis of this effect was only possible using an accurate, high resolution measurement of interface traps.

4.2. The Effect of Interface Traps on MOSC Characteristics[9, 120, 136, 157, 158]

At any gate bias those traps which lie above the fermi level will be empty and those below it will be full. Changes in bias will result in changes in the occupancy of traps. For example if gate bias is swept from flatbands to depletion in a p-type sample, then traps are filled by hole emission, and when sweeping in the opposite direction they empty by hole capture. The charge transfer processes are analogous to those described for bulk traps in chapter 3. However, the time constants which characterise the two processes are very different, because of the different energy distribution and position in the silicon of the two types of defect.

In admittance (Y) measurements, traps are detected by applying a small ac voltage on top of the applied gate (dc) voltage. The ac signal causes both capture and emission processes to occur in one cycle, which affects the measured capacitance (C_m) and conductance (G_m).

Since the charge held in traps (Q_{it}) varies during the ac cycle it will have an associated capacitance C_{it}, such that

$$C_{it} = \frac{dQ_{it}}{d\psi_s} \qquad (4.1)$$

C_{it} will be frequency dependent because the capture and emission processes take a finite time. At sufficiently high frequencies the interface trapped charge cannot follow the ac signal at all, which implies that $dQ_{it}/d\psi_s=0$, and hence $C_{it}=0$. The equivalent circuit of a MOSC at high frequencies is shown in figure 4.1(a). In contrast at sufficiently low frequencies traps follow the ac signal without any delay. In this case the equivalent circuit becomes 4.2(b), where

$$C_{it} = qD_{it}(E) \tag{4.2}$$

where $D_{it}(E)$ is the density of interface traps at the energy E in the silicon bandgap. Bandgap energy is changed by sweeping the gate voltage.

At intermediate frequencies, C_{it} will vary between 0 and the value given by equation 4.2. At these frequencies there is a proportion of the charge which lags behind the ac cycle leading to non-equilibrium distributions of charge. To restore equilibrium, energy which is measurable as an equivalent parallel conductance (G_p), is exchanged with the lattice.

The original model[159] for the influence of interface traps on conductance characteristics implicitly assumed that all interface traps were identical, i.e. all located at one energy in the silicon bandgap. The equivalent circuit for this model is shown in figure 4.2a, for which

$$\frac{G_p}{\omega} = \frac{C_{it}}{(1+(\omega\tau)^2)} \tag{4.3}$$

where τ is the (RC) time constant of the interface trap admittance, such that

$$\tau = \frac{C_{it}}{G_p} \tag{4.4}$$

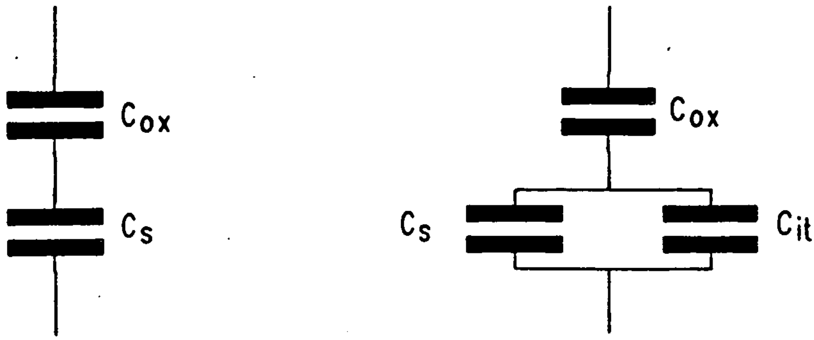


Figure 4.1 Equivalent circuits for the MOS Capacitor with interface traps at (a) high frequency, and (b) low frequency.

An alternative expression for τ in terms of physical quantities is[12]

$$\tau = \frac{1}{c_p N_{sub}} \exp(\pm \nu_s), \qquad + \textit{holes}, \quad - \textit{electrons} \tag{4.5}$$

where c_p is the capture probability, and $\nu_s = \beta\psi_s$ is the normalised band bending. Plotting G_p/ω against $\ln\omega$, or G_p/ω against ν_s gives a singly peaked curve.

In real MOS systems there will always be a distribution of interface states across the bandgap. This is modelled in the equivalent circuit shown in figure 4.2(b), for which

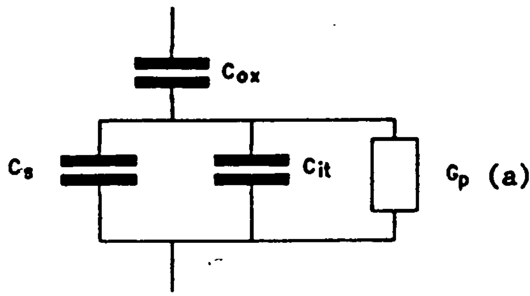
$$\frac{G_p}{\omega} = \frac{C_{it}}{2\omega\tau} \ln(1 + (\omega\tau)^2) \tag{4.6}$$

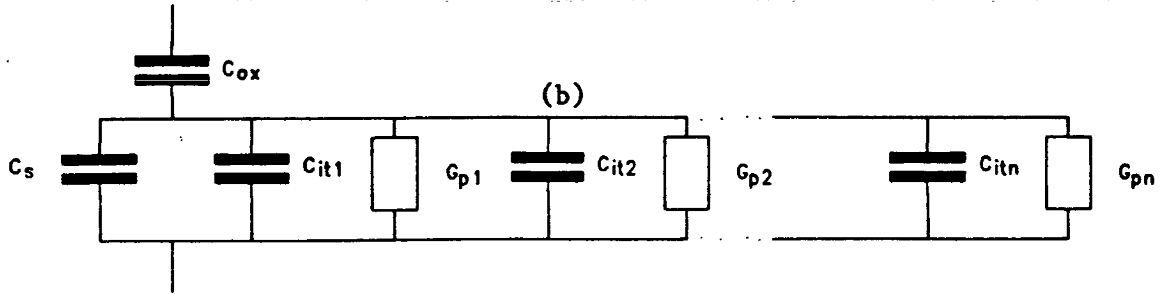
which peaks at higher values of $\omega\tau$ and has a lower maximum than 4.3. However when the $G_p/\omega - \ln\omega$ or $G_p/\omega - \nu_s$ curves of real MOS capacitors are measured, the observed dispersion in the curves is far greater than predicted by either 4.3 or 4.6. The accepted cause of this extra broadening is a random distribution of point charges across the interface which leads to a Poisson distribution of surface potential across the interfacial plane[136]. These charges cause band bending fluctuations at the interface, making the equivalent circuit of the MOSC that shown in figure 4.2(c), for which

$$\frac{G_p}{\omega} = \frac{qD_{it}(E)(2\pi\sigma_s^2)^{-\frac{1}{2}}}{2\xi} \int_{-\infty}^{+\infty} \exp\left(-\frac{\eta}{2\sigma_s^2}\right) \exp(-\eta) \ln\left(1 + \xi^2 \exp(2\eta)\right) d\eta \tag{4.7}$$

where the parameter η is defined by,

$$\eta = \nu_s - \langle \nu_s \rangle \tag{4.8}$$





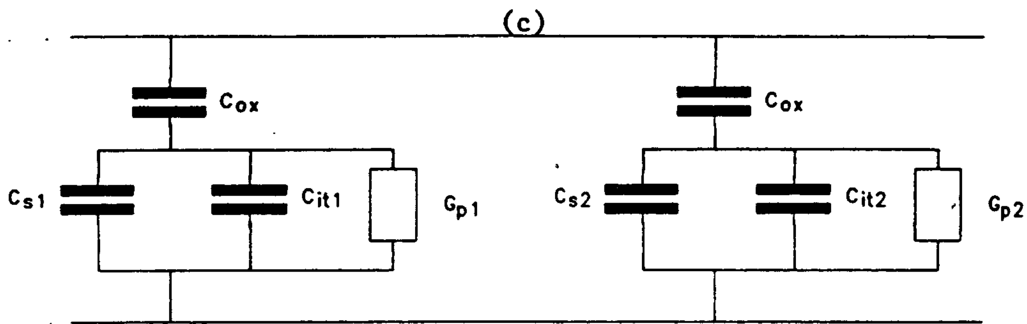


Figure 4.2 Equivalent circuits for the three interface trap models. These are: (a) the single trap model; (b) the distributed trap model; and (c) the distributed trap model with band bending fluctuations.

with $<\nu_s>$ being the mean normalised band bending, and

$$\xi = \omega\tau \; . \tag{4.9}$$

A gaussian distribution of potential variations, which is characterised by a variance σ_s^2, adequately fits the experimentally observed broadening.

Figure 4.3 compares $G_p - \nu_s$ plots using equations 4.3, 4.6 and 4.7, with fixed values for interface trap density and capture probability. Figure 4.4 compares a measured $G_p - \nu_s$ curve with one which was calculated using 4.7. The agreement is very good and gives strong support for the model.

4.3. Determination of G_p from the Measured Conductance(G_m)

The measured conductance (G_m) of an MOS capacitor differs from G_p because of the effects of series resistance, which alters the equivalent circuit from that in 4.5(a) to 4.5(b), and because the oxide capacitance changes the time constant of the circuit[136].

Series resistance is corrected using

$$G_p = \frac{(G_m + \omega^2 C^2) a}{a^2 + \omega^2 C^2} \tag{4.10}$$

where,

$$a = G_m - (G_m^2 + \omega^2 C^2) R_s \tag{4.11}$$

and R_s is determined by one of the methods discussed in section 3.6.1. Figure 4.6 illustrates plots of G_m against V and G_p against V, where it can be readily seen that influence of R_s is far more severe in the conductance plots than it is in the CV plots

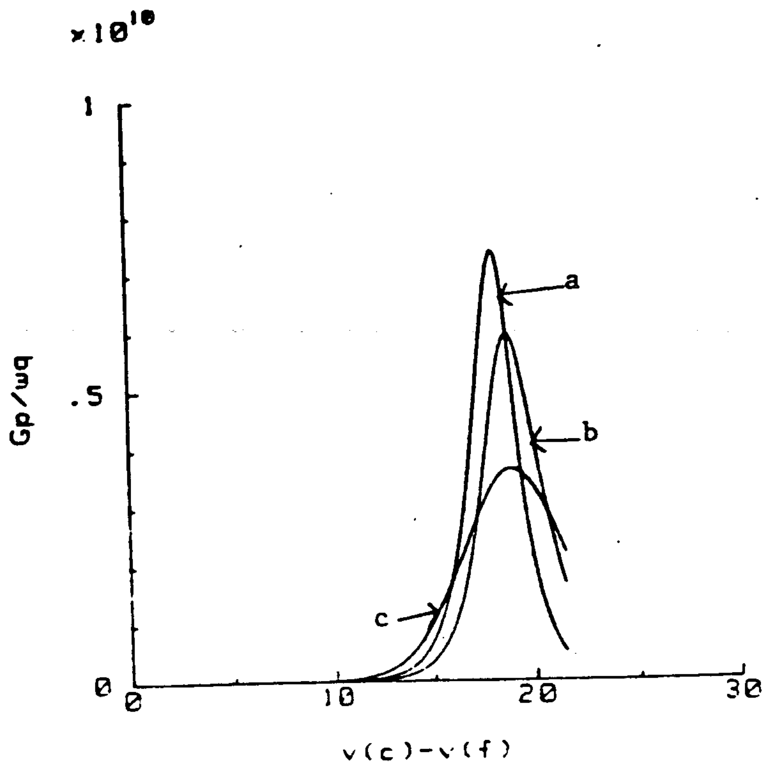


Figure 4.3 Theoretical conductance-band bending plots for the three equivalent circuits shown in figure 4.2. The peaks are: (a) the single trap model; (b) the distributed trap model; (c) the distributed single trap model with band bending fluctuations.

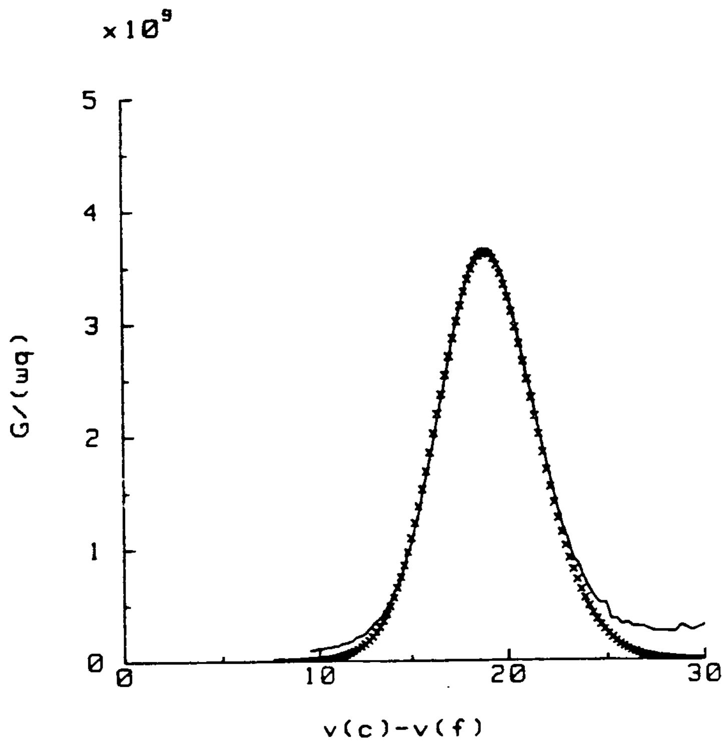


Figure 4.4. Fit of the theoretical conductance-band bending model (crosses), to experimental data.

discussed in chapter 3. For this reason equation 4.10 should always be used to correct data. R_s obviously limits the accuracy with which the interface trap conductance can be extracted and so sample preparation methods which minimise R_s should be adopted, (see chapter 3).

The effect of oxide capacitance on the gate bias and frequency response of the interface trap conductance can be corrected for using

$$G_p^* = \frac{\omega C_{ox}^2 G_p}{G_p^2 + \omega^2 (C_{ox} - C_m)^2} \tag{4.12}$$

where G_p^* is now only due to interface traps. The error introduced by this equation is less for thinner oxides.

4.4. Methods for Determining $D_{it}(E)$ and $\psi_s - V_G$ from Admittance Measurements

4.4.1. The Low Frequency CV Method for Determining $D_{it}(E)$

Low frequency CV measurements[9, 157, 158, 160] can be used to determine both the density of interface traps and their energy distribution in the silicon bandgap, i.e. $D_{it}(E)$ against E. The low frequency CV characteristic by definition has a silicon capacitance described by equation 3.7 and an interface trap capacitance given by 4.1.

Trap energy E_T and surface potential are related by[9]

$$E_c - E_T = E_g/2 + q\psi_s - q\phi_b \tag{4.13}$$

or

$$E_T - E_v = E_g/2 + q\psi_s - q\phi_b \tag{4.14}$$

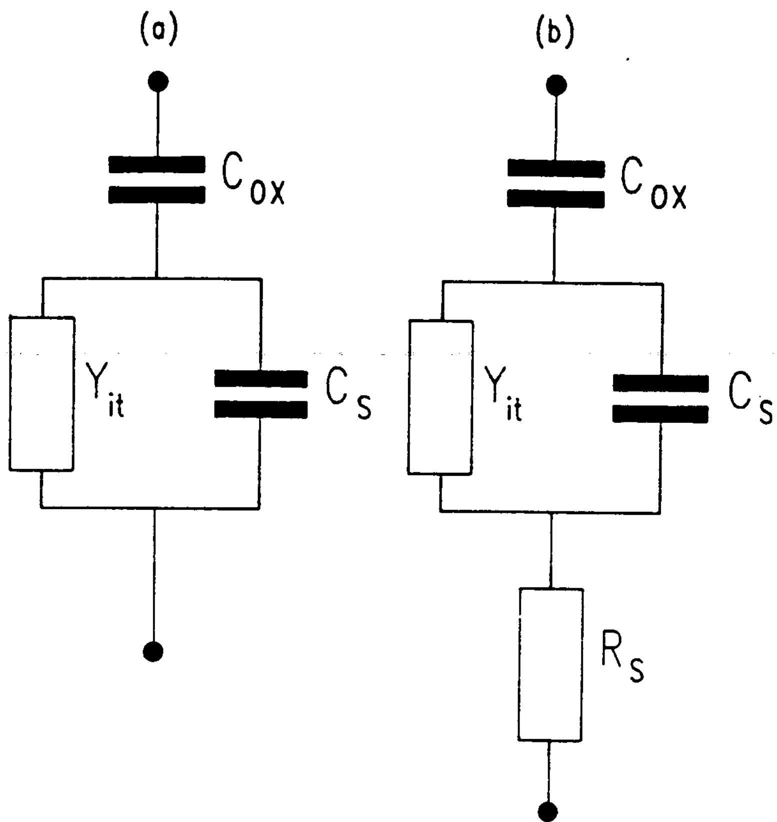


Figure 4.5. Equivalent circuits for MOS capacitors, (a) without series resistance, and (b) with series resistance. N.B. $Y=G+j\omega C$

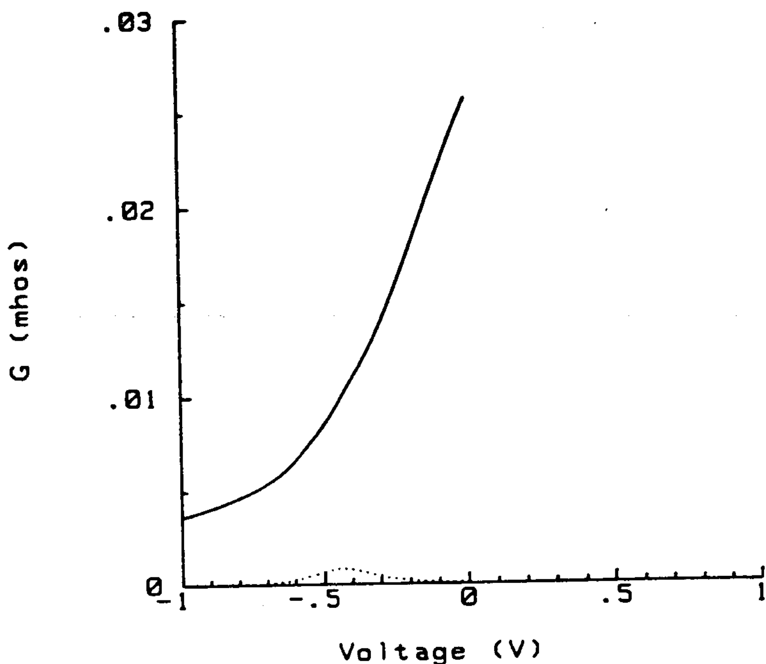


Figure 4.6 (a) Measured conductance (——), and interface trap conductance (......), as a function of voltage, for sample TR1 with 1 MHz. measuring signal. Trap density was 2 x 10^{10}, and series resistance was 20Ω.

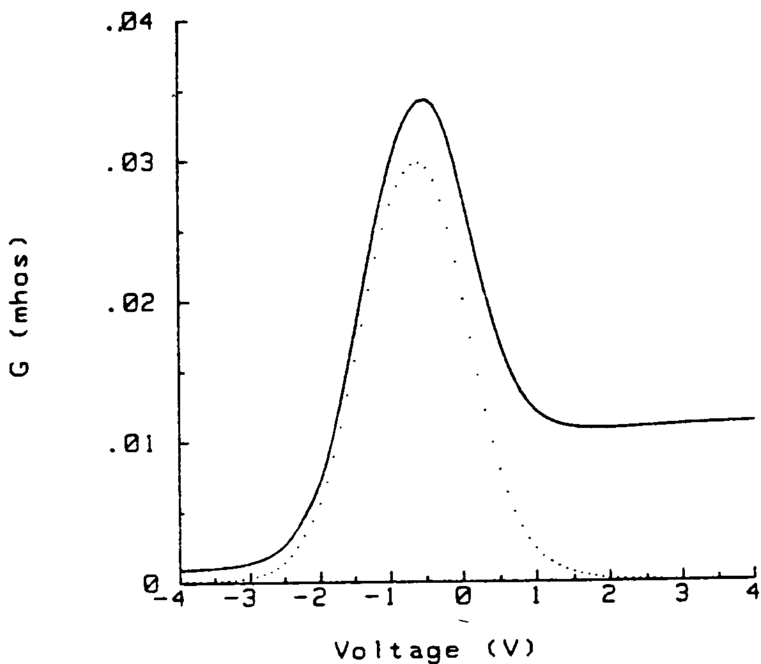


Figure 4.6 (b) Measured conductance (——), and interface trap conductance (.....), as a function of voltage, for sample TR4 with 1 MHz. measuring signal. Trap density was 1 x 10^{12}, series resistance was 20Ω.

and so if the relationship between gate voltage and surface potential can be related, then trap energy as a function of voltage can be found. Berglund[158] showed that

$$\psi_s(V_1) = \int_{V_1}^{V_2} [1 - \frac{C^{LF}}{C_{ox}}] + \psi_s(V_2) \quad (4.15)$$

where C^{LF} is the measured low frequency MOS capacitance, and V_1 and V_2 are arbitrary voltages. Thus by integrating 4.15 the shape of the $\psi_s - V_G$ curve can be determined, and if the additive constant $\psi_s(V_2)$ can be determined then absolute values can be assigned to the x-axis. Several methods for determining $\psi_s(V_2)$ have been proposed[158, 157, 136], all of which are only of limited accuracy. Figure 4.7 shows measured and theoretical $\psi_s - V_G$ plots where $\psi_s(V_2)=0$, because the integration limit V_2 was set equal to V_{fb} which was determined by the method set out in chapter 3.

The equivalent circuit 4.1(a), for an MOS capacitor measured at low frequency has capacitance C^{LF}, given by

$$\frac{1}{C^{LF}} = \frac{1}{C_{ox}} + \frac{1}{C_s + C_{it}} \quad (4.16)$$

where in the ideal case $C_{it}=0$. If theoretical values of C_s are determined, then C_{it} can be found by subtracting C_s from the measured sum of $C_s + C_{it}$, then the trap density is inferred from 4.1. Figure 4.8 shows measured and ideal low frequency CV plots for two samples, and figure 4.9 shows the corresponding plots of $D_{it}(E)$ as a function of silicon bandgap energy.

The frequencies required to obtain accurate values of C^{LF} are typically only a few herz which is too low for measurements to be made on admittance bridge apparatus, and so alternative methods of determining the low frequency characteristic have been developed.

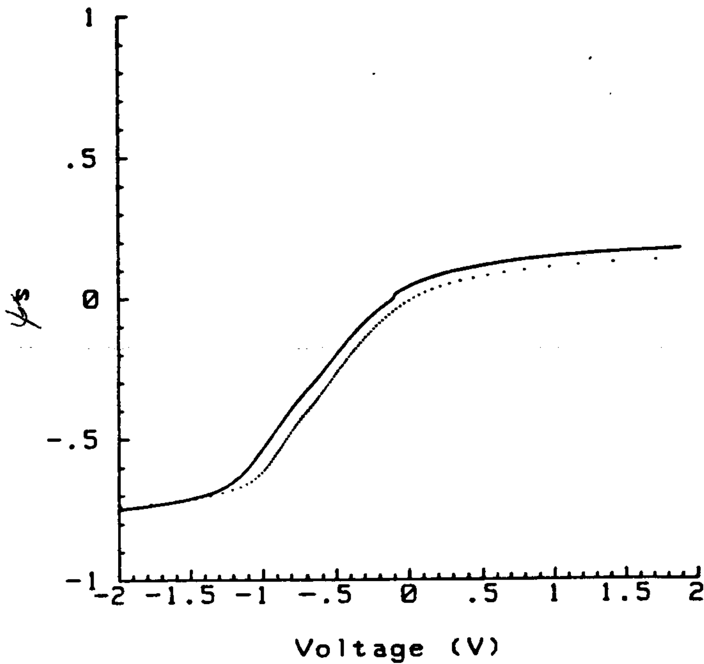


Figure 4.7 (a) Measured (———), and ideal (······), band bending versus gate voltage plots, determined by the low frequency CV method, for sample TR5, which had low minority carrier generation lifetime.

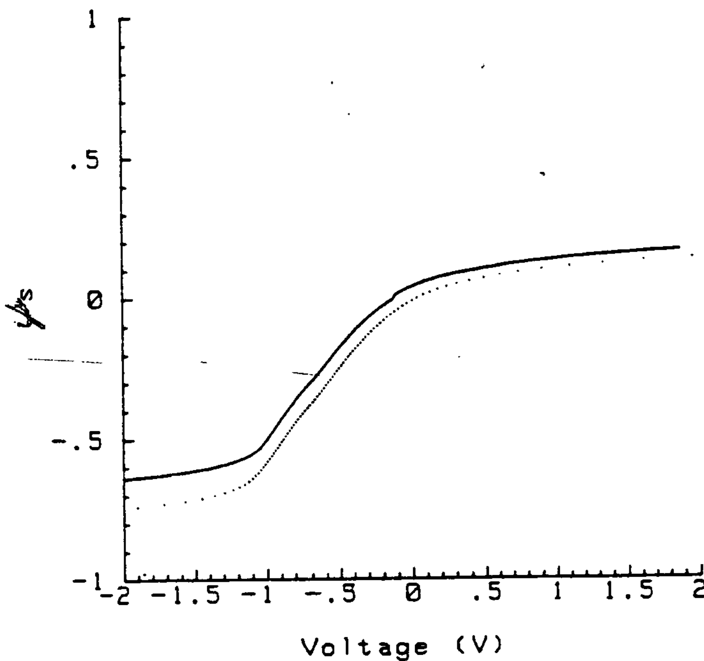


Figure 4.7 (b) Measured (———), and ideal (......), band bending versus gate voltage plots, for sample TR1, which had high minority carrier generation lifetime. In this measurement the carrier lifetime makes it impossible to determine the true low frequency characteristic in inversion, and hence the deviation from the ideal curve shape shown here.

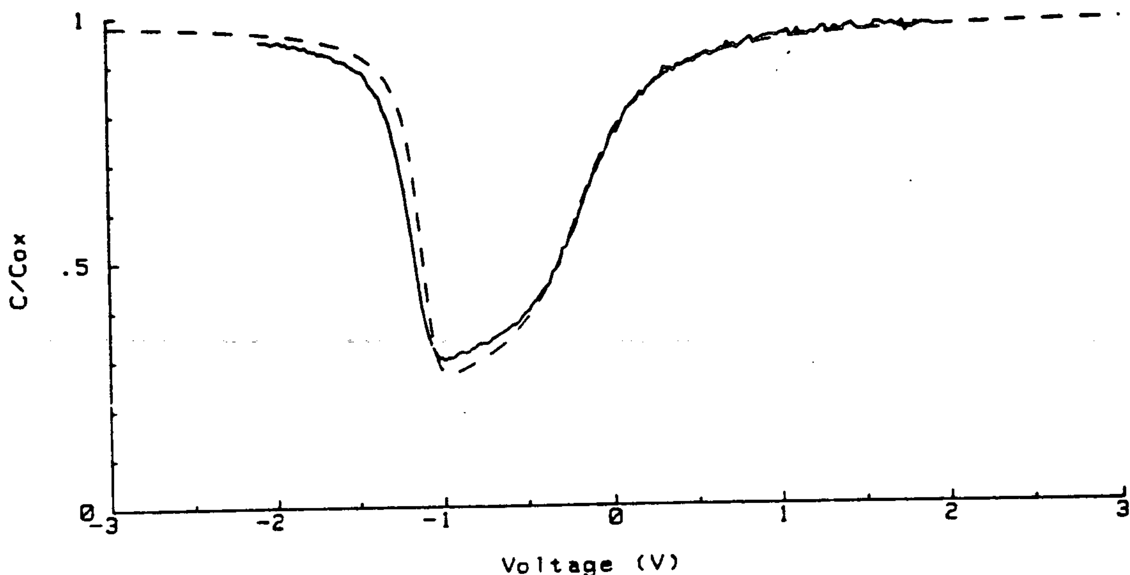


Figure 4.8 (a) Measured (———), and ideal (.....), low frequency CV plots for the sample in figure 4.7(a). This curve was a rarity in the samples fabricated for this thesis, in that a complete low frequency curve could be measured at a ramp rate of 10 millivolts per second.

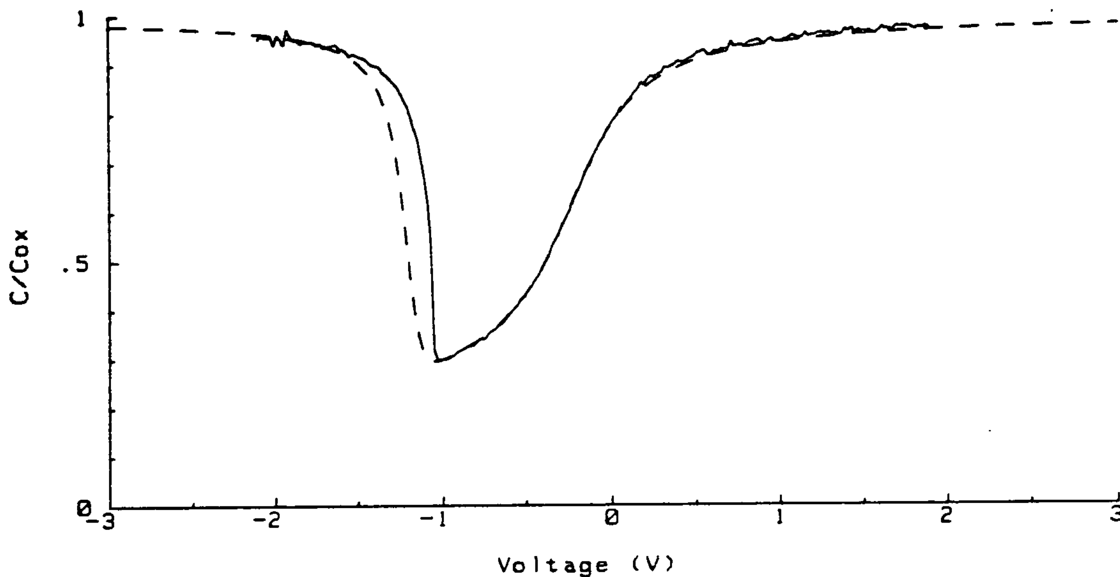


Figure 4.8 (b) Measured (———), and ideal (------), low frequency CV characteristics for sample TR1. This is typical of the result for moderate and high lifetime samples. Even with the lowest practical sweep rate the inversion region cannot be measured in equilibrium.

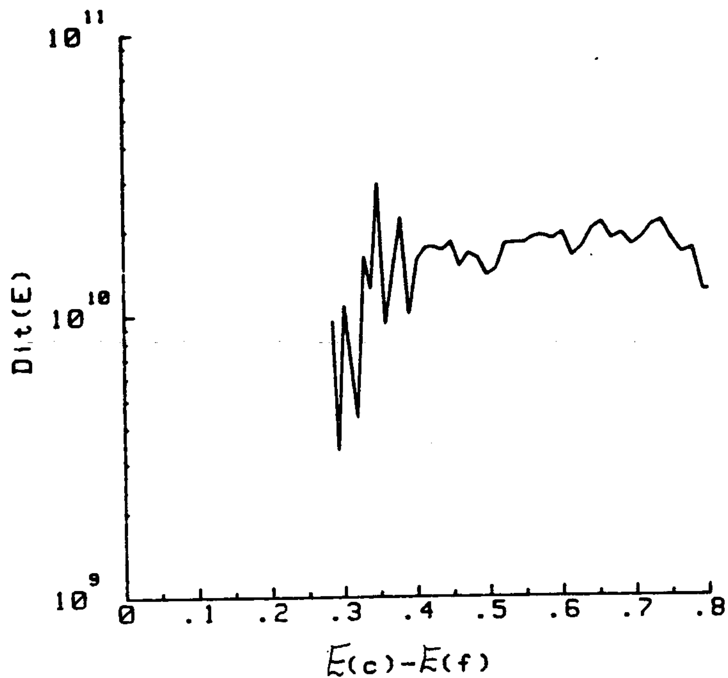


Figure 4.9 (a) Interface trap density determined by the low frequency CV method for sample TR5.

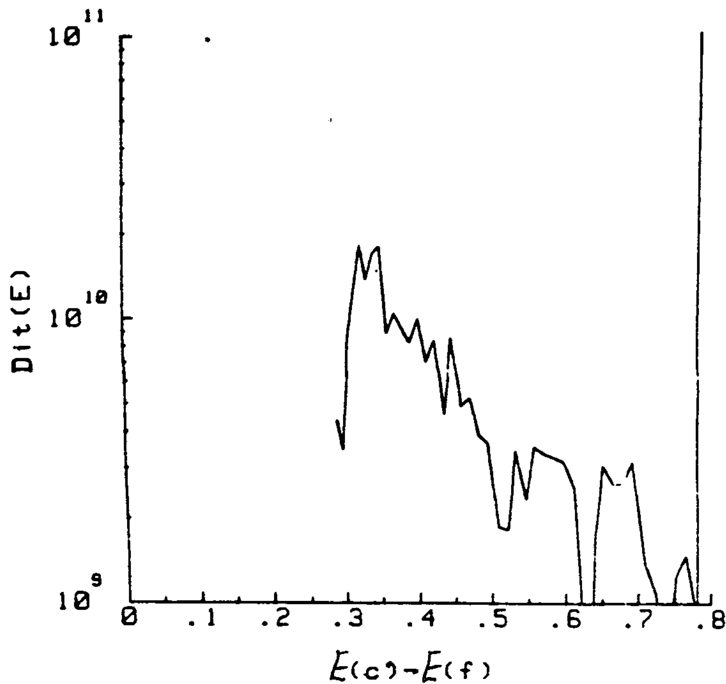


Figure 4.9 (b) Interface trap density determined by the low frequency CV method for sample TR1.

The most common method of determining C^{LF} is the Quasi-Static CV (QSCV) method which was proposed by Kuhn[160], where a linear voltage ramp of slope $\alpha = (dV/dt)$ causes a charging current (I) to flow, in which case

$$I = \frac{dQ}{dt} = \left(\frac{dV}{dt}\right)\left(\frac{dQ}{dV}\right) = \alpha C^{LF} \tag{4.17}$$

Very slow ramps of $50mVs^{-1}$ or less, are required to measure equilibrium values of C^{LF}, and in weak or strong inversion it becomes impossible to measure C^{LF} correctly at even the slowest practicable sweep rates, when minority carrier lifetime is even moderately high. This is why the measured and theoretical curves compare badly in figures 4.7 and 4.8 in strong inversion.

Ziegler and Klausman[161] introduced the QV method which determines $\psi_s - V_G$ directly using a charge sharing method, from which C^{LF} can be determined by integration. To date, reports of the use of the QV technique are rare compared to those using the QSCV method, despite the former's potentially superior accuracy.

The low frequency method is probably the most commonly used method for determining interface trap densities, because in principle it has the potential to measure both $D_{it}(E)$ and $\psi_s - V_G$ over a wider range of bandgap energies than the other method reported here. However, in practice I have found that this potential is never properly realised, principally because the overwhelming majority of capacitors produced at the present time have very high minority carrier lifetimes. In addition, the low frequency CV method has the following disadvantages:

(1) A theoretical CV curve must be calculated, which necessitates the measurement of the doping density for uniformly doped samples or the complete doping profile in the non-uniformly doped case. These profiles require high frequency CV measurements which use different equipment. Errors in doping profiles, or

redistribution of dopants in samples assumed to be uniformly doped can lead to serious errors in the values determined for $D_{it}(E)$ and the $\psi_s - V_G$ curve[9, 143].

(2) The QSCV measurement is very sensitive to noise because it is a very low current measurement. This can be seen in figure 4.8 where capacitance is inferred from currents of approximately 10pA.

(3) C_{it} is measured in the presence of C_s, and so it can only be determined with confidence when it is larger than about $0.1C_s$, bearing in mind that C_s values can easily have an error greater than this from (1) or (2). A converse argument puts an upper limit on the accurate determination of high values of C_{it}.

The problems with the low frequency method mean that in the vast majority of cases the values of $D_{it}(E)$ which are determined are of only very limited accuracy. The method's primary use should be for the determination of $\psi_s - V_G$ for samples with near uniform doping where the integration constant can be determined with reasonable accuracy.

4.4.2. The High Frequency CV Method[120, 127]

At high frequencies, interface traps do not contribute an additional capacitance as occurs in low frequency measurement. However the high frequency CV characteristic is effected by the response of traps to dc bias changes.

When charge is added to the gate of an MOS capacitor, and a proportion of that charge is balanced by the filling or emptying of interface traps, then the surface potential need only change by a lesser amount to bring about a complete balance, than would be the case when no traps were present. Thus interface traps have the effect of stretching out the high frequency CV plot along the voltage axis. An example of this effect is shown in figure 4.10.

The $\psi_s - V_G$ curve can be determined in depletion and weak inversion from the measured CV plot and a theoretical plot of C^{HF} against ψ_s. At any value of capacitance on the measured curve there will be a corresponding point on the theoretical curve which occurs at a surface potential ψ_s, which will also be the surface potential in the measured sample.

Interface trap density is inferred from the degree of stretchout in the measured CV plot along the voltage axis using

$$C_{it} = C_{ox}[(\frac{dV_G}{d\psi_s})-1]-C_s \tag{4.18}$$

and $D_{it}(E)$ is determined using 4.2.

$\psi_s - V_G$ curves and the derivative in 4.18 were originally determined by a manual interpolation and differentiation[120], which is the method that is still suggested in Nicollian and Brews'[9] much more recent book. This makes the high frequency method a tedious and slow procedure. However, it need not be if the analysis is automated by using a computer program to perform the relatively straightforward inverse interpolation and numerical differentiation. It is likely that those who used the high frequency method in the past had access to a computer in order to calculate the ideal high frequency CV and $C^{HF} - \psi_s$ plots, and it therefore seems odd that they did not fully utilise that resource. Figure 4.11 shows measured and ideal $\psi_s - V_G$ curves determined by an automated high frequency CV analysis where the inverse interpolation was performed by a cubic spline method. (Less sophisticated interpolation would probably have been sufficient.) Figure 4.12 shows the interface trap density plot which was determined for the sample in figure 4.11.

The major advantages of the high frequency CV method over the low frequency CV method are that it uses the same equipment as the measurements described in

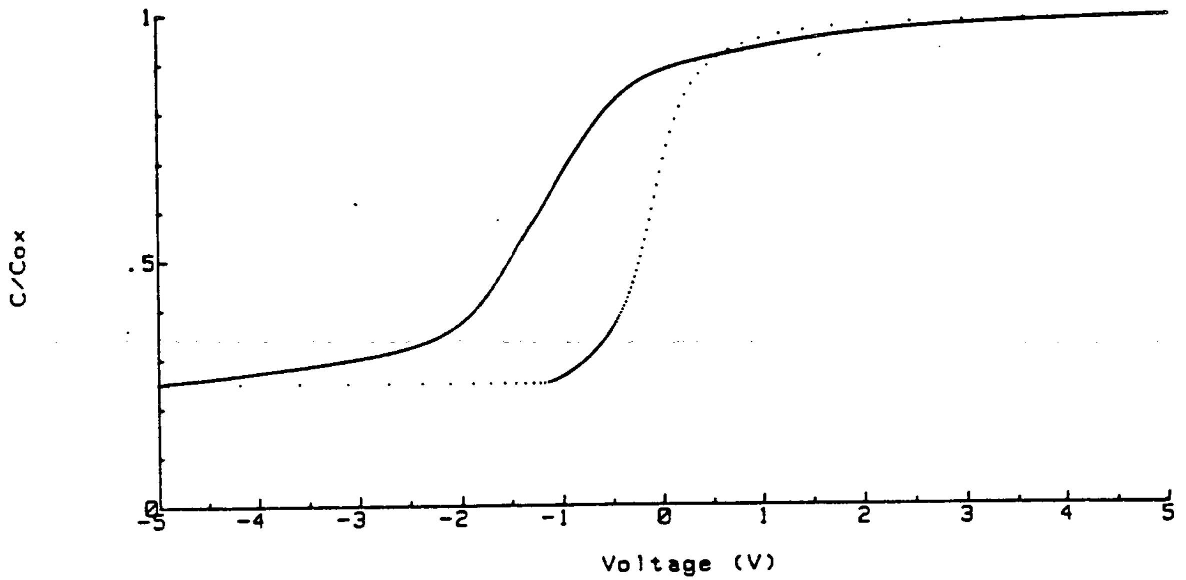


Figure 4.10 The high frequency CV characteristic of sample TR4 which has a high trap density (———), and the ideal (·····) CV characteristic for sample TR4 with no interface traps.

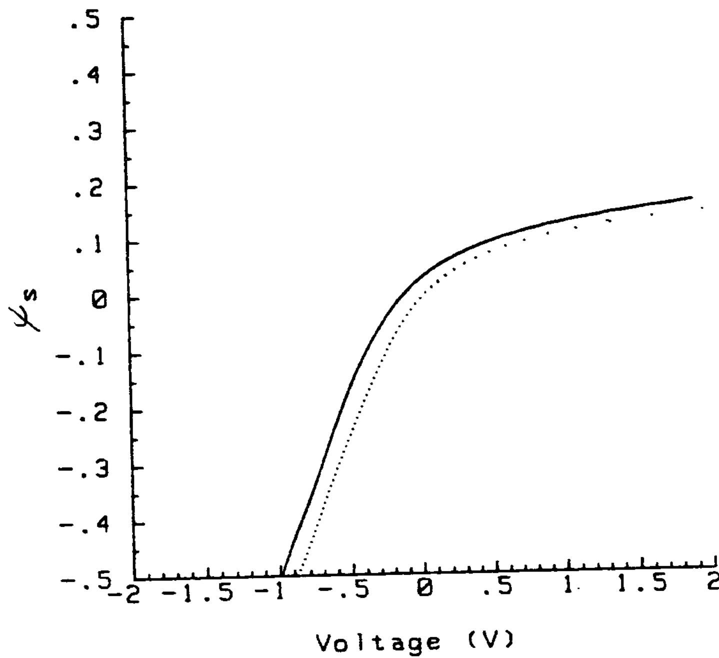


Figure 4.11 Measured (———), and ideal (·····), band bending versus gate voltage plots, determined by the high frequency CV method, for sample TR1.

chapters 3, 5, and 6 and higher trap densities can be measured whilst the lower resolution limit is similar.

Like the low frequency method there remains the necessity to determine a theoretical CV curve and thus errors in doping profiles will cause error in both C_s and the derivative in 4.18. However, because of the better signal-to-noise ratio of the high frequency method compared to the low frequency method, the total error and resolution limit ought to be comparable for both analyses.

$D_{it}(E)$ values from the high frequency method should be treated as only semi-quantitative. However, the $\psi_s - V_G$ curve will be more accurate in depletion not too close to flatbands.

4.4.3. The Combined High and Low Frequency CV Method[162]

If both low and high frequency CV curves are *measured* , then a theoretical calculation of C_s need not be made since both the measured curves contain the same doping profile information, i.e. C_s, and only differ in depletion because of interface trap capacitance. At any voltage in depletion therefore,

$$C^{LF} - C^{HF} = \frac{C_{ox}(C_s + C_{it})}{C_{ox} + C_s + C_{it}} - \frac{(C_{ox}C_s)}{C_{ox} + C_s} = C_{it} \tag{4.19}$$

Inaccurate doping profile data do not contribute any error to values of $D_{it}(E)$ determined in this way. The major errors are noise in the measured data, particularly from a QSCV determination of C^{LF} or round off errors at low trap densities when the difference between the two measured curves is very small. As usual $D_{it}(E)$ is found from 4.1 and the $\psi_s - V_G$ relationship can be found by either the low or high frequency methods. An additional advantage in using 4.19 to determine trap densities is that it applies for any profile. However problems remain in determining the position for

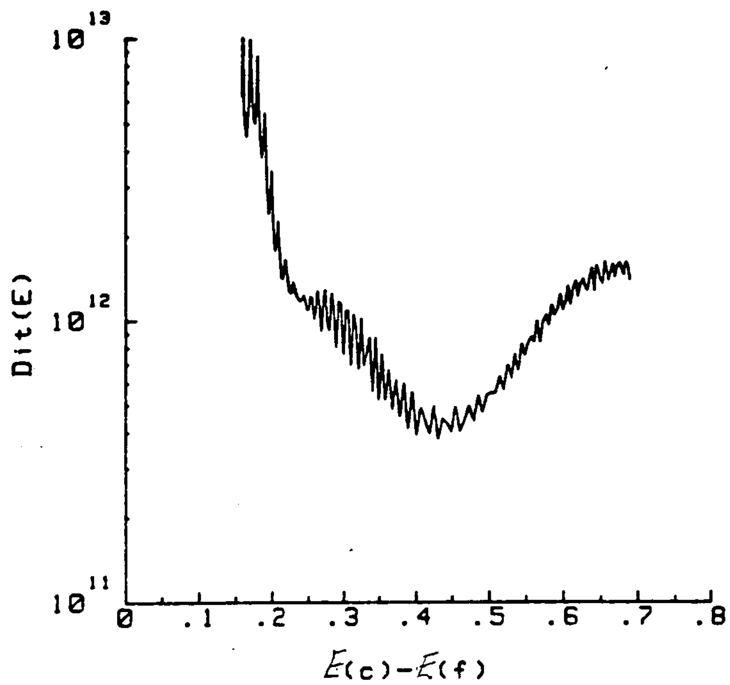


Figure 4.12 Interface trap density as a function of bandgap energy for sample TK1, determined by the high frequency CV method.

non-uniform profiles.

Figure 4.13 compares the trap densities measured by the four different capacitance methods, for samples with differing trap densities. These plots will be discussed in more detail in section 4.5 where the capacitive determination of $D_{it}(E)$ is compared with a conductance method.

4.4.4. The Conductance Methods[12, 136, 163, 164, 165]

Interface trap densities *and* their capture probabilities, c_p can be determined from $G_p/\omega - \nu_s$ or $G_p/\omega - \ln\omega$ plots†. The $\psi_s - V_g$ relationship can be determined by either of the CV methods or the QV method.

Equation 4.7 has shown that the peak shape in the conductance plots is a function of two parameters ξ and σ_s, and therefore a systematic approach to the analysis of curves is required to avoid wasteful trial and error fitting of data.

Nicollian et al.[164] proposed a method for determining $D_{it}(E)$ and c_p from: (i) one low frequency CV curve, (ii) one $G_p/\omega - \ln\omega$ plot and , (iii) GV plots measured at two frequencies. This method is only of very limited accuracy because it assumes that σ_s is a constant, which in the experience of this author is never a satisfactory assumption. This poor assumption results in different values of $D_{it}(E)$ being obtained for different choices of frequency at which the GV data is measured.

Other analyses of conductance data proposed by Simonne[163] and Goetzberger et al.[164] which involve measuring a set of $G_p/\omega - \ln\omega$ plots, do not assume a constant value of σ_s, and are therefore give an unambiguous determination of $D_{it}(E)$ and c_p. One advantage claimed of the Simonne method is that it does not require computer calculations, however the corrections (equations 4.10 and 4.12) must always be

† Before beginning parameter determinations the corrections in equations 4.10 and 4.12 have to be performed.

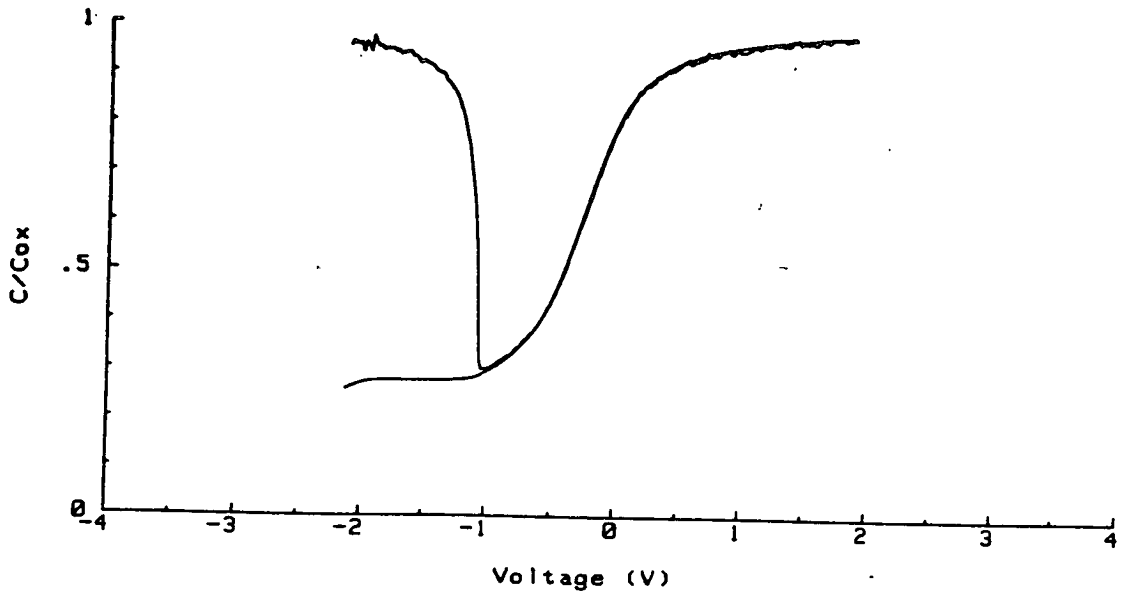


Figure 4.13 (a) Measured high and low frequency CV characteristics for sample TR1.

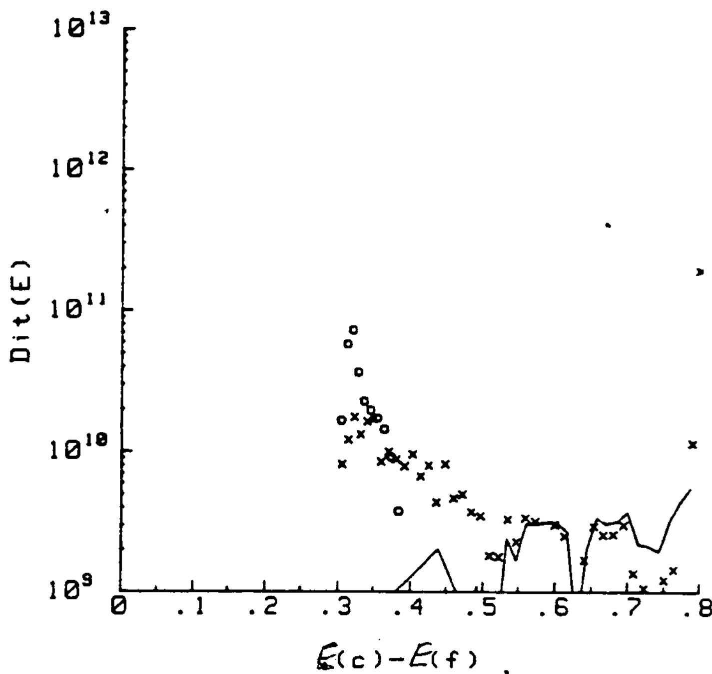


Figure 4.13 (b) Interface trap density as a function of bandgap energy, for sample TR2, determined by the three capacitance techniques which are: the low frequency method (x), the high frequency method (o); and the combined high and low frequency method (———).

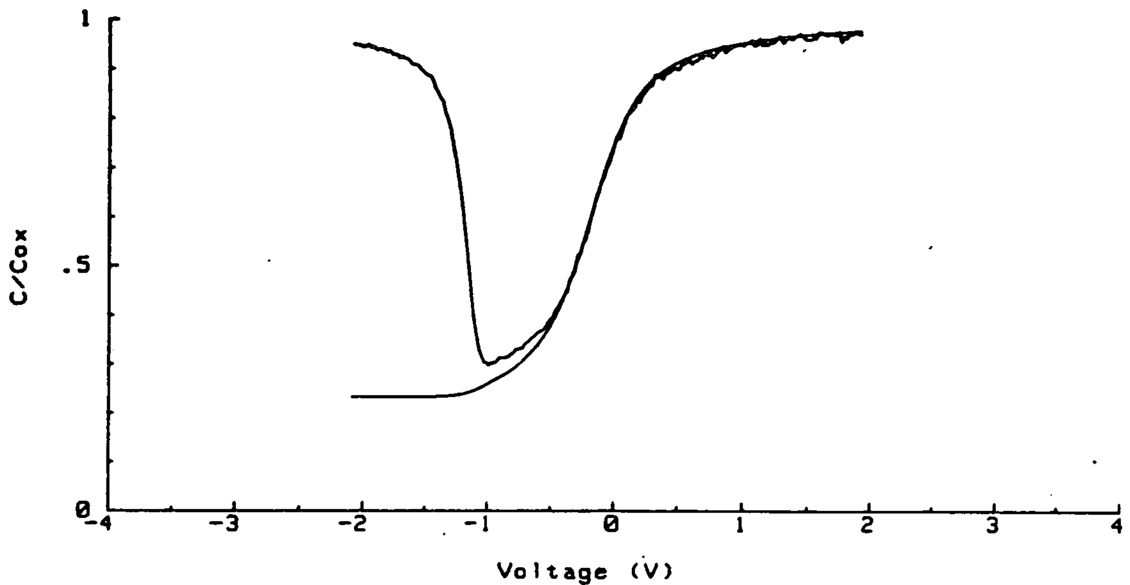


Figure 4.13 (c) Measured high and low frequency CV characteristics for sample TR2.

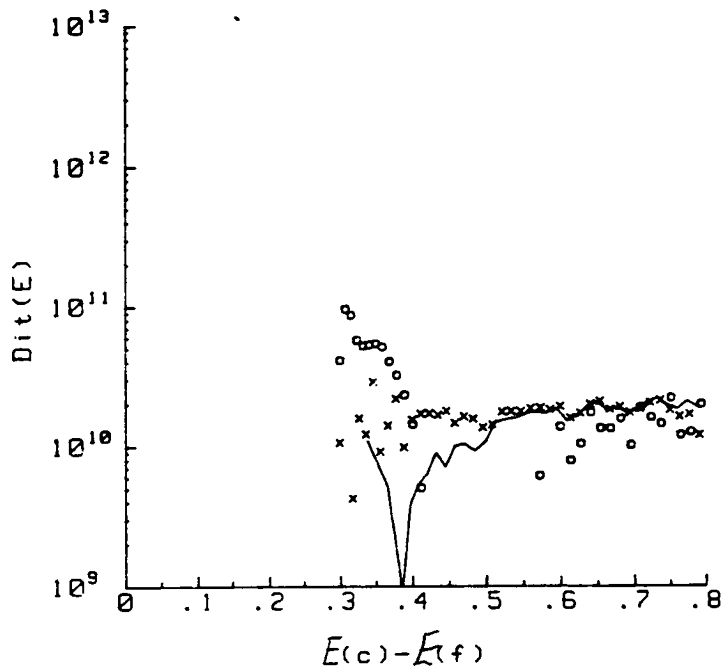


Figure 4.13 (d) Interface trap density as a function of bandgap energy, for sample TR2, determined by the three capacitance techniques which are: the low frequency method (x), the high frequency method (o); and the combined high and low frequency method (_____).

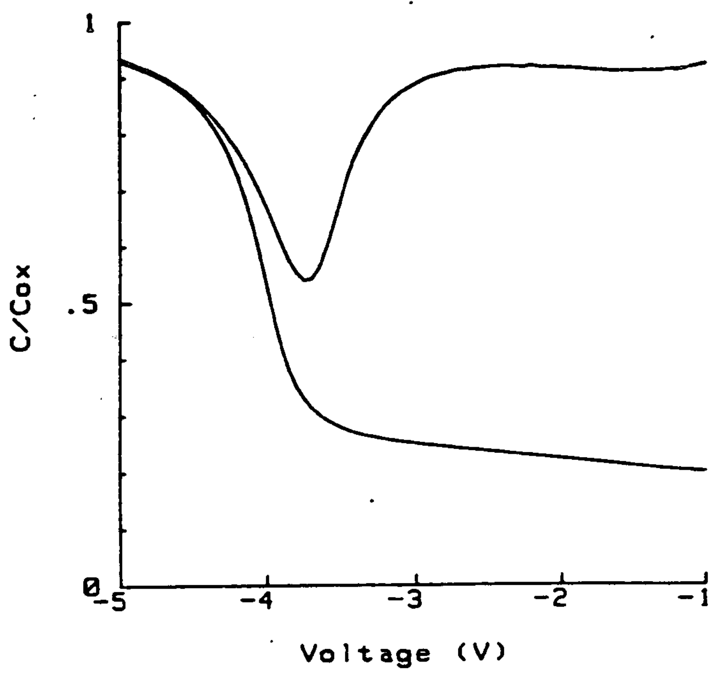


Figure 4.13 (e) Measured high and low frequency CV characteristics for sample TR3.

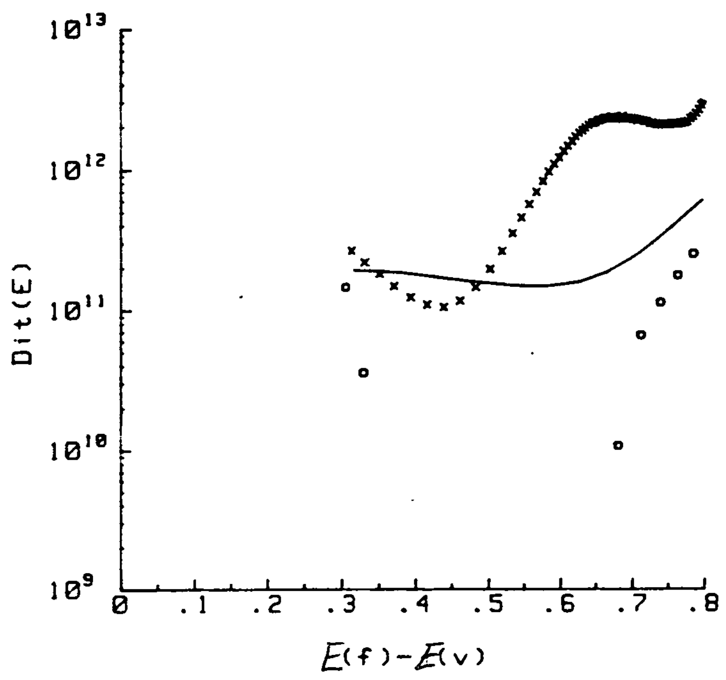


Figure 4.13 (f) Interface trap density as a function of bandgap energy, for sample TR3, determined by the three capacitance techniques which are: the low frequency method (x), the high frequency method (o); and the combined high and low frequency method (———).

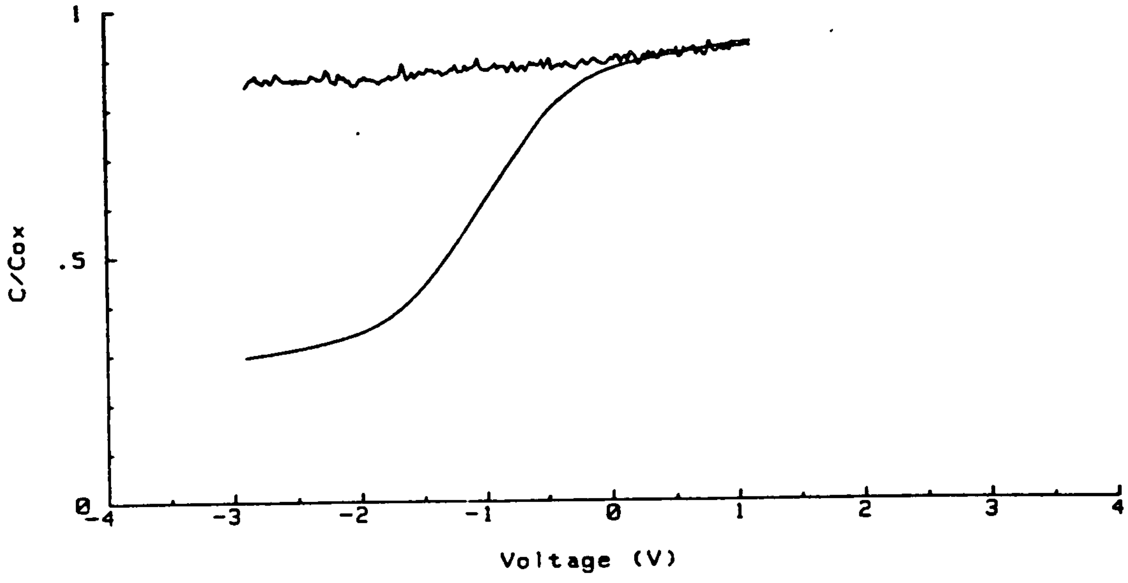


Figure 4.13 (g) Measured high and low frequency CV characteristics for sample TR4.

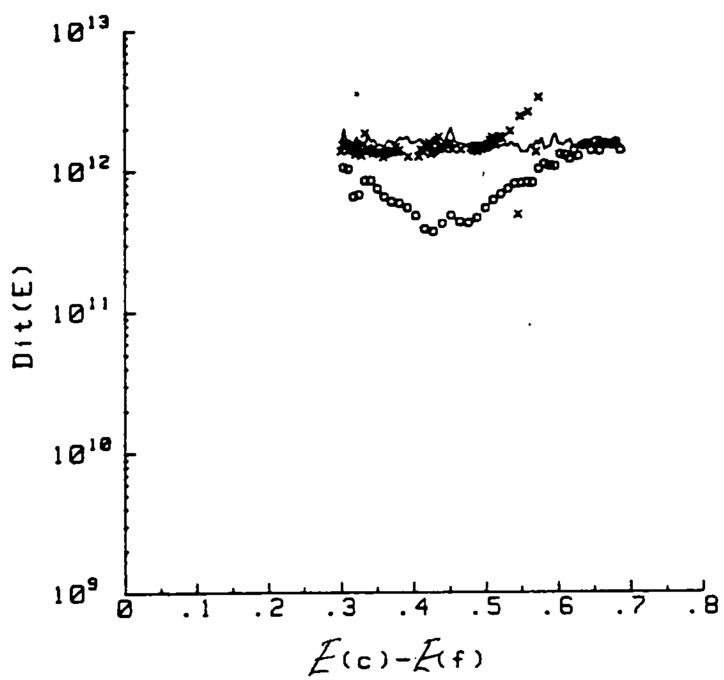


Figure 4.13 (h) Interface trap density as a function of bandgap energy, for sample TR1, determined by the three capacitance techniques which are: the low frequency method (x), the high frequency method (o); and the combined high and low frequency method (———).

made to measured data which would be a laborious task to perform manually, and therefore this is not an advantage particularly at the present time when relatively cheap computers are widely available.

Brews[12] has recently introduced an analysis which can be applied to either $G_p/\omega - v_s$ or $G_p/\omega - \ln\omega$ plots. The introduction of a $G_p/\omega - v_s$ analysis has major advantages, since it requires vastly less data acquisition than the $G_p/\omega - \ln\omega$ methods, because conductance data has to be collected as a function of gate voltage to allow the R_s and C_{ox} corrections to be made. In addition it does not require specialised equipment to measure conductance over a very broad range of frequencies, and consequently most commercially available meters will be sufficient to give at least one very accurate value of $D_{it}(E)$. The new $G_p/\omega - v_s$ and $G_p/\omega - \ln\omega$ analyses both require less data manipulation and are less dependent on accurate determining of peak positions. Peak width measurement is also simplified compared with older methods, and therefore the new analysis ought to provide more accurate values of $D_{it}(E)$ and c_p.

The Brews method is as follows:

(1) $\frac{G_p}{\omega}$ is plotted as a function of v_s or $\ln\omega$.

(2) The maximum in the plot $(G_p/\omega)_{max}$ is measured and its x-axis value (ω_p or v_{s_p}) noted.

(3) $f_w(G_p/\omega)_{max}$ is calculated with f_w equal to one of the following, 0.3, 0.5, 0.6, 0.7, 0.8, 0.9. The difference in the two values on the x-axis at which $f_w(G_p/\omega)_{max}$ occurs is calculated, and then this value is used to determine σ_s from look up tables or charts.

(4) The parameter f_D is determined from a look up table of f_D against σ_s, from which $D_{it}(E)$ is determined using $D_{it} = (G_p/\omega)_{max}/(f_D A q)$.

(5) If capture probabilities are of interest then these found from

$$c_p = \frac{\omega_p}{n_B \epsilon_p} exp(\nu_{sp})$$

where ξ_p is determined from a look up table or plot of ξ_p and σ_s.

This method requires that $D_{it}(E)$ and trap capture cross section be weak functions over a range σ_s and additionally with the $G_p/\omega - \nu_s$ plot that σ_s be a weak function of ψ_s. These assumptions hold in most cases and can be checked quickly by extracting values for parameters at two or more values of f_w, or more rigorously by calculating an entire $G_p/\omega - \nu_s$ or $G_p/\omega - \ln\omega$ plot using extracted data in equation 4.7. This rigorous check was already shown in figure 4.4. Two very rare instances when this assumption fails are shown in figure 4.14.

The samples used in figure 4.13 were analysed by the Brews method so that it could be compared with the capacitance methods in the next section. Figure 4.15 shows the measured $G_p/\omega - \nu_s$ plots for these samples.

4.5. Accuracy of MOS Capacitor Interface Trap Measurements

The average interface trap densities determined by four different measurement methods are compared in tables 4.2 and 4.3. It can be readily seen that although the qualitative agreement between the various methods is quite good the quantitative agreement is generally poor. The following points help determine the relative merits of the different measurement methods.

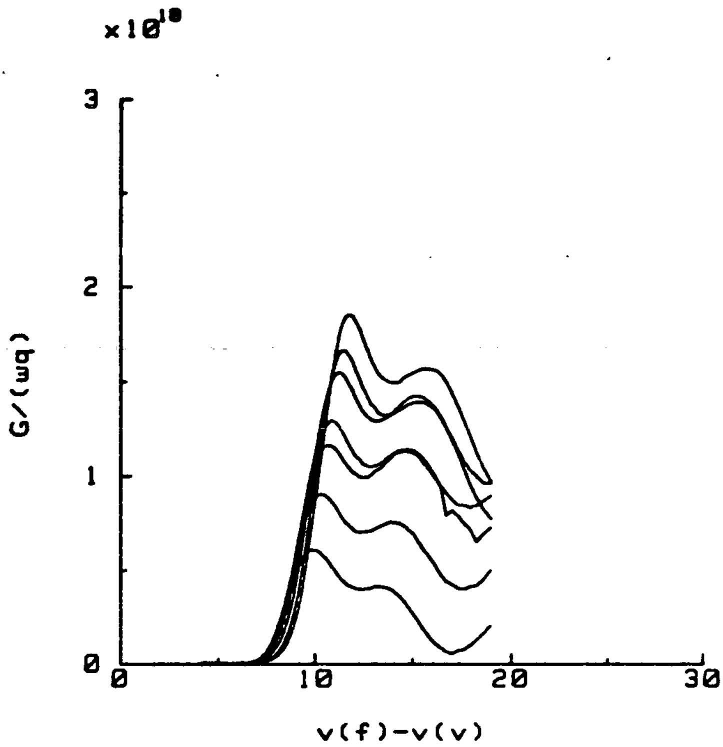


Figure 4.14 (a) Dual peaked GV curve, caused by the creation of an energy level during bias temperature stressing.

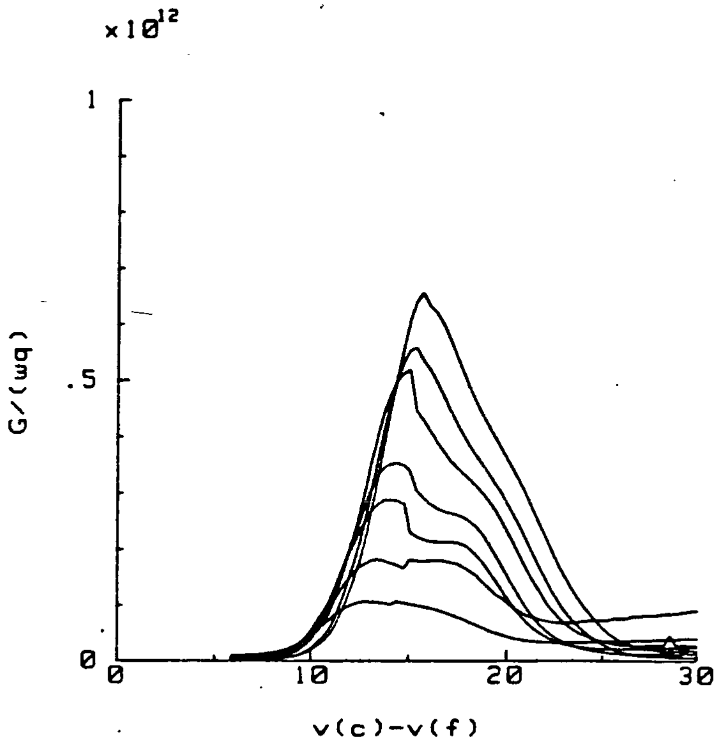


Figure 4.14 (b) Odd GV peaks measured for a very thin (190 Å), gate oxide. The shape of these curves may be due to an uneven oxide thickness creating a gross spread of surface potential at fixed bias.

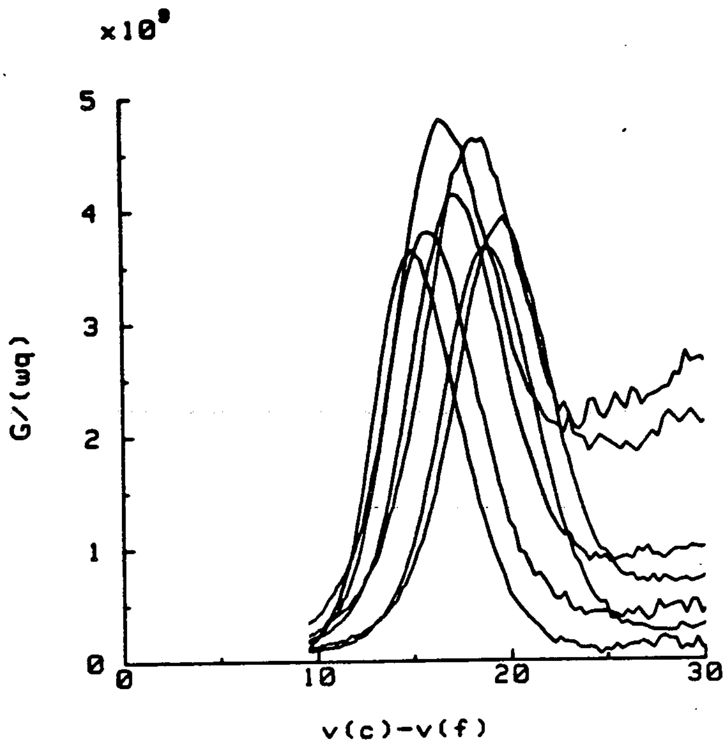


Figure 4.15 (a) Conductance band bending plots for sample TR1 at frequencies of 1MHz, 400 kHz, 200 kHz, 100 kHz, 40 kHz, 20 kHz and 10 kHz. The frequency at which curves were measured decreases from left to right.

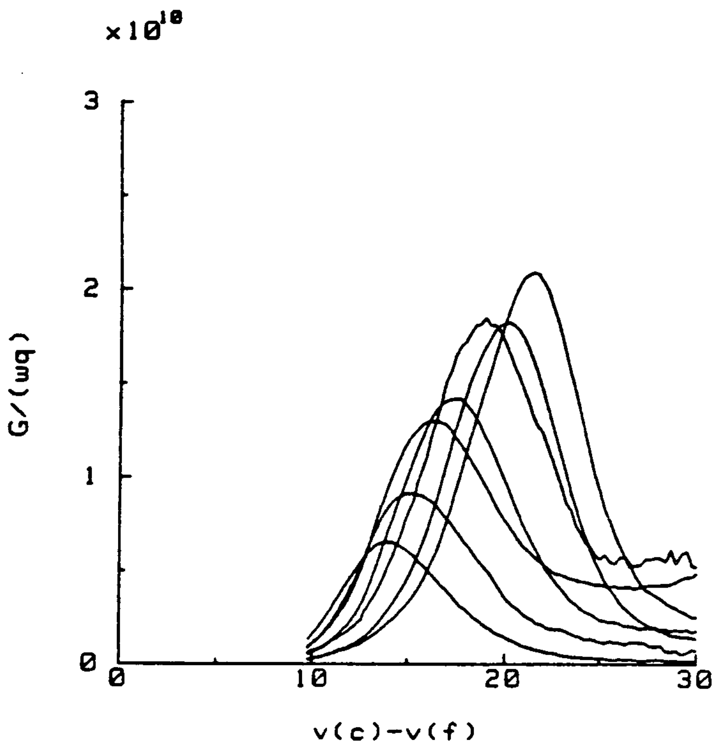


Figure 4.15 (b) Conductance band bending plots for sample TR2 at frequencies of 1MHz, 400 kHz, 200 kHz, 100 kHz, 40 kHz, 20 kHz and 10 kHz. The frequency at which curves were measured decreases from left to right.

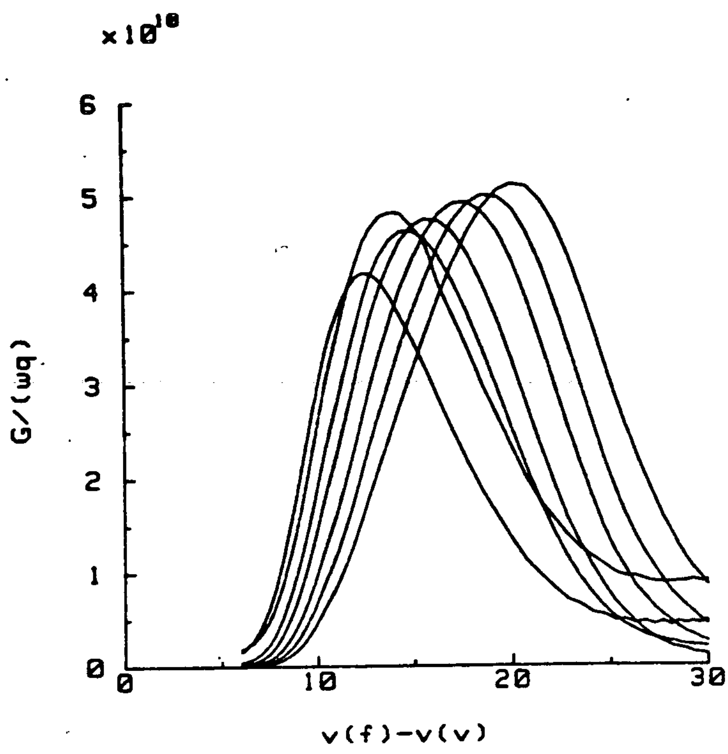


Figure 4.15 (c) Conductance band bending plots for sample TR3 at frequencies of 1MHz, 400 kHz, 200 kHz, 100 kHz, 40 kHz, 20 kHz and 10 kHz. The frequency at which curves were measured decreases from left to right.

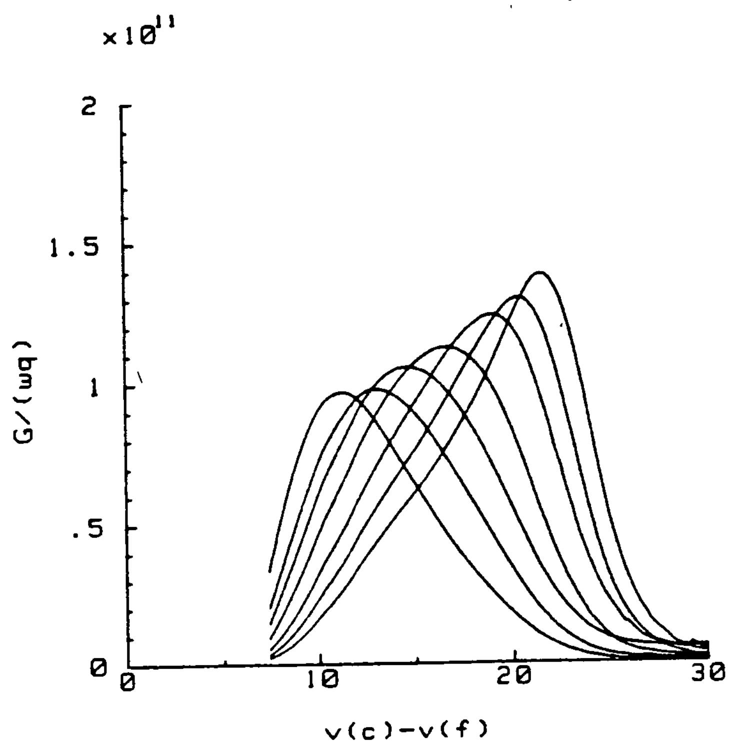


Figure 4.15 (d) Conductance band bending plots for sample TR4 at frequencies of 1MHz, 400 kHz, 200 kHz, 100 kHz, 40 kHz, 20 kHz and 10 kHz. The frequency at which curves were measured decreases from left to right.

Sample	Measurement Method			
	Low Frequency CV	High Frequency CV	High-Low CV	Brew's GV
TR1	$1.5x10^{10}$	$3.1x10^{10}$	ND	$1.3x10^{10}$
TR2	$1.5x10^{10}$	$5.1x10^{10}$	ND	$4.0x10^{10}$
TR3	$1.6x10^{11}$	ND	$1.7x10^{11}$	$3.7x10^{11}$
TR4	$1.4x10^{12}$	$0.7x10^{12}$	$1.6x10^{12}$	$0.7x10^{12}$

Table 4.2 Average trap densities between 0.3 and 0.4 eV from the band edge.

Sample	Measurement Method			
	Low Frequency CV	High Frequency CV	High-Low CV	Brew's GV
TR1	$6.6x10^{9}$	ND	ND	$1.8x10^{10}$
TR2	$1.6x10^{10}$	ND	$8.4x10^{9}$	$4.7x10^{10}$
TR3	$1.4x10^{11}$	ND	$1.7x10^{11}$	$3.7x10^{11}$
TR4	$1.5x10^{12}$	$0.4x10^{12}$	$1.6x10^{12}$	$0.8x10^{12}$

Table 4.3 Average trap densities between 0.4 and 0.5 eV from the band edge.

(1) The Brews conductance method ought to be used as the benchmark by which the capacitance methods are assessed, because it is the only method where the self consistency of measured data can be accurately checked with simulations such as that shown in figure 4.4.

(2) The Brews method requires experimental determination of $\psi_s - V_G$, however it is only sensitive to error in $d\psi_s/dV_G$ and not the absolute values.

(3) The low frequency CV method usually uses the noise sensitive QSCV measurement of C^{LF}, and for high lifetime samples it is not possible to determine C^{LF} in weak and strong inversion. Errors in the low frequency data are also the primary cause of error in the combined method.

(4) Both the low and high frequency CV methods rely very heavily upon the accurate calculation of ideal CV curves.

It can be concluded from this analysis that for quantitative work only GV analyses can provide a reliable determination of trap densities, and that the CV methods should be used only semi-quantitatively. Consequently, in chapter 8, $G_p/\omega - v_s$ plots are used to compare trap densities in a variety of samples. High and low frequency plots were also used to ensure that if any effects occurred outwith the energy range of the GV method they would be detected, (however none were).

4.6. Post Metallisation Annealing of Aluminium-Silicon Gate MOS Capacitors

In this section the rapid and accurate GV method of determining interface trap densities is used to measure very low ($<5\times10^{9}cm^{-2}eV^{-1}$) trap densities. These experiments provide further evidence in support of a trap annealing mechanism[126], namely post metalisation annealing (PMA), which is widely used to reduce midgap interface trap densities in aluminium gate MOS structures. The anneal is thought to involve annihilation of traps by atomic hydrogen species and only occurs with active

metals such as aluminium and magnesium with gold, platinum and other noble metals being ineffective in this role[9, 126]. A silicon nitride layer between an active gate metal and the silicon dioxide inhibits the hydrogen annealing mechanism.[9, 126].

In recent years the MOS fabrication industry has begun to use AlSi for gate electrode and interconnect material to reduce the problem of junction spiking. In this section, differences in trap densities after PMA, for Al and AlSi gates are evaluated[125].

Table 4.1 details the preparation of samples with Al and AlSi gates. Interface trap densities were measured using the Brews conductance method, and the surface potential was determined from the high frequency capacitance measurement.

Figures 4.16(a) and 4.16(b) show the difference in the conductance band bending plots for samples with Al and AlSi gates after a 10 minute PMA in forming gas. The plot for sample Al1 is as expected with peak height decreasing as the frequency is reduced. However the plot for sample AlSi1 shows an increase in peak height with decreasing frequency and the large peaks in 4.16(b) have a narrower width than the corresponding peaks in sample Al1. These changes result in an increased interface trap density towards midgap. Figure 4.17 shows the differences in the interface trap densities for part of the lower half of the bandgap derived from the conductance plots in figure 4.16. The midgap density for unannealed capacitors, which was determined with a gold gate electrode, was 1.2×10^{11} $cm^{-2}eV^{-1}$ therefore a degree of annealing has occurred. Samples grown without the presence of HCl in the oxidising ambient show even poorer annealing. Extending the anneal time from 10 to 40 minutes resulted in a degradation rather than an improvement in trap density as shown in figure 4.18.

These experiments have shown that the presence of a small percentage of silicon in the gate aluminium in an MOS structure clearly interferes with the effective annihilation of midgap interface traps. If the annealing mechanism is that proposed by Deal et. al[9, 126] then the silicon present in the gate is reducing the efficiency with which

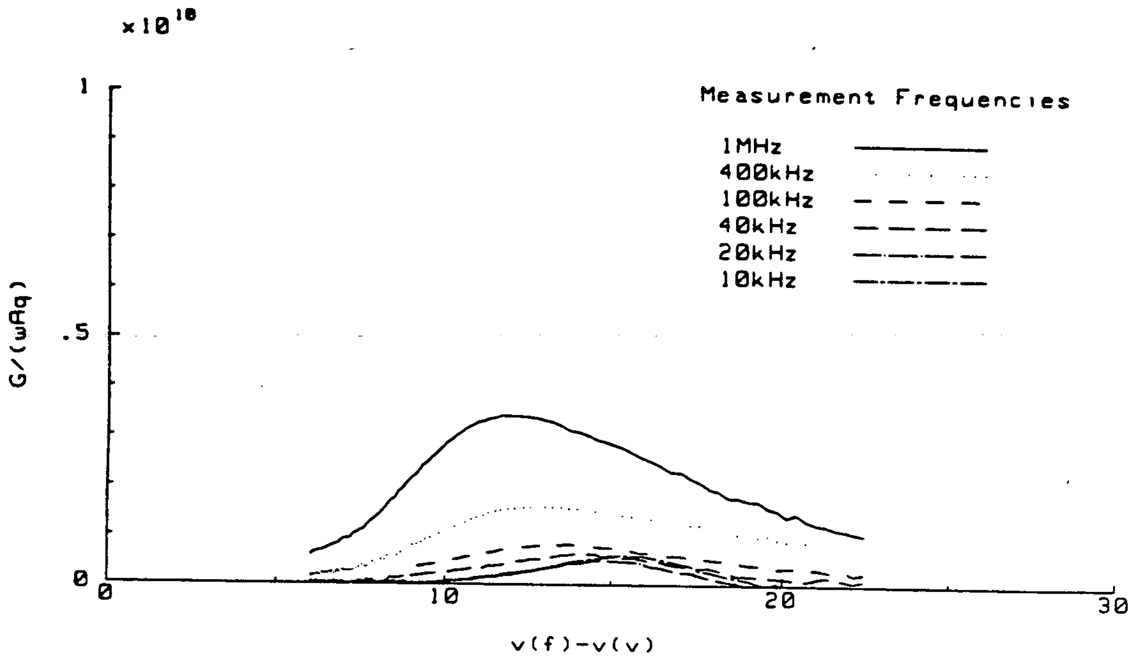


Figure 4.16 (a) Conductance band bending plots for sample Al1.

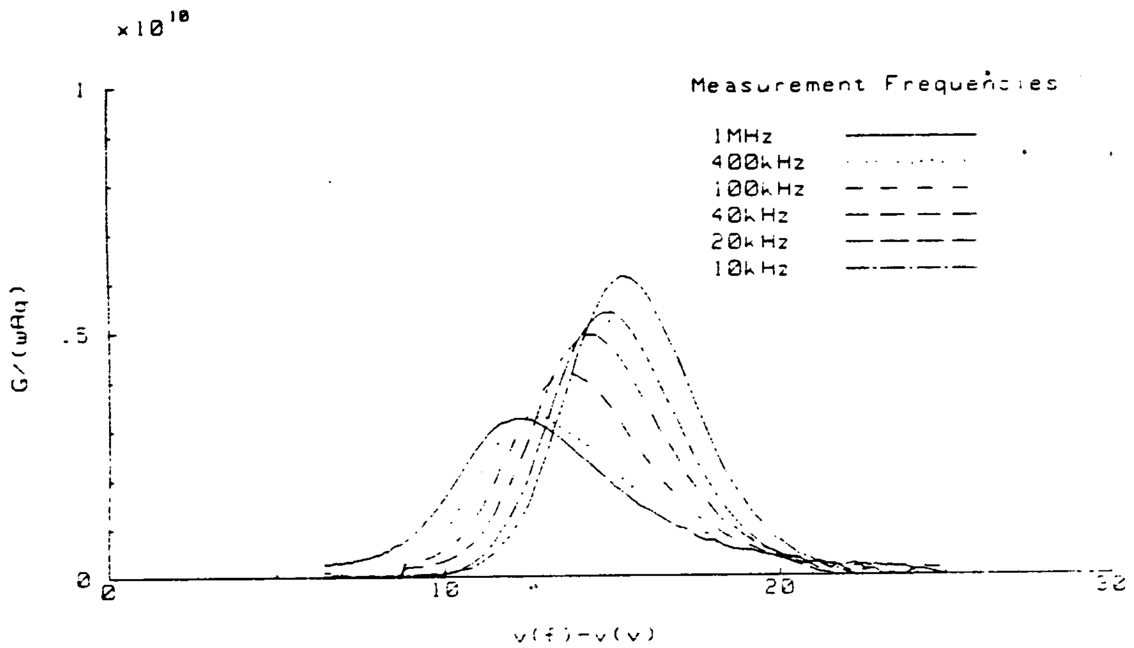


Figure 4.16 (b) Conductance band bending plots for sample AlSi1.

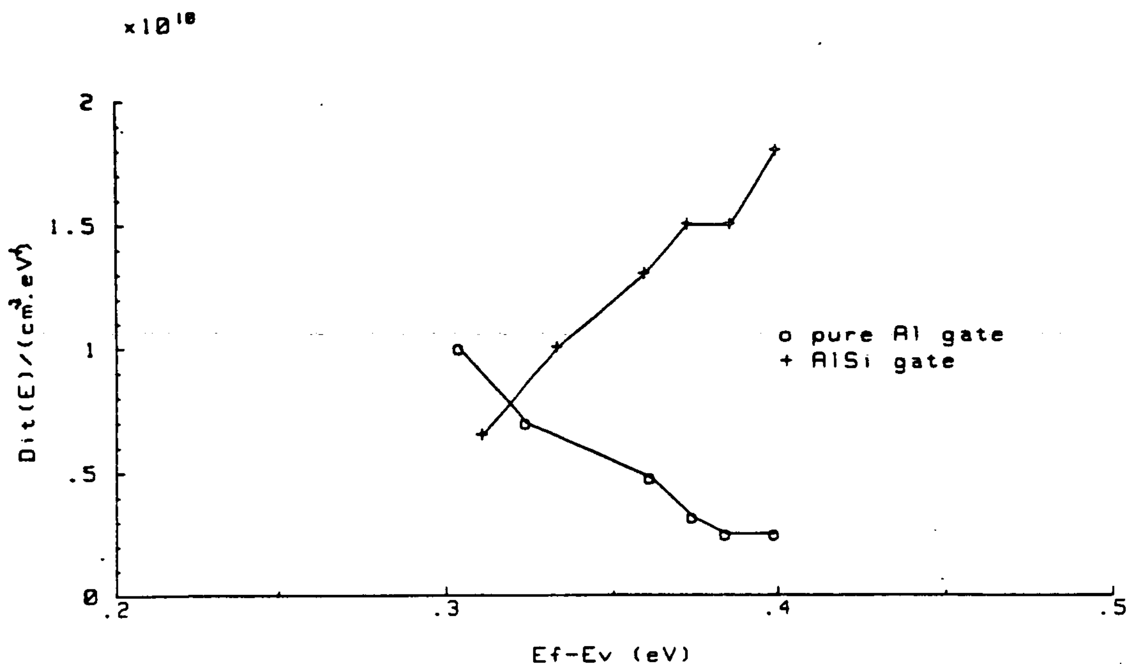


Figure 4.17 Interface trap densities as a function of bandgap energy for the sample Al1 and AlSi1.

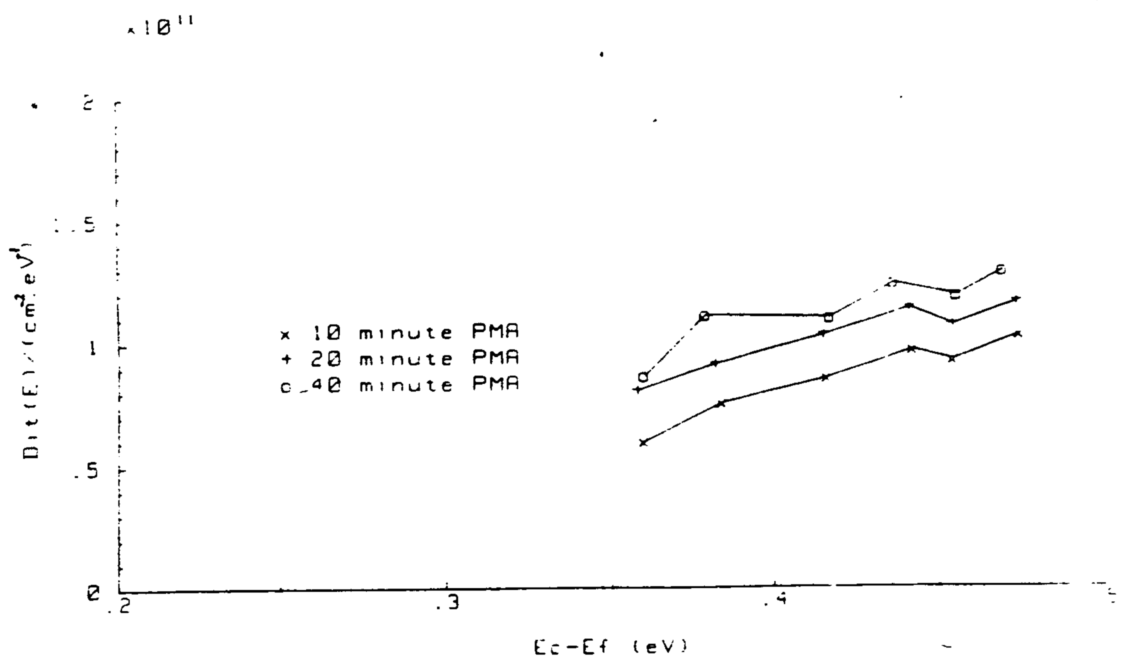


Figure 4.18 Interface trap density after successive post metallisation anneals of sample AlSi2.

atomic hydrogen is produced from its molecular state. This situation is particularly severe when the oxide is grown without HCl. This is probably because in samples with HCl in the ambient, hydrogen is present near the interface[166]. This hydrogen increases the level of atomic hydrogen present during PMA thus improving annealing performance. Interface traps degrade the performance of MOS devices therefore it is important to minimise their level. It is consequently recommended that AlSi should not be used in circuits where there is no problem with junction spiking. If AlSi is to be used, then trap density ought to be minimised by high temperature annealing[9] rather than by using the post metallization technique.

4.7. Conclusions on Interface Trap Meaurements

It has been shown experimentally that three methods of measuring trap density from CV characteristics give results which should be treated as only semi-quantitative. The energy range over which trap densities can be measured by low frequency methods is nowadays limited to depletion energies, because typical values of minority carrier life-time make it impossible to measure a complete characteristic in equilibrium.

A simple modification of a recently proposed GV method, allows full automation of an ac conductance technique for the first time. This method has excellent accuracy over a wide range of trap densities. This automated GV method has been used to provide further experimental evidence in support of the post-metallization annealing mechanism proposed by Deal[126].

CHAPTER FIVE

DOPING PROFILES IN SILICON

5.1. Introduction

Dopant ion distribution is never truly uniform in a semiconductor, particularly at surfaces and interfaces, because of ion implantations, diffusions, and impurity segregation during high temperature oxidations and anneals. Since the doping profile, and dopant type (p or n), control the electrical properties of semiconductor devices it is important to be able to assess these material properties.

Profiling measurements can be divided into two categories:

(a) Methods giving the sum of the ionised and unified dopant ion concentration. This group includes techniques such as secondary ion mass spectroscopy (SIMS) and helium ion, (Rutherford) backscattering spectroscopy (RBS)[167].

(b) Methods giving only the ionised dopant ion carrier concentration. This includes capacitance voltage (CV) and spreading resistance techniques.

The techniques in category (a) are generally only useful for high doping levels ($>10^{16}cm^{-3}$) and use expensive equipment where typical system costs exceed $250,000. These measurements and the spreading resistance method also require special sample preparation, in addition to normal processing. Consequently these methods of profiling are not suitable for cost effective and rapid assessment of large numbers of samples, such as might have to be tested in a process control environment. The rest of this chapter will concentrate on CV profiling techniques. These have the potential of

providing accurate profiles for large numbers of samples, for a system cost of less than $15,000.

The CV characteristics of MOS capacitors[168] and Schottky diodes[169] are widely used to measure doping profiles, (pn junctions[170] can also be used but the analysis is generally more complex and less accurate). The diode structure is favoured when substrate materials such as epi wafers[129] are being assessed or where a thermal oxide has no effect on the doping profile. In most cases of interest oxide does play an important role in determing the doping profile and in these situations the MOS capacitor is the optimum choice of test structure. Section 5.2 examines the accurate implementation of the MOS capacitor CV method and section 5.3 examines the issue of implantation process control.

5.2. Accurate Measurement of the Complete Semiconductor Doping Profile

5.2.1. Data Manipulation

The basic equations proposed by van Gelder and Nicollian[168] which relate MOS CV data to doping profiles are

$$N(w) = 2(q\epsilon_s(d(1/C_m^2)/dV)^{-1} \tag{5.1}$$

and

$$w = \frac{\epsilon_s}{C_D} \tag{5.2}$$

where $N(w)$ is the doping density at the distance w from the oxide-semiconductor interface, and C_D and C_m are the depletion and measured MOS capacitances

Sample Code	Oxide Thickness (nm)	Gate Area (cm^2)	Implant(1) Atoms/cm^3	Implant(2) Atoms/cm^3	Substrate Type
UP*	80	0.0256	0	0	p
UN*	80	0.0256	0	0	n
T18‡	25	0.0024	$6x10^{11}$	$1x10^{12}$	p
T22‡	25	0.0024	$9x10^{11}$	0	p

Table 5.1 Proccessing of Samples for Profiling Experiments

* For further process details see chapter 8 table 2.

‡ For futher process detail see the procces runsheet.

respectively. Pulsed capacitance measurement is used to obtain the 'deep' depletion capacitance, (see chapter 6), at biases beyond V_t, thereby extending the range of the measured profile. Figure 5.1 shows the deep depletion CV characteristics of implanted and unimplanted samples, and figure 5.2 shows the noisy profile which can result from unoptimised data collection.

It is the presence of the derivative in equation 5.1, which is the major source of errors in CV profiling. Kreysig[171] has commented on the inherent 'delicacy' of numerical differentiation where the derivative will always be less accurate than the parent data. In profiling, the difficulty in obtaining accurate derivatives is more acute because the values must be calculated from measured data points and not a known mathematical function. Many years ago Hillibrand and Gold[170] pointed out the essential dilemma that is faced in profiling, viz. "If the impurity distribution varies rapidly, points which are closely spaced in capacitance are needed to minimise averaging of the calculated distribution. On the other hand, with closely spaced data points, differentiation of the experimental data results in very poor accuracy." They were essentially pointing out that derivatives which are too closely spaced are dominated by rounding and random errors. Some workers[148] have suggested that doping profiles measured by CV methods are inherently noisy, whereas their real problem is inaccurate differentiation caused by attempting to obtain profiles from closely spaced data. (Moline[172] claimed that closely spaced data gave accurate values for derivatives. However, his definition of "close" is in fact quite "large"). In the following sections it is shown that accurate profiles can be obtained if the raw capacitance data is collected at an expedient voltage interval.

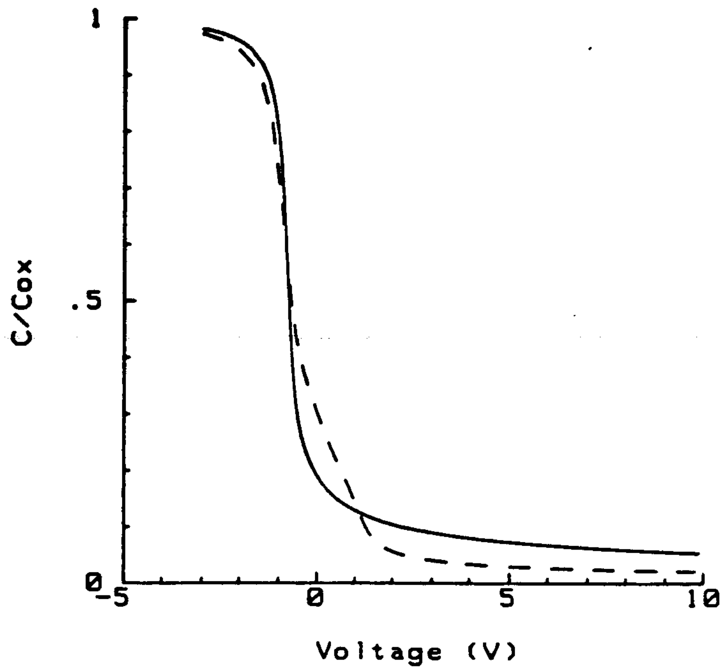


Figure 5.1 Measured deep depletion CV characteristics for a uniformly doped sample (———), and an implanted sample (-----).

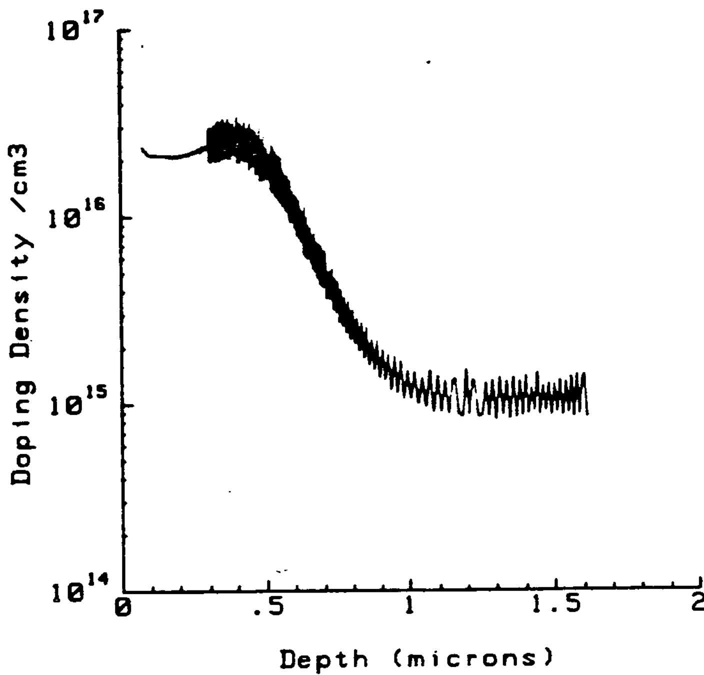


Figure 5.2 A noisy doping profile determined using the pulsed CV method, with poor data acquisition.

5.2.2. Quantitative Error Assessment

To simplify the error assessment it is easier to use the following transformation of equation 5.1,

$$N(w) = \frac{C^3}{q\epsilon_s\left(\frac{dC}{dV}\right)} \tag{5.3}$$

A previous analysis of error in profiles[173, 174], cannot be applied to modern capacitance meters, because it assumes a constant *relative* error, η, which gives a measured capacitance, $C_m = C + \eta C$, where C is the exact capacitance value. In real measurement systems errors arise due to random and rounding errors which give a fixed *absolute* error, ξ, such that $C_m = C + \xi$. With measurement errors equation 5.3 becomes

$$N(w)+dN(w) = \frac{(C+\xi)^3}{q\epsilon_s\frac{d(C+\xi)}{dV}} \tag{5.4}$$

then neglecting terms of order ξ^2 and ξ^3, (5.4) becomes

$$N(w)+dN(w) = \frac{C^3+3C^2\xi}{q\epsilon_s\left(\frac{dC}{dV}+\frac{d\xi}{dV}\right)} \tag{5.5}$$

from which it can be shown that,

$$dN(w) = \frac{3C^2\xi - C^3\frac{d\xi}{dC}}{q\epsilon_s\left(\frac{dC}{dV}+\frac{d\xi}{dV}\right)} \tag{5.6}$$

Now from (5.2) and (5.6) the relative error is obtained, such that

$$dN(w)/N(w) = \frac{3\xi \frac{dC}{d\xi}}{C(1+\frac{dC}{d\xi})} - \frac{1}{(1+\frac{dC}{d\xi})} \qquad (5.7)$$

now ξ takes on a maximum value $\xi_{max} = R$, which is characteristic of the measurement system rounding error and random error, hence the maximum range of the error is 2R. If dC and dV are approximated by ΔC and ΔV respectively, then for the relative error in terms of measured quantities equation 5.7 becomes,

$$\Delta N(w)/N(w) = \frac{\frac{3\Delta C}{2}}{C(1+\frac{\Delta C}{2R})} - \frac{1}{1+\frac{\Delta C}{2R}} \qquad (5.8)$$

When $R << \Delta C$ this simplifies to,

$$\Delta N(w)/N(w) = \frac{3R}{C} - \frac{2R}{\Delta C} \qquad (5.9)$$

which can be further simplified if $\Delta C << C$ to,

$$\Delta N(w)/N(w) = \frac{2R}{\Delta C} \qquad (5.10)$$

i.e. relative error depends only on the relative magnitudes of the capacitance step ΔC , and error R to a first approximation. Using equation 5.10 it can be seen that for 1% error, the condition $\Delta C \geq 200R$ must be satisfied.

An experimental example of the effect that data spacing has on noise in profiles is shown in figure 5.3. In a uniform profile this spacing is not very critical and so large data spacing can be used to obtain very smooth profiles. However, with strongly non-uniform profiles large spacing results in a loss of profile detail. It is therefore necessary to optimise data spacing such that a smooth profile is obtained with maximum detail. A method to achieve this can be devised by rewriting equation (5.1) and substituting finite differences for derivatives, such that

$$\Delta V = \frac{q \epsilon_s N(w) \Delta C}{C^3} \tag{5.11}$$

which with a 1% noise limit becomes,

$$\Delta V = \frac{q \epsilon_s N(w)(200R)}{C^3} \tag{5.12}$$

This algorithm provides a means to fully automate accurate measurement of the doping profile with the CV method. Examples of the voltage patterns resulting from equation 5.12 are shown in figure 5.4 for uniformly doped and implanted samples. Voltage sweep optimisation is implemented in the profile data acquisition and analysis program "PROF". In practice equation 5.12 is used to estimate the next step from the density calculated from the previous two data points. The maximum step is arbitrarily limited to 0.25V in order to prevent skipping through heavy doped regions after lightly doped regions. R must be measured for each range of a capacitance meter. For the HP4275A R is approximately 1-2 times the least significant figure for capacitances larger than 1pF. Figure 5.5 shows three profiles measured with this program. These should be compared with the noise dominated profile in figure 5.2 and that of [148].

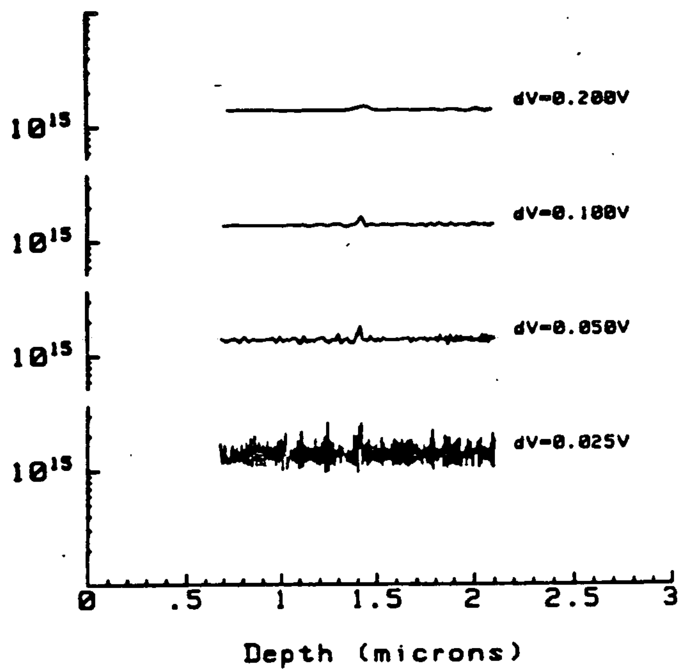


Figure 5.3 A comparison of the noise in a uniform profile measured with different data spacings.

5.2.3. Other Factors Affecting Profiles

For rapidly varying profiles equation (1) is not adequate because the profile gradient creates a space charge within the undepleted portion of the material. Kennedy and O'Brien[175] have shown that the raw profile can be corrected for this effect using,

$$N(w) = N(w) - \frac{kT}{q}\frac{\epsilon_s}{q}\frac{d}{dw}\left(\frac{1}{N(w)}\frac{dN(w)}{dw}\right) \tag{5.13}$$

provided that the depletion approximation[9] is an adequate representation of the depletion layer edge[176, 177, 178, 179]. Since equation 5.13 involves *third* order numerical differentiation, implementation of this correction requires care for noise free profiling. The optimum resolution from equation 5.12 results in too much residual noise for the profile differentiation, and it was found empirically that using twice the estimated optimum spacing combined with a threshold change of 5% for sharp corrections resulted in noise free profiles with reasonable resolution. Figure 5.6 compares two "raw" and corrected profiles.

Another failure of equation 5.1 occurs close to the semiconductor-oxide interface where the depletion layer charge is not solely due to ionised dopant atoms. Ziegler et al.[180] showed that this effect is of importance at distances less than two debye lengths from the interface and they developed a procedure for obtaining the profile right up to the interface. This method is implemented in "PROF" even although it may only be truly valid for uniform profiles[178], in order to allow for an estimate of implantation doses. This approach is the best that can be achieved since the near surface profile is also affected by interface traps which can never be completely eliminated. Figure 5.7 shows an extreme example of the distortion of a measured profile which can occur due to the presence of a high density of interface traps. Brews

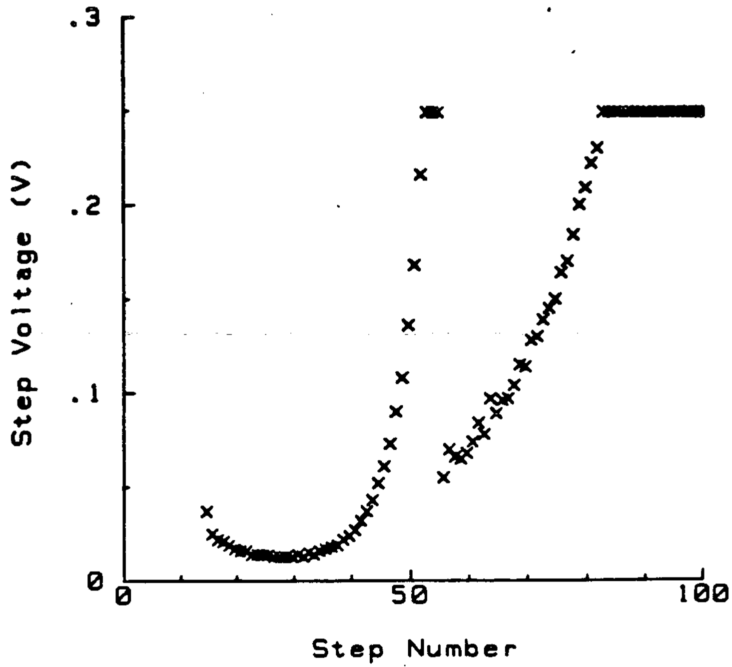


Figure 5.4 (a) Optimised voltage sweep for uniformly doped sample UN. The change in step near the middle of the sweep occurs because of a change in the capacitance meter range.

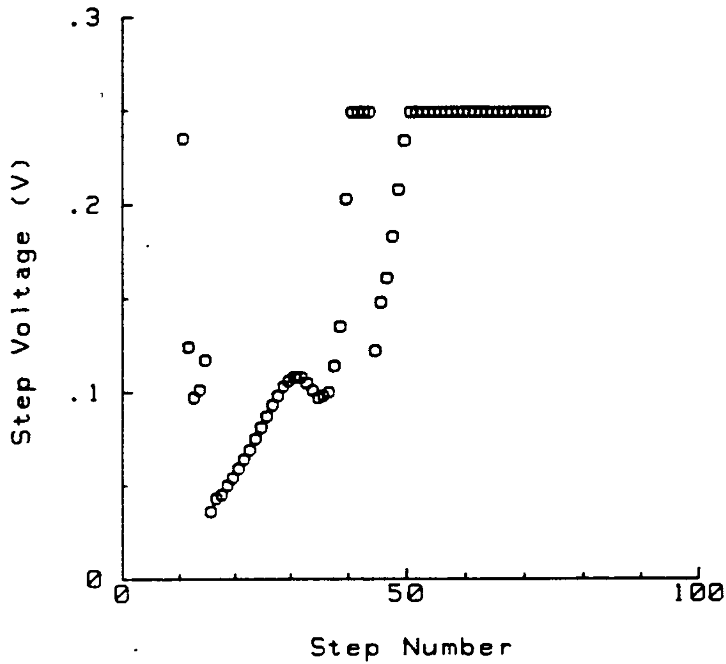


Figure 5.4 (b) Optimised voltage sweep for an implanted sample T22. The change in step near the middle of the sweep occurs because of a change in the capacitance meter range.

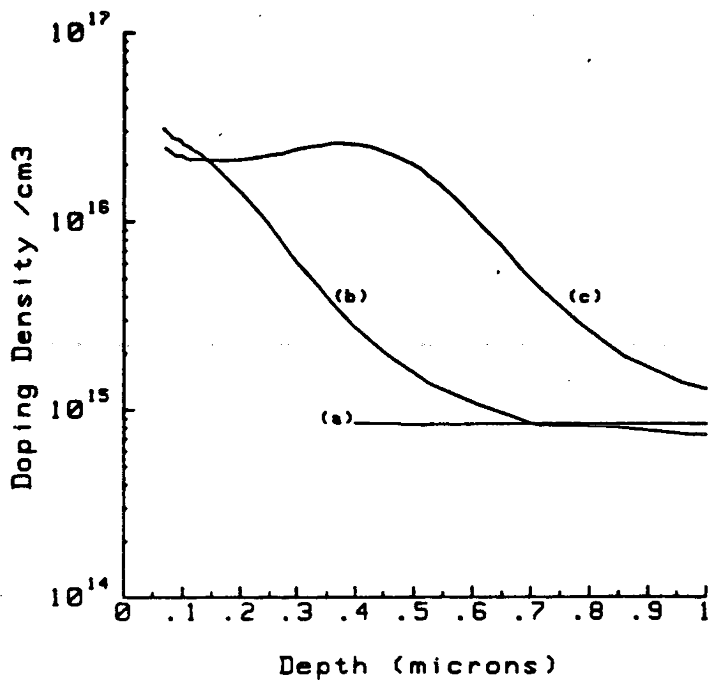


Figure 5.5 Noise free profiles determined by "PROF". Samples are (a) UN, (b) T22, and (c) T18.

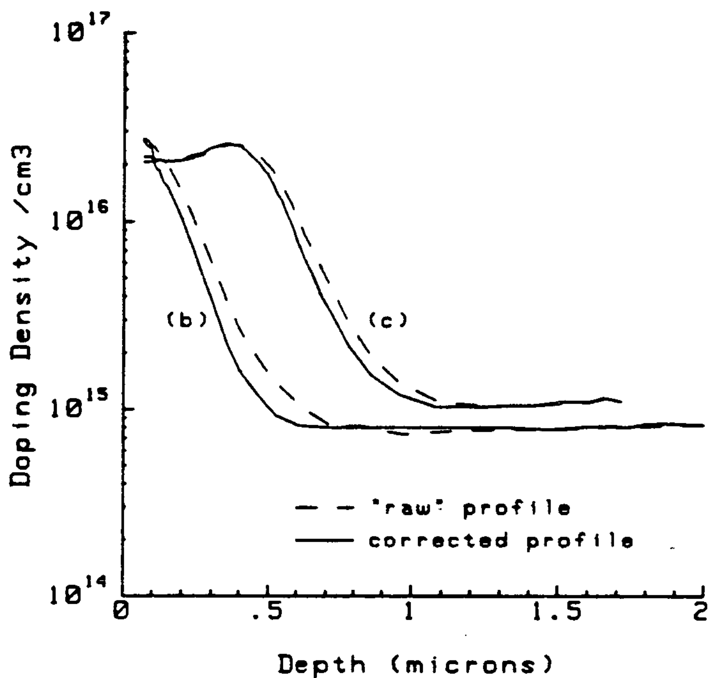


Figure 5.6 Profiles of the implanted samples in figure 5.6 corrected for sharply changing doping density.

proposed a method for correcting profiles for interface trap distortions[181]. However, the accuracy of the correction is limited and requires additional measurements, therefore it is better to minimise trap densities during sample preparation.

Figure 5.8 shows near uniform boron and phosphorus profiles as measured by "PROF". These profiles prove that the assumption of a uniform profile in many of the calculations in other chapters is valid.

Errors in measured C values can occur at voltages higher than V_t because of inversion layer generation during the measurement period[182]. The extent of generation is inversely proportional to minority carrier lifetime, (see chapter 6), therefore it is desirable to use high lifetime samples for profiling measurements. (Alternatively a fast capacitance meter could be used. However this will give higher rounding and random errors!) Figure 5.9 shows the apparent profile for a uniform sample where the carrier lifetime is low. All other samples described in this thesis had lifetimes of 100μS or greater. These showed negligible profile error due to inversion layer generation.

In recent years major efforts have been made to produce IC process and device simulators, with the goal being to allow the development of new IC fabrication recipes and devices without the slow and costly procedure of producing test and evaluation samples[147]. Two complete profiles as determined by "PROF" are compared with results from two simulators in figure 5.10.

As a qualitative check on the accuracy of the differing profiles, values of N_{sub} can be measured using the procedure outlined in chapter 3. This value is the most accurate estimate of the near surface doping which can be determined. Since N_{sub} is the first moment of a nonuniform distribution this quantity can also be determined directly from profiles if the maximum depletion layer width is known. For the measured profiles, the maximum depletion layer width is determined from the equilibrium minimum capacitance, whilst for the SUPREM simulations only, the simulator EPCAP[147] was

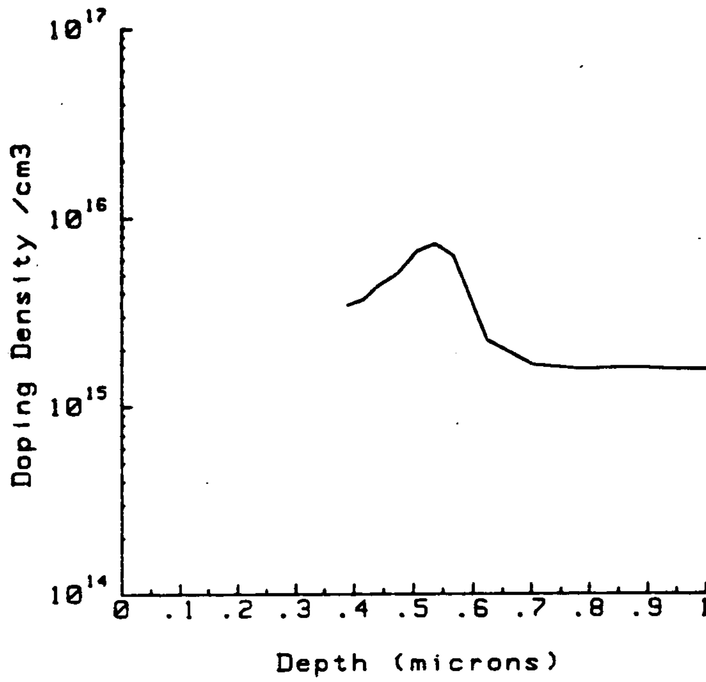


Figure 5.7 Interface trap distortion of a uniformly doped profile.

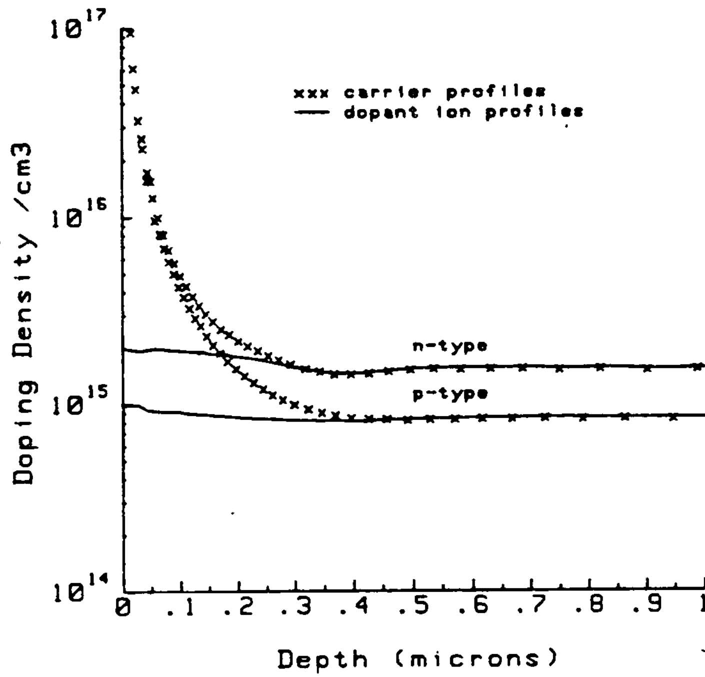


Figure 5.8 Dopant ion and majority carrier charge profiles of uniformly doped p-type and n-type wafers.

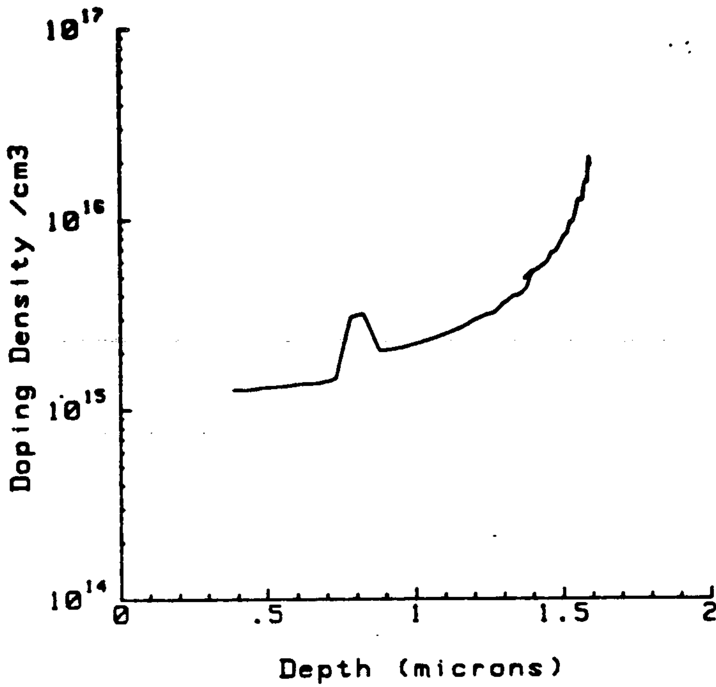


Figure 5.9 Influence of low minority carrier lifetime on a uniformly doped profile. The effect is exaggerated with a 0.5 second delay in the measurement.

Sample	N_{sub}		
	Theory	C_{min}	Dose
T18	2.86	2.24	2.14
T22	2.68	1.96	1.58

Table 5.2 Near surface doping densities.

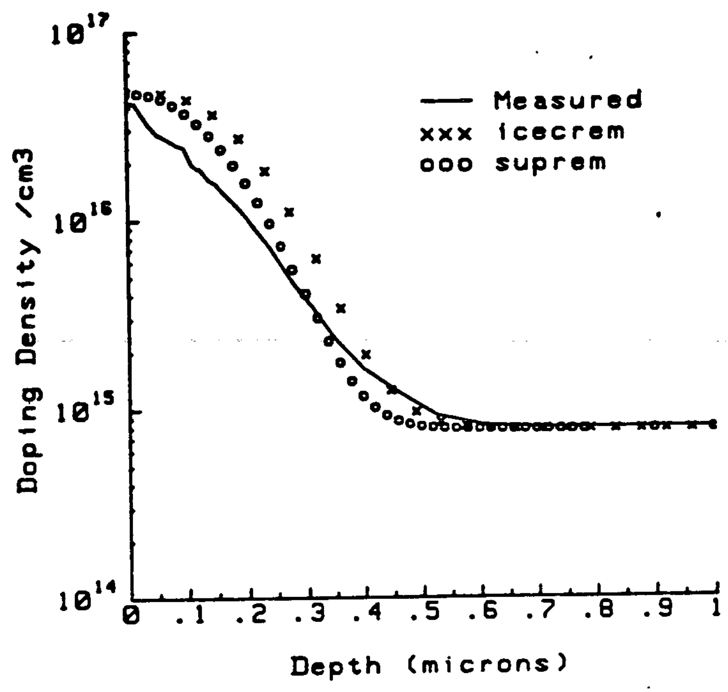


Figure 5.10 (a) A comparison of the profile determined by "PROF" for sample T18, with profiles calculated with process simulators.

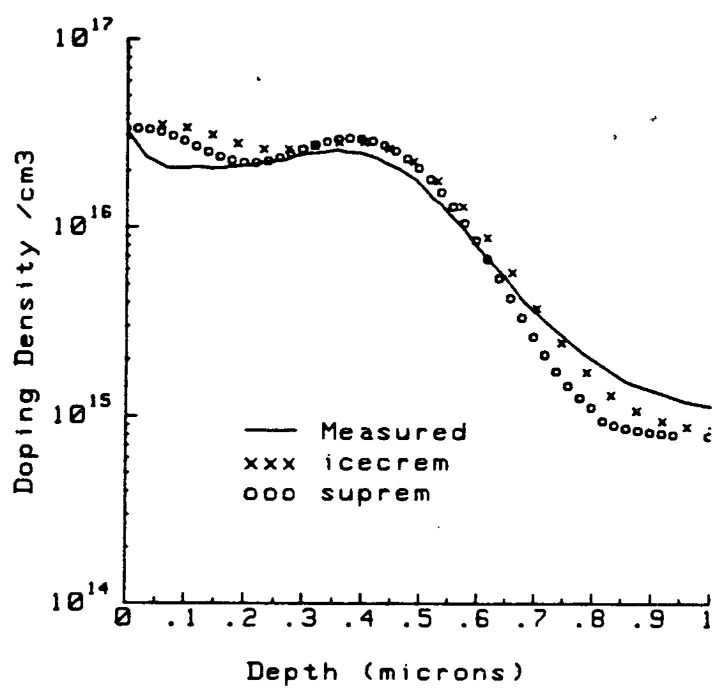


Figure 5.10 (b) A comparison of the profile determined by "PROF" for sample T22, with profiles calculated with process simulators.

used to evaluate this width. Table 5.2 compares the values of N_{sub} and clearly indicates that the surface doping is indeed less than that predicted by SUPREM in both cases.

From these calculations it can be concluded that the measured profile is at least as accurate as the simulated one. Further comparison of the merits of CV profile measurement as compared to profile simulation requires the development of more sophisticated simulators, containing features such as a proper treatment of the behaviour of chlorine-containing oxides.

5.3. Ion Implantation Dose Measurement

The control of ion implant dose and its uniformity across a wafer are critical in obtaining a high yield VLSI IC process[148, 183, 184, 185, 186, 187, 188]. This section assesses the performance of Edinburgh Microfabrication Facility's *Lintott III Ion Implanter* using a CV method, where the dose D is determined using,

$$D = \int_0^w N(w)dw - wN_b \qquad (5.14)$$

where N_b is the bulk doping density.

Deming and Keenan[148] recently proposed the "CdV" method as a MOS capacitor-based implanter uniformity assessment procedure which had better accuracy and utility than resistance measurements or the conventional MOS capacitor dose measurement. It will be shown here that the use of "PROF" to determine profiles overcomes the criticism expressed by Deming and Keenan of equation 5.14.

Figure 5.11 shows a dose uniformity plot for wafer T22 determined using "PROF" and equation 5.14. Uniformity ($\sigma/\bar{x}$), is better than 0.09%, thereby proving both the satisfactory performance of both the measurement and the implanter. This accuracy is

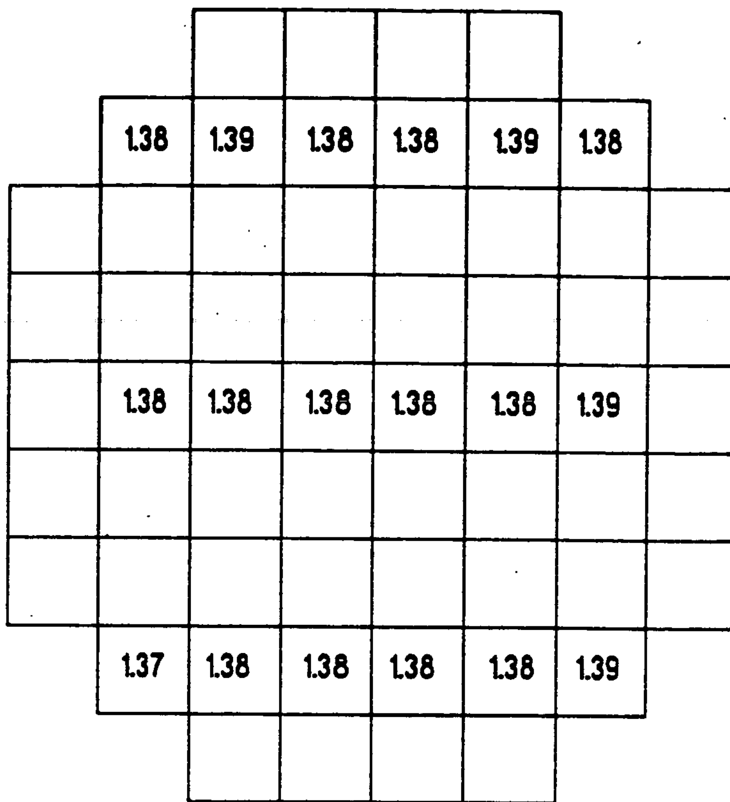


Figure 5.11 Ion implant dose uniformity map for wafer T18.

better than that reported by[148] for an implanter used by IBM, and that determined by Martin[189] for the Edinburgh machine using a spreading resistance measurement. The uniformity of the Lintott III may be better than that of most other systems which use a scanned ion beam[33] because it uses a fixed beam and mechanically scans the wafers[190]

Deming and Keenan used shadow masking to obtain a fast turn around and determined capacitor area from oxide thickness measurements via equation 3.6. The need for such quick turnaround is questionable, (lithography would not take more than 3 hours), since use of (3.6) introduces a spurious correlation between oxide thickness variations and dose variations. In addition many of the implanter problems[188] which occur, such as missed implants, accelerator energy drift, incorrect doses or energies, should be tackled with proper process management and improved implanter design rather than real-time retrospective fault diagnosis. Since the remaining problems occur only infrequently, and have only marginal effects on overall yield[186] the proper role of dose assessment should be as a calibration tool at the time of installation, and after maintainance of an implanter.

5.4. Conclusions

The development of a voltage sweep step algorithm has allowed accurate noise free profiling to be implemented in a fully automated CV measuring system. This system has been used to measure the implantation uniformity of the Lintott series three implanter. Results show better uniformity than was measured for this machine by Martin[189] using spreading resistance measurements. This provides further proof of the proper function of the implanter, and also shows that the CV method of measuring dose uniformity can be very accurate.

CHAPTER SIX

MINORITY CARRIER LIFETIMES IN SILICON

6.1. Introduction

Minority carrier lifetime is an important silicon material property which affects the performance of a wide range of device types. Low lifetime silicon is sometimes undesirable[191, 5, 4]. For example, it degrades solar cell efficiency, increases pn junction and CMOS leakage currents, increases dark current, and reduces transfer efficiency in CCDs, reduces pixel uniformity and brightness in CCD imagers, and reduces refresh time in dynamic random access memorys (DRAMs). In contrast for fast recovery diodes or gate controlled switches (GCS), low lifetime produces a beneficial increase in switching speed[192]. It is therefore important to have the capability of measuring minority carrier lifetimes in order to understand or monitor the effects of device processing.

Ross[191] has outlined a wide range of techniques for measuring lifetime in silicon materials, and stated that there is no sign of lifetime measurement settling to a universal method, and that this was in fact an advantage since different measurements more accurately determine lifetime parameters as they relate to specific devices. For example according to Ross[191], the reverse recovery, or the open circuit voltage decay techniques are suited to measurements of silicon controlled rectifiers or fast recovery diodes, whereas MOS capacitor measurements give the best determination for the wide range of MOS device configurations. The remainder of this chapter will focus on MOS capacitor lifetime measurements. In section 6.2 the leakage current which restores the non-equilibrium MOS capacitor to equilibrium is discussed, and then in

section 6.3.1 the reason for choosing the Zerbst technique in preference to a range of other techniques is dicussed. In section 6.3.2, methods to achieve a time effective and reliable automation technique are proposed, and finally in section 6.4 the Zerbst method is extended to cover the technologically important, (but often ignored), case of lifetime in non-uniformly doped samples.

6.2. Leakage Currents in the Deep depleted MOS Capacitor

MOS lifetime measurements use the deep depleted MOS capacitor to assess minority carrier generation lifetime, τ_g, and minority carrier recombination lifetime, τ_r. The latter has recently received attention[193, 194], because it characterises the deep silicon bulk, (i.e. depths from 10's to 100's of micrometers), whose properties can be modified with intrinsic or extrinsic gettering techniques. However, in this thesis discussion of this parameter is restricted to its effect on τ_g, which can be a direct measure of contamination effects.

Zerbst[14] proposed that two mechanisms were responsible for restoring the deep depleted MOS capacitor to equilibrium. The first is generation via bulk traps at the depletion layer edge which contributes a leakage current I_1, such that,

$$I_1 = \frac{qn_i \Delta WA}{\tau_g} \tag{6.1}$$

where ΔW is the excess depletion layer width beyond the equilibrium inversion region width, and q, n_i and A have their usual meanings. The second current, I_2, arises due to generation via interface traps at the surface where,

$$I_2 = qn_i As \tag{6.2}$$

where s is the effective surface generation velocity which is related to the interface trap density. However it is not worthwhile attempting a direct quantitative correlation between $D_{it}(E)$ and s, because as it will be shown later, s cannot be measured independently.

Schroder and Nathanason[13] detected an unexpected inverse correlation between lifetime and surface generation which they ascribed to a significant peripheral generation component. It meant that an effective generation lifetime τ_{geff} is what is actually measured, where for circular or square electrodes,

$$\tau_g = \tau_{geff}(1-d_o/d)^{-1} \tag{6.3}$$

and

$$d_o = 4s_o t_{geff} \tag{6.4}$$

where d is the diameter or side length of the electrode and s_o is the surface generation velocity of the depleted surface. s_o is in general very much larger than s because inversion layer charge quickly screens the interface under the gate, whilst the area around the periphery of the gate extending to a distance of approximately ΔW remains depleted and consequently has a much higher level of surface generation. This additional source of inversion layer charge and those proposed by Zerbst are illustrated in figure 6.1. Accurate measurement of s_o is very difficult with the MOS capacitor structure, requiring the use of long thin capacitors and a fast capacitance meter[13], therefore it is better to use large area capacitors to ensure $\tau_g \approx \tau_{geff}$, particularly when lifetime is high.

One major source of minority carriers which is often ignored by many reviewers is the inversion layer beyond the gate (ILBG), which was discussed in chapter 3. (The lifetime measurement can be affected by ILBG which is too weak to significantly effect high frequency CV measurements). In the experience of this author, ILBG makes it very difficult to reliably measure lifetime parameters of lightly doped p-type material, unless more sophisticated capacitor structures such as those shown in figure 3.6 are used. Another problem with lifetime measurements is that they are extremely sensitive to light[195], and so use of a double light shield is advisable to ensure complete blackout measuring conditions.

6.3. Techniques for Measuring Generation Lifetime Parameters of MOS Structures

Kang and Schroder[5] have recently reviewed a large number of methods for measuring generation lifetimes, which utilised either C-t transients at fixed bias or non-equilibrium gate bias ramps of MOS capacitors. As in chapter four of this thesis, (when interface trap measuring techniques were compared), they choose an experimental comparison of methods. However, in using only one uniformly doped sample, their conclusions are far from having universal validity. Their results discredit a number of methods which give inaccurate or boundary condition dependent parameters [196, 197, 198, 199, 200, 201, 202, 203, 204, 205], whilst the top of their list of meritorious methods is as follows,

(1) Calzolari[206]

(2) Zerbst[14] †, Trullemans[207] and Kuper[208].

† This reference is in German. References[9, 5, 13] cover Zerbst's work in English.

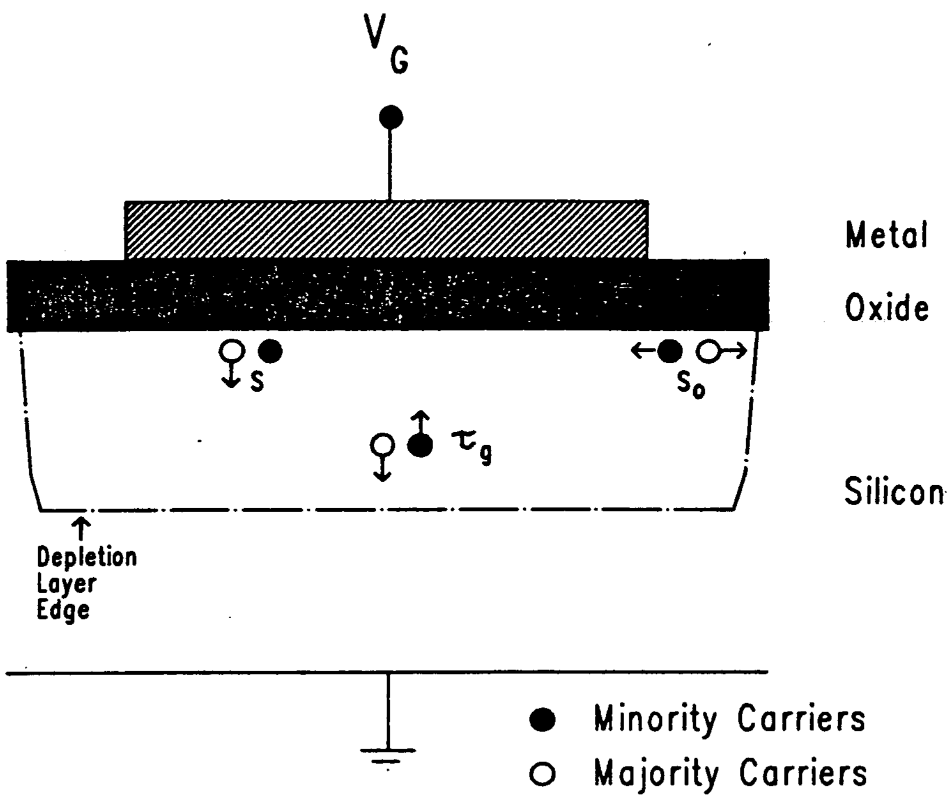


Figure 6.1 Leakage currents which restore equilibrium in the deep depleted MOS capacitor.

(3) Pierret[209, 210, 211] and Taniguchi[212].

This list was drawn on the subjective feel the reviewers had for the ease of instrumentation, measurement time, doping dependence, quantity of data manipulation and parameter accuracy. This author considers that whilst it is valid to discard certain techniques as inaccurate, the ranking of useful techniques is too subjective to be of any practical use. In addition some of the reviewers assertions are questionable. For example they claim that the list is not significantly dependent upon the availability of automation, which this author considers to be of paramount importance, and also a correct merit list will depend very strongly on the type of sample which is to be evaluated and the resources of particular laboratories.

In chapter 7 and 8 of this thesis the Zerbst method as modified by[13] is used to determine τ_g because;

(1) The method is widely accepted as the lifetime measurement benchmark.

(2) τ_g can be accurately determined in the presence of other leakage currents.

(3) Non-ideal behaviour is readily detected in the data analysis, (see later in this section).

(4) The method requires only the measurement of capacitance. Current measurements in high lifetime samples are difficult, e.g. if a MOS capacitor, of area 0.0025 cm^{-2}, is depleted 0.1 microns beyond its equilibrium depletion layer width and lifetime is 1ms, then from equation 6.1 the generation current will be approximately 0.5 picoamps. Such a low current can be very difficult to measure. However, it is possible to measure the difference in capacitance caused by this current with much greater accuracy provided a suitable time interval is chosen.

(5) The method can be implemented on a computer controlled measuring system.

(6) Accurate parameter determination is required.

(7) Long measurement times are not important in the research and development laboratory.

Criticism of the Zerbst method in the literature centres on the fact that the substrate doping profile must be independently determined and that data manipulation is not straightforward. The latter difficulty is easily overcome with a computer controlled measuring system as is the first if a doping profile measurement program is available. Although several methods provide the possibility of determining lifetimes in non-uniformly doped substrates without a determination of the doping profile [200, 202, 206, 210], this is not a real advantage because lifetimes are often correlated to doping density, (see section 6.2.2 and [213]), therefore the interpretation of measured lifetimes is not meaningful without an accompanying doping profile determination.

6.3.1. Accurate Implementation of the Zerbst Method

The Zerbst relationship which applies to the deep depleted MOS capacitor at fixed bias is [9],

$$\frac{-d}{dt}\left(\frac{C_{ox}}{C(t)}\right)^2 = 2\frac{n_i}{N_{sub}}\left(\frac{C_{ox}}{C_f}\frac{(C_f/C(t)-1)}{\tau_{geff}} + \frac{\epsilon_s}{\epsilon_{ox}}\frac{s'}{D_{ox}}\right) \tag{6.5}$$

where C(t) is the measured MOS capacitance at time t. The so called Zerbst plot is a plot of $-d/dt(C_{ox}/C(t))^2 = Z(y)$, against $(C_f/C(t)-1) = Z(x)$. An example of a Zerbst plot is shown in figure 6.2. The slope of the central linear potion of the plot is proportional to τ_g, such that,

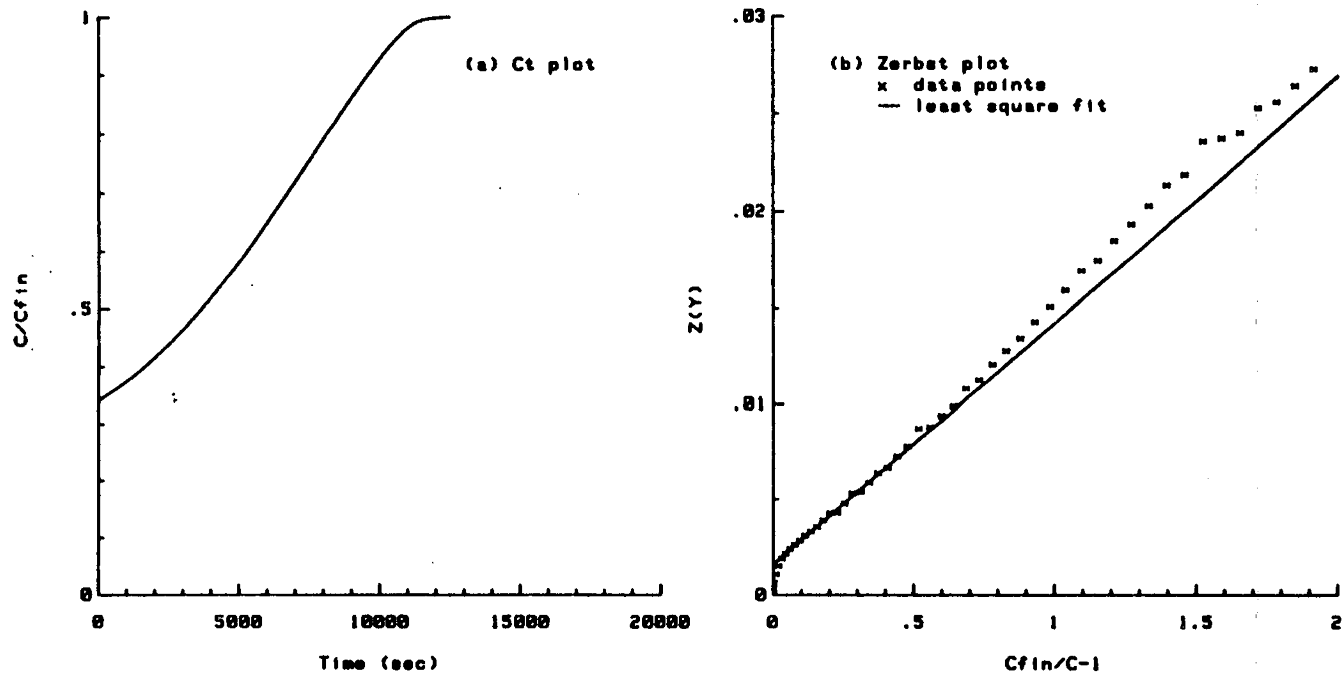


Figure 6.2 Capacitance-time transient and Zerbst plot for a near ideal sample.

$$\tau_{geff} = 2\frac{n_i}{N_{sub}}\frac{C_{ox}}{C_f}\frac{\Delta Z(x)}{\Delta Z(y)} \tag{6.6}$$

The least squares fit line in figure 6.2 delineates this part of the curve. The intercept of an extrapolation of this region with the y-axis yields s', the effective surface generation velocity. It should be noted that the intercept is s', because it relates to the total depth-independent current contribution[5, 193], i.e. it contains contributions from s and s_o, as well as any other depth independent leakage mechanism.

The shape of figure 6.2 is an important feature of the Zerbst method as it gives an indication of how well any particular sample fits the model of a single width-dependent leakage.

Figure 6.3 shows the C-t and Zerbst plots for two adjacent sites on one sample. It is clear from *comparison* of the C-t plots that a defect mechanism is active in the sample with the short transient, but this could have been gleaned by an experienced eye from the Zerbst plot alone. The very steep initial slope in comparison to the near surface slope is indicative of an additional depth-dependent degradation. This type of behavior is most often associated with interface traps. However, this is not the problem here, since high resolution conductance measurements were used to show that interface trap density for both sites was 5 x 10^{10} $cm^{-2}eV^{-1}$, and that the trap density was uniform across the surface. In this case the only plausible explanation is that there is a very active crystal defect in the bulk at depths beyond 1μm, since metal ions associated with lifetime degradation would be expected to diffuse throughout the near surface region.

Since equation 6.5 contains a derivative, care must be taken in acquiring data at an expedient *time* interval to produce a noise free Zerbst plot. The methodology required to achieve this is somewhat different from that which was described for profiling in chapter 5. There the derivative could vary over several orders of magnitude

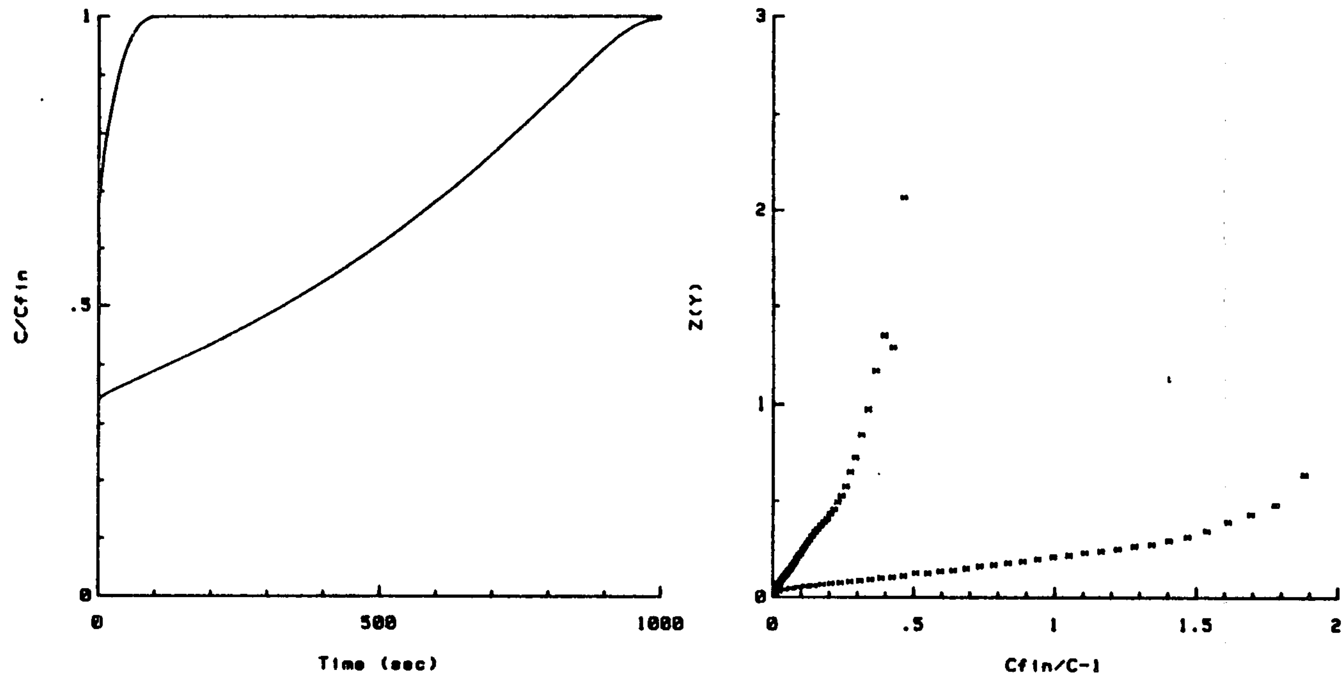


Figure 6.3 Capacitance-time and Zerbst plots for two samples on tne same wafer.

within one measurement and the limits of the independent variable, (voltage) were known. In contrast the derivative is by comparison relatively constant in the C-t measurement, and the upper limit of the independent variable, (time), is not known in advance of the measurement. Software which is currently available for Zerbst analysis[6, 214], requires that the measurement technician estimate the total transient time, t_f, before the C-t measurement commences. If the estimated time is too short then the measurement must be repeated to obtain a complete capacitance transient, or if the estimate is too long then the data may be collected at too coarse an interval to produce a Zerbst plot. Since t_f can vary over a range of at least 0.1 to 10,000 seconds, the amount of time which can be wasted can be very large. It is therefore useful to try to estimate t_f in some way.

One method of achieving this would be to use a simple empirical formula such as,

$$t_f = \gamma(t_1 - t_o)\frac{(C_{fin} - C_o)}{(C_1 - C_o)} \qquad (6.7)$$

where C_o and C_1 are the capacitances measured at times t_o and t_1, and where γ is a constant. The times are chosen to be short compared with the total transient time. In practice a time of 15-60s is a good starting value for $(t_1 - t_o)$. If this is too long, i.e. $(C_1 - C_o) \approx (C_{fin} - C_o)$, then shorter time intervals can be used. This type of approach has been tested using data from 16 samples. Figure 6.4 shows the result for $\gamma = 2$. It is apparent the correlation is quite good. This method requires that certain sample parameters are fixed, viz. substrate doping, oxide thickness and depletion voltage. This situation may well occur in practice, for example in a process control environment, in which case 6.6 will provide a useful algorithm and could be extended to allow γ to be dependent on the magnitude of t_f for greater accuracy.

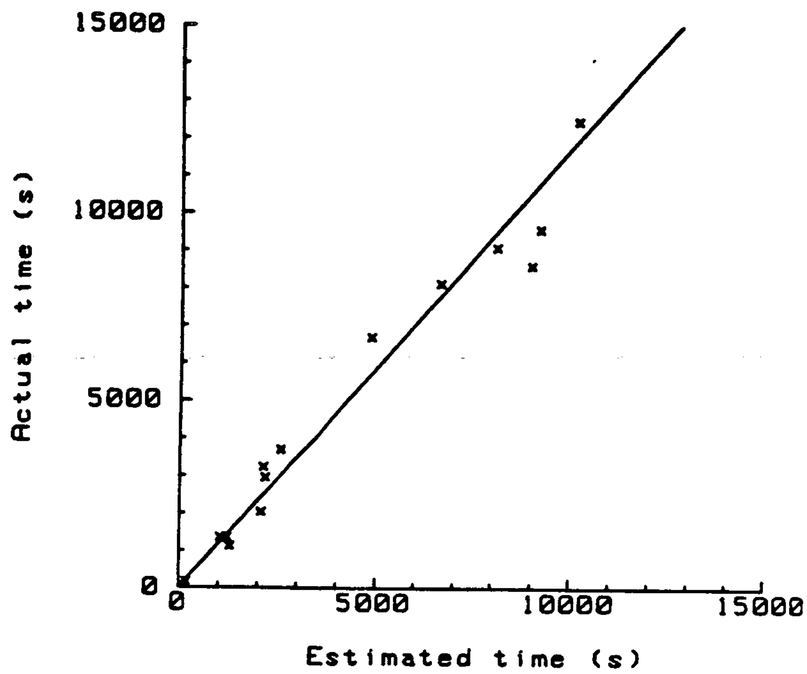


Figure 6.4 Correlation of actual finishing time with the finishing time estimated using equation 6.7.

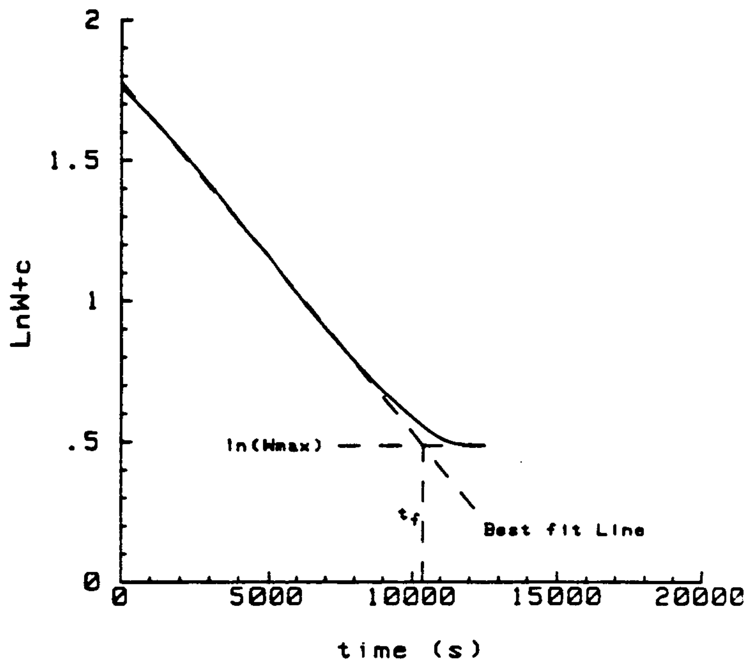


Figure 6.5 Fit of the empirical Rabbani model, (best fit line), to measured data.

The empirical approach may be advantageous in a well defined measuring situation. However, a more general solution to the timing problem would an advantage in C-t measurement software. The basis of a more general solution is the observation of Rabbani and Lamb that a "well behaved"† C-t transient not to close to equilibrium can be described by,

$$W(t) = W_0 e^{-t/K} \quad \text{for } t<t_f \tag{6.8}$$

and,

$$W_f = W_0 e^{-t_f/K} \quad \text{for } t>t_f \tag{6.9}$$

where W(t) can be determined from the measured capacitance using a generalisation of equation 3.23. Thus a plot of lnW(t) against t should be linear with slope -1/K. Figure 6.5 confirms that the empirical Rabbani equations can be an accurate model of the time evolution of the depletion layer width. This plot has been used to determine values for τ_g [201, 215]. However, since the method of parameter determination does not properly separate τ_g from s', the values which are determined are not accurate. Despite this flaw in the τ_g determination, accurate values of t_f can be determined using,

$$t_f = (t_1 - t_0)\left(\frac{lnW_f - lnW_0}{lnW_1 - lnW_0}\right) \tag{6.10}$$

where W_0 and W_1 are the depths measured at times t_0 and t_1 respectively, and W_f is

† "well behaved" is defined as a uniformly doped sample with weak lifetime depth dependence and negligable s_0.

the equilibrium depletion layer width. Of course when the C-t method is used in isolation, W_f is not known until after the measurement. However, for uniformly doped samples it could easily be calculated with sufficient accuracy for use in 6.10. A better approach, and one which is also accurate for non-uniformly doped samples, is to pass W_f as a parameter from a CV measurement program. Equation 6.9 gives an estimate of t_f which is independent of the deep depletion voltage, oxide thickness and substrate doping density, (provided it is uniform). Figure 6.6 shows the correlation of estimated t_f and actual finishing time. The excellent correlation confirms the applicability of this approach. The Zerbst data acquisition and analysis program "LIFE", uses a generalisation of equation 6.9 to arbitrary profiles such that,

$$t_f = \gamma \frac{1}{N(W_o)(W_f - W_o)} \int_{W_o}^{W_f} N(W) dW (t_1 - t_0) \left(\frac{lnW_f - lnW_0}{lnW_1 - lnW_0} \right) \tag{6.11}$$

where N(W) is the doping profile, and γ is a constant whose value is set to take account of τ_{geff} variations in implanted areas. The origin of the weighting factor introduced to 6.11, can be seen by inspection of equation 6.5 where the magnitude of the left hand derivative is inversely proportional to the doping density. There is insufficient data to give accurate values for γ, but reasonable estimates for the implanted samples T18 and T22, (details in table 5.1), would be 0.5 and 1 respectively. This premultiplier reflects the influence of implantation damage reducing the carrier lifetime in implanted areas. The differing γ for different implants is due to differences of dose and energy, which respectively effect the extent of crystal damage and the proximity of the damage to the deep depletion zone. For the two wafers reported here, the implant in T18 is of a higher dose and greater energy thereby increasing its influence on the dynamic recovery behavior.

Once a reliable estimate of t_f has been found the next problem is to space the data to obtain a noise free zerbst plot. Quantitatively it can be shown that when ΔC is

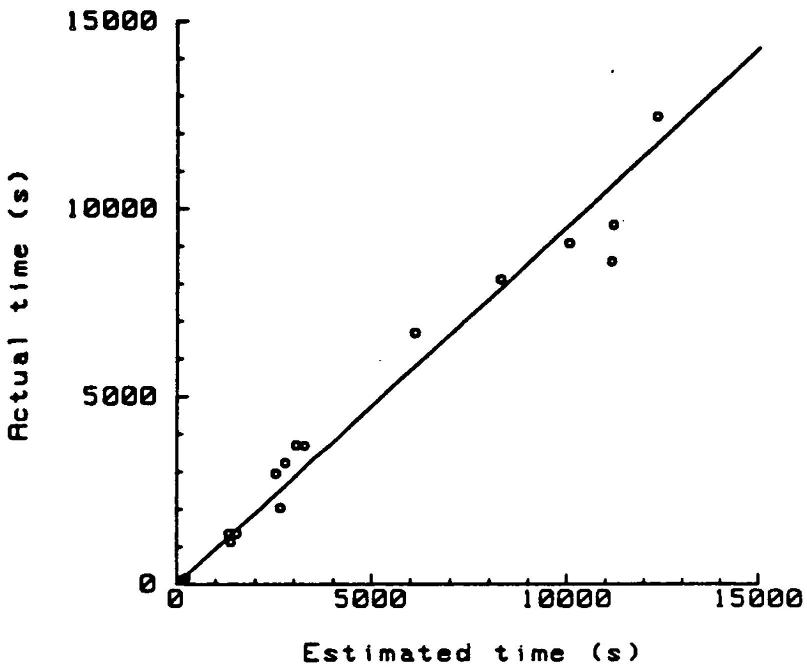


Figure 6.6 Correlation of the actual finishing time with the finishing time estimated using equation 6.10.

used to approximate dC, the noise in the Zerbst plot will be given by,

$$\Delta\tau_g/\tau_g = 2R/\Delta C \qquad (6.12)$$

where R is the characteristic sum of the capacitance meter random and rounding errors, provided that $\Delta C >> C$, which is exactly the same as for profiling. The reason can be seen by rewriting equation 6.3 to obtain,

$$\frac{C(t)^3}{C_{ox}}\frac{dC(t)}{dt} = \frac{n_i}{N_{sub}\tau_{geff}}\left(\frac{1}{C(t)}-\frac{1}{C_f}\right) + \frac{n_i s'}{\epsilon_s N_{sub}} \qquad (6.13)$$

which is of a similar mathematical form to equation 5.3.

It is of course possible to optimise the time step, as was done for the voltage step in profiling. However, a simpler approach is perfectly adequate because the portion of interest in the Zerbst plot is a linear region, and so a reasonably large time step can be taken to avoid noise. In "LIFE" it is simply set to $0.05t_f$. This empirical approach would also appear to be satisfactory for non-uniformly doped samples where lifetime may vary with depth.

To obtain the best estimate of the slope, a best fit line is calculated for the linear part of the plot. (It would be possible to do this with a very noisy plot but noise might mask non-ideal behaviour). Based on 20 experimental observations, "LIFE" assumes that the linear region occurs between 0.1 and 0.3 on the x-axis.

6.4. Generation-Lifetime in Non-Uniformly Doped Semiconductors

The Zerbst analysis described in the last section can be modified to handle arbitrary profiles, by a modification of the Zerbst plot for uniform doping, whereby the y-axis is weighted by a doping factor. This could simply be N(W), but $N(W)/N_{sub}$ gives a more manageable plot.‡ Miyake and Harada[213] proposed an additional modification of the Zerbst x-axis for non-uniformly doped samples, setting it equal to the depletion layer width, W, however since the conventional x-axis is given by,

$$\left(\frac{C_f}{C(t)}-1\right) = \left(\frac{C_f}{\epsilon_s}\right)(W - W_\infty) \tag{6.14}$$

the new axis is simply a translation and then scaling of the conventional axis, and is therefore superfluous.

Figure 6.7 shows a C-t and Zerbst plot using sample T/8. The data acquisition and analysis for these plots were handled by the program LIFE using equation 6.11 to estimate t_f, and profile data from PROF. If the lifetime in sample T/8 were uniform it would be expected that figure 6.7 would resemble figure 6.3, in that the data would all lie close to or slightly above the best fit line. Since this is not the case, lifetime must have a depth dependence. To investigate this phenomena further, lifetime can be estimated as a function of depth using successive pairs of data points in the Zerbst plot to calculate a lifetime from the slope between each pair. Figure 6.8 plots the results of such a calculation, and overlays the doping profile. The inverse correlation between $\tau_g(W)$ and $N(W)$, is similar to that which was observed by Miyake and Harada[213]. An interesting point to note about this result is that T/8 was fabricated using an HCl/O_2 ambient whereas that of Miyake and Harada used pure O_2, and therefore the

‡ The origin of this modification was discussed when deriving equation 6.11.

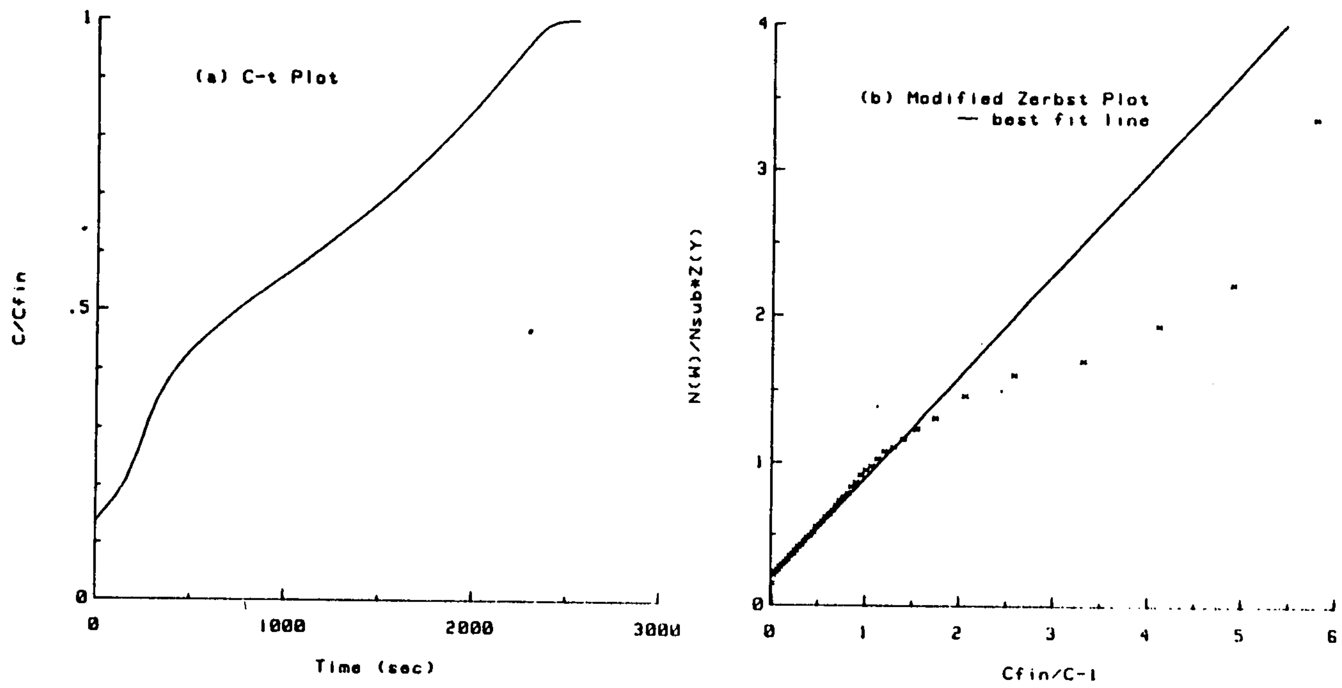


Figure 6.7 Capacitance-time and Zerbst plot for sample T18. The y-axis of the Zerbst plot is scaled to take account of the effect the profile has on the transient response.

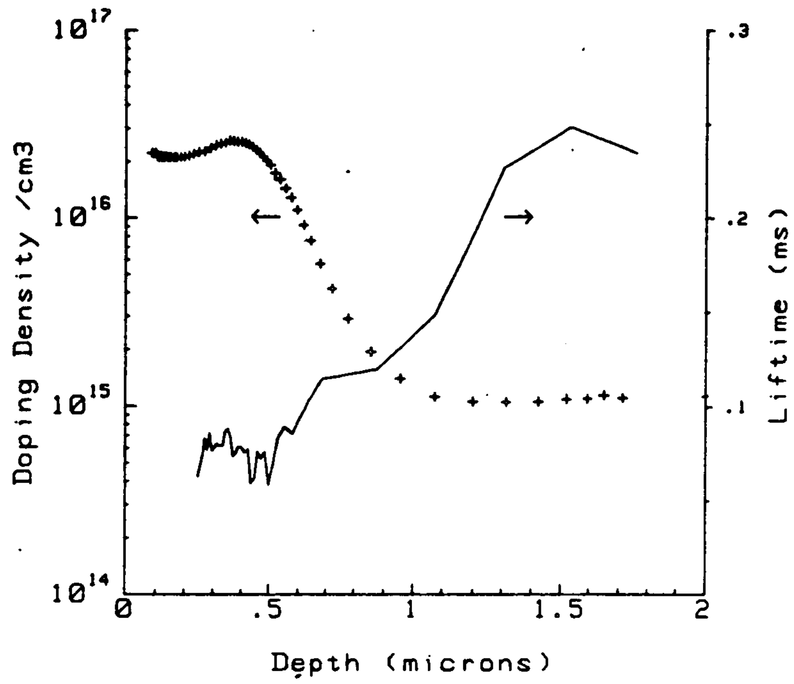


Figure 6.8 Combined plot of doping density and minority carrier generation lifetime for sample T18.

presence of chlorine introduced by this step would not appear to have any gettering effect during post implantation annealing.

The impact of implantation on carrier lifetime would not appear to have been studied by workers who have investigated the multitude of gettering phenomena in silicon, and since commercially important structures do use implants. This omission may cast doubts in the claimed effectiveness of certain gettering procedures when applied to manufacturing processes. The major reason for the omission was probably the difficulty in applying the Zerbst technique to non-uniform samples. The program LIFE overcomes this difficulty, by accessing profile information from PROF, and using the time-effective data acquisition routines described in this chapter. LIFE therefore represents a significant improvement over previously available tools for measuring minority carrier generation lifetimes.

CHAPTER SEVEN

EDUCATES (EDinburgh University CApacitor TEst Software)

7.1. Introduction

EDUCATES is an HP-BASIC† software suite which drives HP4061A test system hardware. The range of measurement programs which EDUCATES incorporates, provides the user with a more detailed picture of the electrical properties of Metal Oxide Semiconductor capacitors, than has previously been possible, whilst at the same time it reduces, (in some cases to zero), the level of user input to measurements. EDUCATES is primarily designed as a research and development tool, but its ease of use and accuracy mean that it could also be useful in a production environment.

The failings of software packages which are currently available were outlined in chapters 1 and 3, and specific features which limit the accuracy of particular measurements, or make them difficult to apply have been discussed in chapters 3-6. In this chapter the structure of EDUCATES' software, its special features, and its use are outlined in section 7.2. Unfortunately, however, the utility of EDUCATES can only be fully appreciated during a hands-on session with the software. (This would be particularly illuminating to those readers who have experience of other MOS capacitor measurement software). The system hardware is described in section 7.3, and finally in section 7.4 the current status and possible developments of EDUCATES are discussed.

† The mention of BASIC makes most "serious" programmers wince. It should be emphasised that HP-BASIC is a very advanced language, particularly for instrument control applications, and should not be likened to BASIC as it is more widely implemented.

7.2. EDUCATES Software

7.2.1. Program Access and Structure.

The EDUCATES software package consists of eight measurement and analysis programs, a menu program and a number of utility programs. It is written in HP-BASIC 3.0 and runs on an HP9127 microcomputer with an HP9133 disk drive. Conversion to other HP series 9000 computer/disk drive combinations, or to level 2.1 or 4.0 BASIC, is straightforward*.

Figure 7.1 outlines how EDUCATES is accessed from power up of the computer and test system. The BASIC 3.0 language loads automatically after switching on, and then the computer automatically loads and runs an "AUTOSTART" program, which prompts the user for the current date and time. Once this is done the EDUCATES menu is automatically loaded and run. (If the system is already powered up the menu is accessed by pressing user softkey 1).

Figure 7.2 shows the skeletal program structure which is common to all the measurement programs. (There are of course considerable differences in detail, and sometimes the order of steps is altered, or steps are added or omitted). This program structure is an important feature of EDUCATES, since it allow most of the programs to be run in one of four† modes, viz.,

* EDUCATES has evolved from software which was originally written on an HP9835A microcomputer, which ran basic a predesessor of BASIC 2.0 with only a tape loop mass storage system and no VDU graphics. It went through an intermediate phase on HP9826 and HP9816 microcomputers, (running BASIC 2.1), before reaching its current implementation.

† Two of the programs have only three modes because they can irreversibly change the sample during the measurement, and consequently they have no fully automated mode.

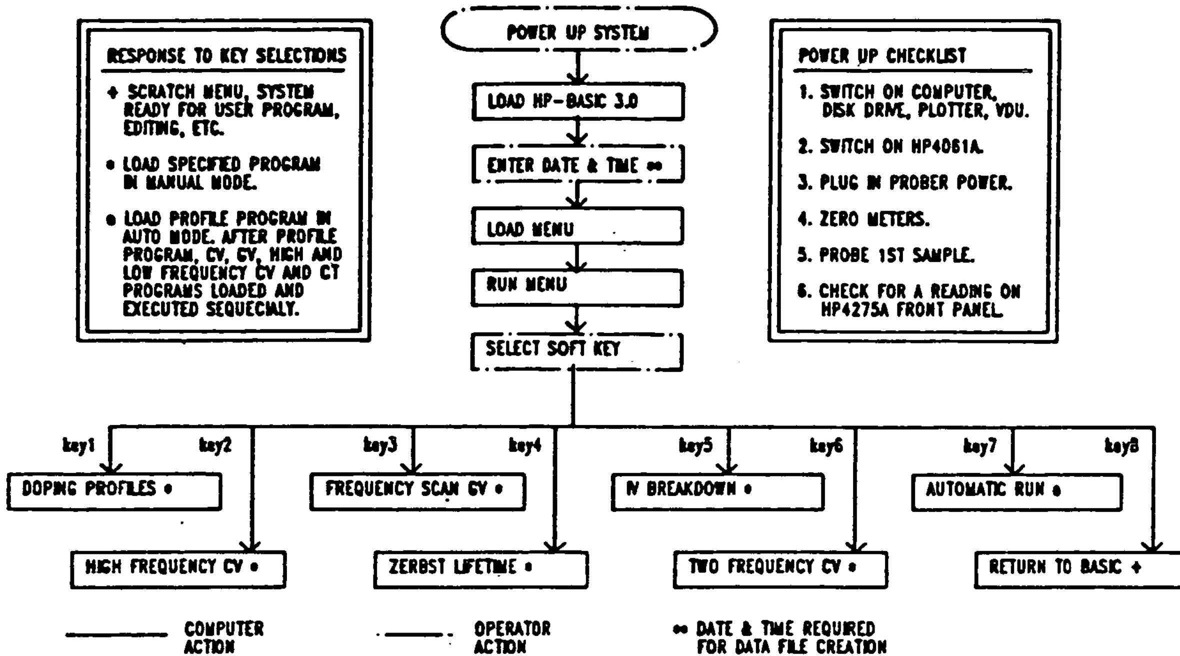


Figure 7.1 Loading and running EDUCATES software.

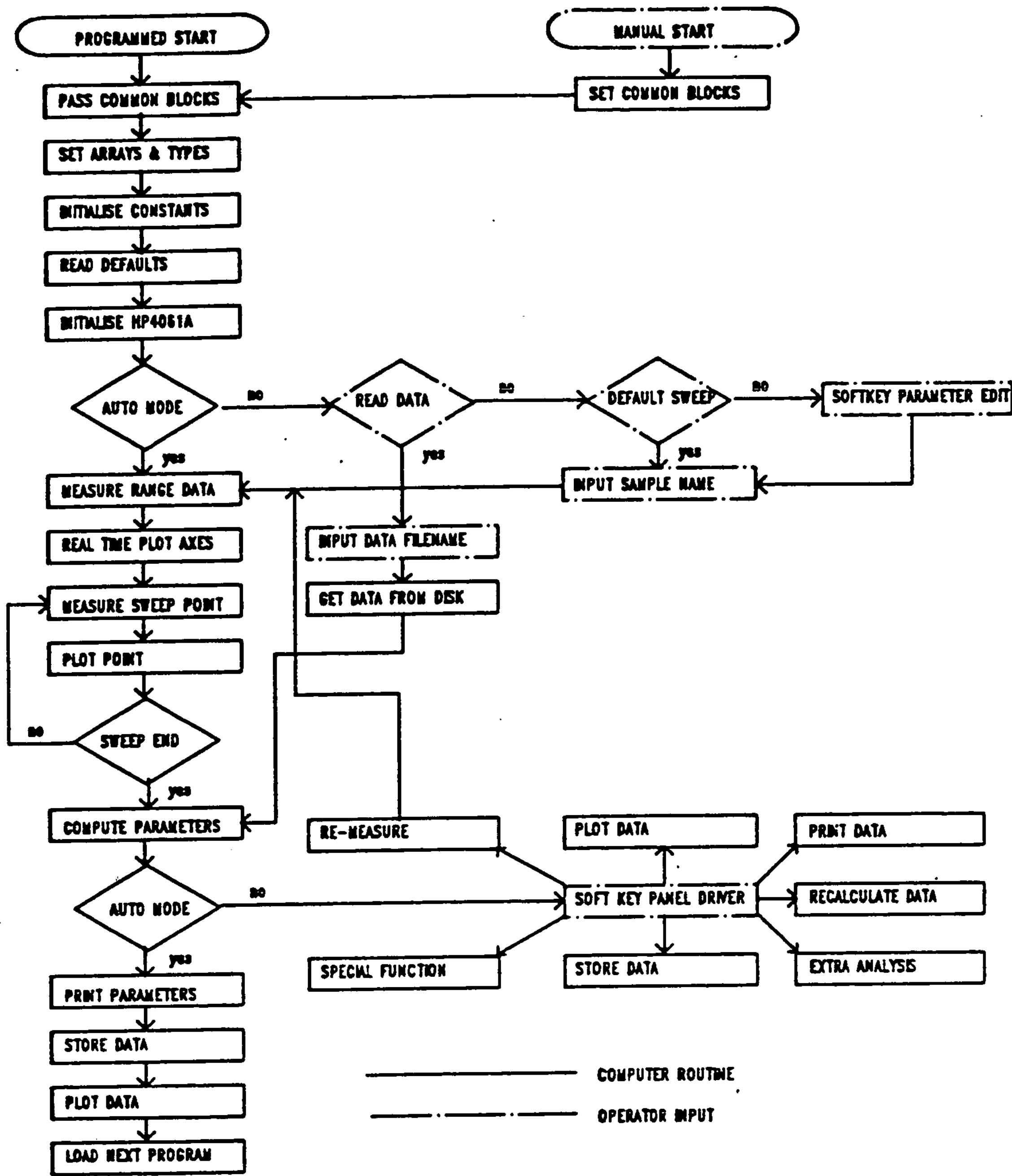


Figure 7.2 Flow diagram that is typical of an EDUCATES measurement program.

(1) *Automatic:* no user inputs are required.

(2) *Default Manual:* user inputs only the sample identifier.

(3) *Custom Manual:* user can specify special measurement or sample details. An interactive edit and display feature, means that only those parameters which need changing are accessed.

(4) *Read:* data is read from disk rather than measured.

A large amount of the input which other software requires does not vary significantly, or at all, from sample to sample. EDUCATES avoids the tedious input of invariant parameters by having default values stored with each program, whilst the interactive edit feature still allows a fully customised measurement when it is required. This structure gives a high level of flexibility, with minimal demands placed on the user, and so saves time and reduces the scope for user errors.

Each program in EDUCATES features real time plotting of the data acquisition step, in order to provide a visual assertion of the correct execution of the measurement. This can save considerable amounts of time, if for example vibration jolts a probe from a small bonding pad during a C-t measurement.

After the measurement sequence hes been completed, a wide range of plots and printouts can be generated. The majority of plots in chapters 3-6 and in chapter 8 are from EDUCATES.

The final feature which is standard throughout EDUCATES is an HP Basic Statistic and Data Manipulation[216] (BSDM) data file format. This allows both replotting and recalculating of data within EDUCATES, and also access to BSDM software, which allows the user to edit, list, plot or perform array statistics on data.

7.2.2. Measurement Programs

Although the structure of EDUCATES greatly adds to its utility, the more fundamental improvement which it represents over previous MOS capacitor analysis software packages, arises from the following;

(1) An improved understanding of the limitations and range of applicabilty of each measurement.

(2) Practical experience of the difficulties in implementing measurements.

(3) The extension of methods to allow analysis of non-uniformly doped samples.

(4) The development of algorithms which facilitate the accurate automation of measurements.

(5) The integration of methods to provide a complete parameter list, and full measurement sequence automation.

In this section a brief description of the function of each program within EDUCATES is given, and this is followed by an overview of how measurement integration is used to produce a fully automated analysis. More details of how each program functions can be obtained from their extensively "commented" listings.

Program "CVMEASURE"

Space Requirement: 66 Kbytes

Accessed Files: STANDARD or user defined with CVPREPARE.

Utilities: CVPREPARE

CVMEASURE performs the classic high frequency CV measurement. The EDUCATES version is an update of a program with the same name which has been reported previously[134], and which was developed for process control work at Hughes

Microelectronics (Glenrothes) Ltd. It has been adapted to allow for the stagnant inversion layer response, using both the pulsed and the single point C_{min} method of determining N_{sub}. When combined with the program "PROF" an improved estimate of the pulsed value of N_{sub} is obtained from an average of several measurements. For calculations the lowest estimate of N_{sub} is used, since when either the value of C_{min} is not quite at equilibrium, or when the pulsed measurement is effected by low minority carrier lifetime, the estimate of N_{sub} will be higher than the true value.

CVMEASURE provides a more rigorous assessment of the high frequency CV characteristic than is typically implemented elsewhere by analysing the curve at several points in depletion, rather than only one, and by plotting the full theoretical CV plot for that sample. The latter provides a visual assurance of the suitability of the analysis, and quickly highlights sample abnormalities, such as high interface trap densities. For uniformly doped samples, the difference between the average near surface doping, (from C_{min}), and the bulk doping density, (from pulsed CV), is computed to determine the deficit of charge[9] at the interface (Q_n/q), and the associated voltage shift, (V_n) that this will cause in the high frequency CV curve, compared to the CV curve of a completely uniform sample.

CVMEASURE differs from most other EDUCATES programs in that it uses files to set sweep and sample parameters. The files are created by the program CVPREPARE which is a highly fault tolerant and user friendly program. This was designed to enable even those unfamiliar with the use of a computer, (eg. process operators), to perform this most basic of all MOS capacitor measurements.

Program "PROF"

Space Requirement: 78 Kbytes

Files Accessed: GFUNCS

Utilities: GFUNCSGEN, (creates GFUNCS)

PROF determines the semiconductor doping profile using the pulsed CV method with the enhancements described in chapter 5. For non-uniformly doped samples the implant parameters[147] dose, range, standard deviation of range, peak depth and concentration in the silicon and surface concentration are all computed. Routines are also included which allow implementation of the "CdV" method[148], though these are not normally executed. PROF accesses the data file GFUNCS, which contains tables of the "g" functions described in [180], which are required for determination of the near surface profile. GFUNCS can be recreated if required by the program GFUNCSGEN.

PROF also calculates values of V_{fb} and C_{fb} for *all* profiles using the method described by Reuter[182]. This evaluation of V_{fb} can be used to extend the GV and dual frequency CV determinations of $D_{it}(E)$, to non-uniformly doped samples.

Program "GV"

Space Requirement: 122 Kbytes

Files Accessed: None

Utilities: None

Program GV determines the interface trap density as a function of bandgap energy, $D_{it}(E)$ using an adaptation of the Brews conductance method[12]. This may be the first ever fully automated implementation of a conductance method for

determining interface trap densities. (This method was pioneered at Bell labs but they do not have this capability!). The current version of the program operates best with 7 frequencies, (1MHz, 400kHz, 200kHz, 100kHz, 40kHz, 20kHz and 10kHz), on the HP4275A, but can be readily adapted to other frequencies using HP4276 or HP4277 LCZ meters. However, the use of the latter meters is not recommended except for high ($> 5 \times 10^{10}$ cm^{-2} eV^{-1}) trap densities because of the poorer resolution of these instruments.

At the end of the measuring sequence the interface trap capture probability, time constant, interface trap capacitance (C_{it}), and silicon capacitance (C_s) are all computed[12].

The analysis of the conductance-band bending peaks is performed at two fractions of the peak maximum to provide a simple evaluation of the validity of the analysis. A more rigorous test of validity can be performed using extracted parameters as the inputs to a conductance peak simulator which is part of "GV".

At present the application of GV to non-uniformly doped samples should be restricted to samples with interface trap densities $<1 \times 10^{11}$ cm^{-2} eV^{-1}, because in non-uniformly doped samples band bending is determined by assuming that the high frequency capacitance is a satisfactory approximation for the low frequency capacitance. In the samples on which this routine has been evaluated, (T18 and T22), this is an excellent approximation since the interface trap capacitance is less than 0.1% of the silicon capacitance. The method may be less satisfactory where the oxide thickness is significantly greater, in which case the low frequency determination of band bending will be more accurate. At present there is no facility to access low frequency data in GV.

Program "HFLF"

Space Requirement: 117 Kbytes

Files Accessed: PROF data files if user specified

Utilities: None

HFLF measures high frequency CV and Quasi-Static CV, (QSCV), curves for the determination of $D_{it}(E)$, by the high frequency, low frequency and combined high and low frequency methods which were reviewed in chapter 4. This allows a range of trap densities to be measured over a potentially wider energy range than the GV method. (Chapter 4 explained why this potential is only rarely fully realised). HFLF measures the high frequency CV curve over a narrower voltage range with a finer voltage step than CVMEASURE, in order to ensure that both capacitance characteristics are measured at exactly the same voltage points, thereby avoiding any interpolation errors in the combined frequency method. For non-uniformly doped samples, only the combined frequency method can be implemented.

A routine is included in HFLF for the determination of doping profiles from equilibrium CV measurements[9], which includes a facility for correcting interface trap distortion[181], and in addition results of these measurements can be compared with pulsed profile determinations with a routine which accesses PROF data files. Code is present for removing interface trap stretchout from CV curves using the method of Nicollian and Brews[9]. However, because experience has shown this to be of little use due to errors in low frequency capacitance values, the routine is not normally executed.

Program "LIFE"

Space Requirement: 47 Kbytes

Accessed Files: PROF data file for non-uniformly doped samples

Utilities: None

LIFE determines the minority carrier generation lifetime from a capacitance-time measurement and a Zerbst plot. Chapter 6 has discussed the methods which are required to apply this method in a time efficient manner, and to assess non-uniformly doped samples. LIFE uses PROF data files when a non-uniformly doped sample is being measured. A facility to superimpose a plot of the profile and lifetime as functions of depth allow a rapid assessment of correlation of these two functions.

The HP4275A meter will give the best results with LIFE, when the total transient time exceeds 40 seconds, whilst for transients in the range 10-40 seconds, the HP4277A which measures capacitance more quickly, should be used. If the recovery time is less than 10 seconds then the HP4277A can still be used but the transient has to be measured several times to acquire enough data for C-t and Zerbst plots. (This has been tested down to 0.1 second total transient duration, though accuracy is uncertain).

Program "BTMEASURE"

Space Requirement: 35 Kbytes

Files Accessed: BTEST or user defined by BTPREPARE

Utilities: BTPREPARE

BTMEASURE measures the shift which is caused by mobile ions in the high frequency CV characteristic after bias temperature stress. The use and structure of this

program as implemented in EDUCATES remains unchanged from when it was previously reported[134]. Its major advantage over other mobile ion determinations is that by measuring shift over a range of depletion capacitances, a check is made to ensure that the shift is parallel and hence due to mobile ions and not any other defect mechanism. Bias-temperature stress measurements are performed on a modified Wentworth "TC100 tempchuck", which is separate from the main probe box used by all other measurements. Care should be taken to ensure that the probe does not slip off, or scratch the capacitor gate, during this measurement, due to either thermal expansion, or vibration which is present at the initial stage of cooling.

BTMEASURE cannot be run as part of the EDUCATES fully automated set up, because the measurement can irreversibly damage the sample.

Program "IVBREAK"

Space Requirement: 20 Kbytes

Files Accessed: None

Utilities: None

IVBREAK performs a ramped voltage sweep on a capacitor until a voltage is reached where the gate dielectric breaks down and allows a significant gate to substrate current to flow[154, 217, 218]. After breakdown, a capacitor cannot be used for further testing, therefore IVBREAK is not part or the automated EDUCATES test sequence. In EDUCATES this program is designed simply for checking oxide integrity, should this be suspect, before making capacitance measurements. (Neither the program nor the HP4061A and prober hardware are well suited to statistical oxide

breakdown studies).

7.2.3. Automatic Mode

The EDUCATES automatic mode allows a sequential execution of PROF, CVMEASURE, GV, HFLF and LIFE. This allows the user to obtain an extensive sample characterisation with only one input into EDUCATES. Thus the user can set a measurement sequence going and let it run unattended for several hours, whilst continuing with other work.

The automatic mode is made possible by the combination of a set of individual programs, each one of which has full internal automation, (this is achieved by routines such as voltage sweep optimisation in PROF), and by a forward pass of parameters from one program to the next. Table 7.1 lists these parameters which EDUCATES passes via global common blocks. (Nw(*) is passed in a data file).

The automatic mode was specifically designed for use in experiments described in chapter 8 of this thesis, in order to provide as complete a picture as possible of the samples. It would be a simple matter to reconfigure the automatic mode with a different measurement set. All this requires is a resetting of the next measurement loaded within each program. It should be noted that any sequence should always begin with PROF because of the importance of the profile in all measurements, (even if only to verify that the profile is indeed uniform). The next in the sequence ought to be CVMEASURE, because this helps GV, HFLF and LIFE measurements to be performed over the correct voltage range, and an accurate estimate of transient time in LIFE. The order of the other measurements is not important.

Table 7.1 Automation Parameters			
Parameter	**Source Program‡**	**Dependent Programs**	**Function**
Auto	MENU (2)	All	Set program mode to automatic
S$	MENU (1)	All	Sample identifier. This appears in all printed output and data file names.
Areag	MENU (1),(2) PROF (3)	All	Calibrates parameters to absolute values. If oxide thickness is input Areag is calculated by PROF.
Uni	PROF (4)	All	Flag which identifies profile as uniform (1), or non-uniform (0).
Dopden	PROF (3) CVMEASURE (3)	All	Sets silicon bulk potential
Nw(*)	PROF (3)	LIFE	Allow for profile in t_f estimate.
Flatv	PROF (3) CVMEASURE (3)	GV * HFLF * LIFE †	* Adjusts sweep range limits and is used as an integration constant. † Adjusts depletion voltage.
Cming	CVMEASURE (3)	LIFE	t_f estimate
Cvfil$	MENU (2)	CVMEASURE	file containing measurement parameters

‡ Parameter Origins are: (1) Operator Input, (2) Default, (3) Calculated, (4) Empirical

7.3. EDUCATES: System Hardware

EDUCATES software is specifically designed to run on HP4061A hardware. This system consists of an HP4275A LCR meter and an HP4140B picoammeter, either of which can be connected to the device under test, (DUT), via the HP16057A switching matrix, which is controlled by the HP4083 switching controller. All the instruments are controlled by a microcomputer using an IEEE-488 parallel interface. Figure 7.3 is a schematic wiring diagram for the system. In this section the major modifications to the system, (compared with that described in the system manual), are listed. The changes are:

(1) The instrument driver software in EDUCATES allows either the HP4276A or HP4277A LCZ meters to be substituted for the HP4275A. This expands the frequency range over which measurements can be performed.

(2) The probe station lights are connected to the general I/O port of the HP4083, which can switch them on or off. This facility is used actively in the measurement programs HFLF and CVMEASURE.

(3) The system computer is now an HP9217, which has 1Mbyte of RAM, an RS232 interface and an IEEE-488 interface.

(4) The mass storage is an HP9133 "hard" disk, which has been initialised in 6 volumes, and one 3½ inch floppy drive. The HP-BASIC language system is stored in hard disk volume 0, EDUCATES is stored in volume 1, and data files are stored in volume 3. The floppy is used to back up the hard disk.

(5) The system printer is now an HP "Thinkjet" printer

(6) The output from the RS232 interface can be switched to a line on the VAX 750 mainframe computer. The program "EMULATOR" provides a pseudo terminal and data file transfer facility.

This new system provides improved measuring facilities, an excellent programming

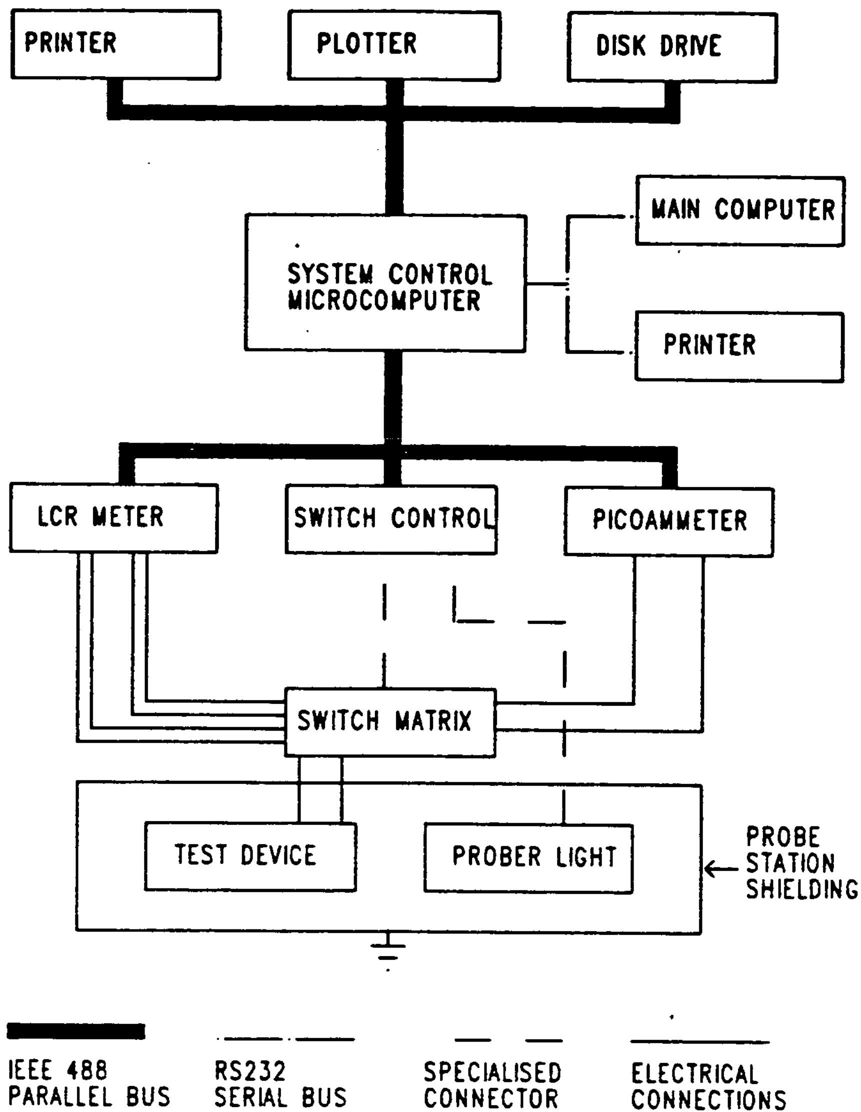


Figure 7.3 Wiring diagram for test system hardware.

environment, and flexible output channels for data.

7.4. EDUCATES: Discussion

7.4.1. Software Status

The functionality of EDUCATES has been rigorously tested and debugged on more than thirty uniformly doped wafers, and two wafers which had implants which are typical of state-of-the-art fabrication processes. Therefore the system is ready for general use, although at the present time it would be desirable to further rationalise the system graphics and handling of user inputs in custom manual mode. Once these straightforward but time consuming tasks have been performed, the writing of a user manual for the software will be greatly simplified.

7.4.2. Future Development

The majority of EDUCATES' software could be used in future to provide wafer maps of parameters, if an automated wafer prober replaced the current manual one. (Lifetime mapping is not feasible, because for example it would take 7 *months* to measure only one six inch wafer with a thousand capacitors with average lifetime in excess of one millisecond).

The replacement of the present chuck with one which was thermally controllable via an IEEE-488 interface would allow the measurement of additional parameters such as the recombination lifetime and the mobile ion activation energy.

From October 1986 onwards, Jim Walls will carry out a CASE PhD project in association with Hewlett Packard Ltd., with the aim of producing an "expert" capacitor measuring system, i.e. a system which can not only perform measurements automatically but can also intelligently interpret results and the validity of tests. EDUCATES

is a suitable foundation from which to develop such a system.

7.5. Conclusions

EDUCATES has lifted capacitor testing to a new level of sophistication, whilst also removing the requirement of having a skilled technician supervising measurements.

CHAPTER EIGHT

The Electrical Effects on MOS Capacitors of Trace Dissolved Impurities in HF.

8.1. Introduction

The importance of silicon wafer cleaning and cleanliness during processing, which are necessary because of the extreme sensitivity of devices to ultra trace levels of contamination was discussed in chapter 2. In this chapter the discussion of ultra trace impurity effects is extended and the purity requirements for processing chemicals is examined experimentally.

It is worthwhile to note at this point that contamination levels which are of importance in silicon are far lower than are ever dealt with in conventional inorganic chemistry, and so table 8.1 defines the concentration ranges which are dealt with in this chapter.

Since ultra trace levels are of importance, all sources of contamination are tightly controlled, even when there is no evidence that this is necessary. A case in point of this type of control is the specifications for chemicals used in device processing. Table 8.2 shows a specification for 50% HF acid as defined by the Semiconductor Equipment Manufacturers Institute, (SEMI), and from two chemical producers. All three specifications are different, begging the following question from the device manufacturer.

Which specification meets my process's requirements?

It is difficult, (if not impossible), to provide a reasoned answer, because the specifications mainly reflect the ability of the chemical manufacturer to analyse large numbers

Concentration/S Atoms/cm^{-3}	Concentration Atoms/(Si Atom)	Range
10^{22}	1	Silicon Crystal Atom Density.
10^{13} - 10^{17}	1 p.p.b. - 10 p.p.m.	"trace"
10^{10} - 10^{13}	1 p.p.t. - 1 p.p.b.	"ultra trace"

Table 8.1 Definition of concentration ranges.
p.p.m.=parts per million, p.p.b.=parts per billion, p.p.t.=parts per trillion.

Table 8.2 Specifications for MOS Grade 50% HF

Element	Organisation		
	SEMI	Riedel de-Hahn	Merck
Antimony	0.03	0.01	0.02
Arsenic	0.03	0.05	0.05
Aluminium	0.05	0.05	0.05
Barium	0.5	0.1	0.1
Boron	1.0	0.02	0.05
Cadmium	1.0	0.01	0.05
Calcium	1.0	0.5	1.0
Chromium	0.01	0.01	0.02
Cobalt	0.5	0.01	0.02
Copper	0.05	0.01	0.02
Gallium	0.05	0.02	0.02
Germanium	1.0	*	*
Gold	0.5	0.02	0.1
Iron	0.5	0.1	0.5
Lithium	1.0	0.02	0.02
Magnesium	0.5	0.1	0.2
Manganese	0.5	0.01	0.05
Nickel	0.1	0.01	0.02
Potassium	1.0	0.1	0.1

N.B continued overleaf.

Table 8.2 Continued			
Element	Organisation		
	SEMI	Riedel de-Hahn	Merck
Silver	0.1	0.02	0.02
Sodium	1.0	0.2	0.2
Strontium	1.0	0.02	0.1
Tin	1.0	0.02	0.1
Zinc	1.0	0.05	0.1
Lead	0.1	0.02	0.05
Berylium	†	0.01	0.02
Bismuth	†	0.02	0.1
Indium	†	0.02	0.02
Molybdenum	†	0.01	0.05
Platinum	†	0.02	0.2
Thallium	†	0.02	0.05
Titanium	†	0.1	0.1
Vanadium	†	0.01	0.05
Zirconium	†	0.01	0.1

* Not Specified by Producer

† SEMI specifies these elements to have the same limit as lead.

of impurity concentrations simultaneously, using an inductive coupled plasma spectrometer, rather than being based on known effects of specific impurities. The origins of the concentration limits set in specifications are a further indication of the limited usefulness of such data, since they are simply set to an arbitrary multiple of the spectrometer's resolution limit for rare impurities; a value slightly above the normal measured level for common impurities; or match values available from competitors. For example Calcium in 50% HF, (table 8.2 Riedel), has a limit of 1.0 p.p.m. because it is a common impurity in the feedstock from which MOS grade HF is distilled, whereas gold has a limit of 0.02 because it is rare[219].

Currently there is still demand for tighter specification of process chemicals, because of a fear of the unknown, due to the complete absence of any data relating dissolved impurity levels in liquids to electrical effects in devices. Reducing impurity levels below that which is currently available may be unnecessary and will certainly be expensive, and therefore it is valuable, as this chapter attempts to do, to quantify the electrical effect, (if any), that specific impurities have.

8.2. Measuring Impurity Effects

In the past, the limited quantity of work which has been published on the contamination and cleaning of wafers with liquid chemicals has relied almost totally on chemical analysis. In most cases the techniques which are used, such as ESCA, Auger or SIMS, lack sufficient resolution for detecting the ultra-trace concentrations which can effect device performance. For example Schwettmann[74] et al observed enhanced oxide growth rates for certain samples in a cleaning experiment, but they could not detect any species on the wafer surface using ESCA and Auger to explain the phenomenon. It has often been the case that in order to compensate for the lack of resolution of chemical analysis that the contamination used is very gross. For example

Schmidt[220] used the following contamination of wafers when evaluating cleaning procedures.

(1) 30 minutes in 6 litres of aqua regia in which a copper disk; a copper plated stainless steel disk; and two gold plated headers had been placed.

(2) 10 minutes in 4 litres of city water to which 20g of $NaCO_3$; 20g of NaCl; and 20g of KOH had been added.

It is extremely doubtful whether the chemistry of cleaning with contaminations several orders of magnitude greater than might feasibly occur in a clean room, can be extrapolated to ultra-trace chemistry.

Since the chemistry has such severe limitations, the chemical contaminationn experiments described in this chapter are assessed via their effect on MOS capacitor electrical parameters. This approach itself would not have been feasible without the development of the capacitor measurement techniques described in chapters 3-7. Electrical measurements have the additional advantage that they only detect the important electrically active impurities. Results of electrical experiments also provide a useful data base for diagnosing the cause of poor electrical parameters that are detected during routine process monitoring.

8.3. Preparations for Contamination Experiments

It was not feasible to examine the specifications of the 40 or more MOS grade chemicals which are on the market, and in fact looking at only one is quite a major task. It is therefore important to select a chemical for analysis which represents the severest test of the specifications. 5% hydrofluoric acid, (HF), was chosen for this purpose for the following reasons:

(1) It is the most extensively used chemical, (other than DI water), in most fabrication recipes.

(2) It is used in the critical pre-oxidation cleaning step.

(3) Its chemical reactions directly expose the silicon surface.

Even with only one acid the total quantity of tests which can be devised is vast if all elements are to be assessed. This work aims to make a start by using elements which are commonly found in the wafer fabrication environment. In order to test HF it was first necessary to establish an MOS capacitor recipe which would produce capacitors with quasi-ideal properties. Table 8.3 is a brief description of the recipe which was found to fulfill this requirement. Over a six month period this recipe consistently produced samples with high minority carrier generation lifetimes and low oxide charge densities. There was one major difficulty with the process at that time however, because p-type substrates had been used. These were found to be extremely sensitive to the ILBG effect, (chapter 3.7), which made measurement of the high frequency CV characteristic and minority carrier lifetime unreliable. To overcome this problem all subsequent wafers used n-type substrates where the ILBG effect does not normally occur.

Schwettmann et al.[74] and Matlock[73] have reported that the final step in a pre-oxidation cleaning sequence can have a marked effect on oxide growth rates and possibly also degrade electrical parameters. Since the cleaning sequence in table 8.3 leaves the silicon surface hydrophobic, whereas the industry standard clean, i.e. acidic RCA clean (H_2O:HCl:H_2O_2,5:1:1), leaves the surface hydrophilic, it was decided to check whether the more standard clean produced better capacitors than the clean ending with 5%HF.

1. Starting material: 3" 100 n-type silicon wafers. Substrate doping density 1.5 x 10^{15} cm^{-3}. Wafers are not subject to heat cycling or back surface damage to enhance gettering.

2. Pre-oxidation clean is $HCl/H_2O_2/H_2O$ (1:1:5) for 10 minutes, a five cycle rinse in a dump rinser, a 60 second dip in 5% HF followed by DI rinsing and spin drying. (Cold fast-circulation DI.)

3. Oxidation in dry oxygen at 950°C with 5% HCl for 2 hours followed by an in-situ anneal in N_2 for 20 minutes. Single wall quartz furnace tube.

4. Metal deposition in an e-beam evaporator and lithography to pattern MOS capacitor gates with a guard ring at 3 microns.

5. Back oxide strip and metal deposition for ohmic contact.

6. Post metal anneal at 435°C in forming gas for 20 minutes.

Table 8.3 MOS Capacitor Fabrication Recipe

Clean	Flatband Voltage (V)	Interface Trap Density ($x10^{10}cm^{-2}eV^{-1}$)	Minority Carrier Lifetime (ms)
RCA	-0.2	1.5	2.1
RCA/HF	-0.2	1.5	2.5

Table 8.4 Electrical Parameters for Different Cleans

This check was performed with 4 wafers, two of which followed table 8.3 exactly, whilst for the other two the final 5%HF dip was omitted. The results of this experiment are given in table 8.4, and shows there is no significant difference between the cleans, and therefore table 8.3 produces capacitors of at least the same quality as the industry standard. Thus a useful basis for contamination experiments was established.

The results of the cleaning experiment reported here are somewhat different from those previously reported. In the current experiments two radically different cleans both produced quasi-ideal MOS capacitors with the same oxide thickness. It may be that the differences reported by previous authors were caused by peculiarities of their processing environment such as; the DI water quality; the cleanliness and type of cleaning equipment; the quantity and form of rinsing; furnace cleanliness; or the purity of process gasses.

8.4. Contamination Experiments

8.4.1. Preparation of Contaminated 5% HF and Oxidation of Samples

Contamination experiments were performed by adding trace quantities of metal ions to the 5% HF used in step 2 of table 8.3. Sample codes, the type and quantity of metal ions and the method of rinsing samples are detailed in table 8.5. One bottle of HF was used throughout to ensure the base liquid did not contribute any unwanted effects. The source of metal ions was 1000 p.p.m. atomic spectroscopy grade metal nitrates. This liquid form of contaminant minimised handling thereby significantly reducing the chance of any additional inadvertent contamination during preparation of contaminated HF. The 1000 p.p.m. nitrates were stabilised with a small excess of nitrate, which when added to HF produced a maximum concentration of 2.5×10^{-6} molar HNO_3. This nitric concentration is far too low to cause silicon etching and as

experimental results will show, it is the metal ion which controls the contamination behavior. Prior to each experiment the oxidation furnace was cleaned with 3 dummy oxidation runs. Each dummy oxidation consisted of six hours with O_2/HCl(5%), followed by two hours with pure O_2, at 1050°C. Teflon tweezers were used to load wafers into the furnace after cleaning.

For all the iron contaminated samples, and for three of the chromium contaminated samples, wafers were oxidised simultaneously. (They were of course cleaned and rinsed separately). The control, (uncontaminated) wafer was placed nearest the gas inlet, with the lightly contaminated wafer 3 inches downstream and the more heavily contaminated sample a further 3 inches downstream. Despite using this set up, it became evident from preliminary results that cross contamination of wafers was occurring. To overcome this effect the wafers used for subsequent experiments were processed sequentially.

8.4.2. Measurements

The following electrical measurements were performed on each of the sample wafers,

(1) Pulsed high frequency CV determination of doping profile (4)†.

(2) High frequency CV determination of flatband voltage (4).

(3) GV measurements at six frequencies for the determination of interface trap densities (2).

(4) High frequency CV and quasi-static CV determination of interface trap densities (2).

Sample Code	Contaminant Metal	Contaminant Added (p.p.m.)	Rinse Method‡
FCA	-	-	I
FP5A	Fe	0.5	I
F5PA	Fe	5.0	I
FCB	-	-	D
FP5B	Fe	0.5	D
F5PB	Fe	5.0	D
CRC	-	-	D
CP5	Cr	0.5	D
C5P	Cr	5.0	D
CRC2	-	-	D
CP1	Cr	0.1	D
CAC	-	-	D
CAP5	Ca	0.5	D
CA5P	Ca	5.0	D
NIC	-	-	D
NP5	Ni	0.5	D
N5P	Ni	5.0	D
PBC	-	-	D
PBP5	Pb	0.5	D
PB5P	Pb	5.0	D
COC	-	-	D
COP5	-	-	D
CO5P	-	-	D

Table 8.5 Levels of Sample Contamination and Rinsing

‡ I=30 seconds immersion rinse.
D=5 spray rinses of 45 second, with water filling above the wafers, then being dumped to the waste line after each cycle.

(5) Capacitance-time determination of minority carrier generation lifetime (4).

(6) Bias temperature stress CV measurement of mobile ion density (1).

† The number in brackets indicates the minimum number of measurements in each case.

The merits and implementation of these measurements have been discussed in chapters 3-6, and the software for their implementation described in chapter 7. Although the measurements reported here were not all performed with the EDUCATES software in its current form, the differences which have been made do not in any way effect the parameter values reported here, since the changes are simply associated with crash proofing and streamlining.

8.4.3. Results of Contamination Experiments

Before looking at the results of each metal ion contamination experiment it is useful to quote the tolerance levels for electrical parameters which Huff[153] has compiled for Ultra Large Scale Integrated, (ULSI) circuits. These are shown in table 8.6. The ULSI limits provide one scale for assessing the results of contamination experiments which are shown in table 8.7. The other scale which is of importance, is the range of values determined for control samples in those samples where they were processed separately from contaminated wafers. The latter provides a semi-quantitative measure of the combined effects of experimental error and inherent experimental variability. Neither of these factors can be accounted for rigorously, since for example it is easily possible to list dozens of experimental variables which could all effect MOS capacitor electrical parameters.

Table 8.6 ULSI Parameter Limits	
Parameter	**Value**
V_{fb}	±0.1V
$D_{it}(E)$	$2\times10^{10}cm^{-2}eV^{-1}$
τ_g	300μs-1ms
Q_m	0

Sample	Oxide Thickness (A)	Flatband Voltage (V)	Interface Trap Density ($cm^{-2}eV^{-1}$) $x10^{10}$	Minority Carrier Lifetime (ms)	Lifetime Range (ms)	Flatband Voltage Shift (V)
FCA	830	-0.22	0.5	1.33	1.01-2.93	-0.2
FP5A	842	-1.20	0.4	2.14	1.33-4.30	-1.7
F5PA	860	-2.60	0.5	1.43	0.83-2.18	-4.6
FCB	780	-0.15	1.7	1.75	1.30-2.35	-0.2
FP5B	782	-0.31	0.6	0.85	0.73-1.12	-0.3
F5PB	779	-1.45	0.6	0.70	0.58-0.82	-1.3
CRC	780	-0.18	1.8	0.26	0.14-0.35	0.0
CRP5	777	-0.10	1.8	0.51	0.41-0.62	0.0
CR5P	790	-0.10	1.9	0.44	0.37-0.55	0.0
CRC2	765	-0.13	1.9	1.86	1.21-1.96	0.0
CRP1	771	-0.13	1.8	3.45	2.70-5.34	0.0
CAC	775	-0.12	2.0	0.9	0.75-1.26	0.0
CAP5	768	-0.10	1.9	1.1	0.81-1.42	0.0
CA5P	772	-0.10	2.0	1.1	0.58-1.58	0.0
NIC	777	-0.12	2.3	0.44	0.35-0.49	0.0
NP5	778	-0.14	1.8	0.70	0.64-0.78	0.0
N5P	788	-0.06	1.9	1.43	0.58-1.90	0.0
COC	779	-0.17	2.7	2.93	1.19-5.38	0.0
COP5	772	-0.14	2.5	1.28	0.95-1.95	0.0
CO5P	773	-0.11	2.2	1.38	1.08-1.57	0.0
PBC	790	-0.11	1.7	2.08	1.88-2.35	0.0
PBP5	794	-0.13	2.5	0.39	0.12-0.59	0.0
PB5P	789	-0.16	2.3	0.64	0.37-0.70	0.0

Table 8.7 Measured Parameter Values

When the experimental results are compared with table 8.6 it is apparent that the vast majority of results meet the ULSI specification and are within the range of control data. This type of result is referred to hereafter as ***insignificant***. It will be shown in following sections that the few results around the margin of this regime, i.e. up to twice the range, can all be explained by assuming that the control samples do not cover the entire spectrum of process variability. Results outwith the marginal regime are referred to hereafter as *gross*.

It is important to note that the variability of results in table four arises due to the processing, and is not a reflection, as occurs with chemical analysis, that the measurements are resolution, or noise limited.

The measurements of minority carrier lifetimes that are reported here suffer from the same limitations as have been reported elsewhere [5,155,221], namely that the spread of lifetimes which are determined is large, (at least an order of magnitude), even for control samples. The reason why this is so can be seen by examining the chemical concentrations responsible for the degradation of lifetime. For example, using the data of Richou et al.[84], a capacitor with 5 parts per trillion of gold in silicon will have a minority carrier lifetime of 1 millisecond, whereas with 50 parts per trillion its lifetime would be 100 *micro* seconds. When such small concentrations are of importance, it becomes impossible to control the processing environment in such a way as to guarantee cleanliness at these levels, and hence the interpretation of experimental results will not normally be straightforward.

8.4.3.1. Calcium contaminations

Calcium contamination at levels of up to 5p.p.m. in 5% HF have an insignificant effect on MOS capacitor electrical parameters, i.e. all parameters are well within the limits of the ULSI specification and also the range of control data. Thus there would appear to be no interaction between calcium in the acid and an electrical degradation of an MOS structure.

8.4.3.2. Iron Contaminations

The two experiments performed with iron contamination, show that this metal ion can cause large negative flatband voltage shifts due to fixed oxide charge. The magnitude of these degradations is proportional to the quantity of metal ion added to the 5% HF.

These effects are unlikely to be due to any inadvertent contamination, since only the two iron contaminated samples which were processed more than one month apart, had high fixed charge densities and high mobile ion densities. Nothing of this magnitude was observed for any other sample throughout the course of all the other experiments, and in particular the samples CRC, CRP5 and CR5P were all processed between the time of the first and second iron experiment. (The flatband voltage which is more negative for sample FCA, than any other control sample, and the measurable mobile ion densities for samples FCA and FCB are almost certainly due to cross contamination during oxidation). In addition to the circumstantial evidence, flame emission spectroscopy was used to show that the level of sodium contamination, (the most likely candidate), was actually less in iron contaminated solutions than in chromium contaminated solution.

It is not certain that it is iron which directly gives rise to the oxide charge levels. It could be speculated that iron causes an enhanced deposition of mobile ions onto

silicon.

The oxide thickness of wafers given only a short immersion rinse, was proportional to the level of iron added to the 5% HF. Extensive rinsing removed this effect whilst the electrical degradations were still present so iron can be ruled out as a cause of enhanced oxidation rate. The involvement of the other five metals in oxidation rate enhancement can also be discounted since they show significant oxide thickness variation. It remains a mystery however as to the exact cause of the rate enhancement.

The low interface state densities in iron samples is also an enigma. It may be that the conductance peak is not properly resolved, because it is excessively wide, due to large interfacial band bending fluctuations produced by the high level of oxide charges. Alternatively the low density might be real, and caused by an ion, (either iron or sodium), enhancing the annealing, or preventing the creation of interface traps. Evidence that the latter might be possible comes from Hillen[109] who observed a lower interface trap density in a sample deliberately contaminated with sodium than in a control sample.

8.4.3.3. Nickel Contaminations

There are slight anomalies in the results for Nickel contamination experiments. However, these can be simply explained, thereby allowing nickel to be categorised as having an insignificant effect on parameters.

The anomalies are that sample N5P has by far the least negative flatband voltage of any sample, and also that the lifetime in sample N5P is higher than for NP5, which in turn has higher lifetime than NIC. The latter effect can be explained by assuming that either NIC or the furnace was inadvertently contaminated at the start of the day's experimentation. Then the subsequent oxidations, (using HCl), progressively cleaned the furnace, thereby giving rise to the observed pattern of a progressive improvement

in lifetimes during the day. This unfortunate complication reqüires explanation, but this is not required for the conclusion to be drawn that lifetime is not degraded by nickel, simply from the fact that for sample N5P the lifetime was high.

The flatband voltage of sample N5P can be regarded as insignificant since there is no correlation between contamination levels in the liquid and flatband voltages.

8.4.3.4. Chromium Contaminations

Two experiments were performed with chromium. Both produced results to indicate that oxide thickness, flatband voltage and interface trap density were insignificantly effected by chromium contamination. The first experiment did not provide a useful answer as to whether there was a lifetime degradation effect, since all samples showed relatively low lifetime, and because the samples, (CRC, CRP5 and CR5P), were oxidised simultaneously the arguments used to explain the nickel result do not apply.† To clarify the lifetime question a second chromium experiment was performed with a control (CRC), and a sample contaminated with 0.1 p.p.m. chromium (CRP1). This lower concentration, which is still well above the specification level, was used in case there was indeed a lifetime degradation, in which case the lower concentration would cause less furnace contamination. In the event this caution was unnecessary as can be seen from table 8.6. The low lifetimes in the earlier experiment must have had the same origin as that which effected the nickel control sample.

An interesting feature of samples with lower lifetime, (i.e. <100*micro* S), is that there no lifetime depth dependence[101]. It might be expected that if contamination entered from the silicon surface that there would be lower lifetimes nearer the surface.

† It was concluded in an earlier report of this work[101], that chromium degraded minority carrier lifetime. However, with the broader range of data now available from subsequent experiments, this conclusion is now considered to be invalid.

8.4.3.5. Cobalt Contaminations

Cobalt contaminations at levels of up to 5 p.p.m. have insignificant effects on electrical parameters. There is slightly more spread than for calcium contaminations. However, these are still well within the ULSI specification.

8.4.3.6. Lead Contaminations

As with four of the other metal ions investigated here, lead contaminations at concentrations of up to 5 p.p.m. have an insignificant effect on electrical parameters. The one anomaly in the lead experiment is that whilst the average lifetimes of samples PBP5 and PB5P are markedly lower than PBC, PB5P is has higher average lifetime than PBP5. The lack of a correlation between metal ion concentration in HF and lifetime, suggests that inadvertent contamination of PBP5 had occurred and carried over to PB5P to a small extent.

8.4.4. Contamination Experiments: Discussion

From the beginning of this work it was suspected that the minimum levels of impurities deemed neccessary for MOS grade chemicals were significantly over specified. The experiments reported here show this to be the case.

The very first experiment in which 5%HF replaced the acidic RCA clean as the final pre-furnace clean, immediately proved that the quality of HF currently available is perfectly adequate, because it can be used in a process which produces quasi-ideal MOS capacitors.

Subsequent experiments were performed with 5% HF contaminated with one of six metal ions. These showed that with only one exception, concentrations of metal ions of up to 5 p.p.m. can be present and quasi-ideal capacitors produced. The exception to this is iron, which gives rise to high fixed and mobile ion charge densities. Iron

degradation can be reduced by rinsing, and fortuitously, it can be very efficiently removed with the acidic RCA clean[89]. Thus although iron shows up as a problem in this type of experiment, in practice with conventional pre-furnace cleans it will be of lesser importance.

Two general points can be extracted from the data in table 8.6. The first is that even extensive DI water rinsing does not always reduce surface contamination to tolerable levels. However, it does appear to eliminate the species which is responsible for oxide growth rate enhancement. Secondly, the presence of HCl in the oxidising ambient does not guarantee that samples will meet the ULSI specification, i.e. contrary to certain popularly held beliefs, HCl does not cure all contamination problems. When considered together, these two results suggest that it is essential to ensure that the surface of wafers is contamination free prior to DI rinsing, and also that the furnace must be clean prior to oxidation.

8.4.5. Contamination Experiments: Conclusions

It was never the intention of this project to analyse all possible contamination effects. Rather it was hoped to show with a limited range of elements that electrical analysis of degradations of MOS capacitors was a tool which could be successfully used to separate those impurities which were of any relevance in MOS grade chemical specifications, from those whose inclusion was probably superfluous. The results of experiments are now summarised.

Iron atoms at concentrations of 5 p.p.m. or less, when added to 5% HF used as the last step in a pre-oxidation cleaning sequence, will produce oxides with high levels of fixed charge and mobile ion charge. In complete contrast, concentrations of up to 5 p.p.m. of five other elements, namely calcium, chromium, cobalt, lead and nickel, produce no significant degradation of electrical properties. Five parts per million is

between ten and fifty times greater than the levels of the impurities which are given in MOS grade chemical specifications, so it can be concluded that these specifications are more than adequate. Indeed it is questionable as to whether the inclusion of many elements in the specification is even necessary.

CHAPTER NINE

CONCLUSIONS

This thesis returned to two subjects which have been of constant interest during the development of MOS integrated circuit technology, namely the effects of trace impurities and MOS capacitor measurements.

These subjects have been widely reviewed but only rarely applied at anything more than the most basic level. This has been because in practical situations it is difficult to produce accurate results for a large number of samples.

Chapters 3, 4, 5 and 6 examined a range of MOS capacitor measurements. The issues of sample preparation, measuring equipment, the choice of measurement technique, and optimised data acquisition were all considered. The result of these examinations was that they allowed accurate parameter values to be determined quickly and routinely. A number of new analyses were also described.

In chapter three the most basic MOS capacitor measurement, namely the high frequency CV method was described firstly for the ideal case, and then for the real case where defects distort the CV characteristics.

The difficulty of resolving low levels of oxide charge, or charge densities in non-uniformly doped samples was discussed. This was followed by a practical example of how the high frequency CV method can be used to control the oxidation steps of a fabrication process. It was then shown that a capacitive contact should not be used for quantitative measurements. Two new methods of determining parameters of high minority carrier lifetime samples were described.

In chapter 4 the contributions of interface states to MOS capacitance and conductance were discussed and the use of CV and conductance characteristics for the determination of interface state densities was explained.

An experimental application of four measurement techniques for the measurement of a wide range of trap densities, clearly showed that only a conductance method originally devised by Brew's and appended to and automated by this author can consistently give accurate results. An additional benefit of using the conductance method is that non-ideal interface behaviour such as 2-D non-uniformity, or single level stress induced trapping states are readily observable.

To end chapter 4 an automated conductance analysis of Aluminium and Aluminium-Silicon gate MOS capacitors, showed a difference of post metalisation annealing behaviour. This observation provided further evidence in support of the model of the annealing process proposed by Deal [126].

Chapter 5 showed that the relatively inexpensive pulsed high frequency MOS capacitor CV measurement can be used to accurately determine semiconductor doping profiles. It was shown that the profile can be measured free of noise with high spatial resolution using a digital capacitance meter if the rounding errors of the equipment are taken into account when setting up the voltage sweep.

A procedure for sweep optimisation was described which can be implemented during the measurement of samples with arbitrary profiles.

Several factors which reduce the accuracy of the pulsed CV method were discussed. It was also shown that the CV method can be successfully used to assess the accuracy of process simulators, and for wafer mapping of ion implantation dose.

In chapter 6 arguments were forwarded in support of the Zerbst method as the most practical method for the accurate determination of minority carrier generation lifetime. New methods for implementing this time consuming measurement in a time

efficient manner were described. Finally, chapter 6 assesses the effect of ion implantation on carrier lifetime. The doping profile and lifetime are shown to be strongly correlated. This result is of interest because in practical applications the leakage of implanted samples is of considerable importance.

In chapter 7 a software package given the name EDUCATES (Edinburgh University Capacitor Test Software) was described. This package provides highly automated measurement of MOS capacitor electrical characteristics. One option provides for the sequential automated measurement and analysis of the pulsed high frequency curve, the equilibrium high frequency characteristic, the low frequency CV characteristic, conductance characteristics as a function of voltage and frequency, and also capacitance time curves. Samples can be either uniform or implanted. The only requirement of the operator of these tests is to probe the sample and enter a sample identifier and the capacitor gate area.

This software significantly reduces the 'learning curve' time for an experimenter new to the subject.

The specifications for trace impurities in the hydrofluoric acid used for pre-oxidation cleaning of silicon was investigated in chapter 8 using EDUCATES software. There was no significant effect on electrical parameters for concentrations of up to 5ppm of calcium, chromium, cobalt, nickel and lead. Experiments with iron contamination showed an increase in fixed and mobile oxide change. This may be due to iron enhancing sodium deposition onto silicon. The deleterious effect of iron which occurred in the experiments reported here should not occur if more conventional cleaning recipes are used.

It was concluded that hydrofluoric acid specifications are satisfactory at present and that considerable simplification of specifications could be achieved by disregarding certain elements. The experimental procedures developed herein could be used to

assess a wider range of impurities and chemicals.

References

1. J. M. Robertson, "Microelectronic Circuit Fabrication," pp. 12-25 in *Impact of Microelectronics Technology*, ed. M. Jack, Edinburgh University Press, Edinburgh (1982).
2. B. E. Deal, *US Patent Number 3,426,422*.
3. A. S. Grove, B. E. Deal, E. H. Snow, and C. T. Sah, "Investigation Of Thermally Oxidised Silicon Surfaces Using Metal-Oxide-Semiconductor Structures," *Solid St. Electron.* **8** pp. 145-163 (1965).
4. W. M. Gosney, "Dynamic Storage Time Measurements on Metal Oxide Semiconductor Random Access Memory Circuits," pp. 58-70 in *Lifetime Factors in Silicon, ASTM STP 712*, American Society for Testing Materials (1980).
5. J. S. Kang and D. K. Schroder, "The Pulsed MIS Capacitor," *Phys. Stat. Sol. (a)* **89** pp. 13-43 (1985).
6. *HP4061A Semiconductor/Component Test System*, System Library vol. 1. 1980.
7. B. J. Gordon, "A Microprocessor Based Semiconductor Measurement System," *Solid St. Tech.*, pp. 43-47 (July 1978).
8. C. G. Shirley, "A Computer Controlled CV Characterisation System," *Semiconductor International*, pp. 81-97 (July 1982).
9. E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics And Technology*, Wiley, New York (1982).
10. S. M. Sze , *Physics of Semiconductor Devices (2nd Edition)*, Wiley, New York (1979).

11. D. G. Ong, *Modern MOS Technology*, McGraw-Hill (1984). Chapter 4

12. J. R. Brews, "Rapid Interface Parameterisation Using a Single MOS Conductance Curve," *Solid St. Electron.* **26**(8) pp. 711-716 (1983).

13. D. K. Schroder and H. C. Nathanson, "On the Separation of Bulk and Surface Components of Lifetime Using the Pulsed MOS Capacitor," *Solid St. Electron.* **13** pp. 577-582 (1970).

14. M. Zerbst, "Relaxationseffekte an Halbleiter-Isolator-Grenzflachen," *Z. Agnew. Phys.* **22** pp. 30-33 (1966). (In German)

15. J. S. Kilby, "Invention of the Integrated Circuit," *IEEE Trans. Electron. Devices* **ED-23**(7) pp. 648-654 (1976).

16. J. A. Hoerni, *Planar Silicon Devices*, paper presented at the Electron Devices meeting Washington, October 1960

17. L. C. Parrillo, "VLSI Process Integration," pp. 445-504 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

18. A. Gundlach, *Private Communication*, Mr Gundlach worked on some of the earliest British MOS fabrication lines and has stayed with the industry through to the VLSI era. He has a wealth of knowledge on many aspects (particularly relating to the practical implementation) of MOS fabrication from its inception to the present day. He is currently Operations Manager at the Edinburgh Microfabrication Facility.

19. J. A. Appels, E. Kooi, M. M. Paffen, J. J. H. Schlorj, and W. H. C. G. Verkuylen, "Local Oxidation of Silicon and its Application to Semiconductor Technology," *Phillips Res. Repts.* **25** pp. 118-132 (1970).

20. D. M. Brown, "The Self Registered MOSFET: A Brief Review," *Solid St. Tech.* **15** pp. 33-37 (April, 1972).

21. J. Sarace, R. E. Kerwin, D. L. Klien, and R. E. Edwards, "Metal-Nitride-Oxide Field Effect Transistors with Self-Aligned Gates," *Solid St. Electron.* **11** pp. 653-660 (1968).

22. P. S. Burggraaf, "Si Crystal Growth Trends," *Semiconductor International* **7**(10) pp. 54-59 (October 1984).

23. C. W. Pearce, "Crystal Growth and Wafer Preparation," pp. 9-49 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

24. D. A. McGillis, "Lithography," pp. 267-300 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

25. M. C. King, "Principals of Optical Lithography," pp. 42-83 in *VLSI Electronics-Microstructure Science (Vol. 1)*, ed. N. G. Einspruch, Academic Press, London (1984).

26. L. E. Katz, "Oxidation," pp. 131-167 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

27. E. A. Irene, "Thermally Prepared SiO_2 Films for VLSI," *Semiconductor International* **8**(6) pp. 92-97 (1985).

28. M. Hamsaki, "Effect of Adding HCl to the Oxidising Ambient on Characteristics of Metal-Oxide-Semiconductor Capacitors," *J. Appl. Phys.* **52**(5) pp. 3484-3490 (1981).

29. A. C. Adams, "Dielectric and Polysilicon Film Deposition," pp. 93-128 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

30. W. Kern, "Deposited Dielectrics for VLSI," *Semiconductor International* **8**(7) pp. 122-129 (July 1985).

31. D. B. Fraser, "Metallization," pp. 347-383 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

32. P. S. Burggraaf, "Advances in Metallization," *Semiconductor International* **8** pp. 72-79 (November 1985).

33. T. E. Siedel, "Ion Implantation," pp. 219-264 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

34. V. G. K. Reddi and A. Y. C. Yu, "Ion Implantation for Silicon Device Fabrication," *Solid St. Tech.* **15** pp. 35-41 (October 1972).

35. J. Sainsbury, "Applications of Ion Implantation in Semiconductor Processing," *Solid St. Tech.* **19** pp. 31-43 (November 1976).

36. J. C. C. Tsai, "Diffusion," pp. 169-217 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

37. C. J. Mogab, "Dry Etching," pp. 303-344 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

38. A. R. Reinberg, "Dry Processing for Fabrication of VLSI Devices," pp. 2-36 in *VLSI Electronics-Microstructure Science (Vol. 2)*, ed. N. G. Einspruch, Academic Press, London (1984).

39. D. L. Flamm, V. M. Donnelly, and D. E. Ibbotson, "Basic Principles of Plasma Etching for Silicon Devices," pp. 190-252 in *VLSI Electronics-Microstructure Science (Vol. 8)*, ed. N. G. Einspruch, Academic Press, London (1984).

40. B. Gorowitz and R. J. Saia, "Reactive Ion Etching," pp. 298-340 in *VLSI Electronics-Microstructure Science (Vol. 8)*, ed. N. G. Einspruch, Academic Press, London (1984).

41. R. T. Fulks, C. J. Russo, P. R. Hanley, and T. I. Kamins, "Rapid Isothermal Annealing of Ion Implantation Damage using a Thermal Radiation Source," *Appl. Phys. Lett.* **39**(8) p. 605 (1981).

42. F. W. Voltmer, "Manufacturing Technology for MOS VLSI," pp. 1-41 in *VLSI Electronics-Microstructure Science (Vol. 1)*, ed. N. G. Einspruch, Academic Press, London (1984).

43. N. Matsukawa, H. Nozawa, J. Matsunga, and S. Kohyama, "Selective Polysilicon Oxidation Technology for VLSI Isolation," *IEEE Trans. Electron. Devices* **ED-29**(4) pp. 561-566 (1982).

44. K. Y. Chiu, J. L. Moll, K. M. Cham, J. Lin, C. Lange, S. Angelos, and R. L. Tillman, "The Sloped Wall SWAMI : A Defect Free Zero Bird's Beak Local Oxidation Process for Scaled VLSI Technology," *IEEE Trans. Electron Devices* **ED-30**(11) pp. 1506-1511 (1983).

45. K. L. Wang, S. A. Saller, W. L. Hunter, P. K. Chatterjee, and P. Yang, "Direct Moat Isolation for VLSI," *IEEE Trans. Electron Devices* **ED-29**(4) pp. 541-547 (1982).

46. P. S. Burggraaf, "The Case for Computerised Diffusion Control," *Semiconductor International* **4** p. 37 (1981).

47. P. H. Singer, "Advances in Diffusion Furnace Technology," *Semiconductor International* **8**(5) pp. 276-283 (May 1985).

48. J. C. Maliakal, D. J. Fisher, and A. Waugh, "Trends in Automated Diffusion Furnace Systems," *Solid St. Tech.* **27** pp. 105-109 (1984).

49. S. K. Lai, "Very Thin Gate Insulators," pp. 417-426 in *Semiconductor Silicon/1981*, ed. T. Takeishi, Electrochemical Society (1981).

50. C. J. Russo, *VLSI Applications of Rapid Isothermal Annealing*, Varian Semiconductor Equipment Group Technical Report 16, January 1982, available on request

51. M. Current and A. Yee, "Ion Implantation and Rapid Annealing of 125mm Wafers," *Solid St. Tech.*, pp. 197-202 (October 1983).

52. P. S. Burggraaf, "Rapid Wafer Heating: Status 1983," *Semiconductor International* **6**(12) pp. 69-74 (1983).

53. L. Risch, C. Werner, W. Muller, and A. W. Wieder, "Deep Implant, 1 micrometer MOSFET Structure with Improved Threshold Control for VLSI Circuitry," *IEEE Trans. Electron. Devices* **ED-29**(4) pp. 601-606 (1982).

54. P. S. Burggraaf, "Silicide Technology Spotlight," *Semiconductor International* **8**(5) pp. 292-298 (May 1985).

55. D. Kahng, "A Historical Perspective on the Development of MOS Transistors and Related Devices," *IEEE Trans. Electron. Devices* **ED-23**(7) pp. 655-657 (1976).

56. B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," *J. Appl. Phys.* **36**(12) pp. 3770-3778 (1965).

57. J. D. Plummer and B. E. Deal, "Thermal Oxidation: Kinetics, Charges, Physical Models and Interaction with other Processes in VLSI Devices," in *Process and Device Simulation for MOS VLSI Circuits*, Nato Advanced Study Institute, Sogesta, Italy (July 1982).

58. R. R. Razouk, L. N. Lie, and B. E. Deal, "Kinetics of High Pressure Oxidation of Silicon in Pyrogenic Steam," *J. Electrochem. Soc.* **128**(10) pp. 2214-2220 (1981).

59. L. N. Lie, R. R. Razouk, and B. E. Deal, "High Pressure Oxidation of Silicon in Dry Oxygen," *J. Electrochem. Soc.* **129**(12) pp. 2828-2834 (1982).

60. E. Bussmann, "High Pressure Oxidation in n-Channel MOS Technology," *Semiconductor International* **6** pp. 162-166 (May 1983).

61. E. A. Irene, D. W. Dong, and R. J. Zeto, "Residual Stress, Chemical Etch Rate and Density Measurements on SiO_2 Films Prepared Using High Pressure Oxygen," *J. Electrochem. Soc.* **127** pp. 396-399 (1980).

62. E. A. Irene, "The Effects of Trace Amounts of Water on the Thermal Oxidation of Si in Oxygen," *J. Electrochem. Soc.* **121** p. 1613 (1974).

63. A. G. Revesz and R. J. Evans, *J. Phys. Chem. Solids* **30** p. 551 (1969).

64. E. A. Irene, "Evidence for a Parallel Path Oxidation Mechanism at the $Si\text{-}SiO_2$ Interface," *Appl. Phys. Lett.* **40** pp. 74-75 (1982).

65. W. G. Spitzer and J. R. Ligenza, "The Mechanisms for Silicon Oxidation in Steam and Oxygen," *J. Phys. Chem. Solids* **17** p. 196 (1961).

66. J. Ligenza, "Effect of Crystal Orientation on Oxidation Rates in High Pressure Steam," *J. Phys. Chem.* **65** p. 2011 (1961).

67. , "TCA Requirements for Thin High Integrity Oxides," J. C. Schumacher Company Product Application Note 14 (1985). Available on request

68. D. W. Hess and B. E. Deal, "Kinetics of Thermal Oxidation of Silicon in O_2/HCl Mixtures," *J. Electrochem. Soc.* **124**(5) pp. 735-739 (1977).

69. B. E. Deal, "Thermal Oxidation Kinetics of Silicon in Pyrogenic H_2O and 5% HCl/H_2O Mixtures," *J. Electrochem. Soc.* **125**(4) pp. 576-579 (1978).

70. B. E. Deal, D. W. Hess, J. D. Plummer, and C. P. Ho, "Kinetics of the Thermal Oxidation of Silicon in O_2/H_2O and O_2/Cl_2 Mixtures," *J. Electrochem. Soc.* **125**(2) pp. 339-346 (1978).

71. C. P. Ho, J. D. Plummer, S. E. Hanson, and R. W. Dutton, "VLSI Process Modelling - SUPREM III," *IEEE Trans. Electron Devices* **ED-30**(11) pp. 1438-1453 (1983).

72. I.G. McGillivray, J.M. Robertson, and A.J. Walton, "Electrical Effects of Chemical Purity on MOS Capacitor Electrical Parameters," *Electrochem. Soc. Extended Abstracts* **86-1**(201) pp. 288-289 (1986).

73. J. H. Matlock, "Material Defect Factors Affecting Device Performance," pp. 332-361 in *Silicon Processing ASTM STP 804*, ed. D. C. Grupta, American Society for Testing Materials (1983).

74. F. Schwettmann, K. L. Chang, and W. A. Brown, "Variation of Silicon Dioxide Growth Rate With Pre-Oxidation Clean," *Electrochem. Soc. Extended Abs.* **78-1**(276) pp. 668-669 (1978).

75. E. A. Irene, "Silicon Oxidation Studies: Some Aspects of the Initial Oxidation Regime," *J. Electrochem. Soc.* **123**(10) pp. 1708-1714 (1978).

76. H. Z. Massoud, J. D. Plummer, and E. A. Irene, "Thermal Oxidation of Silicon in Dry Oxygen: Growth-Rate Enhancement in the Thin Regime II. Physical Mechanisms," *J. Electrochem. Soc.* **132**(11) pp. 2693-2700 (1985).

77. H. Z. Massoud, J. D. Plummer, and E. A. Irene, "Thermal Oxidation of Silicon in Dry Oxygen: Growth-Rate Enhancement in the Thin Regime I. Experimental Results," *J. Electrochem. Soc.* **132**(11) pp. 2685-2693 (1985).

78. H. Z. Massoud, C. P. Ho, and J. D. Plummer, *Stanford University Technical Report TRDXG501-82*. July 1982.

79. S. A. Schafer and S. A. Lyon, "New Model of the Rapid Initial Oxidation of Silicon," *Appl. Phys. Lett.* **47**(2) pp. 154-156 (1985).

80. V. K. Samalam, "Theoretical Model for the Oxidation of Silicon," *Appl. Phys. Lett.* **47**(7) pp. 736-737 (1985).

81. S. Aptekar, M. Fernandes, and D. L. Kwong, "An Empirical Rate Law for Thin Oxides," *J. Electrochem. Soc.* **132**(6) pp. 1448-1450 (1985).

82. Stephan K. Lai, "Dependence Of Thin-Gate Oxide Properties On Processing," pp. 260-272 in *Silicon Processing ASTM STP 804*, ed. D. C. Grupta, American Society For Testing Materials (1983).

83. P. Donovan, "," pp. 175 in *Fundamentals of Silicon Integrated Technology (vol. 1)*, ed. R. M. Burger, (1967).

84. F. Richou, G. Pelius, and D. Lecrosnier, "Thermal Generation of Carriers in Gold Doped Silicon," *J. Appl. Phys.* **51**(12) pp. 6252-6257 (1980).

85. P. W. Morrison and R. J. Yevak, "Clean Rooms for VLSI Fabrication," *Semiconductor International* **8**(5) pp. 208-214 (May 1985).

86. J. E. Lawrence and H. R. Huff, "Silicon Material Properties for VLSI Circuitry," in *VLSI Electronics-Microstructure Science (Vol. 5)*, ed. N. G. Einspruch, Academic Press, London (1984).

87. W. Kern, "Purifying Si and SiO_2 Surfaces with Hydrogen Peroxide," *Semiconductor International* 7 pp. 94-99 (April, 1984).

88. W. Kern and D. A. Poutinen, "Cleaning Solution Based on Hydrogen Peroxide For Use in Silicon Semiconductor Technology," *RCA Review*, pp. 187-206 (June 1970).

89. W. Kern, "Radiochemical Study of Semiconductor Surface Contamination II. Deposition of Trace Impurities on Silicon and Silica," *RCA Review*, pp. 235-264 (June 1970).

90. W. Kern, "Radiochemical Study of Semiconductor Surface Contamination I. Adsorption of Reagent Compounds," *RCA Review*, pp. 207-233 (June 1970).

91. D. Burkmann, "Optimizing the Cleaning Procedure for Silicon Wafers Prior To High Temperature Operations," *Semiconductor International*, pp. 103-116 (July 1981). also available as FSI corporation TR 117

92. P. S. Burggraaf, "Wafer Cleaning: State of the Art Chemical Technology," *Semiconductor International* 4(7) pp. 71-100 (July 1981).

93. W. Thurston and R. Seaman, and A. D. Weiss, "Wafer Cleaning Update," *Semiconductor International* 7(4) pp. 82-85 (April 1985).

94. J. M. Duffalo and J. R. Monkowski, "Particulate Contamination and Device Performance," *Solid St. Tech.* **27** (March 1984).

95. H. Muraoka, K. Kurosawa, H. Hiratsuka, and T. Usami, "Cleaning Solutions Based on Choline," *Electrochem. Soc. Extended Abs.*, (238) pp. 570-572 (1981).

96. D. S. Becker, W. R. Schmidt, C. A. Peterson, and D. C. Burkmann, "Effects of Various Chemistries on Silicon Wafer Cleaning," FSI TR 251 (22 June 1984).

97. D. S. Becker, "Atomic Absorption Study of an FSI B Clean's Ability to Remove Iron From a Silicon Wafer Surface," FSI TR 241 (16 April 1984).

98. D. C. Burkmann, W. R. Schmidt, C. A. Paterson, and B. R. Phillips, "Removing Residual Metals and Particulates," Proceedings of the Technical Programme Semiconductor International 1983, Birmingham, UK (September 1983).

99. K. D. Beyer and R. H. Kastl, "Impact of Deionized Water Rinses on Silicon Surface Cleaning," *J. Electrochem. Soc.* **129**(5) pp. 1027-1029 ().

100. A. Hawrily, Private Communication

101. I. G. McGillivray, J. M. Robertson, and A. J. Walton, "Effects of Trace Impurities on MOS Capacitor Electrical Parameters," pp. 999-1010 in *Semiconductor Silicon 1986*, ed. B. Kolbesen, Electrochemical Society, Pennington NJ (1986).

102. G. Failes and A. Hamilton, Private Communications

103. I. G. McGillivray, "Electrical Assessment of Wafer and Furnace Cleaning," EMF/Hughes Technical Report (October 1985).

104. R. Iscoff, "Water Purification Criteria for Semiconductor Manufacturing," *Semiconductor International* **8**(3) pp. 66-71 (March 1985).

105. P. F. Schmidt and C. W. Pearce, "A Neutron Activation Analysis Study of the Sources of Transition Group Metal Contamination in the Silicon Manufacturing Process," *J. Electrochem. Soc.* **128**(3) pp. 630-637 (1981).

106. M. Parikh and U. Kaempf, "SMIF: A Technology for Wafer Cassette Transfer in VLSI Manufacturing," *Solid St. Tech.* **27** pp. 111-115 (July 1984).

107. B. D. Foster and R. E. Tressler, "Silicon Processing with Silicon Carbide Furnace Components," *Solid St. Tech.*, pp. 143-146 (October 1984).

108. A. Stolmeyer, Private Communication

109. M. W. Hillen, "Charge Trapping Near the Si/SiO_2 Interface In Semiconductor Devices," *PhD. Thesis*, University of Groningen, Netherlands, (1981).

110. K. H. Zaininger and F. P. Heiman, "The CV Technique as an Analytical Tool-Part I," *Solid St. Tech.*, pp. 49-56 (May 1970).

111. K. H. Zaininger and F. P. Heiman, "The CV Technique as an Analytical Tool-Part II," *Solid St. Tech.*, pp. 46-55 (June 1970).

112. F. J. Fiegel, "Characterisation of Dielectric Films," pp. 146-215 in *VLSI Electronics: Microstructure Science (vol. 6)*, ed. N. G. Einspruch, Academic Press, New York (1981).

113. P. S. Burggraaf, "CV Plotting, CT Measuring And Dopant Profiling: Applications And Equipment," *Semiconductor International*, pp. 29-42 (October 1980).

114. A. S. Grove, *Physics and Technology of Semiconductor Devices*, Wiley, New York (1967).

115. C Kittel, *Introduction to Solid State Physics, 5th Edition*, Wiley, New York (1976).

116. J. L. Moll, *Inst. Radio Engrs. Wescon Convention Record* **3** p. 32 (1959).

117. W. G. Pfann and C. G. Garrat, "Semiconductor Varactors Using Surface Space Charge Layers," *Proc. Inst. Radio Engrs.* **47** pp. 2011-2012 (1959).

118. D. R. Frankl, "Some Effects of Material Parameters on the Design of Surface Space Charge Varactors," *Solid St. Electron.* **2** pp. 71-76 (1961).

119. R. Linder , "Semiconductor Surface Varactor," *Bell Syst. Tech.* **41** pp. 803-831 (1962).

120. L. M. Terman, "An Investigation of Surface States at a Silicon/Silicon Dioxide Interface Employing Metal-Oxide-Silicon Diodes," *Solid St. Electron.* **5** p. 285 (1962).

121. J. R. Brews, "A Simplified High-Frequency MOS Capacitance Formula," *Solid St. Electron.* **20** pp. 607-608 (1977).

122. B. E. Deal, "Standardised Terminology For Oxide Charges Associated With Thermally Oxidised Silicon," *J. Electrochem. Soc.* **127**(4) pp. 979-981 (1980). also IEEE-ED, ED-27,606-608(1980)

123. *Federal Standard 209B*, Clean Room Specification (USA)

124. *British Standard, BS 5295 (Class 1)*, Clean Room Specification (UK)

125. I. G. McGillivray, A. J. Walton, and J. M. Robertson, "Post-Metalisation Annealing of Aluminium Silicon ," *Electron. Lett.* **21** pp. 973-974 (1985).

126. B. E. Deal, E. L. Makenna, and P. L. Castro, "Characteristics of Fast Surface States Associated with SiO_2-Si and Si_3N_{44}-SiO_2-Si Structures," *J. Electrochem. Soc.* **116** p. 917 (1969).

127. K. H. Zaininger and G. Warfield, "Limitations Of The MOS Capacitance Method For The Determination Of Semiconductor Surface Properties," *IEEE. Trans. Electron. Devices* **ED** pp. 179-193 (1965).

128. A. Lederman, "Vacuum Operated Mercury Probe for CV Plotting and Profiling," *Solid St. Tech.*, p. 121 (August 1981).

129. D. L. Rehrig and C. W. Pearce, "Production Mercury Probe Capacitance-Voltage Testing," *Semiconductor International*, p. 151 (May 1980).

130. S. Kar, "Determination Of Si-Metal Work Function Differences By MOS Capacitance Technique," *Solid St. Electron.* **18** pp. 169-181 (1975).

131. *Equipment Manufacturers*, All types of equipment are regularly advertised in journals such as Semiconductor International, Solid State Technology, Microelectronics Manufacturing and Test. Semiconductor International publishes an annual list of suppliers each December. All three of the journals mentioned here are available free to those with 'purchasing power'.

132. K. H. Zaininger, "Automatic Display of MIS Capacitance Versus Bias Characteristics," *RCA Review*, pp. 341-359 (September, 1966).

133. A. Goetzberger, "Ideal MOS Curves for Silicon," *Bell Syst. Tech. J.* **45** pp. 1097-1122 (1966).

134. I. G. McGillivray, "A Computer Controlled CV Measurement and Analysis System," EMF/Hughes Report (October 1984).

135. E. H. Snow, A. S. Grove, B. E. Deal, and C. T. Sah, "Ion Transport Phenomena in Insulating Films," *J. Appl. Phys.* **36** p. 1664 (1965).

136. E. H. Nicollian and A. Goetzberger, "The Si-SiO_2 Interface Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique," *Bell Syst. Tech. J.* **46** pp. 1055-1133 (1967).

137. A. J. Walton, *Process Monitoring and Control*, To be published

138. J. Grosvalet and C. Jund, "Influence of Illumination on MIS Capacitances in the Strong Inversion Region," *IEEE Trans. Electron Devices* **ED-14**(11) pp. 777-780 (1967).

139. E. H. Nicollian and A. Goetzberger, "Lateral AC Current Flow Model for Metal-Insulator-Semiconductor Capacitors," *IEEE Trans. Electron Devices* **12** pp. 108-117 (1965).

140. A. H. Marashak, "On The Threshold And Flatband Voltages For MOS Structures With Polysilicon Gate and Non-Uniformly Doped Substrate," *Solid St. Electron.* **26**(4) pp. 361-364 (1983).

141. B. Deal, M. Sklar, A. S. Grove, and E. H. Snow, "Characteristics of Surface State Charge (Qss) of Thermally Oxidised Silicon," *J. Electrochem Soc.* **114**(3) pp. 266-274 (1967).

142. E. A. Arnold, J. Ladell, and G. Abowitz, "Crystalographic Symmetry of Surface State Density in Thermally Oxidised Silicon," *Appl. Phys. Lett.* **13** pp. 413-416 (1968).

143. G. Baccarani, S. Solmi, and G. Soncini, *Alta Frequenza* **40** p. 301E (1971).

144. M. Hamsaki, "Radiation Effects On Thin-Oxide MOS Capacitors Caused By E-beam Evaporation Of Aluminium," *Solid St. Electron.* **26**(4) pp. 299-303 (1983).

145. C. T. Sah, "Characteristics of Metal-Oxide-Semiconductor Transistors," *IEEE Trans. Electron. Devices* **ED-11** p. 324 (1964).

146. R. C. Jaeger, F. H. Gaensslen, and S. E. Diehl, "An Efficient Numerical Algorithm For The Simulation Of MOS Capacitance," *IEEE Trans. CAD of Integrated Circuits And Systems* **CAD-2**(2) pp. 111-116 (1983).

147. A. G. Buttar, *CMOS Process Simulation*, PhD Thesis (Edinburgh University) 1986.

148. R. O. Deming and W. A. Keenan, "Low Dose Ion Implant Monitoring," *Solid State Tech.* **28**(9) pp. 163-167 (September 1985).

149. I. G. McGillivray, "A Comparison of the Quality of Oxide and Oxide-Nitride Gate Dielectrics," EMF/Hughes Report (October 1984).

150. M. Kuhn and D. J. Silversmith, "Ionic Contamination and Transport of Mobile Ions in MOS Structures," *J. Electrochem. Soc.* **118** pp. 966-970 (1971).

151. N. J. Chou, "Application of Triangular Voltage Sweep Method to Mobile Charge Studies in MOS Structures," *J. Electrochem. Soc.* **118** p. 601 (1971).

152. M. Yamin, "Charge Storage Effects in Silicon Dioxide Films," *IEEE Trans. Electron. Devices* **ED-12** pp. 88-96 (1965).

153. H. R. Huff and F. Shimura, "Silicon Material Criteria for VLSI Electronics," *Solid State Tech.* **28** pp. 103-118 (March 1985).

154. D. R. Wolters and J. J. van der Schoot, "Dielectric Breakdown in MOS Devices Part I. Defect Related and Intrinsic Breakdown," *Phillips J. Res.* **40** pp. 115-136 (1985).

155. H. R. Huff, "Minority Carrier Lifetime: Correlation with IC Process Parameters," *J. Electrochem. Soc.* **126** pp. 1142-1147 (1979).

156. A. Goetzberger , "Behavior of MOS Inversion Layers at Low Temperatures," *IEEE Trans. Electron Devices* **14** pp. 787-789 (1967).

157. E. Dijkstra, "The Assimilation of Quasi-Static CV Measurements," Technical Report, Twente Uninversity, Netherlands (May 1983).

158. C. N. Berglund, "Surface States at Steam-Grown Silicon-Silicon Dioxide Interfaces," *IEEE Trans. Electron Devices* **ED-13** pp. 701-705 (1966).

159. E. H. Nicollian and A. Goetzberger, "MOS Conductance Technique for Measuring Surface State Parameters," *Appl. Phys. Lett.* **8** pp. 216-219 (1965).

160. M. Kuhn, "A Quasi-Static Technique For The MOS CV and Surface State Measurements," *Solid. St. Electron.* **13** pp. 873-885 (1970).

161. K. Ziegler and E. Klausmann, "Static Technique For The Precise Measurement Of Surface Potential And Interface State Density In MOS Structures," *Appl. Phy. Lett.* **26**(7) pp. 400-402 (1975).

162. R. Castange' and A. Vallipe, *Surf. Sci.* **28** p. 557 (1971).

163. J. J. Simonne, "A Method to Extract Interface State Parameters from the MIS Parallel Conductance Technique," *Solid State Electron* **16** pp. 121-124 (1973).

164. E. H. Nicollian, A. Goetzberger, and A. D. Lopez, "Expedient Method of Obtaining Interface State Properties from MIS Conductance Methods," *Solid St. Electron.* **12** pp. 937-944 (1969).

165. A. Goetzberger, E. Klausmann, and M. J. Shulz, "Interface States on Semiconductor/Insulator Surfaces," *CRC Critical Reviews in Solid State Science*, pp. 1-43 (January 1976).

166. J. R. Monkowski, M. D. Monkowski, I. S. T. Tsong, and J. Stach, "Hydrogen/Chlorine Distributions in Silicon Dioxide: Detection and Model," pp. 245-259 in *Silicon Processing ASTM STP 804*, ed. D. C. Grupta, American Society for Testing Materials (1983).

167. R. B. Marcus, "Diagnostic Techniques," pp. 507-550 in *VLSI Technology*, ed. S. M. Sze, McGraw-Hill, Singapore (1983).

168. W. V. Gelder and E. H. Nicollian, "Silicon Impurity Distribution as Revealed by Pulsed MOS-CV Measurements," *J. Electrochem. Soc.* **118**(1) pp. 138-141 (1971).

169. C. O. Thomas, D. Kahang, and R. C. Manz, "Impurity Distribution in Epitaxial Silicon Films," *J. Electrochem. Soc.* **109**(11) pp. 1055-1061 (1962).

170. J. Hillibrand and R. D. Gold, "Determination of Impurity Distribution in Junction Diodes from Capacitance-Voltage Measurements," *RCA Review*, pp. 245-252 (1960).

171. E. Kreysig, *Advanced Engineering Mathematics (4th edition)*, Wiley, New York (1979).

172. R. A. Moline, "Ion-Implanted Phosphorous in Silicon: Profiles Using CV Analysis," *J. Appl. Phys.* **42** pp. 3553-3558 (1972).

173. I. Amron, "Errors in Dopant Concentration Profiles Determined by Differential Capacitance Measurements," *Electrochem. Tech.* **5** pp. 94-97 (1967).

174. P. Blood, "Capacitance Voltage Profiling and the Characterisation of III-V Semiconductors Using Electrolyte Barriers," *Semiconductor Science and Technology* **1**(1)(1986).

175. D. P. Kennedy and R. R. Obrien, "On the Measurement of Impurity Atoms by the Differential Capacitance Technique," *IBM J. Res.* **13** p. 212 (1969).

176. D. J. Bartelink, "Surface Depletion and Inversion in Semiconductors with Arbitrary Doping Profiles," *Appl. Phys. Lett.* **37** pp. 220-223 (1980).

177. D. J. Bartelink, "Limits of Applicability of the Depletion Approximation and its Recent Augmentation," *Appl. Phys. Lett.* **38** pp. 461-463 (1981).

178. D. J. Bartelink and R. Tremain, "Exact Formulation of MOS CV Profiling," *IEEE Trans. Electron Devices* **ED-27**(11) p. 1831 (1979). Abstract MP-B4

179. C. L. Wilson, "Correction of Differential Capacitance Profiles for Debeye Length Effects," *IEEE Trans. Electron Devices* **ED-27**(12) p. 2262 (1980).

180. K. Ziegler, E. Klausmann, and S. Kar , "Determination Of The Semiconductor Doping Profile Right Up To Its Surface Using The MIS Capacitor," *Solid St. Electron.* **18** pp. 189-198 (1975).

181. J. R. Brews, "Correcting Interface-state Errors in MOS Doping Profile Determinations," *J. Appl. Phys.* **44**(7) pp. 3228-3231 (1973).

182. S. Lin and J. Reuter, "The Complete Doping Profile Using MOS CV Technique," *Solid St. Electron.* **26**(4) pp. 343-351 (1983).

183. W. A. Keenan, W. H. Johnson, and A. K. Smith, "Advances in Sheet Resistance Measurements for Ion Implant Monitoring," *Solid State Technology* **28**(6) pp. 143-148 (June 1985).

184. B. J. Gordon, "On Line Capacitance Voltage Doping Profile Measurement of Low Dose Implants," *IEEE Trans. Electron Devices* **ED-27**(12) pp. 2268-2272 (1980).

185. J. R. Golin and J. A. Glaze, "High Resolution Dose Uniformity Monitoring of Ion Implanters," *Solid State Tech.* **27**(9) pp. 289-295 (September 1984).

186. G. A. Gruber, "Ion Implant Testing for Production Control," *Solid State Tech.* **26**(8) pp. 159-167 (August 1983).

187. D. S. Perloff, J. N. Gan, and F. E. Wahl, "Dose Accuracy and Doping Uniformity of Ion Implantation Equipment," *Solid State Tech.* **24** p. 112 (February 1981).

188. R. K. Berglund, "Practical Approach to Monitoring of Implanted Layers," *Semiconductor International* 7(5) pp. 155-159 (May 1985).

189. B. Martin (Phillips, Redhill England), Private Communication

190. D. Aitken, "The Design Philosophy for a 200 kV High Current Ion Implanter," *Nuclear Instr. Methods* **139** pp. 125-134 (1976).

191. B. Ross, "Survey of Literature on Minority Carrier Lifetimes in Silicon and Related Topics," pp. 14-28 in *Lifetime Factors in Silicon, ASTM STP 712*, American Society for Testing Materials (1980).

192. T. Mamine, H. Hayashi, T. Matsushita, T. Yanada, O. Kumagai, K. Nishiyama, and K. Kaneko, "Control of Lifetime in Silicon by Implantation of Iron," pp. 119-135 in *Lifetime Factors in Silicon, ASTM STP 712*, American Society for Testing Materials (1980).

193. D. K. Schroder, "Effective Lifetimes in High Quality Silicon Devices," *Solid St. Electron.* **27**(3) pp. 247-251 (1984).

194. D. K. Schroder, J. D. Whitfield, and C. J. Varker, "Recombination Lifetime Using The Pulsed MIS Capacitor," *IEEE Trans. Electron Devices* **ED-31**(4) pp. 462-467 (1984).

195. D. K. Schroder, "The Concept of Generation and Recombination in Semiconductors," *IEEE Trans. Electron Devices* **ED-29** pp. 1336-1338 (1982).

196. F. P. Heiman, "On the Determination of Minority Carrier Lifetime from the Transient Response of an MOS Capacitor," *IEEE Trans. Electron Devices* **14** pp. 781-784 (1967).

197. J. S. T. Huang, "Bulk Lifetime Using an MOS Capacitor," *Proc. IEEE* **58** pp. 1849-1850 (1970).

198. Y. Kano and A. Shibata, "On The Determination of Minority Carrier Lifetime and Surface Recombination Velocity from the Transient Response of MOS Capacitors," *Jap. J. Appl. Phys.* **11** pp. 1161-1164 (1972).

199. A. P. Gorban, V. G. Litovchenko, and D. N. Moskal, "Investigation of the Generation Characteristics of Metal-Insulator-Semiconductor Structures," *Solid State Electronics* **18** p. 1053 (1975).

200. G. Kaplan, "Determination of Minority Carrier Lifetime and Surface Recombination Velocity by Pulsed Hysteresis CV Method," *Solid State Electronics* **23** pp. 513-514 (1980).

201. K. S. Rabbani and D. R. Lamb, "A Quick Method for the Determination of Bulk Generation Lifetime in Semiconductors from Pulsed MOS Capacitor Measurements," *Solid State Electronics* **24** pp. 661-664 (1981).

202. S. T. Lin, "Simultaneous Determination of Minority Carrier Lifetime and Deep Doping Profile Using a Double Sweep MOS-C Technique," *IEEE Trans. Electron Devices* **30** pp. 60-63 (1983).

203. P. Tiwari, B. Bhaumik, and J. Vasi, "Rapid Measurement of Lifetime Using a Ramped MOS Capacitor Transient," *Solid State Electronics* **26** pp. 695-698 (1983).

204. S. R. Hofstein, "Minority Carrier Lifetime Determination from Inversion Layer Transient Response," *IEEE Trans. Electron Devices* **ED-14** pp. 785-786 (1967).

205. C. R. Viswanathan and T. Takino, "Minority Carrier Generation Time and Surface Recombination Velocity from Q-t Measurements," *IEEE Trans. Electron Devices* **ED-25** pp. 817-821 (1978).

206. P. U. Calzolari, S. Graffi, A. M. Mazzone, and C. Moranda, *Alta Frequenza* **41** p. 848 (1972).

207. C. Trullemans and F. van der Wiele, "Bulk Lifetime Determination by a Voltage Step Response Method," *Solid State Electronics* **21** pp. 561-564 (1978).

208. P. Kuper and C. A. Grimbergen, "Determination of Generation Lifetime from Non-Equilibrium Linear Sweep Current and Capacitance Measurements on an MOS Capacitor," *Solid State Electronics* **21** pp. 549-553 (1978).

209. R. F. Pierret, "A Linear Sweep MOS-C Technique for Determining Minority Carrier Lifetimes," *IEEE Trans. Electron Devices* **ED-19** pp. 869-873 (1972).

210. R. F. Pierret and D. W. Small, "A Modified Linear Sweep Technique for MOS-C Generation Rate Measurements," *IEEE Trans. Electron Devices* **ED-22** pp. 1051-1052 (1975).

211. D. W. Small and R. F. Pierret, "Separation of Surface and Bulk Components in MOS-C Generation Rate Measurements," *Solid St. Electron.* **19** pp. 505-511 (1976).

212. K. Taniguchi, "Graphical Technique to Determine Minority Carrier Lifetime and Surface Generation Velocity Using a Triangular Voltage Sweep CV Method," *Solid State Electronics* **21** pp. 1057-1061 (1978).

213. M. Miyake and H. Harada, "A New Method for Non-Uniformly Doped MOS Capacitor Ct Characteristics in Lifetime Evaluation," *Jap. J. Appl. Phys.* **20** pp. L797-800 (1981).

214. *Computerised Semiconductor Measurement System*, Materials Development Corporation, Chatsworth CA

215. C. S. Yue, H. Vyas, M. Holt, and J. Borowick, "A Fast Extrapolation Technique for Measuring Minority Carrier Generation Lifetime," *Solid State Electronics* **28** pp. 403-406 (1985).

216. *HP4062B Semiconductor Parametric Test System*, System Library vol. 2

217. D. R. Wolters and J. J. van der Schoot, "Dielectric Breakdown in MOS Devices Part II. Conditions for the Intrinsic Breakdown," *Phillips J. Res.* **40** pp. 137-163 (1985).

218. D. R. Wolters and J. J. van der Schoot, "Dielectric Breakdown in MOS Devices Part III. The Damage Leading to Breakdown," *Phillips J. Res.* **40** pp. 164-192 (1985).

219. W. Sievert, Private Communication

220. W. R. Schmidt, "Evaluation of FSI Clean, Aqua Regia Clean, Fuming Nitric Acid Clean and Piranha+HF Clean Using SIMS," FSI TR 173 (21 January 1982). Update #2 (13/5/82)

221. K. M. Eisele and E. Klausmann, "Effects of Heavy Metal Contamination from Corrosive Gas and Dopant Handling Equipment in Silicon Wafer Processing," *Solid St. Technology* **27** pp. 177-180 (1984).