OPTIMISATION STRATEGY FOR SEMICONDUCTOR TECHNOLOGY DEVELOPMENT IN A VOLUME MANUFACTURING ENVIRONMENT

Mark Redford

Doctor of Philosophy in Electrical Engineering
The University Of Edinburgh
2000

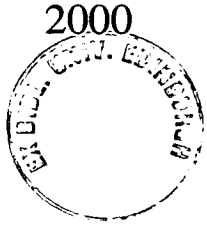


Table of Contents

Abstract vii

Signed Declaration viii

Acknowledgements ix

Dedication xi

Chapter 1: Introduction 1

1.1 Background 1

1.2 Thesis Outline 5

1.2.1 Thesis Motivation 5

1.2.2 Thesis Content 6

References 9

Chapter 2: Technology Computer Aided Design (TCAD) and Design of Experiments (DOE) 14

2.1 Introduction 14

2.2 Technology CAD 15

2.2.1 Circuit Simulation 15

2.2.2 Process and Device Simulation 18

2.3 Design of Experiments (DOE) 23

2.4 Design for Manufacturability (DFM) Strategy 24

2.4.1 Introduction 24

2.4.2 Automation Software 25

2.5 Summary 29

References 30

Chapter 3: Device Physics of a pn Junction 36

3.1. Introduction 36

3.2 The pn Junction Depletion Region 36

3.2.1 The Maximum Electric Field in a PN Junction 37

3.2.2 The P-N+ Junction 42

3.2.3 Forward and Reverse Bias 43

3.3 Avalanche Breakdown 44

3.3.1 Introduction 44

3.3.2 Theory of Avalanche Breakdown 45

3.4 Influence of Field Plates on Breakdown Voltage 46

3.5 Breakdown Drift 51

3.6 Reduced SURface Field (RESURF) Principle 52

3.7 Summary 58

References 59

Chapter 4: The JHVH Fabrication Process 61

4.1 Introduction 61

4.2 Semiconductor Processing 62

4.2.1 Photolithography 62

4.2.2 Implantation 63

4.2.3 Diffusion 66

4.2.4 Deposition 67

4.3 The JHVH Process 68

4.4 Process Challenges 78

4.5 Summary 79

References 80

Chapter 5: JHVH-450 Process Development 83

5.1 Introduction 83

5.2 Strategy 83

5.3 Process Options 85

5.4 Review of Process Simulation Deck 85

5.4.1 Process Simulator 85

5.4.2 Process Information 87

5.4.3 Comparison of Simulated and Silicon Results 89

5.5 Review of Device Simulation Deck 92

5.5.1 Device Simulator 92

5.5.2 Medici Input Deck 93

5.6 Process Optimisation 104

5.6.1 Introduction 104

5.6.2 Optimisation Methodology 104

5.6.3 Epitaxy Only Process 105

5.6.4 Combined Epitaxy and Implant Process 111

5.7 Silicon split lot results 118

5.7.1 Epitaxy Process 118

5.7.2 Epitaxy and Epitaxy/Implant split lot 120

5.8 Process Recommendation 122

5.9 Interaction of Post Fabrication Processing on Breakdown Voltage Drift in High Voltage RESURF Devices 122

5.9.1 Introduction 122

5.9.2 Problem Definition 123

5.9.3 Experimental Procedure 125

5.10 Summary 130

References 132

Chapter 6: JHVH-700 Process Development 135

6.1 Introduction 135

6.2 JHVH-700 Process 136

6.3 Covariance Based Interpolation 139

6.3.1. Introduction 139

6.3.2 Covariance 139

6.4 Process Optimization 141

6.4.1 Introduction 141

6.4.2 Optimisation Methodology 141

6.4.3 Combined Epitaxy and Implant Process 142

6.4.4 Breakdown Distributions 148

6.4.5 Experimental Results From Silicon 150

6.5 Observations 151

6.6 Application of Covariance Fits to Experimental Data 153

6.6.1 Introduction 153

6.6.2 Background 153

6.6.3 Improving the Fit to Experimental Data 154

6.8 Summary 160

References 162

Chapter 7: Conclusions and Further Work 165

7.1 Conclusions 165

7.2 Future Work 167

References 170

Appendices 171

Appendix 1 .....SUPREM-3 Program: Epitaxy Process 172

Appendix 2 .....SUPREM-3 Program: Epitaxy & Implant Process 184

Appendix 3 .....Medici Program 196

Appendix 4 .....Papers Associated with this Thesis 200

Abstract

The time to market (TTM) for integrated circuits is key to the success for any technology and the products manufactured using it. Today, it is no longer considered economical to continue to debug a technology once it has been released to manufacturing as market share and the development return on investment, are both impacted.

This thesis addresses this by proposing a strategy to optimise a given technology. It uses a combination of Design Of Experiments (DOE), Response Surface Methodology (RSM) and Technology Computer Aided Design (TCAD). The breakdown voltage of an LDMOS technology is optimised, which is well known to be a challenge to model using design experiments and provides a stringent test of the strategy. Covariance models are employed to improve the model fits to the TCAD data with the justification being that the TCAD results have no random error associated with them.

This strategy was successfully applied to three technologies where results were comparable to measured data from one silicon iteration and saved at least six months in development, as well as to several other projects within the Analog Process Technology Development group at National Semiconductor, Greenock, Scotland.

Signed Declaration

I declare that this thesis has been solely composed by myself and that this work is my contribution as a member of the Analog Process Technology Development group of National Semiconductor Corporation.

Department of Electrical Engineering, The University of Edinburgh

26 February, 2000

Acknowledgements

I would like to thank Professor A.J. Walton for his encouragement and enthusiasm throughout my PhD. Thanks to Dr. Ian Underwood for his valuable inputs and comments. Thanks also to my employer, National Semiconductor (UK) Ltd in Greenock, Scotland for sponsoring my PhD.

Also, thanks to Dr. Martin Fallon who acted as my industrial supervisor and gave me good advice throughout and particularly when writing up my thesis. He enjoyed "winding me up" too. Though I got him with some "counter wind ups" !!

I would also like to thank my staff of the Analog Process Technology (APTD) group in Greenock, Scotland. In particular, and not in any order, except for, perhaps the amount of alcohol that they are able to consume: Dr. Martin "Silky" Fallon, Dr. Z.Alan "Logic" Shafi, Dr. Alastair "Shout" McGibbon, Frank "Biffa" Boyland, Dave "True Blue" McAlpine. Jim "Ah @#%&" McGinty, Rikki "Rocky" Boyle, Ian "Thompy" Thomson, Derek "Little Buddha" Corcoran, Jeremy "Man In Black" Smith, Dr. Yakub Aliyu and J. Ian "Batman" Doohan for being able to keep everyone entertained at other people's expense. APTD were a great bunch of guys. We achieved a great deal. Such a shame it had to change.

Thanks go to my bosses at the initiation of my PhD: Bruce Diamond, my UK boss, who supported me and taught me much. For that I thank you. To Dr. Reda R. Razouk, my US boss, from whom I learnt a great deal. I thank you.

Thanks too to the MEng students – Neil Rankin, Chris Roxburgh, Chun Hsuing Ng, Keith Findlater, Eric Volk, Kristen Waite, Fraser Brodie, Stuart Leathley, and Paul McKernan. I'm glad I sponsored you. You did well !

I would like to acknowledge and thank Dr. Zheng Jia Zhen for his valuable inputs on the technology details of Chartered Semiconductor's technologies. I owe you some silicon though it has to be a designed experiment!

Lastly, but by no means least, to my wife Carol, who put up with so much whilst I wrote up my thesis. I love you. Finally to my kids Mark and Khym who'll be pleased that's its finally over. I can now spend Sunday with them!

Dedication

I dedicate this PhD thesis to my late father, John Smith Redford, 1925-1998, without whose encouragement, I probably would not have achieved half as much as I have done academically and professionally.

Thanks Dad !!

Chapter 1: Introduction

1.1 Background

The majority of the development of integrated circuits has been focussed upon digital applications. Within this area, the technology has steadily developed to operate at lower voltages and current levels. The continual push to increase transistor density is the driver that has facilitated the reduction in the operating voltage to levels of 3.3V and below [1]. As the transistors have been designed to maximise the number per unit area and thereby increase circuit functionality, the current handling capability of the devices is typically restricted to less than a milliampere.

During the 1980s, it became apparent that systems costs were dominated [2] by the power section of the system. This high cost was primarily due to the power bipolar-based technologies developed during the 1970s. These high voltage power transistors had to be fabricated with a relatively thick epitaxial layer in order to support the high operating voltages and this limited their current gain to less than 10, even when a Darlington topology was utilised. The power bipolar transistor is a current controlled device that requires substantial input current to both maintain it in its on state, and also to switch it quickly from on to off. Bipolar transistor based power circuits are, therefore, composed of numerous stages of discrete devices resulting in a high system weight, size and cost. These challenges were addressed with a new power technology that evolved in the 1980s based upon MOS devices used for CMOS integrated circuits. New types of MOS controlled power devices emerged which were capable of withstanding higher voltages and currents. Among these devices are the power MOSFET and MOS insulated gate

transistor [3] which were of interest because of their high input impedance. The low input gate currents required to control these devices enables components used in the gate drive circuits to be small enough to be integrated. This created an opportunity to significantly lower the number of components, resulting in a reduction in system cost, size and weight. In addition the number of interconnects are also reduced which in turn enhances system reliability.

The applications for power semiconductor technology are distributed over a very large range of power level and the range of applications is shown in figure 1.1.

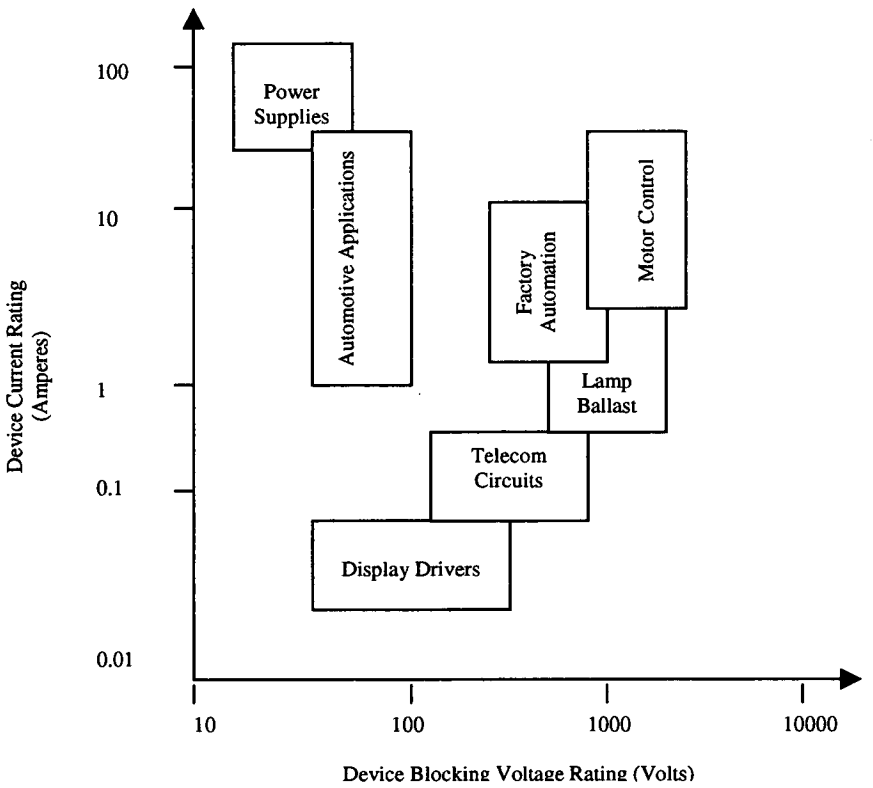


Figure 1.1 Power semiconductor technology applications

Figure 1.1 highlights the fact that no single technology can address all or even portions of application areas. In fact, even within a given application, it is often difficult to define the optimum technology.

Power integrated circuits can be split into two broad categories: (a) "Smart" [4-10] discrete vertical power structures with sense or protection elements and logic or analog circuits built on the same integrated circuit; and (b) High voltage integrated circuits [11-21] in which conventional logic and analog circuits are combined with high voltage lateral devices. The smart power process is mainly targeted to automotive electronics; a large cost and productivity improvement can be obtained by reducing wiring harness in cars and hence power control systems become attractive. The implementation of such systems requires the availability of low cost smart switches that can be distributed throughout the car to run electrical components such as motors. High voltage integrated circuits are typically targeted to applications such as motor control, telecommunications, power supplies, and displays. The power spectrum over which this power technology can be applied ranges from 50 to 1200V with currents of 1 to 100 amps.

One of the challenges in fabricating high voltage integrated circuits is that thick epitaxial layers must be used in order to achieve the desired breakdown voltage. However, the resulting fabrication process requires some very long diffusion cycles [22,23] of the order of 24+ hours. This results in poor performance of low voltage transistors [22] as well as the consumption of large areas of the integrated circuit by the deep diffusion distances. The invention of the RESURF transistor [24,25] enabled high voltage transistors to be

realised on relatively thin epitaxial layers. Several workers [26-30] have used this approach to achieve the desired transistor breakdown of their respective technologies. The use of advanced processing techniques [31] have further optimised the device from a manufacturability perspective. The advantage of this RESURF technology is that the breakdown voltage can be easily scaled [24,25,32] without having to make significant process changes. Further developments are ongoing particularly in the isolation area with the objective of increasing the packing density whilst maintaining the high voltage capability. Several approaches are being investigated [33-37]

All electronic circuits require power supplies and over the last 10 years significant advances in their design has taken place. The most important of these has been the change from linear power supplies to those that operate on a switching basis [38] - switch mode power supplies (SMPS). The main reason for the move for SMPS is their greater efficiency, which is typically 80-90% in comparison to 30-40% for linear power supplies. This greatly reduces the cooling requirements, and allows a much higher power density. The adoption of SMPS technology has been facilitated by the advent of the power MOSFET together with the availability of high-speed diodes and improved magnetic materials. The worldwide market for SMPS is estimated at $5B, of which 80% is being designed in-house.

1.2 Thesis Outline

1.2.1 Thesis Motivation

The time to market for new products is absolutely vital in today's competitive market place. Now it is no longer competitive to continue to debug a technology when its has transferred from development to manufacturing. This has a direct impact on the Return Of Investment (ROI) for the technology development as well as command of market share plus the perceived public relations with customers and investors. In order to address the time to market constraint a methodology is required to facilitate a timely development of a new technology whilst at the same time give a sense of confidence that the technology itself will yield. This requires a Design For Manufacturability (DFM) strategy where Technology CAD (TCAD) can be used with Design Of Experiments (DOE) to find the optimum process conditions which will yield a technology that meets specifications in a manufacturing environment.

The topic of this research is to devise realistic strategies within a manufacturing environment, which will determine an optimum process for a technology using response surface methodology (RSM). This includes assessing the manufacturability of the technology by predicting the parametric distributions that would be seen during manufacturing, and then comparing the results with those obtained from silicon. In some designed experiments there is potential for 'outliers' which results in a poor response surface (RSM) model. As TCAD has no experimental error, using a covariance based interpolation, which perfectly fits response surfaces to data can be justified and its practical application investigated.

Such an approach would lead to the creation of a methodology for process optimisation and manufacturability assessment for any given technology.

The technology chosen to demonstrate the proposed optimisation methodology was National Semiconductor Corporation's Switch Mode Power Supply (SMPS) development. National's goal was to provide a complete power supply solution to their customers. This was to be achieved through providing a solid state solution plus the necessary software to facilitate the design of the power supply without having to be an expert SMPS designer. This was to take out the need for any in-house design and their targeted market was forecast to be worth $100M.

1.2.2 Thesis Content

Chapter 2 reviews Technology CAD (TCAD) and Design of Experiments (DOE) and the software packages available today both from academia and commercial vendors. The interface and automation of TCAD and DOEs which reduces the possibility of errors is discussed. This is followed by a discussion on the software used to predict distributions from the resultant response surfaces

Chapter 3 discusses the basic device physics, focusing on junction breakdown and in particular the avalanche breakdown mechanism. The influence of a field plate to reduce the effects of junction curvature is discussed, highlighting how drift in breakdown can be observed when high electric fields are present in pn junctions utilising field plates. Finally, the RESURF principle of realising high voltages in relatively thin epitaxial layers

is reviewed along with the operation of both the field plate and the RESURF principle through the use of device simulation.

Chapter 4 reviews many of the processing steps used in the manufacture of semiconductor devices. This provides the background information for subsequent chapters. The JHVH process flow is described using cross sections of the RESURF device. Some of the challenges associated with the manufacturing of high voltage analog processes are also discussed.

Chapter 5 proposes and demonstrates an optimisation strategy for the two possible process options under investigation. As part of the proposed methodology, the manufacturability of both flows is assessed using a Monte Carlo simulation approach. This determines the variability of the process variables, and uses the contour plots obtained from the DOE results to predict the breakdown distributions. The results are compared to measurements and a process is recommended. Finally, the interaction of process and packaging is investigated.

Chapter 6 uses the same optimisation strategy to that described in chapter 5 but this time applied to a 700V process. However, the process had simulated data points where the polynomial surface was unable to fit the data. Covariance fitting was applied to provide a gentle distortion of the trend so that the response describes the distribution of data. The results of the covariance analysis were compared to those from the standard polynomial approach. A comparison of the predicted breakdown voltage distributions and the

corresponding measured results is made. Covariance is also applied to experimental data, which included one point that had a large residual error but was not considered as being an outlier.

Chapter 7 presents the conclusions and a final discussion of the results.

References

1. International Technology Roadmap for Semiconductors, Semiconductor Industry Association, 1999 Edition
2. M.S. Adler, K.W. Owyang, B.J. Baliga, R.A. Kokosa, "The evolution of power device technology", IEEE Trans Electron Devices, Vol. ED-31, pp 1570-1591, November 1984
3. B.J. Baliga, D.Y. Chen, "Power transistors", New York: IEEE Press, 1984
4. R.S. Wrathall, "The design of high power solid state automotive switch in CMOS-VDMOS technology", IEEE Power Electronics Specialists Conference Digest, pp 229-233, 1985
5. A. Andreini, C. Contiero, P. Galbiati, "A new integrated silicon gate technology combining bipolar linear, CMOS logic, and DMOS power parts", IEEE Trans. Electron Devices, Vol. ED-33, pp 2025-2030, December 1986
6. B. Murari, "Smart power technology and the evolution from protective umbrella to complete system", IEEE Int. Electron Device Meeting Digest, pp 9-15, 1995
7. C. Contiero, "BCD technologies for VLSI smart power ICs", 25th ESSDERC Conference, September 1995.
8. H.J. Gregory, A. Sadovnikov, D. Farrenkopf, P.J. Cacharelis, "A 0.72μm smart power technology using non-self aligned DMOS and up/down-diffused HV wells", European Solid-State Device Research Conference, Bordeaux, France, pp.384-387, Sept,1998.

9. C.Contiero, P. Galbiati, M. Palmieri, L. Vecchi, "Characteristics of a 0.6μm bipolar-CMOS-DMOS technology combining VLSI non-volatile memories", Proc IEEE Int. Electron Device Meeting Digest, paper 18.1, 1996

10. J. Smith, A. Tessmer, L. Springer, P. Madhani, J. Erdeljac, J. Mitros, T. Efland, C.Y. Tsai, S. Pendharkar, L. Hutter, "A 0.7μm linear BiCMOS/DMOS technology for mixed signal power applications", IEEE Bipolar/BiCMOS circuits and technology meeting, pp 155-157, Minneapolis, MN, USA, September 1997

11. J.D. Beascom, "A high performance high voltage lateral PNP structure", IEEE Electron Device Meeting Digest, pp 424-427, 1985

12. A.Y. Tang, R. Johnston, P.J. Meza, "A dual buried layer technology for the fabrication of high voltage NPN devices compatible with a 1.5 micron epitaxial bipolar process", IEEE Int. Electron Device Meeting Digest, pp 298-301, 1985

13. T. Yamaguchi, S. Morimoto, "Process and device design of a 1000V MOC IC", IEEE Trans Electron Devices, Vol. ED-29, pp 1171-1178, August 1982

14. V. Rumennik, D.L. Heald, "Integrated high and low voltage CMOS technology", IEEE Int. Electron Device Meeting Digest, pp 77-80, 1982

15. M.J. Declercq, J.D. Plummer, "Avalanche breakdown in high voltage DMOS devices", IEEE Trans Electron Devices, Vol. ED-23, pp 1-4, January 1976

16. A.W. Ludikhuize, "High voltage DMOS and PMOS in analog ICs", IEEE Int. Electron Device Meeting Digest, pp 81-84, 1982

17. J.D. Plummer, J.D. Meindl, "A monolithic 200V CMOS analog switch", IEEE J. Solid State Circuits, Vol. SC-11, pp809-817, December 1976

18. G.M. Dolny, O.H. Schade, B. Goldsmith, L.A. Goodman, "Enhanced CMOS for analog-digital power IC applications", IEEE Trans Electron Devices, Vol. ED-33, pp 1985-1991, December 1986

19. S. Krishna, J. Kuo, I.S. Gaeta, "An analog technology integrates bipolar, CMOS and high voltage DMOS transistors", IEEE Trans. Electron Devices, Vol. ED-31, pp 89-95, January 1984

20. G.A.J. Amaratunga, F. Udrea, R.A. McMahon, "Power integrated circuits: devices and applications", IEEE Bipolar/BiCMOS circuits and technology meeting, Minneapolis, MN, USA, pp 75-79, September 1999

21. A.F.J. Murray, W.A. Lane, "Optimisation of interconnection induced breakdown voltage in junction isolated ICs using biased polysilicon field plates", IEEE Trans. Electron Devices , Vol. 44, No. 1, January 1997

22. B. Murari, "Power integrated circuits: problems, trade-offs and solutions", IEEE J. Solid State Circuits, Vol. SC-13, pp 307-319, June 1978

23. B. Zojer, R. Koban, R. Petschader, W. Sereinig, "A 150V subscriber line interface circuit (SLIC) in a new BiCMOS/DMOS technology", IEEE Journal of solid state circuits , Vol. 32, No. 9, September 1997

24. J.A. Appels, H.M.J. Vaes, "High voltage thin layer devices (RESURF DEVICES)", IEEE Int. Electron Device Meeting Digest, pp 238-241, 1979

25. J.A. Appels, M.G. Collet, P.A.H. Hart, H.M.J. Vaes, J.F.C.M. Verhoeven, "Thin layer high voltage devices (RESURF devices)", Phillips Journal of Research, Vol. 35, No. 1, pp 1-13, 1980

26. T. Efland, S. Malhi, W. Bailey, O. K. Kwon, W.T. Ng, M. Torreno, S. Keller, "An optimised RESURF LDMOS power device module compatible with advanced logic processes", IEEE Int. Electron Device Meeting Digest, pp 237-240, 1992

27. Y. Koishikawa, M. Takahashi, H. Yanagigaawa, T. Kuriyama, "Double RESURF device technology for power ICs", NEC Research & Development, Vol. 35, No. 4, pp 438-443, October 1994

28. A.W. Ludikhuize, "A versatile 700-1200V IC process for analog and switching applications", IEEE Trans. on Electron Devices, Vol. 38, No. 7, pp 1582-1589, July 1991

29. M.F. Chang, G. Pifer, H. Yilmaz, E.J. Wildi, R.G. Hodkins, K. Owyang, M.S. Adler, "Lateral HVIC with 1200V bipolar and field effect devices", IEEE Trans. Electron Devices, Vol. ED-33, pp 1992-2001, December 1986

30. R.K. Williams, L.T. Sevilla, E. Ruetz, J.D. Plummer, "A DI/JI compatible high voltage multiplexer", IEEE Trans. Electron Devices, Vol. ED-33, pp 1977-1984, December 1986

31. E.J. Wildi, P.V. Gray, T.P. Chow, H.R. Chang, "Modeling and process implementation of implanted RESURF type devices", IEEE Int. Electron Device Meeting Digest, pp 268-271, 1982

32. S. Colak, "Effects of drift region parameters on the static properties of power LDMOST", IEEE Trans Electron Devices, Vol. ED-28, No. 12, December 1981

33. A.Q. Huang, X. Li, N.X. Sun, "Novel high voltage SOI structures", IEEE Bipolar/BiCMOS circuits and technology meeting, Minneapolis, MN, USA, September 1999

34. Y.K. Leung, S.C. Kuehne, V.S.K. Huang, C.T. Nguyen, A.K. Paul, J.D. Plummer, S.S. Wong, "Spatial temperature profiles due to non-uniform self heating in LDMOS's in thin SOI", IEEE Electron Device Letters, Vol. 18, No. 1, January 1997.

35. Y.K. Leung, A.K. Paul, K.E. Goodson, J.D. Plummer, S.S. Wong, "Heating mechanisms of LDMOS and LIGBT in ultrathin SOI", IEEE Electron Device Letters, Vol. 18, No. 9, September 1997.

36. F. Gonzalez, V. Shekhar, C.K. Chan, B. Choy, N. Chen, "Fabrication of a 300V, high current (300mA/output), smart-power IC using gate controlled SCRs on bonded (BSOI) technology, Proc IEEE Int. Electron Device Meeting Digest, paper 18.3, 1996

37. A. Nakagawa, Y. Yamaguchi, N. Yasuhara, K. Hirayama, H. Funaki, "New high voltage SOI device structure eliminating substrate bias effects", Proc IEEE Int. Electron Device Meeting Digest, paper 18.4, 1996

38. H.W. Whittington, B.W. Flynn, D.E. Macpherson, "Switch mode power supplies", Research Studies Press, 1992.

Chapter 2: Technology Computer Aided Design (TCAD) and Design of Experiments (DOE)

2.1 Introduction

The rapid evolution and growth of integrated circuit technology has impacted society arguably more than any other technological development of the 20th century. Integrated circuits (ICs) are in everything from computers, communications and information processing to automotive, residential and recreational applications. It is difficult to think of an application that does not use ICs. Therefore, this expanding and highly competitive market requires more accurate methods and tools to enable process and circuit designers to meet market requirements for their processes and products.

In the competitive environment, time to market for a new product is crucial in the quest to win market share. As stated previously it is no longer feasible to continue to debug process technologies after manufacturing release. Therefore, a methodology/strategy [1-7] is required that reduces the development cycle time. The strategy utilised in this work is a Design For Manufacturability (DFM) approach using Technology CAD (TCAD) and Design Of Experiments (DOE) where process and device simulation is used in conjunction with DOEs to produce response surfaces (Response Surface Methodology) [8-13]. The advantage of RSM is that the number of experimental runs and cycles of learnings can be minimised whilst still enabling the mapping of the complete parameter space. The resulting response surface can be used to help optimise the control factor settings in a manufacturing environment. This could also lead to a reduction in the number of electrical tests performed at end of line and, potentially, manufacturing zero

losses. Several workers [8-11,13] have automated this procedure in order to reduce the possibility of error as well as the overall cycle time of the work.

Further work has led to linking TCAD and DOE to factory modelling [6,14]. This allows a dynamic and stochastic analysis of a manufacturing facility that is unavailable using conventional analyses. This includes the ability to identify, for example, production bottlenecks. Adding this capability allows the assessment of new technology introduction into a manufacturing line. A cost analysis can also be performed to provide another metric, by which, to evaluate process options during development.

Today, TCAD is an important tool in the technology development process and is used by almost every semiconductor company [15] in their development and research activities. It is also used, in some cases, in a manufacturing environment [15] where it has been successfully used to address manufacturing challenges [16-18] such as yield improvement. TCAD has also been used to facilitate a reduction in cycle time of a technology transfer [19,20] from one fabrication facility to another.

2.2 Technology CAD

2.2.1 Circuit Simulation

Around 30 years ago, circuit design engineers typically used bread boarding of prototypes before committing their designs to IC processes. As the applications of these new circuits became more complex, computer simulation evolved as a more practical way to predict circuit performance. By the 1970s, the circuit simulator had progressed from a useful

design tool to a critical one, replacing the bread boarding of prototypes. The SPICE [21] program, or "**S**imulation **P**rogram with **I**ntegrated **C**ircuit **E**mphasis" developed at the University of California, Berkeley, became the tool of choice among the entire circuit design community. Since the 1970s SPICE has been further developed as technology has advanced. It has since been commercialised and is now available from several Technology CAD [22-24] companies along with those [25-28] who provide software that support the models contained within the circuit simulator. SPICE models that represent various device types have been developed, and have been incorporated into the software used by engineers to extract the models for the SPICE simulator. A Graphical User Interface (GUI) is employed for ease of use. An example of the models available is shown in figure 2.1, which is from Avant!'s Aurora [27] parameter extraction

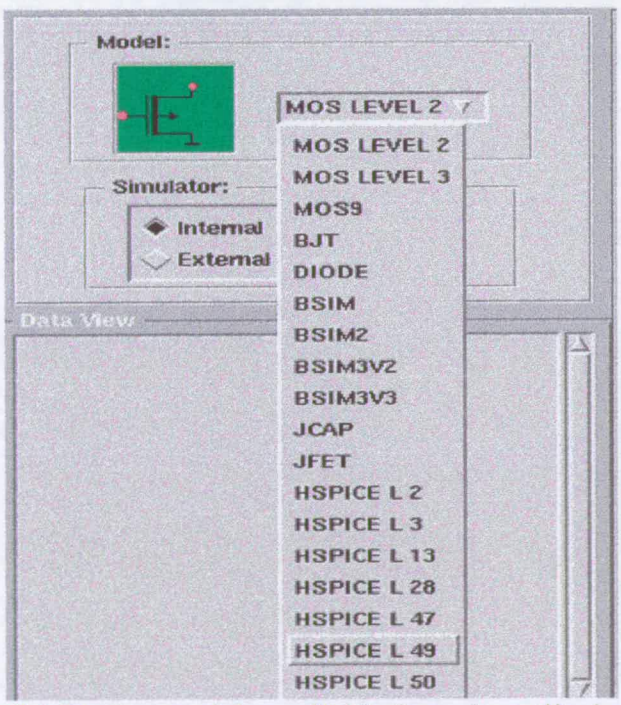


Figure 2.1 Device models available from Avant!'s Aurora [27]

software package. The resulting models are used by the SPICE program or equivalent to model the integrated circuit performance. Examples for bipolar and MOS technologies

are shown in figures 2.2 and 2.3 respectively, where the solid line is the model result and the circles are the measured data. This data was produced using Silvaco's UTMOST software [28].

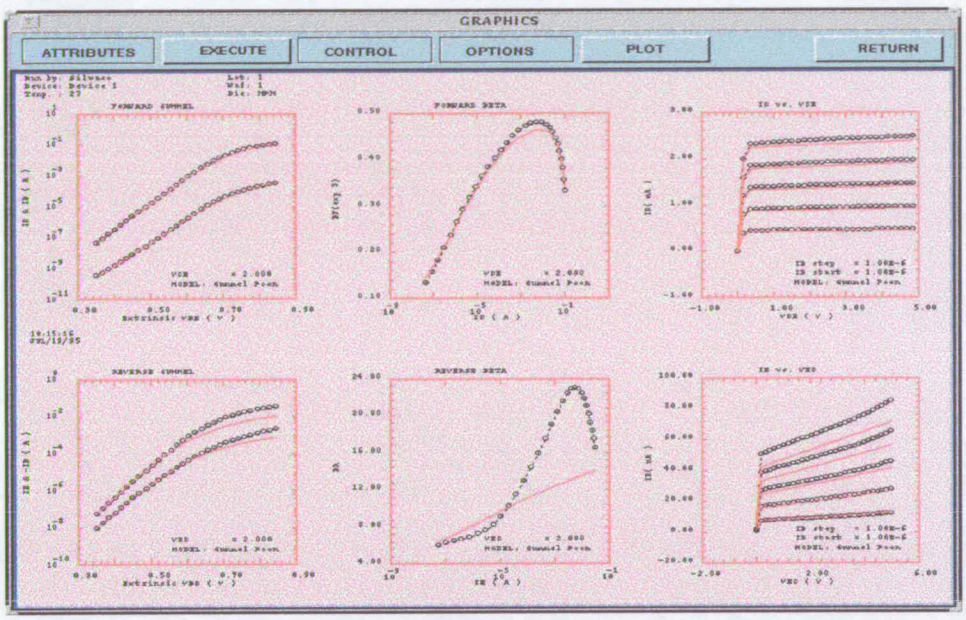


Figure 2.2 Bipolar transistor: model and measured data comparison from Silvaco's UTMOST [28]

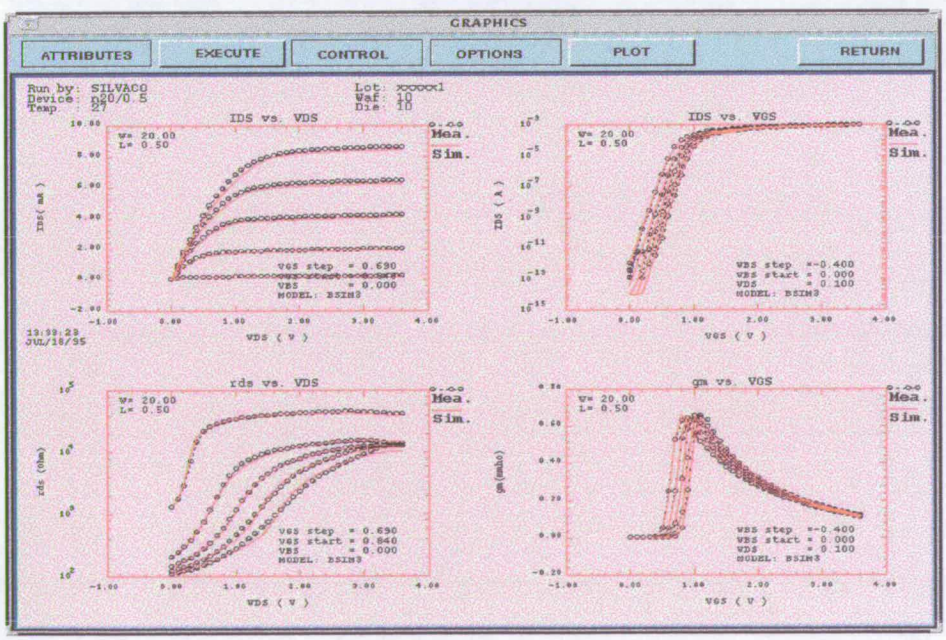


Figure 2.3 MOS transistor: model and measured data comparison from Silvaco's UTMOST [28]

As would be expected the better the fit the more accurate the resulting circuit simulation(s) with the basis for this being if a single component is modeled well, then a circuit containing many of the same components can be modelled accurately.

2.2.2 Process and Device Simulation

The evolution of process and device simulation tools and techniques is closely linked to the development of efficient computers and mathematical models. For example, during the 1960s and 1970s, such tools were sufficient to extract current voltage characteristics but were not frequently used for transient device analysis as it was not economical to perform such simulations.

By the mid-1970s, the significant role of processing technology in determining electrical characteristics was evident. Many important interrelated process and device characteristics were studied and identified by means of a computer. A unified process and device simulator, SITCAP (**SI**mulator for **T**ransistors to **CA**lculate **P**arameters) was developed at the University of Leuven, Belgium [29]. The program inputs process information such as processing times and temperatures along with simple mask geometries, and outputs IV and CV characteristics as well as selected SPICE model parameters. Following several years of experience with SITCAP, researchers at the University of Stanford, California began work on developing a more complete process simulation program. They built on the ideas obtained from SITCAP, SPICE and CASPAR (**C**omputer **A**ided **S**emiconductor **P**rocessing analysis including **E**pitaxial **R**edistribution), a process analysis program developed jointly by Bell laboratories and

Lehigh University [30]. Stanford University developed the SUPREM program (**S**tanford **U**niversity **PR**ocess Engineering **M**odels), and the first version was released in 1977 [31]. Further enhancements took place and the models within SUPREM were enhanced and other models added. This was based on work undertaken at Stanford and other research facilities worldwide. Versions II and III were released in 1978 [32] and 1985 [33], respectively. SUPREM was first commercialised by TMA (now Avant!) followed by several companies such as Silvaco. Such companies improved the efficiency of the program, enhanced the graphics and added graphical user interfaces to make the software more user friendly. Examples of the outputs obtained from SUPREM are shown in figures 2.4 and 2.5.

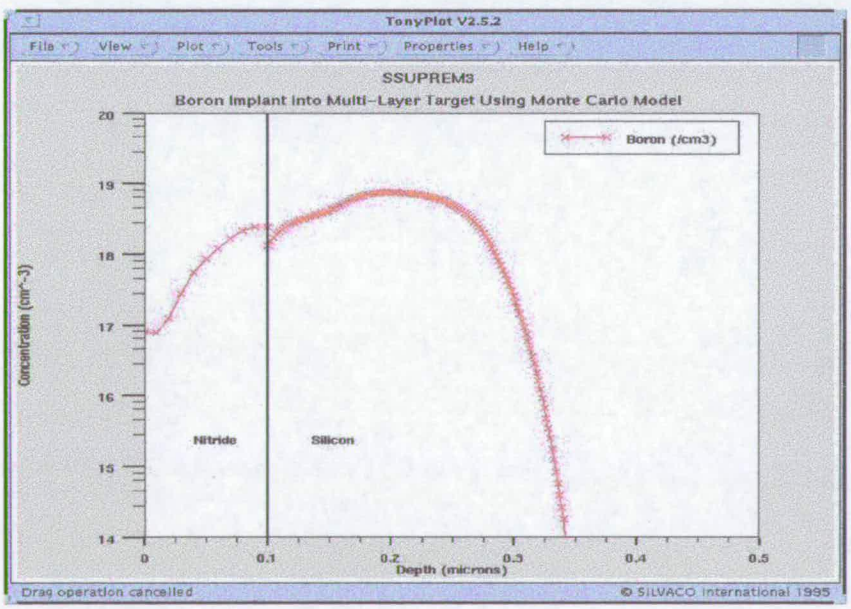


Figure 2.4 SUPREM3 implant dopant profile obtained from SSUPREM3 [34]

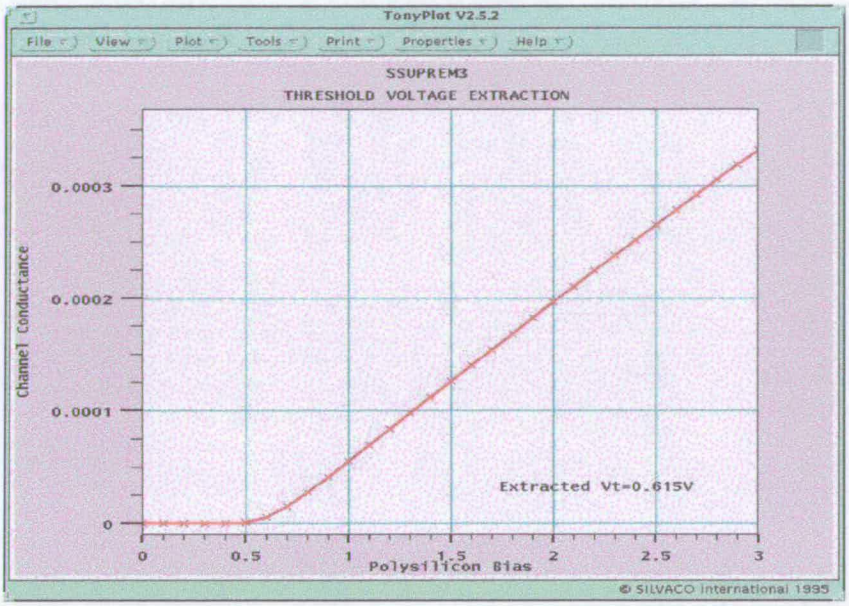


Figure 2.5 MOSFET voltage threshold, V_t, obtained from SSUPREM3 [34]

Figure 2.4 shows an implanted doping profile after the implant through a nitride film. Figure 2.5 shows the IV output capability of SSUPREM3 [34] where the threshold voltage of a MOSFET is calculated.

As technology continued to advance and dimensions and junction depths continued to decrease, two dimensional (2D) computer tools were required. Further work at Stanford resulted in the 2D-simulation program SUPREM-IV [35]. This allowed more complex analyses of semiconductor processes in two dimensions where, for example, dopant diffusion can be investigated in the x and y directions. The models used in SUPREM-IV were more complex, and the trade-off was increased execution time as simulations in 2D were more CPU intensive. Software vendors such as Avant! [36] and Silvaco [37] commercialised SUPREM-IV improving the graphics and software efficiency in order to enhance simulation speed though the execution times remained longer than the 1D

simulations. An example of the capability of TSUPREM4 [36] is shown in figures 2.6 and 2.7 where the simulation of a plasma enhanced chemical vapor deposition (PECVD) is compared to that obtained from silicon.

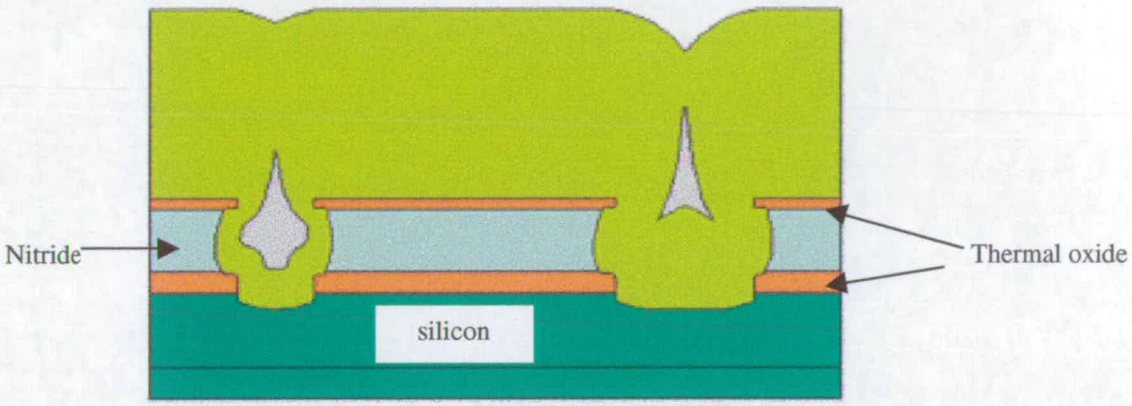


Figure 2.6 TSUPREM4 [36] 2D simulation of PECVD process step

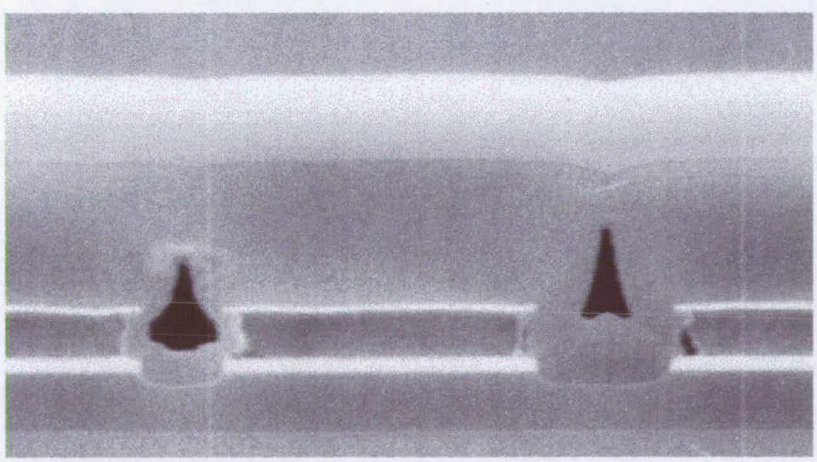


Figure 2.7 SEM silicon cross section of a PECVD process step

It can be seen in this case that the simulator has accurately modelled the observed results. In parallel to the development of process simulation tools, device simulation programs were also being developed. Researchers at Stanford University developed SEDAN

[(**SE**miconductor **D**evice **AN**alysis) program [38]. This led to further development of device simulation tools and researchers at Stanford developed PISCES (**PoI**sson **C**ontinuity **E**quation **S**olver), a 2-dimensional device simulator [39]. These tools took the output from process simulation tools such as SUPREM which in turn allowed device designers to study many areas of the device operation that could not be measured, for example, depletion regions. As computer technology improved, it became economic to perform transient analysis to study problems such as CMOS latch-up [40].

As technology advanced into the deep sub micron area, three-dimensional effects started to become significant, and so modelling software with a 3D capability was developed to enable the device designer to analyse such device types. Figure 2.8 shows the potential contours in a MOSFET under bias conditions indicated in the figure. This was produced using Avanti's Davinci [41].

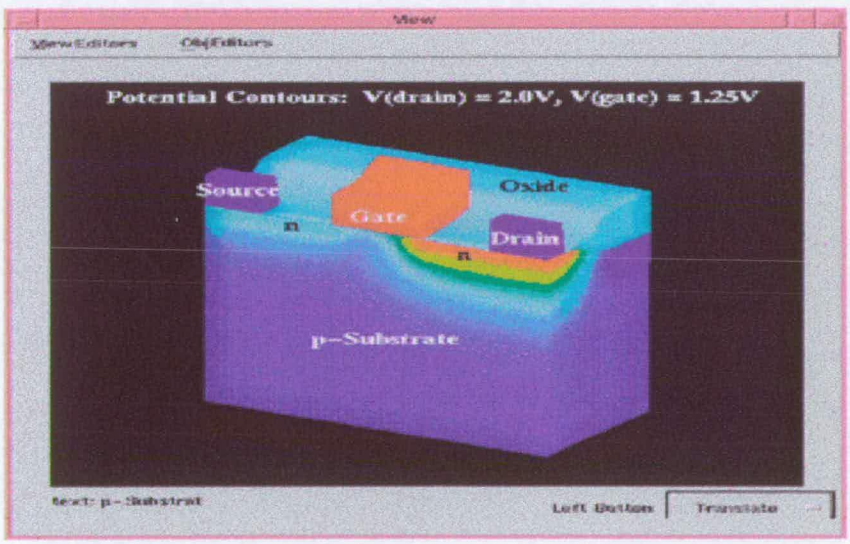


Figure 2.8 3D Davinci [41] simulation of MOSFET under bias

2.3 Design of Experiments (DOE)

Design of experiments (DOE), often referred to as response surface methodology (RSM) and experimental design, is a technique for designing experiments in an efficient way and therefore obtaining the maximum amount of information possible. Traditionally, experiments have been conducted by varying the control parameters haphazardly according to theory or prior results/knowledge. An iterative process would then be applied until the correct or best results were obtained. However, there are many obvious problems using this type of approach [42]:

(a) It is possible that the best results obtained are in fact a local optimisation. For example, to maximise yield, there maybe more than one maximum value in the design space, but with the traditional approach there is no guarantee that the maximum found is the true maximum.

(b) There is no limit on the number of experiment runs required to reach the optimum value.

(c) It is hard to determine which control factors are the major contributors to the response.

In a designed experiment, the engineer decides which parameters to vary, their range, and which responses to record. An experimental design is then chosen and generated, usually from a set of standard designs that have desirable statistical properties [43,44]. Usually the number of runs available for a design is limited and thus designs must be economical as well as allowing for good model fits and providing a means to test the fit. If there are

many control factors, a screening experiment maybe performed to economically eliminate those factors that are not significant before the main experimental design is defined. It is also noted that using TCAD tools, the only limit to the number of experiments that can be simultaneously run is the number of TCAD software licenses and computers available to run the simulations.

Following the experiment, the recorded data can then be used to build an empirical model for the response(s) using regression. Usually a second order response equation is used which will contain quadratic and interaction terms for some or all of the control factors [45]. If the model approximates the true response closely enough then it maybe used to predict the responses. Clearly an important part of the analysis is to determine the goodness of the model fit to the data if it is to be used in a predictive mode. Using simulation there is little, if any, experimental noise because the results of the simulation are repeatable. Once an optimum operating point in the design space is found a confirmation run [46] must be run to check that the prediction was correct.

2.4 Design for Manufacturability (DFM) Strategy

2.4.1 Introduction

As discussed, the use of TCAD is an important tool for the development and understanding of semiconductor processes and device operation. However, if technologies are to be developed which are robust to normal manufacturing variations the simulators themselves must be used in a structured manner. One such methodology [1-7] that can be used is to combine TCAD with DOE. This produces response surfaces [8-13,

47-49] which can then be used to explore the complete parameter space. The major advantage of using DOEs is that all combinations of control factor settings to be simulated are known from the start and therefore it is possible to perform all simulations in parallel.

One potential challenge with this approach, when using stand-alone process and device simulators and DOE software, is ensuring the integrity of the data that is used to create the response surfaces. This is very important since one simulation data point which is incorrect has the potential to badly distort the resulting response surface. Automating the whole procedure removes the potential source of error whilst improving cycle time.

2.4.2 Automation Software

In order to automate this procedure TCAD companies [50,51] and other workers in industry and academia [8-11,53] have produced software that automates the entire simulation cycle using Graphical User Interfaces (GUI). Figure 2.9 shows Avanti's TWB framework [51]. Such software not only links TCAD and DOE but also facilitates the building of the TCAD decks themselves as well as providing library and statistical analysis capabilities.

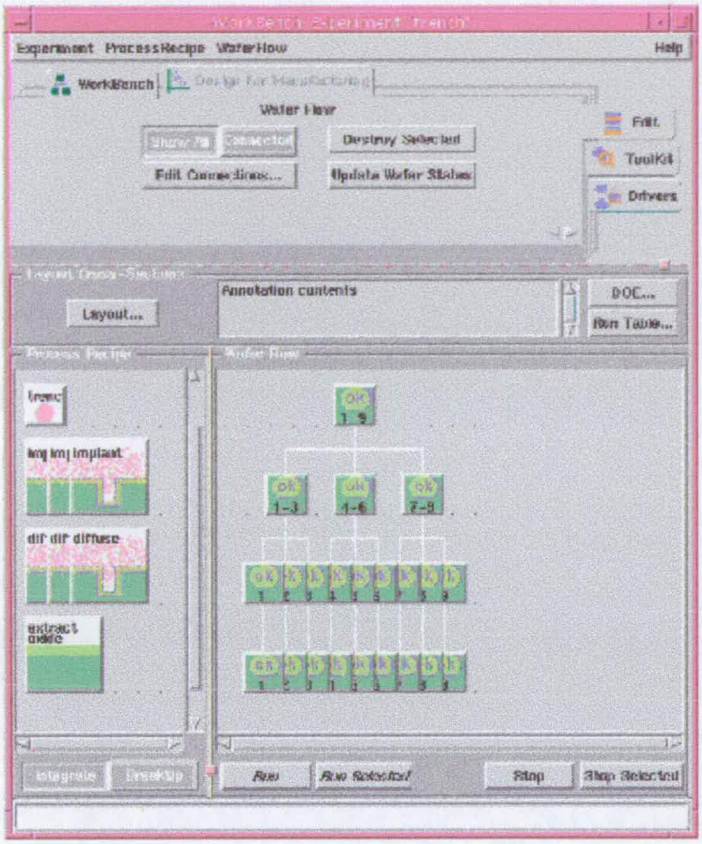


Figure 2.9 Avanti [52] TWB GUI

Walton et al [5] developed Calphurnia [52] which interfaces TWB and RS/1 [42] to automate the optimisation process, and an example of the user interface is shown in figure 2.10. A schematic diagram indicating the role played by Calphurnia in interfacing TWB and RS/1 is shown in figure 2.11.

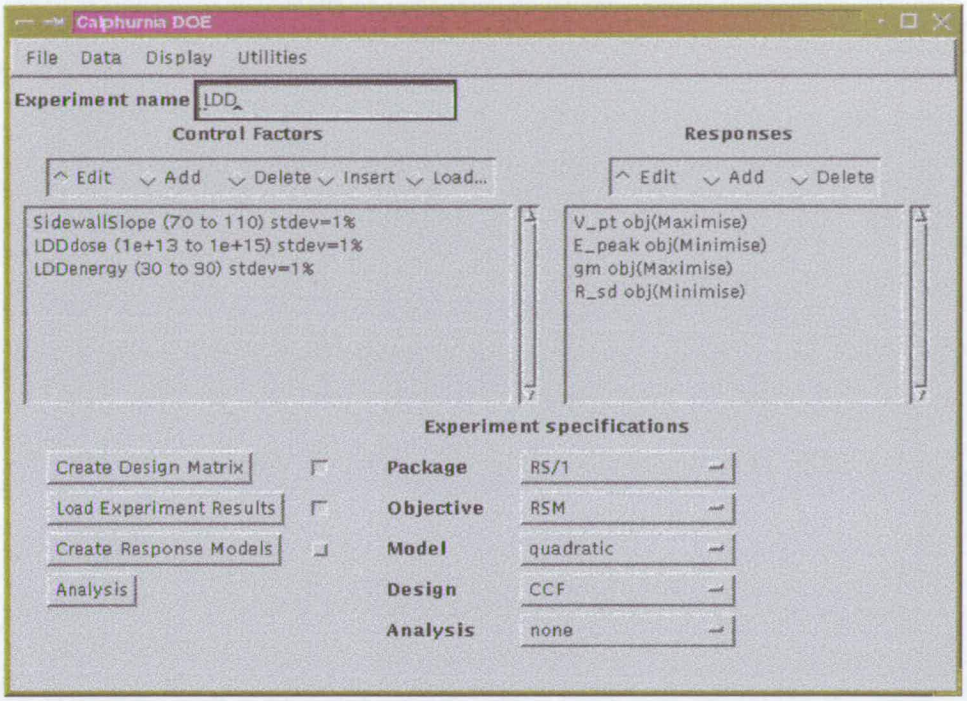


Figure 2.10 The Calphurnia GUI

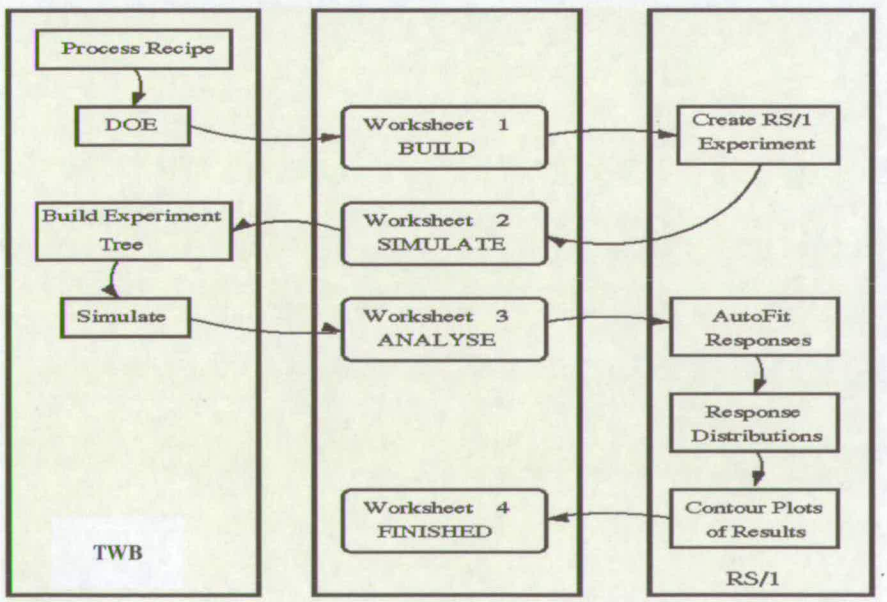


Figure 2.11 The role of Calphurnia in integrating the different software tools

Calphurnia allows the TCAD engineer to simulate processes and fit response surfaces within a GUI environment, thereby ensuring that the data transfer between the simulators and the DOE software is not subject to error. Calphurnia can also use the response surfaces to automatically generate response distributions resulting from control factor variations [9,47,48] that are normally observed in manufacturing. The capability of generating response distributions is an important element of a DFM strategy since it provides an insight into the manufacturability of the technology under consideration. Calphurnia assumes that the control factor distributions are Gaussian in nature as shown in figure 2.12.

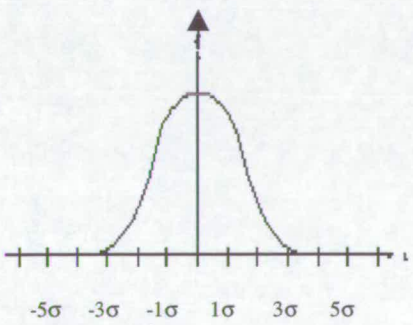


Figure 2.12(a) The normal distribution

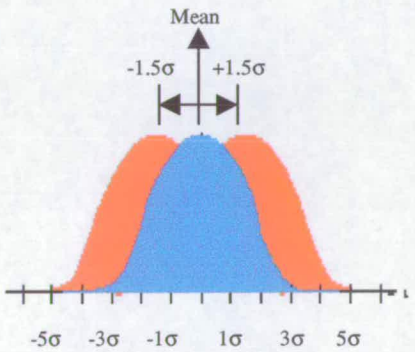


Figure 2.12(b) The normal distribution with a 1.5 shift

Calphurnia allows any distribution function for the control factor variation and provides an option for the control factor mean to be shifted by 1.5σ, mimicking that which is regularly observed in manufacturing. This is depicted in figure 2.12. Once the standard deviations of the control factors have been defined it is then possible to generate their distributions as shown in figure 2.13.

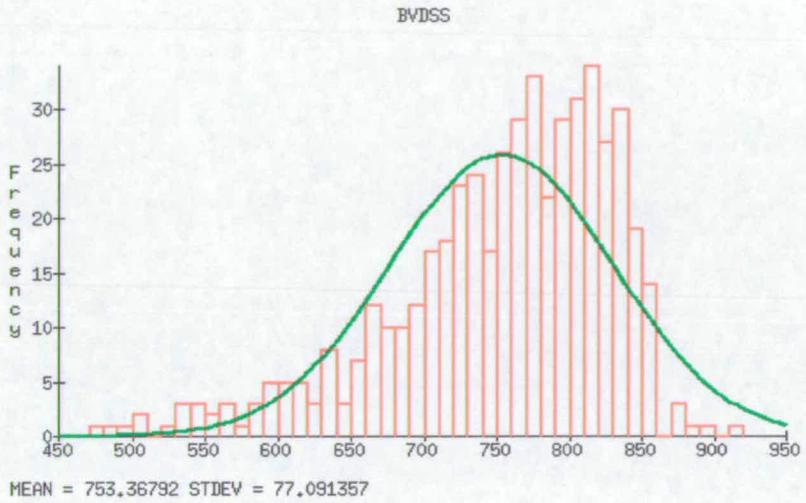


Figure 2.13 Distribution calculated from response surface [48]

2.5 Summary

Chapter 2 has given an overview of the history of technology computer aided design (TCAD) and discussed the application of design of experiments (DOE). In addition, the combination of these techniques as part of a design for manufacturability (DFM) strategy has been reviewed along with integrated software packages (both commercial and academic) that reduce the possibility of error. Calphurnia [52], allows this to be done through integration with the commercial software package TWB [51]. This adds the ability to automatically generate experimental designs as well as analyse the response surfaces automatically. Response surfaces are then used to automatically generate response distributions resulting from control factor variations [9,47,48] that are normally observed in manufacturing. This is an important element of a DFM strategy since it provides an insight into the manufacturability of the technology under consideration. Calphurnia is used in the course of this work.

References

1. M. Redford, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, J.P. Elliott, G.A. Allan, A.J. Walton, "DFM Strategies In Technology Development", National Semiconductor Technology & Innovation Conference, pp 283-284, July 1995.
2. M.R. Kump, S.W. Mylroie, W.J.C. Alexander, and A.J. Walton, "Use of Simulators to Assist in the Design of Processes for Manufacturability", 1990 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, pp 15-21, Danvers, Massachusetts, USA, 11-12th Sept 1990.
3. R. Cartuyvels, R. Booth, L. Dupas, and K. De Meyer, "Process technology optimisation using an integrated process and device sequencing system", Proc ESSDERC 92, pp 507-510, 1992.
4. M.I. Newsam, M. Fallon, and A.J. Walton, "A Total TCAD Environment for DFM", IEE Colloquium on Improving the Efficiency of IC Manufacturing Technology, No. 1995/153, pp.1/1-1/4, Edinburgh, 12th April 1995.
5. A.J. Walton, M. Fallon, M.I. Newsam, R.S.Ferguson, D. Sprevak, G.A. Allan, and J.P. Elliott, "A Total TCAD Strategy for DFM in IC Technology Development", IEE Proc. Science, Measurement and Technology, vol. 144, no. 2, pp. 63-68, March 1997.
6. V.K. Nilsen, A.J. Walton, "Integration of factory simulation with TCAD process and device simulation within a total TCAD approach to DFM", Proc Microelectronic manufacturing technologies, pp 186-196, May 1999.
7. G.J. Gaston, A.J. Walton, "The integration of simulation and response surface methodology for the optimisation of IC processes", IEEE Trans Semiconductor Manufacturing, pp 22-33, Feb 1994.

8. A.J. Walton M. Fallon M.I. Newsam, R.S. Ferguson, D. Sprevak, "Using RSM Techniques to Contour Plot Parameters Related to Response Distributions of Semiconductor Processes and Devices", Proc ESSDERC 94, pp 829-832, April 1995.
9. A.R.Alvarez, B.L.Abdi, D.L.Young, H.D.Weed, J.Teplik, and E.R.Herald, "Application of statistical design and response surface methods to computer-aided VLSI device design", IEEE Transactions on Computer-Aided Design, vol. 7, no. 2, pp 272-287, 1988.
10. D.L.Young, J.Teplik, H.D.Weed, N.T. Tracht, and A.R. Alvarez, "Application of statistical design and response surface methods to computer-aided VLSI device design II: Desirability Functions and Taguchi Methods", IEEE Transactions on Computer-Aided Design, vol. 10, no. 1, pp 103-115, 1991.
11. M. Cecchetti, M. Lissoni, C. Lombardi, and A Marmiroli, "Process analysis using RSM and simulation", ESSDERC 92, pp 511-514, 1992.
12. R. Cartuyvels, R. Booth, S. Kubicek, L. Dupas, and K. De Meyer, "A Powerful TCAD System Including Advanced RSM Techniques for Various Engineering Optimization Problems", SISDEP 93, Vienna.
13. D.S. Boning and P.K. Mozumder, "DOE/OPT -A System for Design of Experiments, Response-Surface Modeling, and Optimization Using Process and Device Simulation", IEEE Transactions on Semiconductor Manufacturing, Vol. 2, pp 233-244, 1994.
14. V.K. Nilsen, A.J. Walton, "Automated cost of ownership analysis for process development", in proc Microelectronic manufacturing technologies, pp 112-120, May 1999.

15. J. D'Ignazio, "TCAD tools leap hurdles to broader use", Semiconductor International pp 99-106, April 1996.

16. D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis of Mixed-Signal Manufacturability with Statistical Technology CAD (TCAD)", IEEE Trans. on Semiconductor Manufacturing, Vol 9, no 4, pp 478-488, 1996.

17. J. McGinty, G. Long, J.K. Kibarian, M. Redford, "Implementation of Statistical TCAD in a Manufacturing Environment", National Semiconductor Technology & Innovation Conference, pp 230-233, September 1996.

18. M. Redford, M. Fallon, J. McGinty, A. J. Walton, M.I. Newsam, G. Michalowicz, C. McGregor, "Identifying And Quantifying Variability Within The LMDMOS 300A Process: A Case Study", IEE Colloquium on Improving the Efficiency of IC Manufacturing Technology, No. 1995/153, pp.10/1-10/4, Edinburgh, 12th April 1995.

19. V.K. Nilsen, A.J. Walton, J. Donnelly, G. Horsburgh, R. Childs, "Implementation of a TCAD based system to aid process transfer", Proc 4th International workshop on statistical metrology, Kyoto, pp 54-57, June 12th 1999.

20. J. Smith, "VIP-1+ 6" transfer process flow", NSUK internal report, July 1998.

21. L.W. Nagel, "SPICE 2, A computer program to simulate semiconductor circuits", ERL Memorandum ERL-M520 University of California, Berkeley, May 1975.

22. SmartSPICE, Advanced analog circuit simulator, Silvaco International, Santa Clara, California, USA.

23. HSPICE, Avant! Corporation, Freemont, California, USA.

24. Eldo, Mentor Graphics, San Jose, California, USA.

25. BTA BsimPro, Berkley Technologies Associates, Santa Clara, California, USA.

26. ICCAP, Hewlett Packard/Agilent, Palo Alto, California, USA.

27. Aurora, Avant! Corporation, Freemont, California, USA.

28. UTMOST, Silvaco International, Santa Clara, California, USA.

29. H.J. DeMan, R. Mertens, "SITCAP – A simulator of bipolar transistors for computer-aided circuit analysis programs", IEEE ISSCC Technical Digest, Paper 9.2, pp 104-105, Feb 1973.

30. P.H. Langer, J.I. Goldstein, "Impurity redistribution during silicon epitaxial growth and semiconductor device processing", J. Electrochemical Society, Vol. 121, No. 4, p563, April 1974.

31. D.A. Antoniadis, S.E. Hansen, R.W. Dutton, A. Gonzalez, "SUPREM-IA program for IC process modeling and simulation", Stanford Technical report, No. 5019-1, May 1977.

32. D.A. Antoniadis, S.E. Hansen, R.W. Dutton, "SUPREM-IIA program for IC process modeling and simulation", Stanford Technical report, No. 5019-2, June 1978.

33. C.P Ho, J.D. Plummer, S.E. Hansen, R.W. Dutton, "VLSI process modeling SUPREM-III", IEEE Trans. ED, Vol. ED-30, No. 11, pp 1438-1453, Nov 1983.

34. SSUPREM3, Silvaco International, Santa Clara, California, USA.

35. M.E. Law, "Two dimensional numerical simulation of dopant diffusion in silicon", PhD thesis, Department of electrical Engineering, Stanford University, Jan 1988.

36. TSUPREM4, Avant! Corporation, Freemont, California, USA.

37. SSUPREM4/Athena, Silvaco International, Santa Clara, California, USA.

38. Z. Yu, R.W. Dutton, "SEDAN III-A generalized electronic material device analysis program", Stanford University Electronics Laboratories Technical Report, July 1985.

39. M.R. Pinto, C.S. Rafferty, H.R. Yeager, R.W. Dutton, "PISCES-IIB", Stanford University Integrated Circuits Laboratory Technical Report, 1985.

40. M.E. Law, C.S. Rafferty, R.W. Dutton, "New n-well fabrication techniques based on 2D process simulation", IEEE IEDM Technical Digest, paper 20.1, pp 518-521, Dec 1986.

41. Davinci, Avant! Corporation, Freemont, California, USA

42. RS/Discover Primer, Domain Software Solutions, Chelmsford, MA, USA.

43. Experimental Design (Part 1), Domain Software Solutions, Chelmsford, MA, USA.

44. Experimental Design (Part 3), Domain Software Solutions, Chelmsford, MA, USA.

45. Exploring Data Trends (Module 2), Domain Software Solutions, Chelmsford, MA, USA.

46. Private communication, Dr Jake Sredni, Resource Optimization Inc.

47. M. Redford, M. Fallon, A.J. Walton, Z.A. Shafi, J. McGinty, I. Underwood. "High Voltage RESURF Process Development using Covariance Based Response Surfaces" Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 76-86, May 18-21 1999.

48. M. Redford, M. Fallon, Z.A. Shafi, N.S. Rankin, J. McGinty, A.J. Walton, "High Voltage RESURF DMOS Process Development Using DFM Techniques A Case Study", in Proc 3rd International Workshop on Statistical Metrology, pp 78-81, June 7, 1998.

49. M. Fallon, G. Martin, M. Redford, R. Prasad, T. Powell, H. Gregory, J. McGinty, "UPD Characterization Using Desirability Functions and Assessment of Process Variability", National Semiconductor Technology & Innovation Conference, pp 30-33, September 1996.

50. Virtual Wafer Fab (VWF), Silvaco International, Santa Clara, California, USA.

51. TWB, Avant! Corporation, Freemont, California, USA.

52. M.I. Newsam, "Calphurnia V2.0 Manual", Dept. of Electrical Engineering, University Of Edinburgh, Scotland.

Chapter 3: Device Physics of a pn Junction

3.1. Introduction

The pn junction is a fundamental building block of integrated circuits. In the LDMOS device under investigation here it is the pn junction breakdown that needs to be optimised. As a precursor to the main research topic, the breakdown theory associated with a basic silicon pn junction will be explored.

3.2 The pn Junction Depletion Region

In order to simplify the analysis we will assume uniform doping and an abrupt pn junction. In the p-type region, $N_A \gg N_D$, and in the n-type region, $N_D \gg N_A$, where N_A and N_D are number of acceptors and donors per unit volume for the p and n type materials respectively. Figure 3.1 shows a schematic of a pn junction and the resulting depletion region, with figure 3.2 showing its corresponding energy band diagram. It can be observed that contact potential is given by, $V_D = V_1 + V_2$.

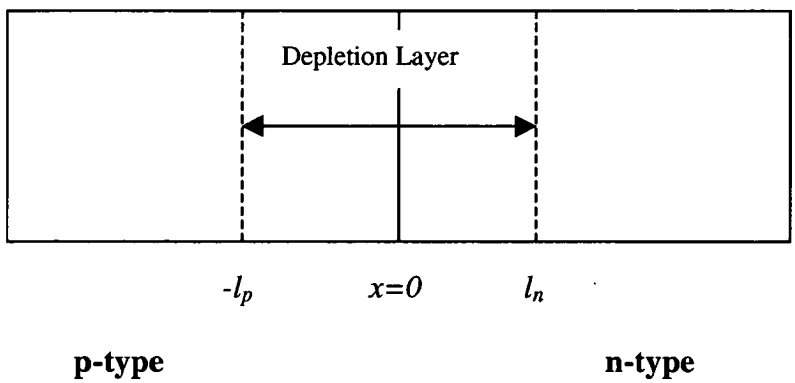


Figure 3.1 Depletion region widths l_n and l_p for the n and p type sides of the junction

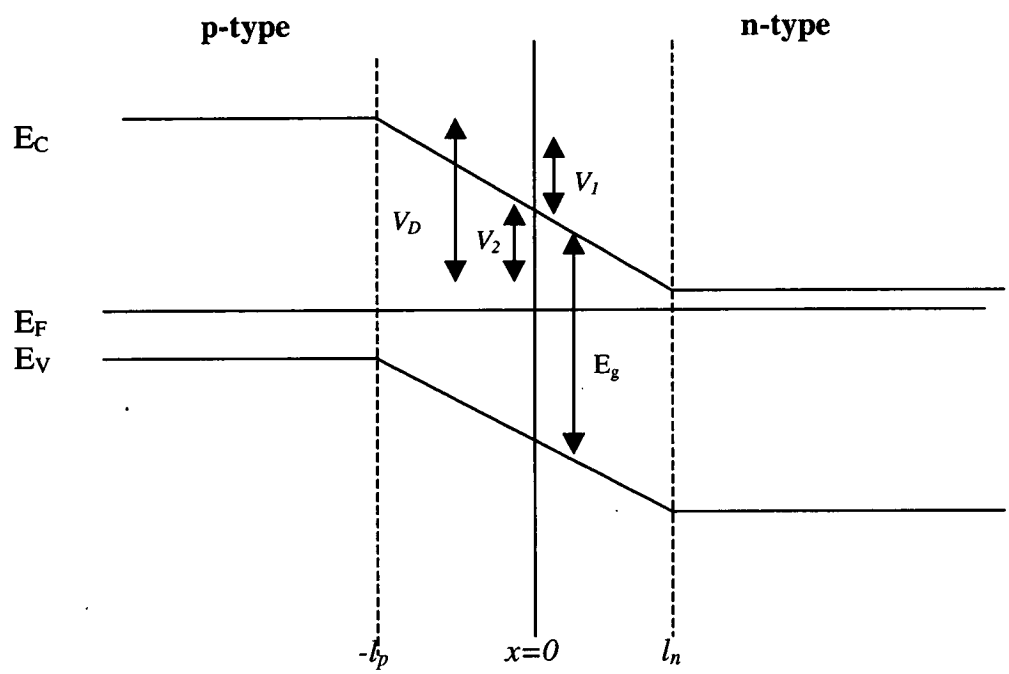


Figure 3.2 Energy band diagram for a pn junction, showing contact potential, $V_D=V_1+V_2$

3.2.1 The Maximum Electric Field in a PN Junction

To determine the maximum electric field across the pn junction the starting point is Poisson's equation. In 1D this is given by [1]:

$$\frac{\partial E}{\partial x}=\frac{\rho}{\varepsilon_o\varepsilon_r} \tag{3.1}$$

where E is the electric field at any point x, ρ is the charge density (Cm^{-3}), ε_o is the permittivity of free space and ε_r is the permittivity of silicon.

The electric field is given by,

$$E=-\frac{\partial V}{\partial x} \tag{3.2}$$

where V is the voltage across the pn junction

Combining (3.1) and (3.2) gives

$$\frac{\partial E}{\partial x} = -\frac{\partial^2 V}{\partial x^2} = \frac{\rho}{\varepsilon_o \varepsilon_r} \tag{3.3}$$

In equilibrium, the charge is the same on either side of the junction, as shown in figure 3.3. That is, the total charge per unit area on either side must be equal and opposite, as shown in figure 3.3

$$N_A l_p = N_D l_n \tag{3.4}$$

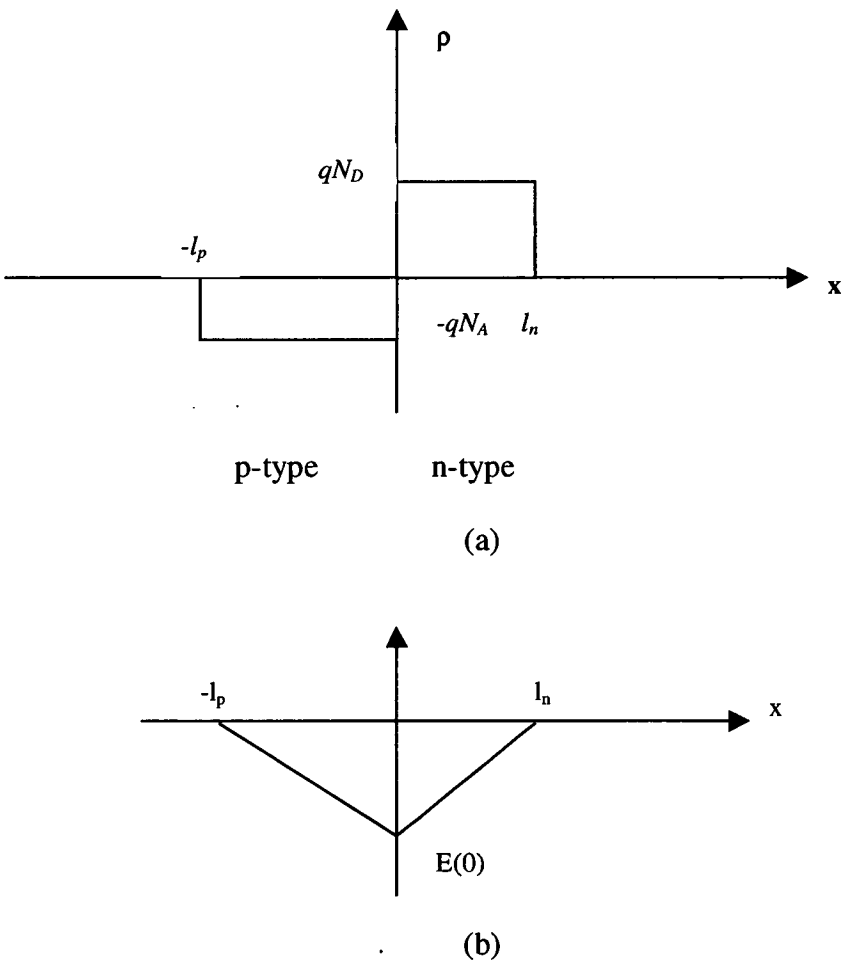


Figure 3.3 pn junction (a) Charge distribution and (b) electric field at equilibrium

3.2.2 Analysis of an abrupt PN junction at thermal equilibrium.

For the depletion region, applying Poisson's equation to the p-type side ($-l_p < x < 0$) gives

$$\frac{\partial E}{\partial x} = -\frac{\partial^2 V}{\partial x^2} = \frac{\rho}{\varepsilon_o \varepsilon_r} = -\frac{qN_A}{\varepsilon_o \varepsilon_r} \tag{3.5}$$

And, for the n-type side ($l_n < x < 0$)

$$\frac{\partial E}{\partial x} = -\frac{\partial^2 V}{\partial x^2} = \frac{\rho}{\varepsilon_o \varepsilon_r} = \frac{qN_D}{\varepsilon_o \varepsilon_r} \tag{3.6}$$

Integrating equation (3.5),

$$-E = \frac{\partial V}{\partial x} = \frac{qN_A x}{\varepsilon_o \varepsilon_r} + A \tag{3.7}$$

When x= - l_p, $E=0$, so $\partial V/\partial x = 0$

Therefore,

$$A = \frac{qN_A l_p}{\varepsilon_o \varepsilon_r} \tag{3.8}$$

$$-E = \frac{\partial V}{\partial x} = \frac{qN_A}{\varepsilon_o \varepsilon_r}(x + l_p) \tag{3.9}$$

Integrating the electric field, E, gives the potential distribution across the depletion region

$$V = \frac{qNA}{\varepsilon_o \varepsilon_r}\left[\frac{x^2}{2} + l_p x\right] \tag{3.10}$$

When $V = V_1$, x = - l_p, which gives

$$V_1 = \frac{qN_A}{\varepsilon_o \varepsilon_r}\left[\frac{l_p^2}{2} - l_p^2\right] \tag{3.11}$$

$$V_1 = -\frac{qN_A l_p^2}{2\varepsilon_o \varepsilon_r} \tag{3.12}$$

Similarly, for equation (3.5), it can be shown that for $0 < x < l_n$

$$-E = \frac{\partial V}{\partial x} = -\frac{qN_D}{\varepsilon_o \varepsilon_r}(x - l_n) \tag{3.13}$$

Which enables V to be determined

$$V = -\frac{qN_D}{\varepsilon_o \varepsilon_r}\left(\frac{x^2}{2} - l_n x\right) \tag{3.14}$$

The voltage, V_2 where $x = l_n$ is given by

$$V_2 = \frac{qN_D}{2\varepsilon_o \varepsilon_r} l_n^2 \tag{3.15}$$

Thus the junction potential, V_D, is given by:

$$V_D = V_1 + V_2$$

$$V_D = \frac{q}{2\varepsilon_o \varepsilon_r}\left(N_A l_p^2 + N_D l_n^2\right) \tag{3.16}$$

When $x = 0$, equations (3.9) and (3.13) can be equated

$$-E(0) = \frac{qN_A l_p}{\varepsilon_o \varepsilon_r} = \frac{qN_D l_n}{\varepsilon_o \varepsilon_r} \tag{3.17}$$

Reducing equation (3.17) to give:

$$N_A l_p = N_D l_n \tag{3.18}$$

Combining equations (3.12), (3.15), and (3.18)

$$\left|\frac{V_1}{V_2}\right| = \frac{N_D l_n^2}{N_A l_p^2} = \frac{l_n}{l_p} \tag{3.19}$$

The width of the depletion layer can be obtained by combining equations (3.16) and (3.18)

$$V_D = \frac{q}{2\varepsilon_o \varepsilon_r}\left(\frac{N_D N_A^2 l_p^2}{N_D^2} + N_A l_p^2\right) = \frac{q N_A^2 l_p^2}{2\varepsilon_o \varepsilon_r}\left(\frac{1}{N_A} + \frac{1}{N_D}\right) \tag{3.20}$$

Which gives,

$$l_p = \sqrt{\left(\frac{2V_D \varepsilon_o \varepsilon_r}{q} \times \frac{N_D}{N_A(N_D + N_A)}\right)} \tag{3.21}$$

$$l_n = \sqrt{\left(\frac{2V_D \varepsilon_o \varepsilon_r}{q} \times \frac{N_A}{N_D(N_A + \mathrm{N}_D)}\right)} \tag{3.22}$$

Hence, the depletion width, W, is given by

$$W = l_n + l_p \tag{3.23}$$

$$W = \sqrt{\frac{2V_D \varepsilon_o \varepsilon_r (N_A + N_D)}{q N_A N_D}} \tag{3.24}$$

Thus any change in V_D using an external bias will modify the depletion width. From figure 3.2 it can be observed that the maximum electric field is located at the junction and can be determined by substituting in equation (3.21) or (3.22)

$$E_{\max} = \frac{qN_D l_n}{\varepsilon_o \varepsilon_r} = \frac{qN_A l_p}{\varepsilon_o \varepsilon_r} = \sqrt{\frac{2qV_D}{\varepsilon_o \varepsilon_r} \times \frac{N_A N_D}{N_A + N_D}} \tag{3.25}$$

3.2.2 The P-N+ Junction

From equation (3.25) it can be seen that for a pn junction, where $N_D \gg N_A$, the maximum electric field, E_{max} can be approximated by:

$$E_{\max} = \frac{qN_A l_p}{\varepsilon_o \varepsilon_r} = \frac{qN_A}{\varepsilon_o \varepsilon_r} \sqrt{\frac{2V_D \varepsilon_o \varepsilon_r N_A}{q}} \tag{3.26}$$

Which gives,

$$l_n = \sqrt{\frac{2V_D \varepsilon_o \varepsilon_r}{q} \times \frac{1}{N_A}} \tag{3.27}$$

In this case,

$$E_{\max}^2 = \frac{2qN_A V_D}{\varepsilon_o \varepsilon_r} \tag{3.28}$$

And V_D is given by,

$$V_D = \frac{\varepsilon_o \varepsilon_r E_{\max}^2}{2qN_A} \tag{3.29}$$

3.2.3 Forward and Reverse Bias

The application of a voltage across the junction affects the energy band diagram depending upon the polarity and magnitude of the applied voltage. The application of a forward bias (p-type is at more positive bias than the n-type), reduces the barrier and the depletion width, enabling current to flow across the junction.

If a negative voltage is applied to the p-type region, then the potential drop across the region is simply the sum of the built in potential, V_D, and the applied voltage. The depletion width on both sides of the junction increases. If the voltage is large enough the junction will breakdown. There are basically three types of breakdown mechanisms:

(1) Punch-through when the depletion region reaches through to the other side of the pn junction under reverse bias. The phenomenon is sometimes referred to as 'reach through' [1]

(2) Zener breakdown occurs when both the p and n types are highly doped [1] (~10^{18} cm^{-3}). At high doping levels, the n-side of the conduction band appears opposite the p-side valance band. This facilitates tunnelling of electrons from the p-side to the n-side as the depletion region becomes very narrow under reverse bias.

(3) Avalanche multiplication which occurs when, under high electric fields electrons or holes gain sufficient kinetic energy and collide with the silicon lattice thereby releasing another electron. If the freed electron gains sufficient kinetic energy it in turn can release more electrons leading to a chain reaction resulting in junction breakdown. This process is described in more detail in section 3.3.

3.3 Avalanche Breakdown

3.3.1 Introduction

At high reverse applied voltages, the field across the depletion region can result in the generation of electron hole pairs (EHPs). This occurs when an electron entering from the p side of a pn junction is swept across the depletion region by the electric field. As it travels it can be accelerated to a high enough kinetic energy to cause an ionising collision with the silicon lattice. The result is carrier multiplication; the original electron and the generated electrons are swept to the n-side of the junction, with the holes moving towards the p-side. The generated hole can itself gain sufficient kinetic energy to generate another EHP, and so forth. Should this be sustained it is termed an avalanche process since each incoming carrier can initiate a large number of new carriers. When impact ionisation takes place, photons are released and these can be captured by infrared emission photography to highlight the point where the breakdown occurs. Figure 3.4 shows an example of an infra-red emission photograph for a lateral DMOS device.

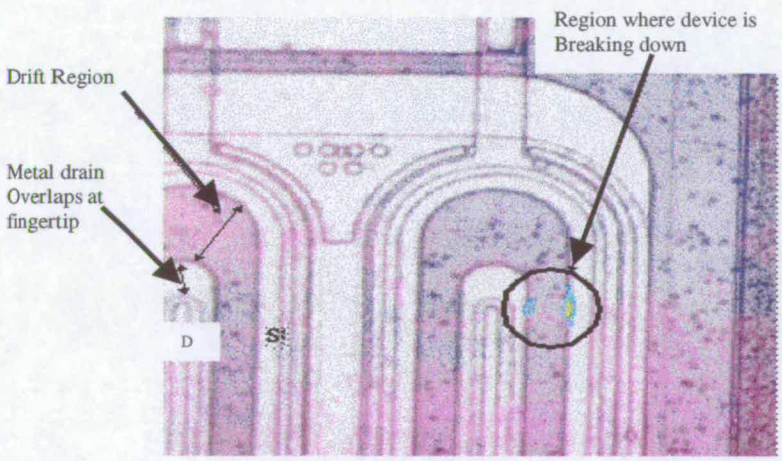


Figure 3.4 Infra-red emission photograph of a lateral DMOS device showing breakdown due to impact ionisation/avalanche mechanisms

3.3.2 Theory of Avalanche Breakdown

Avalanche breakdown has been well documented by numerous researchers [2-6]. The theory of avalanche multiplication [1] can be summarised with reference to figure 3.5

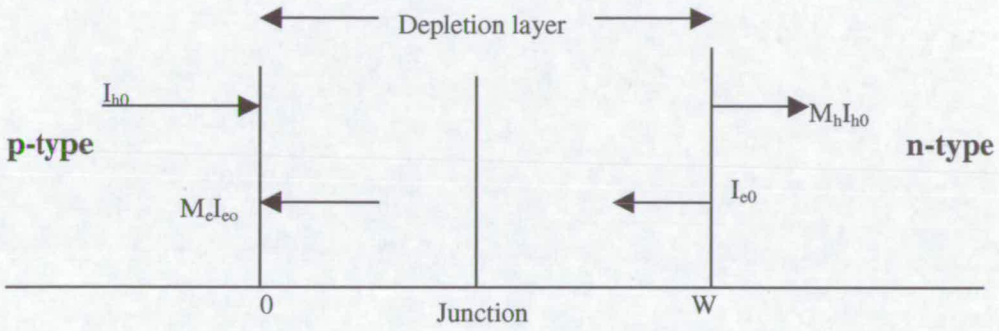


Figure 3.5 Avalanche Multiplication Process of a pn junction

A hole current, I_{ho}, entering the space charge region at x=0, travels a distance W and acquires a new value of M_hI_{ho}, where M is the multiplication factor and is determined by the ionisation rates of the electrons and holes [1]. In the same way, electrons enter at W with a current I_{e0} and exit at x=0 with a new value M_eI_{eo}. The relationship between incremental current and distance for electrons is [1]:

$$\alpha_h I = \frac{dI_e}{dx} - (\alpha_e - \alpha_h)I_e \qquad (3.30)$$

where I is the total current (sum of electron and hole current) and α_e and α_h are the electron and hole ionisation rates respectively.

The multiplication factor for electrons is defined as follows [4]:

$$\int_0^W \alpha_e \exp\left[-\int_0^x (\alpha_e - \alpha_h dx')\right] dx = 1 - \frac{1}{M_e} \qquad (3.31)$$

A similar equation can be written for hole current. Once the electric field is known as a function of distance within the depletion layer, the ionisation integral of the equation

(3.31) can be determined. The breakdown of the junction occurs when M_e becomes infinite. Several sets of ionisation coefficients have been published which allow M to be determined [2-6]. A point to note is that breakdown voltage calculations will seldom be exact, even when numerically solving the ionisation integral for a given impurity profile. This is because the ionisation coefficients themselves are not absolute constants, but are to some extent dependent on the material [1].

3.4. Influence of Field Plates on Breakdown Voltage

The curvature of a finite pn junction drastically reduces the breakdown compared to that obtained from an infinite plane pn junction. Utilising deeper junctions reduces this curvature which increases the breakdown voltage. However, there are limitations to this approach with respect to manufacturability. For example, isolating the device from other devices in an integrated circuit process (chapter 4) can become problematic where junction depths can be of the order of tens of microns.

The peak of the electric field occurs in the region of the curved pn junction, which reduces the avalanche breakdown to a small fraction of the bulk breakdown voltage. The critical peak field E_1 (figure 3.6) can be reduced by applying a field plate, which expands the depletion region at the surface beyond the edge of the field plate.

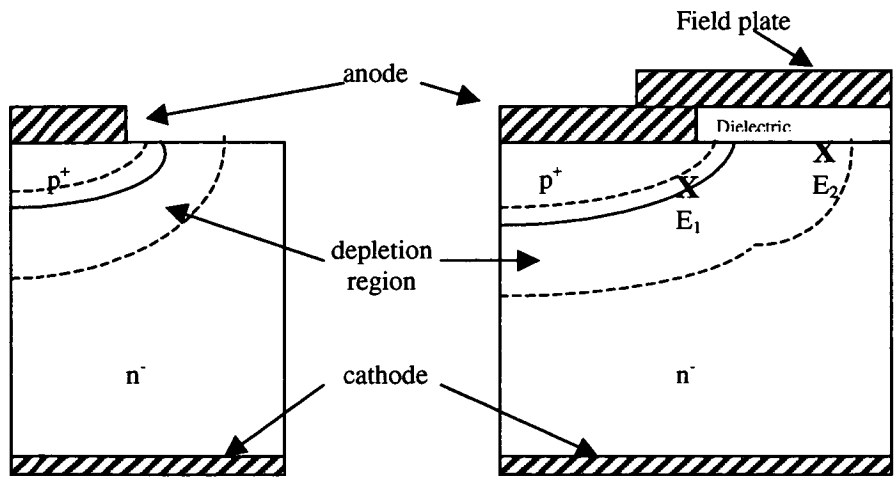


Figure 3.6 Cross section of a planar pn junction without (left) and with a field plate (right)

The field plate has been investigated by Grove et al [7]. The application of field plates has a second beneficial effect because it also protects the surface of the semiconductor against deposition of impurities from the environment. A pn junction with a field plate is shown in figure 3.6 and it can be seen to reduce the curvature of the depletion region. A peak field E_2 occurs at the surface of the semiconductor beneath the edge of the field plate as well as the peak field E_1 in the bulk. Careful design of the field plate results in a higher breakdown voltage than would be possible in a conventional pn junction.

Previous work [8] recognised the fact that the curved boundaries at the edge of diffused regions degraded breakdown voltage because of the field concentration at the edge of the junction [4], and this has become known as the radius of curvature effect.

The use of field plates to alleviate the problems encountered by the radius of curvature effect has also been investigated by Conti and Conti [9]. They showed that the thickness of the oxide beneath the field plate plays an important part in the determination of the breakdown voltage of the junction. Their work indicated that the voltage drop across the oxide increased as a function of oxide thickness. Therefore, the presence of a thin oxide results in field plates that are more effective in reducing the electric field and the resultant avalanche multiplication. As a consequence, the thin oxide thickness is typically increased, from 200 to 800Å in the technologies under consideration in this work, to avoid edge avalanche. This makes it possible for the surface junction breakdown voltage to approach the value of cylindrical junctions without the field plate [4].

The numerical determination of breakdown voltage by Conti and Conti [9] showing that the thickness of the oxide beneath the field plate plays an important role in the determination of the junction breakdown voltage does not lend itself well to the design of planar junctions with field plates. However, O'Neil and Alonas [10] have proposed a simpler solution.

The structure under discussion under reverse bias is schematically shown in figure 3.7. It is assumed that the junction depth of the p^+ diffusion is small compared to the width of the overlapping field plate, W_P, and that the field plate is held at the same potential as the p^+ diffusion.

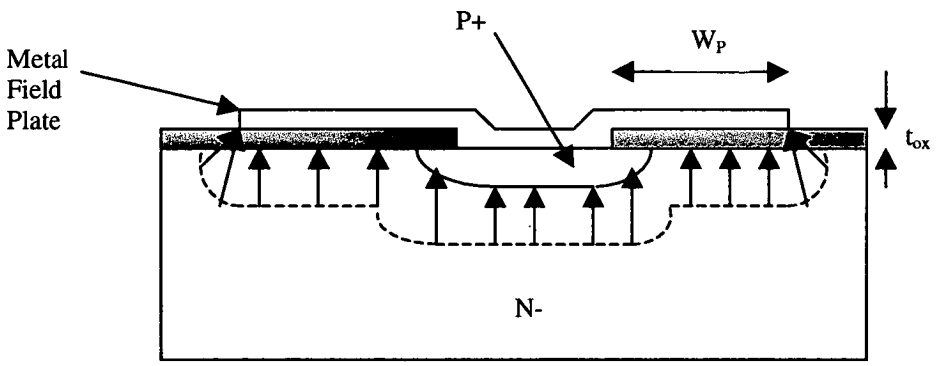


Figure 3.7 p+ n diode with field plate under reverse bias showing field lines from edge of the depletion region to termination at the junction or field plate

It can be seen from figure 3.7 that the field lines crowd at the edge of the field plate, but in the region of the junction the field is more or less normal to the surface of the silicon. The breakdown of a plane junction is given by Sze & Gibbons [11] as

$$V_{B_plane} = 60\left(\frac{N_D}{10^{16}}\right)^{-0.75} \quad \text{Volts} \qquad (3.32)$$

Equation (3.32) is modified in the absence of the field plate by another term which involves the ratio of the junction depth to the maximum depletion width which would occur in the case of plane junction breakdown [4]

$$V_{B_planar} = V_{B_plane}\left(\sqrt{(2+\gamma)\times\gamma}\right) - \gamma \qquad (3.33)$$

where $$\gamma \equiv \frac{x_j}{w_d} \qquad (3.34)$$

The depletion width which would occur at V_{B_plane} is given by [4]

$$w_d = \left(\frac{2\varepsilon_{Si} V_{B_plane}}{qN_D}\right)^{1/2} \qquad (3.35)$$

The field plate area is shown in more detail by figure 3.8 illustrating that the crowding of field lines is similar to the crowding associated with the radius of the junction in the absence of field plates.

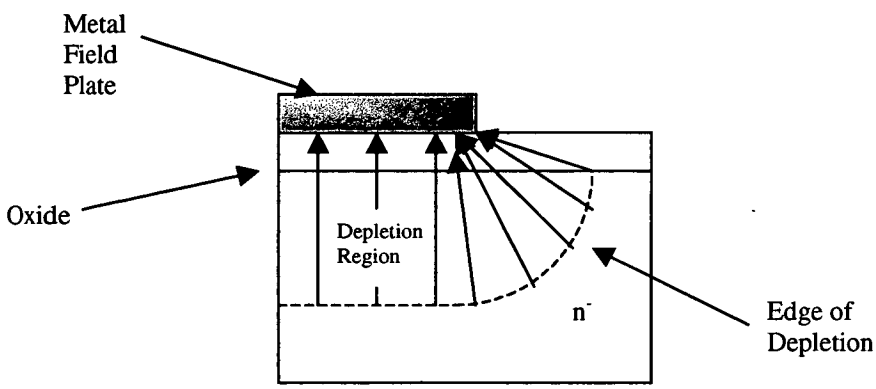


Figure 3.8 Crowding of electric field lines at the edge of the field plate

There are, however, two main differences.

(a) The geometry involves a plane interface between the oxide and the silicon rather than the curved interface in the radius of curvature case.

(b) The field passes through two dielectrics of different permittivities (oxide & depleted silicon).

The requirement that the field lines be continuous through the interface [10] implies that the field in the oxide will be related to the field in the silicon by the ratio of the permittivities of the two materials. Hence the field in the silicon is lower in the oxide by this ratio. The effect of the field plate on the breakdown voltage is similar to the effect of junction depth on breakdown in planar junctions. That is, making the oxide thicker reduces the field crowding effect and thereby increases the breakdown voltage. This effect has been observed by a number of workers [9,12-14]. Therefore, the role of the oxide thickness is similar to that of the junction depth, x_j with the following equivalence

$$x_j \equiv \frac{\varepsilon_{si}}{\varepsilon_{ox}} t_{ox} \tag{3.36}$$

The reduction of the field in the silicon resulting from an increase in oxide thickness is greater than the reduction of the field in the silicon that results from the same increase in the junction depth. An approximation for the breakdown voltage of a planar junction with a field plate in figure 3.6, modified with γ' is given by [10]:

$$\gamma' = \frac{\varepsilon_{Si}\, t_{ox}}{\varepsilon_{ox}\, w_d} \tag{3.37}$$

which gives

$$V_{B_fieldplate} = 60 \left(\frac{N_D}{10^{16}} \right)^{-0.75} \left(\sqrt{(2+\gamma')\times\gamma'} - \gamma' \right) \tag{3.38}$$

3.5 Breakdown Drift

When the junction breaks down, it avalanches at the edge of the field plate, which tends to be located at the edge of the silicon-oxide interface. The electric field here is typically greater than 2.5×10^{5} Vcm^{-1}. This is demonstrated in figure 3.9 where the peak electric field is highlighted in red. This high field imparts electrons with kinetic energy equivalent to heating them to several thousands degrees. With this energy they can then escape from the silicon into the silicon dioxide [15,16]. During this process a negative space charge consisting of trapped electrons builds up at the edge of the oxide. This produces a larger voltage drop across the oxide reducing the breakdown at the silicon surface, consequently increasing the breakdown voltage. However, this effect is limited because when the breakdown voltage at the cylindrical region [4] of the metallurgical junction is reached, the drift stops as the avalanche is no longer localised at the edge of the field plate having moved into the bulk silicon. At this point hot electron injection into the oxide becomes impossible [17].

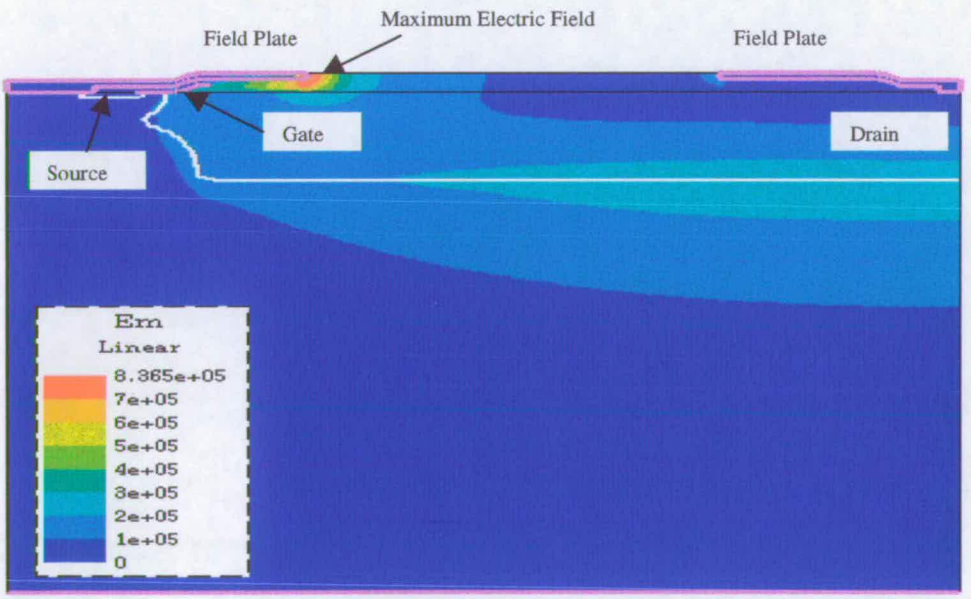


Figure 3.9 Cross section of lateral RESURF DMOS transistor highlighting the maximum electric field, Em, at the edge of the field plate (Medici Simulation)

EX. DIV. EDINBURGH

3.6 Reduced SURface Field (RESURF) Principle

A RESURF device is typically designed to withstand applied voltages of up to several hundred volts. The RESURF technique utilises two-dimensional electric field control near the surface of the device by the careful selection of impurity concentrations and epitaxial thickness. With this technology the device ratings, such as breakdown voltage, can be easily scaled. The epitaxial layer thickness is relatively thin which makes such devices attractive for manufacturing as they can be integrated into an existing IC technology if desired.

The **RE**duced **SUR**face **F**ield (RESURF) concept can be described by considering the simple diode cross section in figure 3.10. The distance between the n^+ and p^+ junctions is L. The electric field for the vertical p^+n^- junction at the surface is E_S, and the electric field in the bulk for the horizontal n^-p^- junction is E_B.

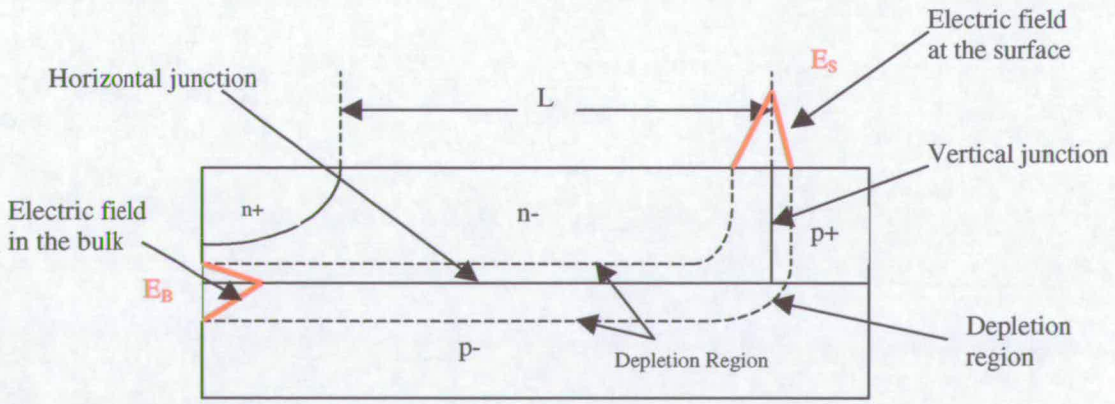


Figure 3.10 Cross section of an epitaxial RESURF diode

Considering the structure in figure 3.10, the n^- epitaxial layer is assumed to be very thick (greater than 5μm) and the distance L (distance between the n^+ and p^+ regions) is much larger than the epitaxial thickness.

The depletion behaviour at the horizontal and vertical junctions can be approximated by the plane junction model. In this case the depletion at the vertical junction can be considered one-sided ($N_A \gg N_D$), whereas for the horizontal junction, the depletion region forms on both sides of the junction. As a result the critical field, E_S, is first reached at the surface (vertical junction). Therefore, if the surface field, E_S, can be reduced then a much higher voltage can be applied before reaching the critical field at the horizontal junction, E_B.

If the epitaxial layer is made thinner, then the influence of the depletion region from the bottom side at the vertical junction then becomes more prominent. This is indicated in figure 3.11 where it can be seen that the interaction of the vertical and horizontal parts of the depletion layer causes it to stretch along the surface a much longer distance than in the plane junction case shown in figure 3.10. This significantly reduces the electric surface field, E_S. Further increase of the reverse voltage totally depletes the epitaxial layer as shown in figure 3.12. In this situation the vertical junction can be considered disconnected from the n^+ contact by depletion from the bottom side.

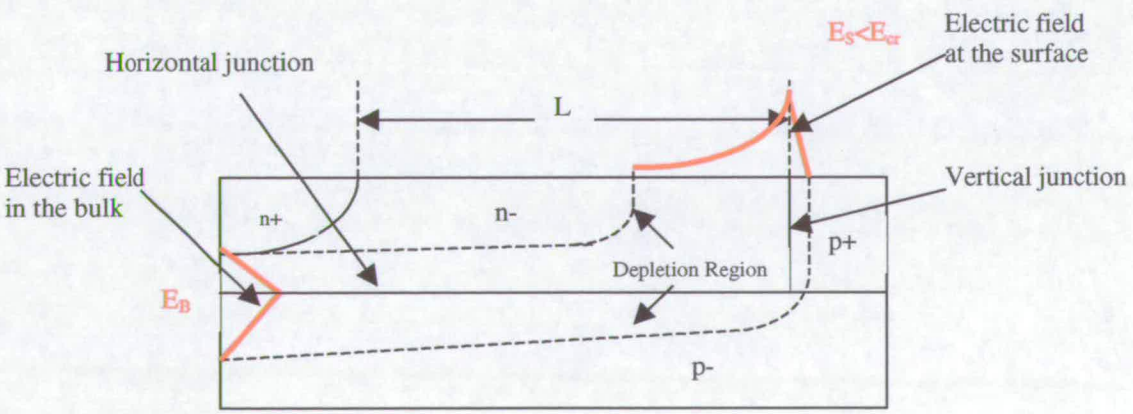


Figure 3.11 RESURF diode structure with thinner epitaxy

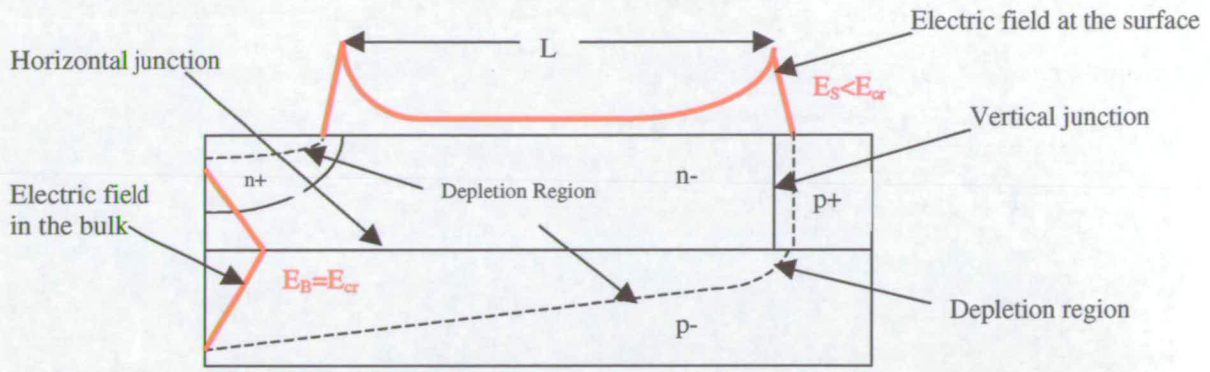


Figure 3.12 Epitaxial RESURF diode structure with thinner epitaxy where the applied voltage totally depletes the n⁻ layer (not to scale)

If complete depletion of the epitaxial layer has to take place before the electric field at the vertical junction can exceed a critical field strength, then from the 1D solution of Poisson's equation, the epitaxial thickness, d_{epi}, and its concentration, N_{epi} must satisfy the following condition [18]:

$$N_{epi}\, d_{epi} < \frac{\varepsilon_o \varepsilon_r E_{cr}}{q\sqrt{1+\frac{N_{epi}}{N_s}}} \qquad (3.39)$$

For the horizontal part of the junction, the theoretical breakdown voltage is given by [18]:

$$BV = \frac{\varepsilon_o \varepsilon_r E_{cr}^2}{2qN_s} + E_{cr} d_1 - \frac{qN_{epi} d_1^2}{2\varepsilon_o \varepsilon_r} \qquad (3.40)$$

Figure 3.11 shows the electric field for this situation. As the breakdown voltage is determined by the integral of the electric field, it can be deduced that increasing the resistivity of the p⁻ substrate, increases the breakdown voltage.

Increasing the voltage after totally depleting the n^- epitaxial will start to deplete the n^+ contact diffusion and at the other side mainly in the p^- substrate. Where the junction curves in the n^+ diffusion, there could be localised enhancement of the electric field, which could lead to breakdown earlier than predicted by equation 3.31.

The RESURF principle can be demonstrated by examining the device cross section shown in figure 3.13. The Medici simulation can be used to show the interaction between the vertical and horizontal depletion regions when the device is put further and further into reverse bias. Figure 3.14(a) shows the depletion regions when 20 volts is applied to the drain with figures 3.14(b) and 3.14(c) showing the depletion regions for 150 and 530 volts. It is also worth noting the simulated depth and width of the device in figure 3.14 are each greater than 60µm, with the drift region 30µm wide. The depth of the device needs to be 60µm to accommodate the depletion spread during its operation.

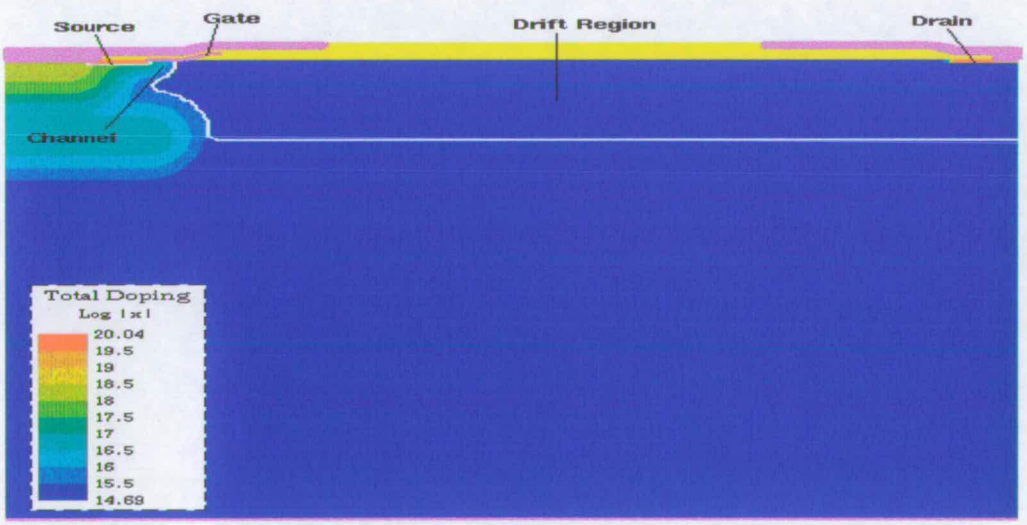


Figure 3.13 Cross section of lateral RESURF DMOS transistor simulated using Medici highlighting the dopant regions and electrodes

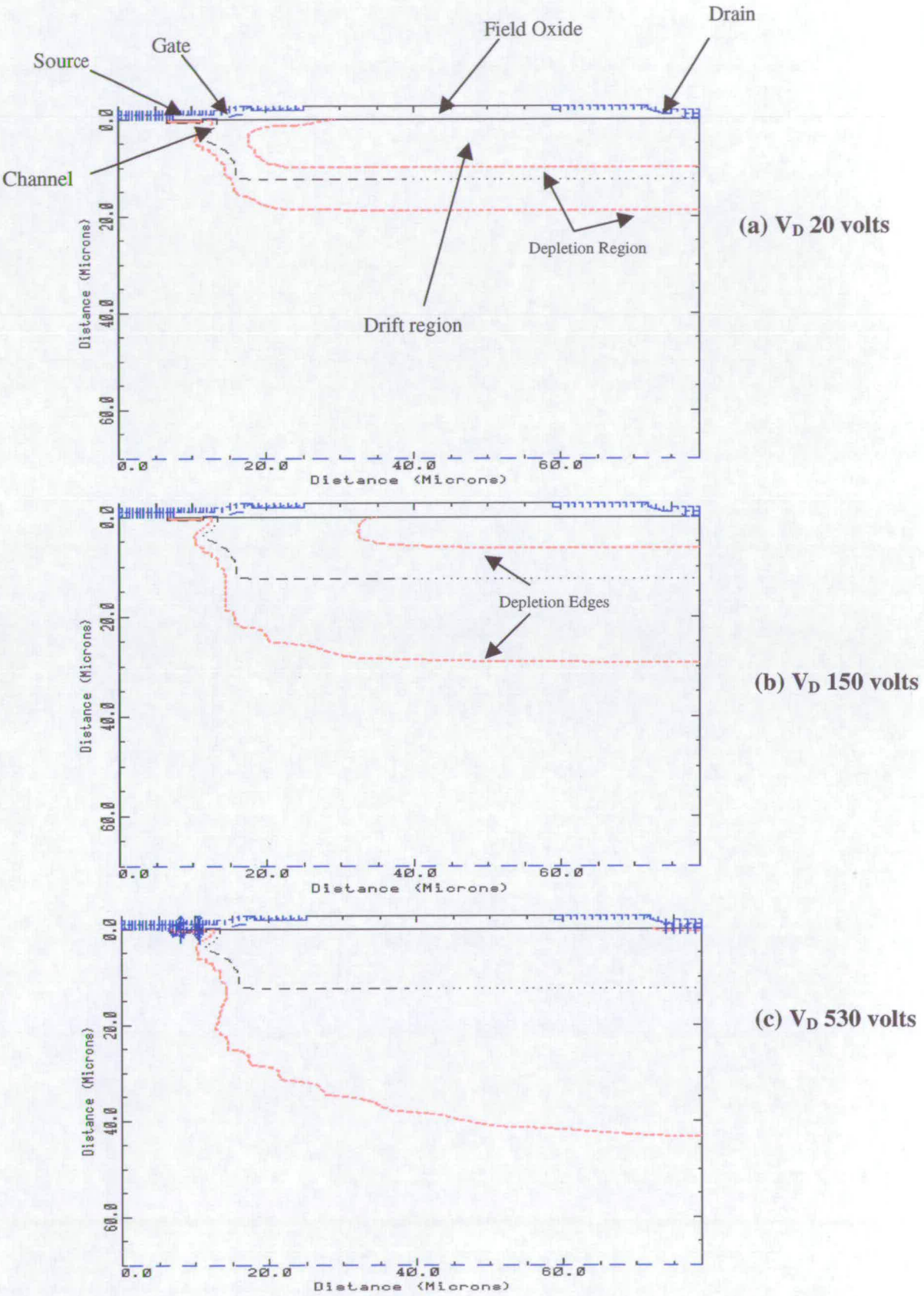


Figure 3.14 Depletion layer interaction from vertical and horizontal p-n junctions with applied drain voltage increasing from 20 to 530 volts

Figure 3.15 shows the results for a Medici simulation showing a cross section illustrating how the applied potential is spread across the n^- region. It can be seen that the potential contours are spread between the source and drain of the structure thus demonstrating the RESURF principle.

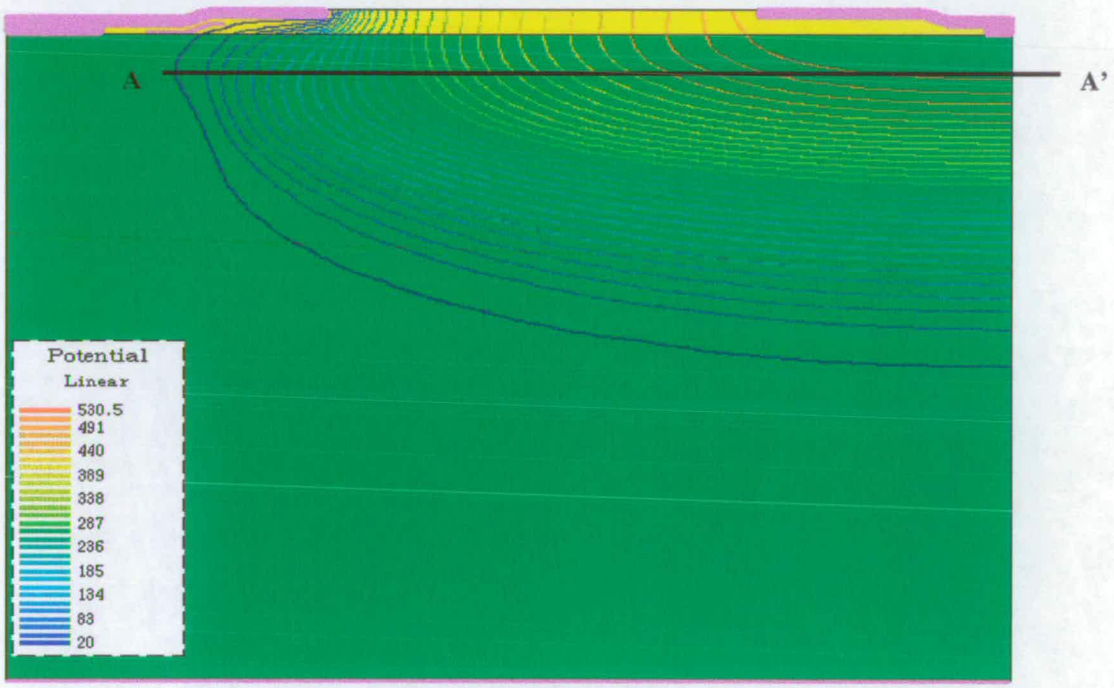


Figure 3.15 Medici device simulation showing applied voltage evenly spread between the source and drain

Examination of the electric field between the source and the drain through points AA' in figure 3.15 is shown in figure 3.16. This shows that the electric field is spread across the drift region and the peak is located beneath the field plate as would be expected for this device design.

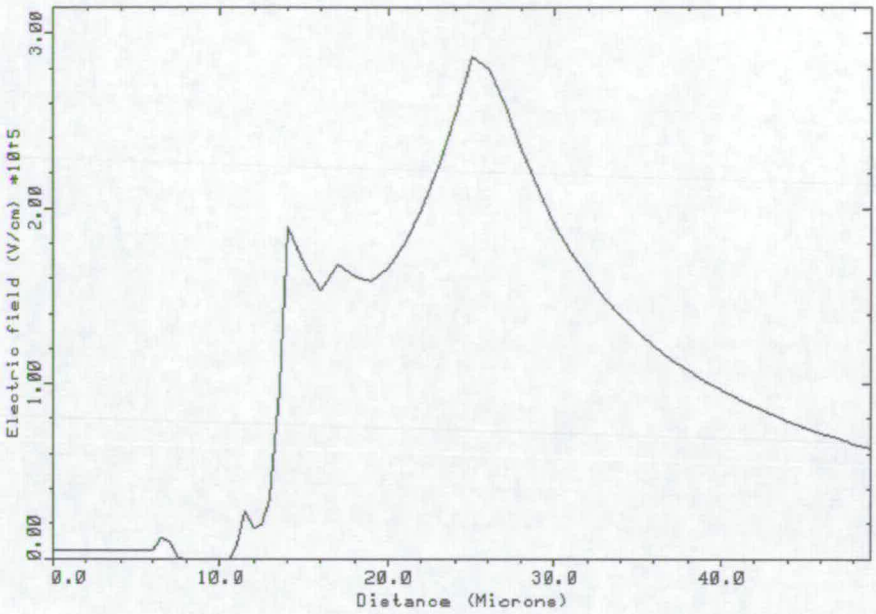


Figure 3.16 Simulated electric field across points AA' of RESURF DMOS transistor shown in figure 3.15

3.7 Summary

Chapter 3 has presented some of the device physics behind the breakdown of a pn junction and in particular the avalanche breakdown mechanism. The influence of a field plate to reduce the effects of junction curvature has been discussed highlighting how drift in breakdown can be observed when high electric fields are present in pn junctions utilising field plates. The RESURF principle of realising high voltages in relatively thin epitaxial layers has been reviewed along with the operation of the field plate through the use of a device simulator. The basic process used for fabricating a RESURF device will be presented in chapter 4, with process and device optimisation discussed in chapters 5 and 6.

References

1. D.J. Roulston, "Bipolar Semiconductor Devices", McGraw-Hill, 1990.
2. R. van Overstraeten, H. de Man, "Measurement of the ionization rate in diffused silicon pn junctions", Vol. 13, pp 583-608, May 1970.
3. A.G. Chynowith, "Ionization rates for electrons and holes in silicon", Phys Rev, Vol. 109, pp 1537-1540.
4. Sze, "Physics of Semiconductor Devices" 2nd edition, John Wiley, 1981.
5. C.T. Sah, R.N. Noyce, W. Schokley. "Carrier generation and recombination in pn junctions and pn junction characteristics", Proc. IRE, Vol. 45, pp 1228-1243, September 1957.
6. J. Conradi, "Temperature effects in silicon avalanche diodes", Solid State Electronics, Vol. 17, pp 99–106, 1974.
7. A.S. Grove, O. Eeistiko, W.W. Hooper, "Effect of surface fields on the breakdown voltage of planar silicon pn junctions", IEEE Trans. Electron Devices, ED-14, pp 157-162, 1967.
8. A.B. Phillips, "Transistor Engineering", McGraw-Hill, 1962.
9. F. Conti, M.Conti, "Surface breakdown in silicon planar diodes equipped with field plate", Solid State Electronics, Vol 15, pp 93–105, 1972.
10. V.P. O'Neil, PG. Alonas, "Relation between oxide thickness and the breakdown voltage of a planar junction with field plate electrode", IEE Trans Electron Devices, Vol. ED-26, No. 7, July 1979.
11. S.M. Sze, G. Gibbons, Solid state Electron, Vol 9, p 831, 1966.
12. A. Rusu, C. Bulucea, "Deep depletion breakdown voltage of silicon-dioxide/silicon MOS capacitors", IEEE Trans on Electron Devices, Vol. ED-26, No. 3, March 1979.

13. W. Feiler, W. Gerlach, "The reverse voltage limit of planar pn junctions with field plates", Archiv fur Elektrotechnik 75, pp 223–229, 1992.

14. R.C. Bassus, E. Falch, W. Gerlach, "Application of the evolution strategy to optimise multistep field plates for high voltage planar pn junctions", Archiv fur Elektrotechnik 75, pp 345–349, 1992.

15. E.H. Nicollian, A. Goetzberger, C.N. Berglund, App Physics Letters, 174, 1969.

16. R. Poirier, J. Olivier, Applied Physics Letters 15, 364, 1969.

17. D.J. Bartlelink, J.L. Moll, N.I. Meyer, Physics Rev 130, 973, 1963.

18. J.A. Appels, M.G. Collet, P.A.H. Hart, H.M.J. Vaes, J.F.C.M. Verhoeven, "Thin layer high-voltage devices (RESURF devices)", Philips J. Research, Vol. 35, No.1, 1980.

Chapter 4: The JHVH Fabrication Process

4.1 Introduction

The Junction isolated High Voltage Hybrid (JHVH) fabrication process is a subset of the LMDMOS process flow [1,2] run in National Semiconductor's 6" wafer fabrication facility in Greenock. The LMDMOS process is a fully integrated smart power technology for both power and automotive applications such as switching regulators and anti-lock breaking systems. The device types available to the designer are NPN and PNP bipolar transistors, vertical and lateral DMOS transistors for power applications, and complementary MOSFET transistors, resistors and diodes. The main process features of the LMDMOS technology are:

1. 18 masking steps
2. 12μm epitaxy thickness
3. Dual layer metal
4. Single layer polysilicon
5. Metal pitch (line+space) 6μm for the first layer metal and 12μm for level two metal
6. 2.5μm minimum geometry for polysilicon contacts

A requirement of the JHVH process is for it to be integrated into the LMDMOS flow so that the increased functionality can be used as an addition to the portfolio of devices produced by the LMDMOS process. The JHVH flow itself differs from LMDMOS in that it has fewer masking layers; 11 compared to 18 and requires only one metal layer. It uses the second layer metal as it is thicker than level one metal and can, therefore, carry the greater currents required by power transistors. All other process features remain the same.

4.2 Semiconductor Processing

The main process steps used in semiconductor manufacturing are photolithography, implantation, diffusion and deposition. These are used in combination to define dopant distributions within the silicon thereby making up the devices themselves. The following discussion concentrates on the processes used for fabricating the devices which are the subject of this thesis. For details on more advanced processes the reader is referred to references [18-23].

4.2.1 Photolithography

Photolithography is central to the whole fabrication process: transferring patterns from the photo-mask onto the surface of the wafer. During the design stage, alignment tolerances must be taken into account in order to produce good working die.

The photolithography process flow is shown in figure 4.1 [3]. The wafer prime process promotes good adhesion of the photoresist (resist) to the underlying layer. The resist is then applied using a spin track and for the JHVH process, the thickness of the resist is 1.25μm and the bake process is used to remove solvents in the resist. The align, expose, develop, and rinse sequence are the processing steps that transfer the pattern. Variations in masks, device layout, and photo process must be taken into account during the design in order to minimise the effects of any process variation. Projection printing is used at Greenock and this has the advantage that it avoids any contact of the mask plates with the wafer, hence providing an economical way of extending mask lifetime. The masks also have a pellicle (protective cover) attached to them which ensures that any particles that are on the pellicle are not imaged on the

wafer as they are out of focus. A mercury vapor lamp in the ultra violet (UV) light region performs the exposure. Following the exposure the resist is developed and rinsed. Wafers are then inspected and at critical layers CDs (critical dimensions) are measured. Wafers can be reworked at this point if they do not meet the acceptance criteria: for example poor develop, poor spin, CDs out of specification. If a resist mask implant is to be performed (see section 4.2.2), a hard bake prior to the implant is used to drive off any residual solvents resulting in a more rapid implant cycle time (faster pump down in the implanter). Once the photoresist has been developed, the etch process follows. This can be either a wet or dry process depending on the technology. For the technology under consideration, the etches are predominantly wet. A resist strip followed by a final inspection completes the photolithography process.

4.2.2 Implantation

In semiconductor manufacturing, ion implanters are used to introduce the dopant into the semiconductor substrate [3,4]. This is an alternative to high temperature diffusion processing which is used in older technologies. During the implantation a beam of dopant ions is accelerated to energies ranging from a few keVs to several MeVs into the surface of the semiconductor. Depending upon the impurity and energy chosen the depth of the implanted dopant distribution can vary from a few hundred angstroms to several microns.

A block diagram of an implanter is shown in figure 4.2. A gas or vapour containing the required dopant is ionised in the source chamber. The extraction process then

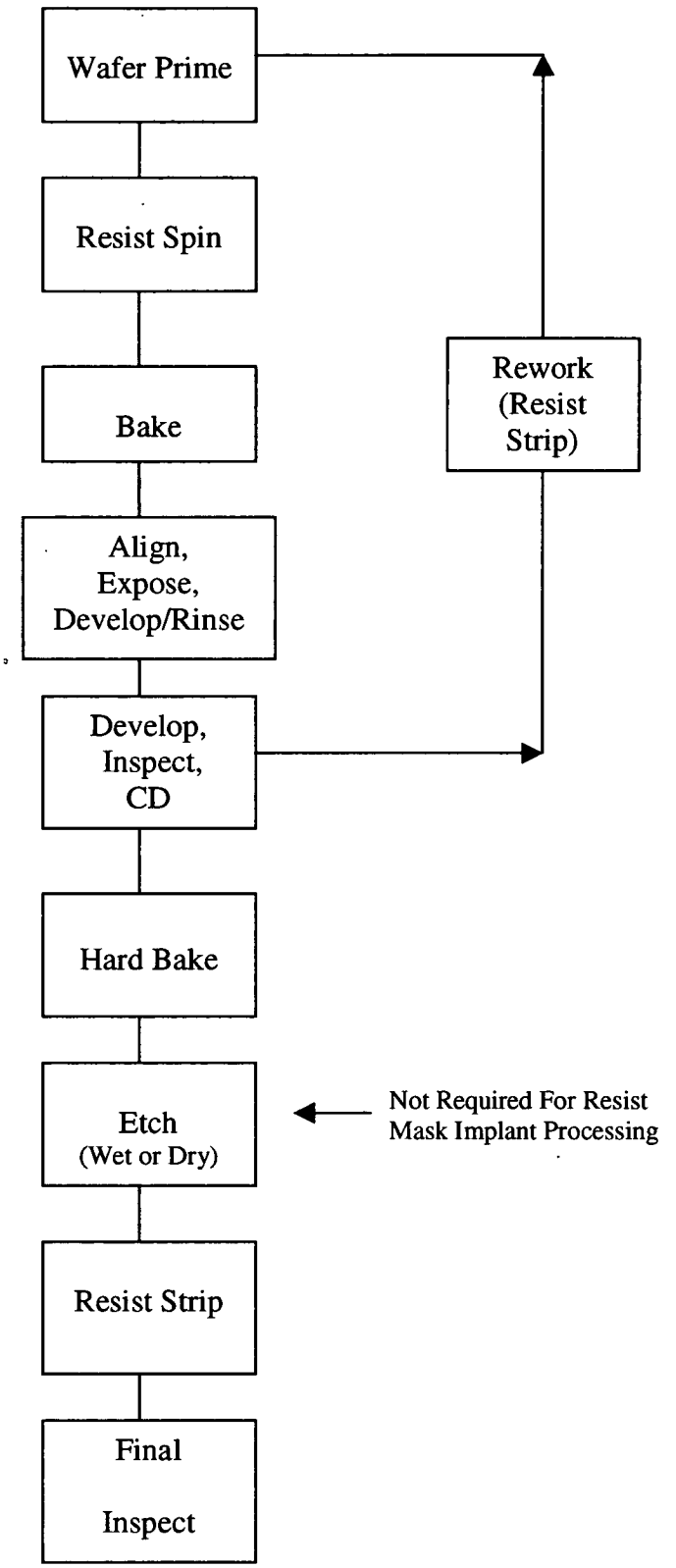


Figure 4.1 Photolithography Flow Chart

accelerates these ions which pass through a mass separator to select only the required species. Further acceleration is then applied after which the ions enter the target chamber.

For lower doses, repeated raster scanning provides good dopant uniformity across a wafer. It should be noted that high current implanters normally use mechanical scan systems and the target chamber typically has automatic wafer handling capability in order to facilitate a good cycle time in a production environment.

The main advantage of implantation compared to furnace doping is that it can be performed at low temperatures. Therefore, doping of layers can be undertaken without adversely affecting previously diffused regions. Also the quantity of dopant introduced into the silicon substrate can be very well controlled to achieve the desired sheet resistances. Films, including photoresist (see section 4.2.1) and oxides, can be used to selectively block the implant.

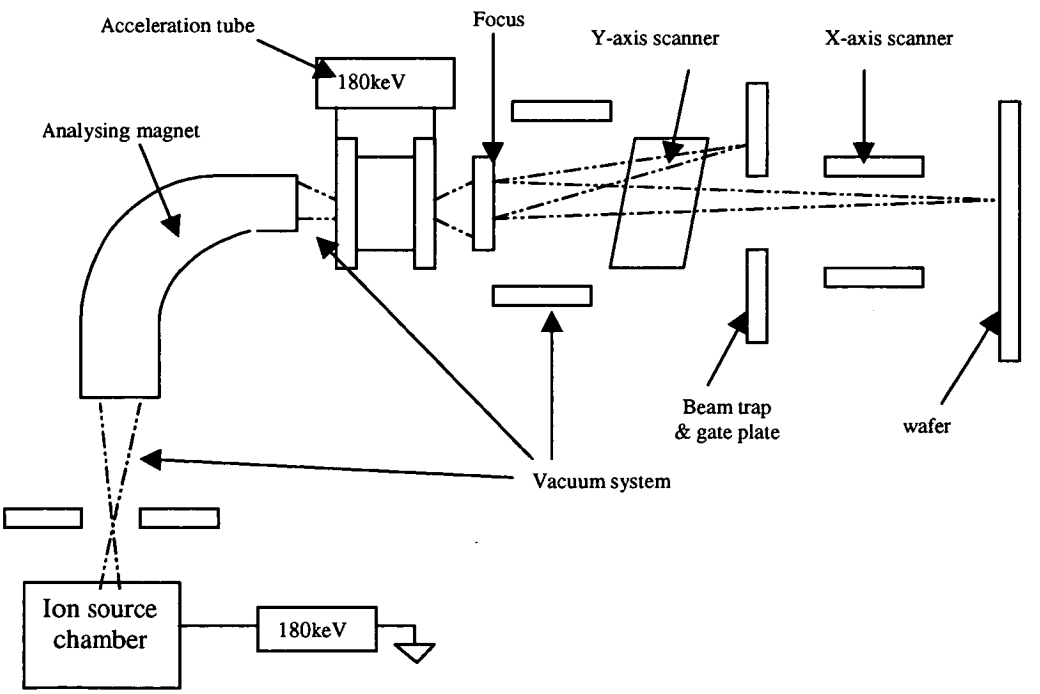


Figure 4.2 Simplified block diagram of an ion implanter

A downside to implantation is the damage caused to the crystal lattice by the collisions between the incident ions and the lattice atoms. Most of this damage can, however, be removed by using a thermal process that anneals the silicon.

Implants are frequently performed through oxides (see section 4.3) in order to minimise the effects of channeling [3], which causes a fraction of the implanted ions to travel a greater distance into the silicon which obviously impacts device performance. A standard method employed to minimise this effect involves tilting the wafer, typically 7°, which reduces the probability of ion channeling along principal crystalline axes.

4.2.3 Diffusion

Diffusion is a high temperature process, which redistributes implanted dopant atoms within the silicon. The dopant atoms are incorporated substitutionally into the silicon lattice [4] and create extrinsic p-type or n-type layers. Dopant atoms diffuse at temperatures greater than 800°C and the rate of movement of the atoms increases with temperature. In large geometry high voltage applications there is no significant diffusion of dopant atoms below 650°C and so for normal device operating temperatures the dopant profiles are fixed.

The diffusion process is typically performed in a furnace where the temperature can be accurately controlled. Great care and attention must be paid when introducing the wafers into the high temperature environment and during the ramp down of the wafers to room temperature, as this can result in stress induced defects [5]. Fortunately modern computer controlled furnaces minimise any temperature gradients across the wafers.

A typical diffusion sequence is as follows:

1. Diffusion clean: The purpose of this is to remove any contaminants and a typical clean is detailed in reference [3]. The wafers are then rinsed in de-ionised water and dipped in a dilute hydrofluoric acid in order to remove any native oxide. This is followed by another rinse in de-ionised water.
2. The wafers are then automatically loaded into the furnace, and the temperature is ramped to the required value with the gases automatically programmed to switch on for predefined times and flow rates. On completion the wafers are removed from the furnace once the temperature has been ramped down.
3. After the wafers have cooled down, parameters such as oxide thickness and resistivity are typically measured to ensure that the required parametrics are within specification.

4.2.4 Deposition

An important method of layer deposition in semiconductor manufacturing is chemical vapour deposition (CVD) for which all the constituents forming the deposited layer are brought into the process equipment as gases. CVD occurs over a wide temperature range (300 to 1200°C) depending upon the layer to be deposited. The energy for the reaction is usually facilitated by heating the wafer [3] and in some cases providing energy through the formation of a plasma [6]. The layer deposited by CVD results from a gas dissociating or two gases reacting on contact with a hot surface. Any bi-products must be gaseous otherwise they will be incorporated into the deposited film

Some examples of processes which use chemical vapor deposition are:

1. Single crystal epitaxial silicon: The integrated circuit is formed on the epitaxial layer. This is a particularly important process in bipolar and smart power technologies.
2. Polycrystalline silicon (polysilicon): This is typically used to form the gate electrode in MOS integrated circuits and emitters in bipolar technologies.
3. Silicon nitride: This prevents oxidation in specific areas where thick oxides are grown for isolation – LOCOS (local oxidation). This can also be used to deposit silicon nitride as the final passivation layer over the completed integrated circuit.

4.3 The JHVH Process

A complete summary of the major process steps in the JHVH process is detailed in table 4.1.

No cleaning or resist strips have been included in this table but key important cleans and strips are mentioned in the following description.

A cross-section of a completed LDMOS transistor configured for a low side switch is shown in figure 4.3. In such applications the source is tied to the substrate and the breakdown of the source to the drain is optimised with the source shorted to the substrate, BV_{dss}.

Step No.	Step Type	Details
1	Substrate	Resistivity 22-32 Ω-cm, <100>, P-type, 150mm diameter
2	Initial Oxidation	Masking Oxide of 8000Å
3	Lithography	Isolation Up Mask
4	Up Isolation Implant	Boron, $1x10^{14}$ cm^{-2}, 150keV
5	Isolation Up Drive/Ox	Repair lattice and activate implant
6	Oxide Etch	Remove Masking Oxide
7	Epitaxy	Silicon Tetrachloride, 11-13µm, Resistivity 10-15Ω-cm
8	Drift Implant oxidation	700Å Thick
9	Drift implant	Phosphorus, $1.2x10^{12}$ cm^{-2}, 120keV
10	Epitaxy Oxidation	Masking Oxide of 8000Å
11	Lithography	P-Well Mask
12	Implant Oxidation	Implant Screen Oxide of 700Å
13	P-Well Implant	Boron, $1x10^{15}$ cm^{-2}, 30keV
14	P-Well Drive	Repair lattice and activate implant
15	Oxide Strip	Remove Masking Oxide
16	Pad Oxidation	1000Å Thick
17	Nitride Deposition	1500Å Thick
18	Lithography	Active Area Mask
19	Field Oxidation	1.3µm Thick
20	ONO Etch	Etch Oxide, Nitride, Oxide
21	Gate Oxidation	600Å Thick
22	Poly Deposition	5500Å Thick
23	Poly Dope	Dope Polysilicon
24	Lithography	Poly Mask
25	P-Body Implant	Boron, $3.4x10^{13}$ cm^{-2}, 80keV
26	P-Body Drive	1400Å Thick
27	Lithography	N+ Source Mask
28	N+ Source Implant	Arsenic, $5x10^{15}$ cm^{-2}, 50keV
29	N+ Source Drive	Activate and Drive Implant
30	Lithography	Shallow P+ Mask
31	Shallow P+ Implant	Boron, $5x10^{14}$ cm^{-2}, 50keV
32	Field Vapox	8500Å Thick
33	Oxide Reflow	1400Å Thick
34	Lithography	Contact Mask
35	Metal Deposition	2µm Thick
36	Lithography	Metal Mask
37	Vapox Deposition	5500Å Thick
38	Lithography	Pad 1 Mask
39	Nitride Deposition	1µm Thick
40	Lithography	Pad 2 Mask

Table 4.1 Major Process Steps in JHVH Fabrication Process Flow

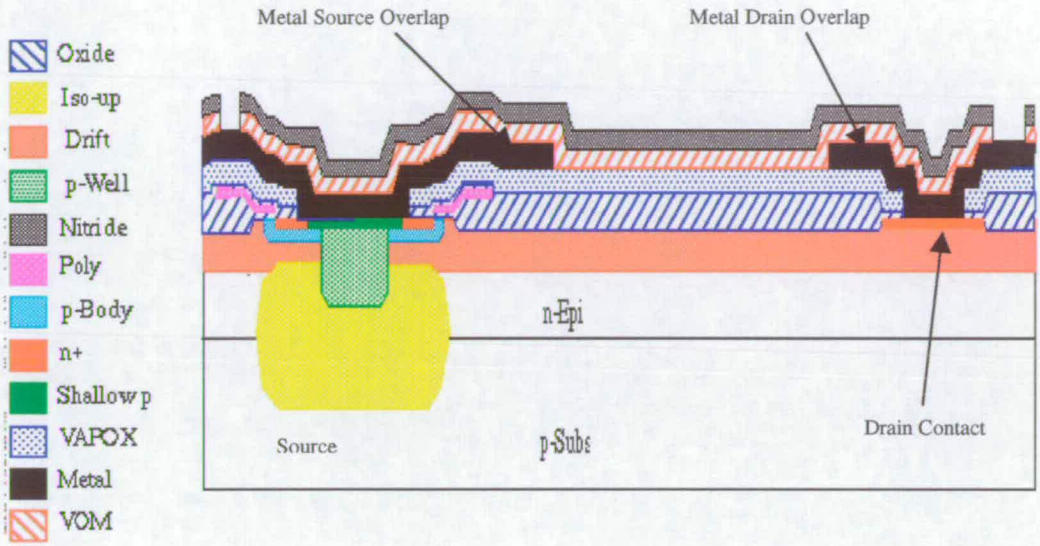


Figure 4.3 Cross-section of a low side lateral DMOS

As discussed in chapter 3, the interaction of the depletion layer in the n- epi and p substrate with the drain-body depletion layer spreads the total drain source potential uniformly along the drift region. If the doping is too low this causes premature breakdown at the drain; too high a level of epitaxial doping causes premature breakdown near the source [7]. In a well-designed RESURF device, only the plane breakdown voltage of the n+/n/p diode under the drain contact limits BV_{dss}.

The buried p-layer under the source of the device in figure 4.1 prevents premature breakdown at the source end, and also shorts the source to the substrate, making the transistor an exclusively low side device. Removing the buried p-layer makes high side operation possible, but reduces BV_{dss}.

The starting material used for the JHVH process is 150mm (6") p-type, <100> with a resistivity 22-32 Ω-cm. The <100> crystal orientation is standard for MOS processes

at National Semiconductor UK Ltd., and is used because it results in the lowest number of surface states at the Si-SiO_2 interface. The masking sequence for this flow is shown in table 4.2. The first process operation is to grow an oxide which will be used as a masking layer. This layer is lithographically patterned and etched to define areas where boron will be implanted ($1x10^{14}cm^{-2}$, 50keV) to form the up isolation region used to connect the substrate to the source of the low side lateral RESURF DMOS transistor. After the implant a high temperature oxidation is used to anneal out the crystal damage caused by the implant as well as drive in and activate the dopant. The resulting cross section of the device after this process stage is shown in figure 4.4.

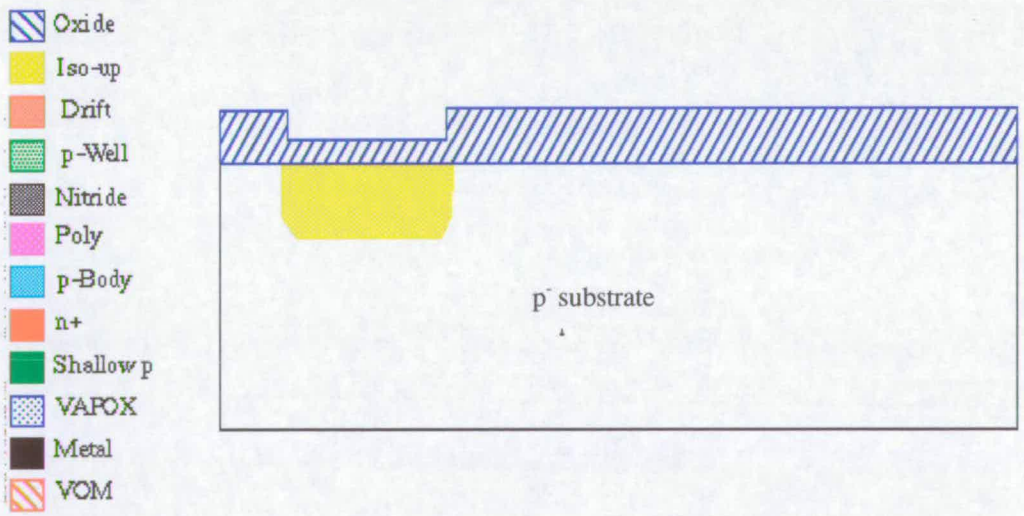


Figure 4.4 Device cross section after up isolation implant and drive

The next step is to etch all the oxide off, and grow an epitaxial layer. One micron of intrinsic epitaxy is grown prior to activating any dopant in the epitaxy reactor. This is to minimise any out diffusion and auto doping of boron from the up isolation implant

Alignment Sequence By Mask Name

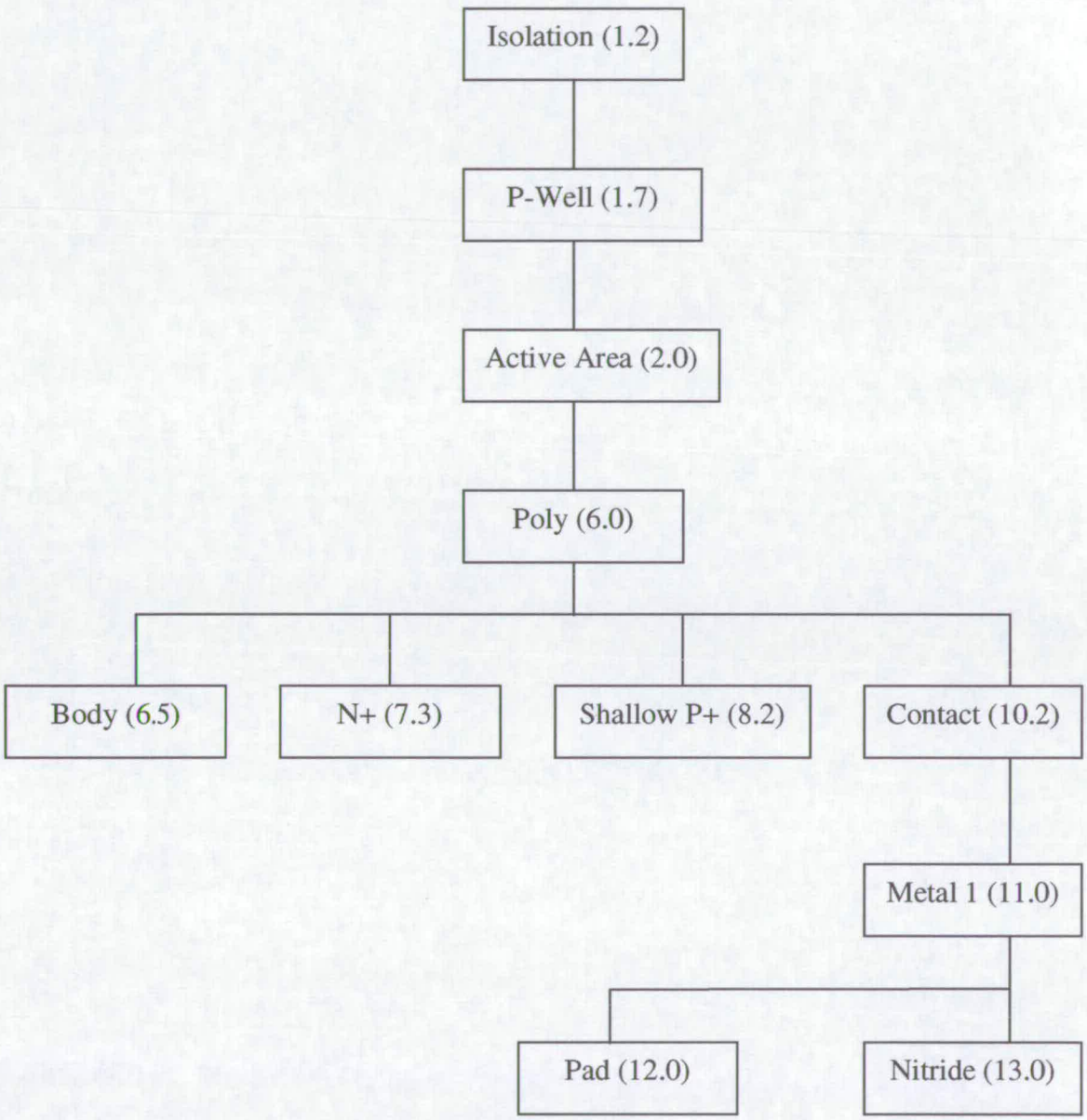


Table 4.2 Alignment Sequence by Mask Name for the JHVH Fabrication Process

during this high temperature process. Should this occur, it could have a detrimental effect on the electrical parameters of other device types as well as the RESURF device itself. Next a thin oxide is grown, after the epitaxial process, which is used as a

mask for the drift implant. This thermal process also diffuses iso-up further into the epitaxial region. The drift implant now follows and is used to control the doping in the epitaxial region so that the RESURF device operates as designed. After this a field oxide is grown which is later used as a mask for the p-well. A cross section of the device at this stage is shown in figure 4.5

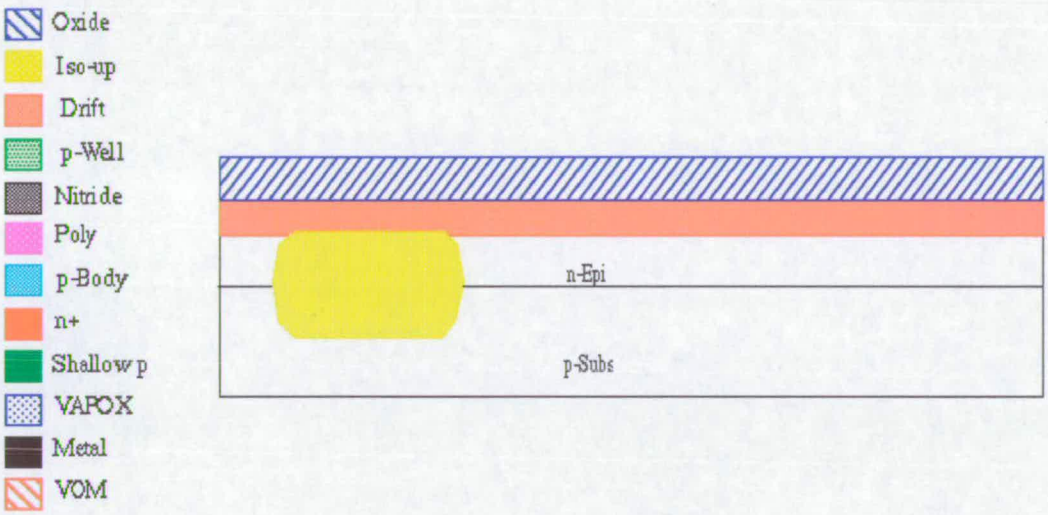


Figure 4.5 Post drift implant and field oxidation

The areas for the p-well are then defined lithographically and oxide etched to define regions for the boron implant. A thin screen oxide is grown to position the peak of the implant as near to the silicon surface as possible and also to reduce channelling effects (section 4.2.2). Following the boron implant, a high temperature diffusion step is then used to drive the implant and the resulting cross section is shown in figure 4.6.

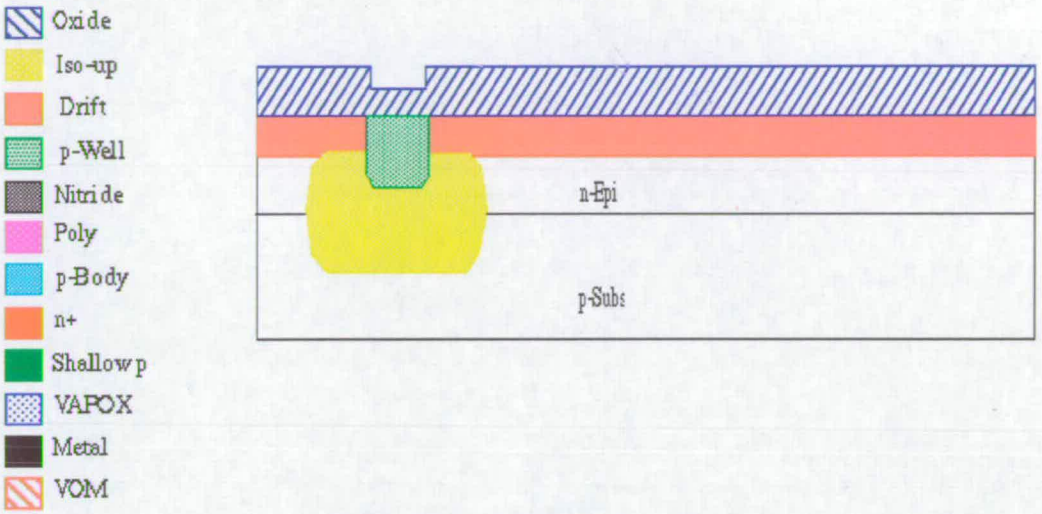


Figure 4.6 Post p-well drive

After stripping the oxide used for the implant mask, a pad oxide is grown followed by the deposition of a layer of silicon nitride. The purpose of the oxide is for stress relief while the nitride is used to prevent the oxidation of certain areas. Lithography is then used to define the oxide and nitride which is then etched to leave island regions in the silicon where the devices will be fabricated. After field oxidation, the resulting cross section is shown in figure 4.7

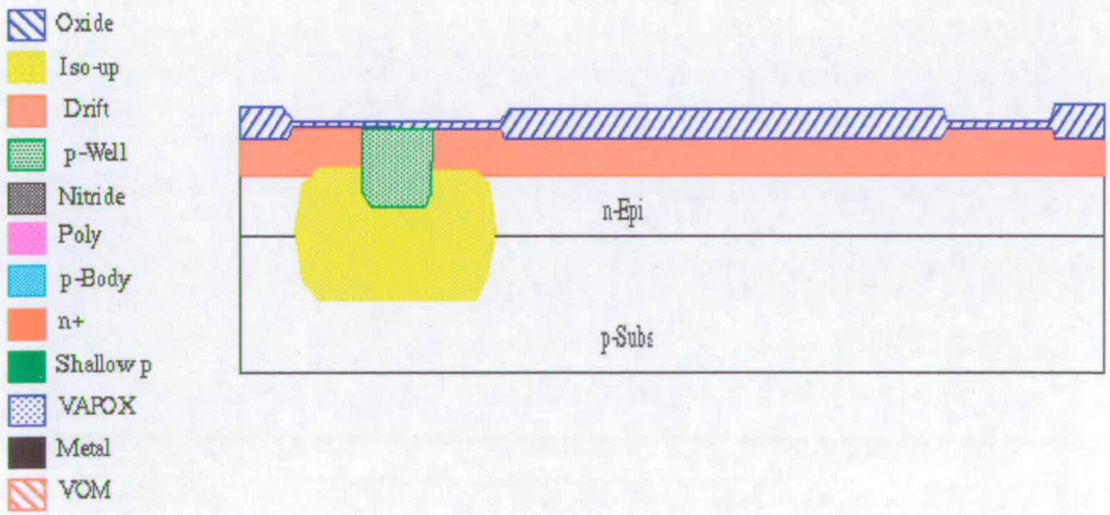


Figure 4.7 Post field oxidation

The nitride and stress relieving oxide are then removed to expose the active area to further processing. The thick field oxide is also etched but the etch time is such that only a small amount is consumed.

The wafer is now thoroughly cleaned in order to prepare for the gate oxidation process which is a critical process stage. Prior to growth of the gate oxide, a sacrificial oxide is grown then stripped to clean up the silicon surface. Another clean is then performed to ensure that all organics and other materials are removed. Dichlroethene (DCE) is used during the oxidation to facilitate the reduction of sodium, which has the potential to cause parametric problems, such as Vt drift, and a reduction in pn junction breakdown voltage. In fact, all furnace processes typically use DCE to reduce sodium content.

Next polycrystalline silicon (poly) is deposited to form the gate of the transistor and then doped with phosphorus using a pre-deposition process. This reduces the sheet resistivity of the poly effectively making it conductive (typically 8Ω/). After this, a deglaze is performed to remove the heavily doped oxide resulting from the pre-deposition stage. There then follows photolithography and etch steps to define the gate.

The next process step is known as the P-body implant. Photoresist is used as the implant mask and after ion implant the resist is stripped off and a high temperature drive is performed which also activates the boron. The n^+ source areas are defined lithographically and the oxide grown during the drive-in stages is etched to form the arsenic implant mask. After implantation, an oxide etch back is now performed to

remove the heavily doped and damaged oxide layer and the poly is then re-oxidised to help to protect the gate region as well as cover all exposed silicon with thermal oxide resulting in the cross section shown in figure 4.8.

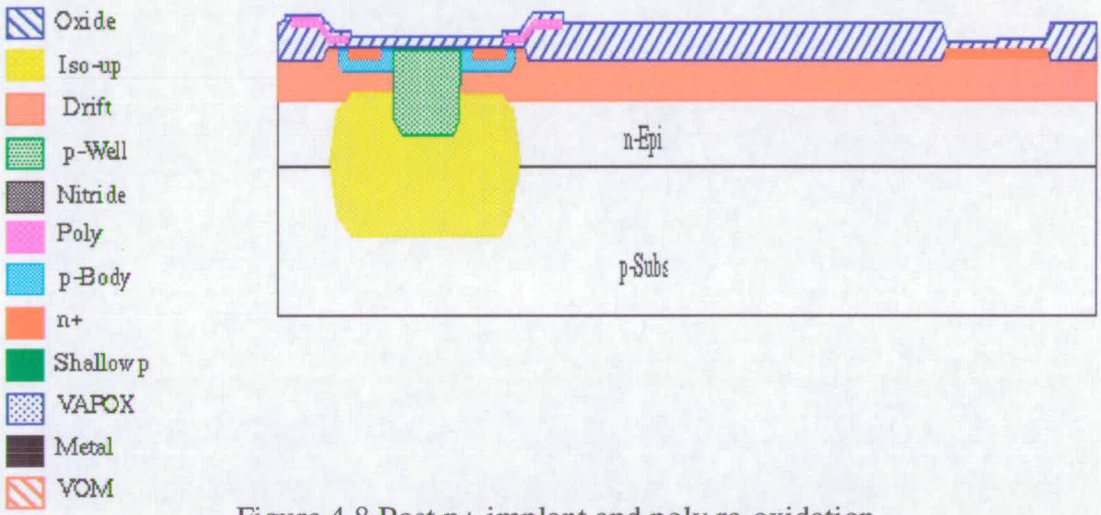


Figure 4.8 Post n+ implant and poly re-oxidation

Following the poly re-oxidation process the areas for shallow p^+ regions are lithographically defined to form the resist mask for the p^+ implant. A thick oxide (vapox) is then deposited which is followed by a re-flow process that densifies the film. This layer is doped with phosphorus which acts as a getter for mobile ions in preparation for connecting up the devices via metallisation.

The next process operation cuts contact holes in the resulting thick oxide as shown in figure 4.9.

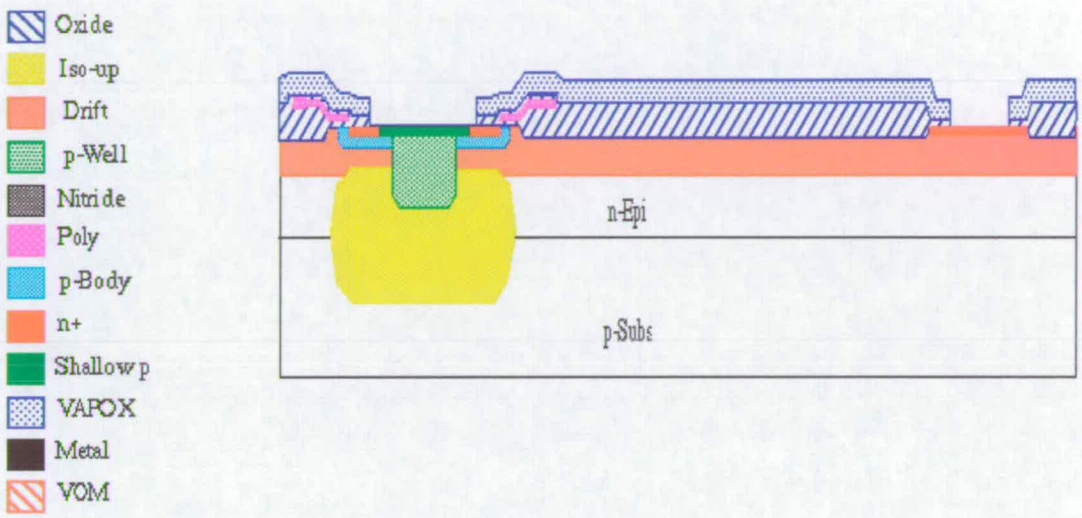


Figure 4.9 Post contact mask

Metal is then deposited and patterned to connect the device to the outside world. In order to protect the circuitry from moisture, scratches and metallics that could impact performance and reliability, protective layers are now deposited. This consists of a thick oxide layer which is etched to expose the aluminium pads. The above process is repeated for a thick layer of nitride to result in the cross section is shown in figure 4.10.

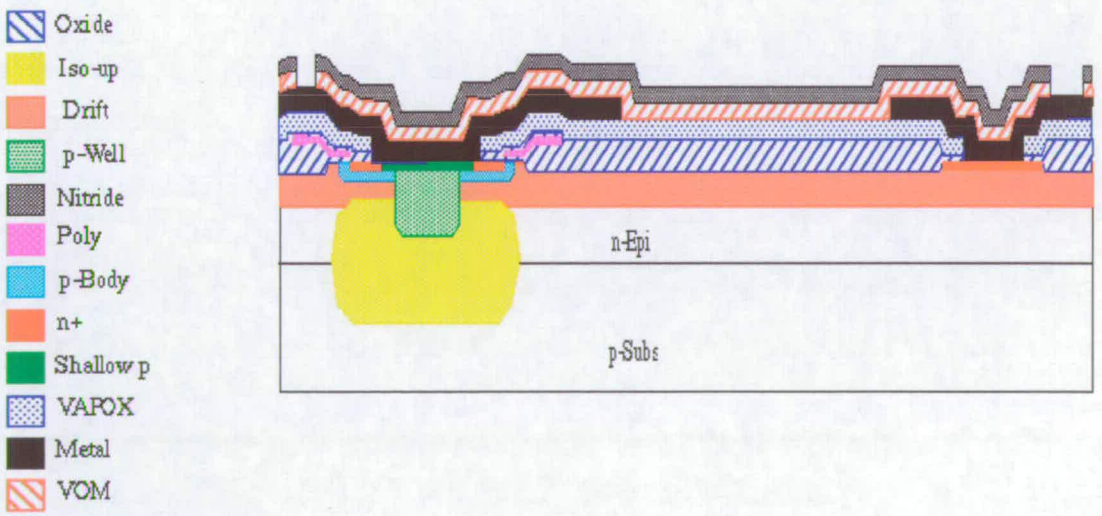


Figure 4.10 Device cross section at the completion of the process

4.4 Process Challenges

There are many challenges that occur in the fabrication of a semiconductor process and almost all have a direct impact on the performance of the resultant devices. These can range from process repeatability (the capability to continually produce identical products day in day out) to reducing the number of particulate defects that land on the wafer.

Procedures can be put in place to monitor many key parts of the fabrication process that detect quality or process control issues [16]. Such procedures can be achieved by both visual inspection and measurement methods [17]. For example, during the photolithography process (section 4.2.1.) visual inspection is performed after develop, (*Develop Inspect)* and at the end of the process, (*Final Inspect).* As indicated in section 4.2.1 photolithography is a key component of the whole fabrication process. Hence, inspections to check registration error, poor develop, etc is crucial and for more advanced VLSI technologies, automated inspection systems are employed. Other parametrics can be measured electrically [8,9] and together can be used to assist in yield analysis and process enhancement projects.

Analysis techniques such as scanning electron microscopy (SEM) [10], scanning transmission electron microscopy (STEM) [11], secondary ion mass spectroscopy (SIMS) [12] and spreading resistance profiling (SRP) [12] are also used to assist in the determination of process related issues which impact device performance.

4.5 Summary

This chapter has reviewed many of the processing steps used in the manufacture of semiconductor devices and provides the background information for the following chapters. The JHVH process flow has been described using cross sections of the RESURF device. Some of the manufacturing challenges have been briefly discussed as have some of the monitors, and failure analysis techniques that can be used to facilitate device debug.

References

1. M. Redford, M. Fallon, J. McGinty, A.J. Walton, M.I. Newsam, G. Michalowicz, C. McGregor, "Identifying and quantifying variability within a BiCMOS process : A case study", IEE Colloquium Improving the Efficiency of IC Manufacturing Technology, No 1995/153 pp10/1 – 10/4, Edinburgh 12th April, 1995.
2. J. McGinty, G. Long, J.K. Kibarian, M. Fallon, P. McKernan, M. Redford, "Implementation of statistical TCAD in a manufacturing environment", National Semiconductor Technology and Innovation Conference, pp 230-233, 1996.
3. Technology Associates, "Semiconductor Technology Handbook", Jan 1985.
4. B.G. Streetman, "Solid State Electronic Devices", Prentice Hall, 1980.
5. M. Redford, "Yield analysis of the DS3487", NSC Internal report, Nov 1985.
6. D.K. Reinhard, "Introduction to Integrated Circuit Engineering", Houghton Mifflin, 1987.
7. E.M.S. Narayanan, G. Amaratunga, W.I. Milne, "A study of the RESURF principle for thin epitaxial layer high voltage integrated circuits", Proc 1992 International symposium on power semiconductor devices, pp172-175.
8. K. Findlater, M. Fallon, M. Redford, "Investigation and application of the buried layer van der pauw and bar resistors", Proc Microelectronic manufacturing technologies, pp 249-255, May 1999.
9. M. Fallon, M. Redford, K. Findlater, M.I. Newsam, "Novel structure to measure emitter-base misalignment", Proc. IEEE International Conference on microelectronic test structures, Vol 10, pp 156-158, March 1997.
10. A. McGibbon, C. McLeod, G. Robertson, M. Redford, K. Yallup, C. Buchanan, P. Cavanaugh, C. Elgey, "The detection and analysis of platinum silicide spiking in

contact structures", National Semiconductor Technology and Innovation conference, pp200-203, September 1996.

11. A. McGibbon, R. Boyle, M. Redford, "Role of STEM as a high resolution failure analysis tool for bipolar technologies", Proc Microelectronic manufacturing technologies, pp 219-226, May 1999.
12. A. McGibbon, K. Yallup, J. Altrip, M. Harper, M. Redford, "A fingerprint methodology for process transfer and dual sourcing", National Semiconductor Technology and Innovation conference, pp 171-174, July 1997.
13. T. Yamaguchi, S. Morimoto, "Process and device design of a 1000V MOS IC", IEEE Trans Electron devices, Vol ED-29, pp 1171-1178, Aug 1982.
14. B.G. Streetman, "Solid State Electronic Devices", 2nd edition, Prentice-Hall, 1980.
15. Private communication, R. Busse.
16. "6 inch Develop/Final inspect masking specification", National Semiconductor (UK) Ltd, Greenock, Scotland.
17. "6 inch Electrical test specification", National Semiconductor (UK) Ltd, Greenock, Scotland.
18. L. Deferm, G. Badenes, "Lateral isolation", IMEC Scientific Report, pp 51-53, 1997.
19. X. Ren, M.C. Ozturk, D.T. Grider, M. Sanganeria, S. Ashburn, "A novel implantation free raised source/drain MOSEFET process using selective thermal chemical vapour deposition of in-situ boron doped SiGe", Materials Research Society Symposium Proceedings, Vol 303, p 37, 1993.
20. L. Van den hove, S. Vanhaelemeersch, "Patterning for 0.18 micron and below". IMEC Annual Research Review Meeting, November 1998.

21. K. Maex et al., "Silicides for 0.13μm and beyond", Advanced Research Review Meeting, IMEC, Leuven, Belgium, October 27-28 1999.

22. S. Arnauts et al., "Ultra clean processing, High-k dielectrics, epitaxial materials, (PE)CVD, ES&H and CMP", Advanced Research Review Meeting, IMEC, Leuven, Belgium, October 27-28 1999.

23. X.W. Lin, D. Pramanik, "Future interconnect technologies and copper metalization", Semiconductor International, October 1998.

Chapter 5: JHVH-450 Process Development

5.1 Introduction

National Semiconductor's first high voltage lateral DMOS process was a single metal variant of the LMDMOS 300B smart power process fabricated in the 6" line in Greenock. This is an 18-mask (3μm minimum geometry) process using projection alignment lithography. Not all the process steps in 300B are required for the lateral DMOS device and so some of these have been eliminated, to result in the less complex, discrete 400V DMOS process described in chapter 4.

There are power supply applications that require devices [1-3] with a breakdown source to substrate, BV_{dss}, greater than 400 volts. In a single ended switch mode power supply, the high voltage transistors experience a peak V_{ds} of twice the rectified AC voltage, plus a voltage spike generated at turn-off. For a 110 VAC, the minimum acceptable BV_{dss} is, therefore, around 400 volts.

In this chapter, two process options will be studied using process and device simulation tools combined with design of experiment tools [4] in order to determine the optimum process conditions. These will be used as examples to demonstrate the manufacturability of the two process options by predicting the breakdown voltage distribution and comparing the results with those obtained from silicon.

5.2 Strategy

High breakdown voltages in relatively thin epitaxial silicon are made possible by using the RESURF (Reduced SURFace Field) principle described in chapter 3. It was required that this device could be integrated into the LMDMOS process, and so the constraints were that the epitaxial thickness had to remain the same and existing

process steps must, whenever possible, be used. It was desired that as few steps should be added that would adversely impact other device parametrics when the high voltage device was integrated into the smart power process flow. Previous work [5] had already indicated the main process variables that control the breakdown voltage were epitaxial thickness, epitaxial doping, and substrate doping.

The strategy used is shown in figure 5.1 where TCAD tools, and process and device simulation is used in conjunction with designed experiments to generate response surfaces for the breakdown voltage of the RESURF device. Parametric distributions could then be generated using Calphurnia [6] as described in chapter 2.

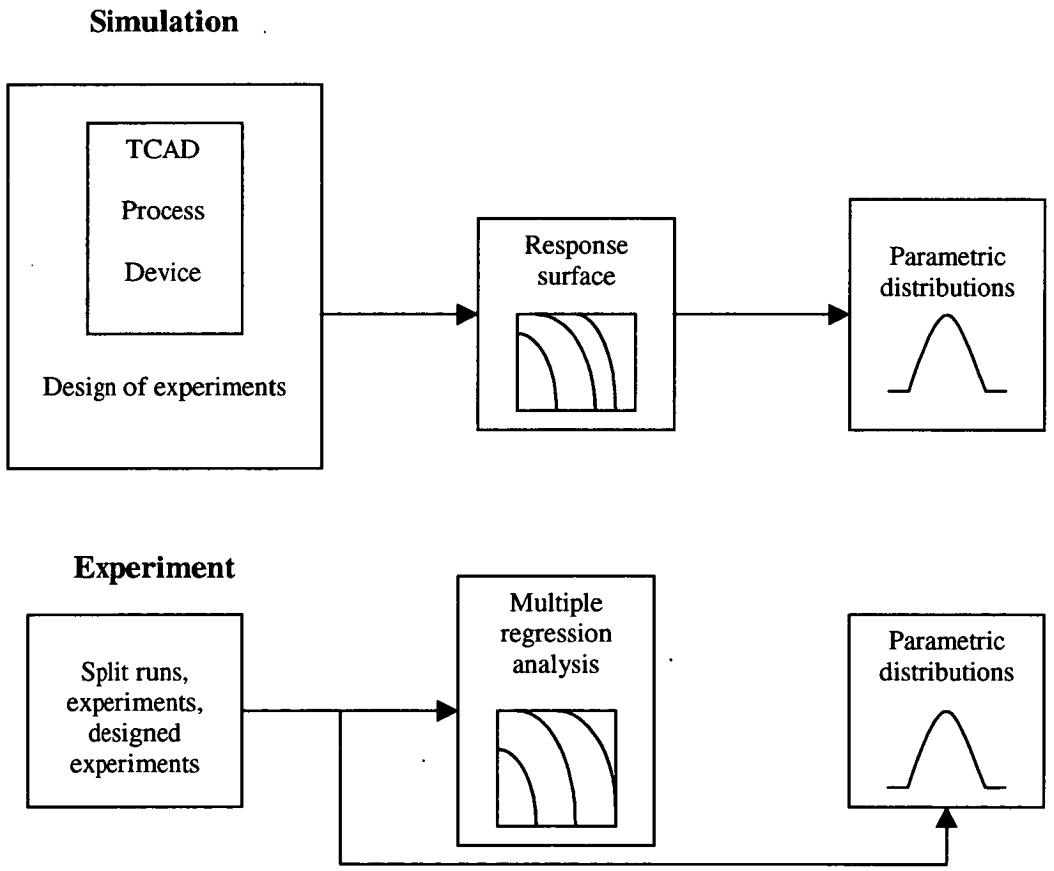


Figure 5.1 Design for manufacturability strategy

The data from silicon split runs and experiments were analysed using multiple regression analysis and the response surfaces compared. The simulated distributions were also compared with measurement from silicon.

5.3 Process Options

The concentration in the epitaxial region is an important control factor for the breakdown voltage. There were two process options available, each with its advantages and disadvantages. The first option uses the epitaxy process itself to set the concentration in the drift region. The second uses a higher resistivity epitaxial process with the concentration in the drift region being set with an additional implant step [7]. The epitaxy process is the cheaper option simply because it has fewer process stages but the implant option has potentially better control on the concentration in the drift region and should be more repeatable. Both options were evaluated and assessment of process repeatability and controllability made.

5.4 Review of Process Simulation Deck

5.4.1 Process Simulator

The process flow was modelled using the one dimensional simulator SUPREM-3 as the execution time is much faster than a two dimensional simulator. The process was based on the LMDMOS 300B calibrated process model [8] and the initial process specification based on [9]. This provided some confidence in the simulation tools used, as the LMDMOS model had been previously successful in predicting parametric distributions and solving technical issues.

The information provided by SUPREM-3 consists of the layers of material in the structure modelled and the distribution of impurities within those layers. For the JHVH-450 process, the list of default coefficients provided with the simulator were used. The impurity profiles resulting from the simulations were later used as inputs to the device simulator as described in section 5.5.

Process Statements

The major process steps available in SUPREM-3 are diffusion and deposition. Figure 5.2 shows the furnace profile for the JHVH-450 initial oxidation process. The corresponding SUPREM-3 code for this furnace is shown in figure 5.3.

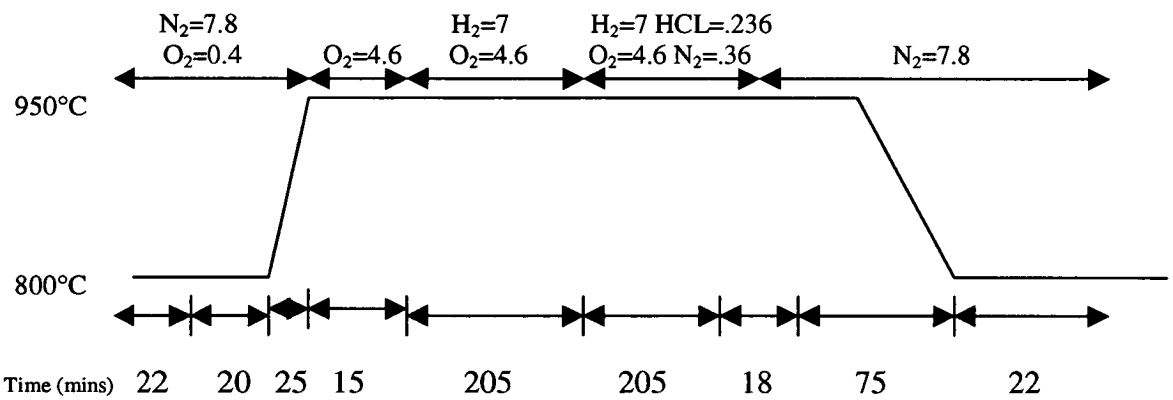


Figure 5.2 Furnace profile for initial oxidation

```
$ Initial Oxidation

assign name=gas1 n.value=0.829

Diffusion temp=800 time=22  press=@gas1 f.n2=7.8 f.o2=0.4
Diffusion temp=800 time=20  press=@gas1 f.n2=7.8 f.o2=0.4
Diffusion temp=800 time=25  press=@gas1 f.n2=7.8 f.o2=0.4 t.final=950
Diffusion temp=950 time=15  press=@gas1 f.o2=4.6
Diffusion temp=950 time=205 press=@gas1 f.o2=4.6 f.h2=7
Diffusion temp=950 time=205 press=@gas1 f.o2=4.6 f.h2=7 f.hcl=0.236 f.n2=0.36
Diffusion temp=950 time=18  press=@gas1 f.n2=7.8
Diffusion temp=950 time=75  press=@gas1 f.n2=7.8 t.final=800
Diffusion temp=800 time=22  press=@gas1 f.h2=7.8
```

Figure 5.3 SUPREM-3 code for JHVH-450 initial oxidation profile shown in figure 5.2

An example of the SUPREM code for a deposition process is given below:

```
Deposition polysilicon temperature=615 thickness=0.55 pressure=4.7e-4
  + xdx=0.00 dx=0.0100 spaces=25
```

where in addition to the process parameters, other parameters are set which define the grid spacing and location within the deposited layer. The layers of the device are built up using a grid, which is one dimensional for SUPREM-3. The maximum number of grid points available in SUPREM-3 is 1000 and in the case of the deposition process, the grid has been controlled using four parameters: **thickness**, **xdx**, **dx**, and **spaces**. **xdx** sets the location of the nominal grid spacing **dx**, relative to the top surface. The number of grid spaces in the deposited layer was set to 25, and 5500Å of polysilicon deposited. This approach was adopted to concentrate the grid points in key areas that impact the device performance such as junctions and a key film thickness such as gate and screening oxides [10].

5.4.2 Process Information

The accuracy of any model is only as good as the information it contains. Great attention to detail is required in order to ensure the information in the process model is

as complete and accurate as possible. For example, cleans as well as native oxide can potentially impact out diffusion of dopant during subsequent diffusion, which leads to sheet resistivities being incorrect. Hence it is important that these factors are accounted for which is especially important [10] otherwise the dopant profiles are used for device simulation when they will fail to predict parametrics correctly.

This being the case, furnace profile information was checked against specifications and furnace control system outputs as the actual thermal budget experienced in realty does vary over the furnace tube (source, centre and handle) as indicated in figure 5.4. In addition, temperature overshoots are also experienced after temperature ramps. The latter was subsequently automated by Fallon et al [11] so that furnace profiles could be automatically downloaded into SUPREM format. This also enabled the source, centre and handle regions of the furnace to be analysed in order to determine that sources of variation are fluctuations of the position along the furnace tube and the resulting impact on device parametrics such as sheet resistivities.

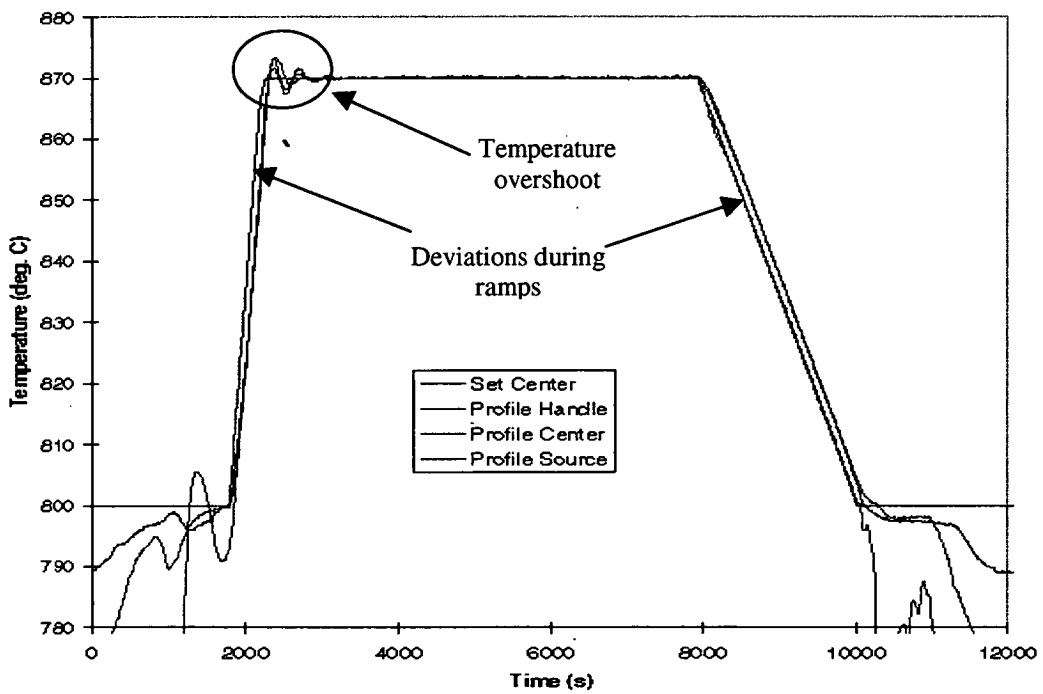


Figure 5.4 Furnace control system output of 3-zone furnace profiles

The outputs of the process model were calibrated to in-line process data such as oxide thickness, and electrical test data (e.g. sheet resistivity). Any fine-tuning in oxide thickness was achieved by adjusting the pressure in the furnace process flow (section 5.4.2). The thickness of the oxides should be those actually measured, and these values may not be necessarily at the centre of the specification. If the oxide thicknesses are correct then any post implant dopant profile will be correct The diffusion coefficients of the dopant species within SUPREM-3 were assumed to be correct. None of the defaults in the SUPREM-3 program were changed for this work with the exception of setting the interface charge, Q_{ss}, in the gate oxide in order to calibrate Vt as was the case for the LMDMOS process models [8,14]

The impurity profiles were calibrated to Spreading Resistance Profiles (SRPs) and, in some cases, Secondary Ion Mass Spectroscopy (SIMS), taking into account measurement error. Therefore, the simulated SRP and plot of chemical concentration were both used in SUPREM-3 to account for this error. The calibration of the process model is crucial for the output from the device model to be of any use in predicting electrical performance. The final test of the process model's capability to predict device performance will be presented in section 5.5.

5.4.3 Comparison of Simulated and Silicon Results

Calibration of the simulation results to silicon was performed by comparing the resulting simulated dopant profiles with those obtained by Spreading Resistance Profiles (SRPs) from silicon. It is worth remembering that an SRP does not show chemical concentration but shows electrical active concentration. Figure 5.5 shows an SRP of the shallow n^+, p-body and n^- drift dopant profile obtained from silicon after

completion of the full process. Comparison to the simulated plot in figure 5.6 shows a shallow n^+ diffusion junction depth and peak concentration of 0.65μm and $7\times10^{19}cm^{-3}$, which is similar to the values obtained from the SRP (figure 5.5) of 0.7μm and $1\times10^{20}cm^{-3}$ respectively. It can also be seen that the n^- drift regions concentrations are approximately the same, $2\times10^{15}cm^{-3}$. The same procedure was used to check all the dopant regions of the device.

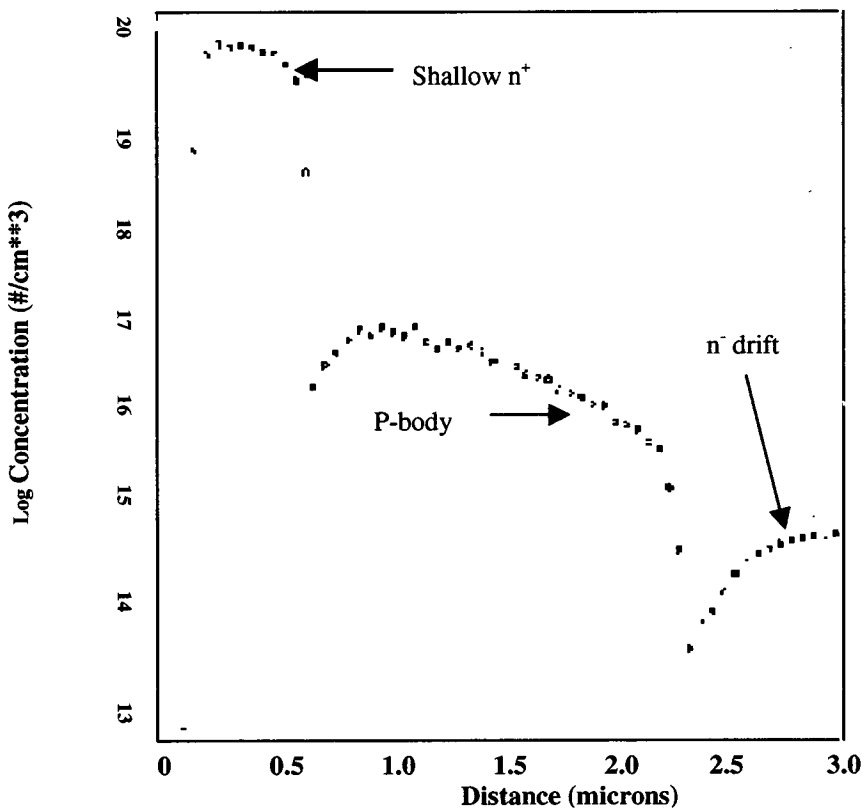


Figure 5.5 SRP of the JHVH-450 process of shallow n^+ into p^-body into the n^- drift region

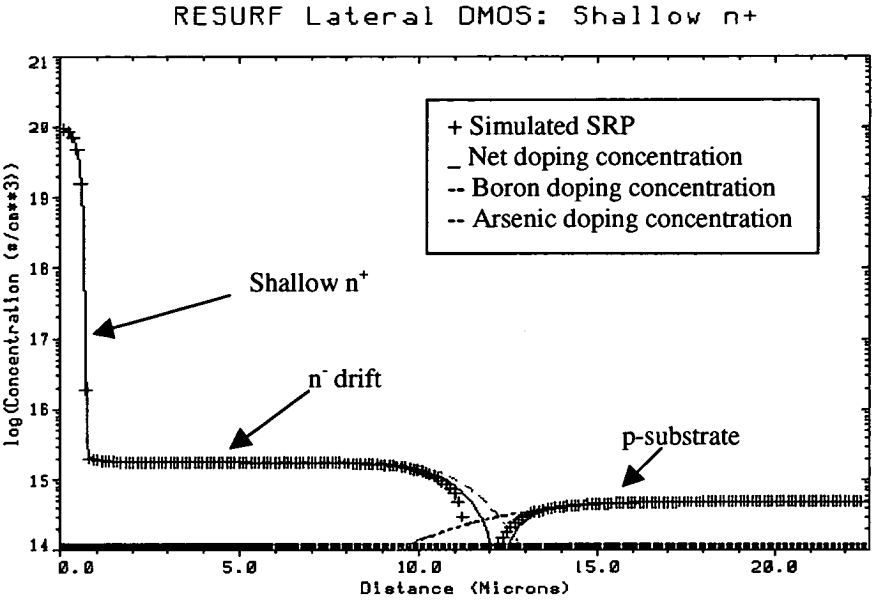


Figure 5.6 Simulated SRP from SUPREM-3 of the shallow n^+ region

Table 5.1 compares the simulated and measured values of sheet resistivity for the main dopant profile regions.

Region	Measured sheet resistance (Ω/•)	Simulated sheet resistance (Ω/•)
Shallow p^+	193.2	141.8
Shallow n^+	24.34	23.06
Drift region	2900	3100
p-body	970.85	823
Deep p^+	112	79

Table 5.1 A comparison of measured and simulated sheet resistance values for JHVH-450 process

From table 5.1 it can be observed that the simulated sheet resistances are up 15% lower than the average values obtained from silicon. As the simulation was being used to predict trends this was considered accurate enough and so no further calibration was undertaken. Interestingly the n- drift region difference in table 5.1 is only 6.5% which is fortunate as this is the key doping region for the RESURF device.

As a consequence of these results, work was initiated to investigate the causes of the observed differences with the objective to improve the accuracy of the simulation results. This work [12] attributed the difference between the simulated and measured sheet resistances to the SUPREM-3 diffusion models being purely analytical, and not taking into account some of the physical phenomenon associated with diffusion. Other work [13] investigated boron activation and segregation issues for the LMDMOS process. The results proved that all boron implants were activated up to concentrations of $1\times10^{19}cm^{-3}$. However, SUPREM-3 does not model boron segregation accurately. It was noted that despite the offsets seen between simulated and measured results, the generic LMDMOS models had been used successfully to predict process and device performance and distributions [14]. Therefore, there was reasonable confidence in the ability of the tools to predict the JHVH-450 device performance.

5.5 Review of Device Simulation Deck

5.5.1 Device Simulator

The electrical device characteristics were modelled using 2D simulator Medici. The outputs produced by SUPREM-3 are used to set-up the device dopant profiles from which the electrical performance can be modelled. The information provided by Medici allows the user to investigate parameters that cannot always be measured on

silicon. These include visualisation of the internal operation of the device such as the movement of depletion regions, and position of breakdowns.

5.5.2 Medici Input Deck

The set-up for Medici is shown in figure 5.7, and the complete deck is given in Appendix 3.

The first task in the device simulation file is to define the grid with sufficient detail in the correct regions which is crucial in order to obtain accurate device simulation results [15].

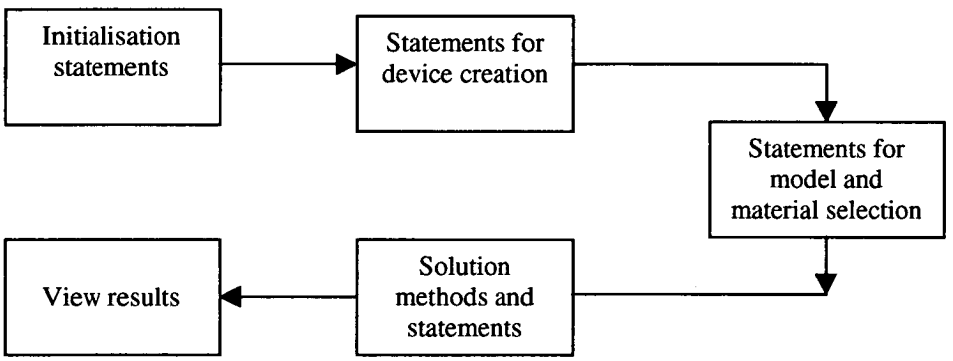


Figure 5.7 Generic structure of Medici input deck

As with other simulation tools there is a trade-off between simulation accuracy and execution time determined by the number of nodes (N) in the grid. There is a point where increasing the number of nodes no longer improves the accuracy and only results in increased simulation time. The number of arithmetic operations necessary to achieve a solution is approximately N^3 [16] and so this issue was considered very carefully when simulating the JHVH-450 transistor due to the length and depth of the structure as can be seen from figure 3.13. The methodology used is presented below. Firstly, the **assign** statement was used to set values to parameters that will vary during

the optimisation procedure. This facilitates modification of specific device and process parameters in a timely and accurate manner.

Device breakdown will occur in one of three areas on the transistor as illustrated in figure 5.8, and discussed in chapter 3. The peak field is likely to occur at the fingertips just before the parallel region begins as the electric field contours crowd [17].

1. Source fingertip
2. Drain fingertip
3. Parallel region

It can be seen that the fingertip regions are effectively two dimensional whereas the parallel region can be considered 1D. A simple option is to model the structure as three separate devices

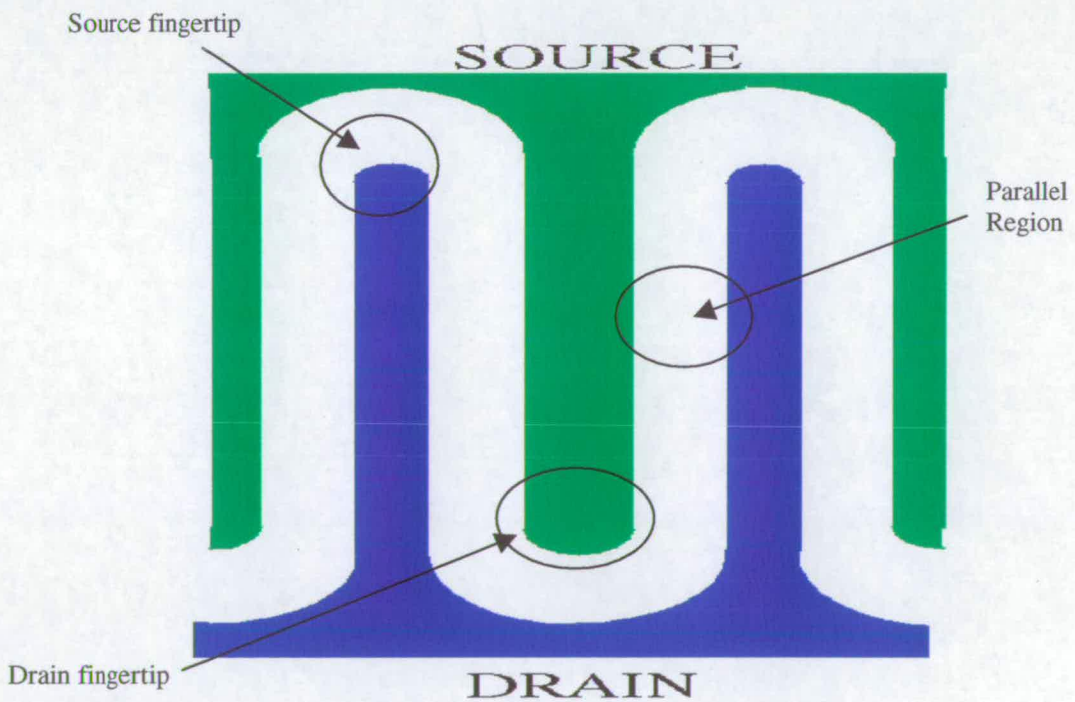
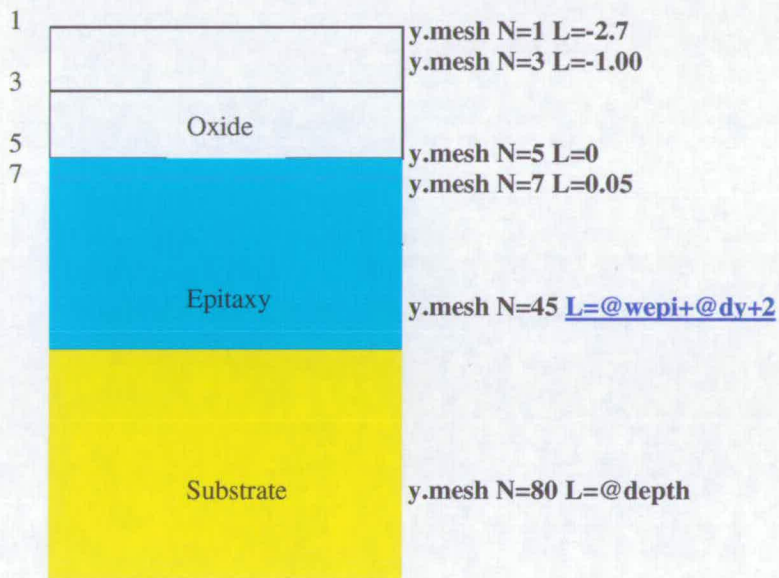


Figure 5.8 Plan view of device layout showing areas where breakdown could occur under reverse bias conditions

and determine the one which yields the lowest breakdown. The challenge simulating the fingertip region is the curved shape. In the device simulator, Medici, there is a

choice of two variables that can be defined after the **mesh** statement. The **Rectangular** method used cartesian coordinates that assumes that the device structure is identical for all values in the z direction and this was very easy to implement for the parallel region. The **Cylindrical** method was used for the fingertips as this models the radial and z direction dependency. The grid was defined using the **x.mesh** and **y.mesh** statements in the x and y directions and the mesh process is explained in figure 5.9.



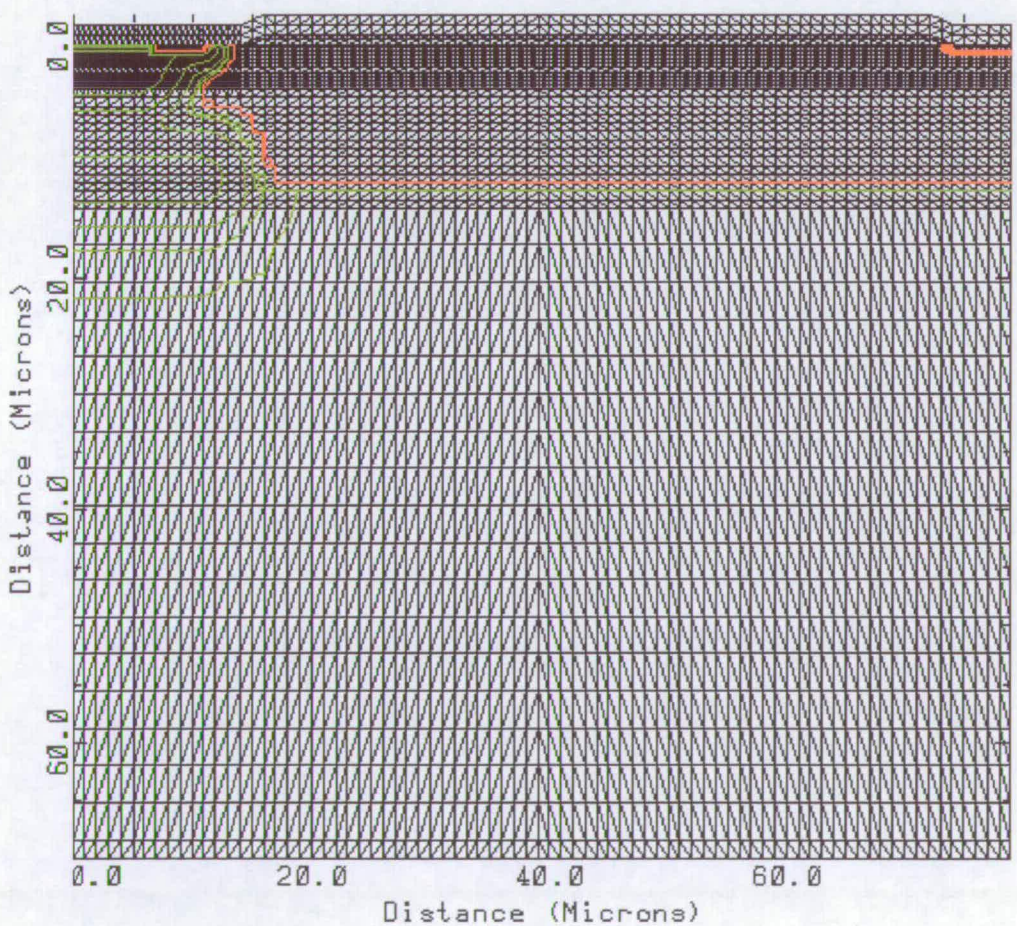
(Not to scale)

Figure 5.9 Mesh set up in the y direction for JHVH-450 device simulation

The first seven nodes were defined at specific depths according to the make up of the device structure. The grid in the y direction for the epitaxial layer and the silicon substrate were defined by the last two **y.mesh** statements. In each command the depth of the region and the number of nodes are defined where **wepi** is the epitaxy thickness, depth is the thickness of the simulation region and **dy** is a variable that is used to allow the epitaxy thickness to be varied. This facilitated automation of multiple simulations generated from designed experiments to assess

manufacturability. Medici automatically generates the grid by adding diagonals to define the triangular elements. The **eliminate** command is used to reduce the number of nodes in regions of little interest. The source and drift region of the transistor which is the area of interest has the greatest number of grid points. This region was determined from initial simulations using a rough grid, which highlighted the breakdown and current flow and the grid then gradually refined to improve accuracy [15]. The final grid is shown in figure 5.10.

Figure 5.10 Final Medici grid used for JHVH-450 simulation



The next step is to define the electrodes of the JHVH-450 device. This was a carefully considered task as the gridding in the oxide is important as modelling the source and drain overlaps impact the breakdown voltage of the device as discussed in chapter 3. The **spread** command was used to reflect the shape of the overlaps and electrodes. An initial approximation was made based on mask layout. Once the device had been fabricated, scanning electron microscope (SEM) photographs allowed this to be modified to best reflect silicon. Figure 5.11 shows a 2D device cross section produced by Medici.

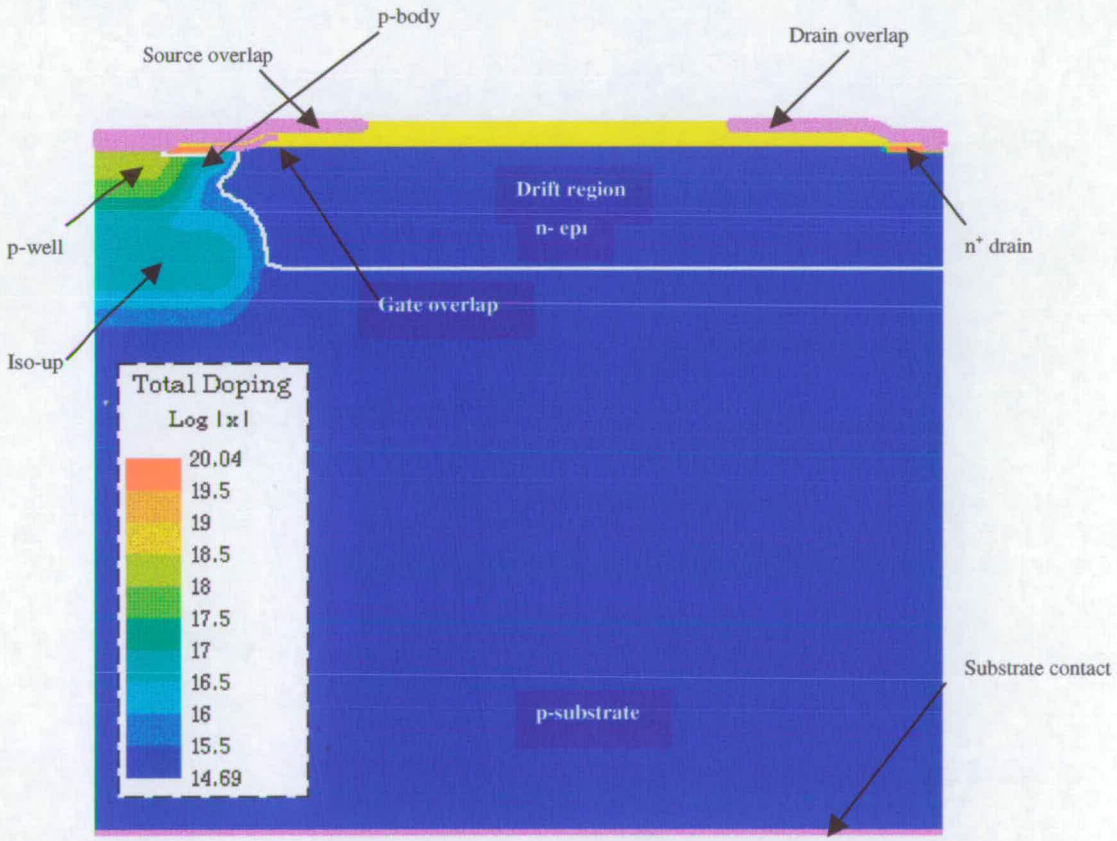


Figure 5.11 2D Medici cross section of a JHVH-450 device

The shape of the sideways diffusion was modelled using the **x.char** parameter which was calculated by multiplying the profile junction depths in order to get 80% of the vertical diffusion. In fact, the amount of sideways diffusion is dependent on the

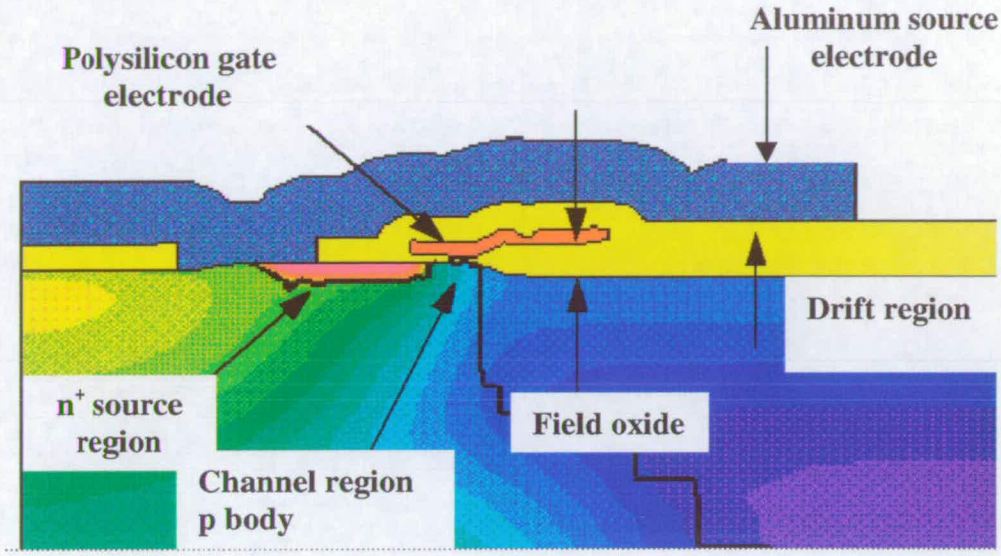
dopant species, relative concentration and junction depth. This was incrementally modified to agree with measured characteristics and figure 5.11 shows the resulting shapes of the various doping regions.

A two-dimensional process simulator such as TSUPREM-4 is obviously more appropriate to simulate 2D related effects and the topography of the poly gate and the metal field plates. From work initiated, Rankin [12] investigated the benefits of using a 2D process simulator. This showed an improvement with experimental measurement of sheet resistivity when compared with SUPREM-3 results (section 5.4.4). In addition the topography of the device was well modelled when compared to the SEM photographs (figure 5.12). However, the simulated breakdowns did not differ greatly from those obtained via 1D process simulation suggesting that this key area of the JHVH-450 device of direct interest to this work was well modelled.

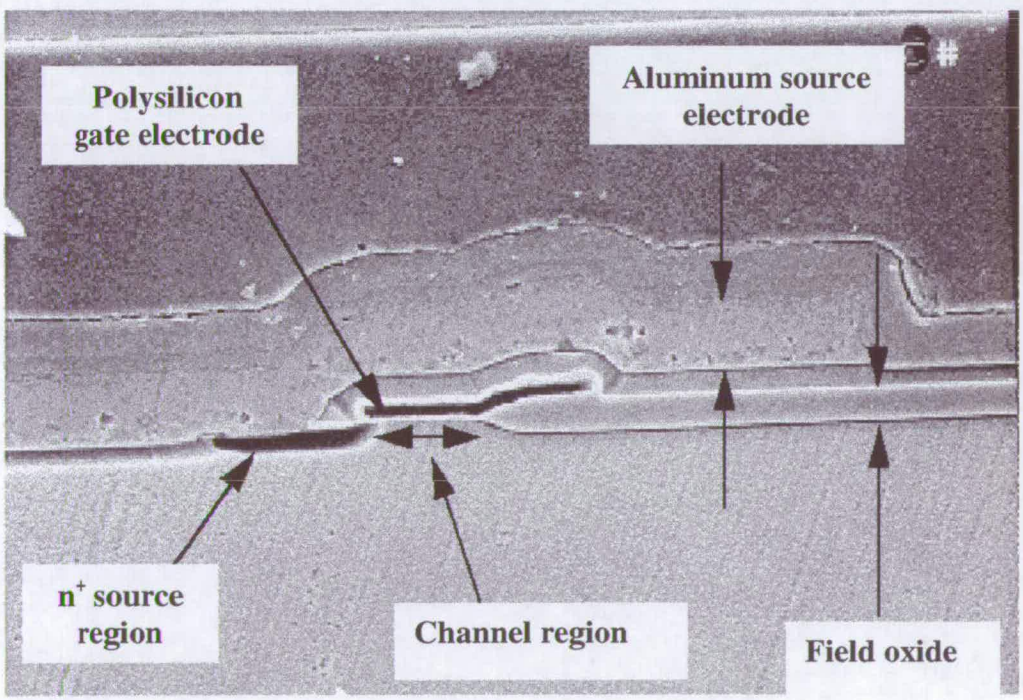
Once the structure was defined, common models to be used in the solution were selected. An example of this for simulating the breakdown voltage is shown below.

Model bgn conmob fldmob srfmob2 auger impact.i

The models selected were those used in previous work [8,14,18]. **Impact.i** specifies that carrier generation due to impact ionisation is included in the solution. The **auger** model models recombination. Phonon scattering at the oxide silicon interface is modelled using the **srfmob2** enhanced surface mobility model.



(a)



(b)

Figure 5.12 Comparison of (a) TSUPREM-4 2D process simulation cross section with (b) a SEM silicon cross section [12]

Once the models have been selected the terminal bias conditions must be set. The source to drain breakdown voltage, BV_{dss}, can be determined from simulation provided both the Poisson and continuity equations are solved. The gate, source and substrate were all connected to ground as shown in figure 5.13 and the drain ramped.

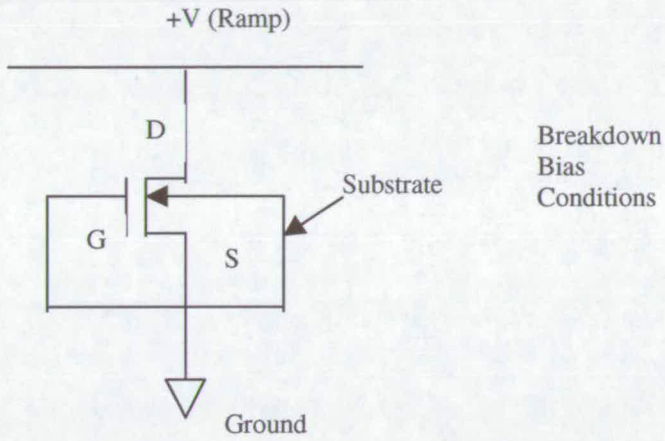


Figure 5.13 Breakdown voltage, BV_{dss}, bias set-up for JHVH transistor

As discussed in chapter 2, Poisson's equation governs the potential in the JHVH-450 device, and a set of equations known as the continuity equations govern the carrier concentrations. The continuity equations [17] describe the time dependent relationship between current density, recombination rate, and distance. They may be derived by taking non-uniform excess carrier concentration and considering the current (of one type of carrier) entering and leaving an incremental strip of width *dx*. These equations are discretised [16] on the simulation grid set up above. The discretisation of the equations uses non-linear algebraic equations that are solved by the **Gummel** and **Newton** iteration methods.

The covergence of these iteration methods to a solution is determined by the accuracy of the initial guess.

symbol gummel carriers=0
method iccg damped
solve initial

The **symbol** statement is used to select the gummel iteration method with zero carriers to reduce the number of equations. The **method** statement was used to control the **gummel** iteration process by selecting the fastest iteration solver, **iccg**, the incomplete Cholsky conjugate gradient method, and damping the numerical ringing in the Poisson iteration which limits the maximum voltage change per iteration in the device [16]. The selection of these parameters allowed an initial solution of Poisson's equation and the current continuity equations. The solution obtained from **solve initial** was simply a neutral charge assumption used to obtain the first bias point. It was used as an initial guess to solve the breakdown biasing conditions of the JHVH device as shown

Symbol newton carriers=2
Solve elec=drain continue c.vstep=0.5 c.toler=0.1 c.vmax=600

It is recommended that the Newton iteration [16] should always be used to solve Poissons's equation and the current continuity equations after the initial solution. To solve for electron and hole current continuity equations two **carriers** were selected. Two methods were chosen to determine the breakdown voltage. The first used the continuation method, which traces the I-V characteristic by automatically selecting the bias step was very easy to use for the large JHVH-450 device but at times, convergence problems were experienced. This method simplified the convergence task by selecting smaller and smaller steps until convergence is obtained. The **c.vstep** parameter controls the size of the initial step and the spacing between computed

points is controlled by **c.toler** and the simulation continued until **c.vmax** is reached. The value of **c.vmax** was typically large enough to be able to view the JHVH breakdown characteristic in figure 5.14. Whilst this method resulted in a converged solution and I-V characteristic, the execution time was inordinately long – up to 8 hours. The second method involves applying 10-volt steps on the drain and solving Poisson's and the continuity equations. If convergence was not achieved then the previous converged solution was noted as the breakdown voltage.

This was found to be much faster and the results were comparable to those obtained from the previous continuation method.

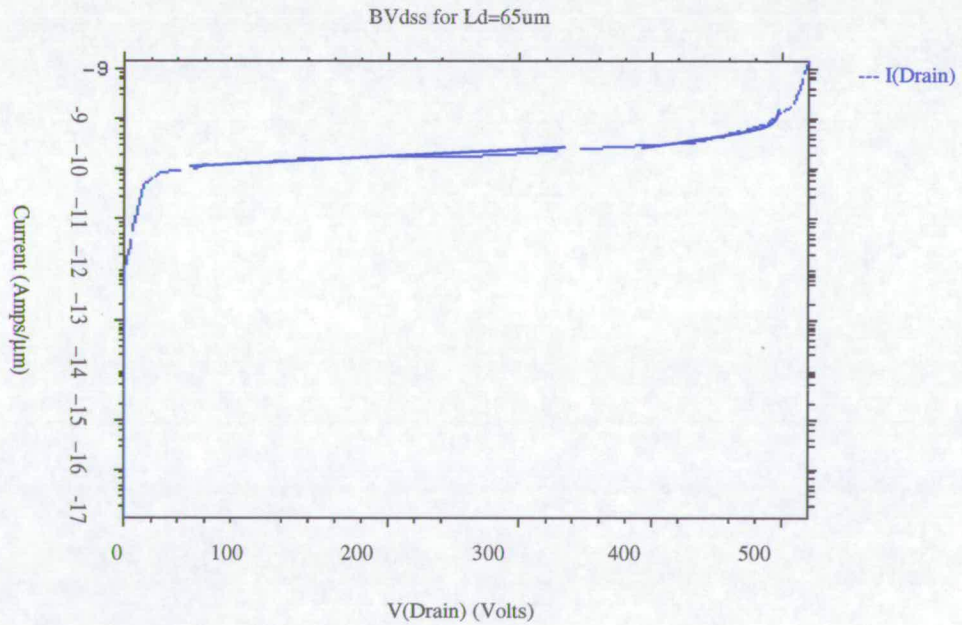


Figure 5.14 Simulated JHVH-450 source fingertip breakdown characteristic

Figure 5.15 shows current flow lines in the device after the device has broken down.

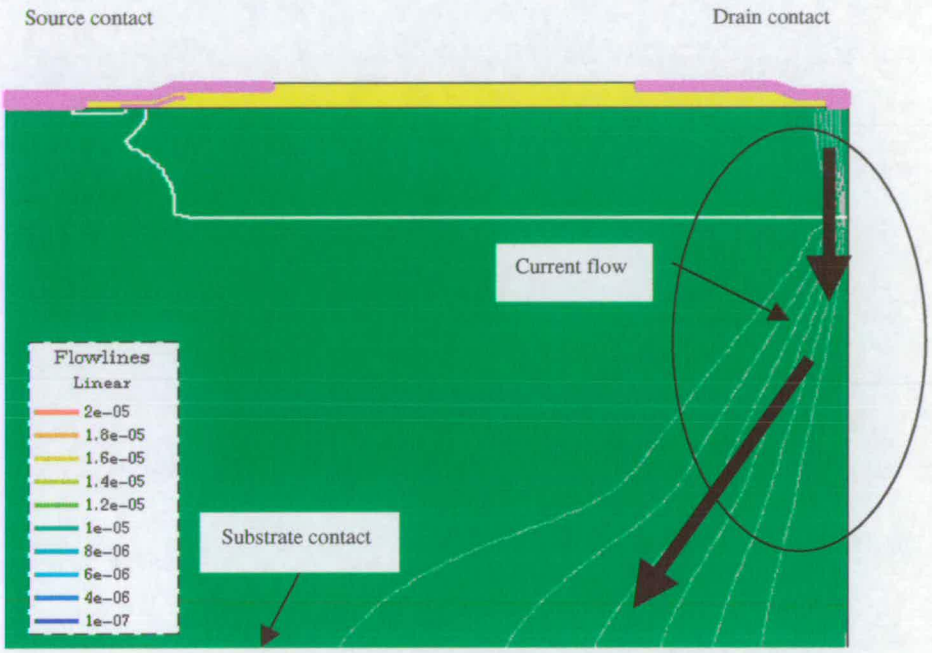


Figure 5.15 2D plot of current flow in JHVH device

The flow lines indicate that the breakdown occurred beneath the drain region of the device and the ionisation generation shown in figure 5.16 confirms this.

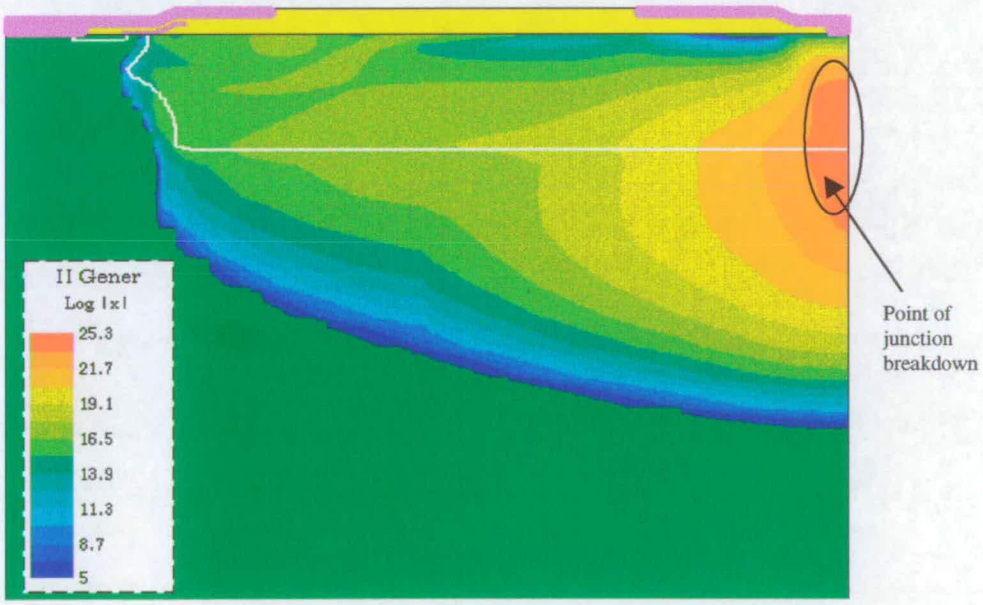


Figure 5.16 2D plot of impact ionisation for JHVH device at breakdown

5.6 Process Optimisation

5.6.1 Introduction

As discussed in sections 5.1 and 5.2, the target breakdown was 450V with a guaranteed minimum of 400 volts. The constraints placed upon the process were that the device had to be able to be integrated into the LMDMOS technology and therefore, the epitaxial layer thickness had to remain the same.

5.6.2 Optimisation Methodology

As covered in chapters 3 and 4 and [4], the process parameters that required to be optimised were:

(a) Epitaxy resistivity

(b) Epitaxy thickness

(c) Substrate resistivity

(d) Implant dose

(e) Implant energy

For the epitaxy only process, there were only three parameters that were applicable, as this process had no implant step. Even though the epitaxy layer thickness was fixed to be the same as the LMDMOS process, the impact of layer thickness on the breakdown voltage and hence the manufacturability over its specification had to be assessed, and interaction with the other process variables. The same was also true for the other variables.

To optimise the variables above to achieve a process with an achievable specification for the breakdown voltage in a manufacturing environment Design Of Experiment

(DOEs), in conjunction with Technology CAD (TCAD) tools were used (see chapter 2). This resulted in a 'window of operation' for the technology. The manufacturability of the technology was then assessed by determining the effect of the variability of each of the control factors on the breakdown voltage. This analysis assumed that the control factors vary in a Gaussian manner. Calphurnia [6] was used to perform a Monte Carlo analysis using the response surface polynomials to predict the breakdown voltage distribution. The variability of the respective process parameters was determined from in-line measurements, electrical test data, and discussions with equipment and process engineers. Previous work [4,8,14,19] provided confidence in the values used.

5.6.3 Epitaxy Only Process

The initial process specification was based on previous work [9] where each process variable had been varied separately. A Central Cubic Inscribed (CCI) designed experiment constructed using RS/1 [20] was simulated and the resulting breakdown voltage, BV_{dss}, noted (table 5.2). The process control factors were varied as follows:

1. Epitaxial resistivity – 2.6 to 4.4Ω-cm
2. Epitaxial thickness – 10.5 to 13.5μm
3. Substrate resistivity – 21.5 to 32.5Ω-cm

A wider range of epitaxial layer thickness than would normally be expected was investigated in order to assess any manufacturability challenges that may occur as a result of integrating the RESURF device into the LMDMOS flow. The breakdown voltage was the minimum of the three simulated values for the source fingertip, drain fingertip, and parallel regions (section 5.5.2 and figure 5.8). The drift region was set at 30μm for the parallel region and 60μm at the fingertips which was simulated with

radial coordinates as described in section 5.5.2. The field plates in the parallel region were 5μm and 8μm at the drain and source respectively. These were extended at the fingertips to 15μm.

Run #	Epitaxy Resistivity (Ω-cm)	Epitaxy Thickness (μm)	Substrate Resistivity (Ω-cm)	BVdss (volts)
1	3	11.1	23.73	470
2	4	11.1	23.73	440
3	3	12.9	23.73	400
4	4	12.9	23.73	520
5	3	11.1	30.27	500
6	4	11.1	30.27	500
7	3	12.9	30.27	380
8	4	12.9	30.27	550
9	2.6	12	27	370
10	4.4	12	27	500
11	3.5	10.5	27	490
12	3.5	13.5	27	450
13	3.5	12	21.5	480
14	3.5	12	32.5	540
15	3.5	12	27	520

Table 5.2 CCI design worksheet for JHVH-450 epitaxy process

The results were analysed in RS/1 and the R^2-adj measure of the quality of the model fit was 0.979. The closer the R^2-adj is to 1, the better the fit and a good fit would be expected since there is no experimental noise associated with simulated data (see chapter 2). The main effects plot of the control factors in figure 5.17 show that for the ranges over which the control factors were varied, the epitaxy resistivity has the greatest impact on the breakdown voltage, BV_{dss}. This is perhaps not too surprising since the charge in the drift region controls the interaction between the vertical and horizontal depletion regions in the RESURF device. It is also worth noting that the contributions of epitaxy thickness and substrate resistivity to BVdss are comparable.

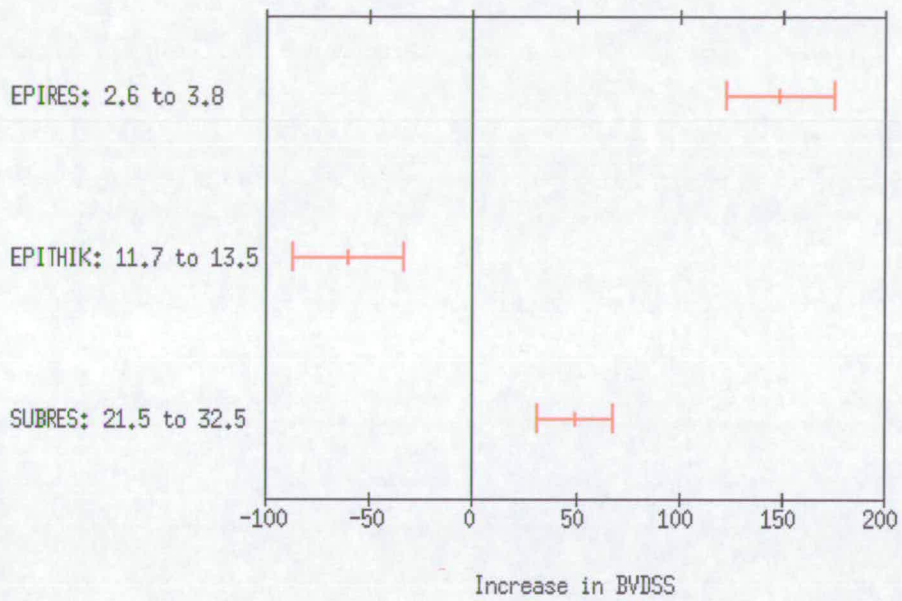


Figure 5.17 Main effects plot for JHVH-450 epitaxy process with 95% confidence intervals

The resulting response surfaces showed that the epitaxy resistivity specification based on previous work [9] of 3.2±0.3Ω-cm is not capable of delivering 400V over the entire resistivity specification as shown in figure 5.18.

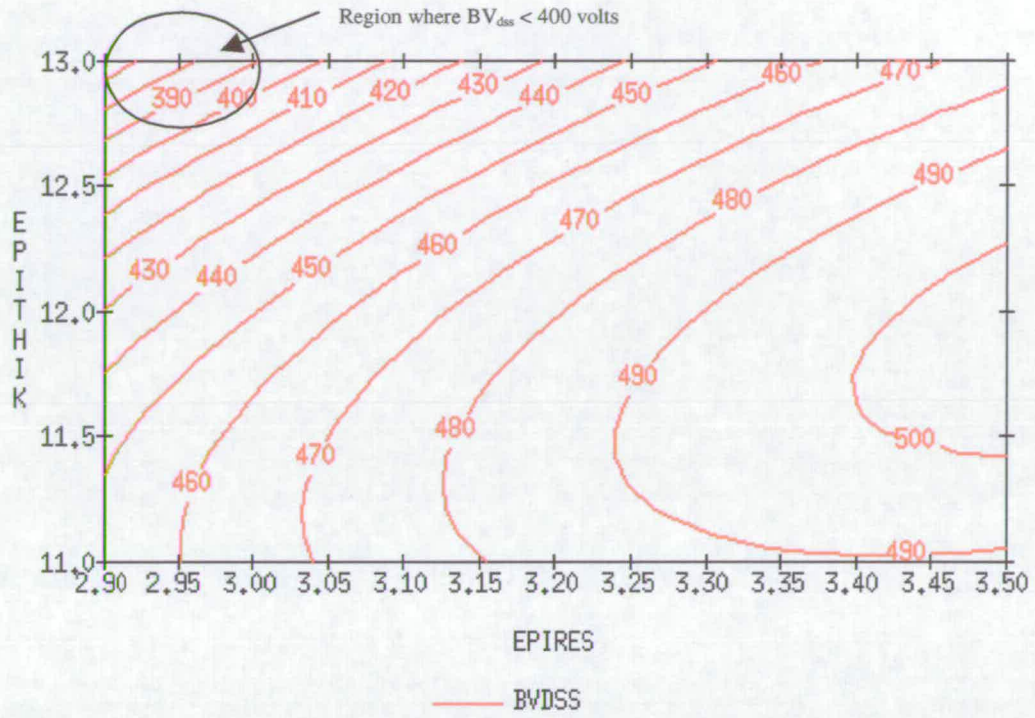


Figure 5.18 Breakdown voltage response surface for JHVH-450 epitaxy process with epitaxy resisitivity varying 3.2±0.3Ω-cm

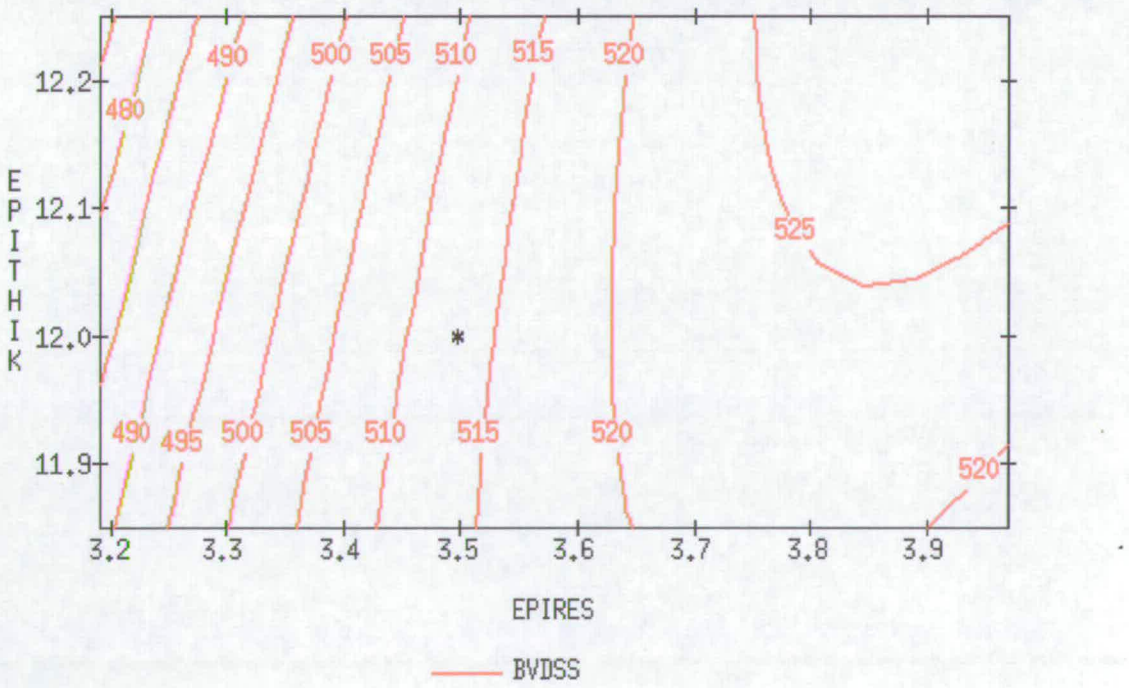


Figure 5.19 Breakdown voltage response surface for JHVH-450 epitaxy process with a workable epitaxy resisitivity process specification

The response surface in figure 5.19 shows a more workable epitaxy specification where BVdss is always greater than the minimum specification of 400 volts. This showed that using experimental design can highlight potential manufacturability challenges before silicon has been fabricated. As mentioned in section 5.6.2, the next step in the process optimisation process was to assess the manufacturability of the chosen process conditions. Calphurnia [6] was used and the design of experiment process variables were varied as follow:

1. Epitaxy resistivity - ±7%
2. Epitaxy thickness - ±5%
3. Substrate resistivity - ±3%

The resulting simulated response distribution is shown in figure 5.20. This distribution shows that the optimised process should yield a manufacturable process

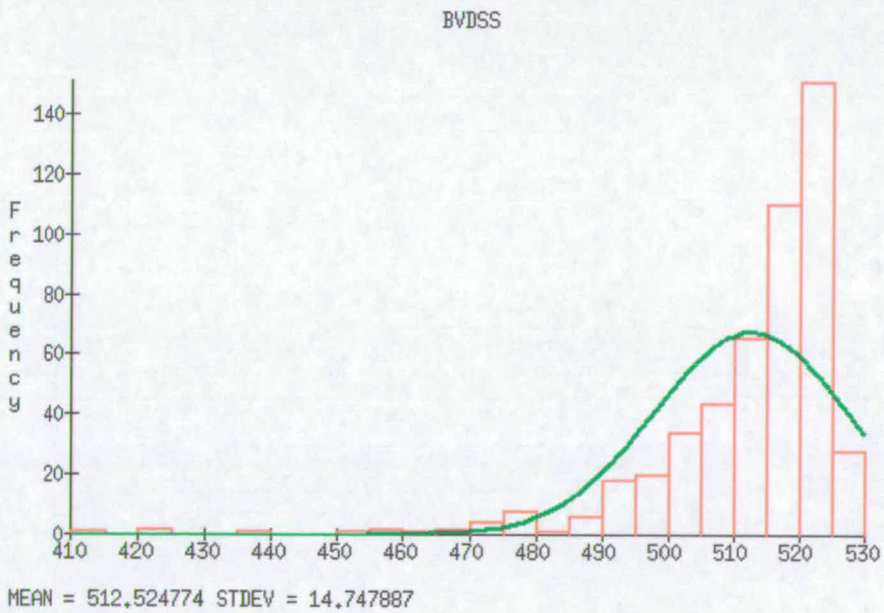


Figure 5.20 Simulated breakdown voltage distribution for JHVH-450

when fabricated on silicon. This, of course, assumes that no other variables affect the breakdown voltage apart from those highlighted in table 5.2 and [5].

When silicon was fabricated and the result was breakdown voltage distribution shown in figure 5.21.

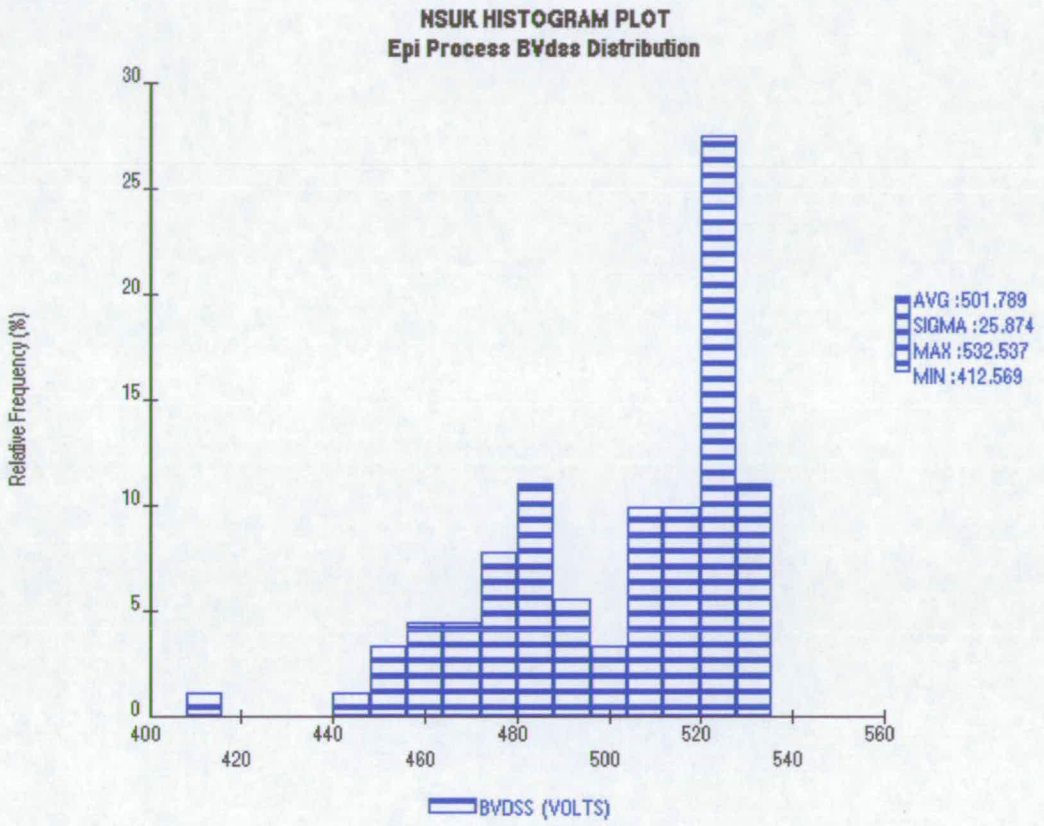


Figure 5.21 Breakdown voltage distribution obtained from silicon for JHVH-450 epitaxy process

A comparison of figures 5.20 and 5.21 is made in table 5.3. The results show a reasonable prediction was achieved from simulation for the mean value but the

	Mean	Standard Deviation
Simulated	512.525	14.747
Silicon	501.789	25.874

Table 5.3 Comparison of simulated and silicon distributions for JHVH-450 epitaxy process

standard deviations are significantly different. This could be due to one of two reasons. Firstly, the variability assigned to the process variables may well be greater than anticipated, or secondly the actual spread from silicon may well have tighter due to the wafers being processed by engineering. The results suggest that the selected control factors are the main ones driving the breakdown voltage for the epitaxy process. It should be noted that in figure 5.21 (silicon) there appears to be a dual distribution. The processing conditions were checked and nothing was found that could suggest a reason for this dual distribution. The resulting recommendation for this process was:

Epitaxy resistivity – 3.3 to 4.3Ω-cm

Epitaxy thickness – 11 to 13μm

Substrate resistivity – 22 to 32Ω-cm

The epitaxy resistivity specification was chosen to allow for a process specification that could not only guarantee the lower breakdown limit of 400 volts but also a unit process that could be controlled by the manufacturing process engineers. Also the epitaxy thickness satisfies the constraint that the RESURF device should be able to be integrated into the LMDMOS flow. The resistivity was one that manufacturing believed could be reproduced with good control

5.6.4 Combined Epitaxy and Implant Process

For this process there was no prior work which could be used to help determine the optimum process settings. However, it was known that then a higher epitaxy resistivity was required in order to have good control of the charge in the drift region using an implant [7], to control the breakdown voltage, BV_{dss}. After some discussion with the manufacturing process engineers, it was decided that the best approach was

to use an epitaxy resistivity of 10-15Ω-cm as this resistivity specification was possible to run in manufacturing without great difficulty and engineering effort. A Central Cubic Inscribed (CCI) designed experiment constructed in RS/1 [20] was then simulated and the resulting breakdown voltage, BV_{dss} (table 5.4). The ranges used for the process variables were:

1. Epitaxial resistivity – 9 to 16Ω-cm
2. Epitaxial thickness – 10.5 to 13.5μm
3. Substrate resistivity – 21.5 to 32.5Ω-cm
4. Drift implant dose – 1.2 to 1.6×10^{12} cm^{-2}
5. Drift implant energy – 115 to 125keV

Run #	Epitaxy Resistivity (Ω-cm)	Epitaxy Thickness (μm)	Substrate Resistivity (Ω-cm)	Drift Implant Dose ($10^{12}cm^{-2}$)	Drift Implant Energy (keV)	BVdss (volts)
1	10.8	11.3	24.25	1.5	117.5	440
2	14.3	12.8	24.25	1.3	122.5	480
3	12.5	10.5	27	1.4	120	460
4	14.3	11.3	29.75	1.3	122.5	480
5	10.8	11.3	29.75	1.3	122.5	490
6	12.5	12	27	1.4	115	470
7	14.3	11.3	24.25	1.5	122.5	460
8	12.5	12	27	1.4	125	470
9	12.5	12	27	1.4	120	470
10	14.3	12.8	29.75	1.5	122.5	420
11	10.8	12.8	29.75	1.5	117.5	390
12	14.3	11.3	24.25	1.3	117.5	450
13	10.8	12.8	24.25	1.5	122.5	410
14	10.8	12.8	29.75	1.3	122.5	440
15	12.5	12	21.5	1.4	120	450
16	14.3	12.8	24.25	1.5	117.5	440
17	10.8	11.3	29.75	1.5	122.5	420
18	10.8	11.3	24.25	1.3	122.5	460
19	12.5	12	27	1.2	120	480
20	16	12	27	1.4	120	470
21	14.3	12.8	29.75	1.3	117.5	500
22	12.5	12	32.5	1.4	120	440
23	12.5	13.5	27	1.4	120	430
24	9	12	27	1.4	120	420
25	10.8	12.8	24.25	1.3	117.5	480
26	12.5	12	27	1.6	120	410
27	14.3	11.3	29.75	1.5	117.5	450

Table 5.4 CCI design worksheet for JHVH-450 epitaxy and implant process

The results were analysed in RS/1 and the R^2-adj was 0.94. A main effects graph in figure 5.22 shows how the various process variables used in the design impact the breakdown voltage BV_{dss}. This indicates that the drift implant dose has the greatest impact on BV_{dss}, while the drift implant energy effect on the breakdown is negligible. Consequently, the drift implant energy was fixed at 120keV thereby simplifying silicon-based experiments and reducing engineering involvement and contributing to a reduction in fabrication and development cycle time.

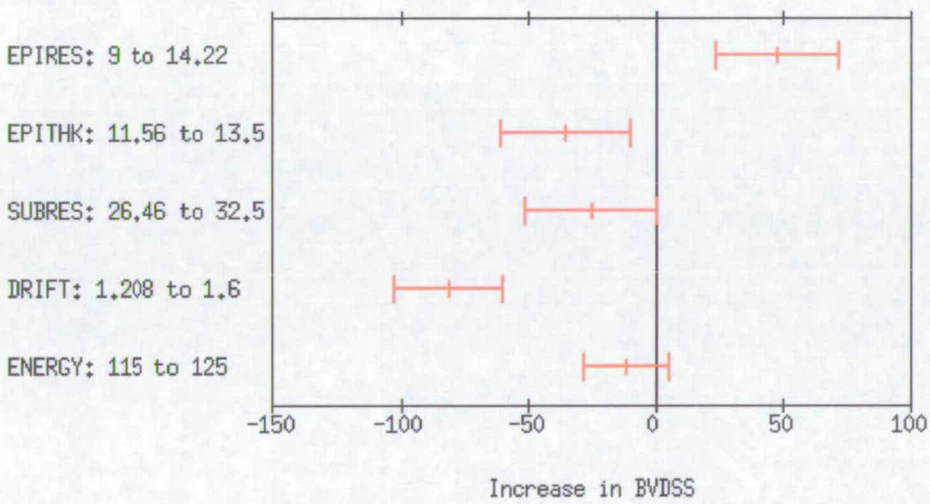


Figure 5.22 Main effects plot for JHVH-450 epitaxy and implant process with 95% confidence intervals

The resulting response surfaces of epitaxy resistivity versus epitaxy thickness with the other control factors fixed are presented in figures 5.23 and 5.24, for different drift implant doses: 1.2, 1.4 and 1.6×10^{12} cm^{-2}. Examination of figures 5.23 and 5.24 indicate that a drift implant dose of 1.2×10^{12} cm^{-2} yields the optimum process. It can be observed that drift implant doses of 1.4 and 1.6×10^{12} cm^{-2} were incapable of consistently delivering a minimum 400 volts BV_{dss} over the range of process variables

highlighted as driving the breakdown voltage, BV_{dss}. Therefore the optimum process conditions were:

Epitaxy resistivity – 10 -15Ω-cm

Epitaxy thickness – 11 - 13μm

Substrate resistivity – 22 - 32Ω-cm

Drift implant dose – 1.2×10^{12} cm^{-2}

Drift implant energy – 120keV

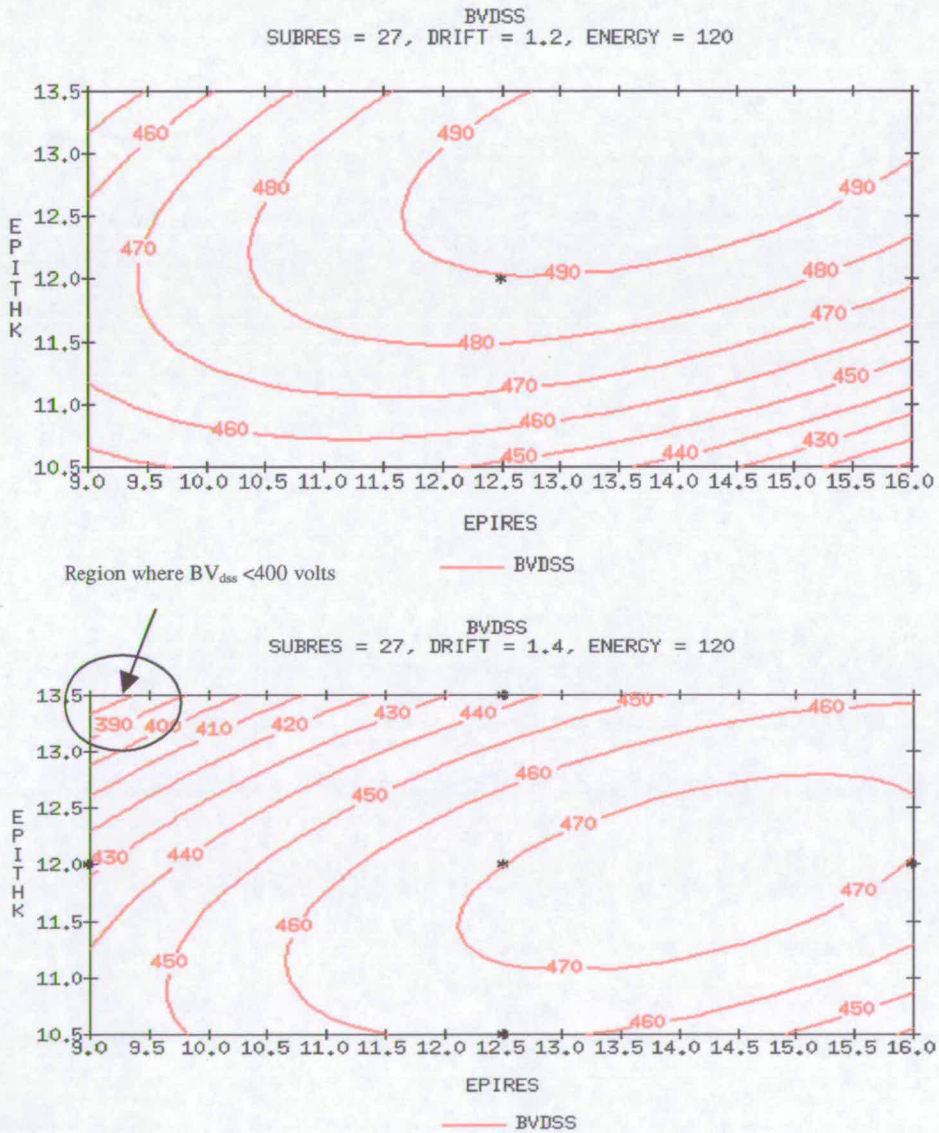


Figure 5.23 BVdss contour plots for JHVH-450 epitaxy and implant process for drift implant doses (a) 1.2 and (b) 1.4×10^{12} cm^{-2}

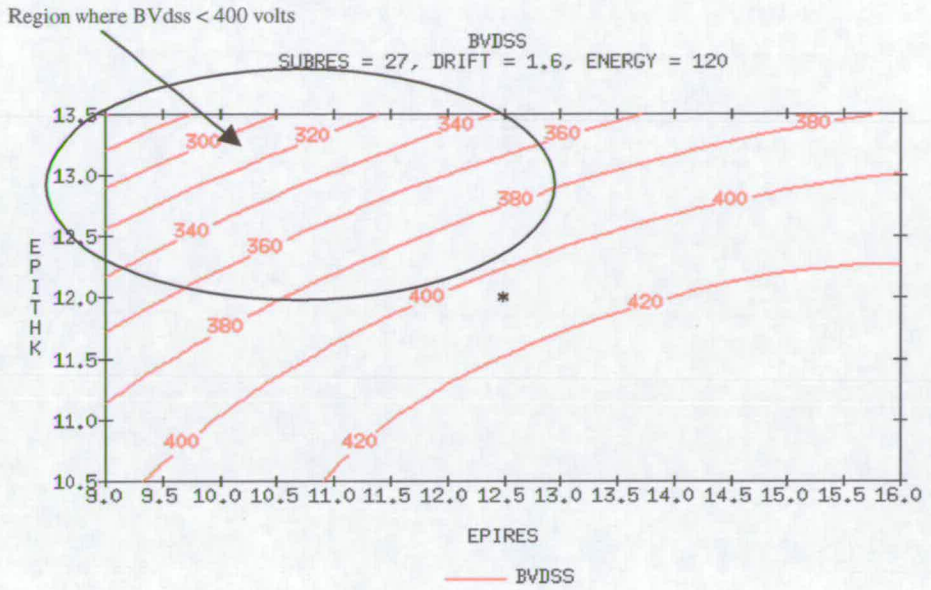


Figure 5.24 BVdss contour plots for JHVH-450 epitaxy and implant process for drift implant dose of 1.6 $\times 10^{12}$ cm^{-2}

The epitaxy thickness satisfies the constraint that the RESURF device should be able to be integrated into the LMDMOS flow. The resistivity was one that manufacturing could reproduce with good control and the substrate resistivity was the same as the epitaxy only process option, and being similar to one that currently ran on another process. The difference being that the desired specification for the JHVH-450 process was tighter and so, existing silicon could be screened if required or the tighter specification ordered for both technologies assuming no cost penalties would be incurred. Lastly, the required implant dose and energy could be easily implemented as these are simply programmed into the implant machine.

As discussed in section 5.6.2, the next process optimisation step was to assess the manufacturability of the chosen process condition. It should be noted that this time five factors are being investigated in comparison to three previously.

Calphurnia was used and the control factors were varied as:

1. Epitaxy resistivity - ±7%
2. Epitaxy thickness - ±5%
3. Substrate resistivity - ±3%
4. Drift implant dose – ±7%
5. Drift implant energy – ±2%

The resulting simulated response distribution is shown in figure 5.25. This simulated BV_{dss} distribution shows that the proposed optimised process should be manufacturable. Again, as with the epitaxy process discussed previously, this assumes that no other process variables control the breakdown voltage apart from those highlighted in table 5.4.

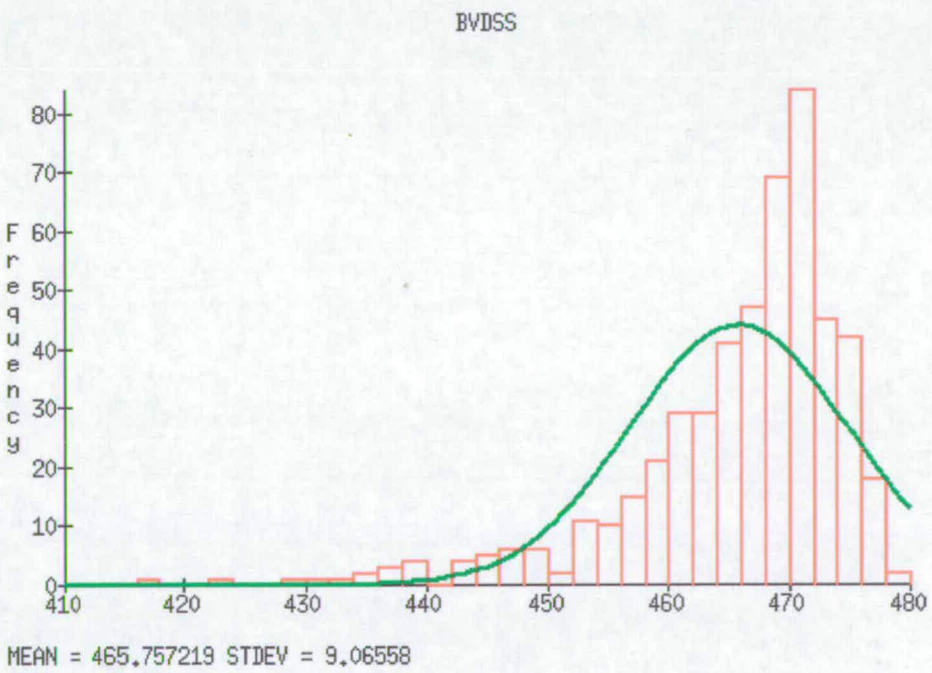


Figure 5.25 Simulated breakdown voltage distribution for JHVH-450 epitaxy & implant process

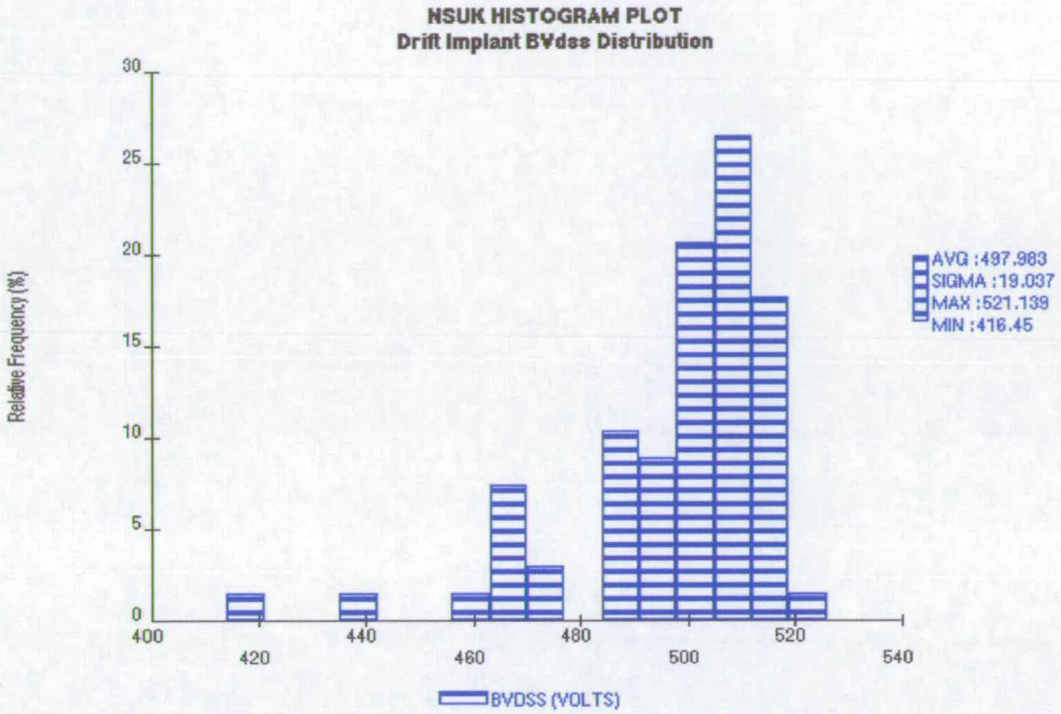


Figure 5.26 Breakdown voltage distribution (silicon) for JHVH-450 epitaxy and implant process

Silicon was fabricated and the resulting breakdown voltage, BV_{dss}, distribution is shown in figure 5.26.

The results are summarised in table 5.5.

	Mean	Standard Deviation
Simulated	465.757	9.0656
Silicon	497.983	19.037

Table 5.5 Comparison of simulated and silicon distributions for JHVH-450 epitaxy and implant process

The results show a reasonable prediction of breakdown voltage has been achieved by simulation. It is again noted that the standard deviation predicted by simulation is much smaller than the measured distribution. The same reasons as those described previously for the epitaxy process were hypothesised. Further investigation identified challenges at the epitaxy process stage which will be discussed further in section 5.7.1.

5.7 Silicon split lot results

5.7.1 Epitaxy Process

Several lots, 208 wafers in total, were fabricated using the epitaxy process which was varied over the range specified in section 5.6.3. The results from these wafers were analysed using the MULtiple REGression analysis (MULREG) available in RS/1, and the results were compared to those from simulation. The epitaxy resistivity and thickness were determined from measurements on test chips. The silicon wafers were screened from available production material in order to ensure the resistivity met the 22-32Ω-cm specification and in this case the resistivity data had a very tight distribution centred at 27Ω-cm.

The epitaxy process varied between the following values:

Epitaxy resistivity – 3.23 to 3.97Ω-cm

Epitaxy thickness – 11.85 to 12.25μm

The data was analysed using multiple regression, the R^2-adj, a measure of how good the model produced fits the data, was 0.789. The resulting contour plot in shown in figure 5.27 and shows how BV_{dss} varies with changing epitaxy thickness and

resistivity. From the MULREG analysis it was noted that the substrate resistivity did not effect the breakdown voltage and was therefore, not included in the model. This was not too surprising because the resistivity of the material used had a very tight distribution.

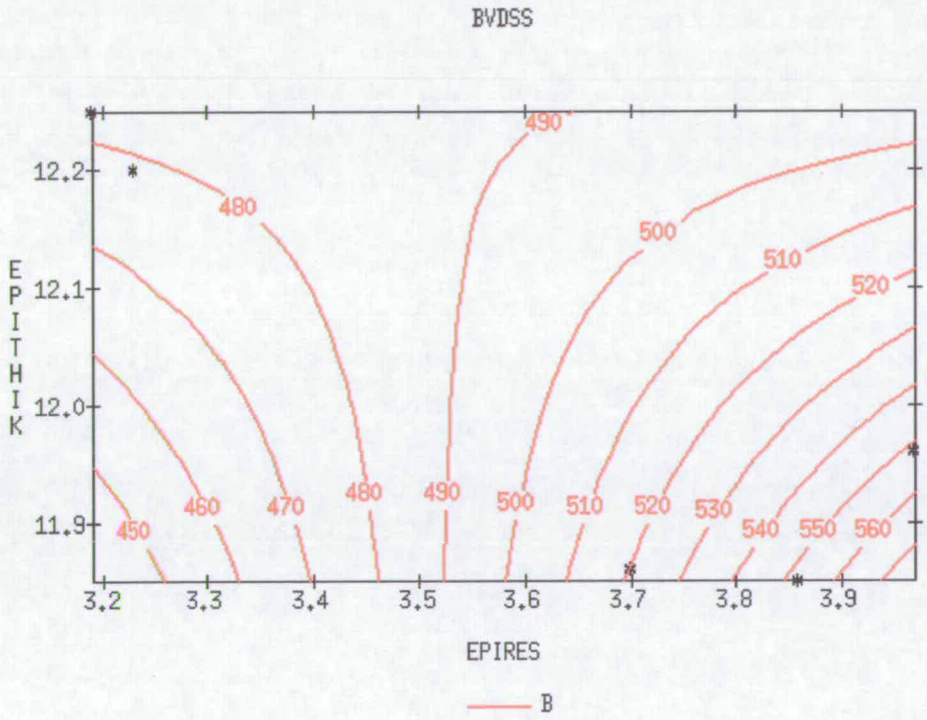


Figure 5.27 Breakdown voltage response surface from silicon runs for the epitaxy JHVH-450 process

Comparing figure 5.27 with figure 5.19 it can be observed that the contours do not match, even though the simulation results for the epitaxy process successfully resulted in a process that consistently yielded greater than 400 volts.

The R^2-adj value of 0.789 indicates a poor fit to the data highlighting that there was perhaps, some experimental noise in the data or errors with measurement tools. A comparison to bench measurements yielded good agreement. However, the epitaxy measurement was questioned. Initial investigations indicated that different silicon

orientation had been used for the production runs. This caused some concern as the silicon orientation on the test chips (¼ wafers), meant that the measurements were not directly comparable. A comparison of SRP results also indicated that there was an opportunity for improvement in the SRP methodology employed [21] and a detailed analysis [22] determined that the measurement left much to be desired. In fact the junction depth being in error by as much as 40%. As a result a new methodology was employed where the test chips were fabricated on wafers with the same orientation as production wafers. The use of the same orientation was supported by the work of Hirmer [23] where he had shown that silicon orientation does impact resisitivity. An overhaul of the SRP methodology was also undertaken and as a result the accuracy the SRPs improved considerably.

5.7.2 Epitaxy and Epitaxy/Implant split lot

The purpose of this process split was to evaluate the differences between the simulated process distributions and those obtained from silicon so that a process could be recommended. The process conditions for both processes were those previously recommended. A total of 24 wafers were processed; 12 wafers per process. The breakdown voltage was measured and the resulting histogram comparing both processes is presented in figure 5.28. The comparison of the two processes is summarised in table 5.6.

Process Option	Average BVdss (Volts)	Sigma (Volts)
Epitaxy	492.787	5.035
Epitaxy (simulated)	512.525	14.747
Epitaxy & Implant	437.849	20.929
Epitaxy/Implant (simulated)	465.757	9.0656

Table 5.6 Comparison of JHVH-450 epitaxy and epitaxy/implant processes

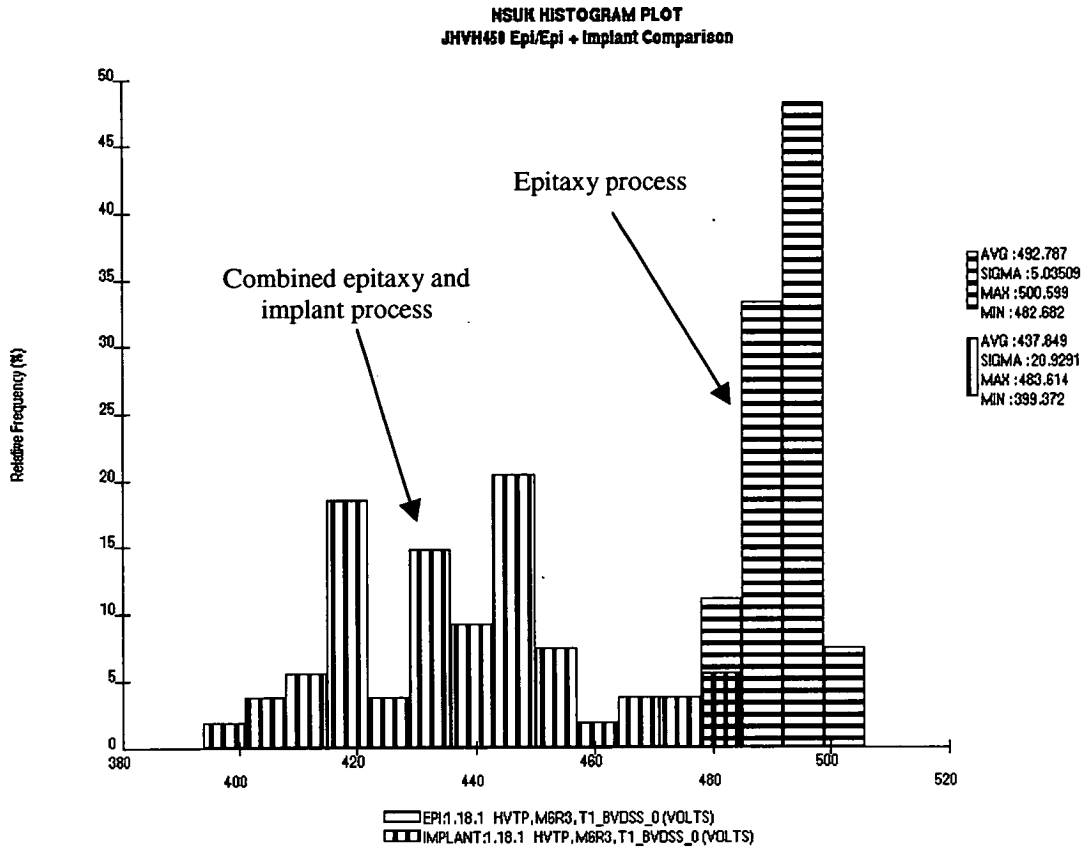


Figure 5.28 Histogram of BV_{dss} for JHVH-450 epitaxy and epitaxy/implant processes

The results show that both processes could meet the required specifications. However, there are some interesting observations. Firstly, the epitaxy process provides a tighter process than the epitaxy and implant flow. This was surprising since the implant control should have controlled the charge in the drift region better than in the epitaxy case where the charge was defined by the epitaxy process itself. The root cause was unknown. It was thought to be related to the higher resistivity epitaxy used for the implant process option, and possibly the epitaxy measurement methodology previously highlighted. The implant step itself was not totally discounted. Further work on characterising this 10-15Ω-cm epitaxy process would be required. Secondly, a comparison of the results to simulation shows comparable results for the average

values predicting a higher breakdown for the epitaxy process. Though the predicted variance from the simulated distributions indicated that the implant process would have yielded a tighter distribution.

5.8 Process Recommendation

The silicon results showed that the optimised processes, for both process options under consideration, did meet the process specification. Comparison of the silicon distributions did reveal some interesting results. The epitaxy process gave a better breakdown distribution than the combined epitaxy and implant. This was the opposite from what was expected, as the implant should provide superior control of the drift region charge and, therefore, the breakdown distribution. Investigation identified opportunities for improvement in both the metrology used in the epitaxy unit process and in the spreading resistance profiling methodology. Once the impact of these improvements are assessed then one would be in a better position to truly study which process option is best. Further work is required for this investigation to take place. It was also noted that there were tails in the simulated distributions for both process options. The reason for this is unknown but needs to be investigated.

5.9 Interaction of Post Fabrication Processing on Breakdown Voltage Drift in High Voltage RESURF Devices

5.9.1 Introduction

Even with all the controls in place during the fabrication process to ensure good quality device performance, there are situations where devices fail to meet their

specifications. One such example is post assembly, where the device is tested in the package [25]. All the main factors controlling the breakdown voltage are known (see chapter 3). However, the stability of the breakdown to processing conditions is not so well understood and the interaction between process and packaging cannot be forgotten.

5.9.2 Problem Definition

The breakdown mechanism is avalanche multiplication (section 3.2) and the criteria used to define the breakdown voltage is a drain current of $I_D = 100\mu A$. An example of a typical breakdown distribution is shown in figure 5.29 and these results are very reproducible.

Reliability testing requires that the devices pass a High Temperature Reverse Bias (HTRB) test which involves biasing the packaged devices with $V_G=V_S=0V$ and $V_D=320V$ at a temperature of 125°C. The breakdown voltage is then measured at room temperature and repeated at intervals until 1000 hours has elapsed. The results of this test is shown in figure 5.30 and clearly show that the breakdown voltage drops and devices fail the 400V minimum limit at the first time point of 24 hours. Additionally the spread in the results increases after HTRB testing has started. Note that these units were not those produced from figure 5.29.

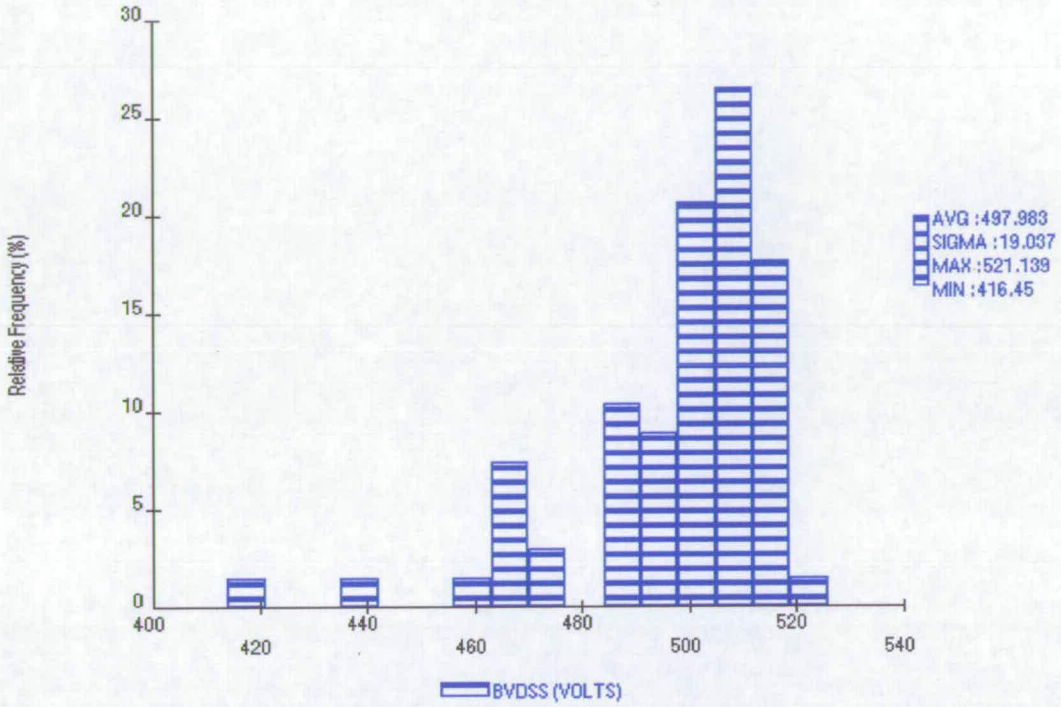


Figure 5.29 Typical voltage at wafer sort

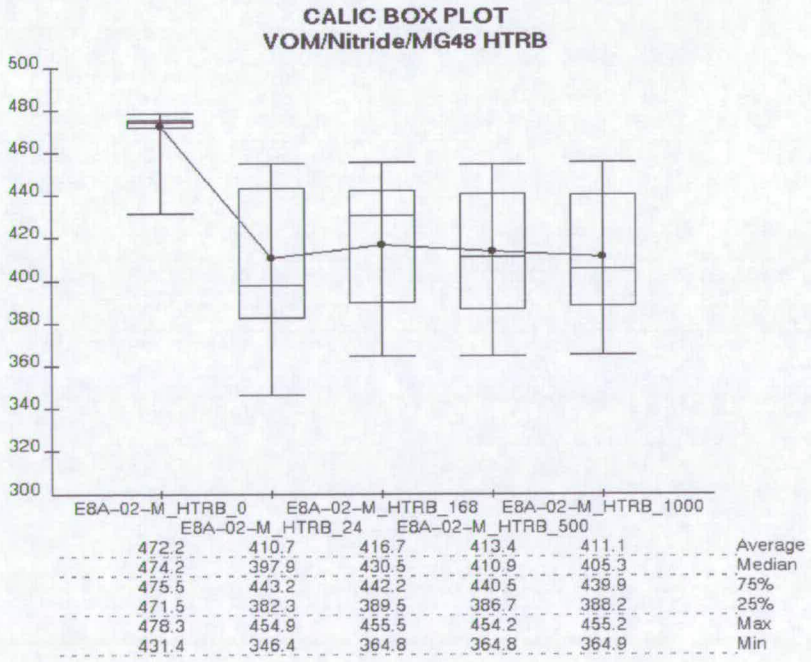


Figure 5.30 HTRB for standard passivation and packaging

5.9.3 Experimental Procedure

One option to solve the breakdown voltage drift problem during HTRB involves increasing the drift region length to give a greater initial voltage before HTRB testing. This however, reduces the area efficiency of the device and was, therefore, not considered viable. Hence, other options were investigated.

A number of screening experiments were designed to help determine the root cause of the voltage drift. These are summarised in table 5.7, together with the summarised effect.

Experiment	Range	Effect on Voltage Drift
Drift Implant Dose	1.2 - 1.6 $\times 10^{12}$	No Significant Effect
Final Alloy	Yes No	No Significant Effect
Mould Compound Type	B17 MG48	No Significant Effect
No Nitride Passivation	Yes No	Large effect
AMT5000 Low H Nitride	Yes No	No Significant Effect

Table 5.7 Experiments and effect on breakdown drift

This shows that the greatest impact was observed with the elimination of the nitride passivation, which is known to affect surface control breakdowns [26-28]. The optimum breakdown voltage is achieved when the peak electric fields in silicon are located beneath the source and drain field plates (see section 3.5); mobile hydrogen will affect this distribution and hence degrade the breakdown voltage. The HTRB results for devices that have no nitride and therefore no mobile hydrogen are shown in figure 5.31.

There are, however, problems with the oxide only passivation scheme. There is no moisture barrier and the yield was typically <40%, with the main fail mode being low breakdown voltage. Failure analysis using SEM cross sections revealed cracks in the oxide passivation layer at the edge of the field plates, as shown in figure 5.32. The wafer was processed through metal etch, which etches metal but not silicon to highlight the location of any pinholes (figure 5.33)

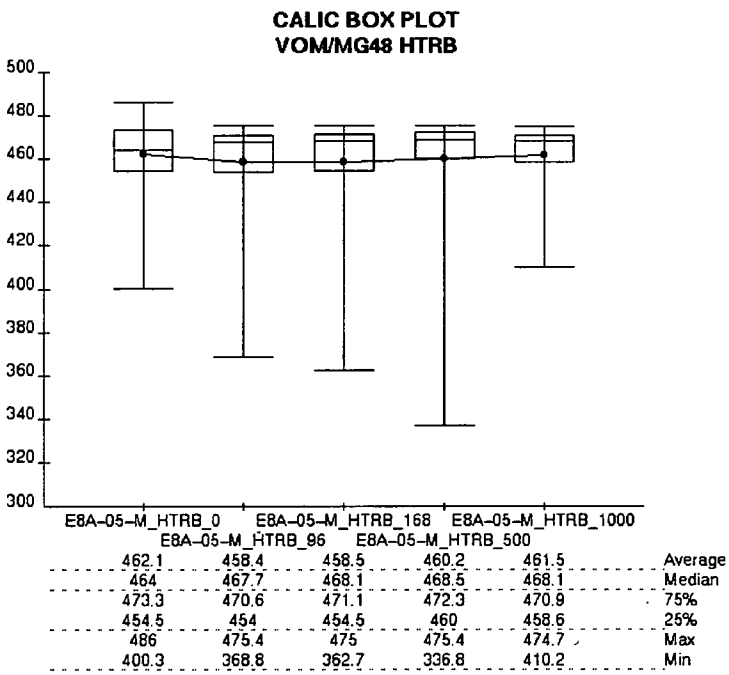


Figure 5.31 HTRB for oxide only passivation

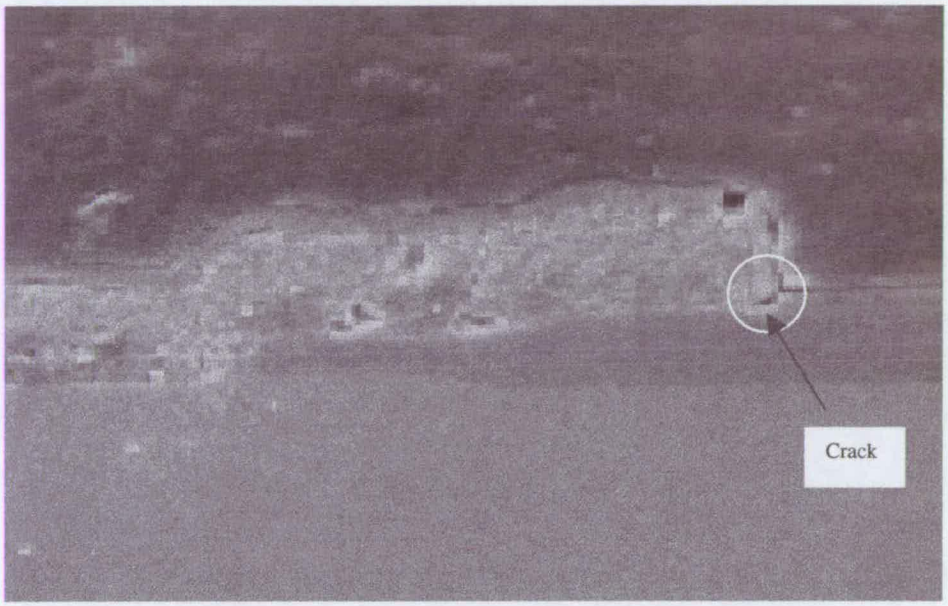


Figure 5.32 SEM of field plate edge for oxide only passivation showing cracks under the field plate

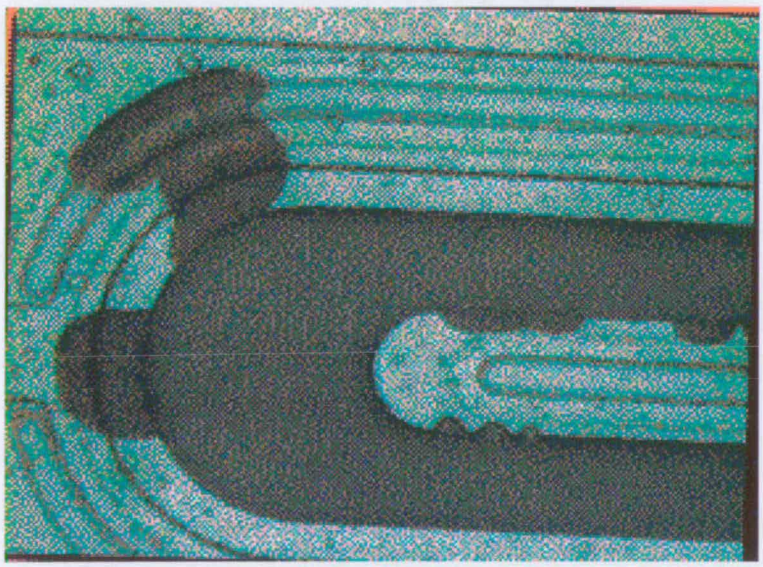


Figure 5.33 Fingertip region of device after KOH etch

Significant areas of metal have been etched from locations corresponding to the edges of the field plates (figure 5.33). This is consistent with the crack observed in figure

5.32, the position of which coincides with the position of the highest electric field in the oxide passivation. A simulation (Figure 5.34) of the same region as the SEM cross section (figure 5.32) confirmed this to be the case. Analyses of devices that had the additional nitride layer showed no cracks. Hence it was concluded that the physical damage in the oxide only passivation was the root cause of the low voltage breakdown failures.

Potential replacements for the nitride film were a low hydrogen film nitride process and an oxynitride process. The low hydrogen showed no improvement as summarised in table 5.7.

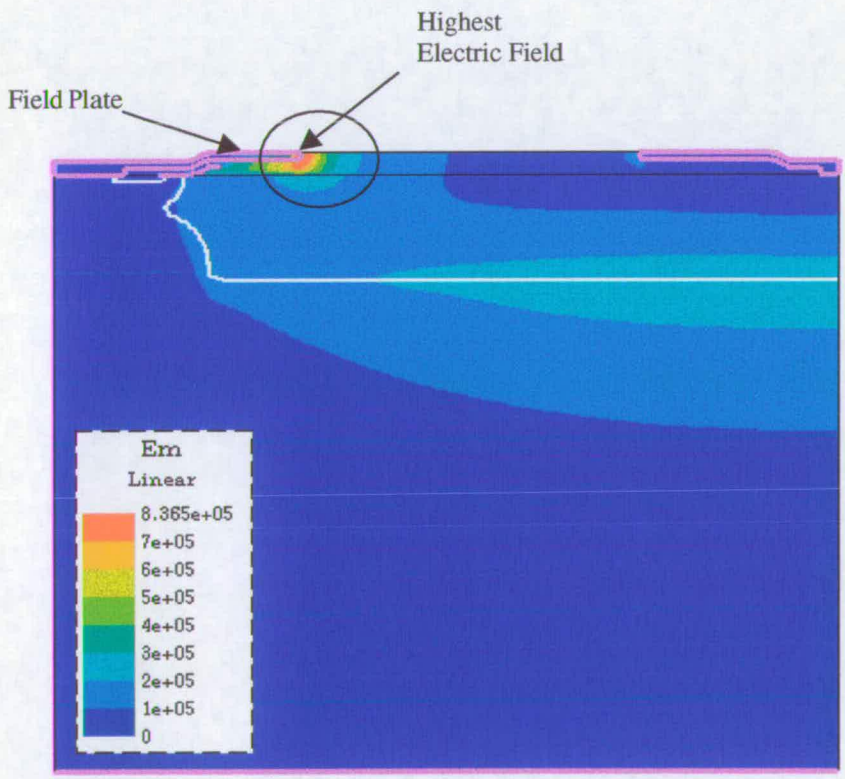


Figure 5.34 Device simulation: electric field distribution around the field plate

However, oxynitride showed very promising results. The best results were for low stress and low hydrogen. Figure 5.35 shows the HTRB results to 168-hour time point. An interesting observation is that the breakdown voltage starts lower than the other results but then increases to the 168-hour time point. Since all previous HTRB data levelled off after this time point, it is considered a good indication of 1000 hours results. However, further work would be required to confirm this. This result highlights that the interaction between process and packaging is an important piece of a design for manufacturability methodology.

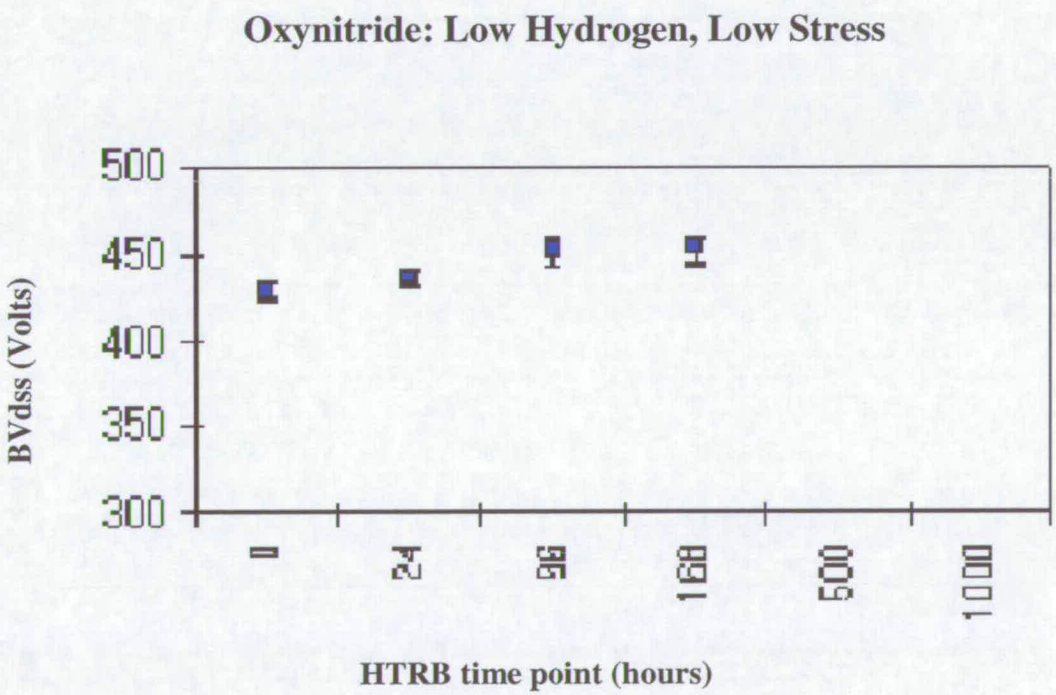


Figure 5.35 HTRB results for oxynitride passivated devices

5.10 Summary

Chapter 5 has demonstrated an optimisation strategy for two process options, both of which resulted in a process that consistently yielded 450 volts and a minimum, guaranteed breakdown of 400 volts. Process conditions were determined using Design Of Experiments (DOE) and Technology CAD (TCAD). The manufacturability of both flows were assessed using a Monte Carlo simulation approach by first determining the variability of the process control factors and using the contour plots obtained from the DOE results to predict the breakdown distributions. Both processes were run on silicon and the resulting distributions were compared to predicted distributions. Silicon results showed that the predicted optimised processes obtained using DOEs and TCAD did meet the desired specification. Comparing the silicon distributions revealed some interesting results. The epitaxy process gave a tighter breakdown distribution than the epitaxy/implant flow. This was the opposite of what was expected, since it was thought the implantation would give a superior control of the drift region charge and therefore, the breakdown voltage distribution. Investigations revealed that the process control methodology used at the epitaxial unit process stage was not optimum and the analytical measurements were extremely inaccurate. As a result there was an overhaul of the spreading resistance profile (SRP) measurement and test chip used was changed to a test wafer with the same crystal orientation as the wafers fabricated at the epitaxial deposition process stage. Also, in the epitaxy/implant process where the epitaxy resistivity was higher (10 - 15Ω-cm), the process control issues highlighted previously questioned the repeatability of the epitaxial unit process. It was determined that further characterisation of this process would be required as it was suspected that the manufacturing process engineers' assessment that this process was easily reproducible needed to be confirmed, if only

to determine its variability in a manufacturing environment. The process recommended was the epitaxy/implant even though the epitaxy flow gave a tighter distribution than the epitaxy/implant. Here it was felt that the tighter control of the charge in the drift region produced by the implant step would result in a better controlled process and would offset the additional cost with the added process step, despite the challenges seen with the epitaxial metrology. It was observed during the analysis that there were tails in the simulated distributions for both process options. Further work to ascertain which process conditions caused this to occur still need to be performed. Finally, the interaction of process and packaging has been discussed. This highlighted the fact that during the optimisation of a technology all facets of the semiconductor manufacturing process need to be considered.

References

1. H.W. Whittington, B.W. Flynn, D.E. Macpherson, "Switch mode power supplies", Research Studies Press, 1992.
2. LM3730B Design Review, National Semiconductor Corporation internal report June 15th, 1995.
3. B.J. Baliga, 'High Voltage Integrated Circuits', IEEE Press, 1984.
4. M. Redford, M. Fallon, Z.A. Shafi, N.S. Rankin, J. McGinty, A.J. Walton, "High Voltage RESURF DMOS Process Development Using DFM Techniques A Case Study", Proc 3rd International Workshop on Statistical Metrology pp 78-81, June 7, 1998.
5. E.M.S. Narayanan, G. Amaratunga, W.I. Milne, "A study of the RESURF principle for thin epitaxial layer high voltage integrated circuits", Proc 1992 International symposium on power semiconductor devices, pp172-175.
6. M.I. Newsam, "Calphurnia V2.0 Users' Manual", Department of Electrical Engineering, University of Edinburgh, Scotland.
7. E.J. Wildi, P.V. Gray, T.P. Chow, H.R. Chang, "Modeling and process implementation of implanted RESURF type devices", IEDM Digest of Technical Papers, pp 268-271, 1982.
8. M. Redford, M. Fallon, J. McGinty, A.J. Walton, M.I. Newsam, G. Michalowicz, C. McGregor. "Identifying and Quantifying Variability Within the LMDMOS 300A Process: A Case Study",IEE Colloquium on improving the efficiency of IC manufacturing technology, No. 1995/153, pp10/1-10/4, University of Edinburgh, April 12th 1995.

9. J.M. McGregor & M. El-Diwany, "450V lateral DMOS device module compatible with LMDMOS technology", Proc, National Semiconductor TCAD Workshop '93, pages 10-13, 1993.

10. M. Redford, S Irving, "Stanford TCAD Workshop", Stanford University, California, USA, August 1993.

11. M. Fallon, K. Findlater, J. McGinty, N. Rankin, A. Yarr, "Automatic furnace downloading to SUPREM format", Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 105-111, Edinburgh, Scotland, May 18-21 1999.

12. N.S. Rankin, M.Eng Hons Project HSP1110, "Characterisation and optimisation of LMDMOS 300B based processes with emphasis on a lateral high voltage RESURF DMOST", December 1995.

13. C. Roxburgh, M.Eng Hons Project HSP1111, "Characterisation of the LMDMOS 300B process", January 1996.

14. J. McGinty, G. Long, J.K. Kibarian, M. Redford, "Implementation of Statistical TCAD in a Manufacturing Environment", National Semiconductor Technology & Innovation Conference, pp 230-233, September 1996.

15. M. Redford, S. Irving, "NSC Technology CAD (TCAD) workshop", August 1992

16. Medici Manual, Avant! Corporation, Freemont, California, USA.

17. D. J. Roulston, "Bipolar Semiconductor Devices", McGraw Hill, 1990.

18. G. Rollins, Avant! Application Note "Medici model selection", Avant! Corporation, Freemont, California, USA.

19. D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis if Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", IEEE Trans. On Semiconductor Manufacturing, Vol 9, No. 4, pp 478-488, Nov 1996.

20. RS/1 Manual, Domain Software Solutions, Chelmsford, MA, USA.

21. J. Altrip, "Spreading Resistance Improvements", NSUK Internal Memorandum, January 1995.

22. C. McGregor, "Epi Monitor Improvement", NSUK Internal Memorandum, December 1994.

23. P. Hirmer, "The Effect of Temperature and Substrate Orientation upon the Incorporation Rate of Boron and Arsenic in Silicon Epitaxial Layer Growth", Journal Electrochem Soc., Vol 136, No. 11, November 1989.

24. F. Conti, C. Conti, "Surface Breakdown in Silicon Planar Diodes Equipped with Field Plate", Solid State Electronics, 1972, Vol. 15, pp 93-105.

25. Z.A. Shafi, R.Busse, R. Orr, J.M. McGregor, M. Redford, "Minimising Breakdown Voltage Drift in High Voltage (>400V) Lateral DMOS RESURF devices", National Semiconductor Technology & Innovation Conference, pp 196-199, September 1996.

26. J. Dunkley, G, Ganschow, D. Hannaman, J. Patterson, S. Willard, P. Gopi, "Hot electron induced hydrogen compensation of boron doped silicon resulting from emitter-base breakdown", Proc IEDM, pp 785-788, 1992.

27. D. Quon, P.K. Gopi, J. Sonek, G.P. Li, "Hot carrier induced bipolar transistor degradation due to base dopant compensation by hydrogen: theory and experiment", IEEE Trans Electron Devices, vol. ED-41, pp 1824-1830, 1994.

28. N.M. Johnson, "Mechanisms for hydrogen compensation of shallow acceptor impurities in single crystal silicon", Phys. Rev B, vol 31, pp 5525-5528, 1985.

Chapter 6: JHVH-700 Process Development

6.1 Introduction

The development of National Semiconductor's first 700V lateral DMOS process was to be a single metal variant of the LMDMOS smart power process produced in the 6" line at Greenock. As with the 400V process, not all the process steps in 300B are required, therefore, the same process stages are eliminated as for the 400V process described in section 4.3. The same unit process steps control the breakdown voltage [1,2]: epitaxy resistivity and thickness, and the start material resistivity.

The purpose of the 700V process is to provide a universal solution for both 110 and 240V power supplies. These applications experience a peak V_{ds} of twice the rectified AC voltage plus a voltage spike generated at turn-off. For a 240 VAC, the minimum acceptable BV_{dss} is, around 700V and as these devices produced by the process will also be able to withstand 110 VAC, they are termed the universal solution.

In this chapter, the combined epitaxy and implant flow was the only process option considered. This was due to the voltage requirements which it was known could not be met by the epitaxy only process as the control in manufacturing of epitaxial dopant concentration required to achieve the desired breakdown voltage could not be achieved. The optimum process settings for the epitaxy and implant process was determined using of design of experiment methodologies in combination with process and device simulation tools [3]. The manufacturability of the resultant process was determined by

predicting the distributions of the breakdown voltage, BV_{dss} and these results were compared to those obtained from silicon.

6.2 JHVH-700 Process

The process was based on the JHVH-400 flow described in chapter 3 with the main differences resulting from the higher breakdown voltage requirements. High breakdown voltages in relatively thin epitaxy are made possible by application of the RESURF (Reduced SURFace Field) principle described previously. It was required that this device could be integrated into the LMDMOS process, which resulted in several constraints. These were that the epitaxial thickness had to remain the same and existing process steps were to be used. It was desired that as few steps as possible should be added that would adversely impact other device parametrics when the high voltage device was integrated into the smart power process flow. A cross section of the RESURF device used is shown in figure 6.1.

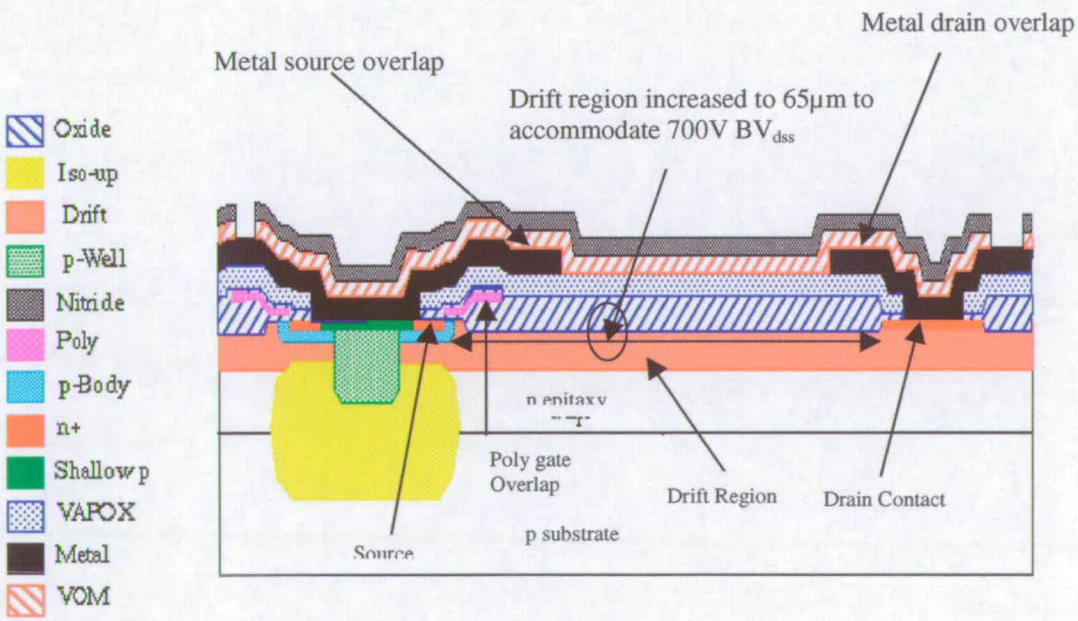


Figure 6.1 Lateral DMOS cross section with source shorted to substrate for JHVH-700

The layouts of the 400 and 700 volt transistors are selectively scaled to accommodate the increased breakdown voltage. Consequently, the drift region length for the JHVH-700 transistor is 65 µm, and the drain and source overlaps are 8µm and 10µm respectively. A tabulated comparison between the JHVH-400 and 700 volt processes is detailed in table 6.1.

		Drift Region (µm)	Field Plates (µm)	
			Source	Drain
JHVH-450	Parallel Region	30	5	8
	Fingertips	60	15	8
JHVH-700	Parallel Region	65	10	8
	Fingertips	130	30	8

Table 6.1 JHVH-450 & 700 transistor layout comparison

Figure 6.2 presents the results for a Medici simulation which shows a cross section of the applied potential and how it is spread across the n^- region at the source fingertip. It can be observed that the potential contours are spread between the source and drain of the structure thus demonstrating the RESURF principle described in chapter 3 and shown in figure 3.15 for the JHVH-400 process. The potential contours can be seen crowding at the edge of the source field plate demonstrating that the field plate performs as designed. This will eventually lead to the device breaking down as is shown in figure 6.3 where the impact ionization contours highlighted in red show the region of maximum impact ionization beneath the edge of the source fingertip.

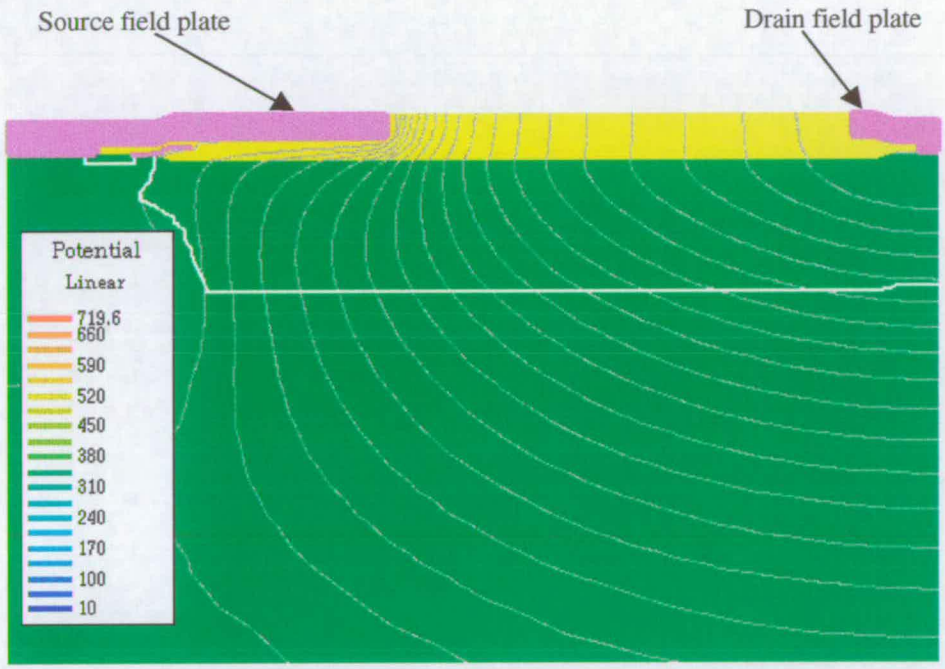


Figure 6.2 Medici device simulation of the source fingertip structure showing applied voltage evenly spread between the source and drain

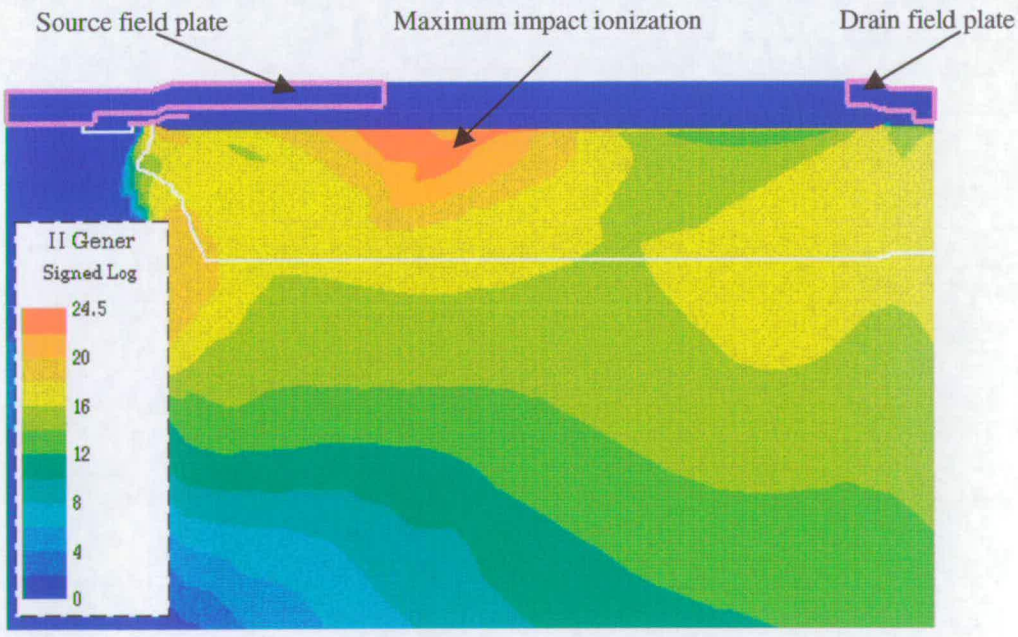


Figure 6.3 2D plot of impact ionization for JHVH-700 source fingertip structure at breakdown

6.3 Covariance Based Interpolation

6.3.1. Introduction

Response surface methodology is a method of accurately predicting the effect of multiple control factor changes on the response(s) of interest. Such response surfaces are polynomial based, typically quadratic, and are used to determine the trend in the response. The region of interest is usually fitted by the polynomial using a least squares approach [6]. This procedure is ideally suited to situations in which the observations are affected by random error where trends rather than precise fits are most meaningful [7]. However, results from simulators do not contain random error and, therefore, models which only determine the trend may miss out the relevant detail. The capability to model wider ranges of control factors is potentially very useful in the initial stages of process development when it is advantageous to fully examine the parameter space. This approach can also be beneficial when performing multi-parameter optimisation. Here optimum control factor combinations may be separated in the parameter space and the ability to model over larger control factor ranges can be useful.

6.3.2 Covariance

One of the problems with using polynomial models is that sometimes they cannot be distorted to fit the data. In this case, it is possible to use a covariance model [4]. Covariance is a method-based on a technique called kriging [8] by both modelling the trend and interpolating the observations so that the fitted model exactly reproduces the observations. Work by Waring et al [4,5] concentrated on the interpolating version of kriging. Their work optimised and verified the covariance technique as applied to VLSI

technologies for optimisations with two process variables. With experimental data [8], the trend is first fitted with a polynomial with the second step being a gentle distortion of the trend so that the response goes through selected data points that are believed to reflect reality as will be discussed in section 6.7. It is important the covariance-based approximation is used carefully and does not alter the surface in such a manner that fluctuations in measurement due to noise are modelled. In this event, the capability of the model to predict the response at other points would be compromised. However, as mentioned previously results from simulators do not contain random error and, therefore, there is justification in fitting the response through all data points.

The response surface for the breakdown voltage often presents a severe challenge for any function to fit as it consists of areas that have very rapidly varying gradients [5]. It can be observed that the variation of BV_{dss} as shown in figure 6.4 presents such a challenge and it is in situations such as this that the covariance approach has potential benefits.

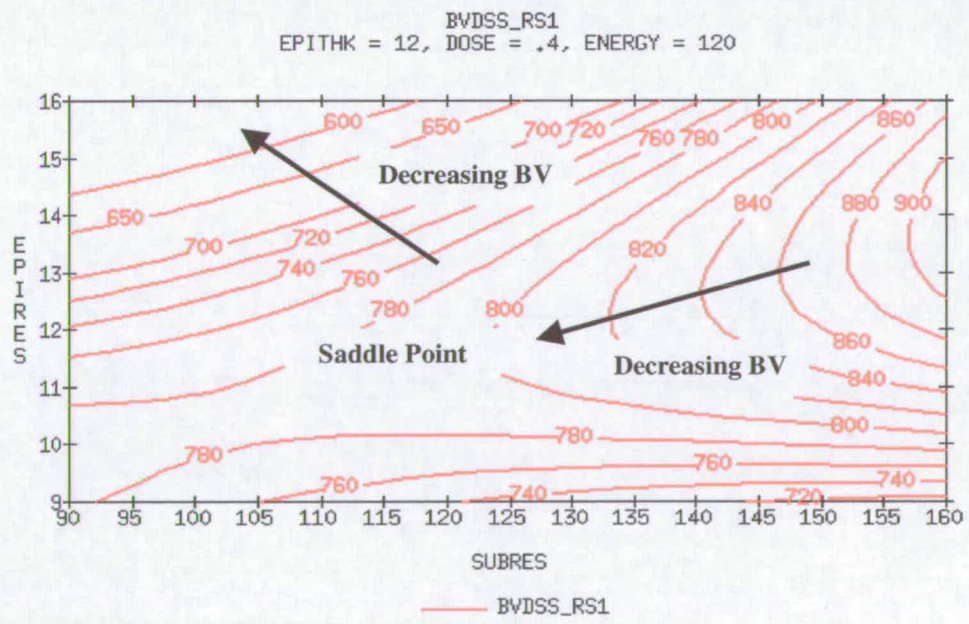


Figure 6.4 BV_{dss} contour plot for simulated data polynomial model

It will be shown later how covariance can improve the fit to get a more representative model for the simulated data.

6.4 Process Optimisation

6.4.1 Introduction

As discussed in sections 6.1 and 6.2, the target minimum breakdown voltage was 700 volts, and the constraints placed upon the process were that the device had to be able to be integrated into the LMDMOS technology. Therefore, the epitaxial layer thickness had to remain the same at 11-13μm. Retaining this thickness specification limited the scale of process change possible, and as only one process option was considered – combined epitaxy and implant. As mentioned previously, the dopant concentration control in the drift region required to achieve the desired breakdown specification could only be performed using ion implantation.

6.4.2 Optimisation Methodology

As covered in chapters 3 and 4 and [2], the process parameter settings that could be optimised were:

(a) Epitaxial layer resistivity

(b) Epitaxial layer thickness

(c) Substrate resistivity

(d) Implant dose

(e) Implant energy

Even though the epitaxial layer thickness was required to be the same as the LMDMOS process, the impact of this parameter on the breakdown voltage from a manufacturing standpoint had to be assessed, along with any interactions with other process variables. This was also the case for all the other parameter settings that were optimised.

In order to optimise the variables described to produce a process with the required specification for the breakdown voltage in a manufacturing environment, Design Of Experiments (DOEs) methodologies were used, in conjunction with Technology CAD (TCAD) tools as discussed in chapter 2. This provided a 'window of operation' for the technology from which the manufacturability of the technology could be assessed by determining the variability of each of the control factors. Using Calphurnia [11] this analysis assumed that the control factors varied in a Gaussian manner. This enabled a Monte Carlo analysis on the resultant DOE response surface to predict the breakdown voltage response distribution. The variability of the respective process parameters was determined from in-line measurements, electrical test data, and discussions with equipment and process engineers. Previous work [2,12-14] provided confidence in the values used. This design for manufacturability strategy is the same as used for the JHVH-450 process as described in section 5.2 and figure 5.2.

6.4.3 Combined Epitaxy and Implant Process

A Central Cubic Inscribed (CCI) designed experiment was constructed using RS/1 [15] and the simulation results for the resulting breakdown voltage, BV_{dss}, recorded in table 6.2. The process settings were varied as follows:

1. Epitaxial resistivity – 10 to 15Ω-cm
2. Epitaxial thickness – 10.5 to 13.5μm
3. Substrate resistivity – 100 to 150Ω-cm
4. Drift implant dose – 0.3 to 0.7×10^{11} cm^{-2}
5. Drift implant energy – 115 to 125keV

A wider range of epitaxial layer thickness than desired was investigated in order to assess any manufacturability challenges that may occur as a result of integrating the RESURF device into the LMDMOS flow. It was decided that the epitaxial resistivity would remain the same as that used for the 400 volt process as this would result in the most efficient use of process equipment if both the 400 and 700 volt processes were to be manufactured in the same line. The breakdown voltage used in the analysis was the minimum of the three simulated values for the source fingertip, drain fingertip, and parallel regions as described in section 5.5.2 and figure 5.11. The drift region was kept at 65μm for the parallel region and 130μm for the fingertips, which were also simulated using radial co-ordinates as described in section 5.5.2. The field plates in the parallel region were 10μm and 8μm at the drain and source respectively. These were extended at the fingertips to 30μm. The resulting worksheet is shown in Table 6.2.

Run #	Subres (Ω-cm)	Epithk (μm)	Epires (Ω-cm)	Dose ($\times 10^{11}$ cm^{-2})	Energy (keV)	BV_{DSS} (Volts)
1	107.5	11.25	10.75	0.4	122.5	800
2	142.5	11.25	10.75	0.4	117.5	860
3	107.5	12.75	10.75	0.4	117.5	780
4	142.5	12.75	10.75	0.4	122.5	860
5	107.5	11.25	14.25	0.4	117.5	720
6	142.5	11.25	14.25	0.4	122.5	800
7	107.5	12.75	14.25	0.4	122.5	740
8	142.5	12.75	14.25	0.4	117.5	800
9	107.5	11.25	10.75	0.6	117.5	840
10	142.5	11.25	10.75	0.6	122.5	580
11	107.5	12.75	10.75	0.6	122.5	740
12	142.5	12.75	10.75	0.6	117.5	580
13	107.5	11.25	14.25	0.6	122.5	820
14	142.5	11.25	14.25	0.6	117.5	860
15	107.5	12.75	14.25	0.6	117.5	820
16	142.5	12.75	14.25	0.6	122.5	860
17	90	12	12.5	0.5	120	760
18	160	12	12.5	0.5	120	860
19	125	10.5	12.5	0.5	120	800
20	125	13.5	12.5	0.5	120	860
21	125	12	9	0.5	120	600
22	125	12	16	0.5	120	780
23	125	12	12.5	0.3	120	720
24	125	12	12.5	0.7	120	600
25	125	12	12.5	0.5	115	840
26	125	12	12.5	0.5	125	840
27	125	12	12.5	0.5	120	840

Table 6.2 Worksheet for CCI designed experiment

The results were analysed in RS/1 and the response surface achieved is shown in figure 6.5. This indicates that there is a 'workable' region that meets the required specification. However, this analysis does highlight that the breakdown voltage will be a challenge to control. The R^2-adj, measure of the quality of the model fit was 0.833.

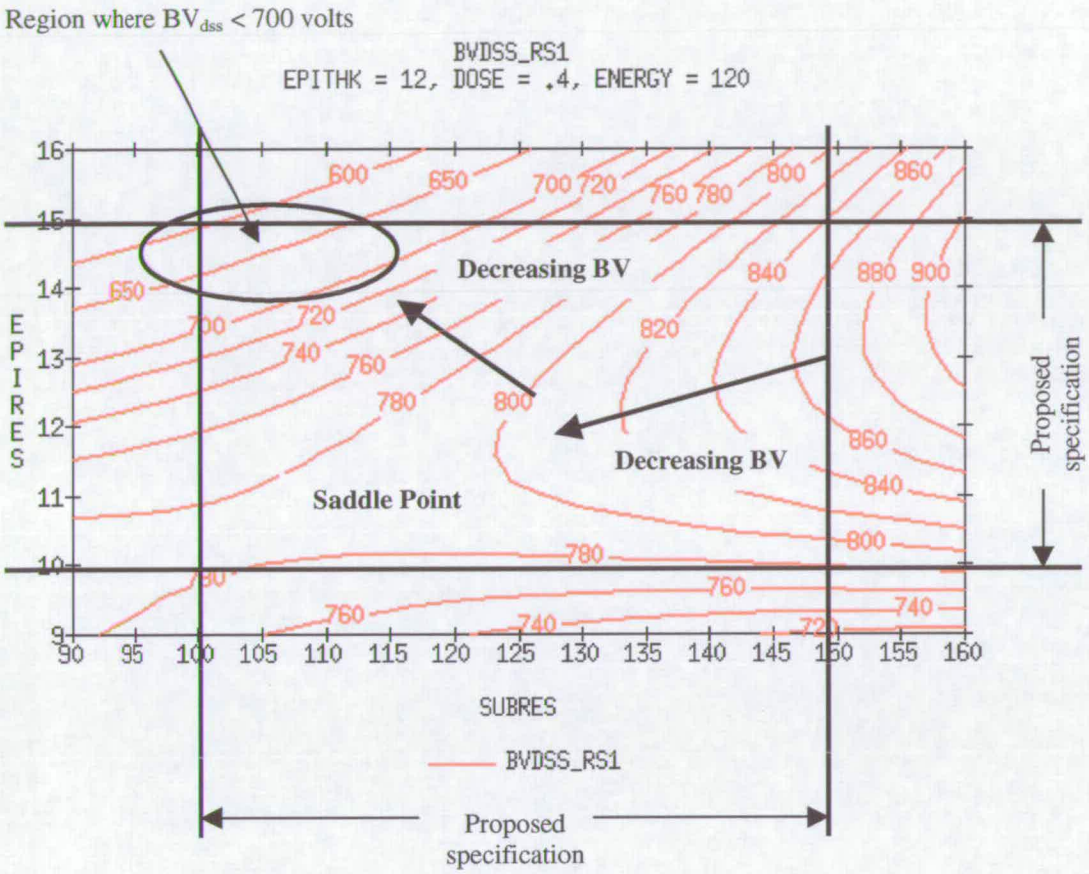


Figure 6.5 BV_{dss} polynomial response surface for the simulated data

Figure 6.6 compares the quadratic model fitted in figure 6.5 with a covariance-based model fitted using Calphurnia [11]. In this case, it is apparent that the polynomial surface has not fit the worksheet data precisely. This is even more evident in the comparison performed for drift implant doses of 0.5 and 0.6×10^{-11} cm^{-2} shown in figures 6.7 and 6.8 respectively. Examination of the response surfaces revealed that for the quadratic model there was a data point where the residual percentage error is in double figures (run 9). The simulation results were double-checked and found to be correct. Remembering that simulations do not contain random error, therefore, there is justification in applying the

covariance procedure to the response surfaces. A comparison of the response surfaces for polynomial and covariance fits identifying the control factor settings for run 9 is shown in figure 6.8. This demonstrates how the covariance effectively distorts the surface so that the response goes through the data point generated by run 9 (840V) whereas the quadratic model gives a value of 800V. It can be observed that the saddle point has disappeared from the area of interest and this highlights how application of the covariance fit has the potential to improve the accuracy of prediction.

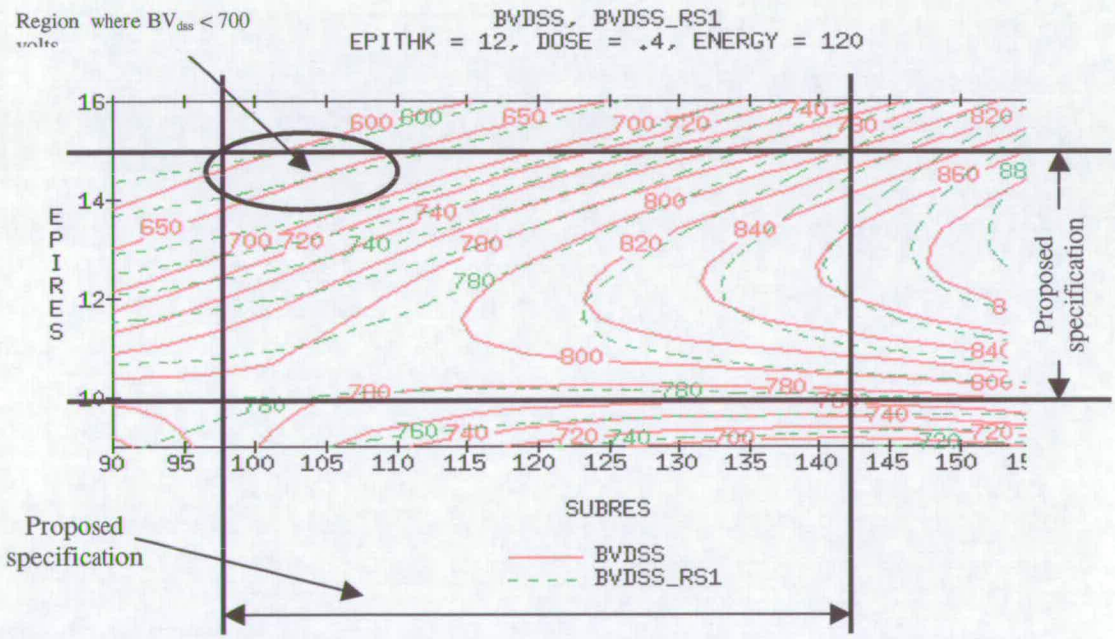


Figure 6.6 Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} for a dose of 0.4×10^{11} cm^{-2}.

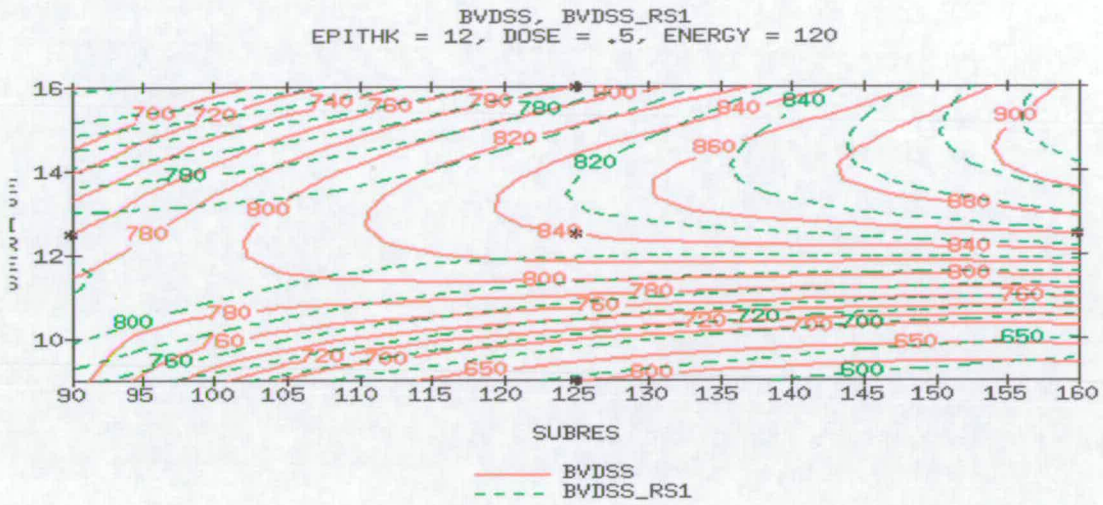


Figure 6.7 Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} for a dose of 0.5×10^{11} cm^{-2}.

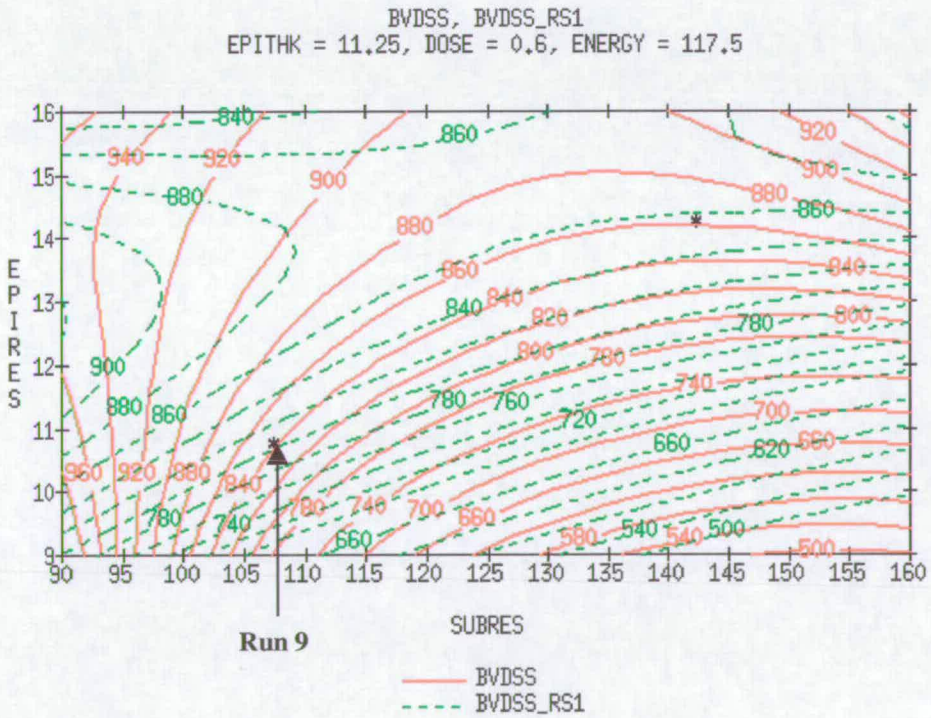


Figure 6.8 Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} indicating the combination of control factors that make up run 9.

6.4.4 Breakdown Distributions

The ability to create accurate response surfaces is an important requirement of Design For Manufacturability (DFM) since they have potential to provide important information on the manufacturability of the process under consideration. Calphurnia [11] can use the response surfaces to generate response distributions resulting from the expected variations in the control factors that are typically seen in a production environment, as described in chapter 5. Using the same variation specified in reference [14] for the control factors, response distributions have been generated using both covariance and polynomial models are shown in figure 6.9 and 6.10 respectively. Table 6.3 compares parameters for the distributions obtained from two different implant doses for both covariance and quadratic based distributions. It can be observed that the covariance means are higher but this is of no statistical significance.

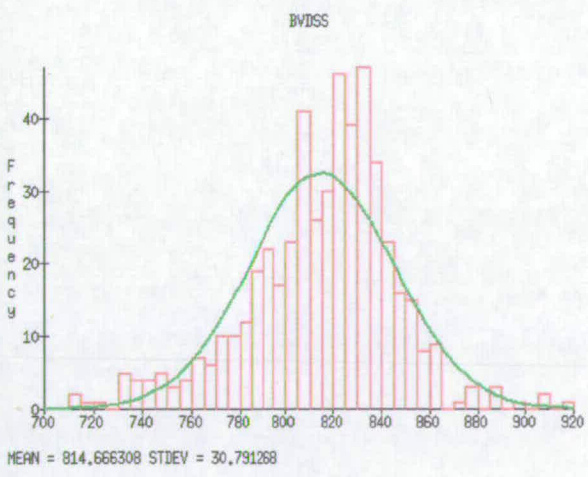


Figure 6.9 VB_{DSS} distribution calculated using the covariance model

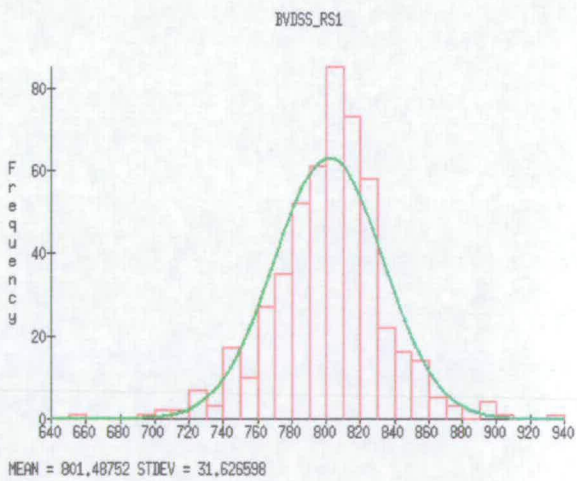


Figure 6.10 VB_{DSS} distribution calculated using the quadratic

	Dose - 0.4×10^{11} cm^{-2}		Dose - 0.5×10^{11} cm^{-2}	
	(Mean) (V)	BV_{DSS} (Std Dev) (V)	BV_{DSS} (Mean) (V)	BV_{DSS} (Std Dev) (V)
Quadratic Model	801	32	812	30
Covariance Model	814	31	825	29
Difference	13	1	13	1

Table 6.3 Comparison of parameters for the distributions obtained for two different implant dose settings. The covariance and quadratic based distributions for the 0.4×10^{11} ions cm^{-2} dose are shown in figures 6.9 and 6.10

Both polynomial and covariance based distributions using the target values for run 9 are shown in figure 6.11 and in this case it can be observed that their shape is different with the covariance distribution exhibiting a longer tail for high breakdowns. Table 6.4 compares the distribution parametrics associated with the two response models. As would be expected from the contours shown in figure 9 there is a big difference in the mean values. However, there is also a significant difference in the variance. If manufacturability is to be considered then the variability of the response needs to be predicted as accurately as possible.

	Run 9 target values	
	BV_{DSS} (Mean) (V)	BV_{DSS} (Std Dev) (V)
Quadratic Model	803	98
Covariance Model	874	140
Difference	71	42

Table 6.4 Comparison of parameters for the distributions obtained from covariance and quadratic based distributions using run 9 control factor settings as the target values. The distributions from which these parameters extracted are shown in figure 12.

Covariance Based Distribution

Polynomial Based Distribution

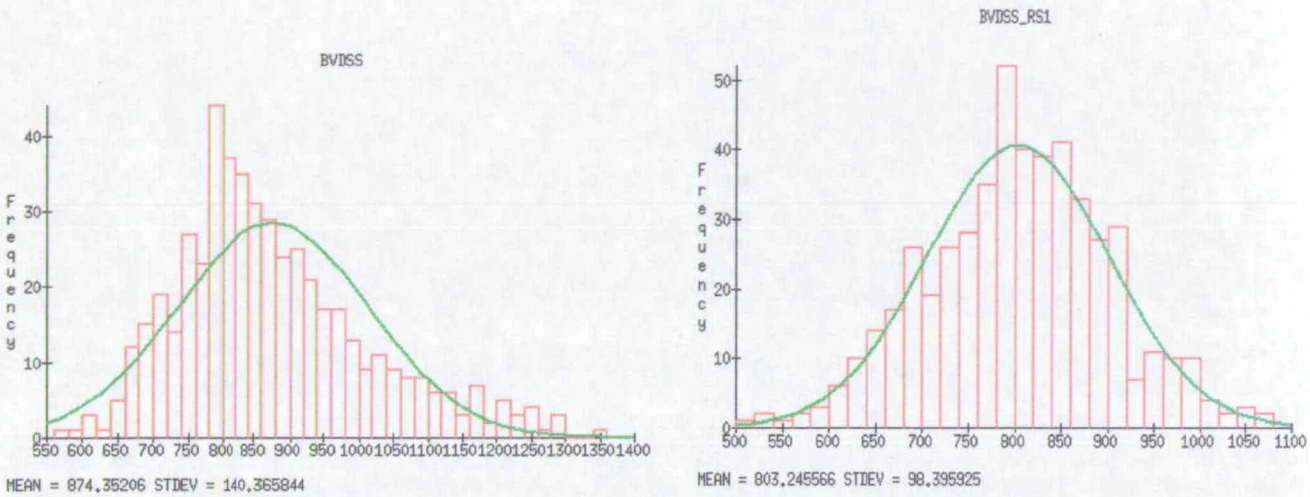


Figure 6.11 VB_{DSS} distribution calculated using the covariance and quadratic model

6.4.5 Experimental Results From Silicon

Figure 6.12 shows the measured distributions for different implant doses of 0.4, 0.5 and 0.6 $\times 10^{11}$ ions cm^{-2} and it can be observed that there appears to be two distributions present in the data. The simulation results predict the 'major' distribution but not the failures that have lower values of breakdown voltage. It would appear that there is a second factor in operation that is not being modelled by the TCAD simulations. It is known that the layout of the device has an impact on the breakdown. Field plates at the source and drain are used to minimise the two and three-dimensional field crowding to help maximise the breakdown voltage. Non-optimal design of these will lead to occasional lower breakdown voltage than would otherwise be expected as would process issues such as field oxide thickness beneath the field plates. These issues obviously need further investigation but what has been identified from the TCAD work is that the

developed process is intrinsically capable of meeting the required specifications and the low measured breakdown voltages are not due to variability in the control factors investigated in the experiment.

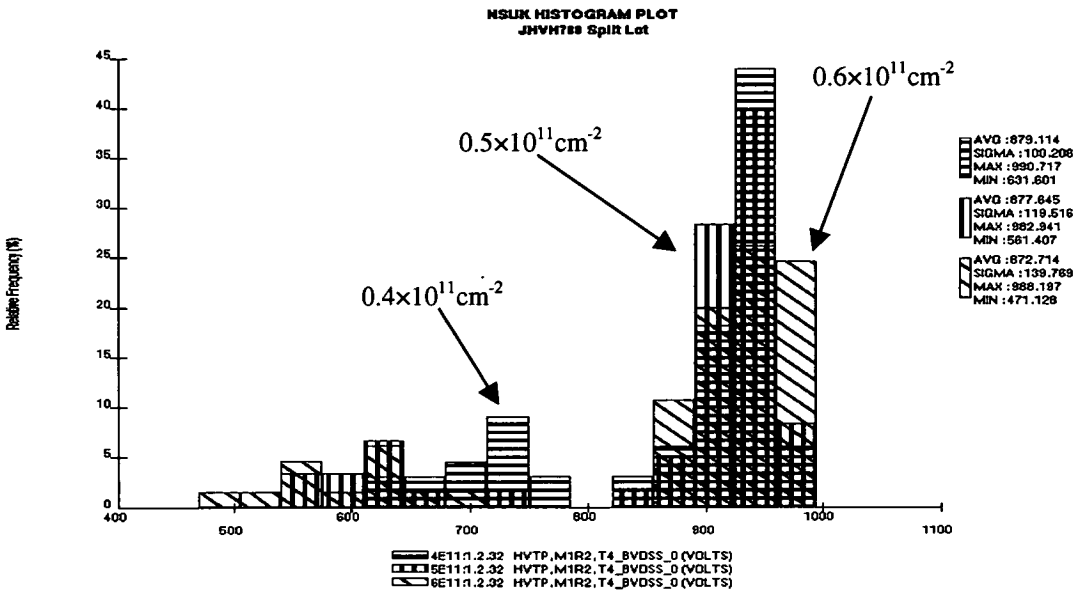


Figure 6.12 Breakdown Distribution from Silicon for Drift Implant Doses

6.5 Observations

The comparison of conventional quadratic polynomial response surfaces with covariance fits has shown that near data points where there are small residual errors the difference in predictive capability is small. However, when the quadratic polynomial exhibits large residual errors near data points the use of covariance based models provides a significant improvement. It has been demonstrated that the covariance model can be successfully fitted to a CCI design with five control factors.

The distributions generated from the TCAD results have confirmed that the variation observed experimentally is caused by factors other than those used as control factors in the CCI experiment. Further work is required to fully identify the causes of lower breakdown voltages.

Such work could include but is not limited to the following areas:

(a) Study of other control factors – field plate overlap at source and drain fingertips

(b) Investigation of non normal control factor distributions when predicting breakdown voltage distributions

(c) The methodology used to simulate BV_{dss} at the source and drain fingertips. This introduces some inaccuracy, which can be seen with the aid of a schematic for the source fingertip in figure 6.13.

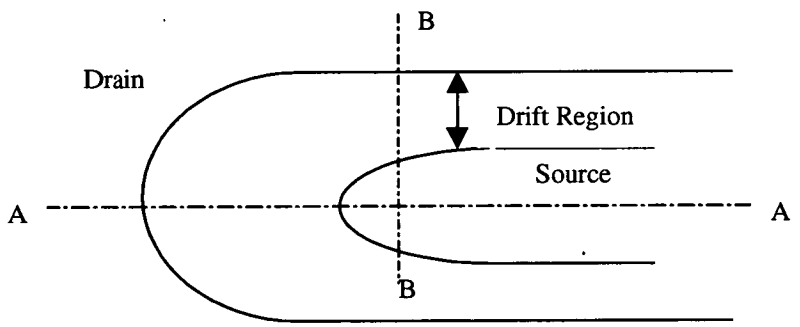


Figure 6.13 Schematic of source fingertip showing changing drift region dimension around source the fingertip

The drift region length is twice as wide at the extreme of the fingertip, section A-A, compared with its length adjacent to the parallel region, section B-B. The simulations performed in this work used the most pessimistic case using cylinderical coordinates with section B-B.

6.6 Application of Covariance Fits to Experimental Data

6.6.1 Introduction

Experimental designs in conjunction with Response Surface Methodologies are important tools that can be employed to help optimise IC processes [17] in a manufacturing environment. A key to the success of this approach is the accuracy with which the model can be fitted to the data. One of the problems with fitting experimental measurements is that there is uncertainty associated with each data point. Consequently, unlike the situation with simulated data, there is no justification in fitting the response surface through all the measurement points [18]. Therefore, the trend is typically fitted using a polynomial based function which effectively smoothes the response if experimental error is present [19]. This section presents a method of improving the fit of response surfaces for experimental data when one or more data points are poorly fitted and there is confidence in the fact that these points are not anomalous.

6.6.2 Background

As mentioned previously there are occasions when a low order polynomial cannot adequately fit one or two data points that are known not to be rogue measurements. Transformations are one method that is traditionally used to improve the fit but, if this is not successful, the only other option is to increase the order of the polynomial. This latter option can, in the limit, result in the response surface ceasing to fit the trend. Another possible solution is to use a covariance-based fit, which ensures that the resulting function passes through all the data points [5]. As discussed previously, this is a quite acceptable approach for computer based experiments as simulation results have no

experimental error associated with them (the same results are obtained every time the program is run). However, this approach is inappropriate for experimental data because it will fit the surface to the errors when only a fit to the overall trend can be justified.

6.6.3 Improving the Fit to Experimental Data

The solution is a two-stage process. First the data is fitted with a polynomial and the problem points identified where there are large residual errors. These points are then examined and if it is decided that they are not rogue measurements then a hybrid polynomial-covariance [4,5] fit is employed. A new response table is created by adding the experimental data points where there are large residuals to the worksheet with the rest of the required data being generated from the polynomial fitted to the original data set. The effect of this is to ensure that the trend is fitted while at the same time distorting this response surface through the data points that were deemed to be poorly fitted by the original quadratic.

This procedure is best illustrated through an example. The worksheet for a full factorial experiment with three control factors and a single response is shown in table 6.5. When the response is fitted with a second order polynomial the Adj R^2 value is 0.989. Table 6.6 presents information on how well each data point is fitted. Taking the traditional approach, the error in the fit may be reduced by applying a transformation. In this case a square root transformation is recommended, and this increases the Adj R^2 value to 0.994.

Run no.	Slope (°)	Dose (cm^{-2})	Energy (keV)	R_{SD} (Ω)
1	70	1e+13	30	1420
2	90	1e+13	30	760
3	110	1e+13	30	425
4	70	5.05e+14	30	1110
5	90	5.05e+14	30	620
6	110	5.05e+14	30	315
7	70	1e+15	30	1020
8	90	1e+15	30	490
9	110	1e+15	30	245
10	70	1e+13	60	1170
11	90	1e+13	60	720
12	110	1e+13	60	415
13	70	5.05e+14	60	1090
14	90	5.05e+14	60	600
15	110	5.05e+14	60	305
16	70	1e+15	60	1000
17	90	1e+15	60	480
18	110	1e+15	60	225
19	70	1e+13	90	1150
20	90	1e+13	90	700
21	110	1e+13	90	405
22	70	5.05e+14	90	1050
23	90	5.05e+14	90	590
24	110	5.05e+14	90	295
25	70	1e+15	90	940
26	90	1e+15	90	475
27	110	1e+15	90	210

Table 6.5 Worksheet for 3 factor 3 level full factorial experiment

Figure 6.13 shows the response surface resulting from the polynomial fit and it can be observed that R_{SD} does not reach 1420Ω when the dose, energy and side-wall slope are 1×10^{13} cm^{-2}, 30keV and 70° in run 1.

Run no	Expt Data	Pol`y RSM model	Absolute Residual Error	% Error
1	1420	1260	160.00	11.27
2	760	752.22	7.78	1.02
3	425	435.97	10.97	2.58
4	1110	1166.25	56.25	5.07
5	620	636.67	16.67	2.69
6	315	320.42	5.42	1.72
7	1020	1050.69	30.69	3.01
8	490	521.11	31.11	6.35
9	245	204.86	40.14	16.38
10	1170	1221.11	51.11	4.37
11	720	719.44	0.56	0.08
12	415	431.11	16.11	3.88
13	1090	1105.56	15.56	1.43
14	600	603.89	3.89	0.65
15	305	315.56	10.56	3.46
16	1000	990	10.00	1.00
17	480	488.33	8.33	1.74
18	225	200	25.00	11.11
19	1150	1160.42	10.42	0.91
20	700	686.67	13.33	1.90
21	405	426.25	21.25	5.25
22	1050	1044.86	5.14	0.49
23	590	571.11	18.89	3.20
24	295	310.69	15.69	5.32
25	940	929.31	10.69	1.14
26	475	455.56	19.44	4.09
27	210	195.14	14.86	7.08

Table 6.6 Residual errors for the polynomial model

In this example there are three data points where the residual percentage error is in double figures (run 1, 9 and 18). We make the assumption that only the data point associated with runs 9 and 18 have significant experimental error associated with them and there is high confidence that the data associated with run 1 is not a rogue measurement. Hence it is very desirable to create a response surface that passes through this point.

As previously proposed, the first stage is to fit the data with a standard polynomial. This function is used to complete the run table shown in table 6.7 for all data points except run 1 for which the original value of experimental measurement is entered. This data is now fitted using covariance which creates a response surface that passes through all the data points and the resulting contour plots are shown in figure 6.14. It should be noted that if there was confidence that all three runs (1, 9 and 18) were just poorly fitted by the polynomial model (i.e. not due to experimental error) it would have been quite possible to select all three data points for the covariance fit.

Comparison of figures 6.13 and 6.14 demonstrates how the covariance-based response surface has only been distorted from the "polynomial based trend" around the single point with the response surface perfectly fitting the other data points (the polynomial surface).

The procedure is designed to enable the method to be used on experimental data and removes the need for transformations. This in turn enables the user to precisely select which measurements are exactly fitted by the model. The practical application of this approach which has been discussed here has clearly demonstrated its feasibility.

Run no	Original data (R_{SD})	Covariance data (R_{SD})
1	1420	1420
2	760	752.22
3	425	435.97
4	1110	1166.25
5	620	636.67
6	315	320.42
7	1020	1050.69
8	490	521.11
9	245	204.86
10	1170	1221.11
11	720	719.44
12	415	431.11
13	1090	1105.56
14	600	603.89
15	305	315.56
16	1000	990
17	480	488.33
18	225	200
19	1150	1160.42
20	700	686.67
21	405	426.25
22	1050	1044.86
23	590	571.11
24	295	310.69
25	940	929.31
26	475	455.56
27	210	195.14

Table 6.7 Modified worksheet for 3 factor 3 level full factorial experiment for the covariance fit

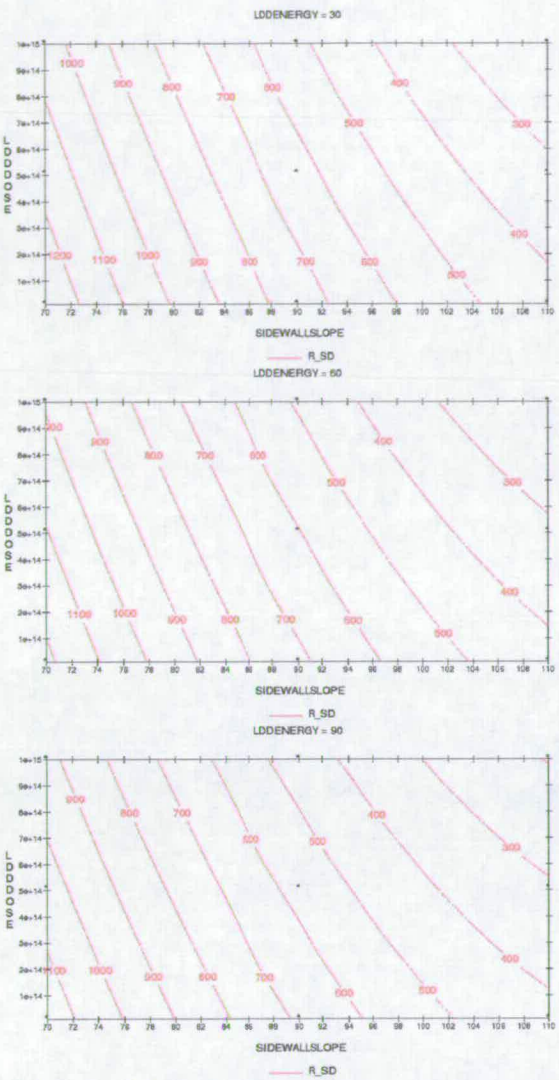


Figure 6.13 Contour plots of response surfaces for a second order polynomial fit to the data in table 6.5

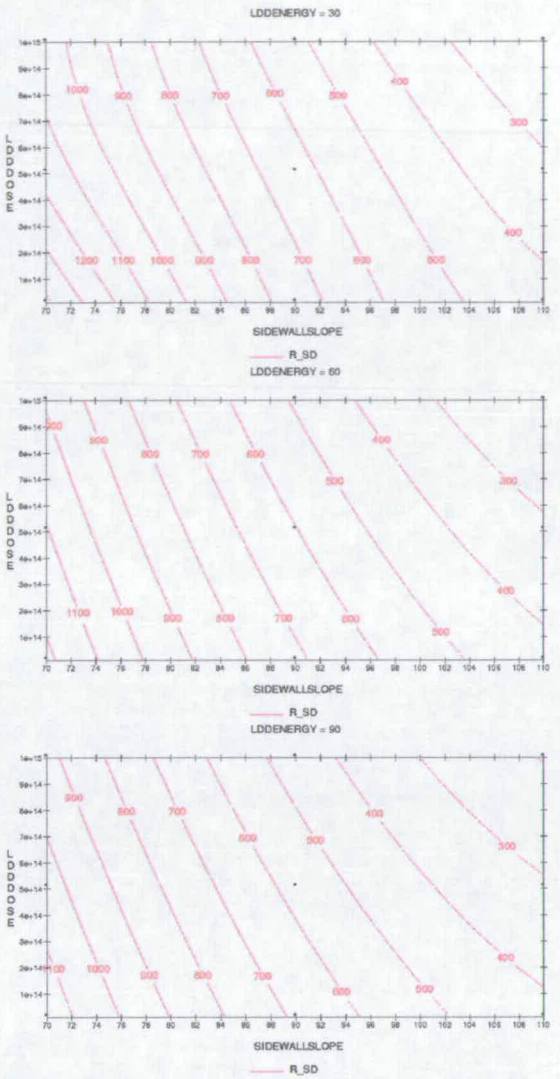


Figure 6.14 Contour plots of response surfaces for a covariance fit to the data in table 6.6

6.7 Summary

Chapter 6 has presented an optimisation strategy similar to that described in chapter 5. However, the JHVH-700 process had data points where the polynomial surface was unable to fit the data. Covariance fitting was applied to provide a gentle distortion of the trend so that the response goes through the data points. As the results from simulators do not contain error there was justification in this approach. The results of the covariance analysis were compared to those from the standard polynomial approach. This showed comparable results for simulated distributions in those regions where the polynomial fit was close to the data points. Where the difference was significant, the predicted distributions were very different. The covariance distribution had a tail in the distribution and was hence non-normal in nature. This shows that the process would be a challenge to control in a manufacturing environment, and that good process control would be required if the process were to yield consistently. Comparison with actual silicon results showed that simulation predicted the major distribution. Experimental data gathered showed two distributions, suggesting the influence of another factor not modelled by TCAD that affects the breakdown voltage. Further experimental and theoretical work would be required to investigate this further, and several approaches were proposed in this chapter. Also, it could be possible that running more silicon would result in the two distributions merging somewhat to create a more normal type of distribution. It has been observed that in integrated circuit manufacturing [20], the data comes in "bunches" or "packets" [21]. In this example, the silicon data maybe from a "packet" that was from the high end of the distribution. In other words, as the amount of silicon run was limited to 48 wafers, so the double distribution found may be due to an insufficient statistical sample.

In the course of this work, it has been demonstrated that the covariance model can be successfully fitted to a CCI design with five control factors

Covariance was applied to experimental data. In the example used, one point that had a large residual error and was not considered to be an outlier. Covariance analysis demonstrated that a good fit could be obtained without compromising the trend in the data and also without having to use transformations.

References

1. E.M.S. Narayanan, G. Amaratunga, W.I. Milne, "A study of the RESURF principle for thin epitaxial layer high voltage integrated circuits" Proc International symposium on power semiconductor devices, pp 172-175, 1992.
2. M. Redford, M. Fallon, Z.A. Shafi, N.S. Rankin, J. McGinty, A.J. Walton, "High Voltage RESURF DMOS Process Development Using DFM Techniques A Case Study", Proc 3rd International Workshop on Statistical Metrology, pp 78-81, June 7, 1998.
3. M.I. Newsam, M. Fallon, A.J. Walton, "A Total TCAD Environment for DFM", IEE Colloquium on Improving the Efficiency of IC Manufacturing Technology, University of Edinburgh, No. 1995/153, pp1/1-1-4, 12th April 1995.
4. T.G. Waring, A.J. Walton, D. Sprevak, R.S. Ferguson. "Application of Covariance Structures to Improve the Fit of Response Surfaces to Simulation Data", IEEE Trans Semiconductor Manufacturing, pp 366-374, 1999.
5. T.G. Waring, A.J. Walton, D. Sprevak, R.S. Ferguson. "The Sensitivity Of The Parameters Of Covariance Based Response Surfaces", Proc 3rd International Workshop on Statistical Metrology 1998, pp 50-53, June 7, 1998.
6. R.H. Myers, D.C. Montgomery, "Response Surface Methodology", Wiley International, 1995.
7. R.H. Myers, A.I. Khuri, W.H. Carter Jr.,"Response Surface Methodology: 1966-1988", Technometrics, vol 31, pp 137-157, May 1989.
8. G. Matheron, "Principles of Geostatistics", Economic Geology, vol 58, pp 1246-1266, Dec 1963.

9. M. Redford, A.J. Walton, D. Sprevak, R.S. Ferguson. "Application of Covariance Based Models to Fit Response Surfaces to Experimental Data", Proc 4th International Workshop on Statistical Metrology 1999, pp 42-45, June 12.

10. RS/1 User's Manual, Domain Software Solutions.

11. M.I. Newsam, "Calphurnia V2.0 Manual", Dept. of Electrical Engineering, University Of Edinburgh, Scotland.

12. M. Redford, M. Fallon, J. McGinty, A. J. Walton, M.I. Newsam, G. Michalowicz, C. McGregor, "Identifying And Quantifying Variability Within a BiCMOS Process : A Case Study", IEE Colloquium on Improving the Efficiency of IC Manufacturing Technology, No 1995/153, pp10/1-10/4, University of Edinburgh, April 12th, 1995.

13. J. McGinty, G. Long, J.K. Kibarian, M. Redford, "Implementation of Statistical TCAD in a Manufacturing Environment", National Semiconductor Technology & Innovation Conference, pp 230-233, 1996.

14. D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis of Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", International Symposium on Semiconductor Manufacturing, pp 85-91, Austin, Texas, August 1995.

15. RS/1 Manual, Domain Software Solutions, Chelmsford, MA, USA.

16. M. Redford, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, J.P. Elliott, G.A. Allan, A.J. Walton, "DFM Strategies In Technology Development", National Semiconductor Technology & Innovation Conference, pp 283-284, 1995.

17. A.J. Walton, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, G.A. Allan, J.P. Elliott; "A Total TCAD Strategy for DFM in IC Technology Development", IEE Proc. Science, Measurement and Technology, 144, No 2, pp 63-68, March 1997.

18. G.J. Gaston and A.J. Walton, "The integration of simulation and response surface methodology for the optimisation of IC processes", IEEE Trans Semiconductor Manufacturing, vol. 7, no. 1, pp. 22-33, Feb 1994.

19. D.C. Montgomery, "Design and Analysis of Experiments", John Wiley, 1991.

20. Statistical Process Control Seminar, Chartered Semiconductor Manufacturing, Singapore, October 5-8th, 1999.

21. Private communication, Dr Jake Sredni, Resource Optimization Inc.

Chapter 7: Conclusions and Further Work

7.1 Conclusions

The optimisation strategy developed in the thesis has been successfully demonstrated in an industrial environment with process conditions determined using a combination of Technology CAD (TCAD) and Design Of Experiments (DOE). This strategy was successfully applied to three different process technologies where they were optimised for breakdown voltage. The breakdown voltage is, by its nature, difficult to model with response surfaces, and hence provided a stringent test of the optimisation methodology. The manufacturability of the flows were assessed using a Monte Carlo approach by first determining the variability of the process and then using the response surfaces obtained from the DOE to predict the breakdown distributions. In the optimisation of the 700 volt process it was demonstrated that the covariance model could be successfully fitted to a Central Cubic Inscribed (CCI) design with five control factors. The covariance method was also successfully adapted to fit experimental data, where a good fit was obtained even when one data point possessed a large residual error. The predicted distributions showed reasonable agreement with silicon results and provided a realistic assessment of the manufacturability of each technology before any volume manufacturing was considered.

In a manufacturing environment it is often not possible to perform rigorous experiments because of constraints, such as time, silicon availability and equipment access. However, with careful experimental and logistical planning the methodology described in this thesis

has been successfully applied to several projects resulting in each case in a significant reduction in development cycle time.

A 22-mask complementary poly-silicon bipolar technology was, based on conventional methods, scheduled to take up to 18 months to transfer from Santa Clara, California to the 6" line in Greenock. Application of the strategy developed in this thesis resulted in first silicon was completed in 12 weeks and yielded 90% [1-3]. Altogether the first three lots could be used as qualification material, and the cycle time required to reach this point was reduced from 18 months to 18 weeks. This work involved extensive pre-work and planning to make this possible. A business opportunity was subsequently identified for this technology where if particular breakdown voltages were increased then a $150M could be quickly realised. However, this had to be achieved whilst ensuring that the existing set of active and passive devices were unchanged. An adaptation of this optimisation strategy [4] that included applying DOE techniques to layout optimisation on mask sets resulted in first silicon achieving the desired goal again resulting in development cycle time being drastically reduced.

The strategy was adapted again for another technology where a schottky device suite was integrated into an analog bipolar flow [5]. Again the desired electrical characteristics were realised on the first iteration of silicon thus providing a competitive advantage to the product line concerned.

In summary, this optimisation strategy became a key component of the process technology development work at National Semiconductor in Greenock.

7.2 Future Work

In the course of this work, several areas of work have been identified for further investigation.

The characterisation of 10-15Ω-cm epitaxy used in the combined epitaxy and implant process should be completed. The manufacturing engineers' assessment of its reproducibility needs to be determined, if only to determine its variability in a manufacturing environment. Once this characterisation is complete, an assessment of the epitaxial process variation can then be made and its impact studied if it is found to be different from that used to create the simulated distributions.

The differences between the experimental and simulated distributions for JHVH-700 indicated that there may be other influential factors, not modelled by TCAD that effect breakdown. One of these, which can be modelled using TCAD, is the effect of the field plates. The simulated DOE could be repeated with the field plate overlaps as an added control factor, and the resultant simulated distributions compared to silicon. The silicon data could also include more runs to increase the size of the data set.

As discussed in chapter 3, careful design of the field plate results in a higher breakdown voltage, it would be worthwhile investigating the use of a nitride field plate as its permittivity is higher than oxide. In this event, it may be possible to further enhance the breakdown voltage using the same nitride thickness.

A conclusion to the process/packaging interaction should be sought. That is, a complete oxynitride passivation evaluation and characterisation should be carried out by performing and analysing the High Temperature Reverse Bias (HTRB) test through 1000 hours. This should be carried out using a DOE on the oxynitride and packaging processes to determine optimum fabrication and packaging processing conditions to optimise the breakdown conditions through the HTRB test.

The behaviour and control in manufacturing of the on-resistance for the RESURF device should also be characterised. As part of this, a combined TCAD & DOE approach to predict the on-resistance distributions should be performed. This would provide another practical application for the methodology. The breakdown voltage and on-resistance are strongly correlated to on another and so have similar control factor sets. If there is any changes as a result of the epitaxial characterisation described above then it should also be applied to the on-resistance distribution simulations.

As the high voltage RESURF device is to be integrated into a smart power process, it would be prudent to study the impact and manufacturability of integrating the RESURF device into the smart power flow. Such a study would include analysis of all key device types.

In conclusion, the work described in this thesis has demonstrated a robust methodology for process characterisation prior to experimental evaluation, with a profound impact on development cycle time. The potential applications of this technique are wide, with clear

benefits already being demonstrated on recent projects that have adapted this methodology.

References

1. J.I. Doohan, M. Fallon, I. Thomson, R. Boyle, J. Norval, "Dry etch uniformity and optimisation using graphical and DOE techniques", Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 165-172, May 18-21 1999.
2. I. Black, "ASM stepper alignment through thick epitaxial silicon films" in proc Microelectronic Manufacturing Technologies, Vol 3741, pp 23-33, May 18-21 1999.
3. A. McGibbon, K.Yallup, J. Altrip, M. Harper, M. Redford. "A Fingerprint Methodology for Process Transfer and Dual Sourcing", National Semiconductor Technology & Innovation Conference, pp 171-174, July 1997.
4. Z.A. Shafi, J. McGinty, M. Redford. "Rapid and Cost Effective Technology Development Using TCAD : A Case Study", Proc Microelectronic Manufacturing Technologies, Vol 3743, pp 234-240, May 18-21 1999.
5. A. McGibbon, Z.A.Shafi, N.Cameron, M.Redford, "DOE Analysis of Layout/Process Interactions in Bipolar Schottky Technology", Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 87-95, May 18-21 1999.

Appendices

Appendix 1: SUPREM-3 - Epitaxy process

```
$ JHVH discrete LDMOS process flow
$ adapted from 300B LMDMOS deck by Jim McGinty

$ epitaxy process
$ target substrate concentration for JHVH450: 5.0e14 (26 ohm-cm)
$ target epi        concentration for JHVH450: 3.5e14 (10-15 ohm-cm)
$ target substrate concentration for JHVH700: 1.0e14 (130 ohm-cm)
$ target substrate concentration for JHVH700: 3.5e14 (10-15 ohm-cm)


title discrete LDMOS process flow


assign name=nsub n.val=5.0e14
assign name=nepi n.val=3.5e14
$assign name=dose n.val=1.2e12
$assign name=energy n.val=120
option ^terminal

$defining the masks

$                          x      y      x      y

mask name=well      rect.1=(2.0,   1.0    3.0,   2.0)
+                   rect.2=(14,    1.0    15,    2.0)
mask name=shallp    rect.1=(4.0,   1.0    5.0,   2.0)
mask name=body      rect.1=(6.0,   1.0    7.0,   2.0)
mask name=nplus     rect.1=(8.0,   1.0    9.0,   2.0)
mask name=nprotect  rect.1=(8.0,   1.0    9.0,   2.0) clear
mask name=isoup     rect.1=(10,    1.0    11,    2.0)
mask name=active    rect.1=(0.0,   0.0    12,    3.0

initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100

$ plot of simulation layout

plot mask name=well bottom=0 left=0 right=20 device=x
+title="SUPREM-3 JHVH layout"
plot mask name=shallp add color=3 line.type=2
plot mask name=nplus  add color=4 line.type=2
plot mask name=active add color=5 line.type=2
plot mask name=body  add color=6 line.type=2
plot mask name=isoup  add color=7 line.type=2
```

```
$masks details
$mask                        name     x.loc        x.char [characteristic length]
$deep p+ into thin ox         reg.1     2.5          2.57
$shallp+                      reg.2     4.5          0.12
$pbody                        reg.3     6.5          1.12
$shaln+                       reg.4     8.5          0.18
$isoup under thinox           reg.5     10.5         2.66
$drift region under Fox  reg.6       12.5
$deepp+ under Fox        reg.7       14.5         2.67


initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100

loop steps=7

assign name=zloc n.value=1.5
assign name=x.loc n.value=(2.5,4.5,6.5,8.5,10.5,12.5,14.5)
assign name=reglabel c.value=reg.1 delta=1
assign name=tstring c1="deepp+(thinox)" c2="shallp+" c3="pbody" c4="shalln+"
+c5="isoup" c6="drift region (Fox)" deepP+ (Fox)"

initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100
+ x.location=@xloc z.location=@zloc

option ^terminal

$$$$$
$ initial oxidation

assign name=gas n.value=0.82

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=205 press=@gas f.n2=.36 f.o2=4.6 f.hcl=.263 f.h2=7
diffusion temp=950 time=18 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4

$$$$$
$isoup implant

mask apply name=isoup
```

```
etch oxide all
implant boron dose=1e14 energy=150 pearson
mask remove

$$$$$



$$$$$ isoup anneal

assign name=gas n.value=0.82

deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=67 press=@gas f.n2=7.8 f.o2=0.6 t.final=1200
diffusion temp=1200 time=50 press=@gas  f.n2=7.8
$dry then wet
diffusion temp=1200 time=42 press=@gas f.o2=3.9
diffusion temp=1200 time=42 press=@gas f.o2=5 f.h2=7 t.final=800
diffusion temp=1200 time=159 press=@gas f.n2=7.8
diffusion temp=950 time=22 press=@gas f.n2=7.8
diffusion temp=800 time=20 press=@gas f.n2=7.8
etch oxide all
etch silicon thick=0.05

$$$$$ epitaxial growth

epitaxy thickness=12 temp=1130 time=12 arsenic=@nepi dx=0.001 xdx=0 spaces=200
grid layer.1 spaces=200 dx=0.050 xdx=0.15 x.min=12
grid layer.1 spaces=350 dx=0.005 xdx=0.10 x.min=12


$$$$$plot the post epi profile

plot active net tic.grid device=x timestamp title="post epi:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ ndrift implant oxide

assign name=gas n.value=0.82
```

```
deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.6 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6 f.hcl=0.2632
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.2632
diffusion temp=950 time=21.5 press=@gas f.h2=7.8
diffusion temp=1200 time=159 press=@gas f.n2=7.8
diffusion temp=950 time=22 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

$$$$$

$mask n+ regions and implant drift region
$mask apply name=nprotect
$implant phosphor dose=@dose energy=@energy boltzman
$mask remove

$$$$$plot the post drift implant profile

$plot active net tic.grid device=x timestamp title="post drift implant:"@tstring
$plot active boron add color=2 line.type=2
$plot active arsenic add color=3 line.type=3
$plot active phosph add color=4 line.type=4
$plot active antimony add color=5 line.type=5

$$$$$ epi reox

assign name=gas n.value=0.83

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263 f.n2=0.36
diffusion temp=950 time=18 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

print layers

$$$$$plot the post drift implant profile
```

```
plot active net tic.grid device=x timestamp title="post epi-reox:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$


$$$$$ Well region

mask apply name=well
etch oxide
mask remove

grid layer.1 xdx=0.0001 x.max=3 min.dx=0.001

$$$$$ implant oxide


assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.6 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6 f.hcl=0.2632
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263
diffusion temp=950 time=21.5 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

grid layer.1 xdx=0.0001 x.max=3 min.dx=0.001


$$$$$ deep + implant

implant boron dose=1e15 energy=30 boltzman

$$$$$ 40s HF clean to remove 50A
etch oxide thick=0.005

$$$$$ deep p+ drive

assign name=gas n.value=0.924
deposit oxide thick=0.002
```

```
diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=44 press=@gas f.n2=7.8 f.o2=0.4 t.final=1200
diffusion temp=1200 time=23 press=@gas  f.n2=7.8
diffusion temp=1200 time=160 press=@gas f.n2=7.8
diffusion temp=1200 time=113 press=@gas f.o2=5 f.h2=7
diffusion temp=1200 time=50 press=@gas f.n2=7.8
diffusion temp=1200 time=84 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

etch oxide all

deposit oxide thick=0.0025 spaces=5

$$$$$plot the post deep p+ drive/well profile

plot active net tic.grid device=x timestamp title="post pwell drive:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5


$$$$$ pad oxidation

assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=10 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=10 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263 f.n2=0.36
diffusion temp=950 time=21 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8


deposition nitride thickness=0.15 spaces=100

$$$$$ mask/etch active area
```

```
mask apply name=active
etch nitride all
mask remove
etch oxide thick=0.059



$$$$$ field oxide


assign name=gas n.value=0.89

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=28 press=@gas f.n2=7.8 f.o2=0.4 t.final=1025
diffusion temp=1025 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1050
diffusion temp=1050 time=160 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=1050 time=20 press=@gas  f.o2=4.6 t.final=950 t.final=1000
diffusion temp=1000 time=525 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.175 f.n2=0.24
diffusion temp=1000 time=15 press=@gas f.n2=7.8
diffusion temp=950 time=17 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

$$$$$ etch ONO

etch oxide thick=0.04
etch nitride
etch oxide thick=0.11

$$$$$plot the post ONO etch profile

plot active net tic.grid device=x timestamp title="post ONO etch:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ sacrificial oxidation

deposit oxide thick=0.002

assign name=gas n.value=0.969

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
```

```
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=7.2
diffusion temp=950 time=10 press=@gas f.o2=7.2 f.h2=6.6
diffusion temp=950 time=10 press=@gas f.o2=7.2 f.h2=6.6 f.hcl=0.329 f.n2=0.45
diffusion temp=950 time=21 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

etch oxide thick=0.1400

$$$$$ gate oxide

assign name=gas n.value=0.97

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=7.2
diffusion temp=950 time=7 press=@gas f.o2=7.2 f.h2=6.6
diffusion temp=950 time=7 press=@gas f.o2=7.2 f.h2=6.6 f.hcl=0.329 f.n2=0.45
diffusion temp=950 time=23 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8


print layers

deposition polysilicon temperature=615 thickness=0.5500
+ xdx=0.00 dx=0.0100 spaces=25 pressure=4.7e-4

$$$$$ poly doping

deposit oxide thickness=0.002

diffusion temp=800 time=22 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11 f.o2=0.4 t.final=940
diffusion temp=940 time=4 press=@gas f.n2=11 f.o2=0.4 t.final=949
diffusion temp=949 time=6 press=@gas f.n2=11 f.o2=0.4 t.final=950
diffusion temp=950 time=0.5 press=@gas  f.n2=11 f.o2=1
diffusion temp=950 time=18 press=@gas f.n2=11 f.o2=1 phosphor=1e21
```

```
diffusion temp=950 time=26 press=@gas f.n2=11
diffusion temp=950 time=75 press=@gas f.n2=11 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=11

etch oxide thickness=0.1600
etch polysilicon
etch oxide thickness=0.0075

$$$$$ pbody mask

mask apply name=body
implant boron dose=3.4e13 energy=80 pearson
mask remove

$$$$$plot the post pbody implant profile

plot active net tic.grid device=x timestamp title="post pbody implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5


$$$$$ pbody drive

$$$$$ deep p+ drive

assign name=gas n.value=1.15
deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=35 press=@gas f.n2=7.8 f.o2=0.4 t.final=1075
diffusion temp=1075 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1100
diffusion temp=1100 time=10 press=@gas  f.o2=3.9 f.hcl=0.073
diffusion temp=1100 time=291 press=@gas f.n2=7.8
diffusion temp=1100 time=30.5 press=@gas f.o2=3.5
diffusion temp=1100 time=30.5 press=@gas  f.o2=3.9 f.hcl=0.073
diffusion temp=1100 time=50 press=@gas f.n2=7.8  t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8


$$$$$plot the post pbody drive implant profile

plot active net tic.grid device=x timestamp title="post pbody implant:"@tstring
```

```
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5



$$$$$ mask/etch n+ source
mask apply name=nplus
etch oxide
mask remove

$$$$$ n+ implant

implant arsenic dose=5.0e15 energy=50 boltzman

etch oxide thickness=0.1100

$$$$$plot the post n+ source implant profile

plot active net tic.grid device=x timestamp title="post n+ source implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ poly reox

assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=28 press=@gas f.n2=7.8 f.o2=0.4 t.final=1025
diffusion temp=1025 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1050
diffusion temp=1050 time=45 press=@gas f.n2=7.8 f.o2=0.4 t.final=
diffusion temp=1050 time=10 press=@gas f.o2=3.9
diffusion temp=1050 time=33 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

$$$$$ shallow p+ implant

mask apply name=shallp
implant boron dose=8.25e14 energy=50 boltzman
```

```
mask remove

$$$$$ undoped LTO

deposit oxide thickness=0.8000



$$$$$ sinker dope

diffusion temp=800 time=22 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11 f.o2=0.4 t.final=940
diffusion temp=940 time=4 press=@gas f.n2=11 f.o2=0.4 t.final=949
diffusion temp=949 time=6 press=@gas f.n2=11 f.o2=0.4 t.final=950
diffusion temp=950 time=0.5 press=@gas  f.n2=11 f.o2=1
diffusion temp=950 time=18 press=@gas f.n2=11 f.o2=1 phosphor=1e21
diffusion temp=950 time=26 press=@gas f.n2=11
diffusion temp=950 time=75 press=@gas f.n2=11 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=11

$$$$$ deposit 900A to make up for short fall during the predep process
deposit oxide thickness=0.1130 spaces=10

$$$$$ deglaze
etch oxide thickness=0.1600

print layer
etch oxide all

$$$$$ savefile for medici input

savefile device file=@reglabel all

electrical
end

$$$$$plot the post n+ source implant profile

initialize structure=@reglabel

plot active net tic.grid device=x timestamp title="JHVH-450 RESURF
+LDMOS:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
```

```
plot active antimony add color=5 line.type=5


$$$$$ Simulated SRP

assign name=loops n.val=200
assign name=extsav c.val=mark
extract thickness layer=1 name=tsi
assign name=depth n.val=0
extract file=@extsav
loop steps=int(@loops-1)
option ^terminal
electrical
end electrical
extract electron y.extract x=0.0 name=econc
extract hole y.extract x=0.0 name=hconc
assign name=net n.val=abs(@econc-@hconc)
extract variable=(@depth @econc @hconc @net)
assign name=depth n.val(@tsi/@loops):(@tsi/@loops)
etch silicon thick=(@tsi/@loops)
l.end
extract close
loadfile file=@reglabel all
plot add data=mark x.col=1 y.col=4 color=1 symbol=4 ^curve

l.end

stop
```

Appendix 2: SUPREM-3 Combined epitaxy & implant process

```
$ JHVH discrete LDMOS process flow
$ adapted from 300B LMDMOS deck by Jim McGinty
$ combined epitaxy & implant process

$ epitaxy process
$ target substrate concentration for JHVH450: 5.0e14 (26 ohm-cm)
$ target epi        concentration for JHVH450: 3.5e14 (10-15 ohm-cm)
$ target substrate concentration for JHVH700: 1.0e14 (130 ohm-cm)
$ target substrate concentration for JHVH700: 3.5e14 (10-15 ohm-cm)

$ drift implant 1.2e122 for JHVH-450
$ drift implant 5e11 for JHVH-700

title discrete LDMOS process flow


assign name=nsub n.val=4.0e14
assign name=nepi n.val=5.0e14
assign name=dose n.val=1.2e12
assign name=energy n.val=120
option ^terminal

$defining the masks

$                          x      y      x      y

mask name=well      rect.1=(2.0,   1.0    3.0,   2.0)
+                   rect.2=(14,    1.0    15,    2.0)
mask name=shallp    rect.1=(4.0,   1.0    5.0,   2.0)
mask name=body      rect.1=(6.0,   1.0    7.0,   2.0)
mask name=nplus     rect.1=(8.0,   1.0    9.0,   2.0)
mask name=nprotect  rect.1=(8.0,   1.0    9.0,   2.0) clear
mask name=isoup     rect.1=(10,    1.0    11,    2.0)
mask name=active    rect.1=(0.0,   0.0    12,    3.0

initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100

$ plot of simulation layout

plot mask name=well bottom=0 left=0 right=20 device=x
+title="SUPREM-3 JHVH layout"
plot mask name=shallp add color=3 line.type=2
plot mask name=nplus  add color=4 line.type=2
```

```
plot mask name=active add color=5 line.type=2
plot mask name=body  add color=6 line.type=2
plot mask name=isoup  add color=7 line.type=2

$masks details
$mask                        name     x.loc        x.char [characteristic length]
$deep p+ into thin ox        reg.1     2.5          2.57
$shallp+                     reg.2     4.5          0.12
$pbody                       reg.3     6.5          1.12
$shaln+                      reg.4     8.5          0.18
$isoup under thinox          reg.5     10.5         2.66
$drift region under Fox  reg.6       12.5
$deepp+ under Fox        reg.7       14.5         2.67


initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100

loop steps=7

assign name=zloc n.value=1.5
assign name=x.loc n.value=(2.5,4.5,6.5,8.5,10.5,12.5,14.5)
assign name=reglabel c.value=reg.1 delta=1
assign name=tstring c1="deepp+(thinox)" c2="shallp+" c3="pbody" c4="shalln+"
+c5="isoup" c6="drift region (Fox)" deepP+ (Fox)"

initialize <100> silicon boron=@nsub thickness=12 dx=.010 xdx=0 spaces=100
+ x.location=@xloc z.location=@zloc

option ^terminal

$$$$$
$ initial oxidation

assign name=gas n.value=0.82

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=205 press=@gas f.n2=.36 f.o2=4.6 f.hcl=.263 f.h2=7
diffusion temp=950 time=18 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4

$$$$$
```

```
$isoup implant

mask apply name=isoup
etch oxide all
implant boron dose=1e14 energy=150 pearson
mask remove

$$$$$



$$$$$ isoup anneal

assign name=gas n.value=0.82

deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=67 press=@gas f.n2=7.8 f.o2=0.6 t.final=1200
diffusion temp=1200 time=50 press=@gas  f.n2=7.8
$dry then wet
diffusion temp=1200 time=42 press=@gas f.o2=3.9
diffusion temp=1200 time=42 press=@gas f.o2=5 f.h2=7 t.final=800
diffusion temp=1200 time=159 press=@gas f.n2=7.8
diffusion temp=950 time=22 press=@gas f.n2=7.8
diffusion temp=800 time=20 press=@gas f.n2=7.8
etch oxide all
etch silicon thick=0.05

$$$$$ epitaxial growth

epitaxy thickness=12 temp=1130 time=12 arsenic=@nepi dx=0.001 xdx=0 spaces=200
grid layer.1 spaces=200 dx=0.050 xdx=0.15 x.min=12
grid layer.1 spaces=350 dx=0.005 xdx=0.10 x.min=12


$$$$$plot the post epi profile

plot active net tic.grid device=x timestamp title="post epi:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5
```

```
$$$$$ ndrift implant oxide

assign name=gas n.value=0.82

deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.6 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6 f.hcl=0.2632
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.2632
diffusion temp=950 time=21.5 press=@gas f.h2=7.8
diffusion temp=1200 time=159 press=@gas f.n2=7.8
diffusion temp=950 time=22 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

$$$$$

$mask n+ regions and implant drift region
mask apply name=nprotect
implant phosphor dose=@dose energy=@energy boltzman
mask remove

$$$$$plot the post drift implant profile

plot active net tic.grid device=x timestamp title="post drift implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ epi reox

assign name=gas n.value=0.83

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=205 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263 f.n2=0.36
diffusion temp=950 time=18 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

print layers
```

```
$$$$$plot the post drift implant profile

plot active net tic.grid device=x timestamp title="post epi-reox:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$


$$$$$ Well region

mask apply name=well
etch oxide
mask remove

grid layer.1 xdx=0.0001 x.max=3 min.dx=0.001

$$$$$ implant oxide


assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.6
diffusion temp=800 time=25 press=@gas f.n2=7.8 f.o2=0.6 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6 f.hcl=0.2632
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=5.75 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263
diffusion temp=950 time=21.5 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

grid layer.1 xdx=0.0001 x.max=3 min.dx=0.001


$$$$$ deep + implant

implant boron dose=1e15 energy=30 boltzman

$$$$$ 40s HF clean to remove 50A
etch oxide thick=0.005

$$$$$ deep p+ drive
```

```
assign name=gas n.value=0.924
deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=44 press=@gas f.n2=7.8 f.o2=0.4 t.final=1200
diffusion temp=1200 time=23 press=@gas  f.n2=7.8
diffusion temp=1200 time=160 press=@gas f.n2=7.8
diffusion temp=1200 time=113 press=@gas f.o2=5 f.h2=7
diffusion temp=1200 time=50 press=@gas f.n2=7.8
diffusion temp=1200 time=84 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

etch oxide all

deposit oxide thick=0.0025 spaces=5

$$$$$plot the post deep p+ drive/well profile

plot active net tic.grid device=x timestamp title="post pwell drive:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5


$$$$$ pad oxidation

assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=4.6
diffusion temp=950 time=10 press=@gas f.o2=4.6 f.h2=7
diffusion temp=950 time=10 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.263 f.n2=0.36
diffusion temp=950 time=21 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8


deposition nitride thickness=0.15 spaces=100
```

```
$$$$$ mask/etch active area

mask apply name=active
etch nitride all
mask remove
etch oxide thick=0.059



$$$$$ field oxide


assign name=gas n.value=0.89

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=28 press=@gas f.n2=7.8 f.o2=0.4 t.final=1025
diffusion temp=1025 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1050
diffusion temp=1050 time=160 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=1050 time=20 press=@gas  f.o2=4.6 t.final=950 t.final=1000
diffusion temp=1000 time=525 press=@gas f.o2=4.6 f.h2=7 f.hcl=0.175 f.n2=0.24
diffusion temp=1000 time=15 press=@gas f.n2=7.8
diffusion temp=950 time=17 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

$$$$$ etch ONO

etch oxide thick=0.04
etch nitride
etch oxide thick=0.11

$$$$$plot the post ONO etch profile

plot active net tic.grid device=x timestamp title="post ONO etch:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ sacrificial oxidation

deposit oxide thick=0.002
```

```
assign name=gas n.value=0.969

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=7.2
diffusion temp=950 time=10 press=@gas f.o2=7.2 f.h2=6.6
diffusion temp=950 time=10 press=@gas f.o2=7.2 f.h2=6.6 f.hcl=0.329 f.n2=0.45
diffusion temp=950 time=21 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8

etch oxide thick=0.1400

$$$$$ gate oxide

assign name=gas n.value=0.97

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=20 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=15 press=@gas f.n2=7.8 f.o2=0.4 t.final=940
diffusion temp=940 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=949
diffusion temp=949 time=5 press=@gas f.n2=7.8 f.o2=0.4 t.final=950
diffusion temp=950 time=15 press=@gas  f.o2=7.2
diffusion temp=950 time=7 press=@gas f.o2=7.2 f.h2=6.6
diffusion temp=950 time=7 press=@gas f.o2=7.2 f.h2=6.6 f.hcl=0.329 f.n2=0.45
diffusion temp=950 time=23 press=@gas f.n2=7.8
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=7.8


print layers

deposition polysilicon temperature=615 thickness=0.5500
+ xdx=0.00 dx=0.0100 spaces=25 pressure=4.7e-4

$$$$$ poly doping

deposit oxide thickness=0.002

diffusion temp=800 time=22 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11 f.o2=0.4 t.final=940
diffusion temp=940 time=4 press=@gas f.n2=11 f.o2=0.4 t.final=949
```

```
diffusion temp=949 time=6 press=@gas f.n2=11 f.o2=0.4 t.final=950
diffusion temp=950 time=0.5 press=@gas  f.n2=11 f.o2=1
diffusion temp=950 time=18 press=@gas f.n2=11 f.o2=1 phosphor=1e21
diffusion temp=950 time=26 press=@gas f.n2=11
diffusion temp=950 time=75 press=@gas f.n2=11 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=11

etch oxide thickness=0.1600
etch polysilicon
etch oxide thickness=0.0075

$$$$$ pbody mask

mask apply name=body
implant boron dose=3.4e13 energy=80 pearson
mask remove

$$$$$plot the post pbody implant profile

plot active net tic.grid device=x timestamp title="post pbody implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5


$$$$$ pbody drive

$$$$$ deep p+ drive

assign name=gas n.value=1.15
deposit oxide thick=0.002

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=35 press=@gas f.n2=7.8 f.o2=0.4 t.final=1075
diffusion temp=1075 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1100
diffusion temp=1100 time=10 press=@gas  f.o2=3.9 f.hcl=0.073
diffusion temp=1100 time=291 press=@gas f.n2=7.8
diffusion temp=1100 time=30.5 press=@gas f.o2=3.5
diffusion temp=1100 time=30.5 press=@gas  f.o2=3.9 f.hcl=0.073
diffusion temp=1100 time=50 press=@gas f.n2=7.8  t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8
```

```
$$$$$plot the post pbody drive implant profile

plot active net tic.grid device=x timestamp title="post pbody implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5




$$$$$ mask/etch n+ source
mask apply name=nplus
etch oxide
mask remove

$$$$$ n+ implant

implant arsenic dose=5.0e15 energy=50 boltzman

etch oxide thickness=0.1100

$$$$$plot the post n+ source implant profile

plot active net tic.grid device=x timestamp title="post n+ source implant:"@tstring
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5

$$$$$ poly reox

assign name=gas n.value=1.15

diffusion temp=800 time=22 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=13 press=@gas f.n2=7.8 f.o2=0.4
diffusion temp=800 time=28 press=@gas f.n2=7.8 f.o2=0.4 t.final=1025
diffusion temp=1025 time=13 press=@gas f.n2=7.8 f.o2=0.4 t.final=1050
diffusion temp=1050 time=45 press=@gas f.n2=7.8 f.o2=0.4 t.final=
diffusion temp=1050 time=10 press=@gas f.o2=3.9
diffusion temp=1050 time=33 press=@gas f.n2=7.8 t.final=950
diffusion temp=950 time=75 press=@gas f.n2=7.8 t.final=800
diffusion temp=800 time=22 press=@gas f.n2=7.8

$$$$$ shallow p+ implant
```

```
mask apply name=shallp
implant boron dose=8.25e14 energy=50 boltzman
mask remove

$$$$$ undoped LTO

deposit oxide thickness=0.8000



$$$$$ sinker dope

diffusion temp=800 time=22 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11
diffusion temp=800 time=15 press=@gas f.n2=11 f.o2=0.4 t.final=940
diffusion temp=940 time=4 press=@gas f.n2=11 f.o2=0.4 t.final=949
diffusion temp=949 time=6 press=@gas f.n2=11 f.o2=0.4 t.final=950
diffusion temp=950 time=0.5 press=@gas  f.n2=11 f.o2=1
diffusion temp=950 time=18 press=@gas f.n2=11 f.o2=1 phosphor=1e21
diffusion temp=950 time=26 press=@gas f.n2=11
diffusion temp=950 time=75 press=@gas f.n2=11 t.final=800
diffusion temp=950 time=22 press=@gas f.n2=11

$$$$$ deposit 900A to make up for short fall during the predep process
deposit oxide thickness=0.1130 spaces=10

$$$$$ deglaze
etch oxide thickness=0.1600

print layer
etch oxide all

$$$$$ savefile for medici input

savefile device file=@reglabel all

electrical
end

$$$$$plot the post n+ source implant profile

initialize structure=@reglabel

plot active net tic.grid device=x timestamp title="JHVH-450 RESURF
+LDMOS:"@tstring
```

```
plot active boron add color=2 line.type=2
plot active arsenic add color=3 line.type=3
plot active phosph add color=4 line.type=4
plot active antimony add color=5 line.type=5


$$$$$ Simulated SRP

assign name=loops n.val=200
assign name=extsav c.val=mark
extract thickness layer=1 name=tsi
assign name=depth n.val=0
extract file=@extsav
loop steps=int(@loops-1)
option ^terminal
electrical
end electrical
extract electron y.extract x=0.0 name=econc
extract hole y.extract x=0.0 name=hconc
assign name=net n.val=abs(@econc-@hconc)
extract variable=(@depth @econc @hconc @net)
assign name=depth n.val(@tsi/@loops):(@tsi/@loops)
etch silicon thick=(@tsi/@loops)
l.end
extract close
loadfile file=@reglabel all
plot add data=mark x.col=1 y.col=4 color=1 symbol=4 ^curve

l.end

stop
```

Appendix 3: Medici Deck

```
$ LDMOS Medici deck
$ Adapted from original program from J.M. McGregor
$

$ JHVH-450
$ parallel region len=49, ldrift=30, dplate=5, splate
$ drain finger tip len=79, ldrift=60, dplate=15, splate=8
$ source finger tip len=79, ldrift=60, dplate=5, splate=15

$ JHVH-700
$ parallel region len=81, ldrift=65, dplate=10, splate=8
$ drain finger tip len=81, ldrift=130, dplate=30, splate=8
$ source finger tip len=81, ldrift=130, dplate=8, splate=30

Title discrete LDMOS – JHVH flow

$set device dimensions
Assign name=len      n.val=49
Assign name=depth    n.val=70
Assign name=wepi     n.val=12

$set concentrations
Assign name=nepi     n.val=3.5e14
Assign name=type     c.val=n-type
Assign name=nsub     n.val=5.00e14

$
Assign name=lshalp n.val=7
Assign name=ldeepp n.val=5
Assign name=lsrc     n.val=4
Assign name=lthinox n.val=3
Assign name=ldrift   n.val=30
Assign name=ldrn     n.val=5
Assign name=splate   n.val=8
Assign name=dplate   n.val=5

$set oxide thicknesses
Assign name=dthin    n.val=0.9
Assign name=dthick   n.val=1.8
Assign name=dg1      n.val=0.06
Assign name=dg2      n.val=1.0

Assign name=feature  n.val=@lsrc+@lthinox+@lshalp
```

```
mesh smooth=1
x.mesh width=@len-(@feature+@lthinox+@splate)-(@dplate+5.0) h1=1.0
x.mesh width=@feature+@lthinox+@splate h1=0.50
x.mesh width=@dplate+5.0 h1=1.0

y.mesh n=1 l=-(@dthick+0.2)
y.mesh n=2 l=-@dthick
y.mesh n=3 l=-@dg2
y.mesh n=4 l=-@dthin
y.mesh n=6 l=-@dg1
y.mesh n=7 l=-0.0
y.mesh n=8 l=0.05
y.mesh width=@wepi h1=0.1 h2=2.0
y.mesh width=@depth-@epi h1=2.0 h2=20

eliminate columns x.max=@feature+@lthinox+splate y.min=4
eliminate columns y.min=10
eliminate columns y.min=20
eliminate columns y.min=40

region name=silicon silicon iy.lo=7
region name=oxide iy.hi=7

$Define electrodes
electrode name=substrate bottom

$Gate
electrode name=gate x.min=@lshalp+@lsrc x.max=@feature iy.lo=6 iyhi=6
electrode name=gate x.min=@feature+2.0 +x.max=@feature+3.0 iy.lo=3 iy.hi=3.0

$Source
electrode name=source x.min=0.0 x.max=@lshalp-2.5 iy.lo=1 iy.hi=7
electrode name=source x.min=8-0.5 x.max=@lshalp+2.5 iy.lo=1 iy.hi=7
electrode name=source x.min=0.0 x.max=@feature+1.05 iy.lo=1 iy.hi=4
$source fieldplate
electrode name=source x.min=@feature x.max=@feature+@lthinox+@splate iy.lo=1 iy.hi=2

$drain
electrode name=drain x.min=@len-5.0 x.max=@len iy.lo=1 iy.hi=7
$drain fieldplate
electrode name=drain x.min=@len-5.0-@dplate-0.5 x.max=@len-5.0 iy.lo=1 iy.hi=2

$ impurity profiles
$ uniform over structure
profile p-type unifrm n.peak=@nsub y.peak=@wepi
profile @type uniform n.peak=@nepi y.peak=0 depth=@wepi
```

```
profile n-type y.peak=0.0 1d.proc file=reg.6

profile p-type x.peak=0.0 width=@ldeepp x.char=2.44 y.peak=0.0 1d.proc file=reg.1
profile p-type x.peak=0.0 width=@lshalp x.char=0.12 y.peak=0.0 1d.proc file=reg.2
profile p-type x.peak=0.0 width=@lshalp+@lsrc x.char=0.12 y.peak=0.0 1d.proc
+file=reg.3
profile n-type x.peak=@lshalp width=@lsrc x.char=0.19 y.peak=0.0 1d.proc file=reg.4
profile n-type x.peak=@len-@ldrn width=8.0 x.char=0.19 y.peak=0.0 1d.proc file=reg.4
profile n-type x.peak=0.0 width=@lshalp+@lsrc x.char=2.53 y.peak=0.0 1d.proc file=reg.5

interface region=(1,2) qf=5e10
save mesh outfile=mesh w.model


$ check the 1d doping profiles
plot.1d doping y.log x.start=3 x.end=3 y.start=0 y.end=30 points color=2
+ title="shallow p+ in deep p+ into iso-up profile"

plot.1d doping y.log x.start=25 x.end=25 y.start=0 y.end=22 points color=2
+ title="drift region into epilayer profile" bottom=1e14 top=1e21 right=23

plot.1d doping y.log x.start=8 x.end=8 y.start=0 y.end=17 points color=2
+ title="shallow n+ in p-body into iso-up profile" bottom=1e14 top=1e21

plot.1d doping y.log x.start=48 x.end=48 y.start=0 y.end=2 points color=2
+ title="shallow n+ in drift and epiprofile" bottom=1e14 top=1e21

plot.2d boundary scale grid l.elect=1 device=x
contour doping log color=2 min=14 max=21
contour doping log color=3 min=-21 max=-14

plot.2d boundary scale grid l.elect=1 device=x y.min=-7
contour doping log color=2 min=14 max=21
contour doping log color=3 min=-21 max=-14
label label="p-substrate" x=@len/2.3 y=@depth/2
label label="n-epi" x=@len/2.1 y=@wepi/1.3
label label="drift region" x=@len/1.9 y=@wepi/3.2
label label="s" x=@lshalp y=-5
label label="g" x=@lshalp+@lsrc+2 y=-5
label label="d" x=@len-2 y=-5

contact name=gate n.polysi

$save structure before analysis
savefile out.file=checkout.tif tif
```

```
$model definition
model btbt phumob fldmob srfmob2 bgn auger impact.i

symbol gummel carriers=0
method iccg damped
solve init



model btbt phumob fldmob srfmob2 bgn auger impact.i

symbol newton carriers=2

$ usingcontinuation method
solve elec=drain continue c.vstep=0.5 c.toler=0.1 c.vmax=520

$save after bv
savefile out.file=checkbv.tif tif all
```

Appendix 4: Papers Associated with this Thesis

1. M. Redford, M. Fallon, A.J. Walton, Z.A. Shafi, J. McGinty, I. Underwood, "High Voltage RESURF Process Development using Covariance Based Response Surfaces", Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 76-86, May 18-21 1999.
2. A. McGibbon, Z.A. Shafi, N. Cameron, M. Redford, "DOE Analysis of Layout/Process Interactions in Bipolar Schottky Technology", Proc Microelectronic Manufacturing Technologies, Vol 3742, pp 87-95, May 18-21 1999.
3. A. McGibbon, R. Boyle, M. Redford, "Role of STEM as a High Resolution Failure Analysis Tool for Bipolar Technologies", Proc Microelectronic Manufacturing Technologies Vol 3743, pp 219-226, May 18-21, 1999.
4. Z.A. Shafi, J. McGinty, M. Redford, "Rapid and Cost Effective Technology Development Using TCAD : A Case Study", Proc Microelectronic Manufacturing Technologies, Vol 3743, pp 234-240, May 18-21, 1999.
5. K. Findlater, M. Fallon, M. Redford, "Investigation and Application of the Buried Layer Van Der Pauw and Bar Resistors", Proc Microelectronic Manufacturing Technologies, Vol 3743, pp 249-255, May 18-21, 1999.
6. M. Redford, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, J.P. Elliott, G.A. Allan, A.J. Walton, "DFM Strategies In Technology Development", National Semiconductor Technology & Innovation Conference, pp 283-284, July 1995.
7. M. Fallon, G. Martin, M. Redford, R. Prasad, T. Powell, H. Gregory, J. McGinty, "UPD Characterization Using Desirability Functions and Assessment of Process

Variability", National Semiconductor Technology & Innovation Conference, pp 30-33, September 1996.

8. M. Redford, M. Fallon, Z.A. Shafi, N.S. Rankin, J. McGinty, A.J. Walton, "High Voltage RESURF DMOS Process Development Using DFM Techniques A Case Study", Proc 3rd International Workshop on Statistical Metrology, pp 78-81, June 7, 1998.
9. D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis of Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", IEEE Trans. On Semiconductor Manufacturing, Vol 9, No. 4, pp 478-488, November 1996.
10. M. Redford, A.J. Walton, D. Sprevak, R.S. Ferguson, "Application of Covariance Based Models to Fit Response Surfaces to Experimental Data", Proc 4th International Workshop on Statistical Metrology, pp 42-45, June 12 1999.
11. D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis of Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", International Symposium on Semiconductor Manufacturing, pp 85-91, August 1995.
12. Z.A. Shafi, R. Busse, R. Orr, J. McGregor, M. Redford, "Minimizing Breakdown Voltage Drift in High Voltage (>400V) Lateral DMOS RESURF Devices", National Semiconductor Technology & Innovation Conference, pp 196-199, September 1996.
13. A. McGibbon, C. McLeod, G. Robertson, M. Redford, K. Yallup, C. Buchanan, P. Cavanaugh, C. Elgey, "The Detection and Analysis of Platinum Silicide Spiking in Contact Structures", National Semiconductor Technology & Innovation Conference, pp 200-203, September 1996.

14. T. Powell, R. Prasad, C. McGregor, H. Gregory, G. Martin, M. Fallon, M. Redford, "Thin Film Characterization of Silicon Carbide Chrome" National Semiconductor Technology & Innovation Conference, pp 45-48, September 1996.

15. A. McGibbon, K.Yallup, J. Altrip, M. Harper, M. Redford, "A Fingerprint Methodology for Process Transfer and Dual Sourcing", National Semiconductor Technology & Innovation Conference, pp 171-174, July 1997.

16. M. Redford, M. Fallon, J. McGinty, A.J. Walton, M.I. Newsam, G. Michalowicz, C. McGregor, "Identifying and Quantifying Variability Within the LMDMOS 300A Process: A Case Study" IEE Colloquium on improving the efficiency of IC manufacturing technology, No. 1995/153, pp10/1-10/4, University of Edinburgh, April 12th 1995.

17. J. McGinty, G. Long, J.K. Kibarian, M. Fallon, P. McKernan, M. Redford, "Implementation of Statistical TCAD in a Manufacturing Environment", National Semiconductor Technology & Innovation Conference, pp 230-233, September 1996.

18. M. Fallon, K. Findlater, M. Redford, "Novel test structure for electrical measurement of emitter-base misalignment", International conference on microelectronic test structures, Vol 10, pp 156-158, March 1997.

19. J.P. Elliott, M. Fallon, A.J. Walton, J.T.M. Stevenson, A. O'Hara, M. Redford, A. Shafi, "Electrical assessment of planarisation for CMP & SOG technologies", International conference on microelectronic test structures, pp 151-155, March 1997.

20. A. McGibbon, C. McLeod, G. Robertson, M. Redford, K. Yallup, C. Buchanan, P. Cavanaugh, C. Elgey, "High spatial resolution microanalysis of platinum silicide contact spiking in a STEM", 1997 Spring MRS, San Francisco.

HIGH VOLTAGE RESURF DMOS PROCESS DEVELOPMENT USING COVARIANCE BASED RESPONSE SURFACES

M. Redford, M. Fallon, A.J. Walton‡, Z.A. Shafi†, J. McGinty I. Underwood‡

Analog Process Technology Development, National Semiconductor, Greenock, Scotland
†Now with Chartered Semiconductor Manufacturing, Singapore
‡ Department of Electronics and Electrical Engineering, University of Edinburgh, Scotland

ABSTRACT

This paper compares the performance of quadratic and covariance models for a RESURF device. TCAD is used in conjunction with a CCI design and both response surfaces and response variances are compared for both models. Experimentally measured distributions are compared with those predicted by TCAD and the differences used to help identify other possible sources of variation.

Keywords: Design of experiments, Technology CAD, RESURF, covariance, Response Surface Methodology, distributions

1. INTRODUCTION

An integrated circuit process consists of several hundred process steps, each one of which has a number of control factor settings. Selecting the optimum settings of these control factors is crucial if, for example, parametric specifications are to be met. In today's competitive environment, time to market for a new product is of paramount importance if market share is to be increased with it now being no longer feasible to continue debugging process technologies after release to manufacture. Using TCAD tools in conjunction with designed experiments to produce response surfaces provides a methodology to help optimize control factor settings in a more timely manner. Analysis of the impact of response sensitivity to control factor variations provides a greater insight into the process as well as identifying the key control factor(s) for in-line monitoring. A response surface containing a plateau region next to a steep cliff is well known to present a severe challenge to the fitting of response surfaces [1] and the response surfaces fitted in this work pose similar problems. This paper presents the application of a covariance based interpolating procedure that ensures the fitted surface passes through all the simulated data points generated using TCAD tools [2].

As mentioned above one element of a Design For manufacturability (DFM) strategy is Technology CAD (TCAD) which can be used to analyse the effect of control factor on the performance of processes and the resulting devices. However, one of the drawback of using TCAD [3] is the simulation time, especially when one factor a time is being varied in a sequential manner. If the simulation conditions are determined using Design Of Experiment (DOE) techniques, then a predictor can be built from relatively few simulations, therefore, avoiding the alternative of running an inordinately large number of simulations. This approach is well-documented [2-5]. With Response Surface Methodology (RSM) the resulting data is usually fitted using a low degree polynomial that is normally a quadratic. Using polynomials for the model ensures that even when experimental error is present the general trend is fitted [6]. However, simulation results do not have random error associated with them and, in this case, there are potential benefits in generating a model which passes through all the data points.

2. TECHNOLOGY OVERVIEW

For high voltage applications Double-diffused MOS (DMOS) transistors have advantages over conventional MOS devices, since the drift region between the drain and the active channel allows high voltage biasing. The structure under investigation in this paper is a lateral DMOS transistor with a thin epitaxial layer, and applies the Reduced SURface Field (RESURF) principle [7]. This distributes the applied drain voltage uniformly across the drift region and a cross section of the device is shown in figure 1. The operation of a RESURF device relies on careful selection of the both the resistivity and thickness of the epitaxial layer [8] as well as the substrate resistivity. Interaction of the horizontal and vertical depletion

region spreads the electrical field relatively uniformly across the surface, thus maximizing the breakdown voltage for a given device size. The peak electric fields are positioned beneath the ends of the field-plates at the surface of the silicon. These field-plates are

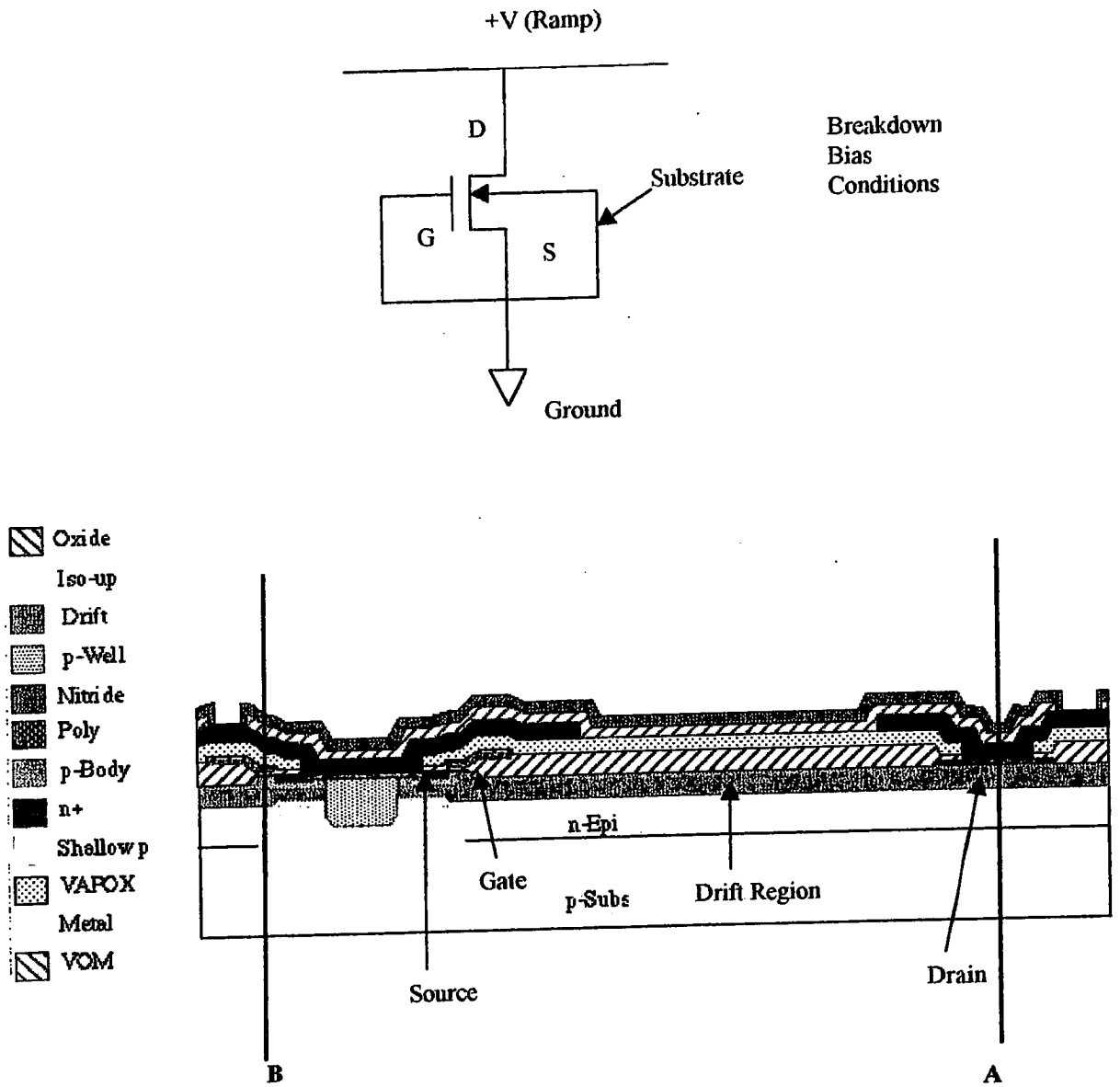


Figure 1. Schematic diagram and cross-section of a RESURF device

designed to minimize the two and three-dimensional electric field crowding which would otherwise reduce the breakdown voltage. The breakdown requirement for the RESURF device examined in this paper is 700V (typical) with a guaranteed minimum of 600 V when $V_{GATE} = 0$ and the source and substrate are connected to each other when the drain voltage ramped. The constraints placed on the development were that the process had to remain as close as possible to an existing SMART power flow in order to facilitate the integration of power transistors thereby simplifying implementation into the manufacturing line. One of the constraints resulting from this approach is that the epitaxy thickness had to be in the 11-13μm thickness range. The fabrication option under consideration in this paper is an epitaxial process with a high resistivity where the doping in the drift region is set using an implant [9]. To meet the required breakdown voltages, a substrate resistivity in the range of 100-150 ohms-cm is required.

3. EXPERIMENTAL PROCEDURE

3.1 EXPERIMENT

The CCI experiment was designed in RS/1. There were five control factors: epitaxy resistivity (Epires), epitaxy thickness (Epithk), drift implant dose (Dose), implant energy (Energy), substrate resistivity (Subres) and one response, breakdown Voltage (BV_{DSS}). The worksheet is shown below in Table 1.

Run #	Subres (Ω-cm)	Epithk (μm)	Epires (Ω-cm)	Dose ($\times 10^{-11}$ cm^{-2})	Energy (keV)	BV_{DSS} (Volts)
1	107.5	11.25	10.75	0.4	122.5	800
2	142.5	11.25	10.75	0.4	117.5	860
3	107.5	12.75	10.75	0.4	117.5	780
4	142.5	12.75	10.75	0.4	122.5	860
5	107.5	11.25	14.25	0.4	117.5	720
6	142.5	11.25	14.25	0.4	122.5	800
7	107.5	12.75	14.25	0.4	122.5	740
8	142.5	12.75	14.25	0.4	117.5	800
9	107.5	11.25	10.75	0.6	117.5	840
10	142.5	11.25	10.75	0.6	122.5	580
11	107.5	12.75	10.75	0.6	122.5	740
12	142.5	12.75	10.75	0.6	117.5	580
13	107.5	11.25	14.25	0.6	122.5	820
14	142.5	11.25	14.25	0.6	117.5	860
15	107.5	12.75	14.25	0.6	117.5	820
16	142.5	12.75	14.25	0.6	122.5	860
17	90	12	12.5	0.5	120	760
18	160	12	12.5	0.5	120	860
19	125	10.5	12.5	0.5	120	800
20	125	13.5	12.5	0.5	120	860
21	125	12	9	0.5	120	600
22	125	12	16	0.5	120	780
23	125	12	12.5	0.3	120	720
24	125	12	12.5	0.7	120	600
25	125	12	12.5	0.5	115	840
26	125	12	12.5	0.5	125	840
27	125	12	12.5	0.5	120	840

Table 1. Worksheet for CCI designed experiment

3.2 TCAD CALIBRATION PROCEDURE

The process flow was modelled using SUPREM-3, and the doping profiles used as input for the device simulator Medici. The initial process specification was based on previous work [10] with model verification having been previously performed on the SMART power process. The TCAD files from the SMART process were adapted and this provided a higher degree of confidence in the accuracy of the simulated results. In addition the process model included actual measured furnace profiles and results were calibrated to oxide thicknesses and measured dopant profiles with transistor threshold voltages V_t being matched by adjusting the interface charge, Q_{ss}. Examples of the profiles produced by SUPREM-3 are shown in figures 2 and 3.

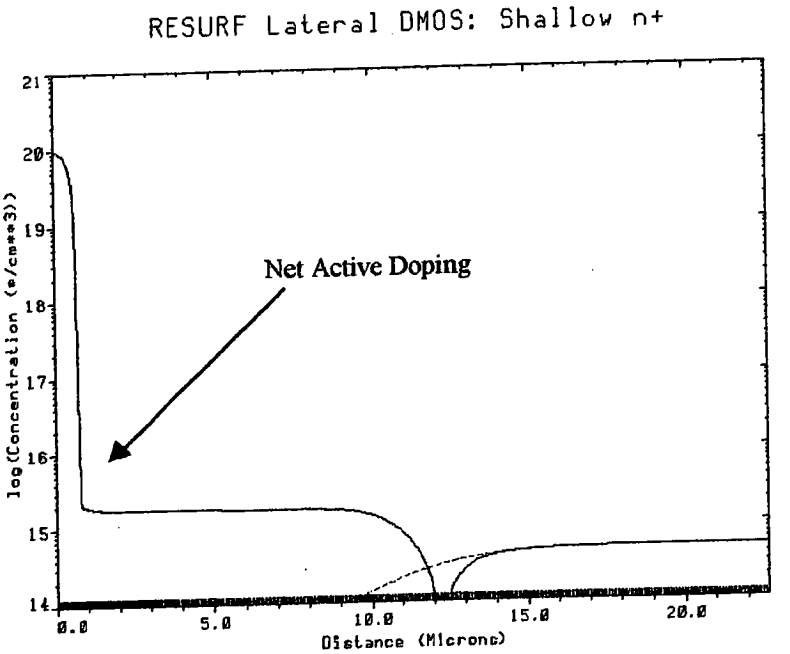


Figure 2. Dopant profile for shallow n+ region (Section A in figure 1) modelled by SUPREM-3

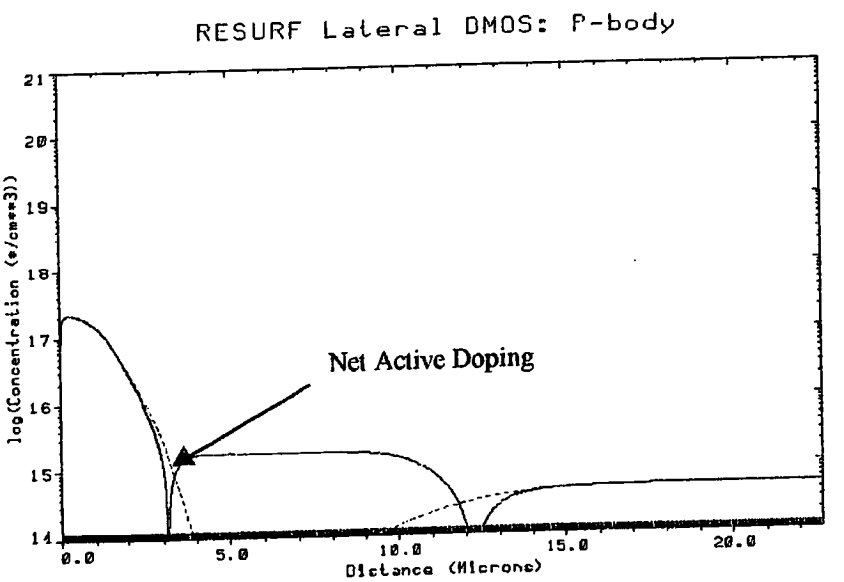


Figure 3. Dopant profile (Section B in figure 1) modelled by SUPREM-3

Figure 4 shows an example cross section of the device where the width of the simulated structure is of the order of 90μm and the depth 75μm. Medici, a 2D device simulator, was used to model the RESURF device that is essentially a 3D in nature and, since the layout of the device affects the breakdown, it is important that this is taken into account. In order to achieve the required accuracy in the minimum time, the structure was split into three distinct regions identified in figure 5. These were:

1. Parallel region modelled using rectangular coordinates.
2. Source fingertips modelled using cylindrical coordinates.
3. Drain fingertips modelled using cylindrical coordinates.

The lowest value of the three simulated breakdown voltages was entered into the worksheet which was used to generate the response surface. This methodology has previously been validated [8] and in section 4 of this paper these results are compared with those obtained experimentally.

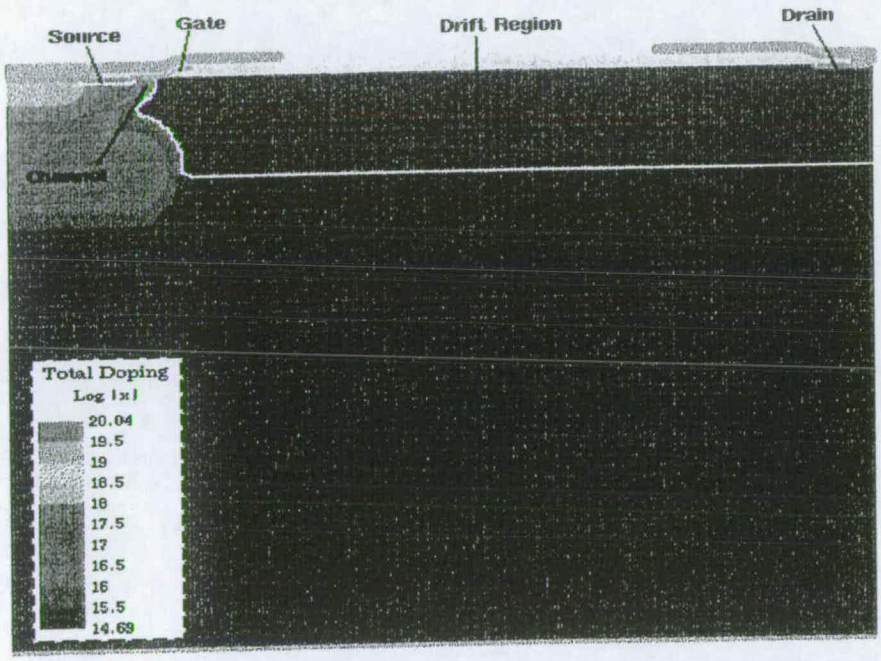


Figure 4. Cross-section of RESURF device as modelled by Medici

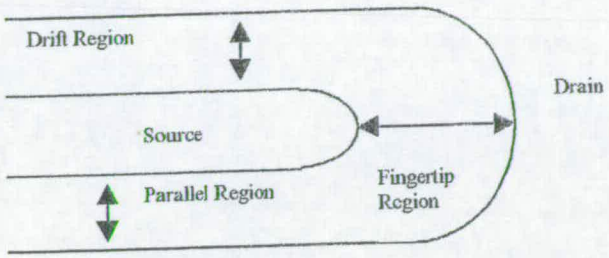


Figure 5. Fingertip region of high voltage transistors

4. RESULTS

4.1 Covariance based interpolation

Conventional response surfaces used in conjunction with designed experiments are polynomial based, typically quadratic. These models provide a smooth surface that on some occasions may miss out the relevant detail. As previously mentioned simulated results have no random error associated with them and so there is a good argument to generate a model that perfectly fits all the data. However, with experimental measurements it is important to fit the general trend as it makes no sense to model random errors. One of the problems with polynomial models is that sometimes they cannot be distorted to fit the data. In this case it is possible to use a covariance model [2]. With experimental data this approach can be considered a two-stage process where the trend is first fitted with a polynomial with the second step being a gentle distortion of the trend so that the response goes through selected data points that are believed to reflect reality. It is important that the covariance based approximation does not alter the surface in such a manner that the measurement noise is slavishly modelled and the capability of the model to predict the response at other points is compromised. However, in this paper only simulated results are fitted with response surfaces and in this case there is justification in fitting the response surface through all data points. The results presented in this paper present a severe challenge for any function as the breakdown voltage consists of areas that have very rapidly varying gradients [1]. It can be observed that the variation of BV_{DSS} as shown in figure 6 presents such a challenge and in situations such as this that the covariance approach has potential benefits.

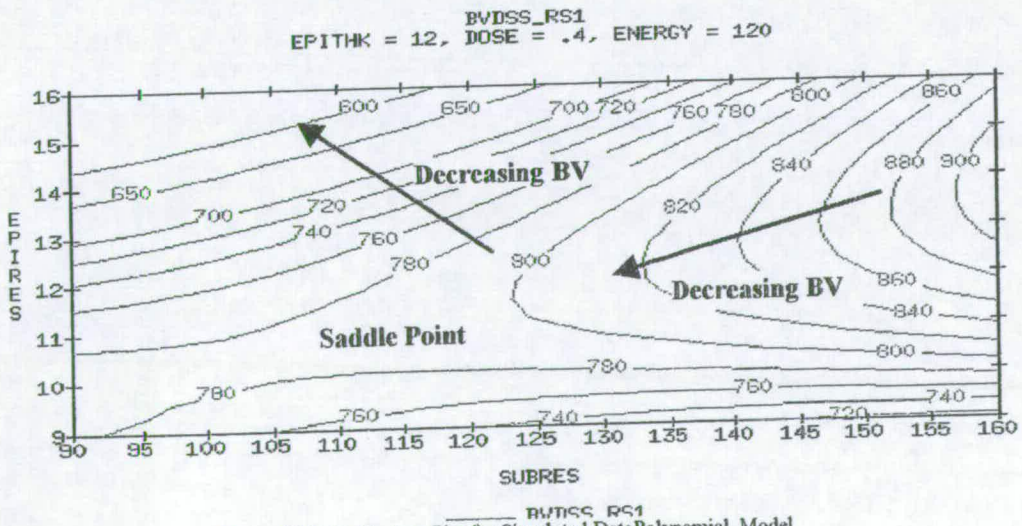


Figure 6 BVdss Contour Plot for Simulated DataPolynomial Model

4.2 TCAD and DOE Procedures

To understand and evaluate the impact of the five-variables or control factors on the breakdown voltage RS/1 software was used to design an experiment. The type of experimental design selected was Cubic Central Inscribed (CCI) which resulted in twenty seven simulation data points. It should be remembered that the simulated breakdown was determined from the minimum of three simulated values: one for portions of the structure where source and drain stripes are parallel to one another; one for the source fingertip; and one for the drain fingertip. The results of RSM analysis shown in figure 6 have already indicated that there is a 'workable' region that meets the required specification. In figure 7 the quadratic model fitted in figure 6 is compared with a covariance based model fitted using Calphurnia [11] and it is apparent that the polynomial surface has been unable to fit the worksheet data. Calphurnia has also been used to fit a covariance based response surfaces for drift implant doses of 0.4 and 0.5×10^{-11} cm^{-2} and a comparison with the contours for a quadratic model are shown in figures 8 and 9. In this example there is one data point where the residual percentage error is in double figures. This data point is run 9. A comparison of the response surfaces for polynomial and covariance fits for the control factor settings for run 9 is shown in figure 9. This demonstrates how the covariance effectively distorts the surface so that the response goes through the data point generated by run 9 (840V) whereas the quadratic model gives a value of 800V. It can be observed that the saddle point has disappeared from the area of interest and this highlights how the improved covariance fit has the potential to improve the accuracy of prediction.

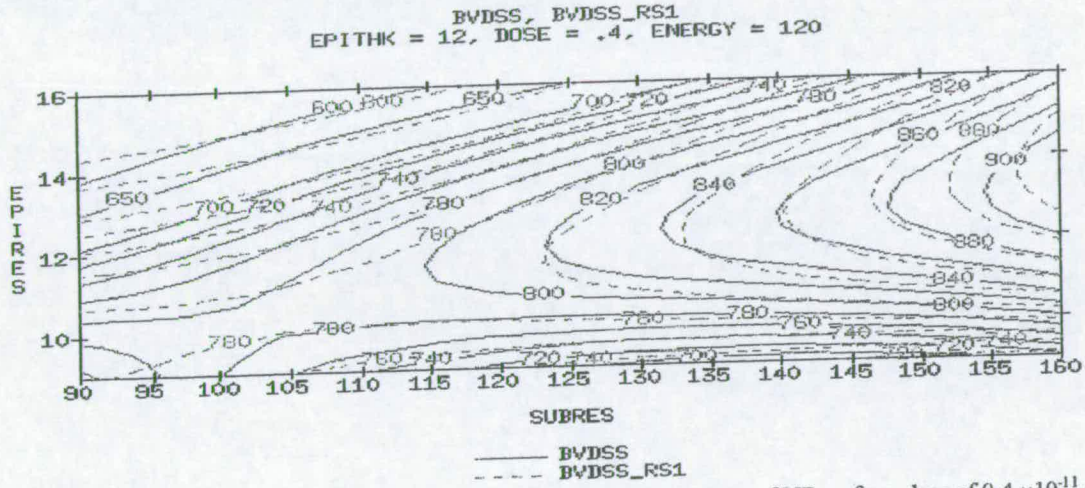


Figure 7. Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} for a dose of 0.4×10^{-11} cm^{-2}.

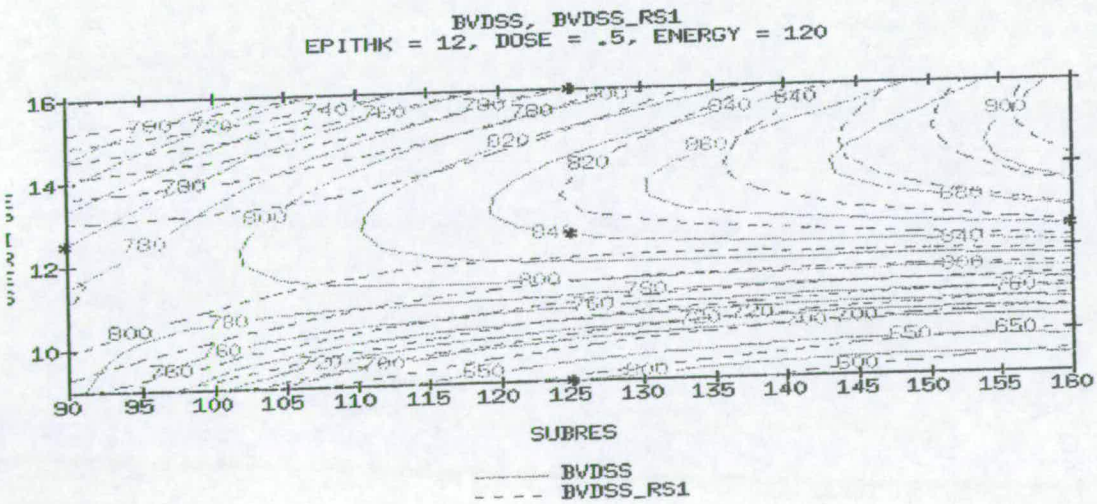


Figure 8. Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} for a dose of 0.5×10^{-11} cm^{-2}.

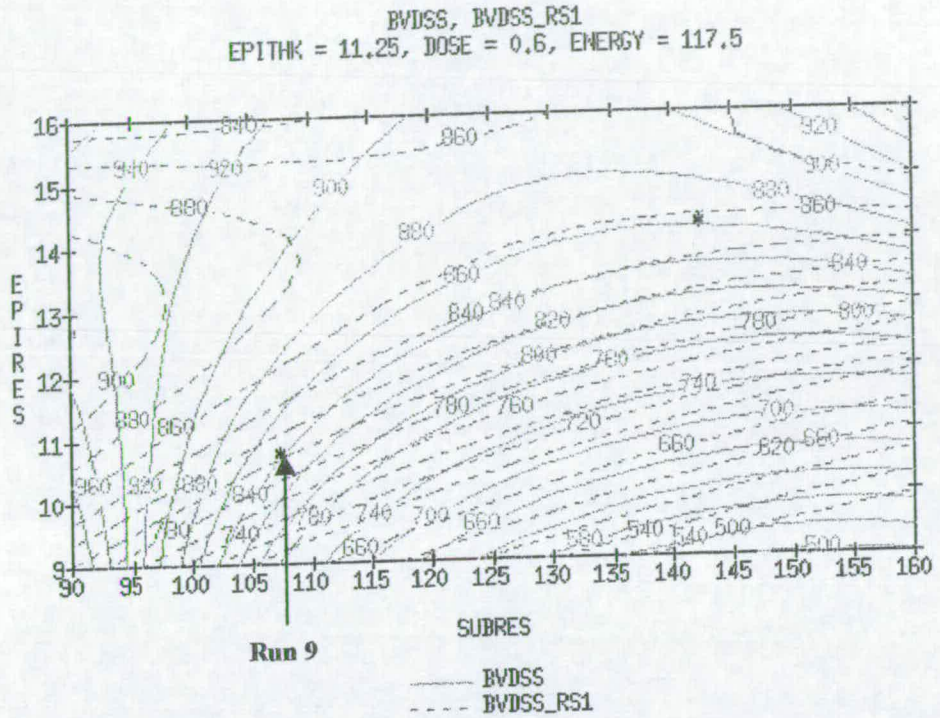


Figure 9. Comparison of polynomial (dashed) and covariance (solid) contours of VB_{DSS} indicating the combination of control factors that make up run 9.

4.3 Breakdown Distributions

The ability to accurately create response surfaces is an important component of Design For Manufacturability (DFM) since they provide important information on the manufacturability of the process under consideration. Calphurnia can use the response surfaces to generate response distributions resulting from the expected variations in the control factors that are typically seen in a production environment. In this example the software makes the assumptions that the control factor distributions are Gaussian in nature, though any distribution function can be used. The standard deviations of all the control factors must first be defined after which it then becomes possible to generate their distributions. These are then used in conjunction with the response surface(s) to create the response distribution(s). Using the standard specified in reference [8] for the control factors response distributions have been generated using both covariance and polynomial models are shown in figure 10 and 11 respectively. Table 2 compares parameters for the distributions obtained from two different implant doses for both covariance and quadratic based distributions. It can be observed that the covariance means are higher but this is of no statistical significance.

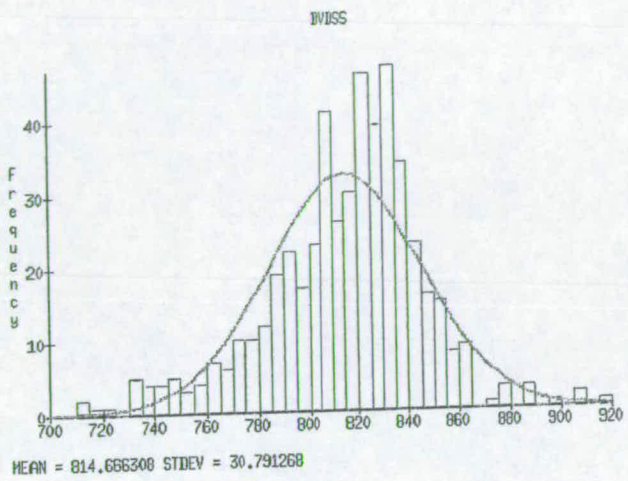


Figure 10. VB_{DSS} distribution calculated using the covariance model

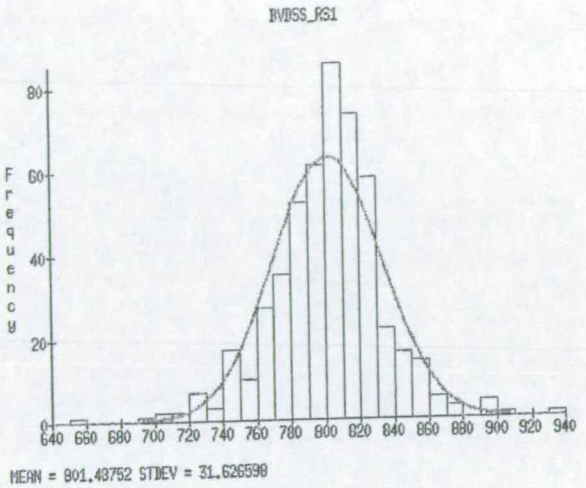


Figure 11. VB_{DSS} distribution calculated using the quadratic model

	Dose - 0.4×10^{-11} cm^{-2}		Dose - 0.5×10^{-11} cm^{-2}	
	(Mean) (V)	BV_{DSS} (Std Dev) (V)	BV_{DSS} (Mean) (V)	BV_{DSS} (Std Dev) (V)
Quadratic Model	801	32	812	30
Covariance Model	814	31	825	29
Difference	13	1	13	1

Table 2. Comparison of parameters for the distributions obtained for two different implant dose settings. The covariance and quadratic based distributions for the 0.4×10^{-11} ions cm^{-2} dose are shown in figures 10 and 11

Both polynomial and covariance based distributions using the target values for run 9 are shown in figure 12 and in this case it can be observed that their shape is different with the covariance distribution exhibiting a longer tail for high breakdowns. Table 3 compares the distributions distribution parameters associated with the two response models. As would be expected from the contours shown in figure 9 there is a big difference in the mean values. However, there is also a significant difference in the variance. If manufacturability is to be considered then the variability of the response needs to be predicted as accurately as possible.

	Run 9 target values	
	BV_{DSS} (Mean) (V)	BV_{DSS} (Std Dev) (V)
Quadratic Model	803	98
Covariance Model	874	140
Difference	71	42

Table 3. Comparison of parameters for the distributions obtained from covariance and quadratic based distributions using run 9 control factor settings as the target values. The distributions from which these parameters extracted are shown in figure 12.

Covariance Based Distribution

Polynomial Based Distribution

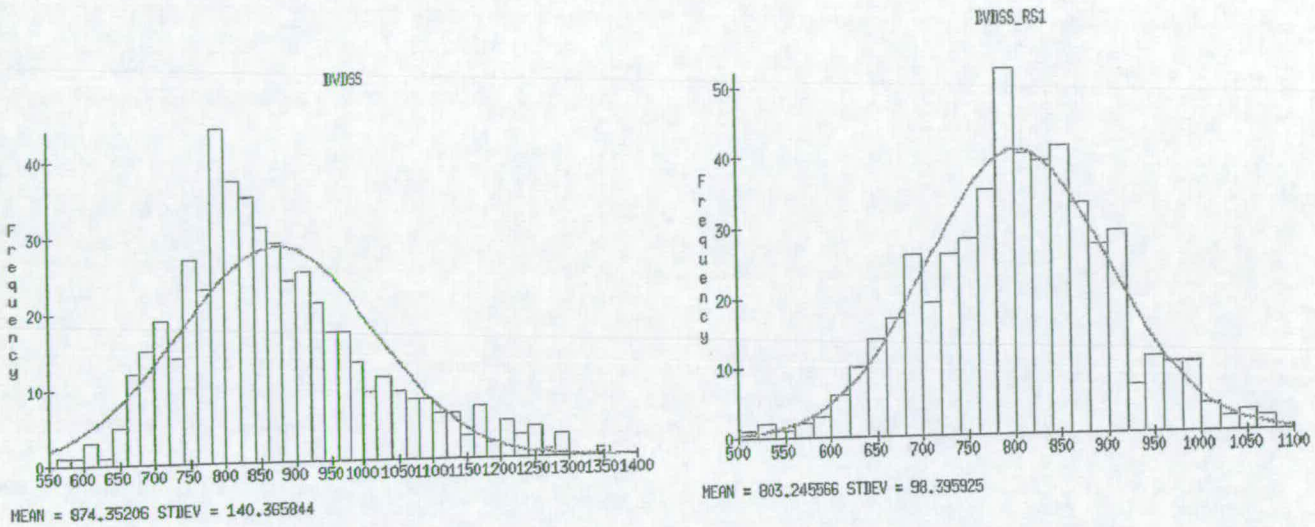


Figure 12. VB_{DSS} distribution calculated using the covariance and quadratic model

4.4 Experimental Results From Silicon

Figure 13 shows the measured distributions for different implant doses of 0.4, 0.5 and 0.6 $\times 10^{-11}$ ions cm^{-2} and it can be observed that there appear to be two distributions present in the data. The simulation results predict the 'major' distribution but not the failures that have lower values of breakdown voltage. It would appear that there is a second factor in operation that is not being modelled by the TCAD simulations. It is known that the layout of the device has an impact on the breakdown. Field plates at the source and drain are used to minimise the two and three-dimensional field crowding to help maximise the breakdown voltage. Non-optimal design of these will lead to occasional lower breakdown voltage than would otherwise be expected as would process issues such as field oxide thickness beneath the field plates. These issues obviously need further investigation but what has been identified from the TCAD work is that the developed process is intrinsically capable of meeting the required specifications and the low measured breakdown voltages are not due to variability in the control factors investigated in the experiment.

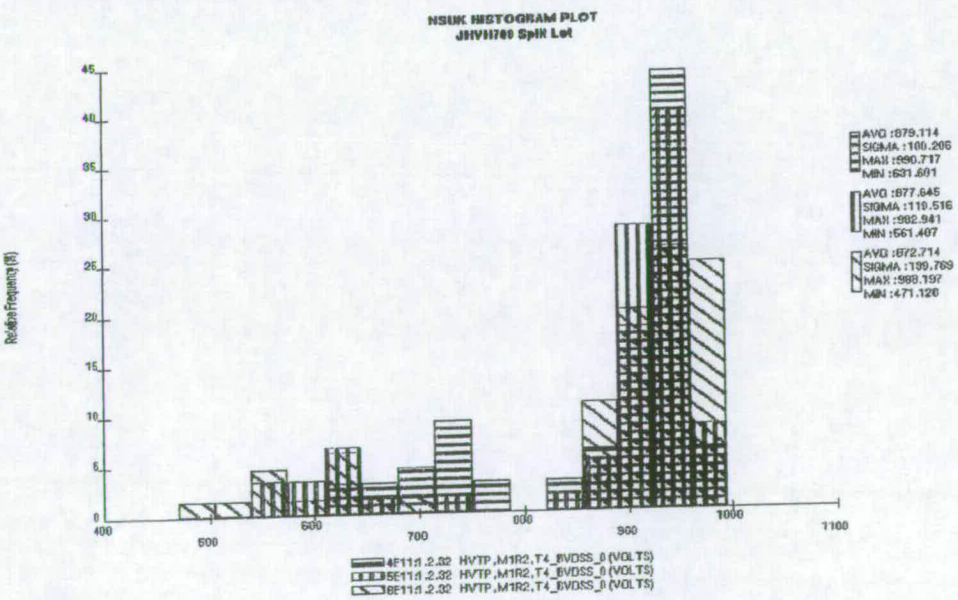


Figure 13 Breakdown Distribution from Silicon for Drift Implant Doses 0.4,0.5,0.6

5. CONCLUSIONS AND OBSERVATIONS

This paper has compared conventional quadratic polynomial response surfaces with covariance fits and shown that near data points where there are small residual errors the difference in predictive capability is small. However, when the quadratic polynomial exhibits large residual errors near data points the use of covariance based models provides an improvement. In the example shown in this paper it has been demonstrated that the covariance model can be successfully fitted to a CCI design with five control factors.

The distributions generated from the TCAD results have confirmed that the variation observed experimentally is caused by factors other than those used as control factors in the CCI experiment. Further work is required to fully identify the causes of lower breakdown voltages.

7. REFERENCES

[1] T.G. Waring, A.J. Walton, D. Sprevak, R.S. Ferguson. "The Sensitivity Of The Parameters Of Covariance Based Response Surfaces" Proc 3rd International Workshop on Statistical Metrology 1998, June 7, 1998

[2] T.G. Waring, A.J. Walton, D. Sprevak, R.S. Ferguson. "Application of Covariance Structures to Improve the Fit of Response Surfaces to Simulation Data" to be published in IEEE Trans Semiconductor Manufacturing.

[3] M. Redford, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, J.P. Elliott, G.A. Allan, A.J. Walton, 'DFM Strategies In Technology Development', National Semiconductor Technology & Innovation Conference, 1995.

[4] A.J. Walton, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, "Using RSM Techniques to Contour Plot Parameters Related to Response Distributions of Semiconductor Processes", Proc ESSDERC '94, 1994, pp 829-832.

[5] M. Fallon, G. Martin, M. Redford, R. Prasad, T. Powell, H. Gregory, J. McGinty, 'UPD Characterization Using Desirability Functions and Assessment of Process Variability', National Semiconductor Technology and Innovation Conference, 1996.

[6] R.A. Bates, H.P. Wynn, "Tolerancing and Optimization for Model-Based Robust Engineering Design", Quality and Reliability Engineering, vol 12, pp 119-127, 1996.

[7] J.A. Appels, M.G. Collet, P.A.H. Hart. H.M.J. Vaes, J.F.C.M. Verhoeven, 'Thin layer high-voltage devices (RESURF devices). Philips Journal of Research, vol 35, no 1, pp1-13. 1980

[8] M. Redford, M. Fallon, Z.A. Shafi, N.S. Rankin, J. McGinty, A.J. Walton, "High Voltage RESURF DMOS Process Development Using DFM Techniques A Case Study", Proc 3rd International Workshop on Statistical Metrology 1998, June 7, 1998

[9] E.J. Wildi, P.V. Gray, T.P. Chow, and H.R. Chang. Modeling and process implementation of implanted RESURF type devices. In IEDM Digest of Technical Papers, pages 268-271, 1982.

[10] J.M. McGregor & M. El-Diwany. 450V lateral DMOS device module compatible with LMDMOS technology. In National Semiconductor TCAD Workshop '93, pp 10-13, 1993

[11] M.I. Newsam, "Calphurnia V2.0 Manual", Dept. of Electrical Engineering, University of Edinburgh, Scotland.

[12] D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis if Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", IEEE Trans. On Semiconductor Manufacturing, Vol 9, No. 4, Nov 1996

DOE ANALYSIS OF LAYOUT/PROCESS INTERACTIONS IN BIPOLAR SCHOTTKY TECHNOLOGY

Alastair McGibbon, Alan Shafi†, Nigel Cameron, Chun Ng, Mark Redford

Analog Process Technology Development, National Semiconductor, Greenock, Scotland
†Now with Chartered Semiconductor Manufacturing, Singapore

ABSTRACT

In this paper, we demonstrate a rigorous and planned design of experiment (DOE) approach to process development with an aim to providing a fully characterised and manufacturable process in a single cycle of silicon. The key to this approach is to examine design rule, design practice and physical attributes and their interactions with the same methodology to ensure that product requirements and variations are fully manufacturable. As a case study, we describe the development and analysis of a Schottky device suite for an analog bipolar technology. We discuss the detailed pre-processing preparation, which has a strong Technology Computer Aided Design (TCAD) element, resulting in a co-ordinated DOE mask set and experimental DOE plan. We then discuss examples of analysed data looking at the individual and combined attributes of design rules, design practices and manufacturing sensitivities. We also describe the way in which the methodology is flexible enough to encourage and enable the development of novel devices (in this case a selective field threshold adjust implanted diode) with a view to immediate implementation on first silicon.

Keywords: Schottky, DOE, manufacturability, RSM.

1. CYCLE TIME REDUCTION IN PROCESS DEVELOPMENT AND IMPLEMENTATION

The continuous drive in the semiconductor industry for faster product implementation at lower cost means that the need for structured process development and implementation methodologies becomes an imperative. The metric required to fulfill this goal is the complete characterisation, using a single cycle of silicon, of a process and its devices in terms of both design collateral (supported and modeled devices, design rules and design practices) andmanufacturability (process sensitivities and operating limits). This includes making sure that the devices and practices recommended to the product designer are fully robust within the manufacturing limits of the process.

The goal outlined above applies equally to both the implementation of a new process, whether it be deep sub-micron CMOS or analog bipolar, or the expansion of process functionality through modular development (e.g. new device implementation, the addition of embedded technologies etc.). A schematic diagram outlining the single-cycle development and implementation methodology is given in Figure 1. Several elements are key to the success of this methodology, primarily those related to the planning and understanding of the process and devices*prior* to any silicon being run: The process itself is modeled and understood as much as possible using Technology CAD (TCAD); detailed discussions are required with the process customer (the designer) to ensure that all device layout and parametric criteria are met (this includes providing the required SPICE device modeling information on first silicon); As much historical data that exists on related processes and devices must be collated. From a manufacturability standpoint, a pareto of risks and sensitivities must be compiled, again through the use of TCAD and process knowledge, but also (if necessary) through the implementation of short loops of silicon around sensitive unit processes. Once the main sensitivities and requirements are understood, a rigorous experimentation plan must be implemented and executed. This involves not only process design of experiments (DOEs), but most critically layout DOEs through the creation of a DOE mask set. It is this that enables the full scope of process manufacturability and design robustness to be investigated on a single cycle of silicon.

The case study given here to demonstrate the layout/process DOE approach is that of the implementation of aSchottky device capability on a super-micron analog bipolar process. The recent introduction of a barrier metal andsilicided contact geometry had enabled the opportunity for a Schottky layout library for designers. However, several new design rules were required, as was the understanding of the impact of existing process modules (for example optional sinker and field threshold adjust - FTA - masks), and the effect of existing manufacturing limits (such as those related to the contact geometries and epilayer control) on Schottky devices. In addition, the opportunity is taken (as in the majority of

development projects) to explore novel devices and applications. However, the same rigorous approach is taken to ensure that successful new devices can be implemented into manufacturing along with those originally planned.

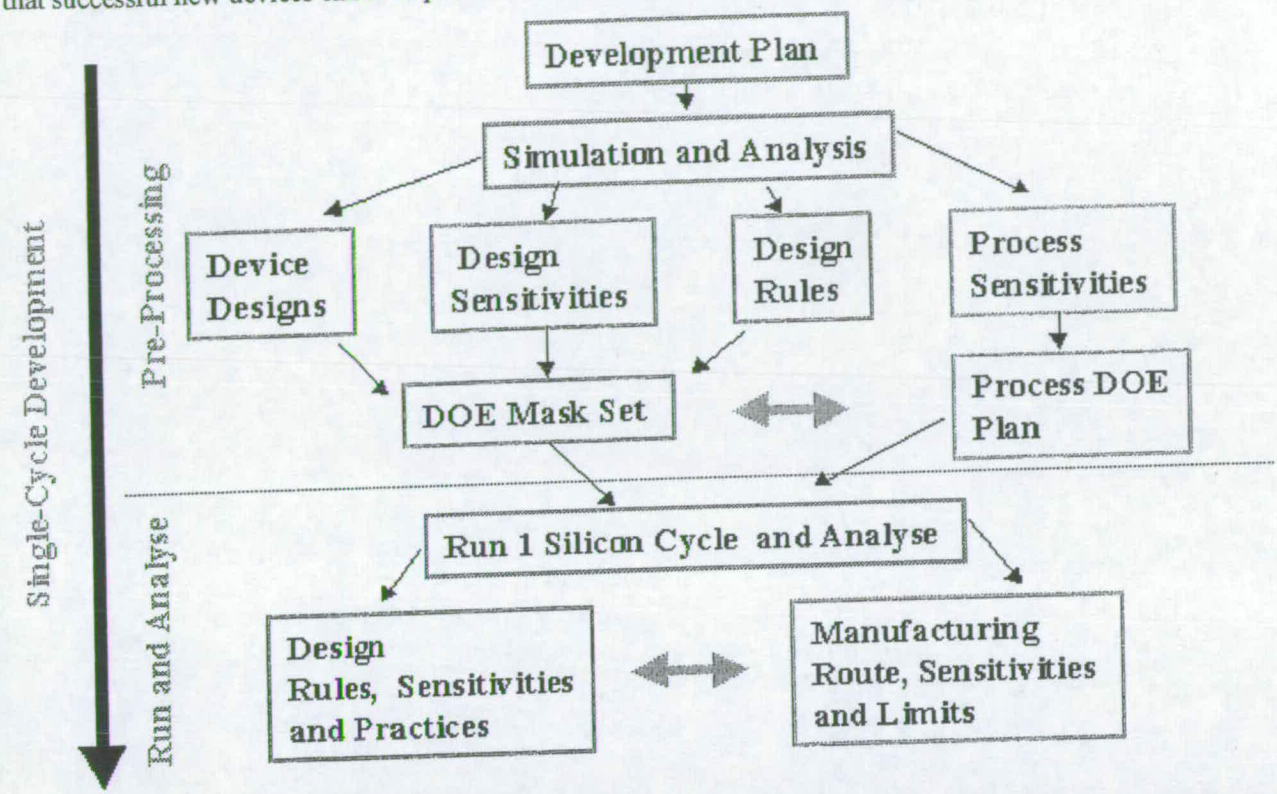


Figure 1: Schematic diagram outlining the principal steps involved in enabling single-cycle development that addresses design rules, design practices and manufacturing sensitivities and their interactions.

2. PRE-PROCESSING PREPARATION

As shown schematically in Figure 2, Schottky-based devices and layouts belong to two main families - guard ring and field plate. In theory, as long as the relative performance of the two is comparable, the latter ispreferable as the device footprint is smaller. However, in practice, it is well documented that current crowding at the edges ofSchottky field plate devices, together with non-perfect interfaces can contribute to degraded reverse breakdown characteristics.

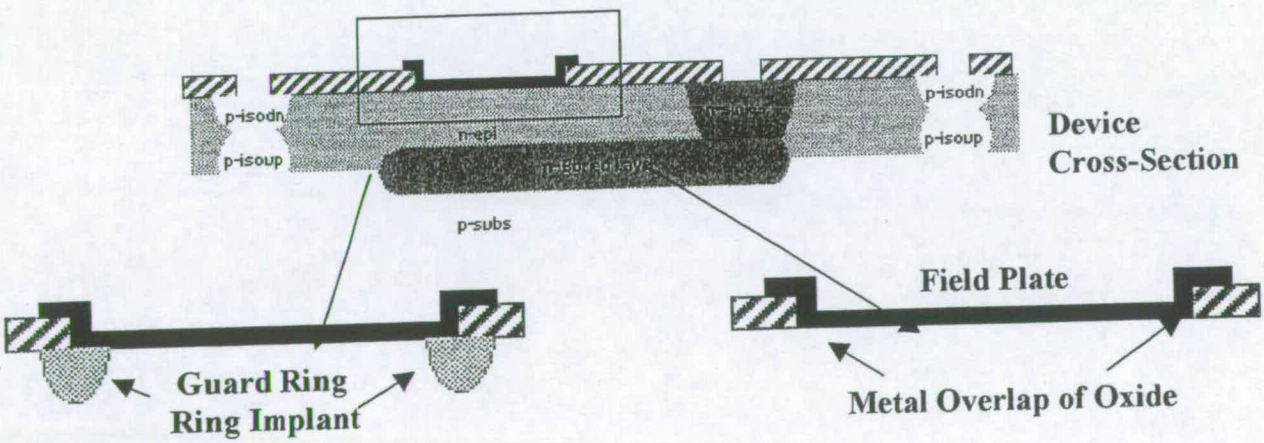


Figure 2: Schematic diagram showing a generic Schottky diode cross-section for a junction-isolated bipolar technology, with details given on the difference between field plate and guard ring structures.

This consequence is avoided in Guard Ring devices, where the implanted ring (in this case using the p-base of the technology) results in a consistent reverse breakdown voltage equivalent to the open-circuit collector-base breakdown of an NPN transistor (BVcbo). In the technology under discussion here, other options available (Figure 2) to the designer include the use or omission of an n+ sinker contact from the cathode to the buried layer, and the use of a selective FTA mask. Even with two such options, the number of possible layouts and rules for a Schottky-based device suite becomes considerable, hence the need for a DOE-based procedure to assist in the characterisation. Although it is diode properties only that are discussed in this paper (as they are equally applicable to clamped Schottky transistors) the DOEs undertaken also examined such effects on transistor clamping performance as a function of layout.

As part of the understanding process required prior to the running of silicon, TCAD simulations were carried out on basic Schottky structures using TMA TWB software, an example of which is shown in Figure 3. Characteristics under consideration using TCAD were physical behaviours such as forward current flow as a function of device layout and the position of reverse breakdown for different diode geometries (e.g. field plate or guard ring) and layouts (e.g. with or without FTA implant). Design rules that could be estimated included the inner guard ring dimension and the distances between the n+ regions and the guard ring that would optimise reverse breakdown. Forward Schottky voltages at a range of currents were also included in the modeling.

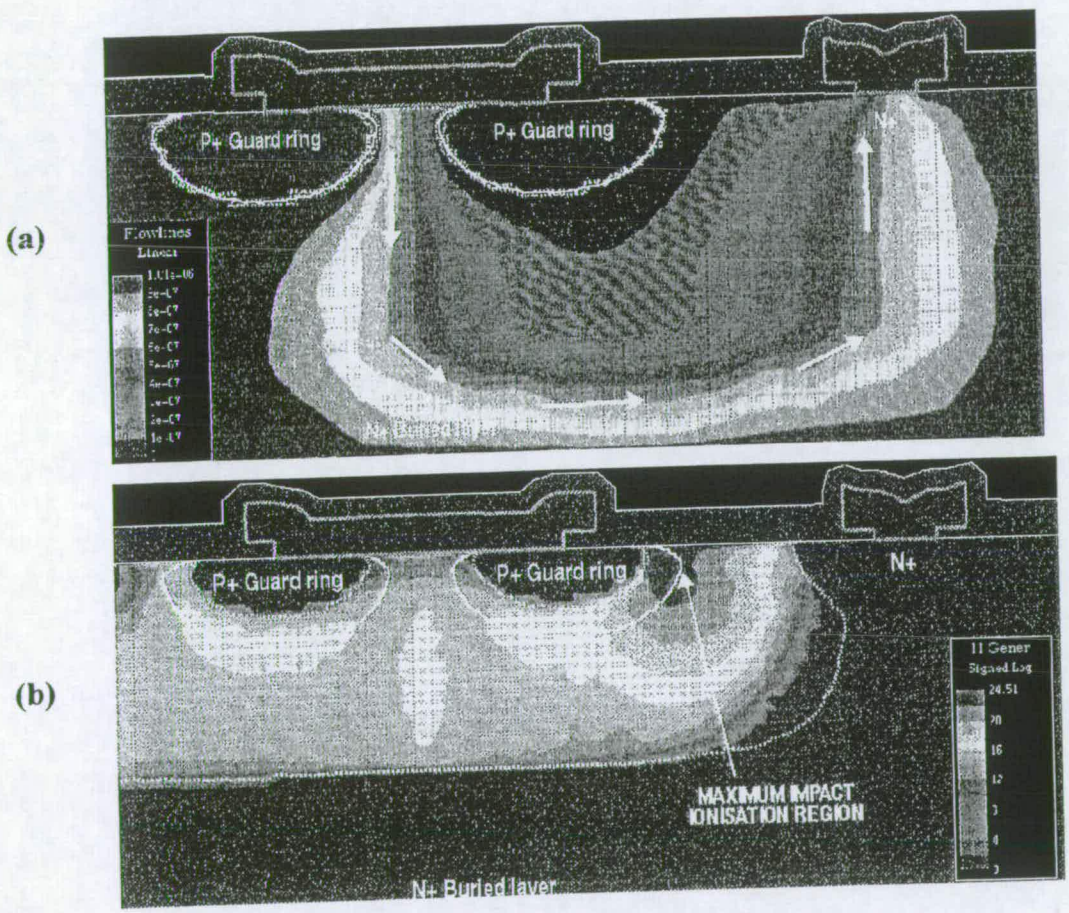


Figure 3: 2-Dimensional TCAD data from a guard ring Schottky diode showing (a) forward bias current flow and (b) the location of the reverse bias breakdown.

A general description of the methodology used to develop the layout/process DOE approach for first silicon success is shown in Figure 4. The first point to note is that the principal DOE responses are chosen carefully at the outset to ensure both commonality between all types of DOEs and applicability to the parametric requirements of the designer.

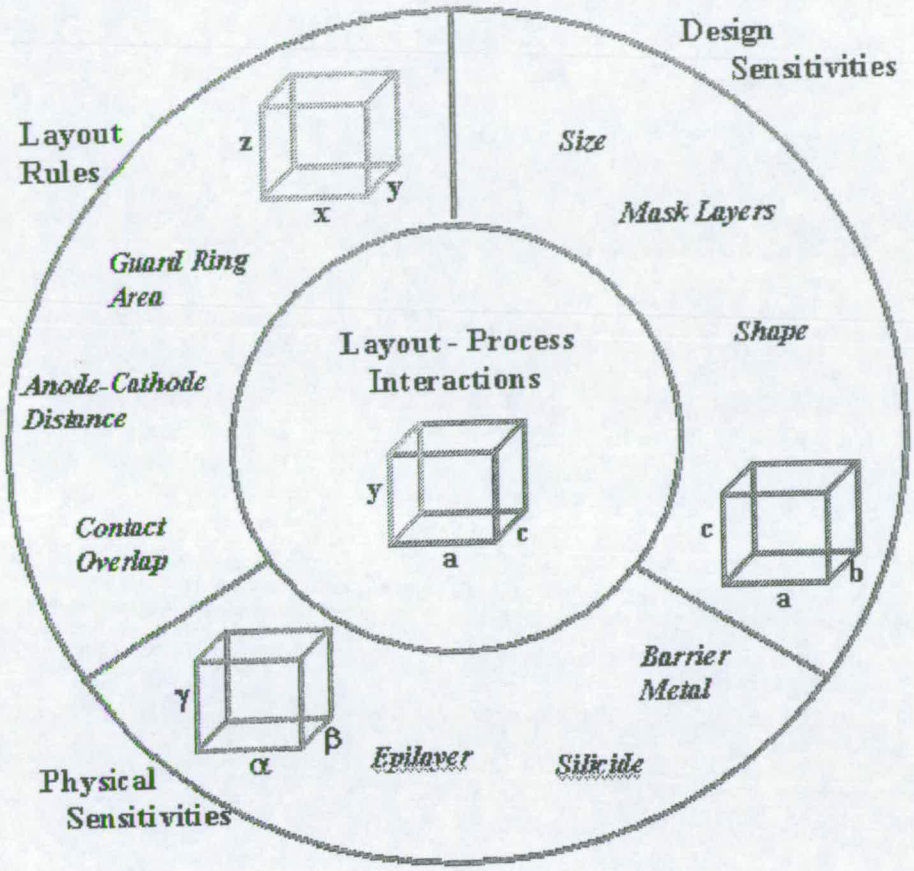


Figure 4: Schematic diagram showing the basic combined DOE approach to process development, with emphasis on key sensitivities for Schottky device development and implementation into manufacturing

For Schottky devices, this is a relatively simple subset of electrical parameters - the forwardSchottky voltage at a number of currents, the reverse bias breakdown voltage, and the reverse bias leakage current at a number of voltages. Also shown in Figure 4 are examples of the principle control factors of interest for layout, design practice and physicalDOEs. For example, the physical parameters in the existing process with the greatest impact on new Schottky device performance were the epilayer resistivity and characteristics related to the silicide and barrier metal deposition. When designing DOEs with both layout and physical control factor components, it was found to be most efficient in terms of cost and cycle time to construct multi-level full factorial layout DOEs - a reason for this is that no matter how many layouts there are, the cost will still be that of one mask set. However, a large number of physicalDOEs will result in both an increased wafer cost component of the development budget and increased silicon manufacturing time.

3. RESULTS AND DISCUSSION

On the completion of experimental silicon DOE processing, the electrical responses described above were measured and documented for each of the layout variations in the development test pattern. As indicated in Figure 4, data was analysed both in terms of the 3 individual groups (design rule, design practice and physical) and their interaction to ensure design manufacturability.

An example of design rule analysis is given in Figure 5: The graphs show simply thebehaviour of the forward Schottky voltage in a guard ring diode at a number of currents (1, 10 and 100μA - which cover the intended operating conditions of the product application of this device) as a function of drawn guard ring inner dimension. Based on TCAD and electrical analyses prior to the running of silicon, the minimum dimension for the reliable operation of this device was chosen as 10x10μm. In normal circumstances, as the guard ring dimension decreases, the forwardSchottky voltage would be expected to increase. However, it can be seen in all three graphs, but particularly that looking at the low current Vs, that this relationship fails at a point around or below that of 8x8μm devices. This is primarily due to guardring side diffusion effects limiting the active Schottky contact region, and is consistent with the TCAD data from the pre-processing preparation.

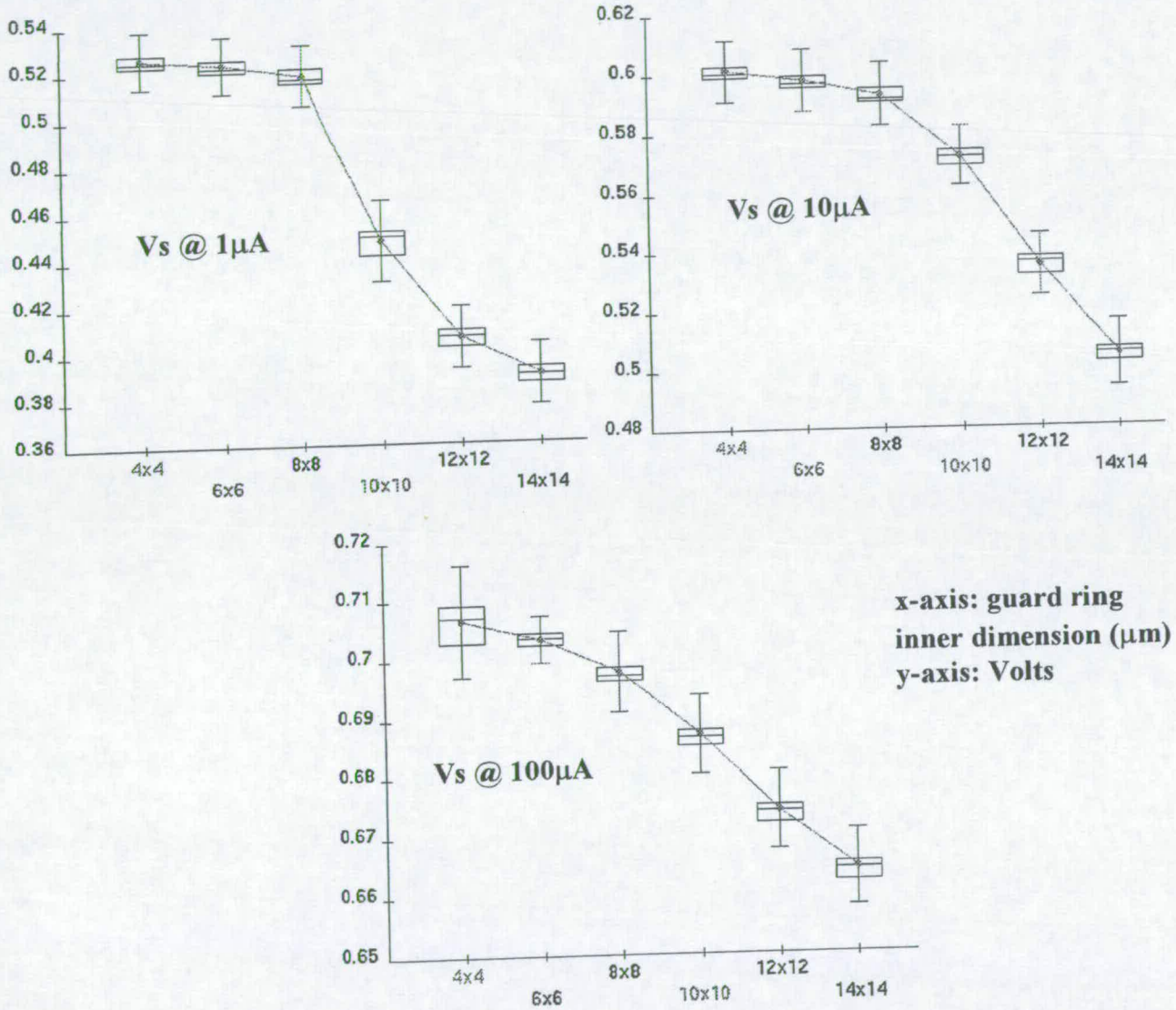
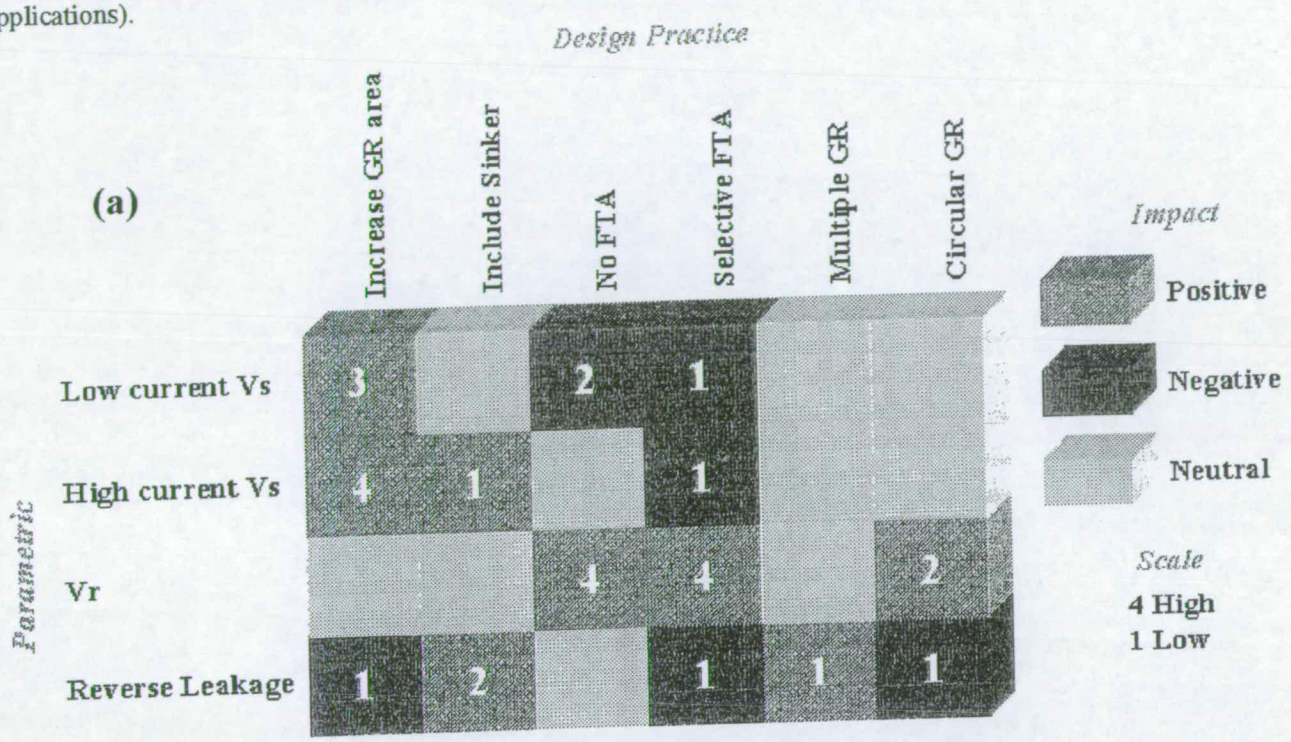


Figure 5: Graphs showing the variation in forward Schottky voltage (Vs) as a function of drawn guard ring inner dimension for currents of (a) 1μA, (b) 10μA and (c) 100μA.

Processing of the design practice DOEs in the development mask set revealed a wealth of information of particular interest to the designer. A schematic diagram showing a summary of design practices (for both guard ring and field plate devices) and the relative impact they have on the key DOE responses is shown in Figure 6. The data can be used can examine not only

major design impacts such as that of active contact area and shape on forwardSchottky voltage, but also secondary effects such as the effect of device geometries on reverse breakdown (which may be used, for example, to design an optimum high voltage device) and the impact of layout on reverse leakage current (which may have a cumulative effect in multiple device applications).



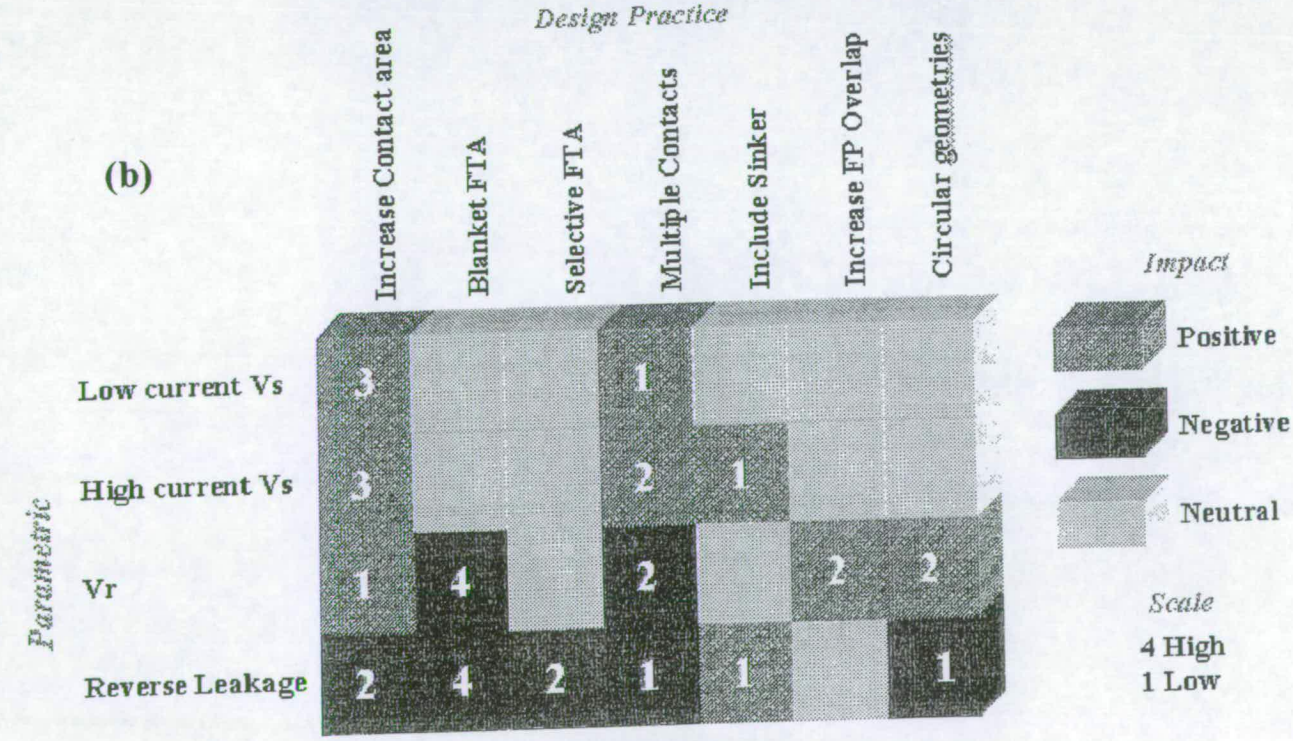


Figure 6: Tables depicting the relative impact of layout design practices on key electrical responses for both guard ring (a) and field plate (b) Schottky devices.

As stated previously in this paper, the rigorous layout approach adopted here also enables the detailed exploration of novel devices with a view to immediate implementation if successful. An example of this is indicated in Figure 6, where it can be seen that by using the FTA implant selectively in a device (in the guard ring or field plate region), the reverse current voltage can be dramatically improved without significant impact to the forward Schottky properties. A more detailed diagram, showing the parametric impact of this effect when compared to blanket implanted or non-implanted devices is shown in Figure 7.

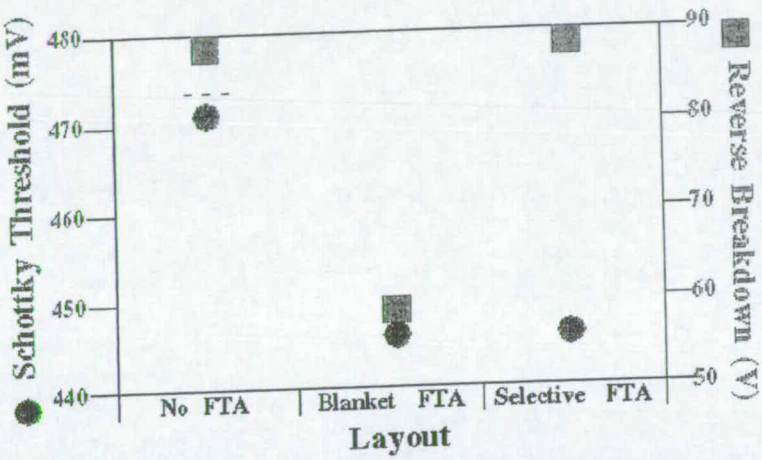


Figure 7: Data showing the relative impact on forward Schottky voltage and reverse breakdown voltage from (left) a Schottky diode with no FTA implant, (centre) one with a blanket FTA implant, and (right) one with a selective FTA implant in the guard ring region.

The most amenable technique for the analysis of the physical DOE data (and their interaction with design rule verification data) was the use of response surface modeling (RSM) methodologies. An example of this approach is given in Figure 8,

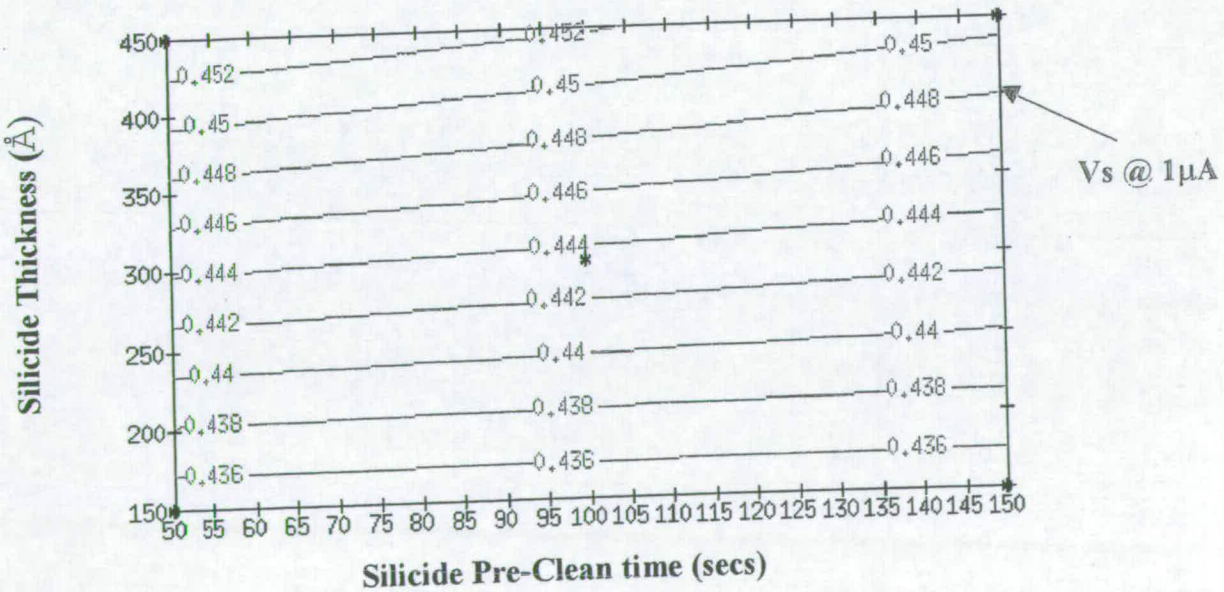


Figure 8: RSM plot showing the effect of silicide pre-clean and thickness on forward Schottky voltage at low currents.

where the variations in low current Schottky voltage for a particular device is examined as a function of silicide thickness and pre-clean time in manufacturing. In this case, the results indicate a low magnitude linear effect on Vs as a function of thickness, with pre-cleans only beginning to impact the response at high silicide thicknesses, well outwith the actual operating window of this particular process, indicating a high degree of robustness.

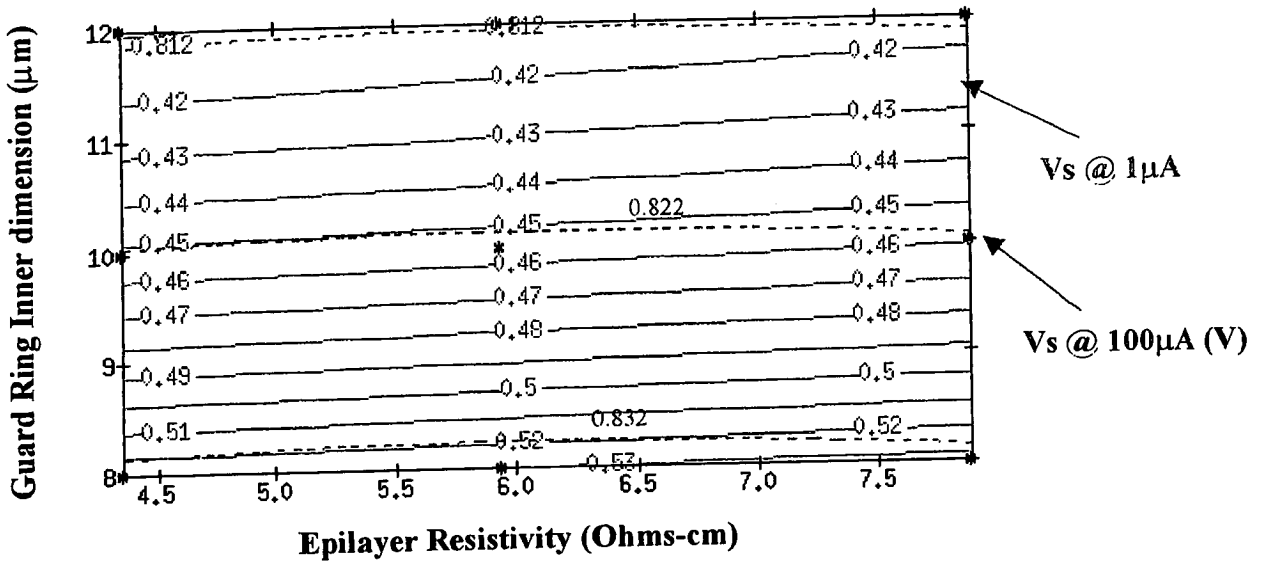


Figure 9: RSM plot showing the effect of drawn guard ring inner dimension and epilayer resistivity on forward Schottky voltage at both low and high currents.

As a demonstration of the way in which design rule sensitivities can be compared against manufacturing limits, we show the RSM analyses of the impact of epilayer resistivity and guard ring inner dimension on forward Schottky voltage (Figure 9) and that of barrier metal resistivity and n+ sinker to guard ring spacing on the reverse bias breakdownvoltage ,Vr (Figure 10).

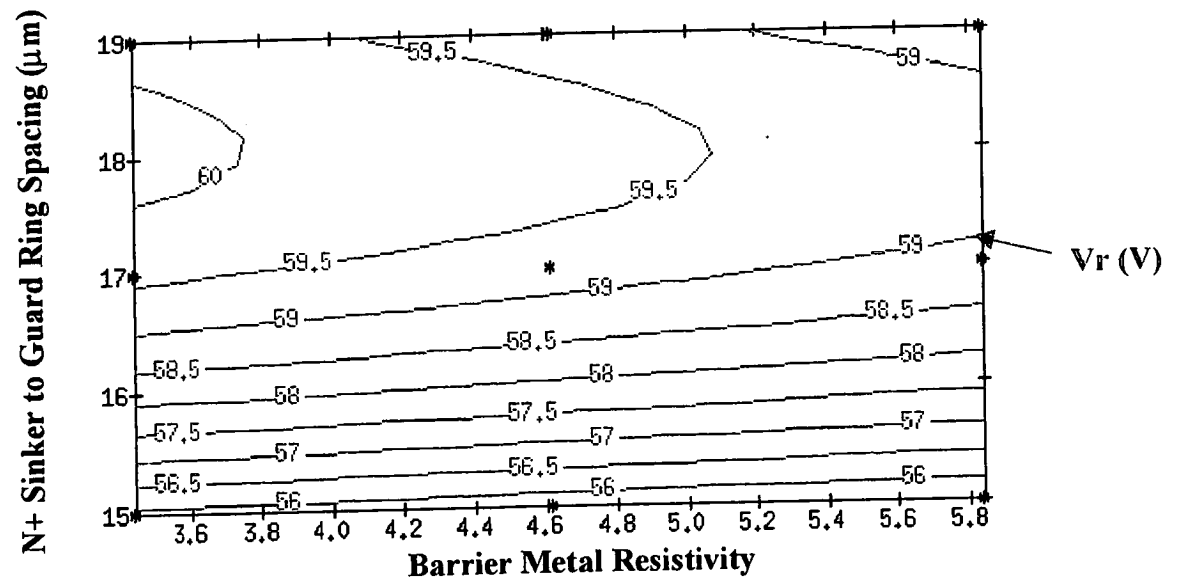


Figure 10: RSM plot showing the effect of n+ sinker to guard ring spacing and barrier metal resistivity on reverse breakdown voltage.

In the former, the data again shows that perceived fab sensitivities are in fact robust for the process in question, as epi resistivity over the range of interest has little impact on device parametrics. The response surface also illustrates the different behaviours of low- (relatively sensitive) and high-current (relatively insensitive) Vs for a particular device. In the analysis of Vr sensitivities in Figure 10, RSM methodologies show very clearly the onset of plateau regions which can be used to determine optimum operating points that will give a stable reverse breakdown over a range of processing conditions. For the device shown here, is can be seen that a sinker to guard ring spacing of 18μm will result in the smallest Vr distribution over the range of operating conditions likely to be encountered in manufacturing.

4. CONCLUSIONS

The data described in this paper has shown that, by employing a rigorous, planned DOE approach to both design and physical attributes of a process development, it is possible to produce a fully characterised and manufacturable process in a single cycle of silicon with minimal cost. In addition, as the example of Schottky device development shows, the methodology has enough flexibility to enable the investigation of new devices (in this case a selective FTA implanted diode family) with a view to immediate implementation in manufacturing.

5. ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the manufacturing and test facilities at National Semiconductor, Greenock, particularly the contributions of David McAlpine and David Colbourne.

THE ROLE OF STEM AS A HIGH RESOLUTION FAILURE ANALYSIS TOOL FOR SEMICONDUCTOR MANUFACTURING TECHNOLOGIES

Alastair McGibbon, Rikki Boyle, Mark Redford

Analog Process Technology Development, National Semiconductor, Greenock, Scotland

ABSTRACT

In this paper, we demonstrate the way in which techniques available on a scanning transmission electron microscope (STEM), particularly Z-contrast imaging and energy-dispersive x-ray microanalysis, can be applied successfully in the characterisation and failure analysis of super-micron semiconductor manufacturing technologies. Following a general description of the techniques, two separate examples are given: Firstly, the detailed characterisation of a low temperature coefficient of resistance SiCCr thin film where a complex microstructure covering a total thickness of ~100Å is revealed and described. Secondly, we describe the way in which STEM was used to detect and observe nanometre-sized PtSi spiking in doped epilayers - the root cause of an NMOS sub-threshold leakage issue.

Keywords: STEM, Z-Contrast, EDX microanalysis, SiCCr thin films, Sub-threshold leakage.

1. STEM AND RELATED MICROANALYTICAL TECHNIQUES

High spatial resolution electron microscopy imaging techniques (TEM) have traditionally been the domain of deep sub-micron technologies. In the cases where these techniques are generally employed, it is primarily crystallographic and/or topological information that is retrieved. However, in many cases in failure analysis, covering both sub- and super-micron technologies, it is a combination of both structural and compositional information that is required by the engineer to fully characterise a particular issue. Here, we demonstrate the way in which a powerful analytical microscope related to the TEM, namely a scanning transmission electron microscope (STEM) can be applied successfully in the quantitative characterisation of failure analysis issues in a manufacturing environment.

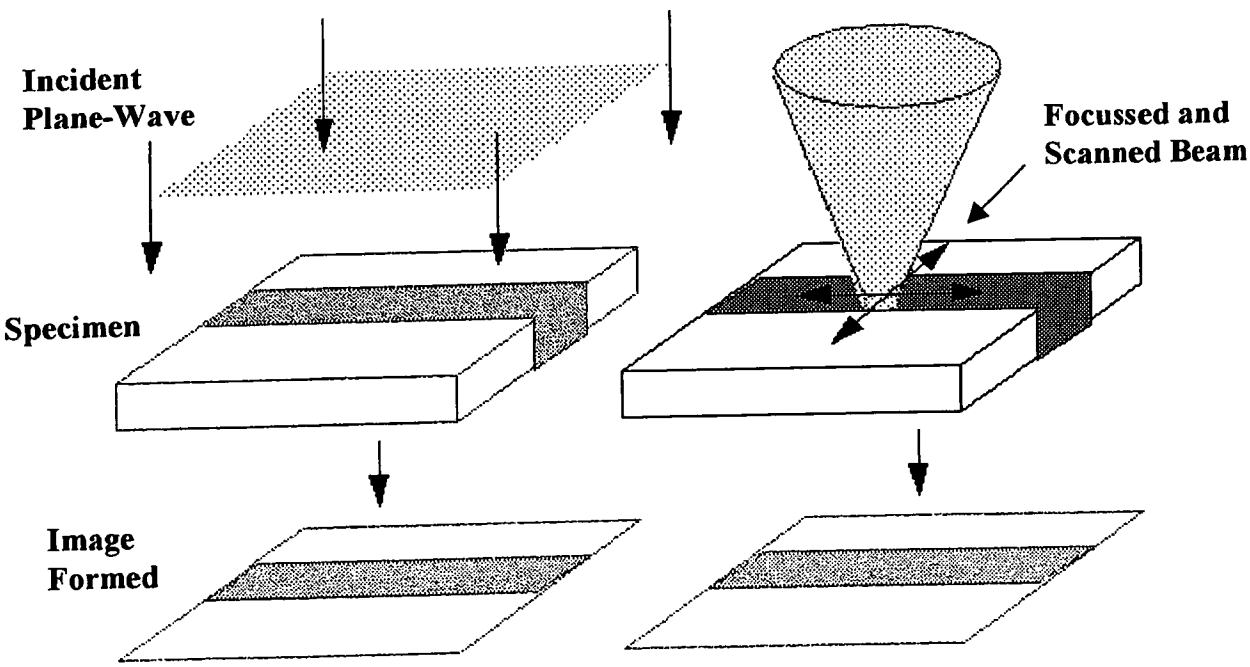


Figure 1: Schematic comparison of image formation between transmission electron microscopy (left) and scanning transmission electron microscopy (right).

STEM is essentially a hybrid of scanning electron microscopy (SEM; low resolution surface imaging and microanalysis) and TEM: As figure 1 shows, unlike conventional TEM, in which an image of a specimen is formed from a plane-wave of high energy (typically 100-400keV) electrons, STEM images are formed by focussing a high current density beam on to the specimen and scanning the beam in a raster pattern to form an image which is equivalent to that seen in a TEM. However, the user has control over the positioning of the beam (which condenses 0.2nA of 100keV electrons into an area of ~5Å in diameter at the specimen surface) which in turn enables the analysis of single points, series of points or line profiles. A number of analytical techniques can be used to investigate the small volume of specimen irradiated by the electron beam. The most suitable and readily interpretable for manufacturing failure analysis issues (shown schematically in figure 2) being energy dispersive x-ray (EDX) microanalysis and annular dark-field (often referred to as Z-contrast) imaging.

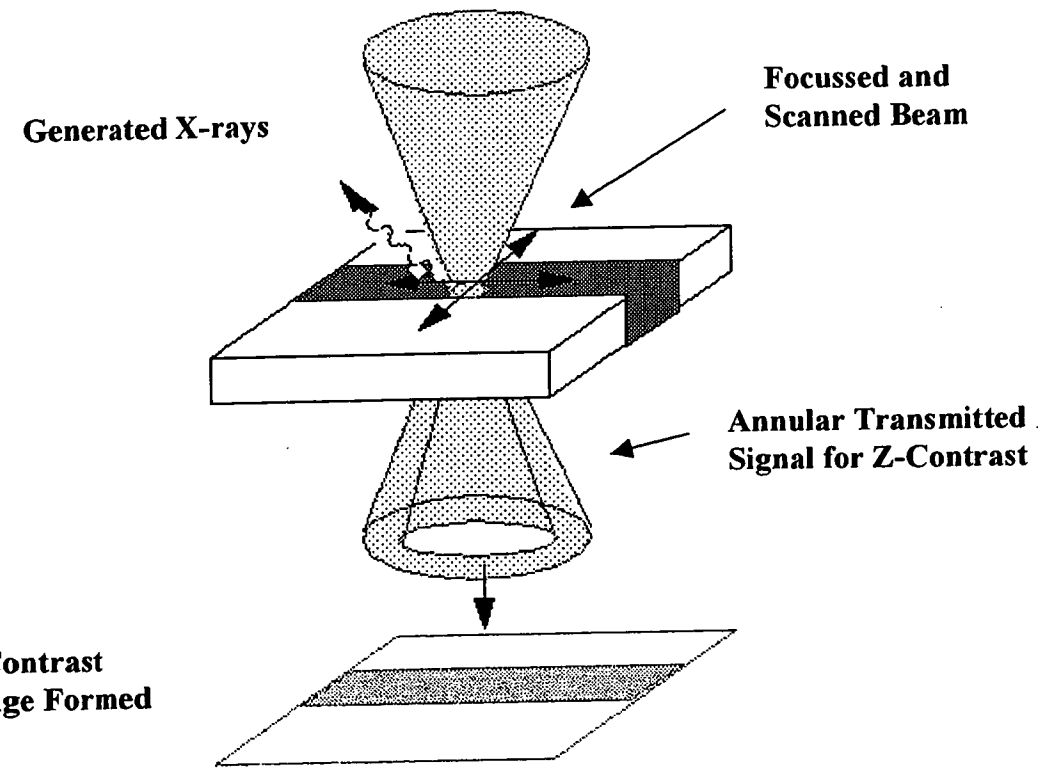


Figure 2: Schematic representation of energy-dispersive x-ray microanalysis and annular dark-field imaging (Z-contrast) in a scanning transmission electron microscope.

EDX microanalysis is commonly employed in SEM to provide quantitative information on elemental composition. This is achieved by detecting x-rays characteristic to a particular atomic species, which are generated by the interaction of the incident electron beam with the specimen. In a STEM, the technique is a very powerful tool, as it can provide the same quantitative information as in SEM, but from volumes of material that can contain only a few hundred atoms.

Z-contrast imaging[1] is a technique that is unique to STEM in which images of the specimen can be formed by collecting only the transmitted electrons that have been scattered through high angles on interaction with the specimen. In a manner similar to Rutherford Scattering, this results (provided major crystallographic orientations are avoided) in a signal with an intensity related to the mean atomic number (Z) of the area irradiated by the electron beam. Unlike TEM images (whose equivalent can also be formed in STEM) which are formed by complex interactions of electron waveforms which are in turn counter-intuitive and very often open to misinterpretation, Z-contrast is a directly interpretable imaging technique. The resolution of the Z-contrast image is limited by the electron probe size, and so by combining this data with EDX information, it becomes possible to probe elemental composition on a sub-nanometer level.

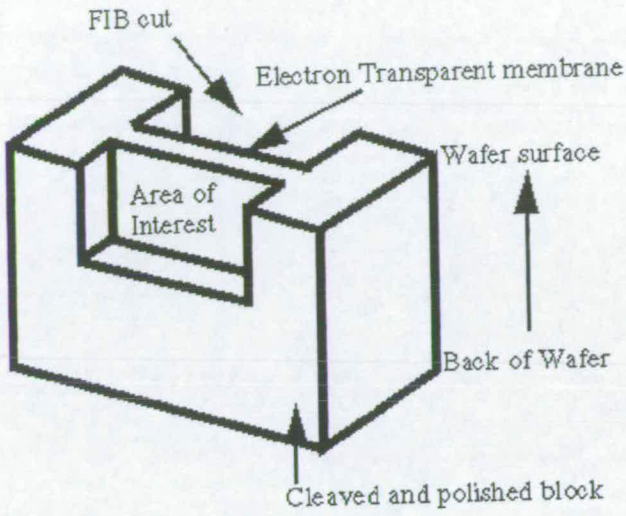


Figure 3: Schematic representation of STEM-ready cross-sectional specimen prepared using focussed ion beam technologies.

The level of specimen preparation required for STEM is the same as for TEM - the region of interest has to be mechanically thinned to form a thin (typically 100 - 1000Å) film to be 'transparent' to the incoming electron beam.

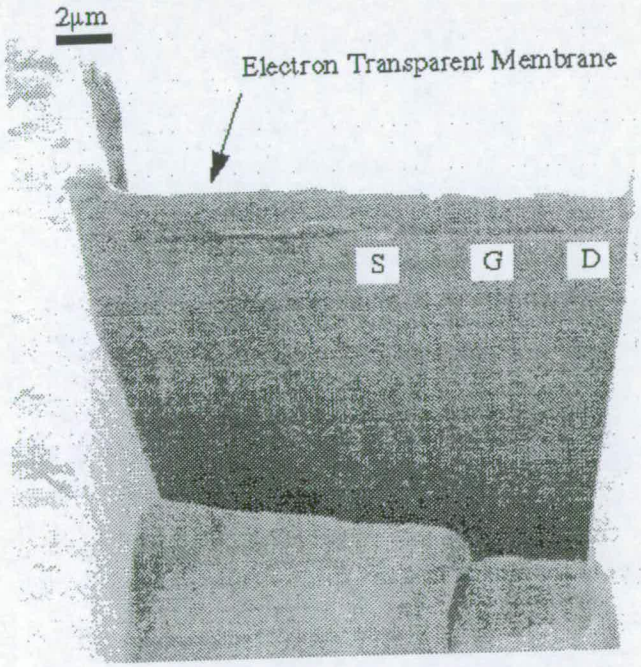


Figure 4: Focussed ion beam (FIB) image of a cross-sectional specimen of an NMOS device prepared for scanning transmission electron microscopy using the FIB technique.

This requirement has traditionally been a limiting factor in the ease with which TEM techniques can be applied in a manufacturing environment with regards both to cycle time and area selectivity. However, recent developments in TEM/STEM specimen preparation have overcome many of the drawbacks by enabling the preparation of thin cross-sectional films using focussed ion beam (FIB) facilities. This approach provides accurate selectivity of regions of interest

and precise control of specimen thickness (not easily achievable using conventional methods). As shown schematically in figure 3, the technique involves the creation by cleaving and mechanical polishing of a small block of material (~1cm wide) containing the region of interest. The FIB is then used to etch out trenches on either side of the region of interest to form an electron transparent film, which is supported by the block edges. The advantage of this technique over conventional TEM specimen preparation is that it is considerably more robust and area-selective. A FIB image of a completed NMOS cross-section (discussed in section 3) is shown in figure 4.

2. CASE STUDY: SiCCr THIN FILM CHARACTERISATION

SiCCr thin films can be used to form very low TCR (temperature coefficient of resistance) resistors with extremely accurate resistor-to-resistor matching properties. As such they play a key role in a number of analog bipolar product applications. However, to ensure long-term manufacturability of these films, and to enable future development based on an understanding of film properties as a function of growth conditions, it is necessary to employ a characterisation technique that will provide a thorough, repeatable fingerprint of the thin film structure.
Conventional TEM analyses of the SiCCr thin film indicated that the microstructure of the film (~100Å thick) was complex, without being able to provide information beyond that. Furthermore, the majority of commonly used techniques such as SIMs, Auger and FTIR are unable to provide 2-dimensional structural and compositional information of a cross-section at the spatial resolution required to analyse the SiCCr thin films. Consequently, analyses of SiCCr thin film cross-sections were carried out on a Vacuum Generators VG HB501 STEM, using both Z-contrast imaging and EDX microanalysis to probe film structure and composition in two dimensions.
The first observation made in the STEM was that there was no degree of crystallinity (single-or poly-) in the SiCCr layers, i.e. that it is either purely amorphous or that it possesses long-range structural order similar to that seen in glasses. Z-contrast imaging (an example of which is given in figure in figure 5a), indicates that there are distinct

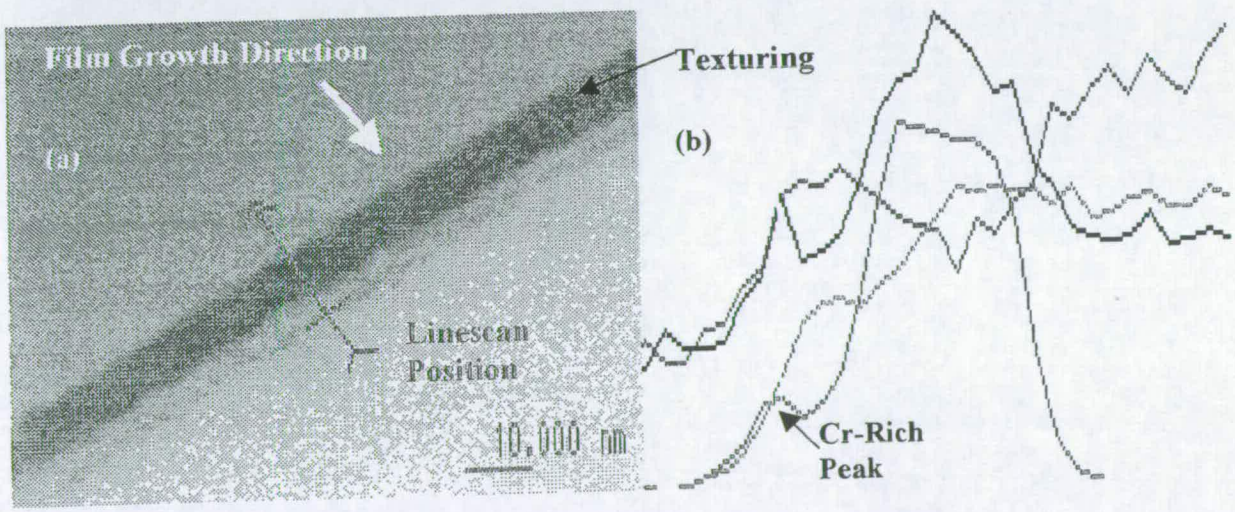


Figure 5: STEM Z-contrast image of a SiCCr thin film (a) showing the location of an EDX linescan used to detect the elemental segregation (b) of Si (green), O (blue), and Cr (red). Integrated intensity is given in black.

variations in elemental composition across the SiCCr region, both laterally along the principal layer, where a form of columnar texturing is observed, and perpendicular to this along the growth direction, where there is evidence for a principal layer (~80Å thick) capped by two smaller layers (each ~15Å thick) the upper of which appears to possess a higher mean atomic number than the lower. To investigate this further, the spatial sensitivity of the STEM was utilised by placing the electron probe at different points on the specimen, and collecting EDX spectra at each point. In addition, EDX line scans were obtained along the film growth direction. A schematic diagram showing where typical spectra were acquired is given in figure 6, with the compositional mapping of the principal elemental species (Si, Cr and O) detectable using EDX across a linescan given in figure 5b.

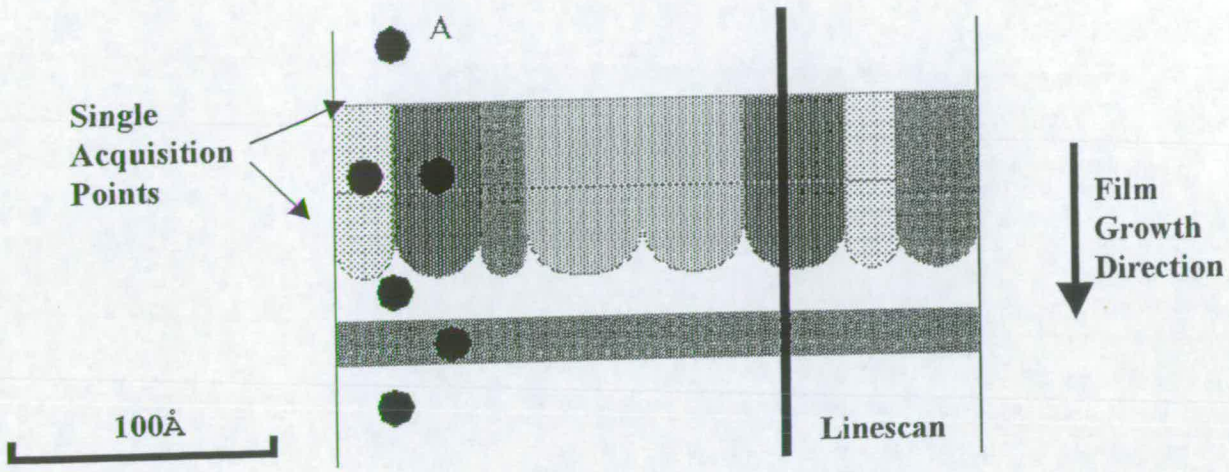


Figure 6. Schematic representation showing typical positions for acquisition of EDX spectra from areas of the SiCCr thin film.

By combining both the Z-contrast and EDX information, it can be concluded that the primary layer observed in the film is primarily SiCCr, but that this is capped by a thin oxide film which is in turn covered by a thin Cr-rich layer. This complex structure, readily observed in STEM and shown schematically in figure 7,

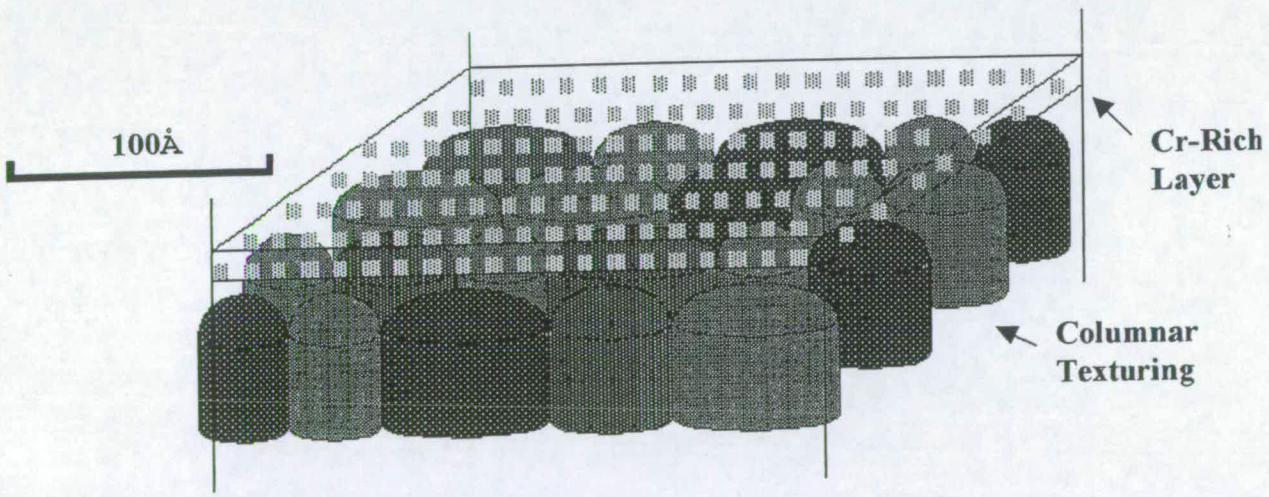


Figure 7. Schematic representation of the SiCCr thin film microstructure based on analysis in the scanning transmission electron microscope.

was found to be repeatable across a number of observed samples using identical growth conditions, and as such can be used as a long-term manufacturability benchmark for the thin film deposition process.

3. CASE STUDY: PtSi SPIKING IN CONTACT STRUCTURES

The discussions in this section centre around the failure analysis of super-micron NMOS devices possessing a platinum silicide contact scheme. It was found that the wider the device, the more likely it was to exhibit subthreshold leakage, which in some cases was as high as 1mA. Detailed probing revealed intermittently poor diode properties (in some cases

direct shorts) across the n+ (As)/ p- source and drain to substrate junctions. However, although device probing indicated that the failure mechanism may be defect related, physical evidence of this was limited, despite the application of several conventional failure analysis techniques such as photoemission, SEM, conventional TEM and preferential etching.

High resolution STEM analysis of FIB-prepared cross-sections revealed the existence of spiking at the device contacts (Figures 8). In these images, spikes of (typically) 50nm in length can be observed emanating from the contact interface with the epilayer. The compositionally sensitive nature of the Z-contrast images suggests that the

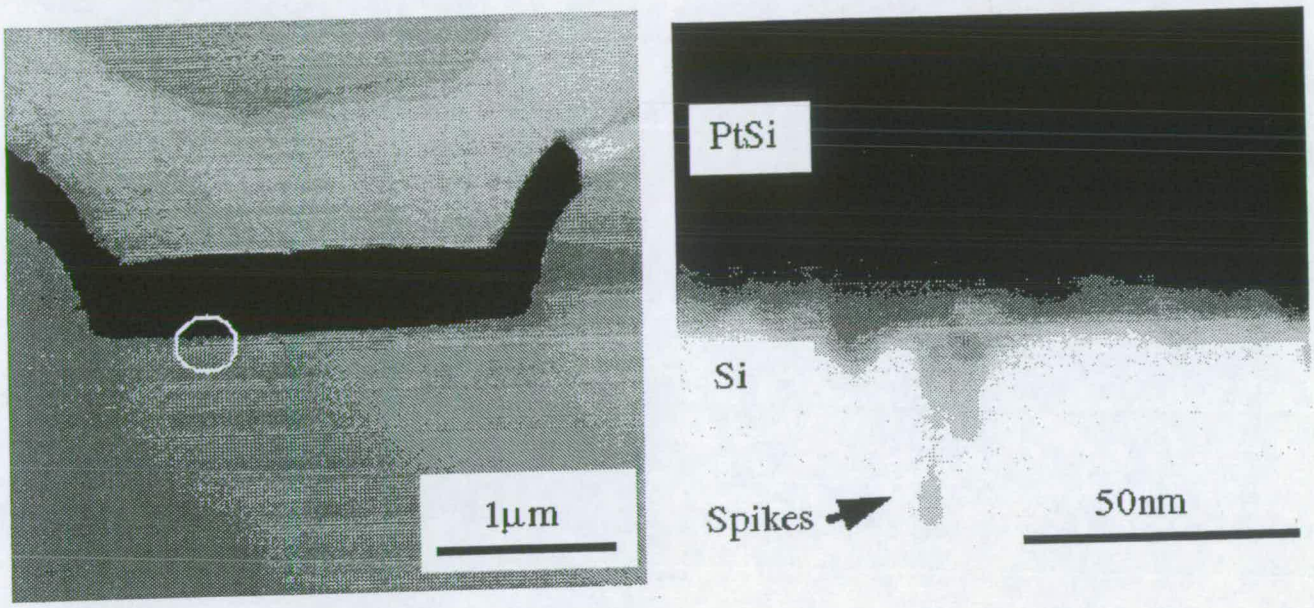


Figure 8. Z-contrast images in a scanning transmission electron microscope showing evidence of spikes at the interface between contact and silicon in a cross-sectional specimen.

composition of the spikes is similar to that of the contact. This is further confirmed by the evidence shown in Figure 9, where STEM analysis of the contact structures was also carried out on a conventional TEM plan-view specimen. The studies again revealed spikes, in this case emanating outwards from the side of the contacts. When taken together with the cross-sectional specimen, this provides valuable information on the general orientation of the spikes.

As discussed previously in this paper, Z-contrast evidence can be pursued further in STEM by acquiring energy-dispersive x-ray (EDX) spectra at specific points along the spikes with a spatial accuracy of ~10Å to determine local compositions. A schematic diagram showing the points at which a typical run of EDX spectra was acquired is shown in Figure 10b. The results from these analyses confirmed that the spikes contain platinum at all points.

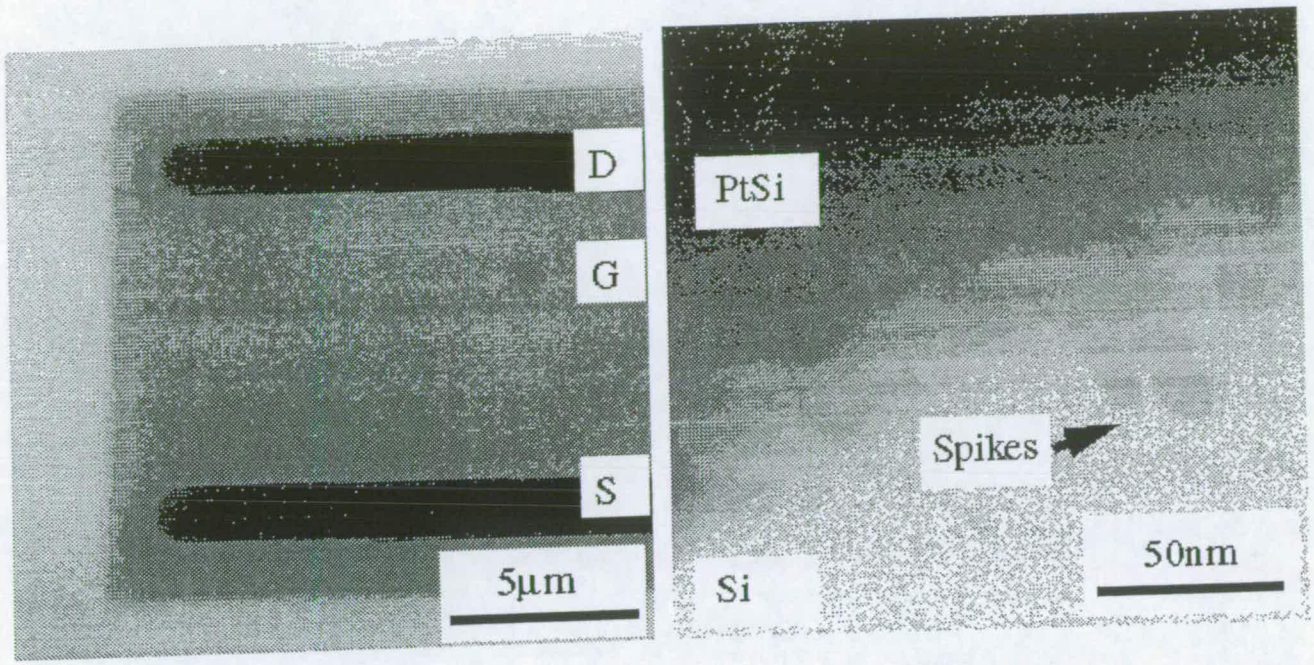


Figure 8. Z-contrast images in a scanning transmission electron microscope showing evidence of spikes at the interface between contact and silicon in a plan-view specimen.

A combination of STEM imaging and microanalysis has revealed the presence of PtSi spikes at the contact/epilayer interface in failing NMOS devices that were not readily observed using other analytical techniques. Although previously studied at silicide/As-doped polysilicon interfaces using TEM[2] observations of Pt spiking at bulk silicon interfaces are rare due to the generally small physical scale of the problem. In general, it can be seen that the interface

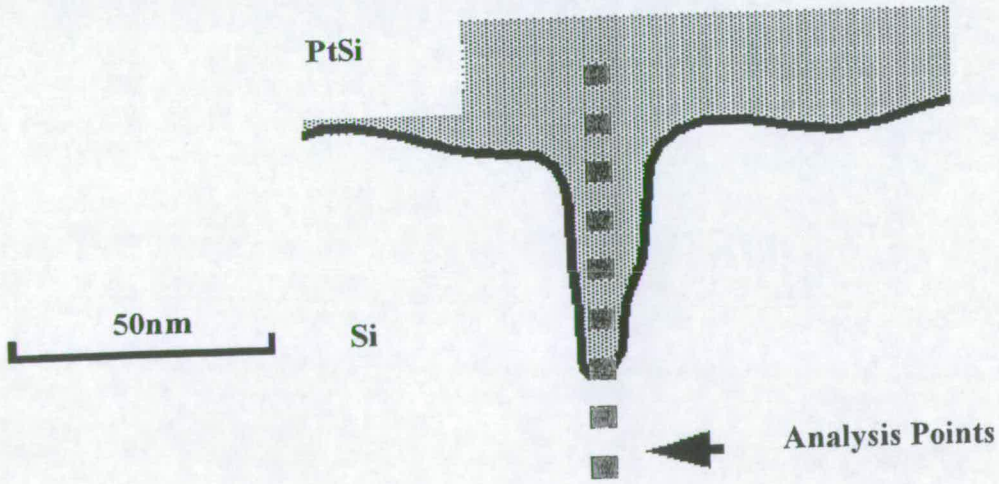


Figure 10. Schematic representation showing typical positions for acquisition of EDX spectra from a Pt spike at the contact/silicon interface.

integrity is poor, with roughness on the 10nm scale. It appears from the images that the interface roughness enables the nucleation of spikes (with a base of typically 10nm) which then penetrate into the silicon. A combination of both the cross-sectional and plan-view information suggests that the preferred direction of spiking is along (111) (i.e. at an angle of 45° into the substrate) and that as many as 10^3-10^4 spikes can exist in a 200μm^2 contact region.

Although the spikes observed were 1/10th of the length required to completely short out the N+ Source drain region in devices studied here, the probability of picking out such a defect without the help of a characteristic photoemission signature (based on electrical test analyses) in a 500Å thick cross-section is about 1 in 2000. In addition, the maximum observable length of a spike in TEM is limited by the specimen thickness (50nm) in all cases except that in which the spike runs exactly in the plane of the cross-section. However, from the evidence seen in the STEM, it became clear that Pt spiking was a problem in wafers where excess Pt has been deposited. These studies resulted in tighter control of the Pt deposition process which in turn eradicated the yield issue.

4. SUMMARY

The case studies described in this paper have demonstrated that high spatial resolution imaging and microanalytical techniques available on a STEM can be applied successfully to the characterisation and failure analysis of manufacturing issues related to super micron technologies.

The ease of interpretation of the techniques described, together with improved cycle time using FIB - based specimen preparation techniques that can be carried out at most manufacturing sites - means that high resolution electron microscopy does now have an important and amenable role to play in failure analysis for semiconductor manufacturing.

5. ACKNOWLEDGEMENTS

The authors wish to acknowledge the support and facilities available at the Department of Physics, University of Glasgow, the Department of Materials Science, University of Liverpool and the Failure Analysis Group, National Semiconductor, Greenock.

5. REFERENCES

1. A. J. McGibbon, S. J. Pennycook and J. E. Angelo, "Z-Contrast imaging", Science **269**, pp. 519-524, 1995
2. G. Grivna, J. Kirchgessner and J. Carlson, "PlatinumSilicide Junction Spiking in Arsenic-Doped Polysilicon Transistors" J. Vac. Sci. Tech. **B 10**, pp. 53-59, 1992

RAPID AND COST EFFECTIVE TECHNOLOGY DEVELOPMENT USING TCAD: A CASE STUDY

Alan Shafi†, Jim McGinty, Martin Fallon, Mark Redford

Analog Process Technology Development, National Semiconductor, Greenock, Scotland
†Now with Chartered Semiconductor Manufacturing, Singapore

ABSTRACT

Rapid and cost effective technology development is required in order to meet competitive time to market constraints for modern semiconductor products. TCAD tools are indispensable for meeting both time and cost targets by pinpointing the required development effort and thereby reducing the development work required. In this study, the requirement is to increase the substrate breakdown voltages of both npn and pnp bipolar transistors for a complimentary vertical bipolar process from 120V to 155V (target is 155V, minimum is 130V). A Design Of Experiment (DOE) approach to both device layout and process technology development is used to ensure that both layout and process sensitivities are rigorously considered in designing-in the manufacturability of the technology at the development stage. TCAD tools are used to show that the increased breakdown voltages can be achieved purely by optimising the device layout. A layout DOE is performed to guarantee that there is no marginality relating to the device layout. Monte-carlo analysis then performed to determine the process steps that the breakdown voltage is most sensitive to. A process DOE is then run to ensure that there is no marginality relating to the process.

1. BACKGROUND

The requirement for this project is to increase the substrate breakdown voltage for both npn and pnp polysilicon emitter transistors. For the npn transistor, it is the collector to substrate breakdown voltage that is important, whilst for the pnp transistor, it is the nwell to substrate breakdown voltage that is important. Furthermore this must be achieved whilst ensuring that the existing set of active and passive devices is still available. A schematic cross-section of the npn and pnp devices is shown in figure 1.

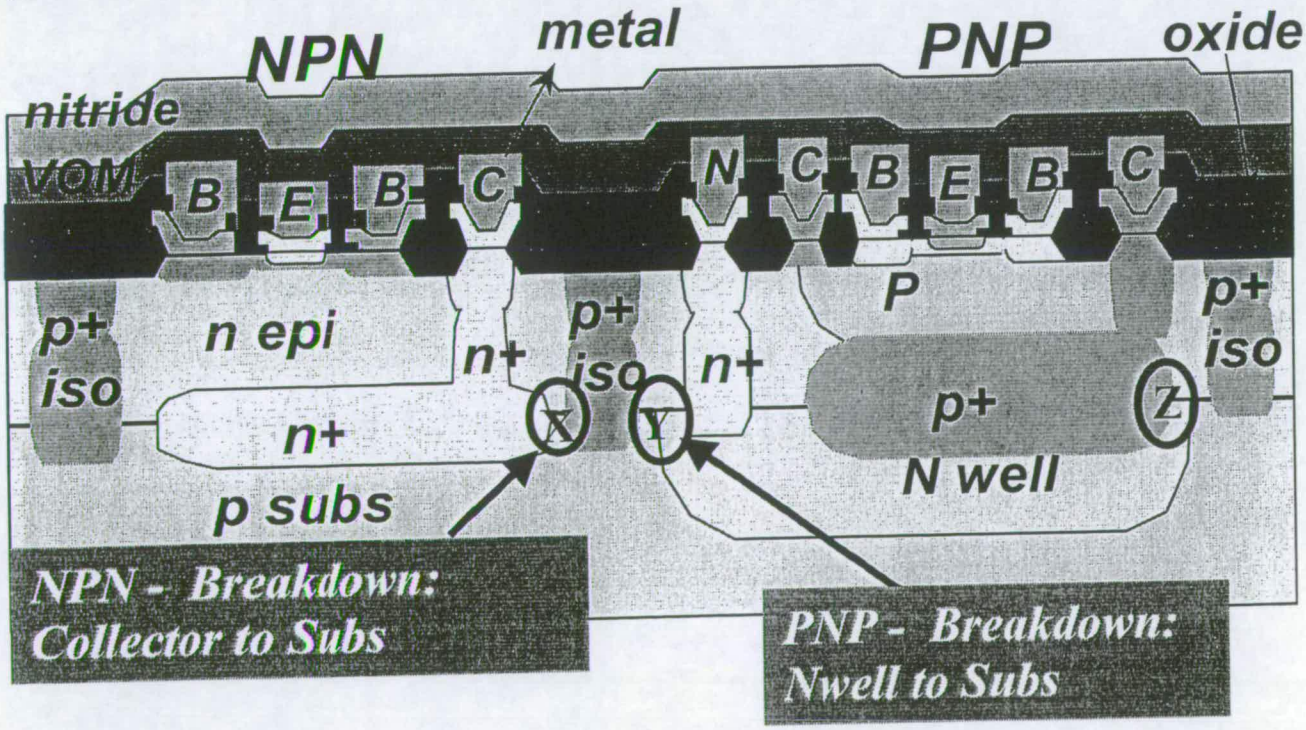


Figure 1. Schematic Cross-section through the npn and pnp devices

By examining the breakdown of existing devices drawn to different design rules, it became clear that there is a layout sensitivity to the breakdown voltage. This is shown in figures 2 and 3 for both npn and pnp devices. In both cases there are two distinct distributions corresponding to devices where the p-isolation ring spacing is changed. It should be noted that the results do not necessarily imply that the target breakdown can be achieved by extrapolation.

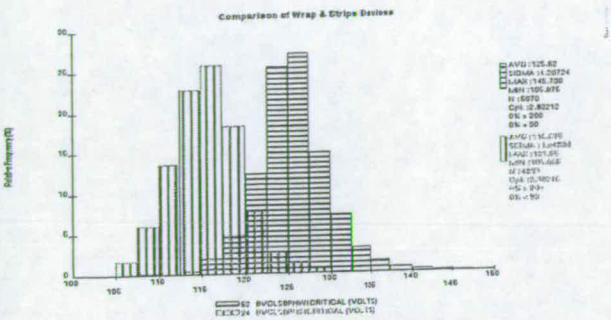


Figure 2. pnp distributions

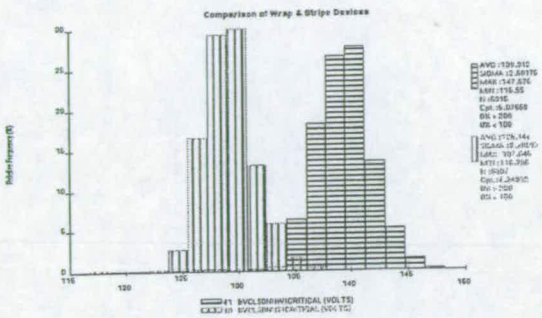


Figure 3. npn distributions

2. INITIAL TCAD EXPOLARATION

TCAD was used to understand the breakdown mechanism and determine whether the new breakdown voltage targets could be met purely by a change in layout. The location of the breakdown is shown in figure 4 for the npn device and in figure 5 for the pnp device. The breakdown of the npn device can easily be increased by increasing the spacing between the p+isolation ring and the n+ buried layer (shown as X in figure1). This is demonstrated in figure 6 which shows the simulated breakdown for the npn device. The required breakdown voltage is met for a p-isolation to n+buried layer spacing of 23 μm. For the pnp device, the target breakdown voltage can also be met as demonstrated by the simulation results in figure 7. Here, the required spacing between the p+isolation ring and the nwell (shown as Y in figure 1) is 19μm. Further TCAD simulations indicate that there is a possibility of punchthrough across the nwell at the edge of the p+buried layer. Some margin may be gained by increasing the n+ring to p buried layer spacing.

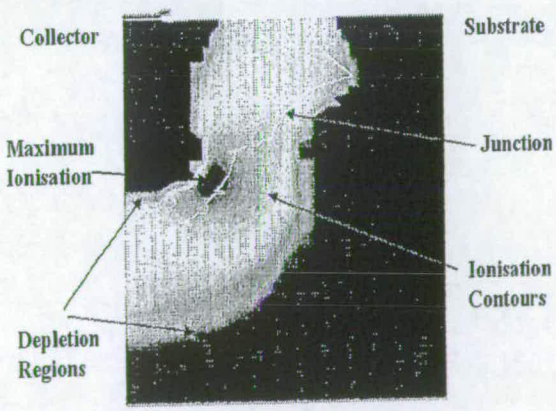


Figure 4. Simulated breakdown location for npn transistors

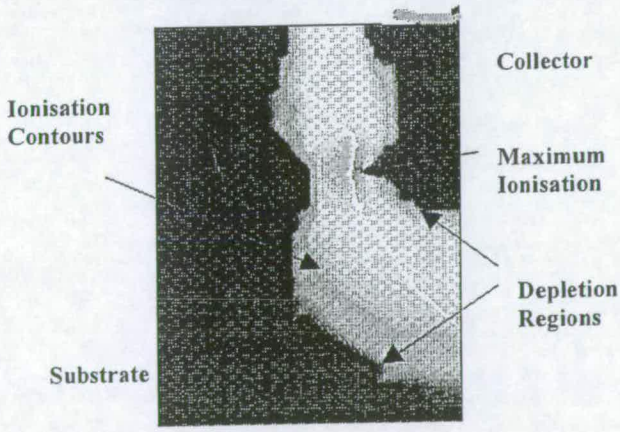


Figure 5. Simulated breakdown location for pnp transistors

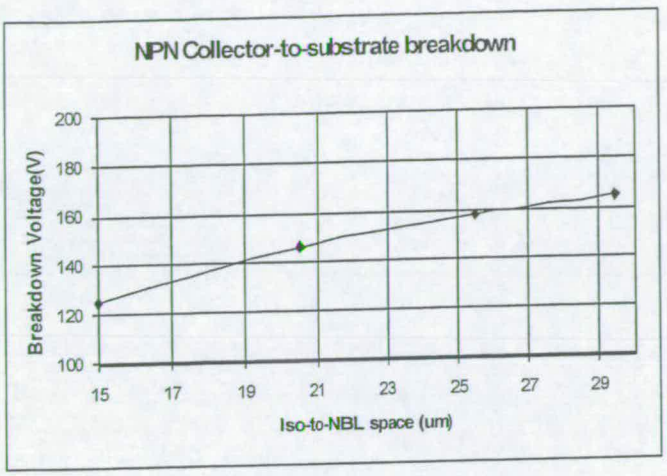


Figure 6. Simulated breakdown for npn transistors

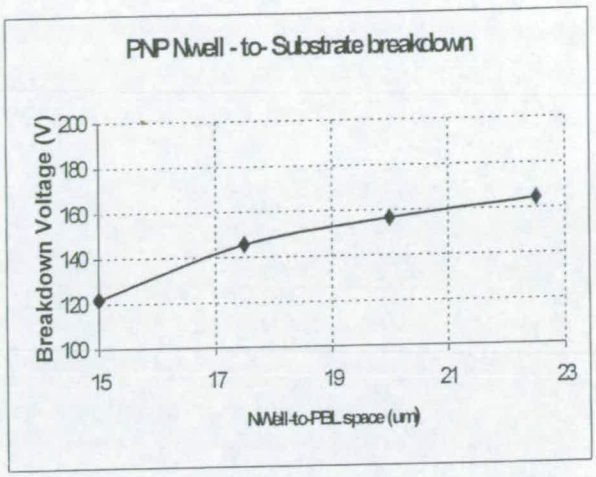


Figure 7. Simulated breakdown for pnp transistors

3. LAYOUT DOE IMPLEMENTATION AND RESULTS.

A layout DOE was implemented on a reticle set. For the case of the npn transistors, there is one variable only. For the case of the pnp transistor, there are two variables. The DOE is summarized in table 1.

Device #	pIso-to-nburied (X)	Device #	pIso-to-Nwell (Y)	Nwell-to-pburied (Z)
1	17.5 µm	4	17.5 µm	23.0 µm
2	20.0 µm	5	17.5 µm	15.0 µm
3	22.5 µm	6	22.5 µm	23.0 µm
		7	22.5 µm	15.0 µm
		8	20.0 µm	20.5 µm

Table 1 – Layout DOE spacings

The results are shown in figure 8 for the npn transistor. Clearly, the required breakdown voltage can be achieved but it is the centre spacing device (#2) that meets the requirements. The results for the pnp transistor are shown in figures 9 (effects plot) and 10 (response surface). Piso-to-nwell (Y) is the only control factor that has an effect. The required breakdown voltage is met again for the centre spacing of the isolation ring.

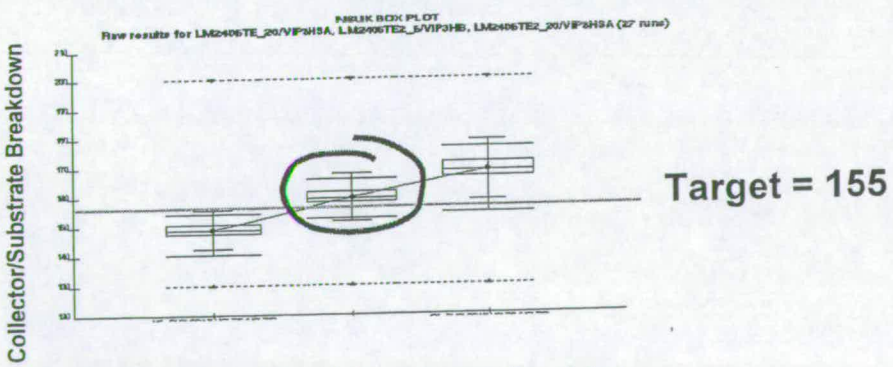


Figure 8. Measured breakdown for npn transistors

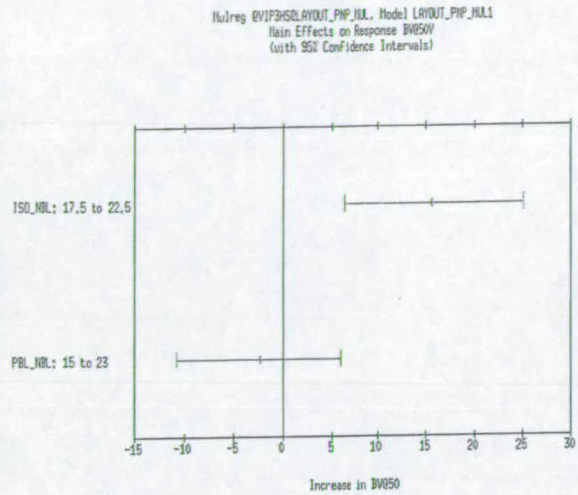


Figure 9. Effects plot for pnp transistors

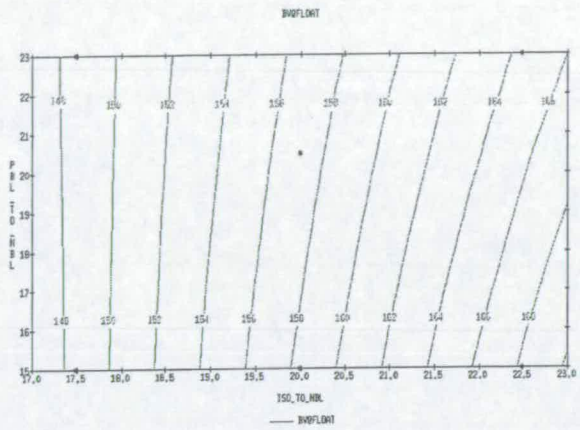


Figure 10. Measured RSM breakdown for pnp transistors

4. PROCESS MANUFACTURABILITY DOE SETUP.

The methodology used to ensure manufacturability of the normal process operating conditions is shown in figure 11. An initial TCAD baseline was established. Process variability was identified for each of the process steps and Monte Carlo sensitivity analysis simulations performed. A multiple regression analysis was then performed on the TCAD results in order to determine the process steps which most influence the responses. It is these process steps that have to be most closely controlled in order to ensure parametric manufacturability of the process. A DOE worksheet was then created and silicon run accordingly. The results were analysed using response surfaces and conclusions drawn regarding the manufacturability of the process.

The results of the TCAD sensitivity analysis are shown in figures 12 and 13 for npn and pnp transistors. Layout related factors have already been considered separately in the layout DOE. The top four process factors were chosen for both npn and pnp transistors and are shown in table 2. Four factor CCF designs with 5 centre points were chosen for both npn and pnp transistors. The ranges in the control factors were chosen to be consistent with realistic ranges in the respective process step.

Top 4 factors for NPN	Top 4 factors for PNP
P Buried Layer Implant Dose	N Sinker Down Dose
P Buried Layer Energy	Epi Thickness
Epi Res	N Sinker Up Energy
N Sinker Down Energy	N Well Dose

Table 2. Control factors for process DOE.

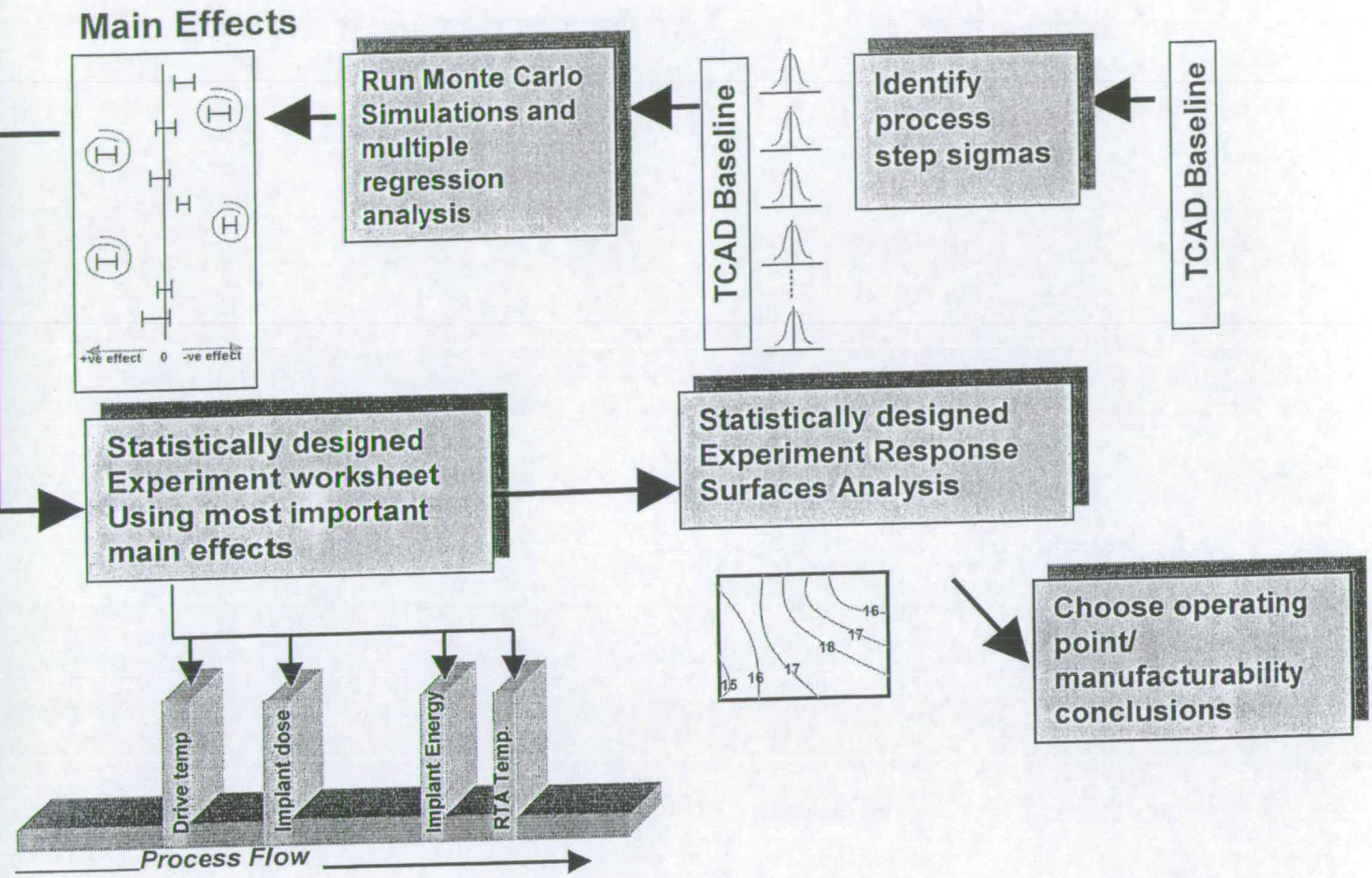


Figure 11. Manufacturability Methodology

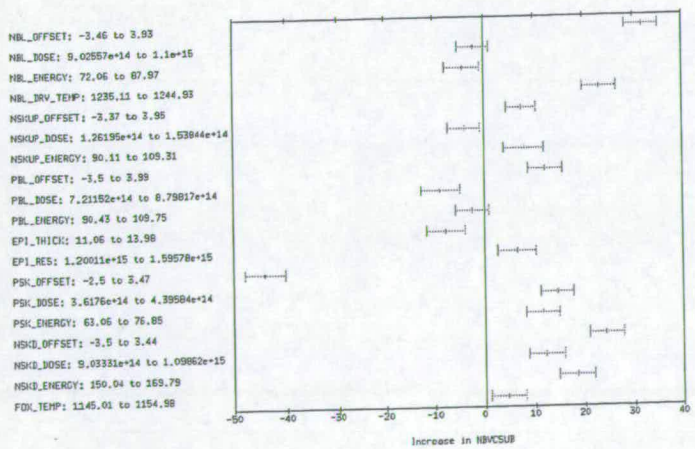


Figure 12. TCAD Sensitivity Analysis for npn transistors

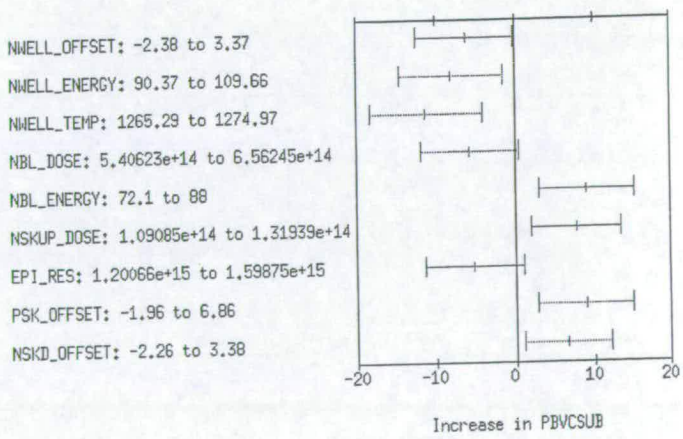


Figure 13. TCAD Sensitivity Analysis for pnp transistors

5. PROCESS MANUFACTURABILITY DOE RESULTS – NPN TRANSISTORS

The centre point wafers which received the standard process show a spread between 152.2 and 160.1 volts which is comparable to the spread in the rest of the experiment - 150.5 to 163.9 volts and is shown in the histogram in figure 14. The highlighted areas are the repeat runs.

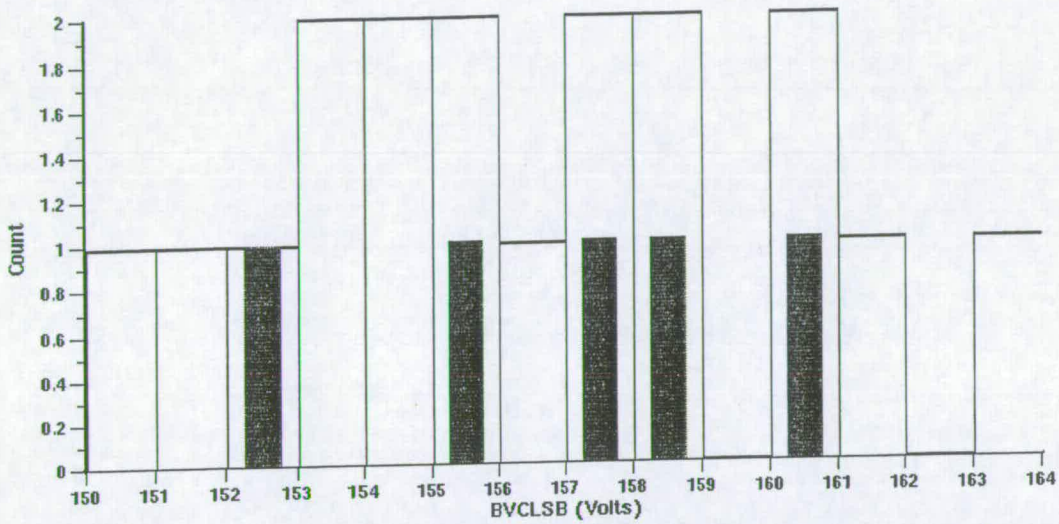


Figure 14. Histogram of npn substrate breakdown results.

There is some evidence of an interaction between the 2 implant energies, such that the effect of changing the N sinker implant energy, when the P Buried layer implant energy is 110keV decreases the breakdown by ~6volts (equivalent to 3 volts reduction from nominal). This is worst case and with ~3volt standard deviation in the breakdown is not seen as a significant problem.

Figure 15 shows the contour plot for the breakdown of the device with increased breakdown versus the 'standard' device (i.e. standard spacing)

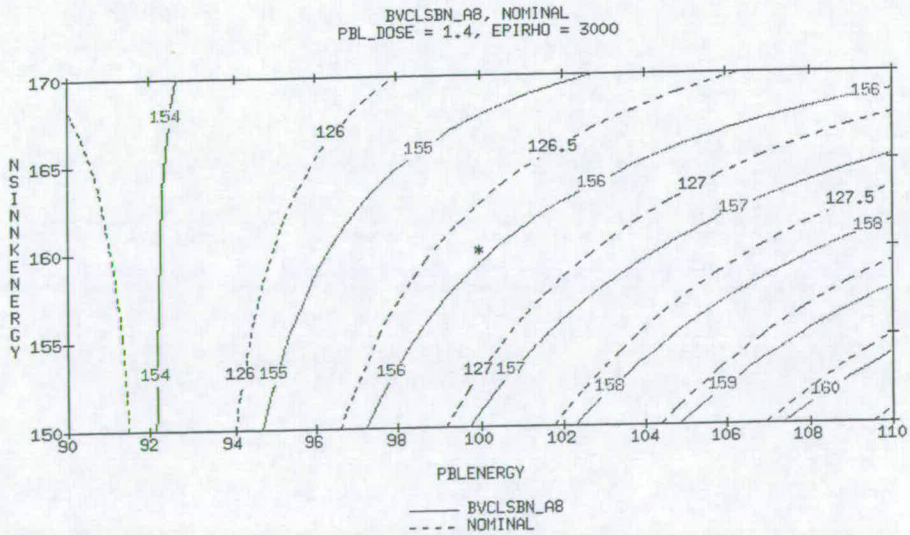


Figure15. Contour plot showing the substrate breakdown voltage versus the standard layout device

6. PROCESS MANUFACTURABILITY DOE RESULTS – PNP TRANSISTORS.

Figure 16 shows a histogram of the results of the process splits with the standard runs highlighted. The distribution is tight and no appreciable difference can be seen. No model could be sensibly fitted to the output data.

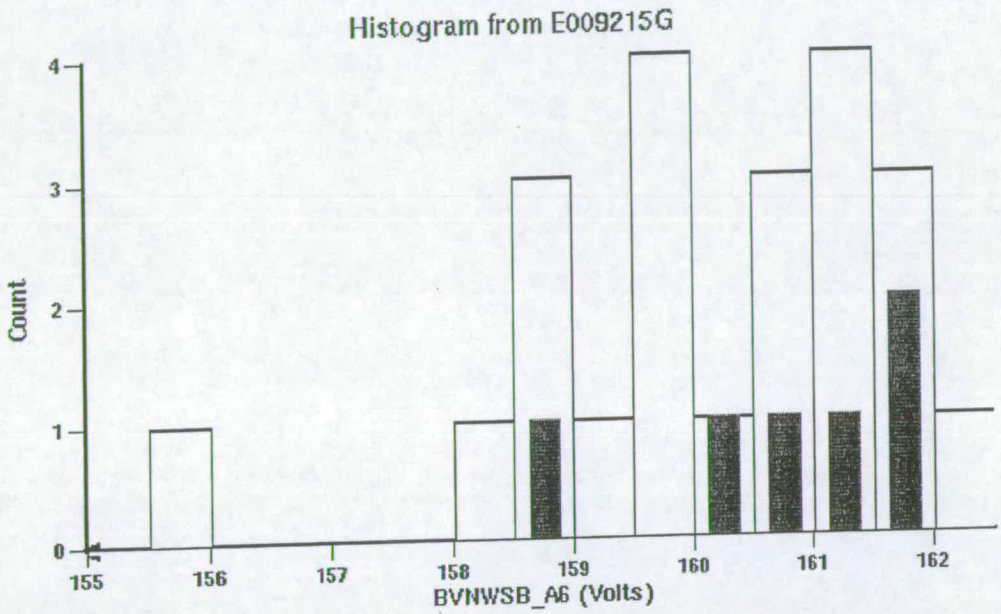


Figure 16. Histogram of pnp substrate breakdown results.

7. PROCESS MANUFACTURABILITY DOE RESULTS SUMMARY.

For the npn devices, although the P Buried Layer energy appears to shift the breakdown, the worst case reduction of ~3volts from nominal is well within the distribution of the breakdown itself and no manufacturing issue is foreseen in running at the increased spacing. The fitted model removed the P buried layer dose since no effect was seen. The process splits were based on previous TCAD results which showed little or no sensitivity of breakdown to process variation. Engineering assessment would predict little variation to breakdown based on these process parameters and so the silicon result confirms the expectation.

For the pnp devices, the process parameters have no impact on the breakdown of the Nwell to substrate for the new high voltage device. Engineering analysis and TCAD predict little or no variation and this was confirmed in silicon.

8.CONCLUSIONS

TCAD has been used to successfully identify a layout change as the means for realizing the increased breakdown requirements. This solution means no extra process flow has to be created in the fab. The required breakdown voltage was achieved in first silicon. Two subsequent DOE lots were run to demonstrate that the technology is fully manufacturable from booth a layout and process perspective.

INVESTIGATION AND APPLICATION OF THE BURIED LAYER VAN DER PAUW AND BAR RESISTORS

Keith Findlater*, Martin Fallon, Mark Redford

Analog Process Technology Development, National Semiconductor, (UK), Ltd., Larkfield Industrial Estate, Greenock, PA16 OEQ, UK

*Now with Edinburgh University, Kings Buildings, Mayfield Road, Edinburgh EH9 3JL.

ABSTRACT

Sheet resistance measurements are often used as process control monitors and are a key part of resistor measurements for SPICE models. These are normally Ohmic in nature. However it can be seen that buried layer resistors are strongly affected by the applied bias conditions in a similar mode to JFET operation. Two standard resistor structures have been studied to investigate this effect: the bar resistor and the van der Pauw. 2D and 3D-device simulation have been used to model the self-modulation and to produce recommendations for optimized routine measurement.

Keywords: pinched base, buried resistor, modulated resistance, van der Pauw, bar resistor.

1. INTRODUCTION

Sheet resistance measurements are one of the most important and widely used process monitoring tools. These measurements are usually performed using the van der Pauw (VDP) resistor test structure[1] or alternatively by four point probe techniques. Sheet resistance is given by:

$$R_s = \frac{1}{\int_W n(z) q \mu(n)} dz \quad \text{.....(1)}$$

where $n(z)$ is the concentration of activated impurities at depth z, q is the electronic charge, $\mu(n)$ is the concentration dependant mobility of the carriers, and W is the width of the layer in the z direction. This ignores any field or temperature mobility dependence.

The information provided by these measurements is extremely useful for the assessment of process variations. During the manufacture of bipolar devices, buried layers such as the collector and particularly intrinsic base are formed. The intrinsic base is a key parameter in determining the transistor gain and is usually a very narrow width structure. Any applied voltage then which modifies the depletion regions of the device alters the width of the base and so changes the resistance. The resistance then is no longer single valued and must be specified along with bias conditions. Further, this may be layout dependent also. Two common layouts used for electrical measurement of the resulting sheet resistance for these layers are the van der Pauw (VDP) or bar resistor test structures. The van der Pauw structure is theoretically superior in this case to the bar resistor due to the elimination of contact resistance errors when using the VDP.

The present work shows that for a junction defined buried layer resistor (figure 1) that this is not the case and seeks to explain the non-idealities using results from 2D- and 3D-device simulation.

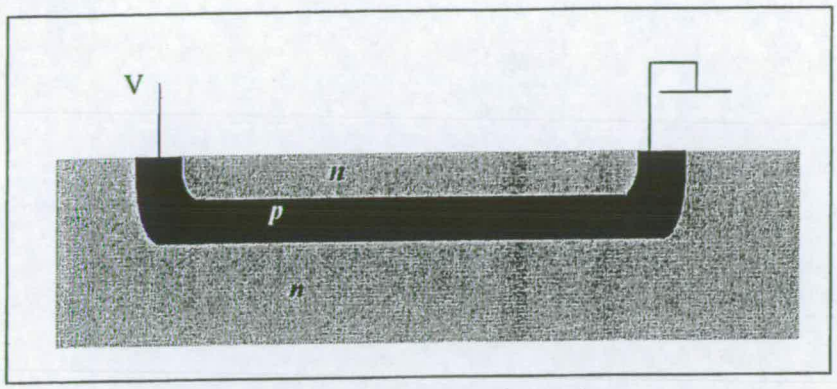


Figure 1. Buried *p*-type bar resistor

2. TEST STRUCTURES

Schematics of the two test structures investigated in this work are shown below (figures 2 and 3).

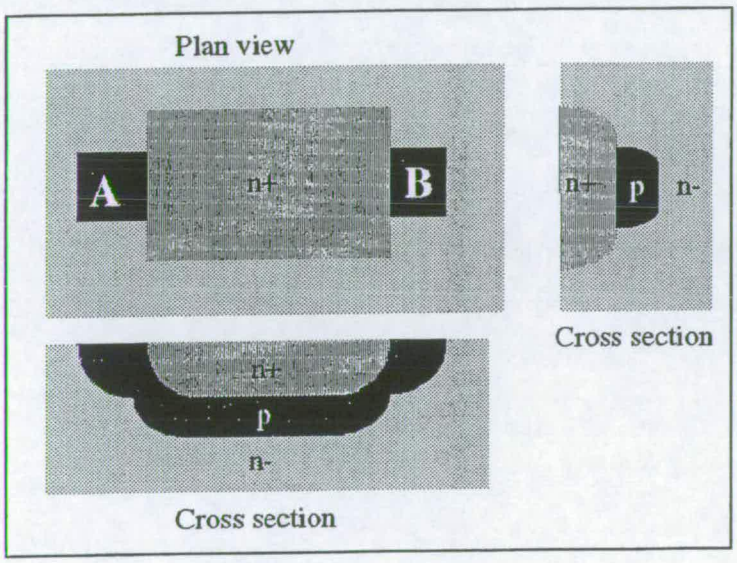


Figure 2. The Bar Resistor

Figure 2 shows the bar resistor in plan view and cross-section. Current is passed from A to B and voltage measured at these points also. The *n*+ layer extends laterally across the buried p layer, with contacts at these ends.

Figure 3 shows the van der Pauw resistor in plan and cross-section. Here the *n*+ region extends on all sides beyond the *p* region with only contact layers of *p* type material extending at the corners.

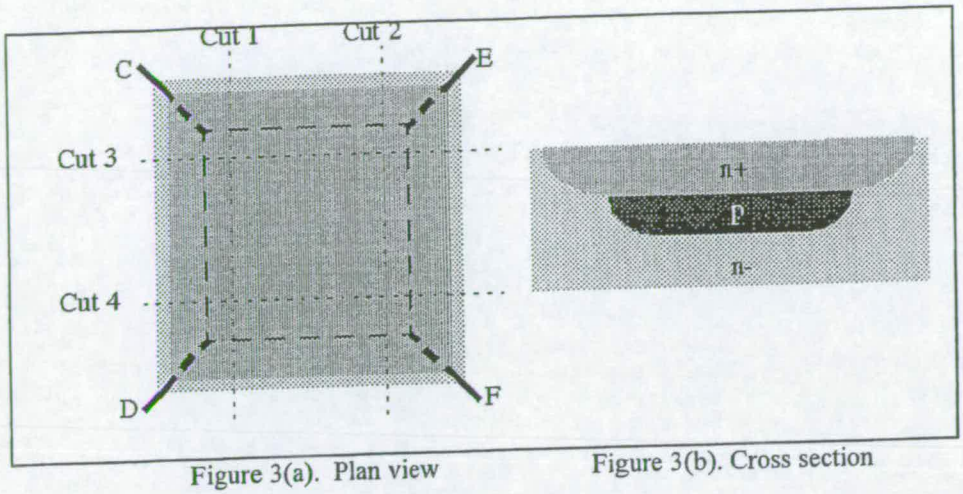


Figure 3(a). Plan view

Figure 3(b). Cross section

In the bar resistor, the sheet resistance R_S (Ω sq.$^{-1}$) is calculated as:

$$R_s = \frac{V_{AB}}{NI_{AB}} \quad .....(2)$$

where V_{AB} is the potential difference between A and B (the ends of the resistor), I_{AB} is the current through the resistor, and N is the number of squares. Note that this technique introduces an error due to the additional extrinsic base regions at the ends of the resistor and the contact resistance, which cannot easily be eliminated.

In the van der Pauw resistor, the sheet resistance is given by the van der Pauw formula:

$$R_s = \frac{\pi}{\ln 2}\frac{V_{DELTA}}{I_{FORCED}} .....(3)$$

where V_{DELTA} is the potential difference between E and F, and I_{FORCED} is the current forced into D and out of C. In this work the VDP structure was slightly modified by chamfering the corners to produce a regular octagon.

3. MEASUREMENTS

Using the HP4156 Semiconductor Parametric Analyser, measurements were taken on fabricated structures. The bar resistor was 10 squares long while the van der Pauw was 160 microns in across the octagon. Figure 4 shows the variation in R_S as the forced current is increased for the van der Pauw structure.

It can be seen that the resistance characteristic is certainly not flat and neither is it linear with current. Additionally, at low forward currents, the resistance exhibits a sharp decrease when the applied forward voltage exceeds the ~600mV diode turn on voltage. In the reverse direction, a breakdown characteristic is observed in figure 5.

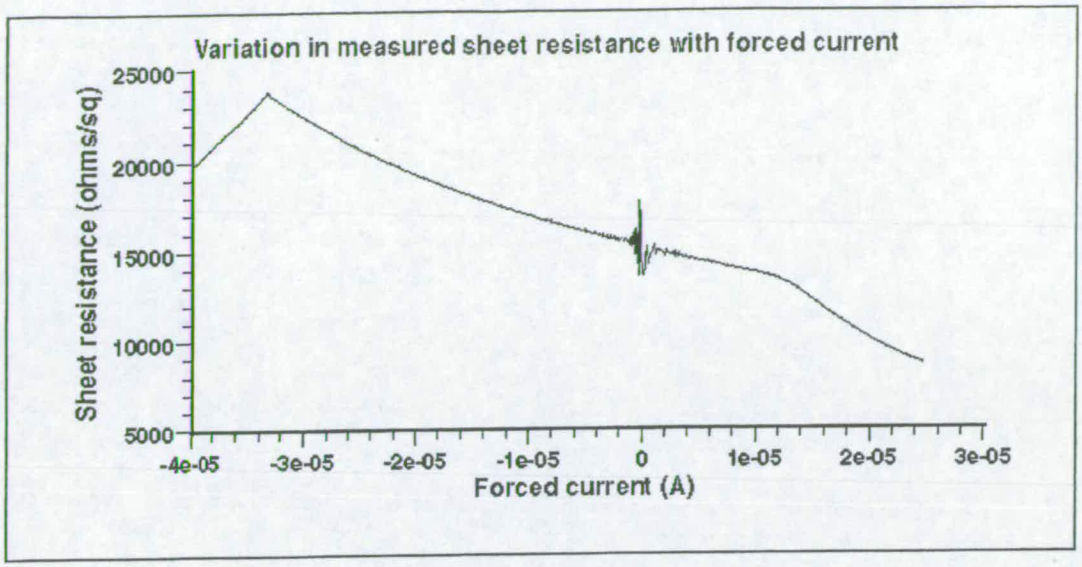


Figure 4. Resistance Measurement from van der Pauw Structure

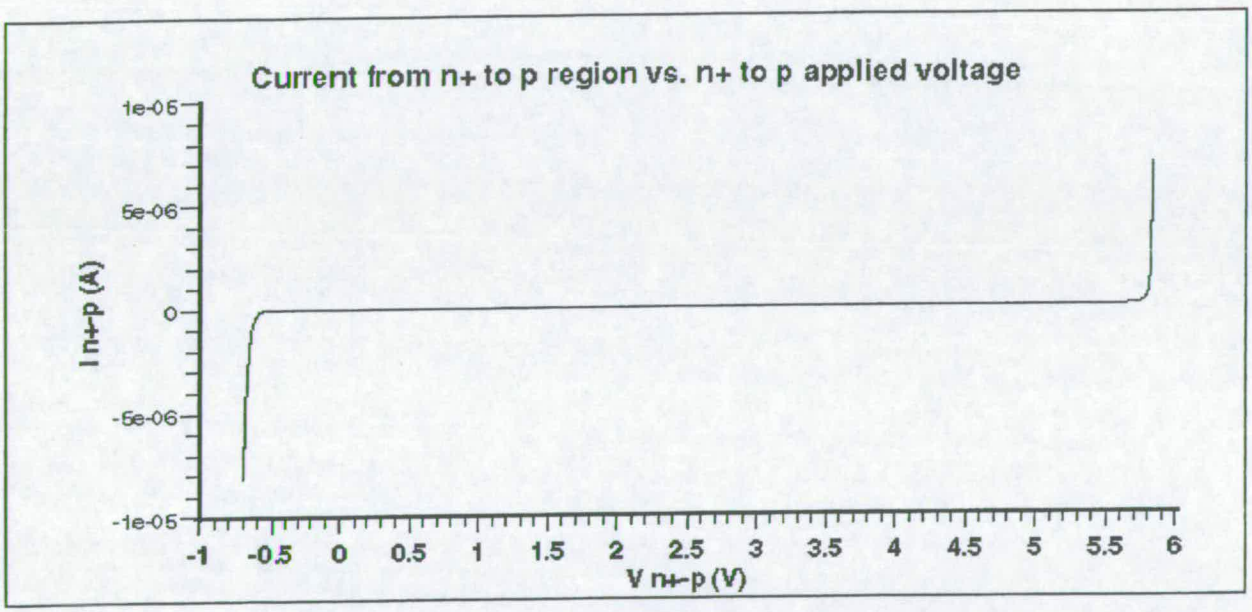


Figure 5. Current results from van der Pauw structure

4. SIMULATION

2D and 3D device simulations were carried out to investigate the non-linear part of the resistance characteristic above. The bar resistor was modelled as a 2D structure using MEDICI[2] and the van der Pauw modelled as a 3D structure using DAVINCI[3]. To investigate the behaviour under different dopant profile conditions, the *n*+ layer was scaled, increasing the depth of the junction and therefore reducing the width of the buried layer. This provided a quick route to altering the critical device dimension without having to re-simulate the entire structure. The simulated sheet resistance was extracted using equations 2 and 3 above. Figures 6 and 7 show the simulated resistance for the two structures.

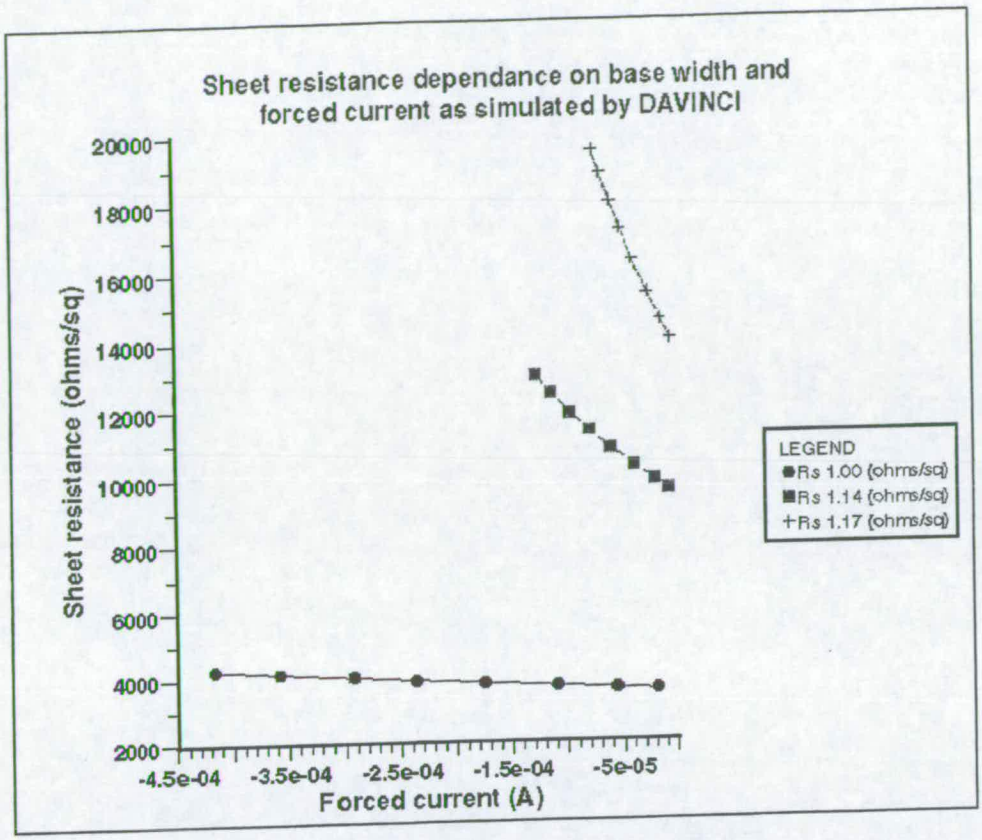


Figure 6. Simulated sheet resistance variation with current for VDP Structure with $n+$ profiles scaled by 1, 1.14 and 1.17

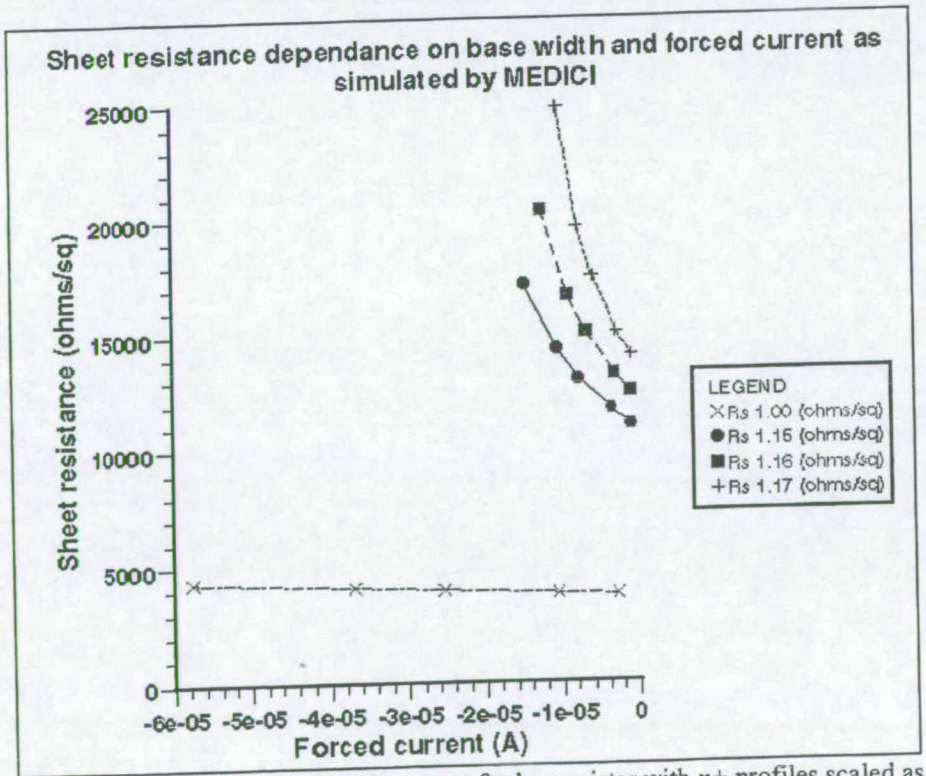


Figure 7. Sheet resistance variation with current for bar resistor with $n+$ profiles scaled as shown.

It can be seen that as the conducting layer gets thinner, the non-linear behaviour of both structures increases. The increase in resistance with reverse bias suggests a narrowing of the conducting channel. This narrowing can only be due to the expansion of the depletion region. Figures 8 and 9 show the narrowing of the conducting layer for the bar resistor as the negative bias increases. A similar effect was observed for the van der Pauw structure.

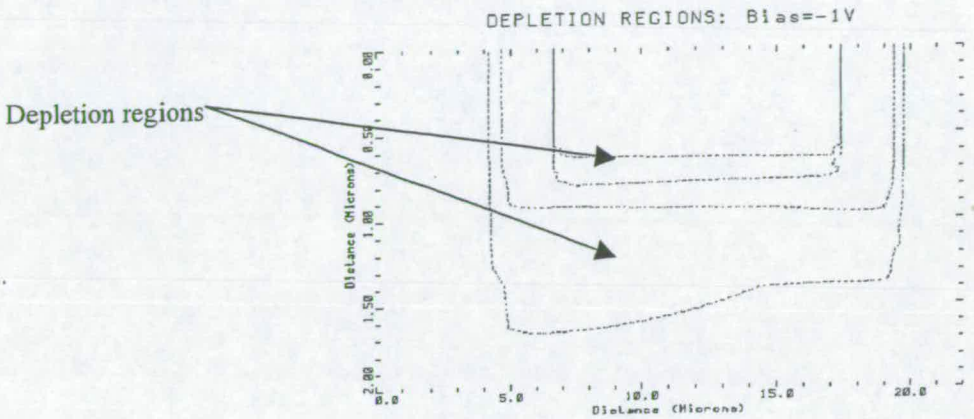


Figure 8. Bar resistor under 1 volt bias showing depletion regions.

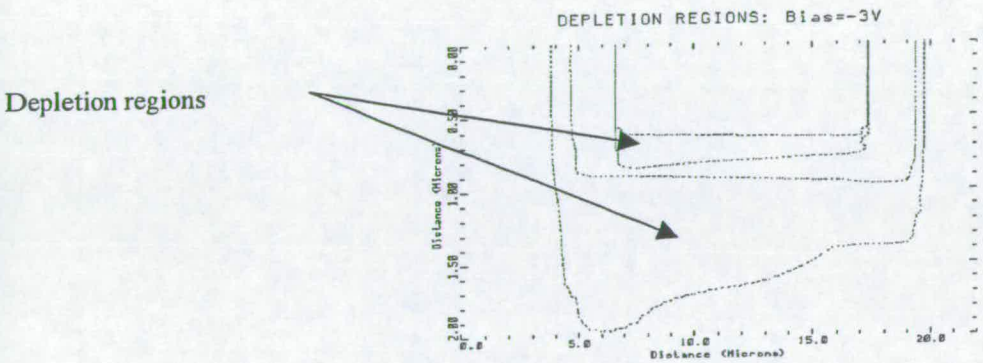


Figure 9. Bar resistor under 3 volt bias showing extended depletion regions and pinched channel.

5. DISCUSSION

Resistors are normally measured by finding the Ohmic regime between low level noise and Joule heating. In the case of buried resistors this approach cannot be applied. The limit is defined by the self-modulating depletion regions. If the total resistance is minimized, the forced current will be larger for a given applied bias resulting in a larger VDELTA and less non-linearity in the current-sheet resistance measurements. Usually the resistance structure (VDP) is drawn large to enable a large current to be passed to produce a larger measured voltage. However, in the case of the buried layer resistor, a large structure will develop a large voltage across the terminals, and enhance the self-modulation. Therefore reducing the size of the van der Pauw is the better approach. The bar resistor, while being a simpler device to measure is just as non-linear as the VDP. IN fact to minimize the error due to the contact resistance and extrinsic base regions it is necessary to maximize the resistance of the intrinsic region which results in increased non-linear behaviour.

Some general guidelines then can be provided based on this work:

1. The structure should have a separate contact to the emitter and substrate regions.
2. The emitter should be drawn outside the base to avoid misalignment problems.
3. The overall resistance of the structure should be minimized.

4. Care should be taken that the measurements do not forward bias the base-emitter junctions, or break down this junction. Additionally the smallest possible current which produces a measurable voltage difference should be used.

6. CONCLUSIONS

The observed non-linearities in the buried layer van der Pauw and bar resistor have been explained as a self-modulating JFET action. Ideally, measurement should take place at 0V bias where the depletion width is at equilibrium, but clearly this is not possible. In any case the forced current should be such that the junctions defining the buried layer are biased below breakdown and with minimal widening of the depletion regions.

For the van der Pauw structure, current crowding effects cause the total resistance of the structure to be much larger than one square of measured sheet resistance. Therefore a smaller structure will accept more current for a smaller bias, and with less modulation of the depletion regions. For the bar resistor, as the number of squares is reduced it will also accept more current, but the errors due to the extrinsic regions and the contacts become more significant relative to the total resistance.

7. REFERENCES

1. 'A Method of Measuring Specific Resistivity and Hall Effects of Discs of Arbitrary Shape,' L. J. van der Pauw, Philips Research Reports, Vol 13, pp1-9, 1958
2. Medici User's Manual, Version 2.1 vol 1, Technology Modeling Associates Inc, 1995.
3. Davinci User's Manual, Version 3.3, vol 1, Technology Modeling Associates Inc, 1997.

DFM Strategies In Technology Development

M. Redford, M. Fallon†, M.I. Newsam†, R.S. Ferguson‡, D. Sprevak‡, J.P. Elliott†, G.A. Allan†, A.J. Walton†

Analog Process Technology Development, Greenock, Scotland,
†Department Of Electrical Engineering, University Of Edinburgh, Scotland,
‡Queens University, Belfast, Northern Ireland.

Introduction

The successful introduction of a new technology or products requires that both I.C. design and process development consider the target manufacturing environment. A Design For Manufacturability (DFM) strategy ensures that all the implications of of design decisions are fully evaluated before any product makes it into volume manufacturing. This requires the involvement of cross functional teams from the beginning of of the development phase to help ensure that assumptions are not which may adversely affect manufacturability. At the start of the development cycle changes to a design or technology are relatively cheap and get progressively more expensive as committments are made to equipment and wafer processing. Since, many crucial decisions are made during the early stages of development it is absolutely key that these decisions are correct. Therefore, it is essential, that the design team have access to the best tools available. Since time is of the essence, these must be accurate, simple to use and also capable of predicting all the important performance parameters and the variability that would be expected during normal operation of the manufacturing facility.This paper will present some of the available tools together together with a framework which fosters a DFM design strategy.

The Strategy

Technology CAD (TCAD) tools such as SUPREM-3, TSUPREM4, and MEDICI are widely used in the development of new processes. However, the traditional approach of varying one process control parameter at a time is very inefficient, and is not certain to result in a global optimum being found. Using simulation tools together with Design Of Experiments (DOE) in conjunction with Response Surface Methodology (RSM) enables a full parameter space to be explored thereby ensuring that the global optimum can be found. The framework, CAESAR, in conjunction with CALPHURNIA is ideally suited to performing this task and because this approach enables simulations to be performed in parallel, significantly reducing the time required to obtain results. The resulting response surfeces can also be used to generate response distributions An example of the resulting output (from the use of process and device simulation in conjunction with RS/1 and Cornerstone) is given in figure 1. In this case the known distributions on the three control factors have been used together with the response surfaces to predict the response ditributions. In this instance, Cornerstone has been employed to display the results and the "paintbrush" used to select the high values of the electric field.This has highlighted the combinations of control factor settings that gave these values together with the other responses associated with these combinations.

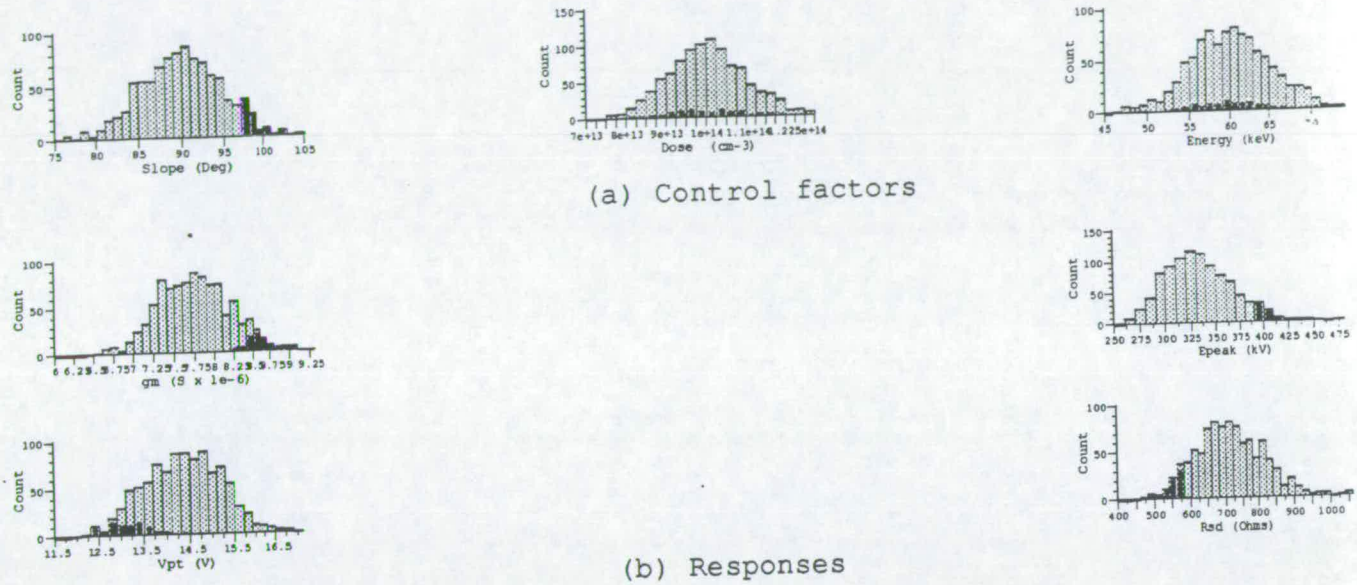
Another essential component of any DFM strategy is to use a similar approach with IC layout. Most designers attempt to minimise silicon area. Figure 2 gives gives an example of a tool that has automatically optimised IC routing to make it more robust to defects and therefore more manufacturable. Interconnect is having an increasingly more important effect upon circuit performance which is somewhat neglected by IC designersbecause of the difficultyin data preparation. Figure 3 shows an example of the automatic generationof 3D data files direct from IC layout. The use of such software enables IC designers to routinely calculate interconnect capacitancesand then use it in the CAESAR-CALPHURNIA DOE/RSM environmentto fully evaluate DFM issues.

The full paper will also present details of the software framework within which the simulation tools and DOE/RSM have been interfaced and also discuss how factory simulation tools must be incorporated in the future. It will also detail the collaboration that has taken place with the software vendors and how this has benefited the company.

1. A.J. Walton, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak. Using RSM Techniques To Contour Plot Parameters Related To Response Distributions Of Semiconductor Processes And Devices, Proc ESSDERC '94, pp829-832.

2. G.A. Allan, A.J. Walton,. Efficient Critical Area Alogrithms And Their Application To Yield Improvement And Test Strategies, 1994 IEEE International Workshop On Defect And Fault, Tolerance In VLSI Systems, Oct 94 pp 88-96.

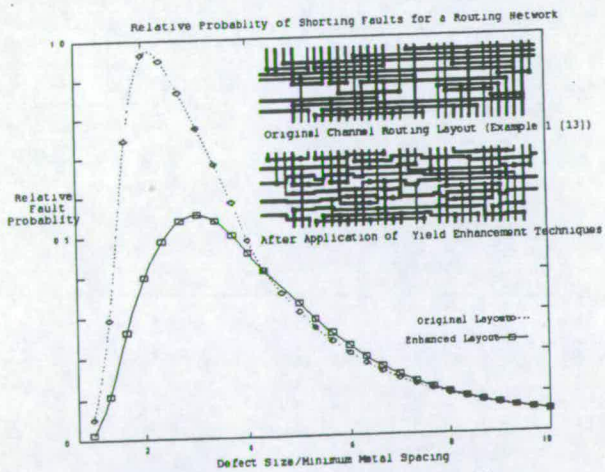
3. J.P. Elliot, G.A. Allan, A.J. Walton. Automating The Calculation Of 3D Interconnect Parasitics. To be presented at the IEE Colloquim On Advanced MOS And Bipolar Devices, Feb 95.



(a) Control factors

(b) Responses

Figure 1. Histograms of control factors and responses created using DOE/RSM related process conditions.



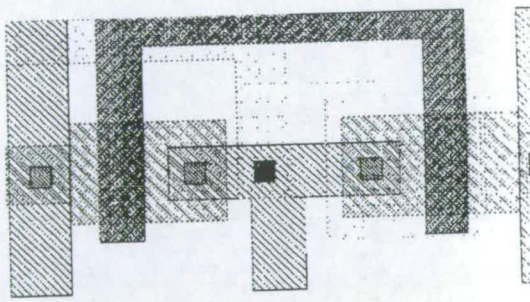
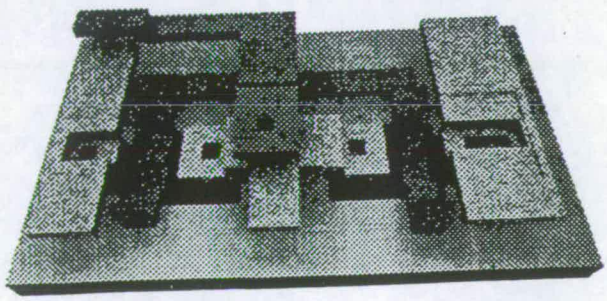
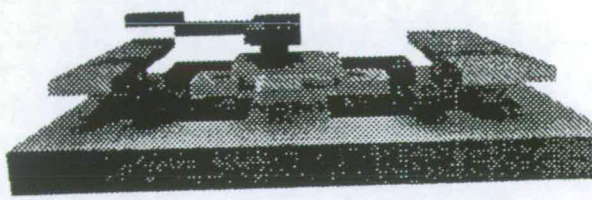


Figure 2. Yield improvement through track displacement (a)



(b)



(c)

Figure 3. 3D circuit representation automatically created from the IC layo

(c) Side view

UPD Characterisation Using Desirability Functions and Assessment of Process Variability

M. Fallon, G Martin, M. Redford, R. Prasad, T. Powell, H. Gregory, J. McGinty

Analog Process Technology Development
NSUK
email: martinf@sinbad.nsc.com

Abstract - Response surface methodology has been applied to unit process development to minimise the sensitivity of the response to the equipment settings and to a full process flow to quantify the effect of Gaussian distributions of the process variables on the responses of interest.

1. 0 Introduction

Two applications of the response surface methodology are presented. In Part I the LFAST transfer was the vehicle for process sensitivity analysis, to develop an emitter process for the 6 inch transfer. In Part II, UPD optimisation was performed using a full DOE and sensitivity optimisation within the RS/1 environment. In both cases an awareness of DFM is evident in developing the new process.

2.0 Part I: Process Variability

The 4 inch LFAST process was transferred to 6 inch and required new furnace recipes to eliminate stress effects on the 6 inch wafers by reducing the push/pull temperatures accordingly. The process was modelled using Suprem3, with the aim of matching dopant profiles, sheet resistances and junction depths between the two processes as shown in table 1. Since a number of possible processes could match the target specification a sensitivity analysis was performed. Six process variables were considered: base dose and energy, base diffusion temperature, emitter dose and energy, and emitter temperature. The six responses were emitter resistance, emitter junction depth, base junction depth, base width, pinched resistance and Gummel number and these are detailed in table 1.

A CCF experiment was designed within RS/1 and the process matrix simulated using Suprem3. The interface between the TMA Workbench tool and RS/1 was provided by Calphurnia, developed at Edinburgh University which provides a seamless link between the 2 software tools, transferring control factors and their limits from TWB to RS/1, building the experiment in RS/1 and providing the information for the process splits to TWB. Once the Suprem3 simulations are completed, the results are transferred by Calphurnia, which utilises a GUI format, into the worksheet in RS/1, and the response surfaces calculated automatically. The response function was calculated and an example of a surface is shown in figure 1. Since there are 6 responses of interest it is very difficult to assimilate information using these contour plots. At this stage the optimiser in RS/1 was employed. However, apart from achieving target response values, Calphurnia can also be used to examine the impact of the variability of the process parameters. Using this feature, Monte Carlo analysis was performed to directly link the response distributions to the process variables by defining the standard deviation of the process variables assuming a Gaussian distribution. Thus, at the chosen points in the experiment workspace, values of each process variable are chosen, constrained to lie within a 3 standard deviations limit of a Gaussian distribution, but chosen at random with respect to one another so producing a corresponding value of each response. Using this technique 500 'simulations' *at each point* in the worksheet can be performed in a few minutes. This is very much faster than performing the full process simulation and makes the Monte Carlo technique possible for complex simulation models.

Control Factors		Responses	
Base Dose	2.44e14	Rpinch	2965
Base Energy	50	Emres	21.52
Base Temp	1000	Xjeb	0.542
Emitter Dose	55	Xjbc	1.060
Emitter Energy	1e16	Base Width	0.518
Emitter Temp	890	Gummel No	1.106e13

Table 1: 4 Inch Parameters

(A Gaussian distribution of the control factors was chosen in this case, though it the software could be modified to accept other distributions and limits other than the 3 sigma ones used here). The control factor distributions can then be displayed in the form of a histogram and the corresponding responses similarly displayed. This is shown in

National Semiconductor Corporation Confidential

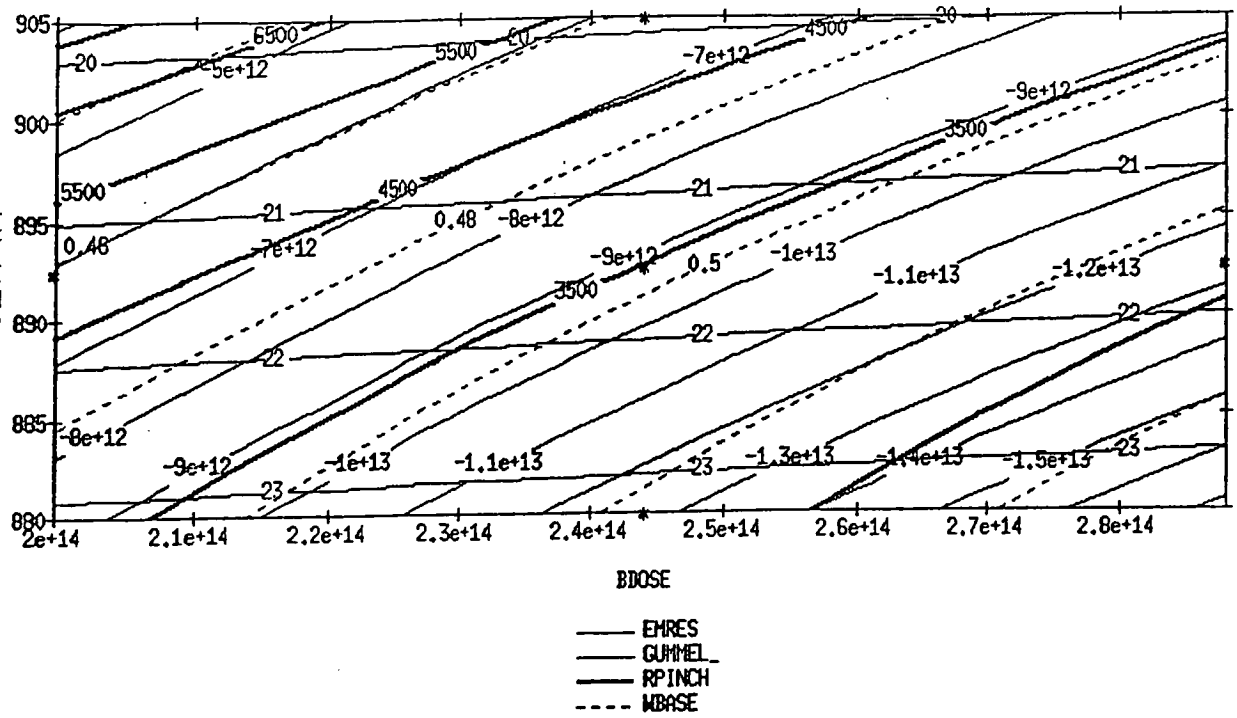


Figure 1: Response Surface

ure 2 for 3 of the control factors and 3 of the sponses.

A feature available within Cornerstone is the ability to aintbrush' regions and highlight the related regions in ner graphs. This is demonstrated by the red (dark) gions in figure 2 where it can be seen that the high val- s of pinched base are related to narrow base width and v Gummel numbers. These are produced by high val- s of emitter temperature, low values of base tempera- e with the emitter energy showing no particular pendency on the control factors. (Similar plots are ailable for the other control factors and responses, but not be shown here because of space limitations). This ormation can be used to identify the existing process iations. Whilst most of the response distributions pear to be Gaussian, it is not necessarily always true, spite using a Gaussian distribution of control factors. leed a striking example is seen in the asymmetry of the *inch* distribution which is mirrored in the current beta tributions. Previous work has identified a strong corre- ion between these 2 parameters, although *Rpinch* is t currently measured at electrical test. The case for cluding a structure to measure this parameter is demon- ated by these simulations. Through the paintbrush util- it is possible to identify the related process variables ich contribute to this skewness which then highlights sensitivity of the pinched base resistance to the con- l factor.

The final stage in the process development is to optimise the process such that the control factors operate in a region whereby routine daily fluctuations produce minimum fluctuations in the responses. This was performed using the Calphurnia software and RS/1. Once the response distributions have been calculated, Calphurnia then calculates the moments of the distributions and the Taguchi signal-to-noise ratios. The second moment, or standard deviation, is a measure of the variability of the response and this can be used in RS/1 as the parameter to drive the optimisation. The target values of the responses were defined by their 4 inch equivalents and the optimiser was run, constraining the target responses between tight limits whilst minimising the standard deviation of the pinched base resistance.

The resulting control factor settings were used to define the first run in the 6 inch process. Two alternatives were considered: a phosphorus and an arsenic emitter process. Of the 2, the arsenic appeared to be much more stable as the phosphorus process showed such a strong dependency on emitter temperature. However the lower junction breakdown voltage prohibited the use of arsenic. After the first silicon was processed further adjustments were required which is indicative of a calibration issue or a limitation in the 1D simulation.

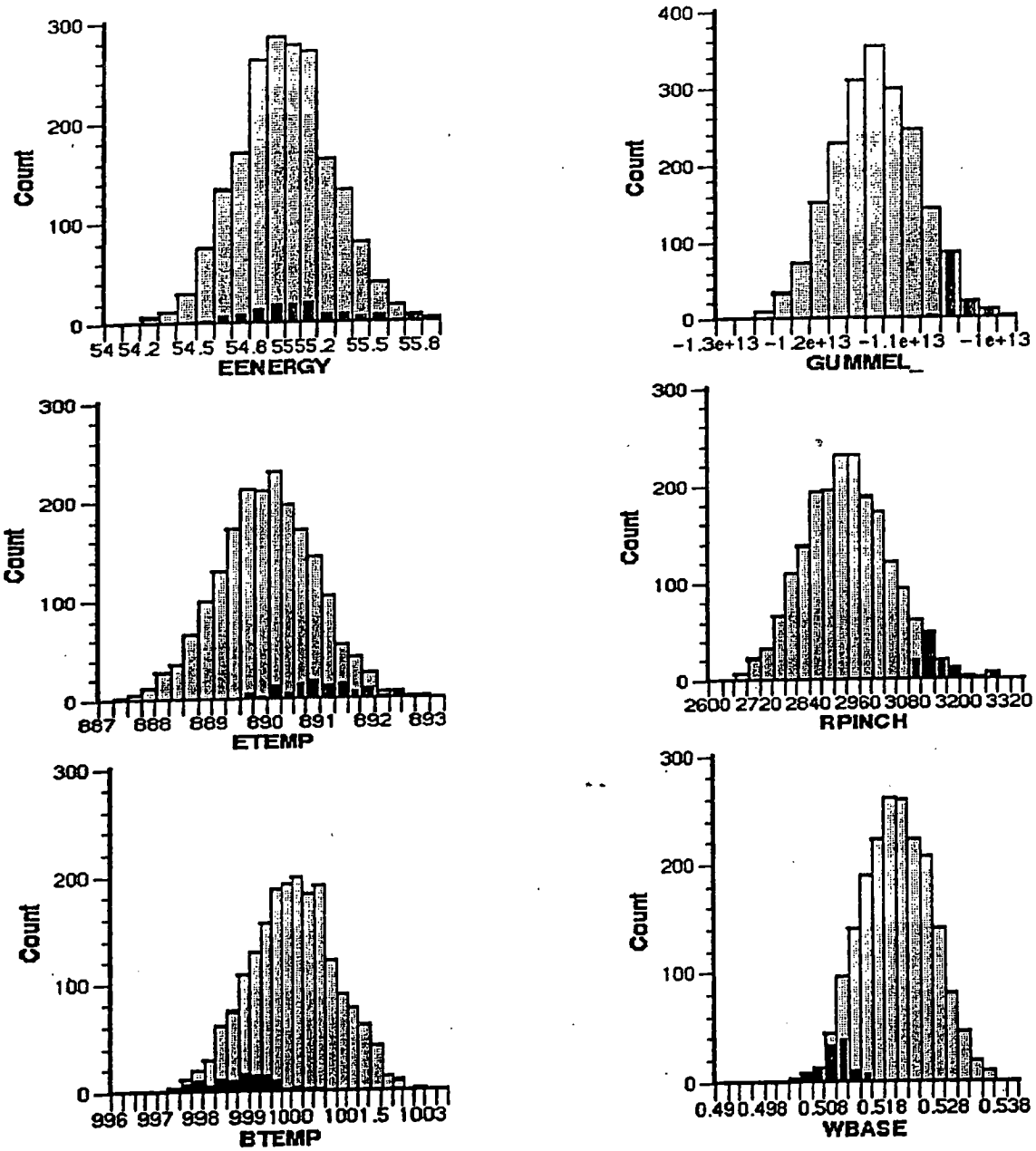


Figure 2 Control Factor and Response Distributions

) PART 2: Unit Process Development

Development of a Carbide Chrome resistor process is derway in NSUK. Two pieces of equipment were estigated as potential routes for depositing Carbide rome. For the sake of brevity only one shall be consid- d here, though the techniques were applied to both. ocess parameters and responses shown in table 2 were ntified and a DOE designed around these control fac- s using RS/1.

Control Factor	Response
Pressure	P-S Res
Temperature	P-A Res
Power	
Time	

Table 2: Varian Control Factors and Responses

4

The process examined was the deposition of the Carbide Chrome and the subsequent alloy, such that the responses were post sputter sheet resistance and post-alloy sheet resistance. A typical response surface for the Varian system is shown in figure 3. Two response values shown are the post sputter resistance (PSRHO) and the post alloy resistance (PARHO). As demonstrated in the last example, the common problem in dealing with response surfaces is that the number of responses and control factors can make visualisation of the problem difficult: more than 4 responses make a single contour plot very cluttered and since only 2 control factors may be plotted at one time, the number of contour plots to be viewed is n! where n is the number of control factors. This quickly becomes unmanageable. There are a number of possible ways round this problem. If the objective is to achieve an optimum process then the RS/1 optimiser may be employed. If the objective is to examine existing processes then the adoption of composite responses becomes a useful tool. The composite response combines the individual responses into one, simultaneously satisfying the requirements of all responses. This is particularly relevant when the responses have to be maximised or minimised. However in this case, the requirement was simply to produce a stable process which meets specification.

Ideally the cross-wafer uniformity would have been a response to be minimised. However the measurement system was not automated and the data was limited to 5 points on a wafer. This is not enough to model uniformity. While 5 points may be enough to model the mean (although this may not always be the case) considerably more data are necessary to model the standard deviation which directly relates to the wafer uniformity. If any attempt is made to optimise a process, then automatic data collection is essential to provide enough data to include uniformity as a response in the experiment. Since this was not the case for this experiment an alternative approach was used.

Each response surface is a function of the control factors and the sensitivity of the response is simply the amount that the response changes as a result of the change in the control factor. Mathematically, it is the first partial derivative of the response to the control factor. Using this definition both responses, PARHO and PSRHO, were differentiated with respect to each control factor using the analytic derivative function available within RS/1. These were summed to produce a single response shown in figure 3 as FABS. This process was automated by coding the procedure in RS/1 using BBN's own RPL. The optimiser was then used to minimise this function. The contours of this function are roughly perpendicular to the contours of the other two responses and since a value of 1100 ohm/sq was the desired post sputter value, the operating point was clearly identified at the intersection of the two contours.

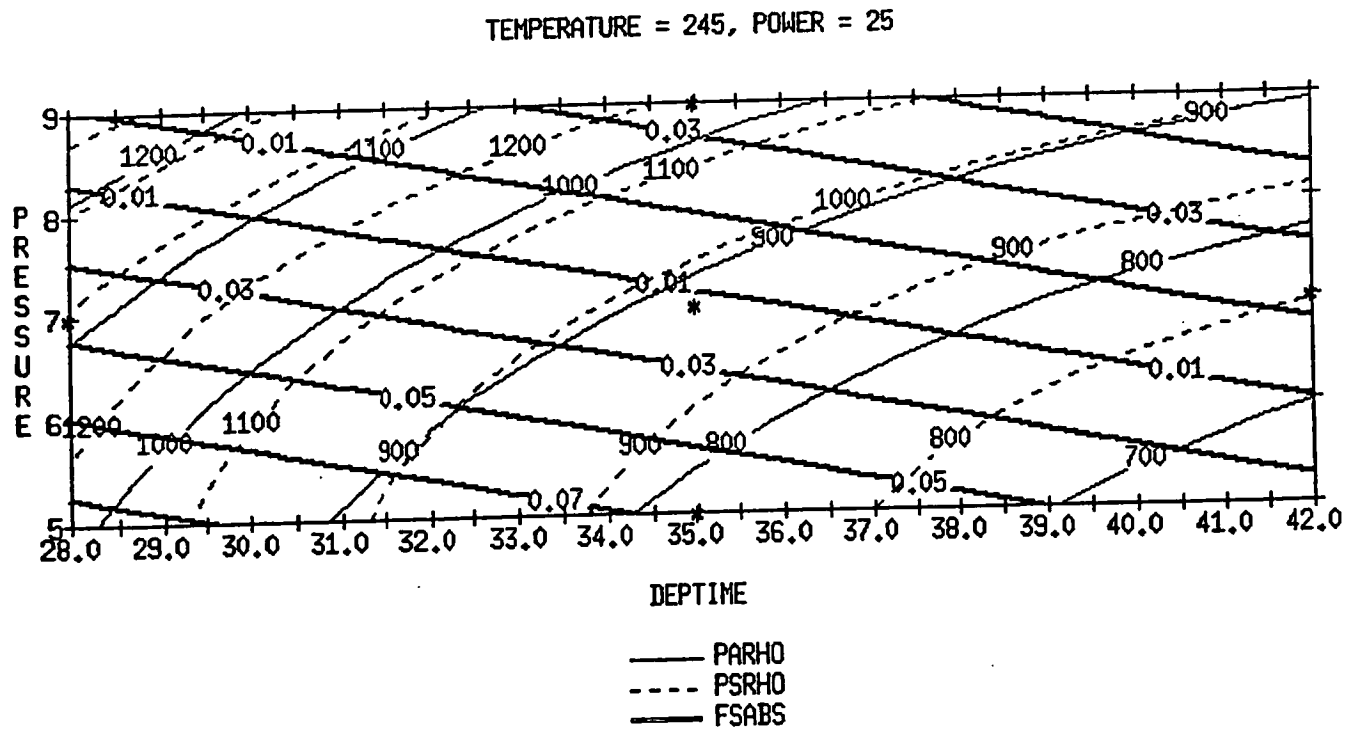


Figure 3: Response Surface for Varian Deposition System

National Semiconductor Corporation Confidential

HIGH VOLTAGE RESURF DMOS PROCESS DEVELOPMENT USING DFM TECHNIQUES A CASE STUDY

M. Redford, M. Fallon, Z.A. Shafi, J. McGinty, N.S. Rankin, A.J. Walton†

Analog Process Technology Development, National Semiconductor (UK) Ltd.,
Larkfield Industrial Estate, Greenock, Scotland,
†Department Of Electrical Engineering, University Of Edinburgh, Scotland

Abstract

Utilising Design For Manufacturability (DFM) techniques, it is possible to reduce the number of split runs of silicon, and the number of cycles of learning, needed in the development of a technology.This ultimately resulting in a process which consistently meets its specifications, and can be manufactured without great engineering effort. The purpose of the work was to develop a process which consistently yielded greater than 400 volt breakdown for applications in the power supply market. Based on the results of the analysis, a process was chosen which was proven after one iteration of silicon, thereby saving on development time and costs. This paper will review the work undertaken to determine the optimum process flow in order to achieve the desired electrical performance, and analysis of the predicted breakdown distribution. A comparison is made to actual distributions obtained from silicon. Then from experimental split lots a sensitivity analysis is performed. A comparison of simulated and actual results are compared. In conclusion, learnings made as a result of the analyses, the advantages and disadvantages of the method/software and methodology used when running and gathering silicon data.

1. Introduction

In today's competitive environment, time to market for new products is of paramount importance in order to win market share. It is no longer feasible to continue to debug process technologies after manufacturing release. Therefore, a methodology/strategy [1] is required that reduces the development cycle time. The strategy utilized in this work is a Design For Manufacturability (DFM) using Technology CAD (TCAD) and Design Of Experiments (DOE) where process and device simulation is used in conjunction with DOEs to produce response surfaces [2]. The advantage of RSM methodology is that the number of experimental runs explores the complete parameter space and the resulting response surface can be used to help optimize the control factor settings. Whilst utilizing DFM techniques to reduce the number of cycles of learning, an analysis of the impact of control factor sensitivity on responses provides a greater insight into, for example, which control factor(s) would require monitoring in-line. This could feasibly lead to moving a reduction in the number of electrical tests performed at end of line, and zero losses.

2. Technology overview

In high voltage applications Double-diffused MOS (DMOS) transistors are more attractive than conventional MOS structures integrated circuits, since the drift region between the drain and the active channel allows high voltage biasing. The structure under investigation is a lateral DMOS transistor with a thin epitaxial layer, and applies the Reduced SURface Field (RESURF) principle [3] which distributes the applied drain voltage uniformly across the drift region. A cross section of the device is shown in figure 1. The operation of a RESURF device relies on careful selection of

the resistivity and thickness of the epitaxial layer [3], and the substrate resistivity. The breakdown requirement for the RESURF device was a guaranteed minimum breakdown voltage of 400 Volts, and 450V target. Two process options were considered: (1) an epitaxial process where the epitaxy concentration sets the charge in the drift region, and (2) an epitaxial process with a higher resistivity where the charge in the drift region is set using an implant [4].

3. TCAD Calibration Procedure

The output and accuracy of a simulator is only as good as the information on which it is based The process flow was modeled using SUPREM-3, and the doping profiles were input into the device simulator Medici. The initial process specification was based on previous work [5]. The simulations were based on models used for the SMART power process from which the flow under investigation was derived, therefore, providing some confidence in the usefulness of the models and hence, TCAD tools used.
The process model was calibrated by including as much process information in the model as possible. This included full diffusion profiles including ramp up & downs, cleans, etc... Oxide thicknesses were calibrated to silicon data, the profiles to SIMs/SRPs, and transistor threshold voltage V_t was calibrated by setting interface charge, Q_{ss}, in the gate oxide.

4. TCAD & DOE Procedure & Results

RS/1 software was used to design an experiment in order to understand and evaluate the impact of the three variables on the breakdown voltage. The layout of the structure under investigation is shown in figure 2. This is essentially a 3 dimensional device. The simulated breakdown was determined from the minimum of three simulated values: one for portions of the structure where source and drain stripes are parallel to one another; one for the source fingertip; and one for the drain fingertip. The results of the RSM analysis showed that the original epitaxy resistivity specification was not capable of delivering 400V over the entire process spread. Another DOE was run but with an increase in the upper limit of the resistivity specification. This resulted in 'workable' epitaxy specification. An example of the resultant contour plots is shown in figure 3. The creation of response distributions is an important piece of the DFM strategy since it provides important information on the manufacturability of the process under consideration. Calphurnia[6] makes the assumption that the control factor distributions are Gaussian in nature, though any distribution function can be used. Once the standard deviations of the control factors have been defined, it then becomes possible to generate their distributions then subsequently the response distribution(s). The RSM contour plots obtained from simulation were analyzed in Calphurnia. The resultant distributions are shown in figure 4. The standard deviations for the control factors were the same as used in [7]. These results are compared with those obtained from silicon, shown in figure 5.

5. Results From Silicon Split Lots

Various split lots were run through fabrication. The results of these splits were analyzed using MULiple REGression analysis (MULREG), and the results were compared to those from simulation. The results from over 200 wafers were used in the analysis, the R^2-adj was 0.789. A comparison of resultant contour plot, figure 6, with that from simulation shows that the contours are not exactly the same, even though the simulation results successfully resulted in a process that consistently gave >400 volts.
Comparison of experimental and simulated results for an alternative epitaxy and implanted process were very similar but again these also resulted in a process that consistently met the electrical specification. In this case the R2-adj fit was 0.4, and a comparison of the associated contour plots has also been made.

6. Conclusions & Observations

There are several conclusions and observations that can be made from this work. First TCAD and DOE have been employed together to sucessfully develop a process with the desired breakdown voltage. In particular Calphurnia has been used to determine which process flow was the most manufacturable. This work has also highlighted inaccuracies with the epitaxy measurement techniques in the 10-15 Ohm-cm range. This was a big issue for the epitaxy and drift implant flow where the variability in the measurement resulted in an R^2-adj of ~0.4 although for the epitaxy process the R^2-adj was 0.789. These results initiated a review of the technique used for measuring epitaxy which has been subsequently changed. The simulated results have used a 2D simulation approximation for a 3D device and a poor modelling accuracy of the device topography (eg poly field plates). However, the reported procedures have proved to be a valuable analysis tool for DFM and this highlights the importance of striking a balance between simulation accuracy and execution time

References

[1] A.J. Walton, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, G.A. Allan, and J.P. Elliott, "A Total TCAD Strategy for DFM in IC Technology Development", IEE Proc. Science, Measurement and Technology , vol. 144, no. 2, pp. 63-68, March 1997.
[2] M. Fallon, G. Martin, M. Redford, R. Prasad, T. Powell, H. Gregory, J. McGinty, 'UPD Characterisation Using Desirability Functions and Assessment of Process Variability', *National* Semiconductor Technology & Innovation Conference, 1996.
[3] J.A. Appels, M.G. Collet, P.A.H. Hart. H.M.J. Vaes, J.F.C.M. Verhoeven, 'Thin layer high-voltage devices (RESURF devices). Philips Journal of Research, vol 35, no 1, pp1-13. 1980
[4] E.J. Wildi, P.V. Gray, T.P. Chow, and H.R. Chang. Modeling and process implementation of implanted RESURF type devices. In IEDM Digest of Technical Papers, pages 268-271, 1982.
[5] J.M. McGregor & M. El-Diwany. 450V lateral DMOS device module compatible with LMDMOS technology. In National Semiconductor TCAD Workshop '93, pages 10-13, 1993
[6] M.I. Newsam, "Calphurnia V2.0 Manual", Dept. of Electrical Engineering, University Of Edinburgh, Scotland.
[7] D.A. Hanson, R.J.G. Goossens, M. Redford, J. McGinty, J.K. Kibarian, K.W. Michaels, "Analysis if Mixed Signal Manufacturability with Statistical Technology CAD (TCAD)", IEEE Trans. On Semiconductor Manufacturing, Vol 9, No. 4, Nov 1996

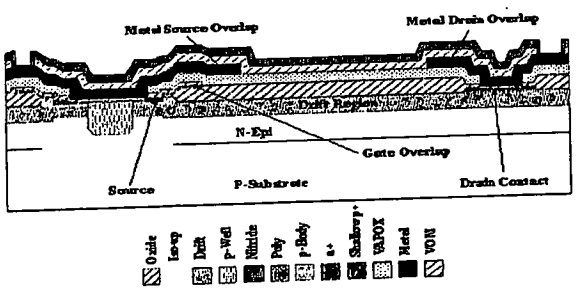


Figure 1
Device Cross Section

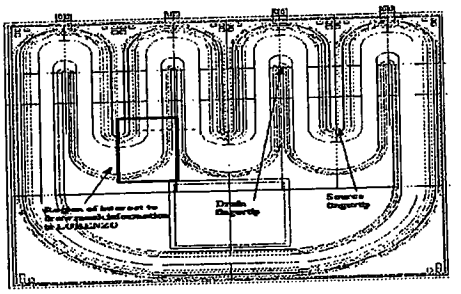


Figure 2
Device Layout

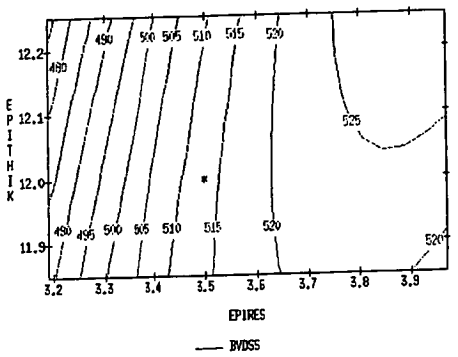


Figure 3
Simulation RSM Contours

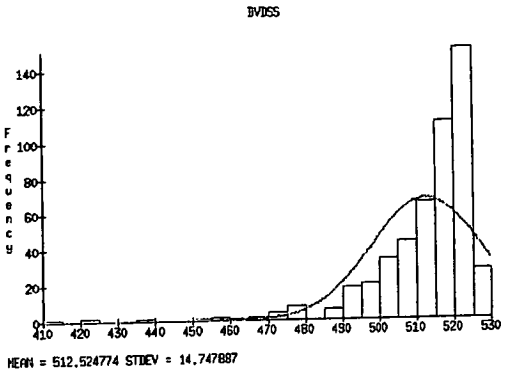


Figure 4
Breakdown Distribution
(Simulation)

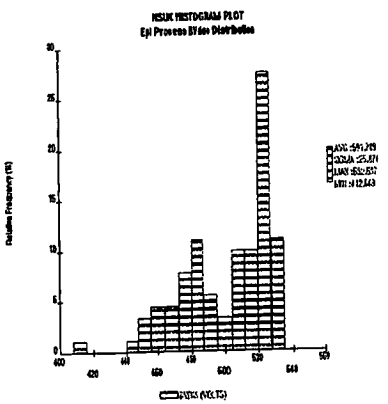


Figure 5
Breakdown Distribution
(Silicon)

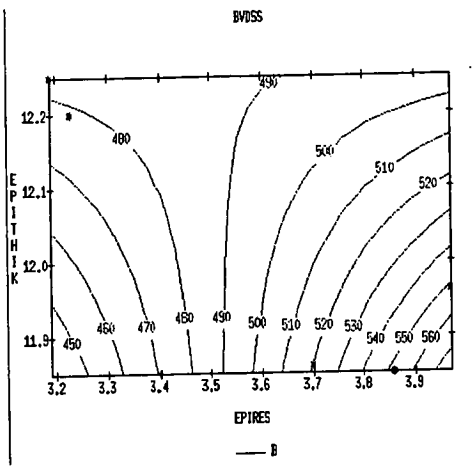


Figure 6
Contour Plot
(Silicon)

Analysis of Mixed-Signal Manufacturability with Statistical Technology CAD (TCAD)

David A. Hanson, *Associate Member, IEEE*, Ronald J. G. Goossens, *Member, IEEE*, Mark Redford, Jim McGinty, John K. Kibarian, *Member, IEEE*, and Kimon W. Michaels, *Member, IEEE*

Abstract—**We have developed a methodology which combines technology CAD (TCAD) simulation with statistical analysis of empirical data to predict and control the manufacturability of IC fabrication processes. As a result, manufacturing tolerance or sigma-based models (also known as worst-case models) can be determined before a significant sample size of fabricated devices can be characterized. Early on in the development cycle, empirical data is collected, and models built from simulated data are refined. These revised models are used to determine process control limits, and optimize in-line and electrical test measurement (E-test) for maximum observability of variation. As the process is stabilized, further refined models are used to perform yield diagnosis and tolerance analysis of circuits. This methodology has been applied to a number of BJT and submicrometer CMOS processes to create predictive sigma-based models, modify the fabrication recipe to meet objective specifications as development proceeds, and finally use them to control the manufacturing line.**

I. Introduction

THE manufacture of integrated circuits (IC's) involves the design of fabrication recipes, high-level behavioral descriptions, circuit netlists, and topological layouts. During the development phase, there is little empirical data available upon which to base SPICE models for concurrent chip design. Later on during the production phase where model control is essential, equipment fluctuations cause variations in the device and interconnect characteristics. To predict and then control manufacturability, it is necessary to understand how these variations impact product performance.

Integrated analog and digital or mixed-signal fabrication processes must meet the objective specifications of both analog and digital devices and compensate for the interactions between them. The high switching speeds of digital circuits create more noise than is desirable for the analog circuits. Conversely, the higher voltages and additional parametric specifications (R_{dson}, g_{mout}) required by analog circuitry results in higher parasitics and extra processing steps.

Design centering and manufacturing tolerance minimization become more crucial in a mixed-signal manufacturing environment. We refer to design centering or the desensitization of a product to process variations as "design for manufacturing" (DFM). Conversely, we refer to tolerance minimization or the robust design ($C_{pk} > 2$) of fabrication processes as "manufacturing for design" (MFD). DFM or MFD typically rely on a statistical characterization of the process [2], [3], [10]. Since traditional statistical characterization techniques are both costly and time-consuming, it is usually undertaken only once, during process qualification.

Since both techniques rely on the collection of a great deal of empirical data, it is often the case that the models are developed after the initial products are designed, but while the process is still unstable. Consequently, approaches which use only empirical data provide statistical information too late to improve the yield for the newest product/process combinations. Moreover, since the characterization is not updated as the process matures (yield improved by learning), new designs for the now older, more stable process are unnecessarily conservative. The result is a waste of expensive fabrication resources: the yield ramp is slow initially, and process capability is underutilized later in the process lifetime.

We have developed a holistic methodology for process characterization, which is both timely and adaptive to data collected during manufacturing. Implementing this methodology provides mechanisms for tightly coupling customer requirements to design, process development, and manufacturing. A number of BJT, CMOS, BiCMOS, and DMOS processes have been characterized by this approach. In this paper, the methodology is first explained and then demonstrated by application to a submicrometer CMOS process and a high-speed complementary BJT process.

II. Overview of Approach

Complexity in fabrication process characterization arises from the following:

- A process contains hundreds of steps, many of which contribute to the variability of the product.
- Only a few in-line and *in situ* measurements are collected, and only a few of the possible sites are measured.
- Many device structures may be used by designers but only some are thoroughly characterized with measured data.
- Many electrical test measurements (E-tests) are collected, but determining the relationship between E-tests and SPICE parameters for a process is a costly, time-consuming procedure.
- Designers quantify the process with SPICE models, but directly measuring SPICE parameters, especially the

Manuscript received January 22, 1996; revised July 2, 1996.

D. A. Hanson and R. J. G. Goossens are with National Semiconductor Corp., Santa Clara, CA 95052-8090 USA.

M. Redford and J. McGinty are with National Semiconductor Ltd., Greenock PA 16 0EQ, U.K.

J. K. Kibarian and K. W. Michaels are with PDF Solutions Inc., San Jose, CA 95110 USA.

Publisher Item Identifier S 0894-6507(96)08116-X.

0894–6507/96$05.00 © 1996 IEEE

parameters affecting ac or transient performance, is costly in a manufacturing environment.

- Many SPICE parameters are correlated so one must account for the interactions between parameters to create realistic worst-case or sigma-based models.

As stated in the introduction, two constraints must be alanced. First, the process needs to be characterized as arly as possible. If one chooses to use only measured data, ample size requirements to assure reasonable confidence in he statistics make it impossible to accomplish characterization n a timely fashion. For example, if confidence intervals for he standard deviation of 85% are required, transistors from 00 or more lots must be characterized. Waiting for 100 lots o be manufactured would make qualification of the process oo costly and time-consuming. To circumvent this problem, echnology CAD (TCAD) simulators coupled with a statistical simulation environment are used to predict device distributions prior to manufacturing.

The characterization process starts by simulating a process recipe and device layout. The results of our simulation include a distribution of SPICE parameters. It is important to note that a deterministic extraction procedure is used. Hence, cause and effect relationships between changes in the device structure (gate length, oxide thickness, etc.) and SPICE parameters can be determined.

The second constraint is that process variations must be monitored to determine the impact on SPICE model parameter distributions. Unfortunately, it is not possible to directly measure SPICE model parameters at wafer electrical test. Even though there have been improvements in the number and variety of *in situ* and in-line measurements, we have found that the observability for controlling the device characteristics is still not sufficient [1]. Often, electrical test measurements are the only choice for monitoring the device characteristics during production.

A technique incorporating principal component analysis (PCA) and pattern recognition has been developed which takes as inputs the simulated distributions of SPICE parameters, electrical tests, and in-lines and determines the following features of the process:

1) The number of sigma-based corners for the process relative to any particular device model and architecture.
2) Which in-lines or E-tests can be used to track the statistical variations in the SPICE parameters.
3) Nonlinear empirical relationships between E-tests and SPICE parameters.

As process development proceeds and in-line and electrical test data is collected, the nonlinear expressions are used to compute how the distributions in SPICE parameters have changed. For example, when the variance of an E-test distribution must be reduced to meet objective specifications, the effect on sigma-based SPICE model cards can be calculated. This same approach is used to control the mature manufacturing process.

In summary, the methodology presented uses simulated data to provide models when little measured data is available. The simulation results also aid the decision of what to measure during production and determine how to interpret the results (see Fig. 1). Our approach to statistical characterization can be thought of as a predictor–corrector strategy. TCAD simulations are used to predict statistical distributions prior to manufacturing a large number of lots. As E-test data are collected, this information is used to update the predicted SPICE performance. This procedure affords both the predictive ability outlined in [3] as well as "accuracy" outlined in [2].

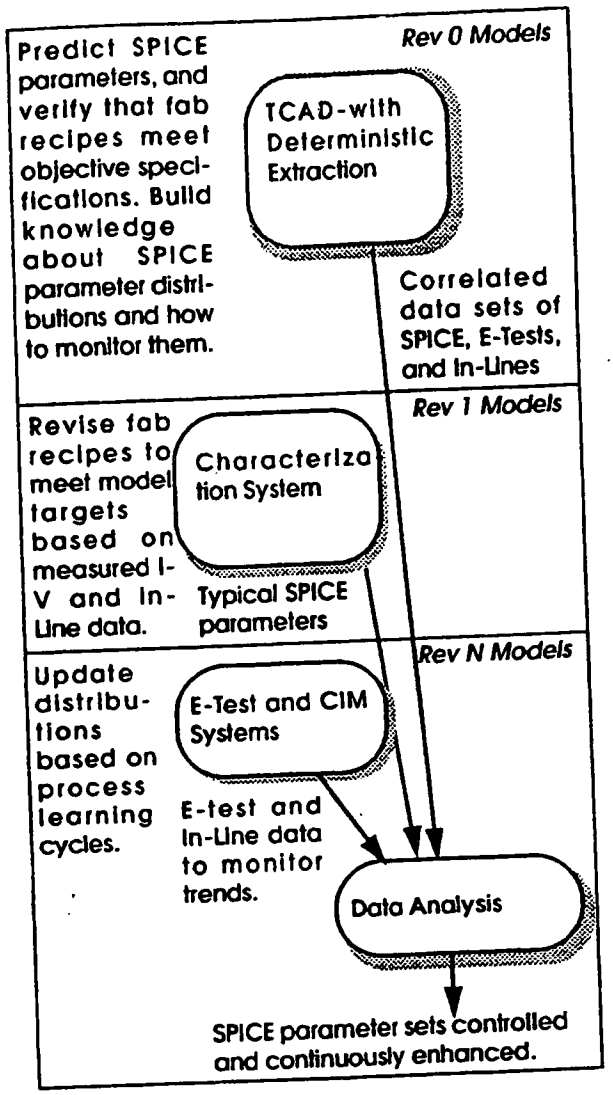


Fig. 1. Using TCAD, characterization, and electrical test data to develop, update, and monitor statistical SPICE model parameter sets.

III. TCAD: Building the Process Knowledge Base

As described above, the statistical variations of SPICE parameters need to be characterized before a large number of lots have been processed. Measuring many wafers from a few lots is not sufficient to accurately estimate factory variability. The distribution of device performance when sampled from a single or few lots is typically much tighter than the distribution sampled from many lots. One solution to this problem is to run statistical TCAD simulations to predict the manufacture of hundreds of lots.

There are two goals for the TCAD simulation. First, determine the statistical distributions of the SPICE model parameters. Second, determine from this data which in-lines and E-test measurements must be measured during production to control the SPICE parameter distributions. This second requirement places constraints on how we extract the SPICE model parameters. Traditional optimization or curve-fitting methods cannot be used since they corrupt the statistics by confounding of independent variables. This confounding of variables makes it difficult to interpret the correlations of parameters which are extracted via optimization. In the next two sections, an overview of the statistical simulation method and physically based SPICE model extraction is given.

A. Experimental Design

A goal for this part of the work is to use process and device simulation to determine the distributions of device characteristics based on the distributions of in-line measurements and equipment settings (process controls). Unfortunately, there are few if any closed-form analytical solutions which fully address these relationships. Since the functional relationships are not well understood, it is not possible to simply transform the in-line and process control distributions to determine the device distributions.

Monte Carlo simulation provides the most accurate method of mapping process control variation to device variation. At the completion of the Monte Carlo simulation, distributions of device characteristics, their intercorrelations, and their correlations with in-lines and process controls can be calculated. To achieve reasonable confidence in the predicted correlations, the number of simulation runs must be fairly large (typically 50–200).

Using simulation, there are two alternatives to predict the statistics of the process. Either one can run a Monte Carlo simulation incorporating physically based TCAD, or one can run a designed experiment (e.g., fractional factorial), build a response surface model (RSM), and perform a Monte Carlo simulation with the RSM to generate statistics. While the DOE technique may be faster, one must either assume the nonlinearities can be modeled with a low-order polynomial, or know *a priori* which relationships are nonlinear and account for this in the experimental design. Fig. 9 illustrates the nonlinear correlation which may encapsulate the interaction between the SPICE parameters. An additional implied assumption when using the DOE/RSM techniques is that one must decide *a priori* which process controls have a significant effect on the variability of the process, and include these controls in the experimental design. Since accounting for the nonlinearities and modeling the statistics are important, we decided to use Monte Carlo simulations to generate our statistics. Other randomized experiments such as Latin hypercubes would also be appropriate.

B. SPICE Model Extraction

In order to provide meaningful statistical information, the SPICE model parameters must be extracted in a physically based manner. Parameter optimization or curve-fitting, as shown in Fig. 2, which can result in a nonunique parameter

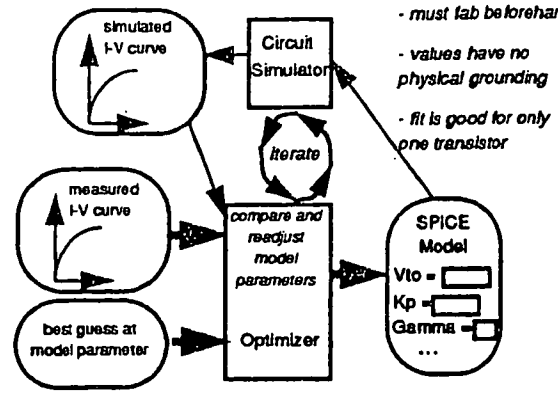


Fig. 2. Flow chart of conventional optimization-based device characterization procedure.

set, is not used. The parameter extraction methodology used incorporates both 2-D dopant layout information and efficient 2-D numerical device simulation performed using SIMOS [17].

A typical SPICE model (for example Gummel–Poon or BSIM3) has approximately 30–45 parameters. As described above, many of these parameters are correlated. Some of these parameters can be thought of as having a physical definition. For example, the BSIM3 parameter R_{dsw} is the parasitic resistance per unit width of the source/drain. Other parameters are more empirical in nature. For example, the BJT Gummel–Poon model parameter *IKF* impacts the current gain degradation at high collector currents. A number of possible physical phenomena can explain the variations in *IKF*. For example, *IKF* may be limited by high injection into the base, or current-induced base widening.

The BSIM3 model is similar to the Gummel–Poon model in that it is comprised of both physically defined and empirical parameters. The parameter T_{ox} (gate oxide thickness) has, of course, a physical definition, whereas the parameter *UC* empirically models mobility degradation due to bulk bias.

For SPICE model generation, we use direct physical extraction routines to extract parameters with a physical basis. For parameters which are empirical in nature, we extract them to fit the part of the *IV* curve for which they are defined. In the example of *UC* for the BSIM3 model, we fit this parameter by applying back bias to the substrate and measuring the change in the mobility in the channel.

The key to using such an approach is the ability to access information about the device structure which is typically difficult to derive from measurements of real devices. For example, when extracting SPICE parameters from a real device, one can measure the *IV*'s and *CV*'s but one does not know the shape of the doping profiles in the gate, L_{dd}, and source/drain. By using process simulation to create the device structure, the exact device structure can be determined. For example, the lateral source/drain diffusion under the gate of the device is available from the concentration profiles. This information can be used to accurately solve for many of the length dependent parameters in the BSIM3 model.

In the course of the work, both full 2-D numerical process device simulation, and a mixture of 1-D numerical process

mulation with 2-D extensions based on analytical approx-
nations have been used. For the data presented in this
aper, the 1-D numerical simulation with 2-D extensions
ased on analytical models was used. Traditionally, the 2-D
pproximations have been basic ratio computations [16]; in
is work the analytical models from FABRICS [4] were used
compute the lateral diffusion based on the 1-D SUPREM3
rofiles. This is explained in some detail in [13].

Since the exact structure for the device is known, informa-
ion from the mesh can be used to gain more insight about the
evice than what is typically available from measurements of
eal devices. Junction depths, depletion widths, and concen-
ration profiles normally determined from destructive physical
neasurements or extracted from an IV curve are available
without additional testing. As a result, with our method there
s a one-to-one relationship between the device structure and
SPICE parameters.

IV. Example of Predictive SPICE Extraction and Statistical Simulation

We have applied our simulation technique to a 0.8-μm CMOS process. In this process, both the n-MOS and p-MOS transistors are formed with L_{dd} regions. The p-MOS device also has a buried channel. In the next sections we summarize the results of this work. While we compare the simulated results with measured data, our goal was to extract conclusions which could be drawn from the data.

A. Predictive SPICE Model Extraction for MOS Transistors

Once entered into pdFab [13], the process was simulated with all control settings at their designed values. BSIM models were generated with pdFab for both the n-MOS and p-MOS transistors. The IV's generated with the SPICE model were compared to curves generated by IMOB33, a National Semiconductor proprietary model (see Figs. 3 and 4). The IMOB33 model was fit to the "typical" devices measured. Since the exact conditions that were used to fabricate the transistor to which the IMOB33 model was fit are not known, it is not possible to compare a goodness of fit of the models to the empirical data. However, there is good agreement shown between the IMOB33 model and the simulated BSIM model suggesting that the "typical" transistor represented the process nominal. The pdFab extraction computation (including the process and device simulation) takes 15 min of CPU time on an IBM RS6000 550.

B. Example of Statistical Simulation for 0.8-μm CMOS

With satisfactory results from the simulation of the designed process, a Monte Carlo simulation of 100 samples was run. Rather than doing a full BSIM extraction for every point, the threshold voltage and $I_{d\,\max}$ (5 V applied to gate and drain) were extracted at every point for both n and p transistors. The Monte Carlo simulation with 100 process steps required eight hours of CPU time on the IBM RS6000 550.

Running the simulation with the standard deviations set to match the in-line measurement specification limits resulted in good agreement between simulated and measured device

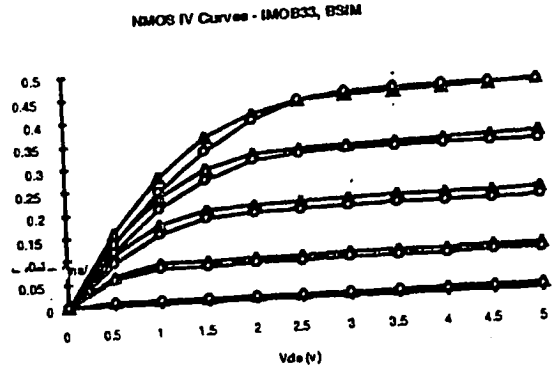


Fig. 3. n-MOS IV curves: IMOB33 and BSIM.

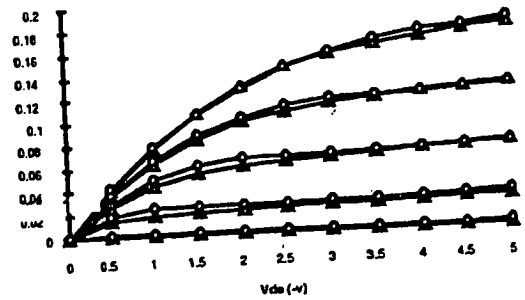


Fig. 4. p-MOS IV curves: IMOB33 and BSIM.

specification limits (less than 5% difference). However, the measured and simulated correlation coefficients did not agree. Upon examination of the measured in-line distributions for T_{ox} and L_{poly}, it was found that the measured T_{ox} distribution was tighter than the specification, and the L_{poly} distribution was larger than the specification. The standard deviations of resist thickness for the L_{poly} mask layer and of the temperature in the gate oxidation furnace were adjusted accordingly and resimulated.

The original standard deviations were based on the equipment specification limits. By examining the distribution of in-line gate oxide thickness, we were able to determine that at National, the equipment was running in better control than the specification. Although the specification limits for the devices did not change significantly, the correlation coefficients did change. The increased linewidth variation caused the correlation between I_{dp} and I_{dn} to increase (larger absolute value of correlation coefficient), while the I_{dp} and V_{tp} correlation decreased. In general, it is preferable if the n- and p-channel transistor drive currents track, so the added correlation due to extra linewidth variation may be useful.

In Table I, the process specification limits are compared to three sigma points for both the measured and simulated data. In general, the simulated distributions are closer to the measured distribution than to the specification limits. On a percentage basis, better control is achieved for the n-MOS transistor than the p-MOS transistor. The measured data consisted of approximately 1000 sample points per parameter collected from WW20 to WW45 in FY93. The data were taken from the MM9309A CS080 flow.

In Table II, the correlation coefficients between V_t and I_d are computed from simulated and measured data. Given the sam-

TABLE I
COMPARISON OF SIMULATED AND MEASURED 3-SIGMA LIMITS TO PROCESS SPECIFICATION LIMITS

Measurement	0.8 μm Specs	Measured Data	*pdFab*
$I_{dn\ max}$	0.50 mA/μm	0.47 mA/μm	0.50 mA/μm
$I_{dn\ mean}$	0.42 mA/μm	0.42 mA/μm	0.45 mA/μm
$I_{dn\ min}$	0.34 mA/μm	0.37 mA/μm	0.39 mA/μm
$I_{dp\ max}$	−0.22 mA/μm	−0.20 mA/μm	0.20 mA/μm
$I_{dp\ mean}$	−0.18 mA/μm	−0.17 mA/μm	−0.18 mA/μm
$I_{dp\ min}$	−0.14 mA/μm	−0.14 mA/μm	−0.15 mA/μm
$V_{tn\ max}$	0.85 V	0.79 V	0.74 V
$V_{tn\ mean}$	0.73 V	0.71 V	0.70 V
$V_{tn\ min}$	0.60 V	0.64 V	0.65 V
$V_{tp\ max}$	−1.10 V	−1.08 V	−1.04 V
$V_{tp\ mean}$	−0.95 V	−0.98 V	−0.93 V
$V_{tp\ min}$	−0.80 V	−0.88 V	−0.83 V

TABLE II
SIMULATED AND MEASURED DEVICE CORRELATION COEFFICIENTS

	Vtn (meas)	Idn (means)	Vtp (meas)	Idp (meas)
Vtn	1.0	-0.61 (-0.73)	0.13 (0.21)	0.12 (0.42)
Idn		1.0	0.0 (0.13)	0.71 (0.8)
Vtp			1.0	0.57 (0.54)
Idp				1.0

TABLE III
SIMULATED AND MEASURED DEVICE AND IN-LINE CORRELATION COEFFICIENTS

	L Poly (meas)	Tox	Nwell Rsh (meas)	Pwell Rsh (meas)
Vtp	-0.03 (-0.21)	-0.05	0.77 (0.59)	-0.08 (-0.10)
Vtn	0.02 (0.39)	0.86	0.33 (0.02)	-0.36 (-0.20)
Idp	-0.76 (-0.76)	-0.11	0.24 (0.21)	0.11 (0.31)
Idn	-0.76 (-0.75)	-0.52	-0.31 (-0.07)	0.21 (0.45)

ple size, all correlations, with the exception of the correlation between V_{tn} and I_{dp}, were within statistical confidence of the measured values. It is not clear why the simulated correlation between V_{tn} and I_{dp} is different from the measured value. It is interesting to note that the correlation between devices (I_{dn} and I_{dp}) is greater than the correlation within the devices (I_d versus V_t). Later this information will be used to generate sigma-based model corners for design.

1) Device Correlations with In-Lines and Process Controls: Correlation coefficients between device characteristics and in-line measurements were also generated from simulated and measured data. Measurements such as poly linewidth CD (L_{poly}), gate oxide thickness (T_{ox}), and sheet resistances (R_{SH}) were considered to be in-line measurements. While L_{poly} and T_{ox} are measured in-line, N_{well}, P_{well}, and R_{SH} were not. While these sheet resistances can be measured in-line, they are typically not measured due to the time delay and

TABLE IV
PERCENTAGE OF DEVICE VARIANCE PREDICTED BY IN-LINES

	Tox	L poly	L poly & Tox	Nwell Rsh
Vtp	0.3%	0.09%	0.4%	59%
Vtn	74%	0.04%	74%	11%
Idp	1%	58%	59%	9.6%
Idn	27%	58%	85%	5.7%

extra effort required. Complicated testing schemes (sometimes destructive) requiring pilot wafers could have been used on a batch (multi-lot) basis.

Due to limitations in the database, it was not possible to determine the measured correlation between T_{ox} and the device characteristics. In general, good agreement between measured and simulated correlation coefficients was found. It is interesting to note that the linewidth variations dominate both I_{dp} and I_{dn} while neither V_{tp} or V_{tn} are strongly correlated with polylinewidth variation. It is the common dependence on linewidth variation that creates the correlation between I_{dn} and I_{dp}.

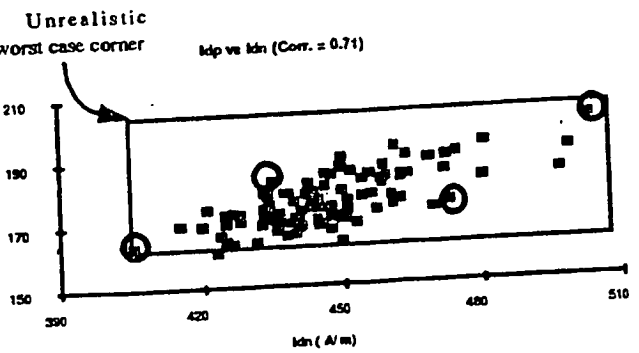
2) Application to Process Control: In-line measurements (e.g., critical dimensions and layer thicknesses) are made during fabrication to monitor the production of integrated circuits. At the completion of all process steps, electrical test measurements are performed on test structures to monitor device performance. Manufacturing engineers compare these measurements to specification limits as a means of predicting whether the fabrication line will produce working circuits at a sufficient yield. Two questions arise when using in-line measurements to monitor the process. First, which in-line measurements should be made (since there is a cost associated with each measurement). Second, what the optimum way is to use those measurements to assure maximum profitability for a process. With the correlation data presented in the previous section, we will address the first question.

The process objective specifications require that gate oxide thickness and poly CD be monitored in the process flow. The goal is to monitor these parameters to assure that the devices meet the specifications agreed to with the circuit design engineers. Since L_{poly} and T_{ox} are uncorrelated, (1) (ρ is the correlation coefficient) can be used to compute the percentage of device characteristic variance that can be explained via in-line measurements. The results are summarized in Table IV.

$$\% \text{ Variance} = \rho^2_{L_{poly}} + \rho^2_{T_{ox}}. \quad (1)$$

After the gate oxidation in-line is measured, a large percentage of the variance of V_{tn} has been determined. Hence, this in-line can be used to control V_{tn}. The poly CD measurement only accounts for 58% of I_{dn}, but if this information were used along with the T_{ox} in-line, 85% of the I_{dn} variation could be predicted.

The observability for the p-MOS transistor is not as good. Only 59% of the I_{dp} variation and 0.4% of the V_{tp} variation can be observed from measuring the two in-lines. This is because a large amount of the p-MOS transistor's variation is controlled by variations in the buried channel dopant con-



g. 5. n-MOS and p-MOS $I_{d\,max}$ correlation plot.

entration. This can be seen by examining the correlation of $_{tp}$ and I_{dp} with the N_{well} sheet resistance. Since an in-ne measurement of the $N_{well}R_{sh}$ after gate oxidation is not vailable, it is not possible to monitor V_{tp} and I_{dp} using n-line measurements. The only way to guarantee that these wo device characteristics are within the specification limits s to explicitly measure them at electrical test. Hence, quality ssurance can only be achieved via direct testing late in the rocess.

3) Application to Sigma-Based Circuit Models: Typically, worst-case models are released to circuit designers which epresent the corners of the process. Corners are defined as egions of the fabrication process where circuits are likely o fall out of specification. For the 0.8-μm process, each ransistor is released with two worst-case models, weak and strong. The weak transistor has both a high (absolute value) threshold voltage and low drive current while the strong transistor has high drive current and low threshold voltage. The n and p models are typically treated as being independent. This implies that the fabrication process can produce weak transistors of one type and simultaneously strong transistors of the opposite type.

By examining the predicted correlations, we see that the n and p transistors are significantly related. Hence, the strong–weak cases are unlikely (see Fig. 5). From the distribution points were selected to represent realistic worst-case or sigma-based corners. For example, the corner which has the strongest n transistor also had the strongest p transistor. The selected sigma-based drive current corners are marked with circles on Fig. 5. Typically, it is assumed that the maximum drive current transistor also has a low threshold voltage while the minimum drive transistor has a high threshold voltage. The implied assumption is that the correlation between V_t and I_d is 1.0. As shown in Figs. 6 and 7, this is not the case. On both graphs we have drawn a line. If the assumption were correct, all data points would have to lie on these lines. In Fig. 6, two cases were circled where the transistor has typical drive current, but the threshold voltages are close to opposite ends of the distribution. BSIM models were generated at the corners circled in Figs. 6 and 7.

4) Conclusion of Simulation Phase: At the end of the simulation phase, a number of conclusions can be drawn. The following has been learned in doing this work:

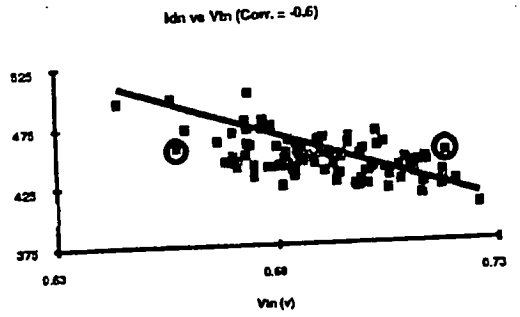


Fig. 6. I_{dn}–V_{tn} correlation plot.

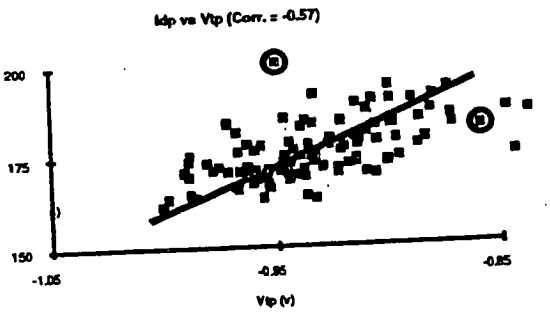


Fig. 7. I_{dp}–V_{tp} correlation plot.

- The 0.8-μm process in-lines do not run at specification, but devices meet specification.

Poly CD varies more than the specification while T_{ox} varies less. As a result, the I_{dn} and I_{dp} are more correlated than they would be if the process were in specification. The V_{tp}–I_{dp} and V_{tn}–I_{dn} are less correlated as a result.

- In-line measurements provide inadequate observability to control the process.

While T_{ox} and poly CD can be used to predict V_{tn} and I_{dn}, they do not provide enough observability to control V_{tp} and I_{dp}. This is because the p-MOS transistor is sensitive to the variations in the dopant distributions in the channel. Without monitoring this aspect of the process, it will not be possible to use SPC to control the device performance.

- Strong–weak combinations are unrealistic worst case corners.

Due to the common dependency on linewidth variations, the n and p drive currents are correlated, while I_d–V_t correlations are less than expected and V_{tn} and V_{tp} are not correlated at all. The large correlation between I_{dp} and I_{dn} should be used by designers while the lack of correlation between I_d and V_t is not covered in existing CAD model files.

V. Multivariate Analysis and Modeling of SPICE Parameters in Terms of Electrical Test Measurements

As described in the overview of the paper, a two-stage approach to DFM is used. In phase 1, we use statistical TCAD simulations with physically base SPICE parameter extraction to predict the distributions of device characteristics, electrical tests, and SPICE models. As demonstrated in the previous example, the results of simulation, in and of themselves provide valuable information.

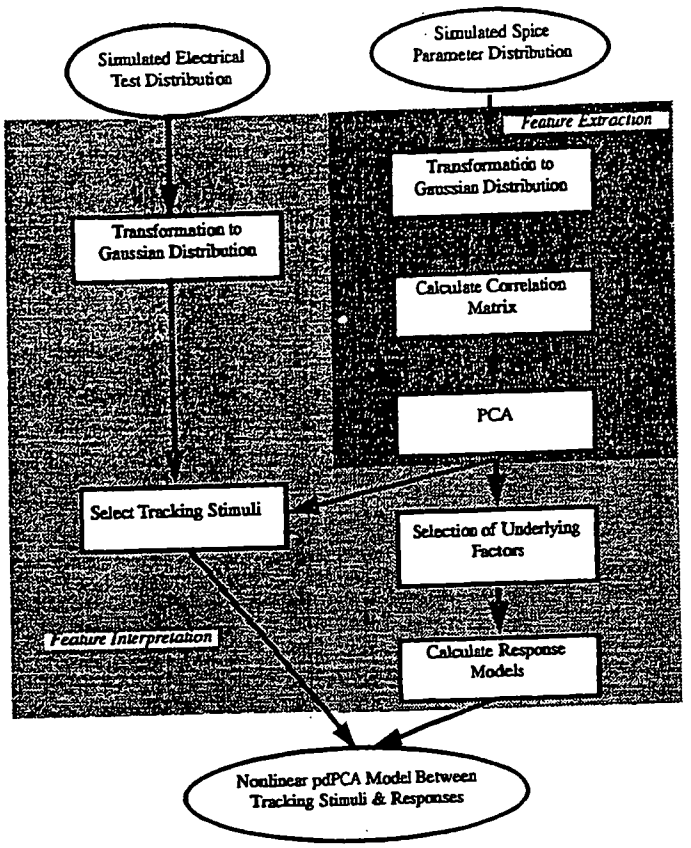


Fig. 8. Algorithm used to determine the relationship between electrical test measurements and SPICE parameters.

However, as in most production situations, a great deal of electrical test data is collected. From this data several things can be determined: how many sigma-based worst case corners there are for the process, which electrical tests are necessary to control each corner, and the functional relationship between the electrical tests and the SPICE model parameters.

Since the true independent variables in the process are process controls, one cannot assume that the electrical test measurements (our view into the process state) are independent variables. Moreover, the relationship between the SPICE parameters and electrical tests can be nonlinear.

To account for these two effects, a pattern recognition approach based on work first described in [5] is used. As shown in Fig. 8, the algorithm has both a feature extraction and interpretation phase.

A. *Feature Extraction: Determining the Number of Underlying Variables*

It is well-known that many device and SPICE parameters are correlated [2], [3], [6]. Often these correlations stem from a common underlying process step. For example, both subthreshold current and maximum drive current are dependent on poly CD and gate oxide thickness variations. Hence, it is expected that these two device characteristics will be correlated.

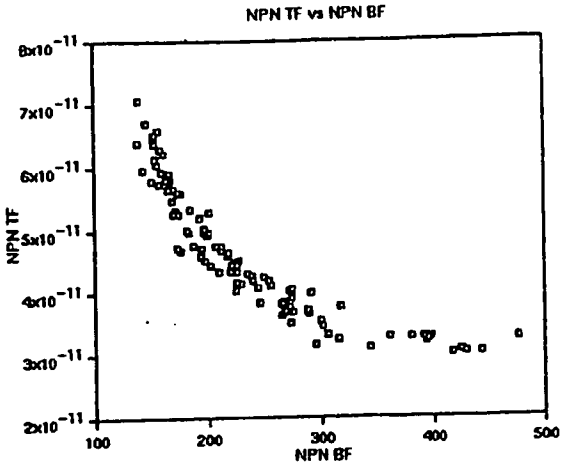


Fig. 9. Scatter plot of B_F versus T_F for a poly emitter n-p-n. Notice that although the two parameters are correlated, the correlation is nonlinear.

Since many of the SPICE parameters are correlated, it is unrealistic to treat them as independent variables [1]. For example there are 34 Gummel–Poon parameters, but there are not 2^{34} possible worst-case corners. Techniques such as principal component analysis (PCA) analyze the correlation matrix of variables to determine the number of independent factors [7]. These independent factors can be thought of as the "features" of the device.

It is important to remember that a correlation between two parameters does not imply that there is necessarily a cause and effect relationship. Hence, analyzing correlations of corrupted data sets must be undertaken carefully [5], [8]. Moreover, when using PCA, one must remember that it incorporates the following assumptions:

- The data used to compute the correlation matrix are sampled from a single distribution. This is often referred to as the assumption of independently and identically distributed (IID). Using Monte Carlo simulation results rather than measured data satisfies this assumption.
- The relationship between two variables can be represented by a linear correlation coefficient.

Due to the complex physical relationships between the underlying device structure and the SPICE parameters, the relationship between two SPICE parameters is rarely linear (see Fig. 9). As a result, the correlations between SPICE parameters can often not be represented adequately with linear coefficients.

Before computing the correlation matrix of the SPICE parameters, nonlinear transformations are applied to assure that a linear correlation coefficient is a good measure of the interactions between any pair of *transformed* SPICE parameters. It is well-known that if two variables have univariate Gaussian distributions, then the joint distributions is also Gaussian. Since a linear correlation coefficient is sufficient to represent the interaction between two random variables that have Gaussian distributions, then by transforming the data to be Gaussian we can assure that the linear correlation coefficient is sufficient.

For most devices, between three and eight factors are sufficient to account for most of the variability in the SPICE models.

B. Feature Interpretation: Building Relationships Between Test Measurements and SPICE Parameters

As stated in the previous section, PCA is used to determine the minimum number of uncorrelated, independent factors that can represent the variability in SPICE model parameters. However, traditional PCA alone cannot determine the physical measurements needed to track each of the independent factors.

Rather than identifying the results of traditional PCA as a set of empirical factors, a physical interpretation for them is useful. Knowing the physical interpretation would enable tracking of how changes in the process are impacting circuit performance. To determine these relationships, a matching algorithm is used to look for similarities between the principal components and E-test or in-line measurements [5].

From a traditional PCA analysis of a BSIM3 model, it is found that there is a factor which accounts for 60% of the variance in C_{GDO} (drain-gate overlap capacitance) and 60% of V_{FB} (flat band voltage). However, the traditional PCA algorithm does not provide any insight into the physical phenomena leading to this result. By applying a matching algorithm, physical insight is gained and an improved explanation of SPICE model variability results.

The matching algorithm looks for similarities between electrical test measurements and principal components. The matching is possible because the electrical tests and components can be compared to each other based on their common correlations to the SPICE parameters. This is achieved without including the electrical test measurements in the original principal component analysis. As can be seen in Table V, the matching algorithm has selected T_{ox} as a tracking parameter. And, the explained variance has increased to about 79% for C_{DG0} and about 75% for V_{FB}.

From this enhanced PCA approach, we can determine the following:

- The minimum set of in-line and E-test measurements which can be used to track statistical variations in the SPICE model parameters.
- A nonlinear functional relationship between the electrical test measurements and SPICE model parameters.

TABLE V
PERCENTAGE OF NFET BSIM3 V2 SPICE PARAMETER VARIANCE MONITORED BY IN-LINES

	Factors				
	Tox	Lpoly	NLdd Conc	N+ Rsh	Sum
Average	28%	25%	14%	7%	74%
PHI	61.1%	0.0%	5.6%	1.8%	68.5%
K1	78.1%	1.3%	0.6%	0.0%	80.0%
K2	75.7%	1.0%	1.1%	0.0%	77.8%
DVT1	31.3%	0.6%	57.7%	2.1%	91.9%
DVT2	61.1	0.0%	5.6%	1.8%	68.6%
DVT0	0.0%	92.9%	4.1%	0.0%	97.0%
VFB	74.6%	3.0%	0.2%	0.7%	78.4%
A0	29.6%	11.0%	7.2%	0.1%	47.9%
KETA	0.5%	94.3%	0.9%	0.0%	95.9%
U0	61.4%	9.7%	11.7%	0.5%	83.3%
UA	26.5%	55.6%	9.1%	0.1%	91.4%
UB	29.3%	31.4%	2.3%	3.4%	66.4%
UC	0.2%	64.1%	5.0%	0.4%	69.8%
ETA	20.0%	26.0%	29.0%	0.1%	75.1%
PCLM	0.4%	29.2%	0.7%	0.1%	30.4%
DROUT	31.3%	0.6%	57.7%	2.1%	91.9%
PDIBL1	1.8%	58.8%	0.3%	1.1%	62.0%
PSCBE2	13.2%	11.6%	0.3%	0.6%	25.7%
DSUB	31.3%	0.6%	57.7%	2.1%	91.9%
CDSC	0.9%	96.4%	0.1%	0.1%	97.4%
CDSCB	0.1%	90.1%	0.4%	0.0%	90.6%
ETA0	0.2%	93.7%	2.5%	0.1%	96.4%
ETAB	0.0%	43.0%	13.5%	2.0%	58.6%
VOFF	32.7%	48.5%	7.1%	0.6%	88.9%
VGHIGH	62.5%	20.8%	8.8%	1.0%	93.1%
VGLOW	0.1%	97.6%	0.0%	0.0%	97.8%
DL	53.1%	0.0%	3.8%	2.1%	59.0%
CGD0	78.8%	0.0%	6.3%	0.3%	85.5%
CGS0	78.8%	0.0%	6.3%	0.3%	85.5%
CJ	15.78	2.2%	47.0%	0.1%	65.1%
MJ	2.6%	0.1%	33.9%	13.1%	49.7%
CJSW	51.2%	0.1%	19.7%	0.1%	72.3%
MJSW	10.4%	0.2%	0.6%	48.3%	59.4%
PB	14.0%	1.8%	51.9%	0.3%	68.0%
PBSW	0.0%	0.0%	0.3%	54.7%	55.0%
JS0	0.0%	1.1%	4.3%	68.1	73.0%
XJ	6.7%	1.2%	77.4%	3.7%	89.0%

VI. EXAMPLES OF MULTIVARIATE ANALYSIS

As described above, we have simulated many processes. Below, the analysis of a submicrometer CMOS process and a high-speed polyemitter complementary BJT process are described. These simulations were analyzed with pdPCA [14] to determine the appropriate in-lines and E-test parameters to control the SPICE model parameters.

A. Example 1: Submicrometer CMOS Process

There are commonly two perspectives from which to view the device variability: in-line or E-test measurements. In-line measurements usually provide structural information about the device, e.g., L_{poly}, T_{ox}. However, these measurements often do not provide the resolution necessary for statistical SPICE model prediction. E-tests, e.g., I_{ds}, V_{th}, *HFE*, are more reliable measurements, but these measurements can often confound effects. For example, variations in V_{th} can depend on T_{ox}, L_{poly}, sidewall spacer width, channel doping, and many other possible causes.

In the simulation work for the CMOS process, BSIM [1] model parameters were extracted. For this part of the paper, we have extracted parameters for the BSIM3 [12] model. The BSIM3 model has much less confounding of physical parameters such as L and T_{ox}. Hence, it is better suited for statistical analysis.

We had two goals for this part of our work. First, determine which in-line measurements are needed to control the SPICE parameters for the device. Second, determine if E-test measurements could be used in place of the in-line measurements. While in-line measurements are a more basic set of parameters, the E-tests are more reliably collected. Moreover, the sample size of E-test measurements is typically much larger.

The in-line measurements, L_{poly} and T_{ox}, were found to be the two most dominant tracking parameters. Collectively,

L_{poly} and T_{ox} in-lines can explain the variability in many of the BSIM3 parameters which control the device threshold (D_{VT0}), body effect (K_1), and mobility (U_A) as shown in Table V. However, to monitor the source and drain electrical characteristics, e.g., capacitance (C_J), and leakage current (P_B, J_{S0}), n^+ sheet resistance and $N_{L_{dd}}$ concentration measures are needed.

Unfortunately the n^+ R_{sh}, and $N_{L_{dd}}$ measurements are not currently measured as in-lines for this process. Clearly, the absence of these in-line measures hinders the ability to predict and then control the SPICE model parameters. While measuring S/D n^+ R_{sh} in-line is possible, it is not desirable. Moreover, in this process there was no valid E-test data from the $N_{L_{dd}}$ test structure. The L_{dd} region has an impact on some of the short channel effects, in particular D_{VT1}, D_{ROUT}, and E_{TA}. To use in-line measurements to control the devices during production, two additional in-lines and possibly an additional mask are needed. This is clearly not feasible.

Instead of the costly addition of in-line measures and mask layers, electrical tests can be used to monitor the BSIM3 parameters during production. The working hypothesis was that with all of the threshold voltage and drive current measurements made, electrical test measurements would accurately span the range of device variability. To test this hypothesis, pdPCA [14] was used to analyze the BSIM3 model with E-tests as the possible controlling factors. As in the last analysis, the third and fourth factors were the $N_{L_{dd}}$ and n^+ R_{sh} (see Table VI). However, the first and second factors were maximum drain current and the drain current when the device is operated in the linear region. While these factors are different than L_{poly} and T_{ox}, they can also span the variability in the SPICE parameters. The reader should be aware that when summed, the first two factors for this model account for the same variability as the first two factors in Table V. However, individually, these factors ($I_{dn\,max}$, $I_{dn\,lin}$) are different from the two factors (T_{ox}, L_{poly}) in the previous model.

The importance of this result is that the possible tracking measurements can be changed and it can be determined if adequate coverage of the device variability is attained. For example, if the $N_{L_{dd}}$ concentration is not monitored, a significant portion of the variance of parameters such as X_J, D_{VT1}, P_B, etc., cannot be assured to be in control. It appears that additional test measurements are needed to control the statistical variations in SPICE model parameters for the devices.

B. Example 2: High Speed BJT Process

To show the generality of our approach, results are presented from analysis of a high speed, polyemitter complementary BJT process. Due to the highly nonlinear dependencies, the resulting distributions of SPICE parameters are rarely Gaussian. Fig. 9 demonstrates that the predicted distribution of B_F (beta) versus T_F (unity gain frequency) for the n-p-n transistor has an extremely nonlinear correlation.

In a manner similar to the first example, statistical simulations were run and the results were used in the multivariate analysis. In this case, it was also found that four significant factors explained most of the SPICE parameter variability. (It is only a coincidence that the number of factors is the same for both the n-FET and the n-p-n.) The matching algorithm found four tracking stimuli which can be easily collected at wafer electrical test (see Table VII). The first factor is clearly the most important. Intrinsic base sheet resistance or peak beta can be used to monitor this factor. The second factor explains the variability in the collector to substrate capacitance and resistance (C_{JS}, V_{JS}, M_{JS}, and R_C), and is represented by the collector/substrate built in voltage. The third factor explains the variability in the SPICE parameters which depend on emitter area variations while the fourth factor tracks to extrinsic base R_{sh} variations.

TABLE VI
PERCENTAGE OF NFET BSIM3 V2 SPICE PARAMETER VARIANCE MONITORED BY E-TESTS

	Factors				
	Id Max	Id,Vd <Vt	NLdd Conc	N+ Rsh	Sum
Average	32%	21%	14%	7%	74%
PHI	36.6%	24.6%	5.6%	1.8%	68.5%
K1	39.1%	40.3%	0.6%	0.0%	80.0%
K2	38.6%	38.2%	1.1%	0.0%	77.8%
DVT1	23.9%	8.1%	57.9%	2.1%	91.9%
DVT2	36.6%	24.6%	5.6%	1.8%	68.6%
DVT0	36.3%	56.8%	4.1%	0.0%	97.0%
VFB	32.7%	45.0%	0.2%	0.7%	78.4%
A0	40.2%	0.6%	7.2%	0.1%	47.9%
KETA	43.2%	51.7%	0.9%	0.0%	95.9%
U0	65.4%	5.8%	11.7%	0.5%	83.3%
UA	75.0%	7.2%	9.1%	0.1%	91.4%
UB	59.7%	1.1%	2.3%	3.4%	66.4%
UC	21.4%	43.1%	5.0%	0.4%	69.8%
ETA	44.5%	1.54%	28.9%	0.1%	75.1%
PCLM	8.1%	21.6%	0.7%	0.1%	30.4%
DROUT	23.9%	8.1%	57.7%	2.1%	91.9%
PDIBLI	33.7%	27.0%	0.3%	1.1%	62.0%
PSCBE2	24.8%	0.2%	0.3%	0.6%	25.7%
DSUB	24.0%	8.1%	57.7%	2.1%	91.9%
CDSC	46.3%	51.0%	0.1%	0.1%	97.4%
CDSCB	32.4%	57.8%	0.4%	0.0%	90.6%
ETA0	39.8%	54.0%	2.5%	0.1%	96.4%
ETAB	17.7%	25.6%	13.4%	2.0%	58.6%
VOFF	77.5%	3.8%	7.1%	0.6%	88.9%
VGHIGH	81.7%	1.7%	8.7%	1.0%	93.1%
VGLOW	40.4%	57.3%	0.0%	0.0%	97.8%
DL	32.7%	20.6%	3.7%	2.1%	59.0%
CGD0	47.6%	31.4%	6.3%	0.3%	85.5%
CGS0	47.6%	31.4%	6.3%	0.3%	85.5%
CJ	4.9%	13.1%	47.0%	0.1%	65.1%
MJ	1.2%	1.44%	34.0%	13.1%	49.7%
CJSW	29.3%	22.2%	19.7%	0.1%	72.3%
MJSW	7.9%	2.7%	0.6%	48.3%	59.4%
PB	4.5%	11.4%	51.9%	0.3%	68.0%
PBSW	0%	0%	0.3%	54.7%	55.0%
JS0	0%	0.9%	4.3%	68.1%	73.0%
XJ	7.3%	0.6%	77.4%	3.7%	89.0%

VII. CONCLUSIONS

A method has been shown which combines TCAD simulation with statistical analysis of empirical data. As a result, sigma-based models can be determined before a significant sample size of devices can be characterized.

TABLE VII
PERCENTAGE OF n-p-n BJT–GP SPICE PARAMETER VARIANCE MONITORED BY E-TESTS

	Factors				
	Int Base Rsh	Col/ Sub Vbi	Emit Area	Ext Base Rsh	Sum
Average	49%	15%	12%	10%	86%
IS	89.4%	0.0%	2.8%	7.2%	99.5%
BF	93.0%	0.2%	0.1%	3.2%	96.5%
IKF	13.5%	62.1%	6.3%	4.4%	86.3%
IKR	90.0%	1.7%	0.1%	0.8%	92.6%
ISE	59.5%	2.6%	8.2%	3.7%	74.1%
ISC	1.9%	0.4%	2.3%	0.6%	5.2%
IRB	69.3%	19.4%	3.7%	0.0%	92.4%
RBM	1.2%	12.9%	5.3%	42.3%	61.7%
RB	89.2%	0.0%	3.6%	7.0%	99.7%
RE	50.1%	5.2%	36.2%	1.5%	93.0%
RC	0.2%	78.4%	8.2%	1.5%	88.3%
CJE	5.8%	15.3%	64.1%	2.7%	87.9%
VJE	62.6%	0.5%	1.9%	0.2%	65.1%
MJE	52.8%	2.0%	8.0%	0.1%	62.9%
TF	97.7%	0.7%	0.0%	0.9%	99.3%
TR	91.6%	4.6%	0.0%	1.4%	97.7%
CJC	23.8%	0.8%	1.5%	51.9%	78.0%
VJC	28.7%	0.6%	9.2%	43.7	82.2%
MJC	46.0%	0.1%	1.7%	45.5%	93.3%
CJS	1.0%	77.3%	3.3%	0.0%	81.6%
VJS	1.0%	82.4%	8.0%	2.6%	93.9%
MJS	1.0%	80.1%	7.8%	3.7%	92.7%
BR	94.6%	0.24%	0.2%	3.2%	98.3%
VAF	97.3%	0.2%	0.1%	0.3%	97.9%
VAR	73.2%	0.3%	0.1%	0.2%	73.8%
NE	92.9%	0.4%	0.0%	3.6%	96.9%
NC	43.3%	0.0%	4.3%	47.3%	94.9%
XTF	53.2%	2.2%	37.6%	5.1%	98.2%
VTF	55.1%	4.1%	0.0%	5.8%	65.1%
ITF	17.0%	5.5%	69.7%	5.2%	97.3%
PTF	89.3%	1.0%	0.2%	8.0%	98.5%
XCJC	0.1%	13.3%	77.5%	4.7%	95.6%

Beyond generating sigma-based models from simulation, algorithms have been developed to use simulated data to analyze measured in-line and E-test data. Finally the minimum set of E-test or in-line measurements to control the SPICE model parameters were determined.

A direct result of this work is the ability to tie together the marketing, manufacturing, and design communities. The mathematical models which span in-line, E-tests, and SPICE model parameters can effectively communicate to design how changes in the process will effect them since they can see the impact in the SPICE model parameters. Moreover, process development can design recipes which lead to agreed-upon objective specifications, and manufacturing can see how variations in the process are affecting design yields. And lastly, marketing can see how critical customer requirements are being met and controlled.

ACKNOWLEDGMENT

The authors would like to express sincere thanks to J. P. Downing for his insights and the phrase "sigma-based models." They would also like to thank C. Skinner for his early support of this work, and E. Mueller, J. Giddings, H. S. Chen, and D. B. Kao for their assistance with the collection and screening of the empirical data.

REFERENCES

[1] D. Hanson *et al.*, "Analysis of the controllability of a sub-micron CMOS process using TCAD," in *Ext. Abst. 1994 Int. Symp. Semicond. Manufact.*, Tokyo, Japan, June 21–22, 1994, pp. 85–89.
[2] P. Yang *et al.*, "An integrated and efficient approach for MOS VLSI statistical circuit design," *IEEE Trans. Computer-Aided Design*, vol. CAD-5, no. 1, pp. 5–15, 1986.
[3] S. Nassif *et al.*, "A method for worst-case analysis of integrated circuits," *IEEE Trans. Computer-Aided Design*, vol. CAD-5, no. 1, pp. 114–131, 1986.
[4] S. Nassif, A. J. Strojwas, and S. W. Director, "FABRICS-II: A statistically based IC fabrication process simulator," *IEEE Trans. Computer-Aided Design*, vol. CAD-3, no. 1, pp. 40–46, 1984.
[5] J. Kibarian, "Statistical diagnosis of IC process faults," Ph.D. dissertation, CMUCAD 90-52, Dec. 1990.
[6] S. L. Sundaram and A. C. Carlson, "Reducing fabrication variability in analog IC technology by the statistical error propagation method using simple test structures," in *Proc. 1990 Int. Conf. Microelectr. Test Struct.*, San Diego, CA, 1990.
[7] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
[8] T. W. Anderson, "Asymptotic theory for principal component analysis," *Ann. Math. Statist.*, vol. 34, pp. 122–148, 1963.
[9] W. Maly, A. Strojwas, and S. Director, "Fabrication based statistical design of monolithic IC's," in *IEEE Int. Symp., Circuits and Systems*, Apr. 1981, vol. 1, pp. 135–138.
[10] E. D. Boskin *et al.*, "IC performance prediction from electrical test measurements," in *Proc. ISMSS*, June 1992, pp. 13–17.
[11] H. K. Gummel and H. C. Poon, "An integral charge control model of bipolar transistors," *Bell Sys. Tech. J.*, June 1970.
[12] J. H. Huang *et al.*, *BSIM3 Manual*, vol. 2, the Regents of the University of California, Berkeley, CA 94720.
[13] *pdFab Users Manual*, vol. 2, ch. 2, PDF Solutions, Pittsburgh, PA, 1993.
[14] *pdPCA Users Manual*, vol. 1, ch. 1, PDF Solutions, Pittsburgh, PA, 1994.
[15] *SUPREM3 Users Manual*, vol. 6, ch. 11, Technology Modeling Associates, Palo Alto, CA, 1993.
[16] *Medici Users Manual*, Technology Modeling Associates, Palo Alto, CA, 1993.
[17] M. S. Obrecht, "SIMOS—A two-dimensional steady-state simulator for MOS devices," *Solid State Electron.*, vol. 34, no. 7, 1993.

David A. Hanson (A'85) received the B.S. degree in chemistry from Chico State University, Chico, CA, in 1978.

In 1979, he joined the technical staff, Fairchild Semiconductor Research Center, Palo Alto, CA, where he worked on charge-coupled device fabrication, and conducted basic studies on the plasma chemistry of halocarbon discharges. From 1982 to 1987, he held several supervisory positions in the area of memory device development and fabrication. In 1988, he began a three-year assignment at SEMATECH, a semiconductor manufacturing research consortium in Austin, TX, where he managed 0.5-μm process development (DFM). He has published more than 15 conference and journal papers and has applied for one patent. He is presently Director of Business Process Improvement for National Semiconductor's product development process.

Ronald J. G. Goossens (M'91) was born and raised in The Netherlands. He received the M.S. degree in particle physics and astronomy in 1980 and the Ph.D. degree in solid-state physics in 1984, both from Utrecht State University, the Netherlands.

After five years in Device Physics with Philips Research Laboratories, Eindhoven, The Netherlands, he joined Stanford University, Stanford, CA, in 1989, initially as an Industrial Visitor and later as a Senior Research Scientist. In 1993, he joined National Semiconductor as Modeling Manager for the Analog Division. From 1994 to 1996, he served on the SEMATECH FTAB for TCAD, the last two years as Industrial Chair. Since June 1996, he has been assigned to the SRC as Director of the Center for Semiconductor Modeling and Simulation (a CRADA between SRC and the National Labs) and as Director of TCAD Sciences.

Dr. Goossens presently serves on the IEDM Program Committee for Modeling and Simulation and on the TCAD-TWG for the 1997 National Technology Roadmap for Semiconductors.

Mark Redford received the B.S. (honors) degree in electronics from the University of Dundee, Scotland.

He has been with National Semiconductor (U.K.) Ltd., Greenock, Scotland, since 1983 and his background is in product device engineering. He now holds the position of Process Technology Development Manager looking at bipolar, CMOS, DMOS, and BiCMOS technologies, down to 0.8 μm, for the analog marketplace. His interests include design for manufacturability involving TCAD and design of experiments. He has published 12 papers to date on the subject of design for manufacturability.

Jim McGinty received the B.S. (honors) degree in electronics from the University of Paisley, Scotland, and the M.S. degree in VLSI design from Bournemouth University, U.K.

He has been with National Semiconductor (U.K.) Ltd., Greenock, Scotland, since 1991 and his background is in TCAD and device physics. He now holds the position of Modeling Team Leader in the Analog Process Technology group. His interests include design for manufacturability, involving statistical TCAD and DOE. He has published six papers to date on the subject of TCAD.

John K. Kibarian (S'83–M'91) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 1986, 1988, and 1991, respectively.

In 1991, he founded PDF Solutions, San Jose, CA, to provide services and software for IC yield improvement. Presently, he is President of PDF Solutions. Prior to his current position, he held summer positions at Texas Instruments, Cherry Semiconductor, and General Motors. His research interests include prediction and diagnosis of IC yields.

Dr. Kibarian received the best paper award from IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING in 1991 for his paper entitled, "Using spatial information to analyze correlations between test structure data." He is a member of Eta Kappa Nu and Tau Beta Pi.

Kimon W. Michaels (S'85–M'92) received the B.S. degree in electrical engineering in 1987, and the M.S. and Ph.D. degrees in electrical and computer engineering in 1989 and 1993, all from Carnegie Mellon University, Pittsburgh, PA.

He is currently Vice President, PDF Solutions, Inc., San Jose, CA, and has previously held summer positions at Texas Instruments and Westinghouse. His research interests include statistical SPICE model characterization and semiconductor manufacturability analysis.

Dr. Michaels is a member of Eta Kappa Nu and Tau Beta Pi.

Application of Covariance Based Models to Fit Response Surfaces to Experimental Data

M. Redford[®], A.J. Walton, D. Sprevak[®], R.S. Ferguson[®]

Department of Electronics and Electrical Engineering
Kings Buildings
University of Edinburgh
Edinburgh
EH9 3JL

[®]Department of Electronics and Electrical Engineering
Ashby Building
Queens University
Belfast

[®]National Semiconductor
Larkfield Industrial Estate
Greenock, PA16 OEQ

Abstract

This paper presents a method of fitting a response surface to experimental data when there are one or more data points that are poorly fitted by conventional polynomial models. The method is based on first fitting the data with a polynomial model and using this to calculate a worksheet for the combinations of control factors that were used in the original experiment. The actual experimental conditions for the poorly fitting points are then substituted into this worksheet and a covariance fit used to fit the data. The resulting surface follows the general trend while also fitting measurement points where there is confidence that there is no significant experimental error.

Introduction

Experimental design together with Response Surface Methodology (RSM) are important tools that can be employed to help optimise IC processes [1]. A key to the success of this approach is the accuracy with which the model can be fitted to the data. One of the problems with fitting experimental measurements is that there is uncertainty associated with each data point. Consequently, unlike the situation with simulated data, there is no justification in fitting the response surface through all the measurement points [2]. Therefore the trend is typically fitted using a polynomial based function which effectively smoothes the response if experimental error is present [3]. This paper presents a method of improving the fit of response surfaces for experimental data when one or more data points are poorly fitted and there is confidence in the fact that these points represent reality.

Background

As mentioned previously there are occasions when a low order polynomial cannot adequately fit one or two data points that are known not to be rogue measurements. Transformations are one method that is traditionally used to improve the fit but, if this is not successful, the only other option is to increase the order of the polynomial. This latter option can, in the limit, result in the response surface ceasing to fit the trend. Another possible solution is to use a covariance based fit which ensures that the resulting function passes through all the data points [4]. This is a quite acceptable approach for computer based experiments as simulation results have no experimental error associated with them (the same results are obtained every time every time the program is run). However, this approach is inappropriate for experimental data because it will fit the surface to the errors when only a fit to the overall trend can be justified.

Application of Covariance Fits to Experimental Data

The solution proposed in this paper is a two stage process. First the data is fitted with a polynomial and the problem points identified. Secondly, if it is decided that these points are not rogue measurements then a hybrid polynomial-covariance [4,5] fit is employed.

In detail the new response data is created as follows. First a quadratic polynomial is fitted in the normal manner using a least square fit. Problem points, for which the residuals are

excessive, must then be identified and it confirmed that these actually represent reality and are not simply erroneous measurements. A new response set of data is created by adding the outliers to the worksheet with the rest of the required data being generated from the polynomial fitted to the original dataset. The effect of this is to ensure that the trend is fitted while at the same time distorting this response surface through the data points that were deemed to be poorly fitted by the original quadratic.

Run no	Slope (°)	Dose (cm^{-2})	Energy (keV)	R_{SD} (Ω)
1	70	1e+13	30	1420
2	90	1e+13	30	760
3	110	1e+13	30	425
4	70	5.05e+14	30	1110
5	90	5.05e+14	30	620
6	110	5.05e+14	30	315
7	70	1e+15	30	1020
8	90	1e+15	30	490
9	110	1e+15	30	245
10	70	1e+13	60	1170
11	90	1e+13	60	720
12	110	1e+13	60	415
13	70	5.05e+14	60	1090
14	90	5.05e+14	60	600
15	110	5.05e+14	60	305
16	70	1e+15	60	1000
17	90	1e+15	60	480
18	110	1e+15	60	225
19	70	1e+13	90	1150
20	90	1e+13	90	700
21	110	1e+13	90	405
22	70	5.05e+14	90	1050
23	90	5.05e+14	90	590
24	110	5.05e+14	90	295
25	70	1e+15	90	940
26	90	1e+15	90	475
27	110	1e+15	90	210

Table 1. Worksheet for 3 factor 3 level full factorial experiment.

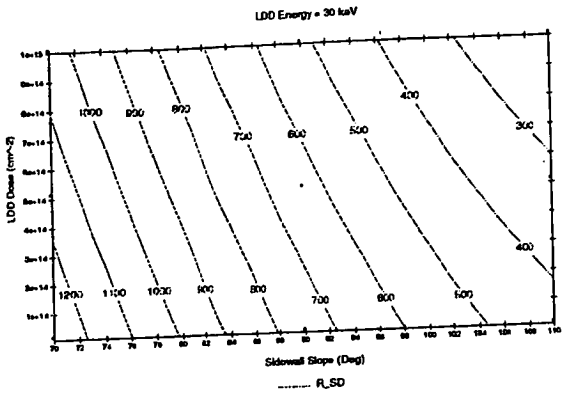
This procedure is best illustrated through an example. The worksheet for a full factorial experiment with three control factors and a single response is shown in table 1. When the response is fitted with a second order polynomial the Adj R^2 value is 0.989. Table 2 presents information on how well each data point is fitted. Taking the traditional approach, the error in the fit may be reduced by applying a transformation. In this case a square root transformation is recommended and this increases the Adj R^2 value to 0.994. Figure 1 shows the response surface resulting from the polynomial fit and it can be observed that it does not reach 1420Ω when the dose, energy and sidewall slope are 1e13 $atoms^{-2}$, 30keV and 70° in run 1.

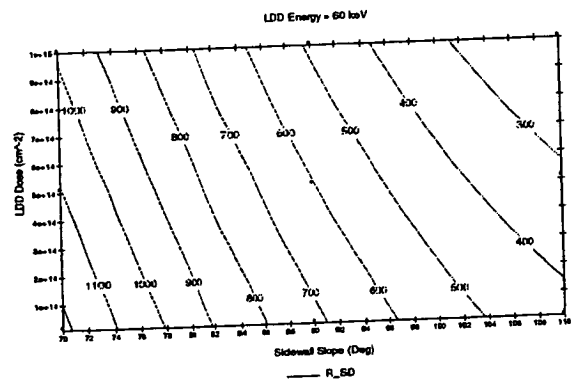
Run no	Expt Data	Poly RSM model	Absolute Residual Error	% Error
1	1420	1260	160.00	11.27
2	760	752.22	7.78	1.02
3	425	435.97	10.97	2.58
4	1110	1166.25	56.25	5.07
5	620	636.67	16.67	2.69
6	315	320.42	5.42	1.72
7	1020	1050.69	30.69	3.01
8	490	521.11	31.11	6.35
9	245	204.86	40.14	16.38
10	1170	1221.11	51.11	4.37
11	720	719.44	0.56	0.08
12	415	431.11	16.11	3.88
13	1090	1105.56	15.56	1.43
14	600	603.89	3.89	0.65
15	305	315.56	10.56	3.46
16	1000	990	10.00	1.00
17	480	488.33	8.33	1.74
18	225	200	25.00	11.11
19	1150	1160.42	10.42	0.91
20	700	686.67	13.33	1.90
21	405	426.25	21.25	5.25
22	1050	1044.86	5.14	0.49
23	590	571.11	18.89	3.20
24	295	310.69	15.69	5.32
25	940	929.31	10.69	1.14
26	475	455.56	19.44	4.09
27	210	195.14	14.86	7.08

Table 2. Residual errors for the polynomial model.

In this example there are three data points where the residual percentage error is in double figures (run 1, 9 and 18). For illustrative purposes in this example we make the assumption that only the data point associated with runs 9 and 18 have significant experimental error associated with them and there is high degree of confidence that the data associated with run 1 is not a rogue measurement. Hence, there is the requirement that create a response surface that passes through this point.

As previously proposed the first stage is to fit the data given in table 1 with a standard quadratic polynomial. This function is then used to generate the run table shown in table 3 for all data points





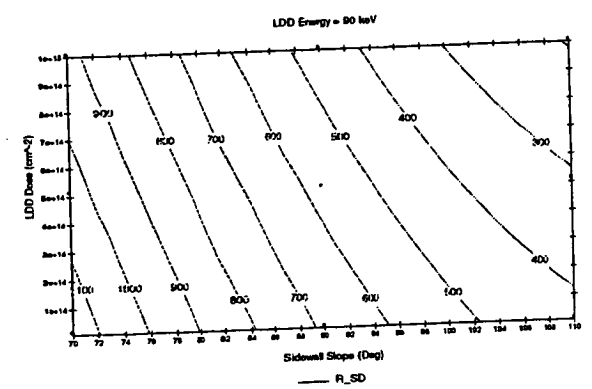
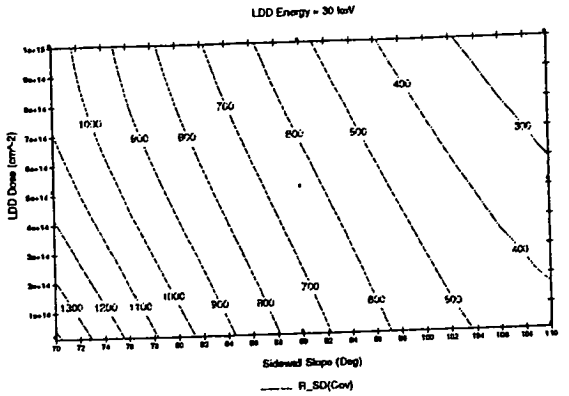
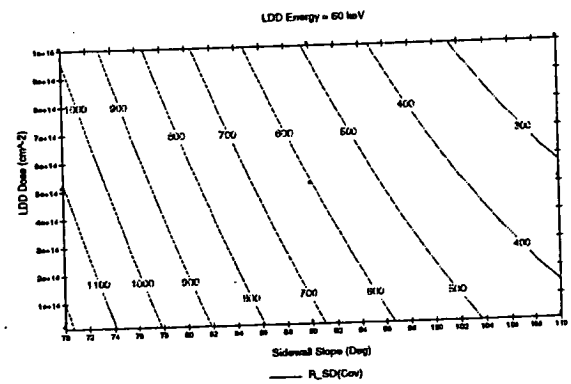


Figure 1. Contour plots of response surfaces for a second order polynomial fit to the data in table 1.





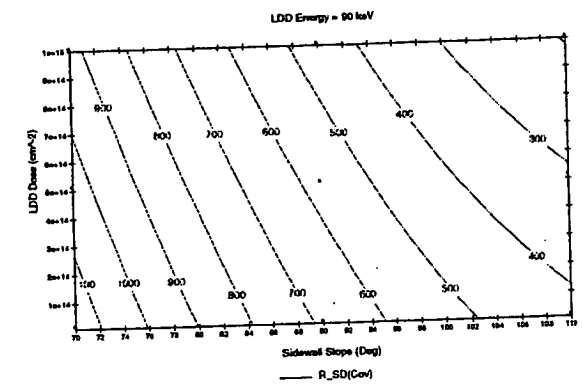


Figure 2. Contour plots of response surfaces for a covariance fit to the data in table 2.

except run 1 for which the original value of experimental measurement is entered. This data is now fitted using covariance which creates a response surface that passes through all the data points and the resulting contour plots are shown in figure 2. Comparison of figures 1 and 2 demonstrates how the covariance based response surface has only been distorted from the "polynomial based trend" at the bottom left-hand corner of the 30kev contour plot with the surface still fitting all the other data points (the polynomial surface).

It should be noted that if there was confidence that all three runs (1, 9 and 18) were just poorly fitted by the polynomial model (i.e. not due to experimental error) it is quite to select all three data points for the covariance fit. Figure 3 shows the results of ensuring that the response surface passes through all three points.

Run no	Original R_{SD} data (Ω)	Covariance R_{SD} data (Ω)
1	1420	1420
2	760	752.22
3	425	435.97
4	1110	1166.25
5	620	636.67
6	315	320.42
7	1020	1050.69
8	490	521.11
9	245	204.86
10	1170	1221.11
11	720	719.44
12	415	431.11
13	1090	1105.56
14	600	603.89
15	305	315.56
16	1000	990
17	480	488.33
18	225	200
19	1150	1160.42
20	700	686.67
21	405	426.25
22	1050	1044.86
23	590	571.11
24	295	310.69
25	940	929.31
26	475	455.56
27	210	195.14

Table 3. Modified worksheet for 3 factor 3 level full factorial experiment for the covariance fit.

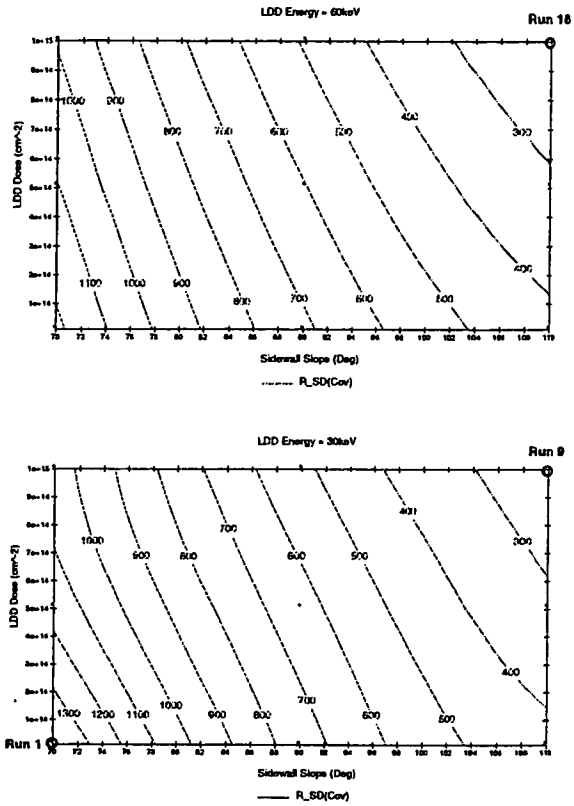


Figure 3. Contour plots which identify the three data points through which the response surface has been distorted. The 90keV plot is very similar to the one in figure 2 and is therfore not shown.

Conclusions

This paper has presented a method of adjusting response surfaces to ensure that they pass through selected data points. The procedure is designed to enable the method to be used on experimental data and removes the need for transformations enabling the user to precisely select which measurements are exactly fitted by the model. The practical application of the proposed approach has been presented as an example.

Acknowledgments

The authors would like to acknowledge the support of EPSRC, National Semiconductor, Mitel, Motorola, Avant! and Domain Solutions.

References

1. A.J. Walton, M. Fallon, M.I. Newsam, R.S. Ferguson, D. Sprevak, G.A. Allan, J.P. Elliott; "A Total TCAD Strategy for DFM in IC Technology Development", IEE Proc. Science, Measurement and Technology, 144, No 2, pp 63-68, March 1997.
2. G.J. Gaston, A.J. Walton; "The integration of simulation and response surface methodology for the optimisation of IC processes", IEEE Trans Semiconductor Manufacturing, 7, No 1, pp 22-33, Feb1994.
3. Montgomery, "Design and Analysis of Experiments", John Wiley , 1991.
4. Waring, A.J. Walton, D. Sprevak, R.S. Ferguson; "The Sensitivity of the Parameters of Covariance Based Response Surfaces", Third International Workshop on Statistical Metrology, pp 50-53, June 7, 1998.
5. Waring, A.J. Walton, R.S. Ferguson, D. Sprevak; "Application of Covariance Structures to Improve the Fit of Response Surfaces to Simulation Data", To be published in IEEE Trans Semiconductor Manufacturing.

Analysis of Mixed-Signal Manufacturability with Statistical TCAD

David A. Hanson, and Ronald J.G. Goossens
davidh@ampere.nsc.com, ronald@ammon.nsc.com
National Semiconductor Corporation, 2900 Semiconductor Drive, Santa Clara, CA 95052-8090
Mark Redford and Jim McGinty
mark@dundee.nsc.com, jim@shark.nsc.com
National Semiconductor Ltd., Larkfield Industrial Estate, Greenock PA 16 0EQ, Scotland
John K. Kibarian, and Kimon W. Michaels
jkk@pdf.com, kimon@pdf.com

Abstract

We have developed a method which combines TCAD simulation with statistical analysis of empirical data. As a result, we can determine worst case models before a significant sample size of fabricated devices can be characterized. As measured data is collected we use models built from simulated data to analyze measured in-line and e-test data. We determine what is the minimum set of e-test or in-line measurements to control the SPICE model parameters. As a result we can provide predictive worst case models, update them based on fabrication data as development proceeds, and finally use them to control the manufacturing line.

1. Introduction

The manufacture of integrated circuits (IC's) involves the design of both fabrication recipe and circuit netlist and layout. During the development phase there is little empirical data available upon which to base SPICE models for early chip design. And, during the production phase where model control is essential equipment fluctuations cause variations in the device and interconnect characteristics. In both phases it is necessary to understand how these variations impact product performance.

We refer to the de-sensitization of a product to process variations as "Design for Manufacturing" (DFM). Conversely we refer to the robust design (Cpk > 1 or 2) of fabrication processes as "Manufacturing for Design". Both approaches, design centering and manufacturing tolerance minimization are competitive imperatives in mixed-signal manufacturing environments.

DFM or MFD typically relies on a statistical characterization of the process [2][3][9]. Since traditional statistical characterization techniques are both costly and time consuming, it is usually undertaken only once, during process qualification.

Since both techniques rely on the collection of a great deal of empirical data, it is often the case that the models are developed after the initial products are designed, but while the process is still unstable. Consequently, approaches which use only empirical data provide statistical information too late to improve the yield for the newest product/process combinations. Moreover, since the characterization is not updated as the process matures (tolerances reduced by learning), new designs for the now older, more stable process are unnecessarily conservative. The result is a waste of expensive fabrication resources: the yield ramp is slow initially, and process capability is under utilized later in the process lifetime.

We are developing a holistic methodology for process characterization, which is both timely and adaptive to data collected during manufacturing. Implementing our methodology provides mechanisms for tightly coupling customer requirements, design, process development, and manufacturing. In this paper we explain our methodology and demonstrate its application to two processes: a submicron CMOS process and a high speed BJT process.

2. Overview of Approach

Complexity in fabrication process characterization arises from the following:

- A process contains hundreds of steps, many of which contribute to the variability of the product.
- Only a few in-line and in-situ measurements are collected.
- When in-line measurements are collected, only a few sites are measured.
- Designers quantify the process with SPICE models, but directly measuring SPICE parameters, especially the parameters affecting AC or transient performance, is costly in a manufacturing environment.
- Many electrical test measurements (e-tests) are collected, but determining the relationship between e-tests and SPICE parameters for a process is a costly, time consuming procedure.
- Many SPICE parameters are correlated so one must account for the interactions between parameters to create realistic worst case models.

As stated in the introduction, two constraints must be balanced. First, we need to characterize the process as early as possible. If one chooses to use only measured data, sample size constraints to assure reasonable confidence in the statistics make it impossible to accomplish statistical characterization in a timely fashion. For example if confidence intervals of 85% are required 100 or more lots are needed. To circumvent this problem we use Technology CAD (TCAD) simulators in a statistical simulation environment to predict device distributions prior to manufacturing.

We start our characterization process by simulating our process recipe and device layout. The results of our simulation include a distribution of SPICE parameters. It is important to note that we use a deterministic extraction procedure. Hence, we can determine cause and effect relationships between changes in the device structure (gate length, oxide thickness, etc.) and SPICE parameters.

Our second constraint is that we wish to monitor process variations to determine the impact on SPICE model parameter distributions. Unfortunately, it is not possible to directly measure SPICE model parameters. Even though there have been improvements in the number and variety of in-situ and in-line measurements, we have found that the observability for controlling the device characteristics is still not sufficient [1]. Often, electrical test measurements are our only choice for monitoring the device characteristics during production.

We have developed a technique incorporating principal component analysis (PCA) which takes as inputs the simulated distributions of SPICE parameters, electrical tests, and in-lines and determines the following features of the process:

1. The number of worst case corners for the process relative to any particular device model and architecture.
2. Which in-lines or e-tests can be used to track the statistical variations in the SPICE parameters.
3. Nonlinear empirical relationships between e-tests and SPICE parameters (see section 4).

As process development proceeds and in-line and electrical test data is collected, we use our nonlinear expressions to compute how the distributions in SPICE parameters have changed. For example, if we notice that the variance of an e-test distribution has been reduced, we can calculate the effect on the worst case SPICE model cards. Later we use this same approach to control the mature manufacturing process.

In summary, our method uses simulated data to provide models when little data is available. The simulations results also help us understand what to measure during production and determine how to interpret the results, see Figure 1. Our approach to statistical characterization can be thought of as a predictor-corrector strategy. We use TCAD simulations to predict statistical distributions prior to manufacturing a large number of lots. As e-test data are collected, we can use this information to update the predicted SPICE performance. We believe that this affords both the predictive ability outlined in [3] as well as "accuracy" outlined in [2].

3. TCAD: Building The Process Knowledge Base

As described above, we need to characterize the statistical variations of SPICE parameters before a large number of lots have been processed. Measuring many wafers from a few lots is not sufficient because we want to accurately estimate factory variability. The distribution of device performance when sampled from a single or few lots is typically much tighter than the distribution sampled from many lots. Our solution to this problem is to run statistical TCAD simulations to predict the manufacture of hundreds of lots.

We have two goals for the TCAD simulation. First, determine the statistical distributions of the SPICE model parameters. Second, determine from this data which in-lines and e-test measurements must be measured during production to control the SPICE parameter distributions. This second requirement places constraints on how we extract the SPICE model parameters. Traditional optimization or curve-fitted methods cannot be used since they corrupt the statistics by confounding of independent variables. This confounding of variables makes it difficult to interpret the correlations of parameters which are extracted via optimization. In the next two sections we present an overview of our statistical simulation method and physically based SPICE model extraction.

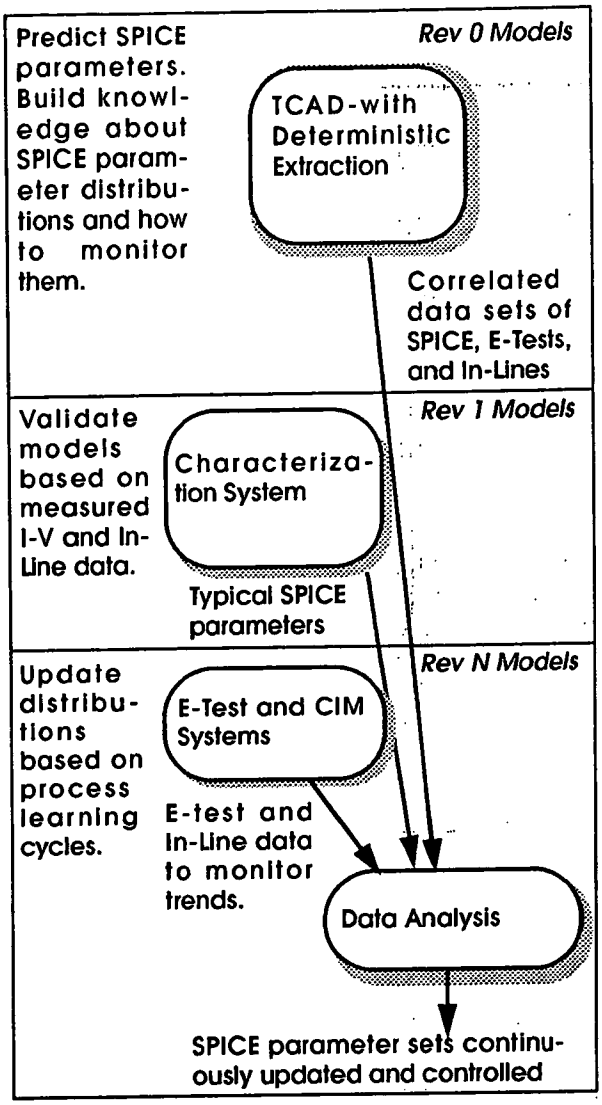


FIGURE 1. Using TCAD, Characterization, and Electrical Test data to develop, update, and monitor statistical SPICE model parameter sets.

3.1 Experimental Design

A goal for this work was to use process and device simulation to determine the distributions of device characteristics based on the distributions of in-line measurements and equipment settings (process controls). Unfortunately there are few if any closed form analytical solutions which fully address these relationships. Since the functional relationships are not well understood, it is not possible to simply transform the in-line and process control distributions to determine the device distributions.

Monte Carlo simulation provides the most accurate method of mapping process control variation to device variation. At the completion of the Monte Carlo simulation, distributions of

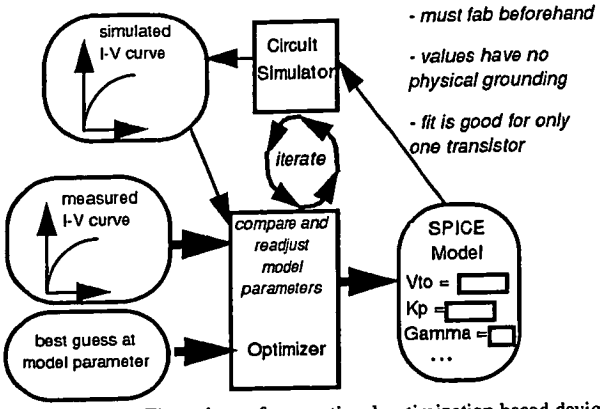


FIGURE 2. Flow chart of conventional optimization based device characterization procedure

device characteristics, their inter-correlations, and their correlations with in-lines and process controls can be calculated. To achieve reasonable confidence in the predicted correlations, the number of runs must be fairly large (typically 50-200).

Using simulation, we have two alternatives to predict the statistics of the process. Either we can run a Monte Carlo simulation incorporating physically based TCAD, or we can run a designed experiment (e.g., fractional factorial), build a response surface model (RSM), and perform a Monte Carlo simulation with the RSM to generate statistics. While the DOE technique may be faster, we must either assume the nonlinearities can be modeled with a low order polynomial, or know a priori which relationships are non-linear and account for this in our experimental design. In addition, we must decide a priori which process controls have a significant effect on the variability of the process, and include these controls in the experimental design. Since accounting for the nonlinearities and modeling the statistics are important to us, we have decided to use Monte Carlo simulations to generate our statistics. Other randomized experiments such as Latin Hypercubes would also be appropriate.

3.2 SPICE Model Extraction

In order to provide meaningful statistical information, the SPICE model parameters must be extracted in a physically-based manner. Parameter optimization or curve-fitting, as shown in Figure 2, which can result in a non-unique parameter set, is not used. The parameter extraction methodology used incorporates both 2-D dopant layout information and efficient 2-D numerical device simulation performed using SIMOS[15].

A typical SPICE model (for example Gummel-Poon or BSIM3) has approximately 30-45 parameters. As described above, many of these parameters are correlated. Some of these parameters can been thought of as having a physical definition. For example, the BSIM3 parameter Rdsw is the parasitic resistance per unit width of the source/drain. Other parameters are more empirical in nature. For example, the BJT Gummel-Poon model parameter IKF impacts the current gain degradation at high collector currents. A number of possible physical phenomena can explain the variations in IKF. For example, IKF may be limited by high injection into the base, or current-induced base widening.

The BSIM3 model is similar to the Gummel-Poon model in that it is comprised of both physically defined and empirical parameters. The parameter Tox (gate oxide thickness) has, of course, a physical definition, whereas the parameter UC empirically models mobility degradation due to bulk bias.

For SPICE model generation, we use direct physical extraction routines to extract parameters with a physical basis. For parameters which are empirical in nature, we extract them to fit the part of the I-V curve for which they are defined. In the example of UC for the BSIM3 model, we fit this parameter by applying back bias to the substrate and measuring the change in the mobility in the channel.

Since we have the exact structure for the device, we can use information from the mesh to gain more insight about the device than what is typically available from measurements of real devices. For example junction depths, depletion widths, and concentration profiles which may normally be determined from destructive physical measurements or extracted from an IV curve. As a result, with our method there is a one to one relationship between the device structure and SPICE parameters.

3.3 Examples of Statistical Simulation: BJT and MOS

We have simulated and extracted parameter sets for a number of BJT and CMOS processes. Here we show simulated statistics for a high speed poly emitter BJT and submicron CMOS process. In general, we have had good agreement between predict and measured statistics [1]. For both processes we used pdFab [12] with TMA-SUPREM3[14] to generate the process and device statistics. In Table 1, we compare the simulated and measured statistics for the 0.8 micron CMOS process. Figure 3 demonstrates that predicted distribution of BF vs. TF for the NPN transistor has an extremely non-linear correlation.

Table 1. Comparison of Simulated and Measured 3-Sigma Limits to 0.8 um Specification Limits.

Measurement	0.8 um Specs	Measured Data	PdFab
Id Max (mA/um)	0.500	0.472	0.499
Id Mean (mA/um)	0.420	0.419	0.447
Idn Min (mA/um)	0.340	0.365	0.393
Idp Max (mA/um)	0.220	0.201	0.201
Idp Mean (mA/um)	0.180	0.168	0.177
Idp Min (mA/um)	0.140	0.135	0.15
Vtn Max (V)	0.850	0.792	0.743
Vtn Mean (V)	0.728	0.714	0.695
Vtn Min (V)	0.600	0.636	0.648
Vtp Max (V)	-1.10	-1.08	-1.04
Vtp Mean (V)	-0.95	-0.977	-0.932
Vtp Min (V)	-0.80	-0.878	-0.825

4. Determining the Number of Worst Case Corners

It is well known that many device and SPICE parameters are correlated [2][3][5]. Often these correlations stem from a common underlying process step. For example, both subthreshold

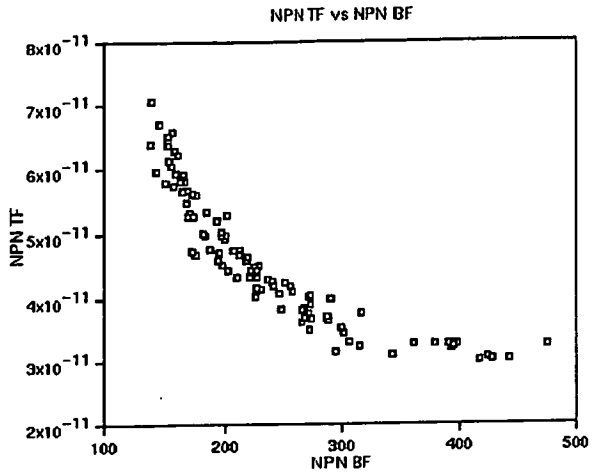


FIGURE 3. Scatter plot of BF VS TF for a Poly Emitter NPN. Notice that although the two parameters are correlated, the correlation is non-linear.

current and maximum drive current are dependent on poly CD and gate oxide thickness variations. Hence, it is expected that these two device characteristics are correlated.

Since many of the SPICE parameters are correlated, it is unrealistic to treat them as independent variables [1]. For example there are 34 Gummel-Poon parameters, but there are not 2^34 possible worst case corners. Techniques such as PCA analyze the correlation matrix of variables to determine the number of independent factors [6].

It is important to remember that a correlation between two parameters does not imply that there is necessarily a cause and effect relationship. Hence, analyzing correlations of corrupted data sets must be undertaken carefully [4][7]. Moreover, when using PCA, one must remember that it incorporates the following assumptions:

- The data used to compute the correlation matrix are sampled from a single distribution (this is often referred to as the assumption of independently, and identically distributed IID). Using Monte Carlo simulation results rather than measured data satisfies this assumption.
- The relationship between two variables can be represented by a linear correlation coefficient.

Due to the complex physical relationships between the underlying device structure and the SPICE parameters, the relationship between two SPICE parameters is rarely linear, see Figure 3. As a result, the correlations between SPICE parameters can often not be represented adequately with linear coefficients.

Before computing the correlation matrix of the SPICE parameters, we apply non-linear transformations to assure that a linear correlation coefficient is a good measure of the interactions between any pair of *transformed* SPICE parameters. It is well known that if two variables have univariate Gaussian distributions, then the joint distributions is also Gaussian. Since a linear correlation coefficient is sufficient to represent the interaction between two random variables that have Gaussian distributions, then by transforming the data to be Gaussian we can assure that the linear correlation coefficient is sufficient.

For most devices we have found that between three and eight factors are sufficient to account for most of the variability in the SPICE models.

5. Building Relationships Between Test Measurements and SPICE Parameters

As stated in the previous section, we use PCA to determine the minimum number of uncorrelated, independent factors that can represent the variability in SPICE model parameters. However, traditional PCA alone cannot determine the physical measurements needed to track each of the independent factors.

Rather than identifying the results of traditional PCA as a set of empirical factors, we would like a physical interpretation for them. Knowing the physical interpretation would enable us to track how changes in the process are impacting circuit performance. To determine these relationships we use a matching algorithm to look for similarities between the principal components and e-test or in-line measurements [4].

From a traditional PCA analysis of a BSIM3 model, we find that there is a factor which accounts for 60% of the variance in CGD0 (drain-gate overlap capacitance) and 60% of VFB (flat band voltage). However, the traditional PCA algorithm does not provide any insight into the physical phenomena leading to this result. By applying our matching algorithm we gain physical insight and improved explanation of SPICE model variability. As can be seen in Table 2 the matching algorithm has selected Tox as the tracking parameter. And, the explained variance has increased to about 79% for CDG0 and about 75% for VFB.

From this enhanced PCA approach we can determine the following:

- The minimum set of in-line and e-test measurements which can be used to track statistical variations in the SPICE model parameters.
- A nonlinear functional relationship between the electrical test measurements and SPICE model parameters.

6. Examples: CMOS and BJT Processes

As described above, we have simulated many processes two of which are described here: a submicron CMOS and a high speed poly-emitter BJT process. From these simulations we were able to generate worst case model files[1]. Next, we present the results of analyzing this data to determine the appropriate in-lines and e-test parameters to control the SPICE model parameters.

6.1 Submicron CMOS Process

We can use two perspectives to view the device variability: in-line or e-test measurements. In-line measurements usually provide structural information about the device (e.g, poly CD, Tox). However, these measurements often do not provide the resolution necessary for statistical SPICE model prediction. E-tests (e.g, Ids, Vth, HFE) are more reliable measurements, but

these measurements can often confound effects. For example, variations in Vth can depend on Tox, L-poly, spacer, channel doping or many other possible causes.

In our original work for the CMOS process, we extracted BSIM [1] model parameters. For this paper we have extracted parameters for the BSIM3 [11] model. The BSIM3 model has much less confounding of physical parameters such as L and Tox.

We had two goals for this part of our work. First, determine which in-line measurements are needed to control the SPICE parameters for the device. Second, determine if e-test measurements could be used in place of the in-line measurements. While in-line measurements are a more basic set of parameters, the e-tests are more reliably collected. Moreover, the sample size of e-test measurements is typically much larger.

We found that when using in-line measurements, the L-poly and Tox were the two most dominant tracking parameters. Collectively L-poly and Tox in-lines can explain the variability in many of the BSIM3 parameters which control the device threshold (DVT), body effect (K), and mobility (U) as shown in Table 2. To monitor the source and drain electrical characteristics e.g, capacitance (CJ), and leakage current (PB, JS0), we need an N+ sheet resistance and NLdd concentration measure. The N+ Rsh, and NLdd measurements are not currently measured as in-lines for this process. Hence, with in-lines alone we cannot predict as well and then control the SPICE model parameters.

We used pdPCA [13] to analyze the BSIM3 model with e-tests as the possible controlling factors. As in the last analysis, the third and forth factors were the NLdd and N+ Rsh. However, the first and second factors are maximum drain current and the drain current when the device is operated in the linear region. While these factors are different than L-poly and Tox, they can also span the variability in the SPICE parameters. The reader should be aware that when summed the first two factors for this model account for the same variability as the first two factors in Table 2. However, individually, these factors are different from the two factors in the previous model.

This is an important result, because now we can change the possible tracking measurements and determine if we have adequate coverage of our device variability. For example, if we do not monitor the NLdd concentration (possibly with a special test structure) a significant portion of the variance of parameters such as XJ, DVT1, and PB cannot be assured to be in control.

6.2 High Speed BJT Process

We preformed a similar analysis on a high speed BJT process. In this case we also found four significant factors (it is only a coincidence that the number of factors is the same for both of these devices). In this case our matching algorithm found four e-test measurements which can be easily collected at wafer electrical test (see Table 4). The first factor is clearly the most important. We found that the intrinsic base sheet resistance or peak beta can be used to monitor this factor. The second factor explains the variability in the collector to substrate capacitance. The third factor explains the variability in the SPICE parameters which depend on emitter area variations while the forth factor accounts for the extrinsic base variations.

7. Conclusions

We have developed a method which combines TCAD simulation with statistical analysis of empirical data. As a result, we can determine worst case models before a significant sample size of devices can be characterized.

Beyond generating worst case models from simulation, we have developed algorithms for using the simulated data to analyze measured in-line and e-test data. We determined what are the minimum set of e-test or in-line measurements to control the SPICE model parameters.

A direct result of this work is the ability to tie together the marketing, manufacturing and design communities. The mathematical models which span in-line, e-tests and SPICE model parameters can effectively communicate to design how changes in the process will effect them since they can see the impact in the SPICE model parameters. Moreover, manufacturing can see how changes in the process are effecting design. And, marketing can see how critical customer requirements are being met and controlled

8. References

(1) D. Hanson et al, "Analysis of the Controllability of a Sub-Micron CMOS Process Using TCAD", Extended Abstracts of 1994 International Symposium on Semiconductor Manufacturing, June 21 - June 22, 1994, Tokyo, Japan, pp 85-89.

(2) P. Yang et al., "An Integrated and Efficient Approach for MOS VLSI Statistical Circuit Design", IEEE Trans. CAD, vol. CAD-5, no. 1, pp. 5-15, Jan 1986.

(3) S. Nassif et al., "A Method for Worst-Case Analysis of Integrated Circuits", IEEE Trans. CAD, vol. CAD-5, no. 1, pp. 114-131, Jan 1986.

(4) J. Kibarian, Statistical Diagnosis of IC Process Faults, Dissertation, CMUCAD 90-52, December 1990.

(5) S.L. Sundaram and A.C. Carlson. "Reducing Fabrication Variability in Analog IC Technology by the Statistical Error Propagation Method Using Simple Test Structures," Proceedings of the 1990 International Conference on Microelectronic Test Structures, San Diego, 1990.

(6) I.T. Jolliffe. Principal Component Analysis, Springer-Verlag, New York, 1986.

(7) T.W. Anderson. "Asymptotic Theory for Principal Component Analysis", Ann. Math. Statist., Vol. 34, 1963, pp 122-148.

(8) Maly, A. Strojwas, S. Director, "Fabrication based statistical design of monolithic IC's", IEEE Int. Symp., Circuits and Systems, vol. 1, pp. 135-138, April 1981.

(9) E.D. Boskin et al. "IC Performance Prediction from Electrical Test Measurements," Proc. ISMSS, pp13-17 June 1992.

(10) H.K. Gummel and H.C. Poon. "An Integral Charge Control Model of Bipolar Transistors", The Bell System Technical Journal, June 1970.

(11) J.H. Huang et al. "BSIM3 Manual", v2 The Regents of the University of California, Berkeley, CA 94720.

(12) "pdFab Users Manual", v2.2 PDF Solutions, Suite 501, 425 N. Craig St., Pittsburgh, PA, 1993.

(13) "pdPCA Users Manual", v1.1 PDF Solutions, Suite 501 425 N Craig St., Pittsburgh, PA 1994

(14) "SUPREM3 Users Manual", v6.11 Technology Modeling Associates, 3950 Fabian Way, Palo Alto CA, 1993.

(15) . M.S. Obrecht, "SIMOS- A Two Dimensional Steady State Simulator for MOS Devices", Solid State Electronics., Vol 34, No 7, 1993.

Table 2. Percentage of BSIM3 v2 Spice parameter variance which can be monitored by In-lines

	Factors				
	Tox	L-poly	NLDD Conc	N+ Rsh	Sum
Average	**28%**	**25%**	**14%**	**7%**	**74%**
PHI	61.1%	0.0%	5.6%	1.8%	68.5%
K1	78.1%	1.3%	0.6%	0.0%	80.0%
K2	75.7%	1.0%	1.1%	0.0%	77.8%
DVT1	31.3%	0.6%	57.7%	2.1%	91.9%
DVT2	61.1	0.0%	5.6%	1.8%	68.6%
DVT0	0.0%	92.9%	4.1%	0.0%	97.0%
VFB	74.6%	3.0%	0.2%	0.7%	78.4%
A0	29.6%	11.0%	7.2%	0.1%	47.9%
KETA	0.5%	94.3%	0.9%	0.0%	95.9%
U0	61.4%	9.7%	11.7%	0.5%	83.3%
UA	26.5%	55.6%	9.1%	0.1%	91.4%
UB	29.3%	31.4%	2.3%	3.4%	66.4%
UC	0.2%	64.1%	5.0%	0.4%	69.8%
ETA	20.0%	26.0%	29.0%	0.1%	75.1%
PCLM	0.4%	29.2%	0.7%	0.1%	30.4%
DROUT	31.3%	0.6%	57.7%	2.1%	91.9%
PDIBL1	1.8%	58.8%	0.3%	1.1%	62.0%
PSCBE2	13.2%	11.6%	0.3%	0.6%	25.7%
DSUB	31.3%	0.6%	57.7%	2.1%	91.9%
CDSC	0.9%	96.4%	0.1%	0.1%	97.4%
CDSCB	0.1%	90.1%	0.4%	0.0%	90.6%
ETA0	0.2%	93.7%	2.5%	0.1%	96.4%
ETAB	0.0%	43.0%	13.5%	2.0%	58.6%
VOFF	32.7%	48.5%	7.1%	0.6%	88.9%
VGHIGH	62.5%	20.8%	8.8%	1.0%	93.1%
VGLOW	0.1%	97.6%	0.0%	0.0%	97.8%
DL	53.1%	0.0%	3.8%	2.1%	59.0%
CGD0	78.8%	0.0%	6.3%	0.3%	85.5%
CGS0	78.8%	0.0%	6.3%	0.3%	85.5%
CJ	15.78	2.2%	47.0%	0.1%	65.1%
MJ	2.6%	0.1%	33.9%	13.1%	49.7%
CJSW	51.2%	0.1%	19.7%	0.1%	72.3%
MJSW	10.4%	0.2%	0.6%	48.3%	59.4%
PB	14.0%	1.8%	51.9%	0.3%	68.0%
PBSW	0.0%	0.0%	0.3%	54.7%	55.0%
JS0	0.0%	1.1%	4.3%	68.1	73.0%
XJ	6.7%	1.2%	77.4%	3.7%	89.0%

Table 3. Percentage of BSIM3 v2 Spice parameter variance which can be monitored by E-tests

	Factors				
	Id Max	Id Vd<Vt	NLDD Conc	N+ Rsh	Sum
Average	**32%**	**21%**	**14%**	**7%**	**74%**
PHI	36.6%	24.6%	5.6%	1.8%	68.5%
K1	39.1%	40.3%	0.6%	0.0%	80.0%
K2	38.6%	38.2%	1.1%	0.0%	77.8%
DVT1	23.9%	8.1%	57.9%	2.1%	91.9%
DVT2	36.6%	24.6%	5.6%	1.8%	68.6%
DVT0	36.3%	56.8%	4.1%	0.0%	97.0%
VFB	32.7%	45.0%	0.2%	0.7%	78.4%
A0	40.2%	0.6%	7.2%	0.1%	47.9%
KETA	43.2%	51.7%	0.9%	0.0%	95.9%
U0	65.4%	5.8%	11.7%	0.5%	83.3%
UA	75.0%	7.2%	9.1%	0.1%	91.4%
UB	59.7%	1.1%	2.3%	3.4%	66.4%
UC	21.4%	43.1%	5.0%	0.4%	69.8%
ETA	44.5%	1.54%	28.9%	0.1%	75.1%
PCLM	8.1%	21.6%	0.7%	0.1%	30.4%
DROUT	23.9%	8.1%	57.7%	2.1%	91.9%
PDIBL1	33.7%	27.0%	0.3%	1.1%	62.0%
PSCBE2	24.8%	0.2%	0.3%	0.6%	25.7%
DSUB	24.0%	8.1%	57.7%	2.1%	91.9%
CDSC	46.3%	51.0%	0.1%	0.1%	97.4%
CDSCB	32.4%	57.8%	0.4%	0.0%	90.6%
ETA0	39.8%	54.0%	2.5%	0.1%	96.4%
ETAB	17.7%	25.6%	13.4%	2.0%	58.6%
VOFF	77.5%	3.8%	7.1%	0.6%	88.9%
VGHIGH	81.7%	1.7%	8.7%	1.0%	93.1%
VGLOW	40.4%	57.3%	0.0%	0.0%	97.8%
DL	32.7%	20.6%	3.7%	2.1%	59.0%
CGD0	47.6%	31.4%	6.3%	0.3%	85.5%
CGS0	47.6%	31.4%	6.3%	0.3%	85.5%
CJ	4.9%	13.1%	47.0%	0.1%	65.1%
MJ	1.2%	1.44%	34.0%	13.1%	49.7%
CJSW	29.3%	22.2%	19.7%	0.1%	72.3%
MJSW	7.9%	2.7%	0.6%	48.3%	59.4%
PB	4.5%	11.4%	51.9%	0.3%	68.0%
PBSW	0%	0%	0.3%	54.7%	55.0%
JS0	0%	0.9%	4.3%	68.1%	73.0%
XJ	7.3%	0.6%	77.4%	3.7%	89.0%

Table 4. Percentage of BJT-GP Spice parameter variance which can be monitored by E-tests

	Factors				
	Int Base Rsh	Col/ Sub Vbi	Emit Area	Ext Base Rsh	Sum
Average	**49%**	**15%**	**12%**	**10%**	**86%**
IS	89.4%	0.0%	2.8%	7.2%	99.5%
BF	93.0%	0.2%	0.1%	3.2%	96.5%
IKF	13.5%	62.1%	6.3%	4.4%	86.3%
IKR	90.0%	1.7%	0.1%	0.8%	92.6%

Table 4. Percentage of BJT-GP Spice parameter variance which can be monitored by E-tests

	Factors				
	Int Base Rsh	**Col/ Sub Vbl**	**Emit Area**	**Ext Base Rsh**	**Sum**
Average	**49%**	**15%**	**12%**	**10%**	**86%**
ISE	59.5%	2.6%	8.2%	3.7%	74.1%
ISC	1.9%	0.4%	2.3%	0.6%	5.2%
IRB	69.3%	19.4%	3.7%	0.0%	92.4%
RBM	1.2%	12.9%	5.3%	42.3%	61.7%
RB	89.2%	0.0%	3.6%	7.0%	99.7%
RE	50.1%	5.2%	36.2%	1.5%	93.0%
RC	0.2%	78.4%	8.2%	1.5%	88.3%
CJE	5.8%	15.3%	64.1%	2.7%	87.9%
VJE	62.6%	0.5%	1.9%	0.2%	65.1%
MJE	52.8%	2.0%	8.0%	0.1%	62.9%
TF	97.7%	0.7%	0.0%	0.9%	99.3%
TR	91.6%	4.6%	0.0%	1.4%	97.7%
CJC	23.8%	0.8%	1.5%	51.9%	78.0%
VJC	28.7%	0.6%	9.2%	43.7	82.2%
MJC	46.0%	0.1%	1.7%	45.5%	93.3%
CJS	1.0%	77.3%	3.3%	0.0%	81.6%
VJS	1.0%	82.4%	8.0%	2.6%	93.9%
MJS	1.0%	80.1%	7.8%	3.7%	92.7%
BR	94.6%	0.24%	0.2%	3.2%	98.3%
VAF	97.3%	0.2%	0.1%	0.3%	97.9%
VAR	73.2%	0.3%	0.1%	0.2%	73.8%
NE	92.9%	0.4%	0.0%	3.6%	96.9%
NC	43.3%	0.0%	4.3%	47.3%	94.9%
XTF	53.2%	2.2%	37.6%	5.1%	98.2%
VTF	55.1%	4.1%	0.0%	5.8%	65.1%
ITF	17.0%	5.5%	69.7%	5.2%	97.3%
PTF	89.3%	1.0%	0.2%	8.0%	98.5%
XCJC	0.1%	13.3%	77.5%	4.7%	95.6%

Minimizing Breakdown Voltage Drift in High Voltage (>400V) Lateral DMOS RESURF Devices

Alan Shafi, Rob Busse*, Richard Orr*, J. Montgomery McGregor#, Mark Redford.

Analog Process Technology Development, NSUK,
#Analog Process Technology Development, NSSC
**Power Management, NSSC*

Abstract - This paper investigates the breakdown voltage drift of a high voltage DMOS RESURF device during high temperature reverse-bias testing. Initial experiments show a severe drift downwards by up to 20% after the first 24 hour timepoint. Experimentation with different passivation materials has identified that optimization of this layer is key to controlling this drift. Replacing the nitride with oxynitride resulted in an upward drift in the breakdown voltage.

Introduction

A discrete Lateral DMOS RESURF[1] device is fabricated using the JHVH-450 Process. The most important requirement for the device is that the breakdown voltage is >400V. The application for the device is the high voltage switch in a flyback configured switch-mode power supply. A cross-section of the device is shown in Fig.1 which was generated from the SUPREM 3 process simulator. The high breakdown voltage is achieved in a relatively shallow epitaxial layer of 12 microns by use of the RESURF (REduced SURFace electric field) technique. An implant into the drift region controls the integrated charge in that region and sets the breakdown voltage. Interaction of the horizontal and vertical depletion regions spreads the electric field relatively uniformly across the surface, thus maximizing the breakdown voltage for the device size. The peak electric fields are positioned beneath the ends of the field plates, at the surface of the silicon. The field plates are present to minimize 2- and 3-dimensional electric field crowding which would otherwise reduce the breakdown voltage. The passivation above the metal consists of two layers. Immediately above the metal is 5500Å of 2.5% P doped, furnace deposited LTO. Above this is 10000Å of plasma deposited nitride. The standard packaging configuration is a five lead TO-220 with MG48 as the mould compound.

Although the main factors controlling the breakdown voltage for these devices, are well understood and documented [2,[3],[4], there is almost a complete absence of literature correlating the stability of the breakdown voltage to the processing. This paper addresses this point.

Experimental Procedure

The mechanism for the breakdown is avalanche multiplication and the criteria chosen to define the breakdown voltage is that a drain current of $I_D = 100\mu A$ is reached. The breakdown voltage of the devices measured at wafer sort is typically 470V for a device which has charge optimized in the drift region. The typical distribution in breakdown voltage is shown in Fig.2. These results are very reproducible. Reliability testing requires that the devices pass HTRB (High Temperature Reverse Bias) testing. This involves biasing the packaged devices with $V_G = V_S = 0V$ and $V_D = 320V$ at a temperature of 125C, followed by measuring the breakdown voltage at room temperature. This procedure is repeated to time-points of up to 1000 hours. Typical results are shown in Fig.3. The breakdown voltage clearly drops rapidly and parts fail the 400V minimum requirement after just the first time-point of 24 hours. In addition, the spread in the results increases dramatically after HTRB testing has started.

Further Experiment and Discussion

One option to increase the initial breakdown voltage is to increase the drift region length to give greater initial saftey margin for the initial breakdown voltage before HTRB testing. This would however reduce the area efficiency and specific on resistance so was not considered viable. Alternatives were therefore investigated.

A number of screening experiments were devised in order to determine the cause of the downward drift in breakdown voltage. Some potential experiments were identified and listed in Table 1, together with summarized effect. There was no significant improvement in the breakdown voltage drift for different drift region implant doses or for alloy ambient. Of the three most suitable mould compounds identified for

2

high-voltage applications, MG48 was marginally the better. The most significant improvements were achieved by eliminating the nitride passivation which is a known source of hydrogen and is known to affect surface controlled breakdown devices [5],[6],[7]. Since the optimum breakdown is achieved where the peak electric fields in the silicon under the source and drain field plates are balanced, mobile hydrogen will affect this distribution and hence degrade the breakdown voltage. HTRB results for devices with no nitride and hence no source of mobile hydrogen are shown in Fig.4. Clearly the mean of the distribution is relatively independent of the HTRB time.

There are two problems with an oxide only passivation scheme. Firstly, there is no longer a moisture barrier and secondly, the yield of these devices was typically <40%, corresponding to low breakdown voltage failures. Further investigation involving SEM cross-sections revealed the presence of cracks in the deposited oxide passivation layer at the edges of the field plates, as shown in Fig.5. As an additional test, the wafer was etched in KOH which will attack the metal but not the oxide. The location of any pinholes will therefore be highlighted by areas of etched metal. Fig 6 shows a top view photograph after etching. Significant areas of metal have been etched from locations corresponding to edges of the metal field plates. This is consistent with the crack observed in Fig.5. The position of the cracks also correspond to the position of highest electric field. Fig.7 shows a simulation of the electric field in the same region as the SEM cross-section shown in Fig.5 which confirms this. Analysis of the oxide layer in devices which have the additional nitride passivation layer show a complete absence of cracks. The physical damage in the oxide only passivated devices is therefore the likely cause of the low breakdown voltage failures.

Potential replacements for nitride film which were considered were a low-hydrogen nitride process deposited by an Applied Materials AMT 5000, and oxynitride. The AMT 5000 nitride resulted in no improvement as summarized in Table 1. However, the oxynitride showed extremely promising results. The oxynitride was deposited with two extremes of stress, -1 and -2 E9 dynes/cm^2 and two hydrogen concentrations, corresponding to refractive indices of 1.7 and 1.87. The most promising results were achieved for the low stress and low hydrogen content split. Fig.8 shows the HTRB data to a 168 hour timepoint. Interestingly the breakdown voltage starts lower than the other results presented here but then increases to the 168 hour timepoint. Since all previous HTRB data has flattened off after this timepoint, it is considered a good indication of 1000 hour results.

Conclusions and Future Work

The passivation scheme has been identified as controlling the drift in breakdown voltage for a high-voltage Lateral DMOS RESURF device. A deposited oxide/oxynitride passivation scheme appears to show the most promise for retaining a moisture barrier and minimizing breakdown voltage drift.

Further experiments are planned around a DOE with drift region charge, oxynitride refractive index and stress, and mould compound as control factors. This would identify the optimum manufacturing window.

References

[1] J.A. Appels, M.G Collet, P.A.H. Hart, H.M.J Vaes J.F.C.M. Verhoeven, 'Thin layer high-voltage devices(RESURF devices). *Philips Journal of Research*, vol.35, no.1, pp1-13. 1980

[2] S. Colak, 'Effects of Drif Region Parameters on the Static Properties of Power LDMOST', *IEEE Trans. Electron Devices*, vol. ED-28, pp.1455-1466, 1981.

[3] S. Colek, B. Singer, E.H. Stupp, 'Lateral DMOS power transistor design', *IEEE Trans. Electron Device Letts.*, vol.ED-1, p.51, 1980

[4] Z. Parpia, C.A.T Salama, 'Optimization of RESURF LDMOS Transistors: An analytical Approach', *IEEE Trans. Electron Devices*, vol. ED-37, pp.789-796, 1990.

[5] J. Dunkely, G. Ganschow, D, Hannaman, J. Patterson, S. Willard, P. Gopi, 'Hot electron induced hydrogen compensation of boron doped silicon resulting from emitter-base breakdown',*Proc. IEDM*, pp.785-788, 1992.

[6] D. Quon, P.K. Gopi, J. Sonek, G.P. Li, 'Hot Carrier induced bipolar transistor degradation due to base dopant compensation by hydrogen: Theory and experiment', *IEEE Trans. Electron Devices*, vol. ED-41, pp1824-1830, 1994.

[7] N.M. Johnson, 'Mechanisms for hydrogen compensation of shallow acceptor impurities in single crystal silicon', *Phys. Rev.B*, vol.31, pp.5525-5528, 1985.

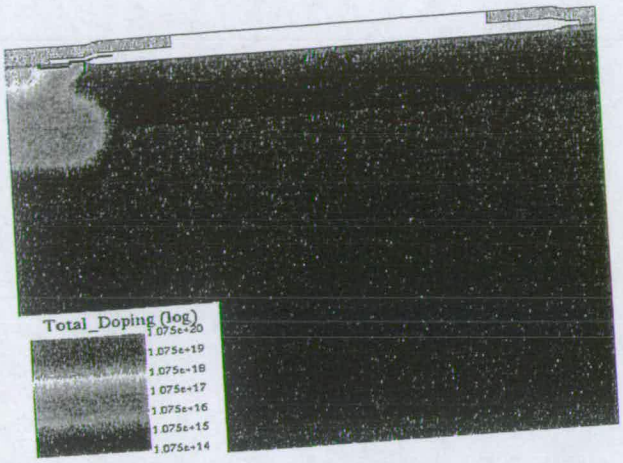


Fig.1 Process cross-section of RESURF device

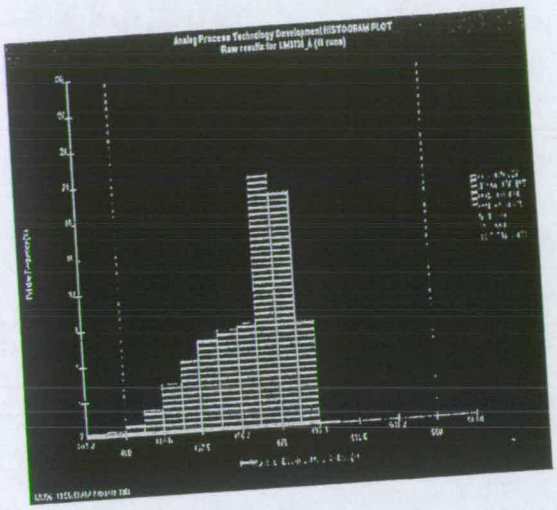


Fig.2 Typical breakdown voltage at wafer sort

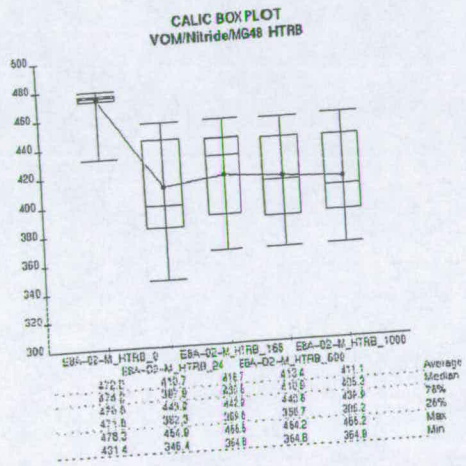


Fig.3 HTRB for standard passivation and packaging

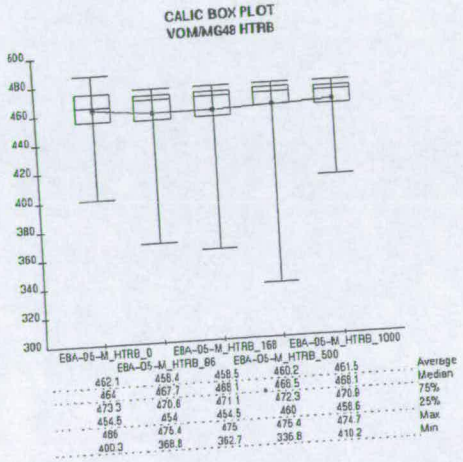


Fig.4 HTRB for oxide only passivation

Experiment	Effect
Drift implant dose	No significant ffect
Alloy ambient -N2, H2/N2, no alloy	No significant effect
Mould Compound Type	No significant effect but best of B17, MG15 and MG48 is MG48
No nitride passivation	Greatly minimizes breakdown voltage drift
AMT5000 Low H nitride	No significant effect

Table 1. Experiments and effect on breakdown drift

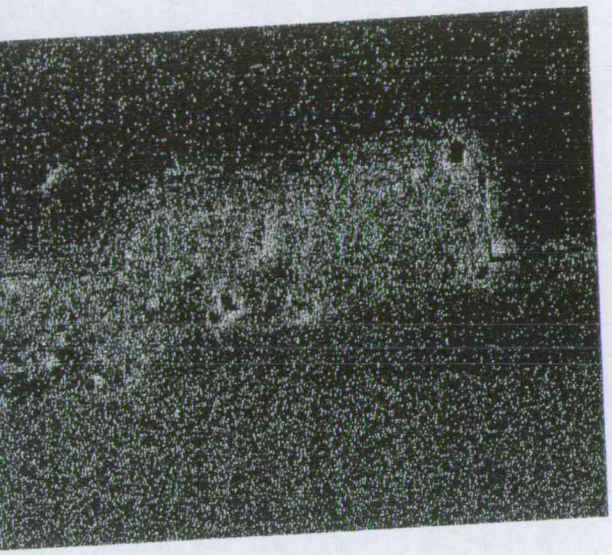


Fig.5 SEM of field plate edge for LTO only passivation showing cracks under the field plate.

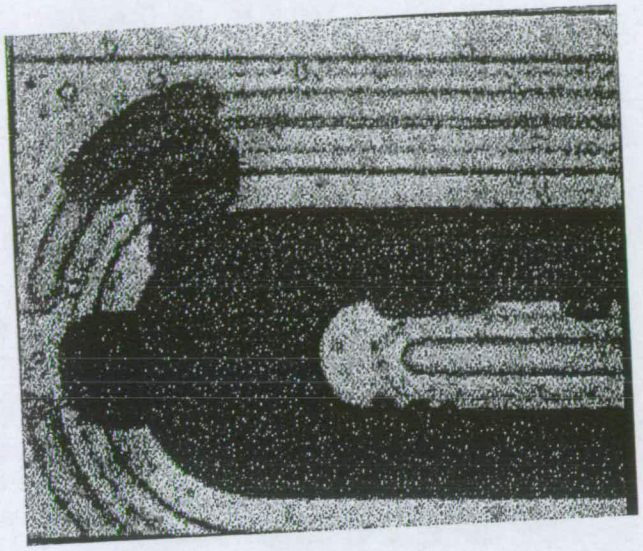


Fig.6 Photograph of fingertip region of device after KOH etch.

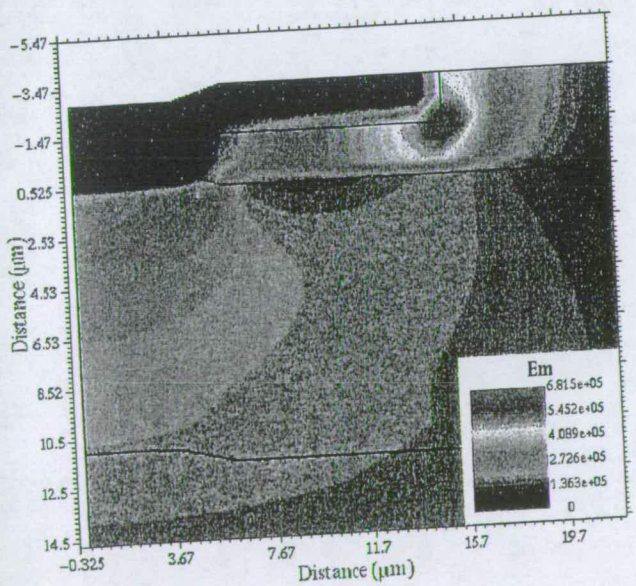


Fig.7 Electric field distribution around field plate

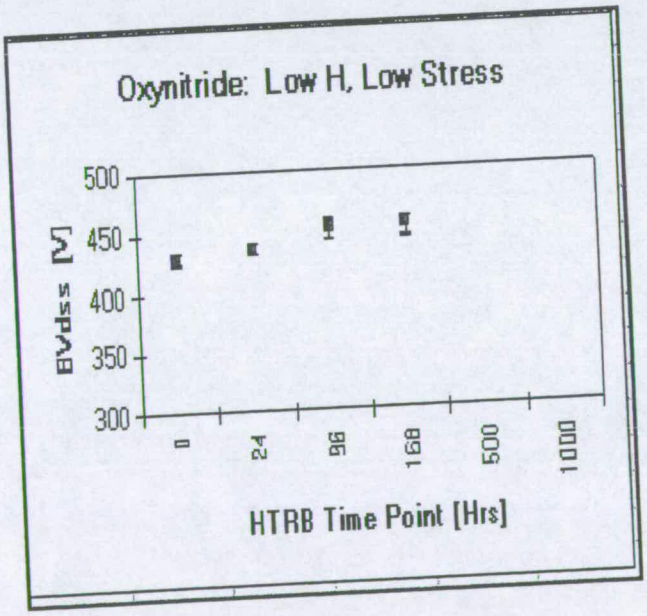


Fig.8 HTRB for oxynitride passivated devices

The Detection and Analysis of Platinum Silicide Spiking in Contact Structures.

Alastair McGibbon*, Calum McLeod**, Gordon Robertson**, Chris Porteous**, Mark Redford*, Kevin Yallup*, Christine Buchanan***, Pat Cavanagh*** and Colin Elgey***

*UK Analog Process Technology Development

**UK BC1.5 Transfer team

***UK Quality Assurance Dept.

Abstract - Please contact authors for technical details.

Introduction

In this paper, we discuss the detection and subsequent observation of PtSi spiking in doped silicon epilayers at contact structures. Spiking can seriously affect electronic properties such as subthreshold leakage in CMOS and emitter-base leakage in bipolar devices. However, although observations have been made in polysilicon (e.g. ref. 1), less work has been carried out on the nucleation and propagation of spikes in doped single crystals, nor does there seem to be an understanding of the way this effect impacts processes run at National Semiconductor. The analysis of spikes, which can be <10Å wide but can propagate as much as a 1µm into the substrate, requires the application of advanced failure analysis techniques. In this paper, we demonstrate that the investigation of yield issues on this scale are now readily accessible. We begin by describing a yield problem related to PtSi spiking that was encountered on a process currently being transferred to the UK. This is followed by a description of the techniques employed to enable analysis of the problem, and a discussion of the results and issues raised by this investigation.

NMOS subthreshold leakage in BC1.5

The discussions in this paper centre around the failure analysis of NMOS devices fabricated using the BC1.5 process. A schematic diagram of a typical NMOS structure in this process is shown in figure 1. BC1.5 is a 1.5µm epilayer, dual metal, single polysilicon BiCMOS process currently being transferred from NSTE to NSUK. In UK silicon, it was found that the wider the device, the more likely it was to exhibit subthreshold leakage, which in some cases was as high as 1µA. In terms of the failure mode in the qualification vehicle, typically 1 in 6 NMOS devices (W=100µm, L=2.5µm) would fail. As is shown in figure 2, isolation and probing of these devices was enabled by focussed ion beam (FIB) etching and Ti sputtering (to form contact pads). Detailed probing revealed poor diode properties (in some cases direct shorts) across the n+(As)/p-(B) source and drain to substrate junctions. A series of split lots run on line attributed the root cause of

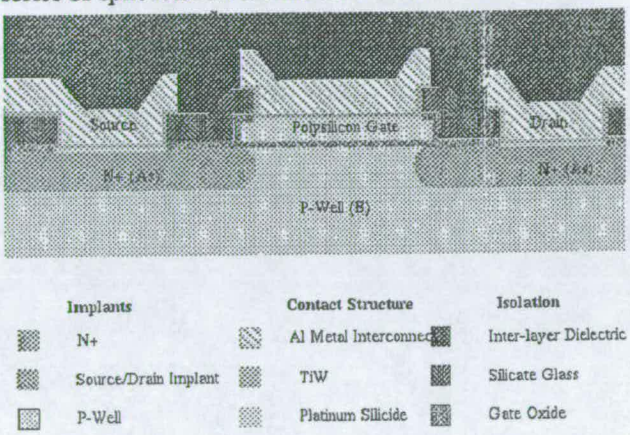


Figure 1. A schematic diagram of an NMOS structure in the BC1.5 process.

failure to the deposition of excess Pt (300Å instead of 150Å) prior to silicidation. However, although device probing indicated that the failure mechanism may be defect related, evidence of this was limited, despite the application of several conventional failure analysis techniques such as photoemission, SEM, conventional TEM and preferential etching.

TEM/STEM analysis of device cross-sections

A more detailed investigation of the failure mode was carried out by performing high spatial resolution (<10Å) dedicated scanning transmission electron microscope compositionally sensitive imaging and microanalysis in a (STEM) at the University of Glasgow. These techniques are not as readily accessible in conventional TEM/STEMS. In a STEM, a high current density electron probe (5Å in diameter) is focussed onto a thin (100-1000Å) film

specimen. The probe can be held stationary to enable the

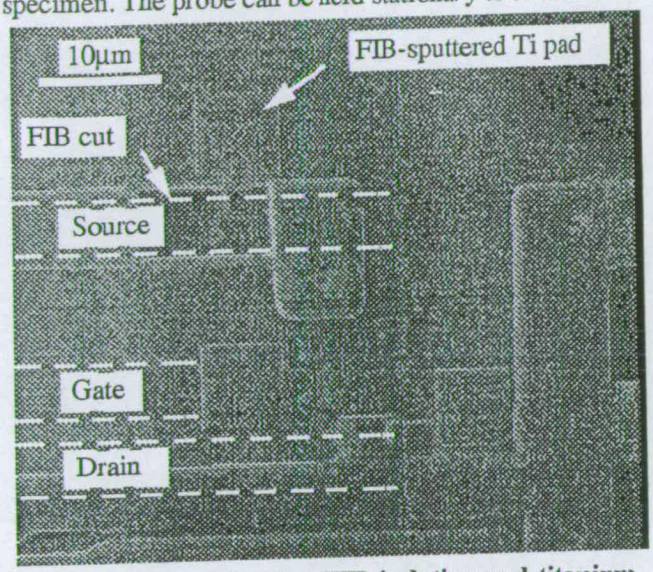


Figure 2. Image showing FIB isolation and titanium pad formation to enable detailed device probing.

analysis of a very small volume of specimen, or scanned in a raster pattern to form an image which is similar to that which can be seen in a TEM, except that image modes can be obtained which are directly interpretable compositional maps of the area of interest. A schematic diagram of probe formation in STEM is shown in figure 3.

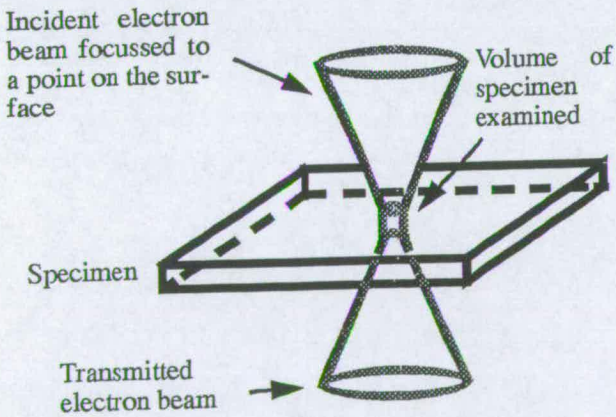


Figure 3. A schematic diagram depicting the focussing of an electron probe onto a thin film specimen in STEM.

To facilitate this analysis, it is necessary to prepare a 500Å thick cross-sectional membrane of a failing NMOS device. This was achieved by the NSUK QA group using a recently-developed TEM specimen preparation method (figure 4) in the FIB facility. This approach provides accurate selectivity of regions of interest and precise control of specimen thickness (not easily achievable using conventional methods). The technique involves the creation by cleaving and mechanical polishing of a small block of material (~100µm wide) containing the region of interest. The FIB is then used to etch out trenches on either side of the region of interest to form an electron transparent film which is supported by the block edges. The advantage of this technique over conventional TEM

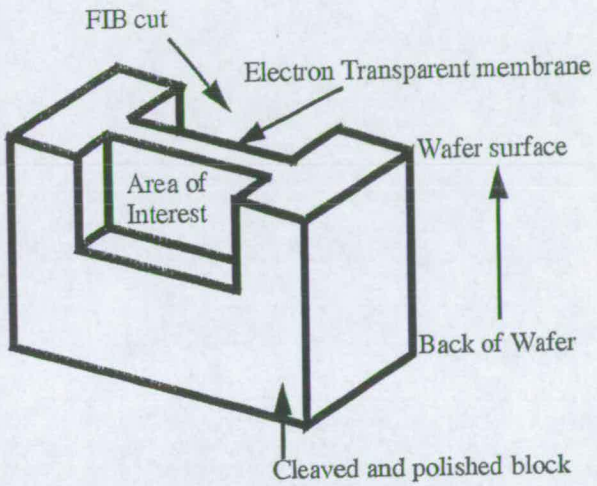


Figure 4. Schematic diagram of a TEM/STEM cross-sectional specimen prepared using a FIB facility.

specimen preparation is that it is considerably more robust and area-selective. A FIB image of a completed NMOS cross-section is shown in figure 5.

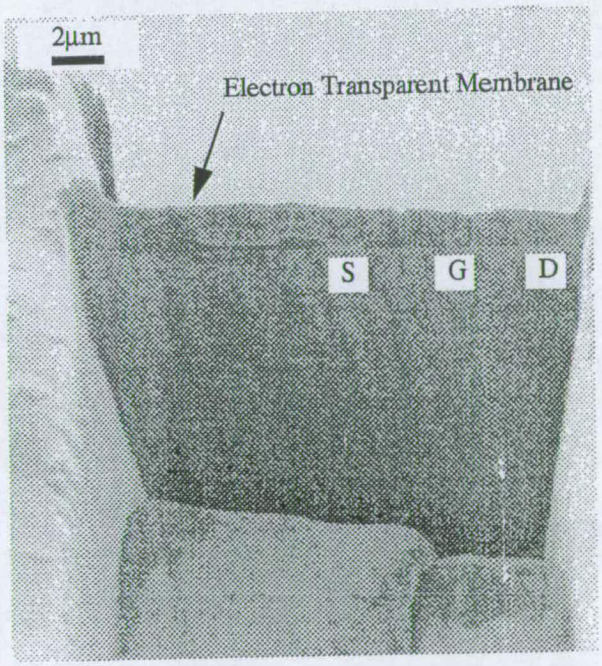


Figure 5. FIB image of a completed cross-section of a BC1.5 NMOS device suitable for STEM analysis.

Analysis of BC1.5 contact structures

High resolution STEM analysis of FIB-prepared cross-sections revealed the existence of spiking in failing NMOS devices. Images of an NMOS structure taken in the STEM are shown in figures 6 and 7. In these images,

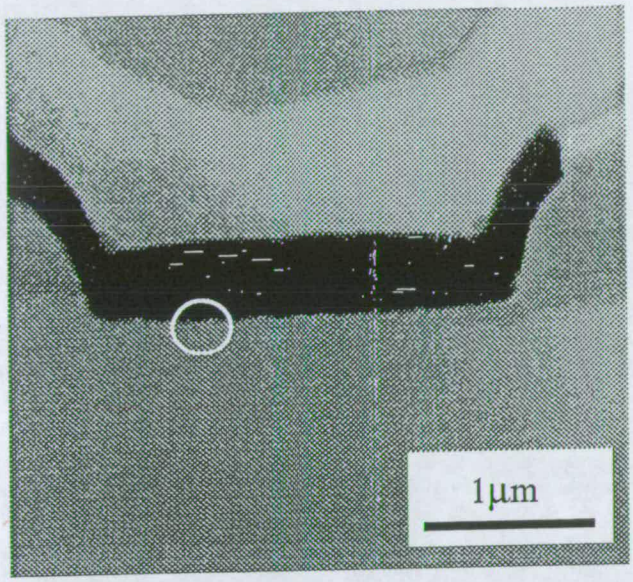


Figure 6: STEM image of a contact structure in an NMOS device. An area containing spikes is circled and is shown in more detail in figure 7.

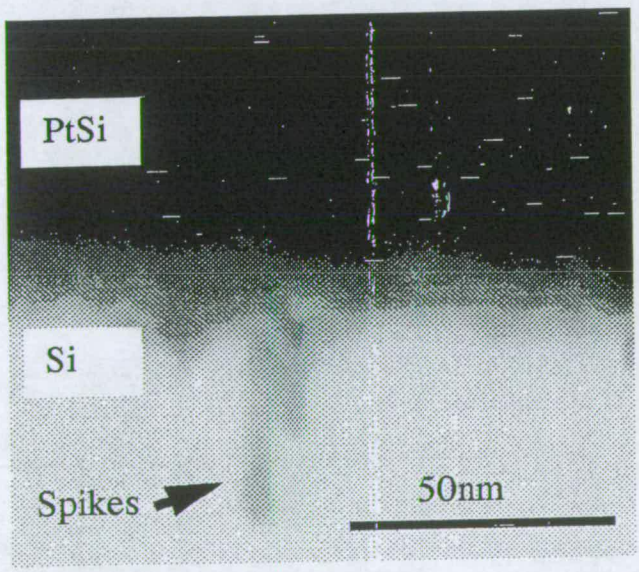


Figure 7: High resolution STEM image of spike at the contact/substrate interface.

spikes of (typically) 50nm in length can be observed emanating from the contact interface with the epilayer. The compositionally sensitive nature of the images (acquired using the technique of high angle annular dark-field or Z-contrast imaging- ref.2) suggests that the composition of the spikes is similar to that if the contact. In STEM, this can be pursued further by acquiring energy-dispersive x-ray (EDX) spectra at specific points along the spikes to determine local compositions. This can be achieved with a spatial accuracy of ~10Å. A schematic diagram showing the points at which a typical run of EDX spectra was acquired is shown in Figure 8. The results from these analyses confirm that the spikes contain Pt at all points.

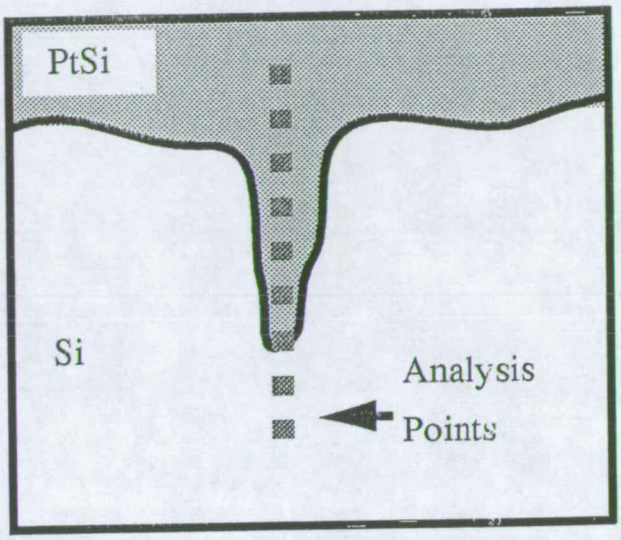


Figure 8: Schematic diagram showing the points at which EDX spectra were acquired from each spike analysed.

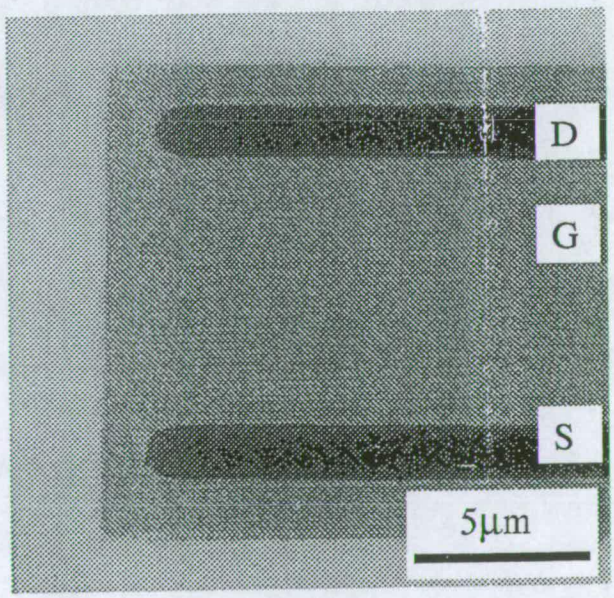


Figure 7: STEM plan-view image of a failing NMOS device.

4

STEM analysis of the contact structures was also

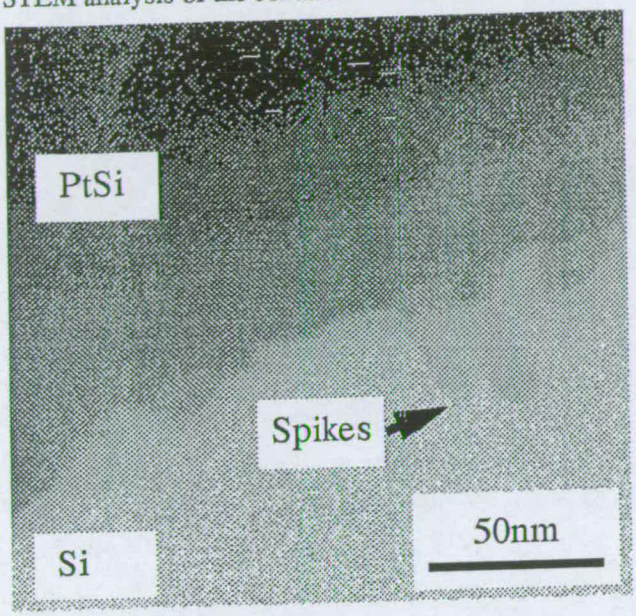


Figure 10: High resolution STEM plan-view image of spikes emanating from the edge of a contact in a failing NMOS device.

carried out on a conventional TEM plan-view specimen (figures 9 and 10). The studies again revealed Pt spikes, in this case emanating outwards from the side of the contacts. When taken together with the cross-sectional specimen, this provides valuable information on the general orientation of the spikes.

Discussion and conclusions

STEM imaging and microanalysis has revealed the presence of PtSi spikes at the contact/epilayer interface in failing NMOS devices. In general, it can be seen that the interface integrity is poor, with roughness on the 10nm scale. It appears from the images that the interface roughness enables the nucleation of spikes (with a base of typically 10nm) which then penetrate into the silicon. However, to pursue the nucleation mechanism theory further, more analyses are required of specimens that have undergone a range of fabrication conditions.

A combination of both the cross-sectional and plan-view information suggests that the preferred direction of spiking is along (111) (i.e. at an angle of 45° into the substrate) and that as many as 10^3-10^4 spikes can exist in a $200\mu m^2$ contact region. Although the spikes observed were 1/10th of the length required to completely short out the N+ Source drain region in BC1.5 devices, the probability of picking out such a defect without the help of a characteristic photoemission signature (based on ET analyses) in a 500Å thick cross-section is about 1 in 2000. In addition, the maximum observable length of a spike in TEM is limited by the specimen thickness (50nm) in all cases except that in which the spike runs exactly in the plane of the cross-section. However, from the evidence seen in the STEM, it is clear that Pt spiking is a severe problem in wafers where excess Pt has been deposited.

Having established a route via which PtSi spiking can be detected and analysed, it is now possible to carry out a general assessment of the quality of silicon/silicide interfaces in processes run at NSUK, which in turn may have a positive impact on yield and reliability.

As a general conclusion, it is also clear that, through the means of a reliable in-house specimen preparation technique, TEM and STEM analysis can now be accessed cheaply, quickly and easily in the investigation of fab yield problems. In addition, by targeting specific electron microscope facilities for specific yield problems, higher quality and more relevant information can be achieved than is very often the case when a single external source is relied upon.

References

1. G Grivna, J Kirchgessner and J. Carlson, J Vac Sci Tech B 10 (1992) 53

2. A J McGibbon, S J Pennycook and J E Angelo, Science 269 (1995) 519

A Fingerprinting Methodology for Process Transfer and Dual Sourcing

Alastair McGibbon, Kevin Yallup, John Altrip, Mhairi Harper and Mark Redford

UK Analog Process Technology Development

Introduction

A major manufacturing challenge for the semiconductor industry is to develop a high degree of flexibility by dual sourcing analog bipolar or Bicmos processes in the same way as can be achieved for digital CMOS. Greater demands for parametric control are placed on analog processes by analog products, e.g. all parts of all transistor characteristics must match, and so dual sourcing becomes a highly complex issue. The goal therefore is to ensure a seamless transfer of products from one site to another, providing matching reliability and final test characteristics. This can be achieved by detailed 'fingerprinting' of the process, enabling a greater convergence between the two sites using detailed information on all aspects of the process. With this goal in mind, we discuss the implementation of a structured methodology for bipolar/Bicmos process characterisation which can be used during process transfers (in accordance with transfer phase review system deliverables) and subsequently for process dual siting and control. This procedure was developed partly from the learnings gained in the course of the BC1.5 Bicmos process transfer from NSUK to NSTE, and evidence of physical and electrical fingerprint matching are given in later sections as an example.

The outcome of this methodology will be, in the first instance, to ensure the successful manufacturing of LB technologies which are high volume commodity bipolar processes. LB300 and LB250 are currently being re-developed at NSUK but at the same time are being transferred to NSTE for dual sourcing as part of the Armadillo project.

Fingerprint Structure and Content

In the course of the BC1.5 transfer (a dual metal, single poly bicmos process) from NSTE to NSUK, detailed electrical and physical fingerprinting of the process was carried out (see figure 1 for examples) as a means of firstly ensuring good physical matching of the two processes, and secondly to highlight any favourable characteristics intrinsic to the process on one site which could be adopted in the other, thus improving process quality on a corporate level. An example of this was considerable yield improvements made by the discovery of NMOS subthreshold leakage (a feature which was not, at the time, measured at electrical test) due to PtSi spiking which was originally highlighted from detailed benching data (1).

At the initiation of the LB transfer between NSUK and NSTE, it was realised that detailed fingerprinting such as that used for BC1.5 could, if implemented in a structured manner act as a key tool for the successful transfer of the process, and ultimately ensure quality control of dual sourcing. A schematic diagram outlining a generic fingerprint structure is shown in figure 2. The fingerprint can be separated into three categories; process, electrical and performance from which ultimately come parametric-intensive documents for process objective specification and quality function deployment. However, the idiosyncrasies of an analog process are such that it requires an analog fingerprint:

1. Process:

The basic process fingerprint can be split into two sections: (a) in line and (b) end of line:

(a) In line:

A fab database of SPC details, together with a library of Goldman profiles for each furnace cycle(2).

SRP profiles from 'golden' wafers dropped off after key drive stages in the process.

SIMS profiles from post-implant test wafers.

SEM cross-sections of CD structures taken post-DI and FI on critical mask layers.

Stained SEM cross-sections to characterise critical alignments (impacted by, for example, epi shift) and plan-views to characterise defect densities (e.g. Sirtl etch prior to epi growth).

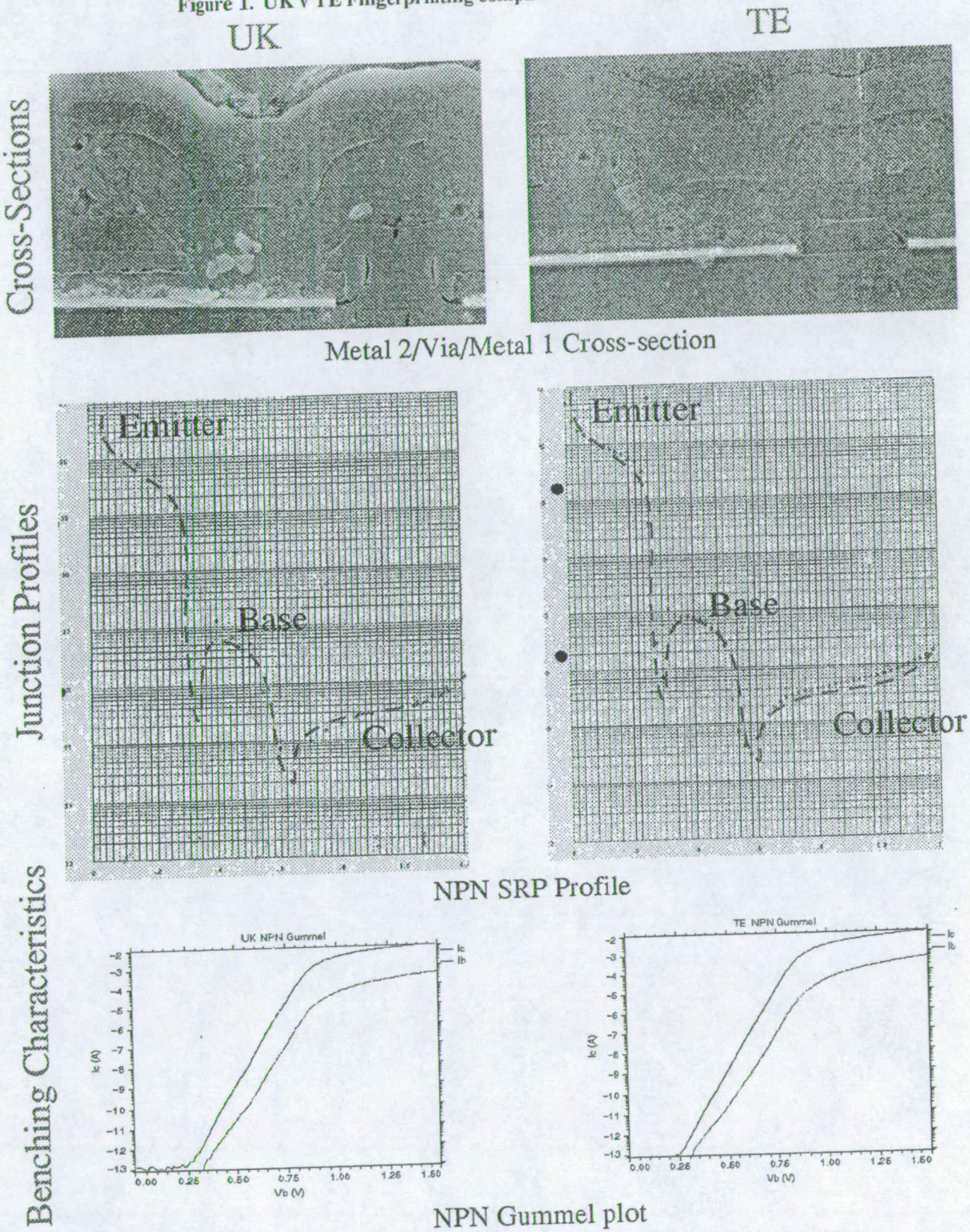
(b) End of line:

SRP profiles of all junctions

SEM cross-sections of key devices and interconnect structures.

National Semiconductor Corporation Confidential

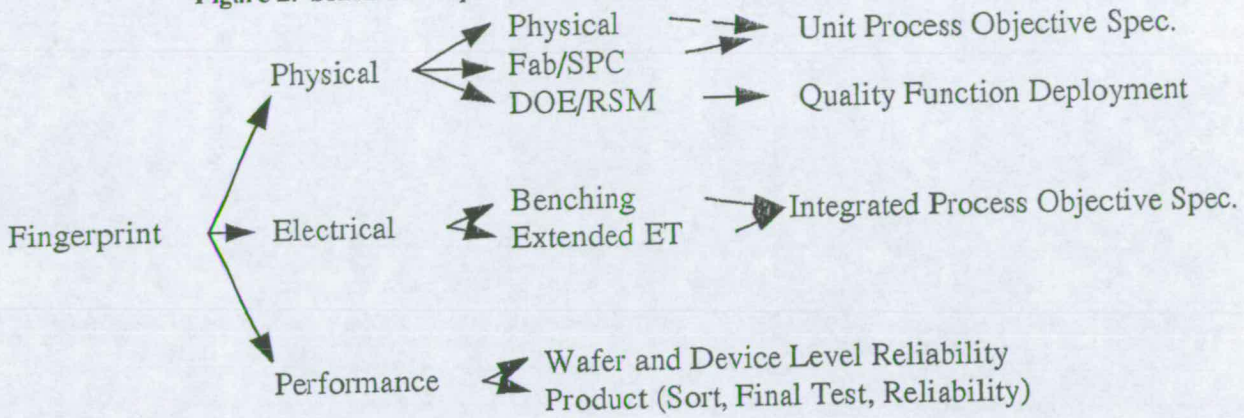
Figure 1. UK v TE Fingerprinting comparisons for the BC1.5 process



Where necessary, the physical fingerprint should also comprise detailed characterisation of additional features which have a strong impact on the performance of the particular process being fingerprinted (e.g. TEM cross-sections looking at sur-

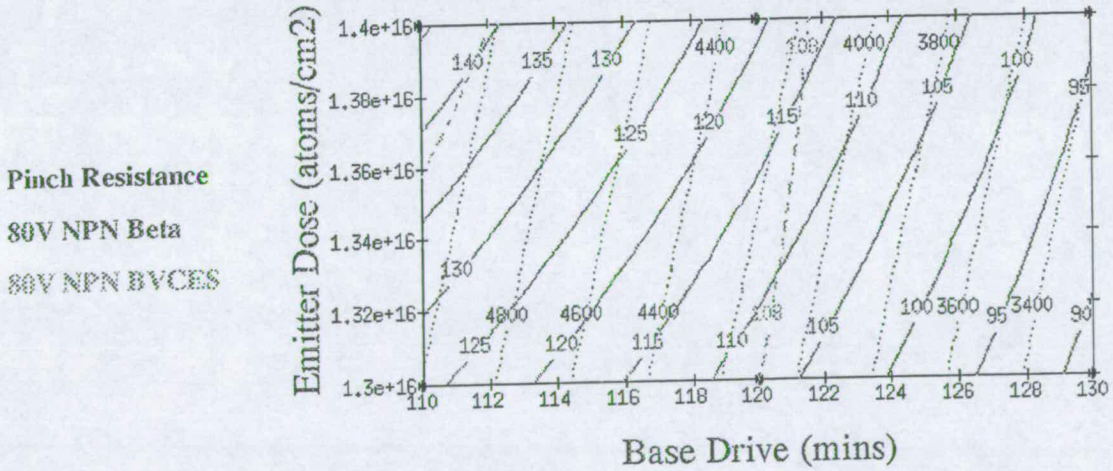
face defect nucleation sites).

Figure 2. Schematic representation of a generic fingerprinting structure



In addition, to ensure manufacturability and a greater understanding of process sensitivities, it is proposed that design of experiment (DOE) techniques (using RS Discover software) and, in particular, response surface methodologies be used to characterise key process stages and interactions such as doses, drives, CDs and spacings. An initial example of this, showing the variation of beta and breakdown in the UK LB300 process as a function of emitter dose and base drive is shown in figure 3. For LB, other key areas include epilayer growth, implant resistor and Cap Ox/MEC Anneal. The aim of this approach is to compare the operational spaces of processes in different sites, ensuring compatibility with a universal set of process design rules and product operation limits

Figure 3. Example of response surface analyses for the UK LB300 process



2. Electrical:

Device parametrics are very important for analog technologies and so electrical fingerprinting must be very thorough:

(a) Benching characteristics:

Measurement over the full device operating range of all principal devices. This includes, for example, gummel beta and Vc-Ic plots for bipolar transistors. The set of devices analysed should be consistent with the supported device list specified by designers for parameter extraction.

Performance at temperature is very important for analog products, and so in LB for example, NPN gummel measurements are carried out over a range of temperatures and the temperature coefficient of resistance is measured for commonly used resistor structures.

(b) Extended ET data:

ET data is effectively a 'snapshot' of device characteristics at set operating points, and so particular process sensitivities highlighted by benching should, when necessary facilitate the addition of more tests at ET.

Additionally, an important aspect of electrical characterisation is the formation of a 'library' of known good wafers which can be used as a reference point when yield issues arise.

3. Performance:

Again, performance data can be characterised in two fields:

(a) Process Reliability:

Device level reliability data based on, for example, 7 alloys testing and lifetime testing on primitive transistors operating at biases which reflect extreme operation conditions in standard products. Ceramic packaging is often used to minimise packaging-induced effects.

Wafer level reliability data, concentrating on key tests which reflect known process sensitivities (e.g. electromigration and mobile ion contamination).

(b) Product Reliability:

Sort, FT and Reliability data from key products supported by the process.

Implementation Methodologies

A key to successful implementation of fingerprinting will be the way in which data is acquired during the course of the transfer (dedicated resources must be available to carry this out), and on what basis decisions are made regarding the degree of matching between the two sites running the process.

It is proposed that at each decision point in the transfer process, matching has been achieved when an agreed set of parametric distributions in the receiving site lie within 1 standard deviation of the mean value in the sending site, and that at the completion of the transfer both distributions satisfy the corporate cpk values. Ultimately, the final decision on process matching will be made on product parametrics. However in the course of the transfer, process objective specifications and ET parametrics will be used at intermediate decision points. During each phase, the wealth of analog data, such as full benching characteristics and junction profiles will be used actively in process 'tuning' to ensure parametric agreement.

Conclusions and Future Opportunities

As process monitoring develops, the fingerprint can be optimised and tailored to monitor sensitivities characteristic to the process of interest. As a dual sourcing tool, it may in future be possible to develop a 'figure of merit' for the site performance of a process by calculating, for example, the mean deviation of an agreed set of parameters (ET and product) from nominal values.

Fingerprinting requires the acquisition of a large data set and so a future challenge for dual sourcing will be to ensure that information is kept up to date at both manufacturing sites. One possible way of managing such data would be through the use of advanced process engineering and control systems such as Starfire by Domain which can provide a platform which can be accessed from multiple sites. In addition, a common web-site, such as that being developed for the LB technologies can be used to record and monitor multiple site fingerprints. Such a site can also act as a home for centralised documentation such as design rules and process flows.

In conclusion, a coherent fingerprinting methodology is well placed to become a highly effective tool for not only highlighting process weaknesses at one or both sites, but also to pinpoint yield improvements which can be transferred between sites, so improving corporate process performance.

References

1. High Spatial Resolution Microanalysis of Platinum Silicide Contact Spiking in a STEM. Alastair McGibbon, Calum McLeod, Gordon Robertson, Mark Redford, Kevin Yallup, Christine Buchanan, Pat Cavanagh and Colin Elgey. 1997 Spring MRS, San Francisco.

2. Martin Fallon et al., this conference.

Identifying And Quantifying Variability Within The LMDMOS 300A Process : A Case Study

M. Redford, M. Fallon†, J. McGinty*, A. J. Walton†, M.I. Newsam†,
G. Michalowicz‡, C. McGregor*

Analog Process Technology Development, Greenock, Scotland,
*Fab 3 Device Physics, Greenock, Scotland,
‡National Semiconductor, Santa Clara, California,
†Department Of Electrical Engineering, Edinburgh University, Scotland.

Abstract

This paper presents a case study illustrating how the Design Of Experiments (DOE) in conjunction with Technology CAD (TCAD) can be used to identify sources of variability within a process that has been tranferred into manufacturing. It will show how this approach can be used to study potential solutions to improve process performance and produce more robust and manufacturable processes.

The case study was carried out on the LMDMOS 300A process running on the 6" line in the UK. This process was observed to have a strong variability in P Channel V_t, which was found to be controlled by epi resistivity [1,2,]. The software employed to design the simulation experiments was RS/1 and the 1D process simulator SUPREM-3 was used within the Caesar framework to produce the contour plots. The corresponding contour plots produced by the DOE/RSM analysis is shown in fig 1 & 2. Fig 1 shows that V_t is highly susceptible to the epi resistivity with there being a significant variation of threshold voltage over the range that the epi resistivity is known to vary. Figure 2 also indicates that the threshold adjust implant has a significant effect on the threshold but this factor is much better controlled. Figure 3 shows the simulated V_t distribution which can be compared with the experimentally observed variation shown in figure 4. The foregoing results indicate that without tighter specifications on the epi resistivity it will be impossible to reduce the V_t variability. The full paper will present further DFM analyses which study the impact of potential process modifications to improve the robustness of the of the P Channel V_t. The results and the effect of these modifications on P Channel V_t control and other electrical characteristics will be presented.

References

1. Epi Monitor Improvement NSUK Memo C. McGregor Dec 94

2. LMDMOS 300 A & B N-Epi Spec Change. G. Michalowicz NSSC Memo Aug94

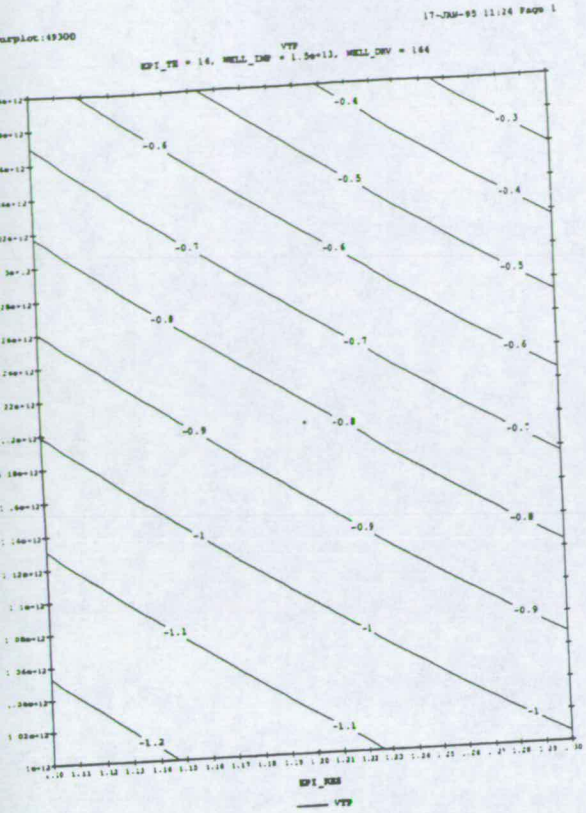


Figure 1

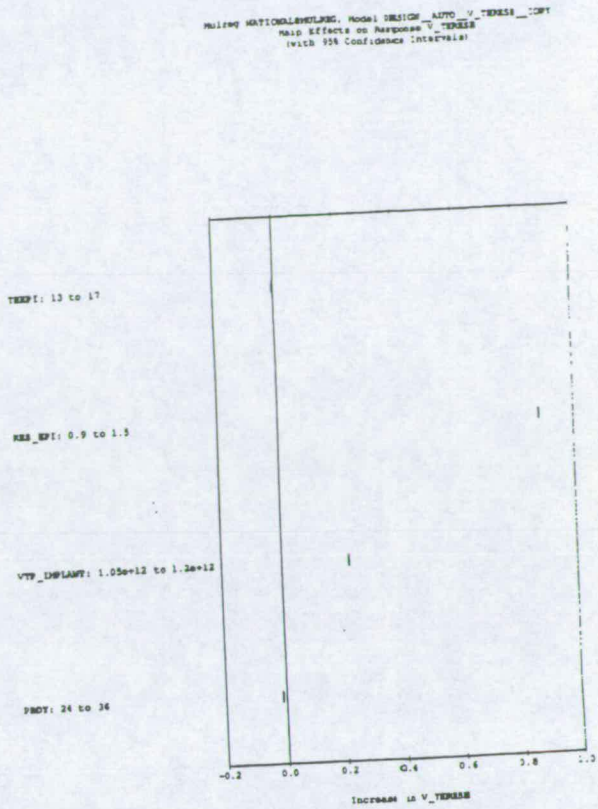


Figure 2

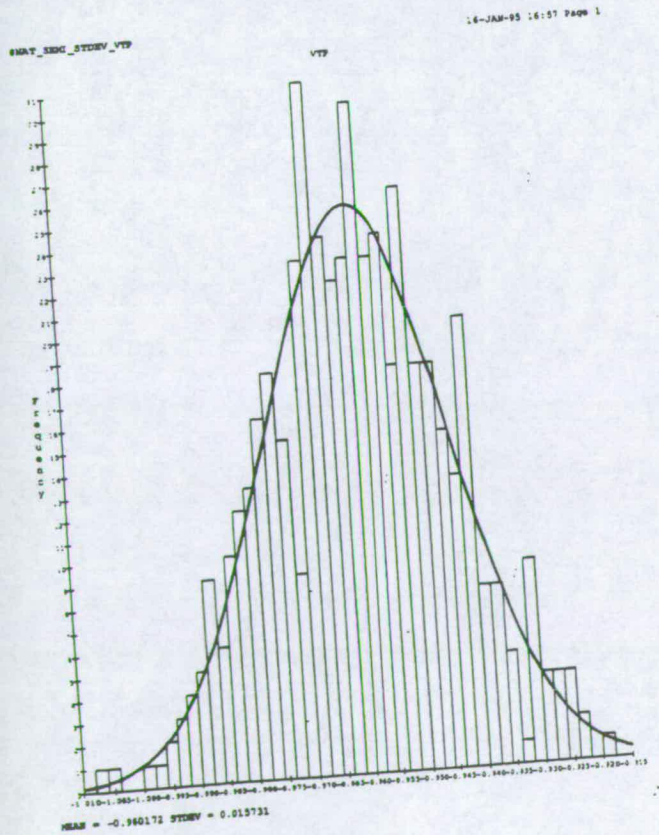


Figure 3

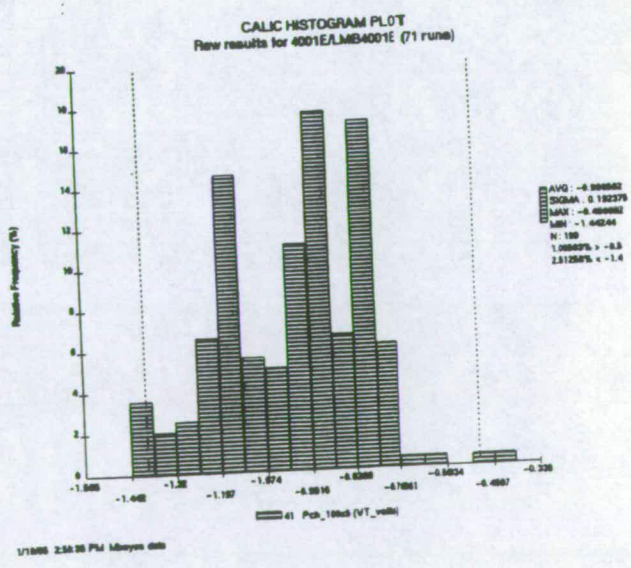


Figure 4

Implementation of Statistical TCAD in a Manufacturing Environment

Jim McGinty, Garrett Long*, John K. Kibarian*, Martin Fallon, Paul McKernan, Mark Redford
Analog Process Technology Development, Analog Division, Greenock, UK
jim@shark.nsc.com
*PDF Solutions Inc., 333 West San Carlos St., Suite 625, San Jose, CA 95110

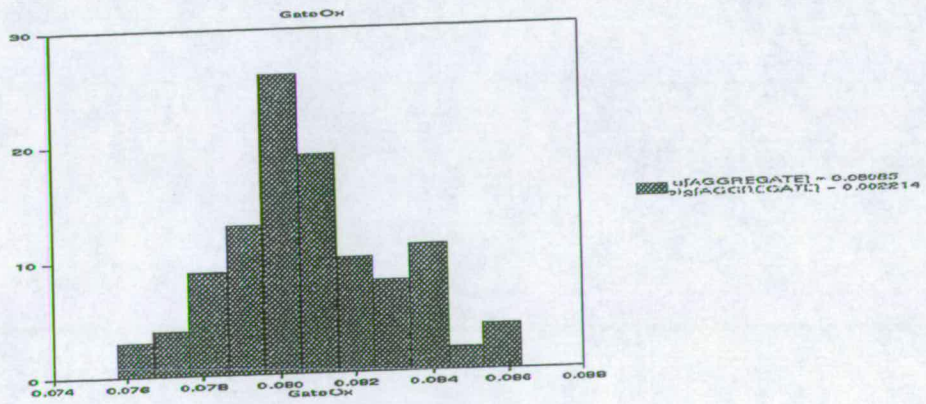
1.0 Abstract

In a drive for higher yielding, more manufacturable processes, a unified approach is needed to ensure success. Concurrently there has been a drive for stable SPICE parameter distributions. In our work we have used a TCAD framework to link together process and device simulation tools and statistical analysis techniques to help improve both yield and our understanding of SPICE distributions. By using Monte Carlo based statistical simulation on a properly calibrated process, it is possible to match distributions of measured Electrical Test, in-line parameters and SPICE parameters to their simulated equivalents. Performing sensitivity analysis gives us insight to which process controls have the greatest impact on parameters of specific interest. Furthermore, by using non-linear Principal Component Analysis (PCA) it is possible to generate SPICE model parameter sets which can be correlated to variations in the manufacturing process. We outline here a technique used in the simulation and calibration of the LMDMOS 300A process run in Fab 3 of NSUK.

2.0 Introduction

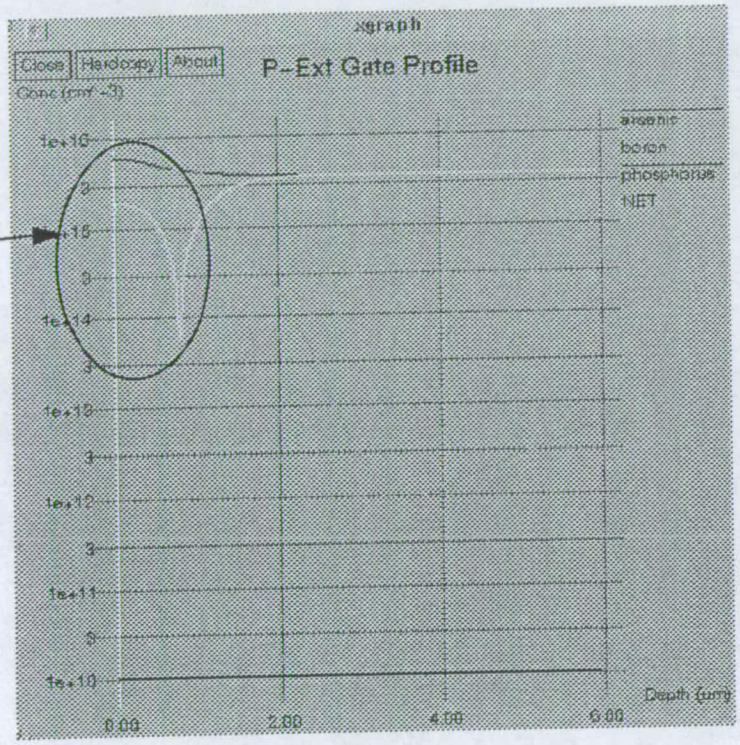
Every process step used in the manufacture of silicon circuits suffers from variation, which can be either random or systematic. By assigning variations to the process controls it is possible to match the in-line distributions seen in fab. The process controls assigned variation in this simulation were temperature, gas flow, pressure, mask size variations and implant dose and energy variations. FIGURE 1 gives a histogram of simulated gate oxide thickness for 150 monte carlo simulations.

FIGURE 1. Simulated Distribution Of Gate Oxide Thickness



After initial variation assignment was complete, a sensitivity analysis was performed. This analysis indicates the individual affects of each process control on final electrical and device parameters. In addition, it can show why a particular device parameter shows extreme behavior.The extended drain PMOS device showed such a behavior. For the threshold voltage, we observe via the sensitivity analysis that a 1-sigma change in either the Vtp dose and epi concentration can have a large impact on the threshold voltage. FIGURE 2 gives the doping profiles of the PMOS device.

FIGURE 2. The doping profiles under the gate of the P-extension device

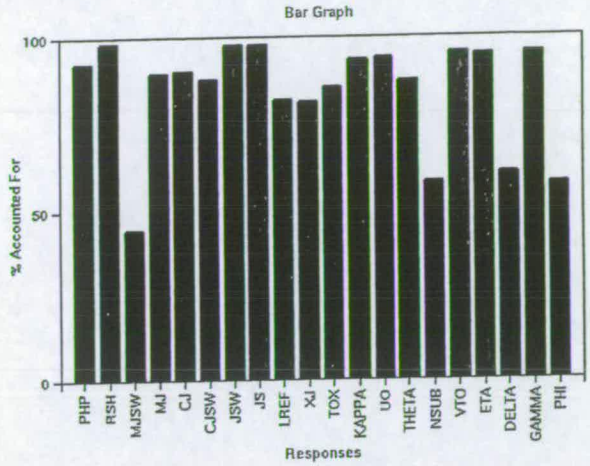


The Boron Vtp implant is just slightly above the epi concentration at the surface (due to pile up).

Once statistical calibration is compete, we have a simulation model which can be used to predict distributions and correlations of electrical test and SPICE parameters. Further correlations can be traced and trends between process parameters and device performance established.

The final step is to use non-linear principal component analysis (PCA). By linking SPICE parameters to electrical test parameters it is possible to extract SPICE model sets based on electrical test data. This PCA technique computes the principal factors for the SPICE distributions and assigns what electrical tests track them. This analysis ensures that the correct number and type of electrical tests are being measured to fully characterize a process. These tracking factors can consequently be measured for every lot and are much less costly than doing a full SPICE extraction.

FIGURE 3. pdPCA result for extended drain PMOS



Tracking factor	% of Total Variance
PMOS Vth	55%
PMOS nsource	15%
PMOS length	9%
Gate Oxide thickness	5%
Total variance accounted for:	84%

In conclusion, a great deal of information can be gained from this process. The sensitivity analysis shows which process controls are critical in controlling device characteristics. The non-linear PCA techniques yield an important mapping of SPICE to electrical test. SPICE parameters can now be extracted far earlier in the development phase, with more confidence and less cost. The TCAD framework also proved out in dealing with a large process with many device types. In addition the procedure we followed is general. The technique outlined here can be applied to any process run within the company and will yield similar valuable information.

Thin Film Characterisation of Silicon Carbide Chrome

T. Powell[1], R. Prasad[2], C. McGregor[2], H. Gregory[2], G. Martin[2], M. Fallon[2], M. Redford[2]

[1] *Process Engineering, Fab 2, U.K.*
[2] *Analog Process Technology Development U.K.*

The material, silicon carbide chrome, is being developed for use as a thin film resistor in precision voltage reference products. The impetus behind the characterisation work was to allow the use of an ellipsometer and interferometer as a fast and reliable film thickness monitor instrument. Calibration curves for the ellipsometer have been produced. To achieve this the refractive index was determined. Subsequently, bulk resistivity values were determined and both parameters appear to depend on the nature of the two deposition techniques used, RF and DC magnetron sputter. Etching experiments and Auger analysis reveal the deposited material to have a non - abrupt interface with the underlying oxide layer.

Introduction

Analog Process Technology Development (APTD), sited in the U.K., is close to completing the development of a new precision voltage reference product, the LM4050A4, an upgrade to the LM4040, for the Data Acquisition Product Group. The new device will replace p-implant resistors with silicon carbide chrome (SiCCr) resistors. These resistors have superior temperature coefficient of resistance (TCR) in the range -40 to -70 ppm / °C compared with implant resistors of the order a few thousand ppm / °C and si-chrome (SiCr) resistors in the region of 300 ppm / °C.

For the purposes of developing SiCCr photo / etch processes, the validity of the measurement instruments and SiCCr physical properties have to be addressed. The use of the interferometer, Nanometrics Nanospec AFT and the ellipsometer, Rudolph Auto EL III, was investigated to measure the SiCCr film thickness. In order to achieve this, the SiCCr refractive index has to be determined. Once this is attained, the SiCCr film thickness can then be determined, allowing establishment of a minimum etch depth and bulk resistivity. This paper covers work on the refractive index, film thickness and bulk resistivity of SiCCr films deposited by two sputter techniques. RF and DC magnetron sputtering are used in the Perkin-Elmer 4400 (P-E) and the Varian 3190 respectively. Only Varian films deposited at 40 °C were investigated. The etch rate for SiCCr, deposited by the Varian, using the standard SiCr LAM 490 SF_6 / O_2 etch process, is determined as well as the layer structure. A more detailed account of this work can be found by contacting the authors.

Refractive index

The ellipsometer can measure the thickness of thin films accurately. However, in order to achieve this, the material's refractive index is required and calibration curves need to be determined. These were determined from absorption experiments and ellipsometer readings of virtual SiCCr substrates. Calibration curves will allow a check on film thickness to be made on production batches and can be coupled with sheet resistance measurements to provide more information on the as-sputtered films. Any change in the refractive index for differing SiCCr film thicknesses implies a change in the material's composition and / or microstructure.

SiCCr films sputtered by the P-E show good correlation for refractive index versus film thickness ranging from 65 Å to 2170 Å. This implies that the material composition / microstructure does not significantly change with film thickness. However, the refractive index for films sputtered by the Varian did not show good correlation with film thickness. Thin film samples upto 400 Å exhibited a different refractive index from a ~1400 Å thick film. Table 1 shows the refractive index of SiCCr determined for films sputtered by both systems.

The ramifications of the varying refractive index for the Varian films are that either the material composition or its microstructure has changed. It may be that the thicker film has a different grain structure arising from possible substrate heating during deposition. In general, the refractive index for films sputtered by the Varian is approximately 80% lower than that for the P-E. A possible explanation is the existence of trapped gas and / or increased silicon / carbon in the sputtered films.

National Semiconductor Corporation Confidential

discounted in the results from Auger analysis. The former point could occur due to the nature of the sputter mechanisms.

Etch rate

The current SiCCr resistor dry etch process used in the development batches is the SiCr process which uses the LAM 490 plasma etchers (SF_6 / O_2). Wafers are etched for 70s, which is sufficient to etch through the SiCCr and into the underlying oxide layer. However, no study until now has been carried out to establish when the SiCCr layer has been etched.

The wafer used in this study mimicked production material that would be used for SiCCr products. This consisted of a Si substrate with SiN_4 then LTO. A Varian SiCCr film was then sputtered onto the wafer and exhibited a nominal sheet resistance of 2060 Ω/square. This value is approximately twice the targeted value for LM4050A4 products. The film thickness may be estimated from sample D's data in table 2. The assumption sheet resistance scales with film thickness is used to estimate the film thickness. However, there is not sufficient data to either refute or support this assumption. The film thickness is estimated as 41 Å.

The Nanospec AFT measured the total film thickness, i.e. of the SiCCr, LTO and nitride thickness, as the wafer was etched. A program was used which assumed an average refractive index of 1.77 for the LTO and nitride films. The refractive index of the SiCCr was not known at the time and so it's refractive index was assumed as 1.77 as well. Calculations were made to account for the true SiCCr refractive index.

The Nanospec AFT film thickness measurements were made at the 5 E.T. sites across the wafer. The measurements were made at 5 second intervals upto a total etch time of 50 seconds. The etch rate was calculated and plotted as a function of etch time, see figure 2. The etch rate data plotted is an average etch rate from the 5 sites. Figure 2 reveals interesting properties of the SiCCr film.

In the first 10 seconds the etch rate falls. This can be understood from the Auger analysis results which show a surface SiO_2 layer mixed with chromium and carbon. This layer is etched at a preferential rate to the pure SiCCr. Once the SiO_2 is etched, the etch rate is a constant, between 10s to 30s, and represents the etching of pure SiCCr. The SiCCr etch rate determined of value 1.26 Å / s, is approximately one sixteenth that of LTO. The plotted 2 Å / s is for a SiCCr film of refractive index 1.77. However, the known value is 2.8, see table 2. This then reduces the etch rate by a factor of 1.77/2.8 to 1.26 Å / s.

After 30s the etch rate begins to climb and eventually there is a sign of the etch rate saturating. This region may be thought of as a transition zone between SiCCr and SiO_2. One would expect the etch rate to be a constant once pure SiO_2 is reached. This is believed to occur beyond 50s. Unfortunately, no more data is available. The transition zone is verified in the Auger analysis results, presented shortly. The consequence of the transition zone is that etching of the SiCCr must continue through to the underlying oxide layer. It should be noted that the development batches have approximately twice the SiCCr film thickness (half the sheet resistance of 2000 ohms /square). It has yet to be confirmed that the transition layer is completely etched. However, if the transition layer was not etched fully and the remaining layer contributed to conduction it might manifest itself in the form of poor resistor matching.

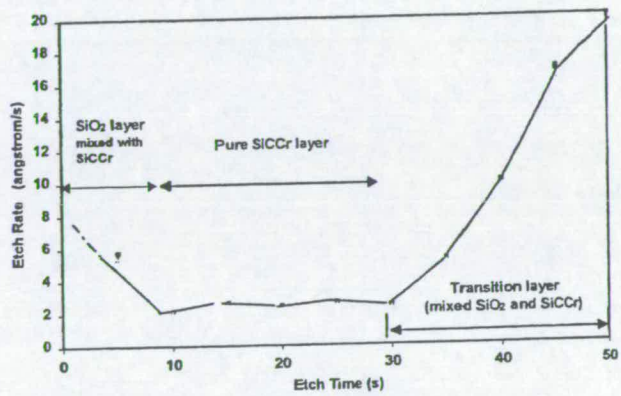


Figure 2 Etch rate as a function of time.

We arbitrarily define the SiCCr film thickness as the difference in the Nanospec AFT film thickness measurement between time = 0s to time = 30s. After time = 30s there is a transition zone where the conductivity is expected to drop as the chromium concentration decays, possibly oxidising. Using the assumed refractive index of 1.77, the thickness is calculated as 90Å. Now if the true refractive index is taken into account of N = 2.8 the corrected thickness is determined as 57 Å. There is some difference between our estimated thickness of 41 Å and the thickness measured by the Nanometrics. In light of this difference, the hypothesis that film thickness scales with sheet resistance may not be true for our SiCCr film. This may be especially true for our SiCCr films because of complications caused by the transition layer.

Auger analysis

Samples were prepared for Auger analysis which was carried out by an outside vendor, Loughborough Consultants Ltd. The purpose of this was two-fold. Firstly, establish that there is a transition layer between SiCCr and SiO_2, i.e. confirm the interpretation of figure 2. Secondly, determine the composition of SiCCr films sputtered by the two systems and try to explain the larger Varian SiCCr bulk resistivity.

A couple of terms have to be explained in order to interpret the results of the Auger analysis of figure 3. The plots show the atomic composition as a function of depth. It should be noted that the depth given should be treated semi-quantitatively as it is estimated from theory. The atomic composition of silicon, carbon, chromium and oxygen are presented. Silicon is found to exist in two states, an oxidised state and in an unoxidised state, labelled Si(o) and Si(e), respectively.

Auger analysis of films sputtered by both systems revealed a general structure as shown in figure 3. There are three regimes: (a) SiO_2 surface layer (b) pure SiCCr layer (c) transition layer between pure SiCCr and SiO_2. This corroborates the interpretation of figure 2. Clearly, there is evidence now for etching through the SiCCr and into the underlying oxide layer, beyond the transition layer.

Comparison of Auger profiles of Varian and P-E SiCCr films revealed no significant difference in the film's composition. The transition layer thickness was determined by carrying out Auger analysis on two samples with different SiCCr deposition times (film thicknesses). The thickness can be extracted using a pair of simultaneous equations and is valid if it is assumed to be a constant regardless of the deposition time. The equations, involve (a) the combined thickness of the surface SiO_2 and the pure SiCCr layers (b) transition layer thickness and (c) the depth at which the chromium concentration decays close to zero.

For P-E SiCCr films the transition layer thickness was determined as 72 Å and the combined SiO_2 and SiCCr layer thickness is calculated as 56 Å. This is a useful result as it points out that the P-E SiCCr film should be over etched by 72 Å, to a total depth of 128 Å for films of sheet resistance 2060 Ω / square.

The higher Varian SiCCr bulk resistivity was postulated earlier on to be explained by one of three means: (a) existence of trapped gas (b) different grain structure (c) increased silicon / carbon content. There is little difference in the atomic composition of silicon and carbon to explain the 75% higher bulk resistivity of Varian SiCCr films. In the light of the Auger analysis, the likely explanation is that Varian SiCCr films have a different grain structure.

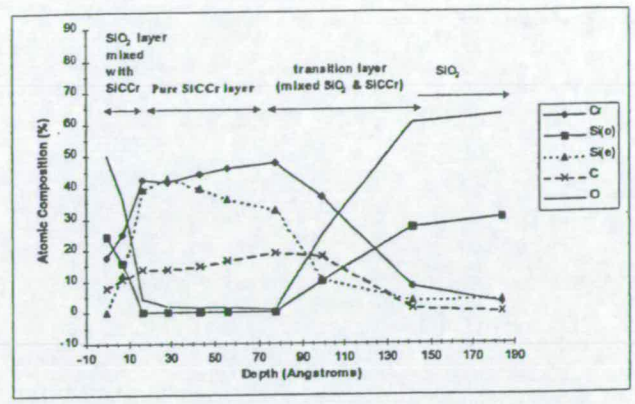


Figure 3 Generic Auger profile of atomic composition as a function of depth.

Summary

We have shown that the Nanometrics interferometer and the Rudolph Auto EL III ellipsometer can measure the SiCCr film thickness. Film thickness of SiCCr layers range in value 60-80 Å for sheet resistances ranging from 700-1200 ohms /square for as-sputtered films, determined by an ellipsometer. The etch rate of Varian SiCCr films, was determined as 1.26 Å/s which is approximately one sixteenth of the LTO etch rate, using the standard si-chrome dry etch process.

The refractive index and bulk resistivity of SiCCr films sputtered by the P-E 4400 and the Varian 3190 have been determined. A lower refractive index and higher bulk resistivity of Varian SiCCr films is observed. Contamination of the films by trapped argon gas is unlikely to be the sole cause, though it may contribute to some of the discrepancy. The possibility of increased silicon / carbon content in the Varian SiCCr films is discounted as there is little evidence from the Auger analysis data to suggest this. However, subtle differences in the silicon / carbon content will contribute to some of the discrepancy. The most likely explanation is that the microstructure is likely to be responsible for the observed discrepancy, with the other two explanations playing a minor role.

Results from etching and Auger analysis reveal three regimes in the SiCCr film. The top most layer is a surface SiO_2 layer mixed with SiCCr, the underlying layer is pure SiCCr followed by a transition layer composed of a SiO_2 mixed with SiCCr. The ramification of this is that an over etch is required in order to remove potential current leakage between resistors.

Proc. IEEE 1997 Int. Conference on Microelectronic Test Structures, Vol 10, March 1997. 156

Novel Structure to Measure Emitter-Base Misalignment

M. Fallon, M Redford
National Semiconductor
Larkfield Industrial Estate
Greenock
Scotland PA16 OEQ
email:martinf@dundee.nsc.com

K. Findlater, M Newsam
Dept of Electrical Engineering
University of Edinburgh
Kings Buildings
Mayfield Road
Edinburgh
Scotland

Abstract
This paper addresses the issue of the alignment between emitter and base of a bipolar process. Current electrical alignment structures usually rely an electrical contact between the layers or a division of one layer by the other. In the case of a bipolar emitter-base structure, neither approach is directly applicable. Here we present an alternative to solve this problem.

1. Introduction

Mask registration has been extensively studied and many alternatives proposed to measure this electrically. In many cases direct electrical contact is made between the layers and in others one layer is 'split' by the other, forming two regions from the initial single layer. In a bipolar transistor the emitter sits inside the base region, making a diode contact but leaving a continuous conducting base region beneath. This then requires an alternative approach to the problem.

2.0 The Test Structure

The layout of the emitter-base region is shown in figure 1a and the corresponding cross-section in figure 1b. Should the emitter region diffuse through the base two isolated base regions would be formed and their respective linewidths measured electrically, with the difference in linewidths yielding the misalignment. However, rather than rupturing the base region, the emitter forms a high resistive region known as the pinched base under the emitter. Thus 3 resistive regions are formed within the base: two low resistance extrinsic and one high resistance intrinsic.

If current is injected as shown in figure 1a there will be a non-uniform current density, depending on the ratio of resistances. Suitably placed voltage taps should be able to detect the degree of non-uniformity and provide information on the registration between the layers. A further degree of flexibility is possible by operating the structure as a JFET, gating the intrinsic base by applying a reverse bias across the emitter-base and the collector-base

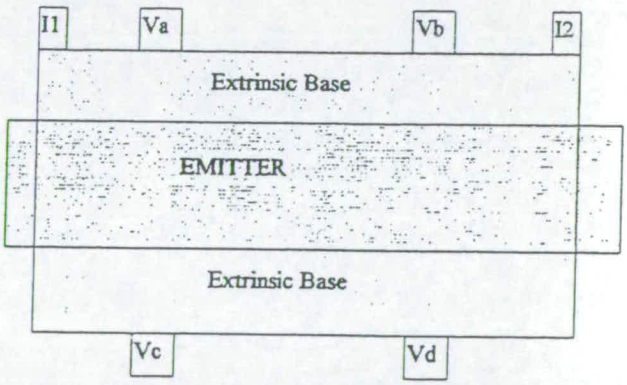


Figure 1a. Layout of bipolar emitter-base region.

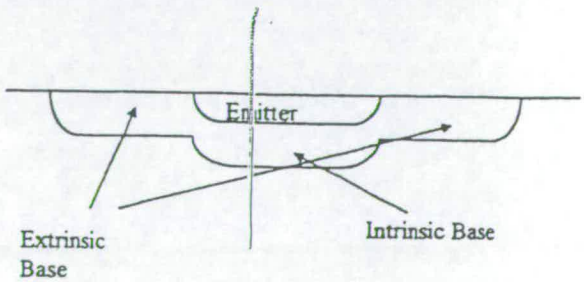


Figure 1b. Cross-section through transistor

junction to widen the depletion layers and narrow the intrinsic base, thus modulating the resistance in that region.

0-7803-3243-1
97CH35914

3.0 Simulations

Simulations using RAPHAEL® were performed to investigate the current distribution. Figure 2 shows the current density plot for a 100Ω/□ extrinsic base and a 5000Ω/□ intrinsic base. The emitter is 3 microns wide and the base 9, providing three base regions each 3 microns wide. The length of the structure is 100 microns. Figure 3 shows the current density through a vertical cut 10 microns from the centre of the device for three different misalignments. The current can be seen to split in the three regions. The majority is in the low resistance extrinsic base region closest to the current injection. Different current conditions exist according to different misalignments.

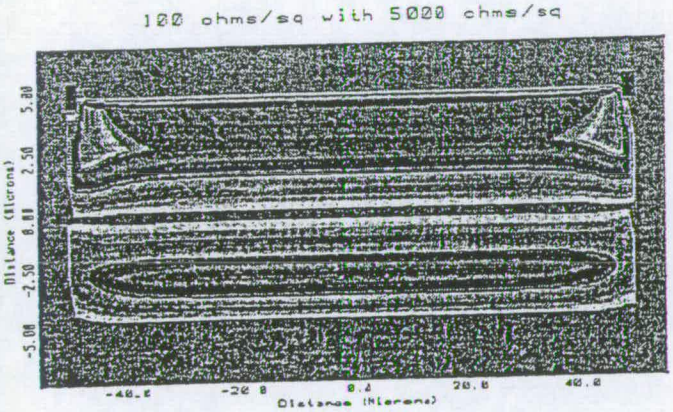


Figure 2. Current density plot

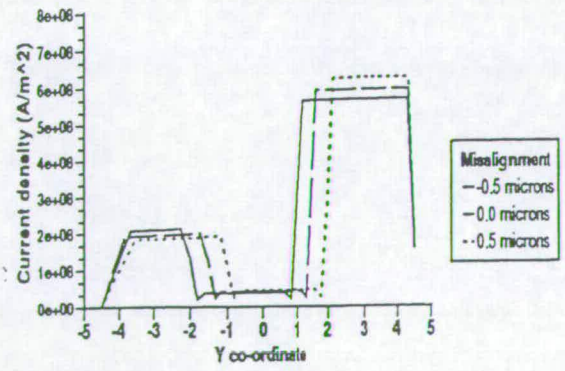


Figure 3. Current density along a vertical cut.

4.0 Experimental Results

Figure 4 shows a photograph of the fabricated device. The voltage taps were placed 200 microns apart. No measurable difference was found during test suggesting equipotentials exist between Va and Vc (and Vb and Vd). If the current is entirely in the x direction then the potential at taps Va and Vc must be equal and similarly with Vb and Vd. Therefore Vab = Vcd and while the current is distributed as shown in the RAPHAEL simulations the measured voltages are identical. The structure was then gated in an attempt to isolate the extrinsic base regions.

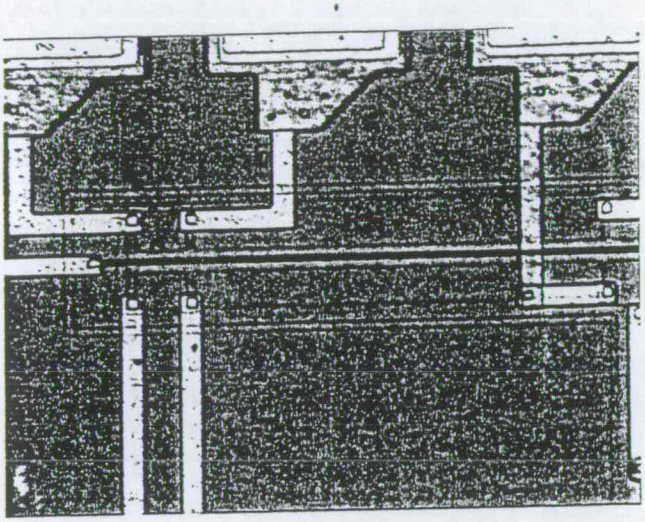


Figure 4. The fabricated device.

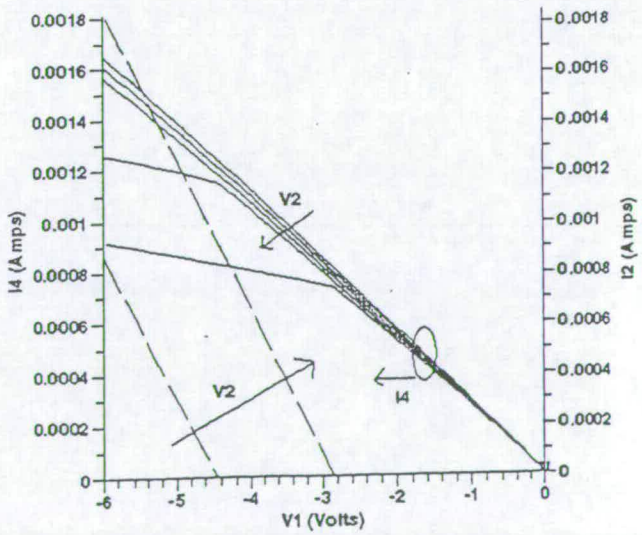


Figure 5. Current through intrinsic base (I4) against voltage applied (V1) for varying emitter-base voltages (V2) and emitter-base current (I2).

Figure 5 shows the relevant IV graphs. To measure the amount of modulation a voltage, V1, was swept form 0 to -6 volts at node C (see figure 1) for a series of voltage

steps (V2) from 0 to 4 volts applied to the emitter region. Currents I4 and I2 were measured between nodes C and A and between the emitter and A respectively. The arrows on the graph refer to increasing V2 values. Only 2 lines are visible for I2 corresponding to voltages of 3 and 4 volts. The first three IV lines for I4 show no current modulation. When the voltage on the emitter is increased to 3 and 4 volts the current through the intrinsic is reduced but only when the emitter base junction starts to pass current. Given these conditions the JFET operation was unsuccessful in controlling the intrinsic region.

5.0 Modified Structure

To overcome the limitations of the test structure, a modified structure is proposed and the layout can be seen in figure 6. Two base regions are defined such that there exists two back-to-back diodes, providing electrical isolation between them. When the emitter is defined the resistance of the base is modulated in a similar fashion as described previously. However each region can now be measured in isolation because of the diode configuration now present. This structure has been modelled using the schematic shown in figure 7, varying the resistances of the extrinsic and intrinsic base regions. The parasitic series resistances have been ignored in the calculation as these are present in each structure and symmetrical. The results are shown in figure 8. It can be seen that this modified structure exhibits sensitivity similar to the standard electrical designs..

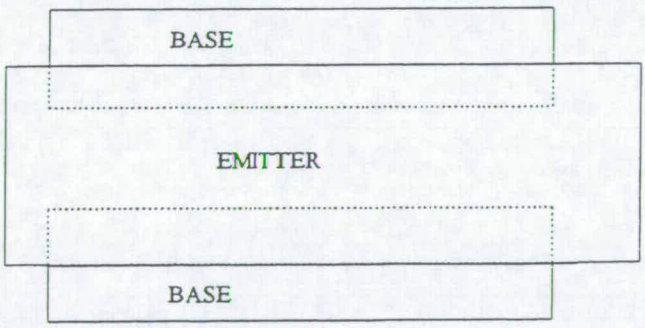


Figure 6. Modified alignment structure. showing separate base regions.

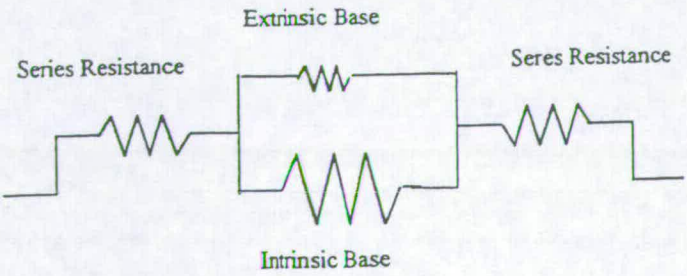


Figure 7. Equivalent circuit of upper base region.

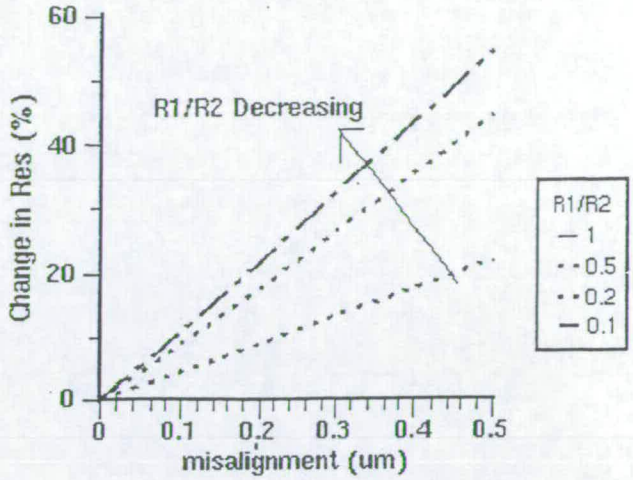


Figure 8. Percentage change in measured resistance for varying intrinsic (R1) to extrinsic (R2) resistance ratios.

6.0 Conclusions

Several variations of a test structure have been considered and presented. The best of these appears to be the modified structure which translates the unique configuration of the bipolar transistor into the more recognisable alignment structure commonly applied today. The gated alignment structure may be dependent on the relative doping levels of the emitter and base and so may be limited in application.

References.

1. RAPHAEL. Technology Modelling Associates.

Electrical Assessment of Planarisation for CMP and SOG Technologies

J.P. Elliott, M.Fallon†, A.J. Walton, J.T.M. Stevenson, A.O'Hara, M. Redford‡A. Shaffi†

Edinburgh Microfabrication Facility
Department of Electrical Engineering
Kings Buildings, University of Edinburgh
Edinburgh, EH9 3JL, UK

†National Semiconductor
Larkfield Industrial Estate
Greenock
PA16 0EQ, UK

Abstract

Experimental measurements of an electrical test structure that can be used to assess the degree of planarisation of inter-layer dielectrics are presented. The test structure consists of two sets of metal combs separated by a dielectric. For each structure the combs on the two layers overlap each other with adjacent structures having the overlap in one direction progressionally offset by 0.2μm. Measurement of Chemical Mechanical Polishing (CMP) and SOG technologies will be compared.

1. Introduction

Simulations have indicated that a proposed structure can be used to electrically determine dielectric planarisation [1]. This paper will compare measurements of both CMP [2] and SOG planarised interconnect.

2. Test Structure

The basic test structure consists of 2 identical metal combs, separated by a layer of dielectric and has previously been described [1,2]. Measurement of capacitance as a function of teeth misalignment enables the degree of planarisation to be assessed.

3. Simulation Results

Two dimensional capacitance simulations have been used to examine the relationship between the inter-comb capacitance and a range of offsets for planarities ranging from 0% to 100%. The simulations were performed using TMA's Raphael software and were carried out for a small section of the structure comprising of 10 teeth from each metal layer. Due to the repetition of this section within the whole structure, the capacitance between the complete combs could be determined by scaling the results for the smaller section by a factor of 10, thus greatly reducing the CPU time. Typical times to simulate the capacitance of this small section for a given offset took around 2 minutes on a Sun IPX workstation. The data describing the structure being simulated was generated semi-automatically by using equations with variable parameters in the Raphael input file. The number of grid points used in each simulation was 15,000.

Figure 1 shows shows the relationship between the comb capacitance and the offset for a slope of 56°. For CMP technology as the dielectric is planarised, the oxide thickness directly above metal1 decreases as the raised oxide is polished away. This accounts for the increase in zero offset capacitance as the dielectric is planarised. It can be observed that this increase in capacitance becomes smaller as the finger overlap reduces. However, the shape of the curve is dominated by the topography of the inter-metal dielectric, which enables the degree of planarity to be extracted. It can be seen from figure 1 that with increasing planarity, the gradient of the central part of the curves decreases progressively from positive to negative.

4. Experimental Work

Four wafers with different degrees of planarisation were processed. Wafers 1 and 3 had a thin base layer of silicon nitride, whilst wafers 2 and 4 had a thicker layer of oxide covered by a layer of silicon nitride. The aluminium forming the lower teeth in the structures was 1μm thick and the inter-metal dielectric was deposited using Electron Cyclotron Resonance (ECR). The conditions

Simulations using 2.4μm teeth on a 6.0μm pitch resulted in slightly larger peaks at maximum misalignment but certainly not approaching the magnitude of that plotted in figure 2(d). Since two layers of ECR oxide were deposited using different processes the dielectric constant of the layers will be different (In fact the two layer can be identified in the cross-sections of figure 3). As a consequence the effective dielectric constant will change depending on the amount of polishing a wafer has received. Again this was simulated but this did not explain the results measured on wafer 4.

6. Conclusions

The test structure presented provides a means of electrically testing the degree of planarisation of inter-layer dielectrics. It depends on the fact that the capacitance between two metal combs will vary as the horizontal offset between them changes, and that this variation will depend on the shape of the oxide separating them. The gradient of a graph of capacitance between the combs as a function of offset can be used to estimate the degree of planarisation with negative gradients indicating planarisation greater than 70%. Further work will be performed to determine the cause of the peak observed at maximum misalignment for the 100% planarised sample. The final paper will also present and compare results for SOG planarisation assessed using the progressional offset capacitance structure.

7. Acknowledgements

The authors would like to acknowledge the support of TMA who supplied the RAPHAEL software used in the simulations and EPSRC (Grant no GR/J45284) for their financial support.

8. References

[1]. J.P. Elliott, M.Fallon, A.J. Walton, J.T.M. Stevenson, A.O'Hara, "A Test Structure for the Measurement of Planarisation", IEEE International Conference on Microelectronic Test Structures, Trento, pp 301-305, March 1996.

[2]. J.P. Elliott, M.Fallon, A.J. Walton, J.T.M. Stevenson, A.O'Hara, "An Electrical Test Structure for the Measurement of Planarisation", submitted to IEEE Trans Semiconductor Manufacturing.

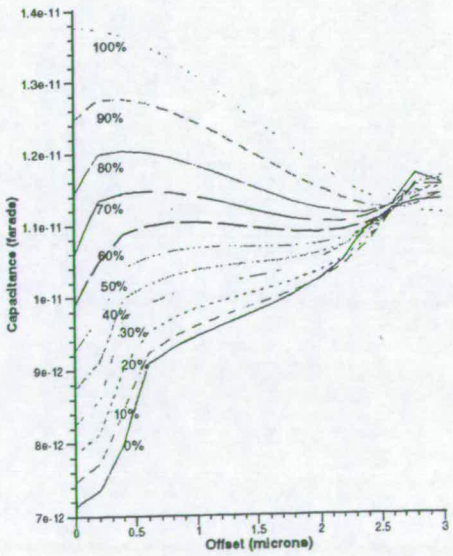


Figure 1. Capacitance as a function of offset for an oxide step of slope 56° for planarities ranging from 0% to 100%.

used for the ECR deposition were selected to ensure complete filling of the trenches between the metal1 teeth and a sufficient oxide thickness for the polishing step that followed. The ECR recipe consisted of an initial deposition using the 'planar' mode (pressure = 7 mTorr, RF power = 280W and table height = 80cm) to ensure that trenches would be filled. This was then followed by a 1 hour deposition in the 'non-planar' mode, to produce the required oxide thickness.

Following the via etch, three of the four wafers were polished, using a Logitech PS2000 system. The slurry used was alkaline colloidal silica, with an expanded polyurethane pad. The pad rotated at 30 rpm, with a downward pressure on the wafer of 0.073 $Kg\,mm^{-2}$. Wafer 1 was not polished, wafer 2 received a 5 minute polish, wafer 3 a 10 minute polish and wafer 4 a 24 minute polish. After polishing, a Nanospec was used to produce a contour map of the oxide thickness of each wafer to determine the global uniformity of the dielectric. Topographical measurements of the wafer surface were also made at individual sites to determine the degree of local planarity using a Dektak surface profiler.

The wafer maps indicated that the oxide thickness was not being polished at a constant rate with variations of 0.25μm on wafer 1, 0.35μm on wafer 2, 0.8μm on wafer 3 and 1.2μm on wafer 4. The rate at which oxide had been removed was found to be greatest at the edges of the wafers. It should be noted that there was a variation in the ECR oxide thickness on the wafers prior to polishing. Local topographical measurements taken with the Dektak surface profiler indicated that the maximum vertical feature sizes of the wafers after polishing were 0.5 - 0.65 μm on wafer 2, 0.2 - 0.5 μm on wafer 3 and 0.02 - 0.06 μm on wafer 4. The Dektak results are useful only as an indication of the relative planarity of the wafers since the 12μm stylus used was larger than individual features on the wafer surface.

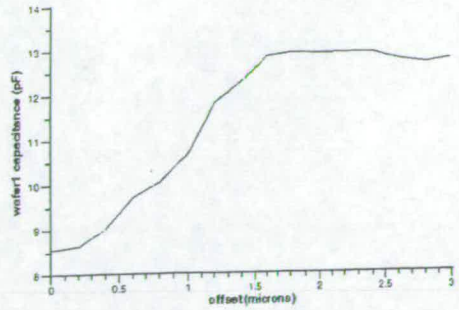
After these measurements 1μm of aluminium was deposited. When patterning this layer, the resist was deliberately overexposed to ensure that no resist remained in the oxide trenches resulting in a reduced metal2 linewidth.

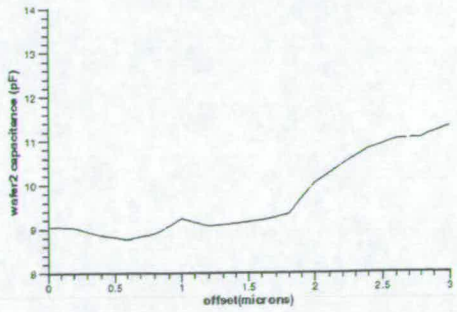
5. Measurement and Experimental Results

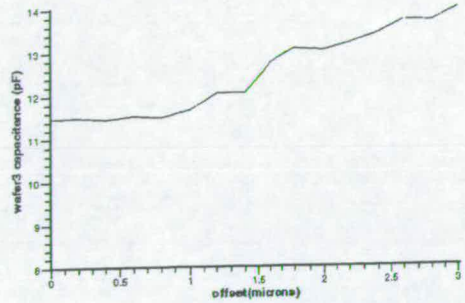
Using the oxide thickness contour maps taken prior to metal2 deposition, sites were chosen on each wafer at which the oxide thickness was similar. The capacitance between the two layers of metal was measured for each set of test structures, and the relationship between offset and capacitance plotted. To ease the comparison of the measured and simulated results, the point was determined where each graph was symmetrical, and the mean of the results to either side of this midpoint plotted. These results are shown in figure 2 with the corresponding wafer cross-sections shown in figure 3.

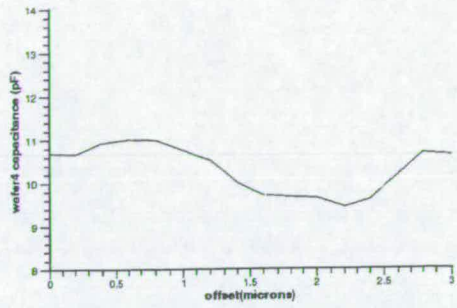
Inspection of the SEM cross-sections indicates that the 56° slope is a reasonable approximation to fabricated structures. Hence, the simulation curves shown in figure 1 have been compared with the curve shapes in figure 2 to determine the degree of planarity. Inspections of figure 2(a) shows a curve which indicates that the planarity is around 0%. Figures 2(b) and 2(c) appear similar to one another and an initial inspection indicates that they lie in the 30% to 50% planarity range. However, figure 2(b) reaches an increase in gradient at a higher offset than that in 2(c) indicating that the planarity of wafer 2 is less than wafer 3. Comparisons with the simulations suggests that the planarity of wafer 2 is around 30%, whilst that of wafer 3 is around 50%. The SEM cross sections of wafers 1, 2 and 3 show a correlation between these results and the observed planarity. Figure 2(d) has a negative slope and before sectioning we estimated that the planarity was in the region of 70% to 80%. However, the cross-section shows that the planarity is actually nearer to 100%. The large increase in capacitance as the teeth approach maximum misalignment is unexpected. This is not predicted by the simulations although it should be noted that there are small peaks at maximum misalignment for the 60-90% curves but nothing of this magnitude.

It should be realised that the simulated structures are at best an approximation to those fabricated. These assumed 3.0μm teeth on a 6.0μm pitch with a dielectric slope of 56°. The shape of the slope is a fairly crude approximation and the teeth on the fabricated structures were 2.4μm wide.









(c) Wafer 3. (d) Wafer 4.

Figure 2. Capacitance as a function of offset.

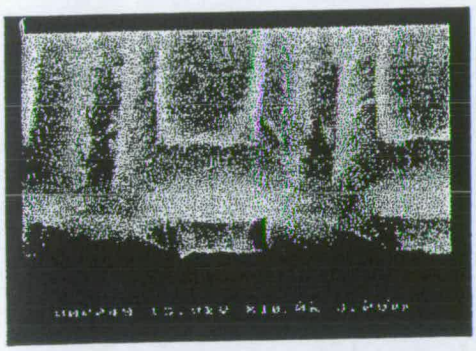


Figure 3(a). Cross-section of wafer 1.

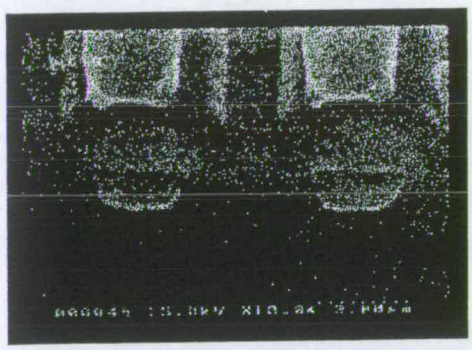


Figure 3(b). Cross-section of wafer 2.

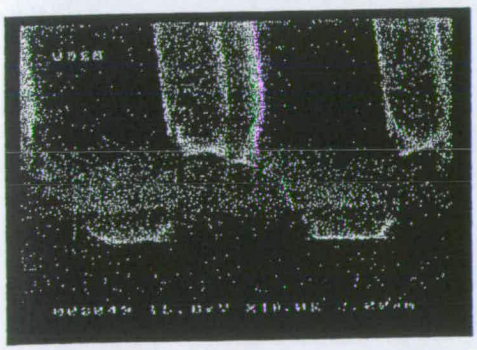


Figure 3(c). Cross-section of wafer 3.

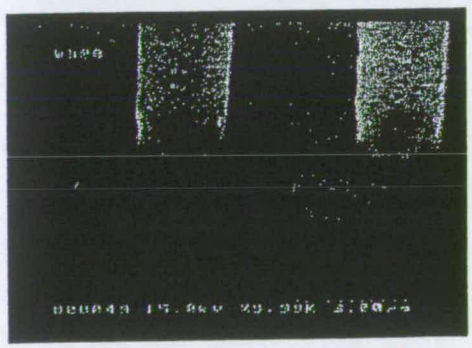


Figure 3(d). Cross-section of wafer 4.

High Spatial Resolution Microanalysis of Platinum Silicide Contact Spiking in a STEM.

Alastair McGibbon, Calum McLeod,Gordon Robertson, Mark Redford, Kevin Yallup, Christine Buchanan, Pat Cavanagh and Colin Elgey
National Semiconductor

Pat Nicholson Stephen McVitie
University of Glasgow

In this paper, we discuss the detection and subsequent observation of Platinum Silicide spiking in doped silicon epilayers at contact structures. Spiking can seriously affect electronic properties such as subthreshold leakage in CMOS and emitter-base leakage in bipolar devices. However, although observations have been made in polysilicon, less work has been carried out on the nucleation and propagation of spikes in doped single crystals.

The analysis of spikes, which can be <10Å wide but can propagate as much as a 1μm into the substrate, requires the application of high spatial resolution (<10Å) compositionally sensitive imaging and microanalysis in a dedicated scanning transmission electron microscope (STEM). To facilitate the analysis of specific device regions, electron transparent cross-sectional membrane of fabricated devices were prepared using a focused ion beam (FIB) facility. This approach provides accurate selectivity of regions of interest and precise control of specimen thickness for compositional microanalysis.

The principal techniques applied in the STEM are energy-dispersive x-ray (EDX) microanalysis and high-angle annular dark-field (Z-contrast) imaging. We discuss the way in which these techniques can be used to complement each other to provide information on spiking nucleation together with defect density, orientation and composition as a function of fabrication conditions. We conclude the paper with a general discussion on the way in which such analyses can be used in future in the general study of silicide/silicon interfaces in ULSI devices.

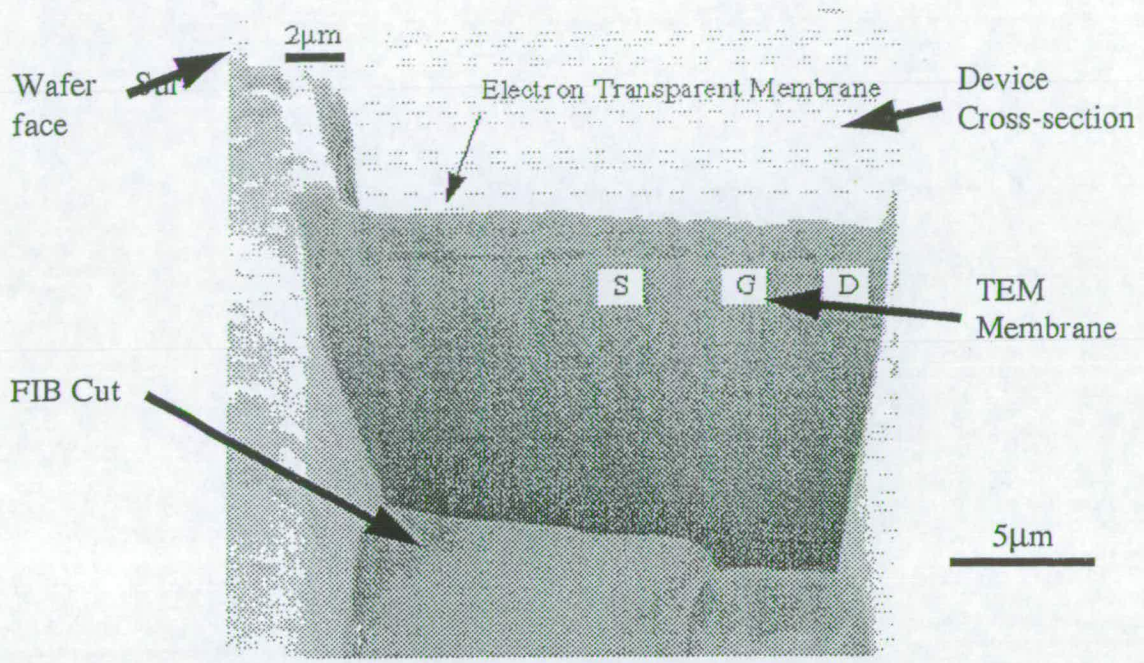


Figure 1. Image of a 500Å thick TEM cross-section of an NMOS device prepared in a FIB.

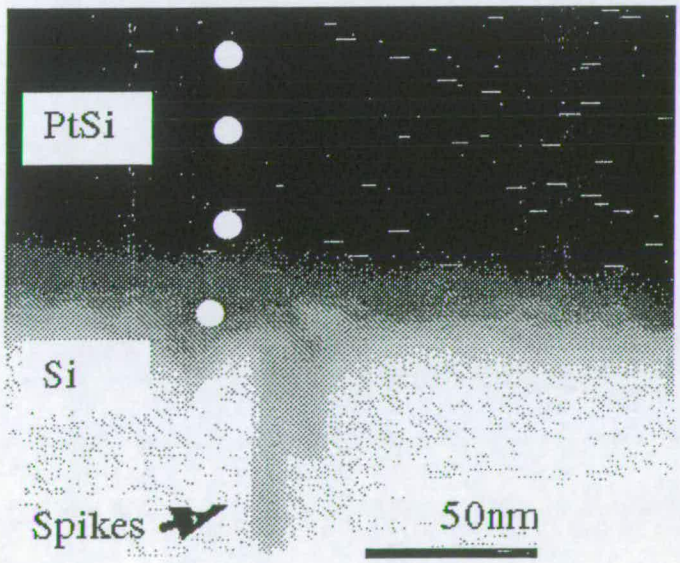


Figure 2. Compositionally sensitive image (high atomic number = low intensity) image of a 50nm long PtSi spike extending from the PtSi/Si interface. The white spots represent the points at which X-ray spectra were acquired to confirm local compositions.