LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH
Ollscoil Átha Cliath

TRINITY COLLEGE LIBRARY DUBLIN
The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

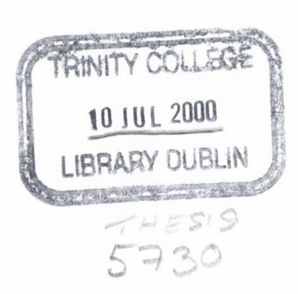
I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Fabrication of Ultra-shallow Boron Junctions using Proximity Rapid Thermal Diffusion

by
Martina Nolan

A thesis submitted for the degree of
Doctor of Philosophy of the
University of Dublin

Department of Electronic and Electrical Engineering
University of Dublin, Trinity College, Ireland


TRINITY COLLEGE
10 JUL 2000
LIBRARY DUBLIN
THESIS
5730

Declaration

I, the undersigned, declare that this thesis is entirely my own work, except where otherwise accredited, and that it has not been submitted for a degree at any other university or institution. I furthermore agree that the library may lend or copy this thesis upon request.

Signed,

Martina Nolan

30th of May 2000

To Richard and my parents

Summary

This work presents proximity rapid thermal diffusion (RTD) as a technique for fabricating shallow *p*-type junctions for LDD devices. Boron-doped spin-on dopant (SOD) is used as a dopant source in proximity RTD. Fourier transform infrared (FTIR) spectroscopy and spectroscopic ellipsometry are initially used to analyse and optimise the SOD dopant source.

Bare and oxidised silicon wafers are doped using the optimised SOD dopant source. The boron junctions are analysed using SIMS and the four-point probe technique. Boron diffusion coefficients are evaluated from the SIMS profiles using Boltzmann-Matano analysis.

Proximity RTD is used to fabricate *p-n* junction diodes. The diodes exhibit excellent current-voltage (IV) characteristics, with near ideal forward characteristics, low leakage currents and sharp avalanche breakdown voltages.

Micro-Raman spectroscopy (MRS), synchrotron X-ray topography (SXRT) and optical microscopy are used to determine the thermal stress and the boron-induced strain that is generated in wafers during rapid thermal oxidation (RTO) and rapid thermal diffusion. The thermal stress deteriorates with time at the peak temperature. The introduction of boron also generates extra strain in the lattice. However, the stress that is generated for very short process times is minimal and does not degrade device performance.

Finally, a LDD process that is compatible with proximity RTD is proposed.

Acknowledgements

I would like to express my gratitude to several people who have helped to make this thesis possible.

Firstly, I would like to sincerely thank my supervisor, Dr. Alan Moore, for giving me the opportunity to undertake this research project and for his invaluable guidance and support. I would also like to thank Dr. Moore for encouraging me to participate in several stimulating international conferences.

I would like to thank Dr. Tania Perova for her endless encouragement and inspiration, and also for her assistance with FTIR analysis. Thanks also to Prof. J. Vij for use of his FTIR spectrometer.

This research project has involved collaborative work with numerous other groups. Above all, I would like to express my sincere gratitude to Prof. Harold Gamble, Queen's University Belfast, for his wonderful ideas and inspiring discussions, and also for providing me with access to all of the facilities in the Northern Ireland Semiconductor Research Centre (NISRC). I would also like to thank all of the friends that I have made in the NISRC for their hospitality and technical assistance, in particular, Dr. Paul Baine, Dr. Liam Quinn, Dr. S.H. Raza, Mr. Denis Campbell and Dr. John Montgomery for helping me to fabricate and test my *p-n* junction diodes.

Many thanks to Prof. John McGilp and Dr. Carlos Beitia, Physics Department, TCD, for their expertise in spectroscopic ellipsometry analysis and for many useful discussions. I would also like to thank Dr. Patrick McNally, Dublin City University, for giving me the opportunity to analyse my samples using synchrotron X-ray topography. Thanks also to Sean Hearne and Dr. Sverre Lidholm, National Microelectronics Research Centre, for their help with SIMS analysis.

Intel Ireland is gratefully acknowledged for their financial support of this research project. Thanks to Philip Moynagh, Dave Selkirk, Sarah McEvoy and Mike O'Connor in Intel for their assistance during the past few years, and in particular, thanks to my mentor Barry Kennedy for his advice and for the interest that he has shown in my work.

Special thanks to my friends Philippe Duval, Gerard Joyce, Vincent Ficet, Stephanie Kremer and Cormac Moore in the Fabrication Lab, TCD, for their support over the years.

Thanks to the staff of the Department of Electronic and Electrical Engineering for their encouragement and assistance.

I would like to thank my parents, Tom and Helen, for giving me the best of everything, including a wonderful education, and for supporting me in everything that I do. And last but not least, I would like to thank Richard Hunter for helping me in so many ways, and for always putting a smile on my face.

Table of Contents

Declaration ii

Dedication iii

Summary iv

Acknowledgements v

Table of Contents vii

List of Figures xi

List of Tables xvi

List of Symbols xvii

Acronyms and Abbreviations xix

Author's Publications xxi

Preface 1

1 MOS Transistor Device Physics **3**

1.1 Introduction 3

1.2 MOS Capacitor 4

1.2.1 Accumulation 5

1.2.2 Depletion 5

1.2.3 Inversion 5

1.2.4 Strong Inversion 5

1.3 Long-Channel MOSFET 6

1.3.1 Subthreshold Currents in Long-Channel MOSFETs 9

1.4 Short-Channel MOSFET 10

1.4.1 Hot-Carrier Effects 10

1.4.2 Drain-induced-barrier-lowering (DIBL) 13

1.4.3 Punchthrough 14

1.5 Lightly Doped Drain (LDD) Devices 15

1.6 Conclusions 18

2 MOSFET Doping Technology 21

2.1 Introduction 21

2.2 Ion Implantation 21

2.3 Rapid Thermal Diffusion 24

2.4 Conclusions 26

3 Proximity Rapid Thermal Diffusion 29

3.1 Introduction 29

3.2 Proximity Rapid Thermal Diffusion Process 29

3.2.1 Experimental Procedure 31

3.3 Removal of Contamination from Spin-on Dopants 33

3.4 Spin-on Dopants 35

3.4.1 Spin-on Dopant B155 35

3.4.2 Spin-on Dopant B153 36

3.5 B153 Doping Effectiveness 37

3.6 Conclusions 40

4 Stabilisation of B153 Spin-on Dopant Source 42

4.1 Introduction 42

4.2 Experimental Procedure 43

4.2.1 Proximity Rapid Thermal Diffusion 43

4.2.2 FTIR Measurements 44

4.2.3 Spectroscopic Ellipsometry Measurements 45

4.2.4 SIMS and Sheet Resistance Measurements 45

4.3 Results and Discussion 46

4.3.1 FTIR 46

4.3.2 Spectroscopic Ellipsometry 50

4.4 B_2O_3 Transfer to the Product Wafer 55

4.4.1 B_2O_3 Step Coverage on the Product Wafer 55

4.4.2 B_2O_3 Layer Removal 55

4.5 Conclusions 55

5 Boron Diffusion & Properties of Proximity RTD Doped Wafers 60

5.1 Introduction 60

5.2 Boron Diffusion 61

5.2.1 Diffusion Mechanisms 61

5.2.2 Fractional Interstitialcy Component 63

5.2.3 Oxidation-Enhanced Diffusion 64

5.2.4 Transient-Enhanced Diffusion 65

5.2.5 Other Anomalous Diffusion Effects 66

5.3 Enhanced Diffusion of Boron from Borosilicate Glass during RTD 66

5.4 Properties of Proximity RTD Doped Product Wafers 67

5.4.1 Sheet Resistance Measurements and SIMS Analysis 67

5.4.2 Diffusion Coefficients 71

5.4.3 Comparison with other Fabrication Techniques 74

5.5 Conclusions 75

6 Electrical Characterisation of *p-n* Junction Diodes 78

6.1 Introduction 78

6.2 Fabrication of *p-n* Junction Diodes 79

6.3 The Ideal Diode Equation 81

6.4 Forward IV Characteristics 83

6.4.1 Surround Diodes (SD) 83

6.4.2 Open Diodes Type 1 (OD1) 87

6.4.3 Open Diodes Type 2 (OD2) 88

6.4.4 High Level Injection and Series Resistance 89

6.5 Reverse IV Characteristics 91

6.5.1 Influence of Junction Curvature on Breakdown Voltage 91

6.5.2 Open Diodes Type 2 (OD2) 94

6.5.3 Surround Diodes (SD) 97

6.5.4 Open Diodes Type 1 (OD1) 99

6.6 Conclusions 100

7 Stress Generated in Silicon during Rapid Thermal Processing 102

7.1 Introduction 102

7.2 Micro-Raman Spectroscopy (MRS) 103

7.2.1 Experimental Procedure 104

7.2.2 Extracting Information from Raman Spectra 104

7.3 Synchrotron X-Ray Topography (SXRT) 105

7.3.1 Experimental Procedure 106

7.4 Stress Generated during Proximity Rapid Thermal Processing 107

7.4.1 Thermal Stress 108

7.4.2 Boron-Induced Strain 111

7.5 Conclusions 114

8 LDD and Proximity RTD Process Integration 116

8.1 Introduction 116

8.2 Conventional LDD Device Fabrication 116

8.3 Limitations of Conventional LDD Devices 118

8.4 Improvements to the Conventional LDD Structure and Proposed Integration of Proximity RTD 120

8.5 Conclusions 124

9 Conclusions and Suggestions for Future Work 126

9.1 Conclusions 126

9.2 Future Work 128

Appendix

A Mask Set for *p-n* Junction Diodes 130

B Process Flow for *p-n* Junction Diodes 134

List of Figures

1.1	(a) Perspective view of a MOS-C, (b) Cross-section of a MOS-C.	4
1.2	(a) Perspective view of a MOSFET, (b) Cross-section of a MOSFET.	6
1.3	(a) I_D-V_{DS} characteristic of a MOSFET being operated at a fixed $V_{GS}>V_T$; (b) V_{DS}=0 V; (c) inversion layer narrowing under moderate V_{DS} biasing; (d) pinch-off, $V_{DS}=V_{SSAT}$; and (e) post pinch-off, $V_{DS}>V_{DSSAT}$.	8
1.4	I_D-V_{DS} characteristic for a long-channel MOSFET.	9
1.5	Cross-section of a PMOS transistor showing hot-carrier processes.	11
1.6	Hot electron injection into the gate oxide in a PMOS transistor.	12
1.7	(a) Reverse biased *pn* junction; (b) Cross-section of a PMOS transistor showing electric field lines from the channel charge terminating on acceptor ions in the source and drain depletion regions.	13
1.8	Cross-section of a LDD PMOS transistor.	15
1.9	Magnitude of the electric field at the Si-SiO_2 interface as a function of distance: L = 1.2 μm, V_{DS} = 8.5 V, $V_{GS} = V_T$ [2]	16
1.10	Device lifetime versus reciprocal stressing drain voltage for a conventional PMOSFET and a LDD PMOSFET [1].	16
1.11	Punchthrough voltage as a function of boron implant dose for a 0.4 μm PMOSFET with and without an LDD structure [3].	17
3.1	Wafer Configuration in Proximity RTD.	30
3.2	Sitesa rapid thermal processor.	31
3.3	Schematic representation of Sitesa RTP. a) Side-view of RTP; b) Plan-view of the quartz support.	32
3.4	FTIR spectra of the source wafers after 15 minutes at various post-spin baking temperatures.	34
3.5	SIMS Carbon and Boron profiles of a product wafer that was doped with B155 during a RTD at 1050 °C for 16 s.	35

3.6 SIMS Boron and Carbon profiles of a product wafer that was doped using B153 in RTD 1050 °C for 16 s. 36

3.7 FTIR spectra of product wafers that were doped at 1050 °C for 16 s. Each SOD source received a different thermal treatment prior to diffusion. 38

3.8 SIMS profiles of junctions that were formed by diffusion of boron from a SOD gel and a cured SOD layer. 39

4.1 IR absorbance spectra of SOD layers which were annealed at 800 °C for various times. 46

4.2 IR absorbance spectra of SOD layers which were annealed at 900 °C for various times. 47

4.3 IR absorbance spectra of SOD layers which were annealed at 1000 °C for various times. 47

4.4 The dependence of the peak intensity of Si-O-Si stretching vibrations band on RTC temperature and time. 48

4.5 The dependence of the peak position of B-O stretching vibrations band on RTC temperature and time. 49

4.6 The dependence of the B-O/Si-O infrared peak ratio on curing time for different temperatures. 50

4.7 The dependence of the refractive index of SOD layers on RTA temperature and time. 51

4.8 Variation in thickness of SOD layers (ellipsometry measurements) with changing RTA temperature and time. 53

4.9 SIMS B concentration profiles of two silicon wafers that were doped by a single dopant source in proximity RTD. 54

4.10 SEM micrographs (x100k) (a) Plasma-etched polysilicon gate structure; (b) Polysilicon gate covered in a conformal B_2O_3 layer after proximity RTD. 56

5.1 Diffusion by the interstitial mechanism. 62

5.2 Diffusion by the substitutional mechanism. 62

5.3 Interstitial-substitutional diffusion by the dissociative mechanism. 62

5.4 Interstitial-substitutional diffusion by the kick-out mechanism. 62

5.5 Interstitialcy diffusion process. 62

5.6 Sheet resistance of doped product wafers as a function of RTD duration at 975 °C and 1050 °C. 67

5.7 Boron concentration profiles of bare silicon wafers that were doped during proximity RTD at 1050 °C for various durations. 68

5.8 Surface boron concentration and junction depth as a function of RTD duration. 69

5.9 SIMS concentration profile of an oxidised wafer that was doped wafer during proximity RTD at 1050 °C for 5 s. 70

5.10 Dependence of boron diffusion coefficient on boron concentration in Si, evaluated from 16 s proximity RTD into bare Si at 1050 °C, and from a 5 s RTD through a 120 Å oxide on Si at 1050 °C. The dotted curve represents the normal diffusion coefficient. 72

5.11 Junction depth at 10^{18} cm^{-3} versus sheet resistance for various shallow boron junction formation techniques. 74

6.1 Mask Layout for Surround Diodes (SD), Open Diodes type 1 (OD1), and Open Diodes type 2 (OD2). 80

6.2 Forward IV characteristics of surround diodes on wafer D5. 84

6.3 Current v's the total *p*-type area of surround diodes on (a) wafer D5 and (b) wafer D4. 85

6.4 (a) Current crowding in a surround diode, (b) As width x increases, the current crowding increases in the diode. 86

6.5 Current v's total area of OD1 diodes in wafers D4 and D5. 88

6.6 Forward bias IV characteristics for OD2 diodes on wafer D5. 89

6.7 ΔV versus I for OD2 278 μm x 50 μm. The slope of the graph is the series resistance, R_s, of the diode. 91

6.8 (a) Planar diffusion process that forms junction curvature near the edge of the diffusion mask, where r_j is the radius of curvature. (b) Formation of the cylindrical and spherical regions by diffusion through a rectangular mask. 92

6.9 Breakdown voltage versus impurity concentration for one-sided abrupt doping profile with cylindrical and spherical junction geometries, where *rj* is the radius of curvature as indicated in Fig. 6.14 [2]. 93

6.10 Breakdown voltages of OD2 278 μm x 50 μm. 93

6.11 (a) Accumulation of electrons at the oxide-substrate interface reduces depletion width to W_1, $W_1 \ll W$ (b) Avalanche multiplication process injects an electron into the oxide (c) Injected electron causes depletion width to change to W_2, $W_2 > W_1$ (d) If avalanche breakdown is repeated, depletion width at oxide-substrate interface increases an saturates to W. 95

6.12 Breakdown voltages for OD2 278 μm x 50 μm on wafer D5. 97

6.13 Shape of the depletion region around (a) SD 666 x 15 μm and (b) SD 100 x 100 μm. 98

6.14 Influence of field-plate voltage on the shape of the depletion region. (a) *p-n* junction with no field-plate (b) *p-n* junction with a field-plate. 100

7.1 Locations of the maxima of the thermally induced resolved stresses and the corresponding slip line directions on a <100> silicon wafer. 103

7.2 Raman spectra were measured at points a-j across each wafer. 105

7.3 Schematic diagram of White Light Back Reflection Topography. 107

7.4 Schematic diagram of White Beam Topography in Section Topography Mode. 107

7.5 Raman shift, Δω, as a function of process duration for RTO and RTD at 1050 °C. 109

7.6 SXRT section topographs: (a) Bare as-received Si reference wafer; (b) Wafer F: RTO for 166 s at 1050 °C; (c) Wafer A: RTD for 5 s at 1050 °C; (d) Wafer C: RTD for 166 s at 1050 °C. 112

7.7 SXRT back reflection topograph of wafer F. 113

8.1 Conventional LDD spacer process. (a) Selective n- implantation; (b) Selective p- implantation; (c) LTO spacer deposition and RIE, followed by a source/drain screen oxidation. Next, a selective n+ implantation is performed; (d) Selective p+ implantation. 117

8.2 Drain current in conventional and LDD MOSFETs [6]. 119

8.3	Disposable spacer process. (a) Undoped polysilicon is deposited as the spacer material. (b) Using resist to cover PMOS areas, the NMOS is formed and followed by an arsenic n^+ implantation. The resist is stripped and the arsenic profile is annealed. (c) A blanket phosphorus n- implant is performed which uses the non-etched polysilicon spacer material to protect the PMOS areas. (d) Using the second patterning step, the NMOS areas are covered with resist and the PMOS poly spacers are formed, followed by the p+ source/drain implant. (e) The spacers are removed and the p- implant is performed. (f) The resist is stripped and the permanent LTO spacers are formed to define the salicided source/drain/gate regions. Conventional processing follows with contact and aluminium metallisation definition. [1].	121
8.4	The dependence of device lifetime on the LDD n- implant dose. The maximum lifetime point is at 5-6x10^{13} cm^{-2} [10].	123
A.1	Mask set for Surround Diodes.	131
A.2	Mask set for Open Diodes type 1.	132
A.3	Mask set for Open Diodes type 2.	133

List of Tables

1.1 Projected LDD junction depths versus gate lengths [4]. 18

3.1 Sheet resistance of 3 wafers that were doped with a single dopant source wafer. 37

4.1 Rapid thermal treatment to convert SOD gel to BSG. 44

6.1 Ideality factor, n, and I_0 for wafers D5 and D4. 84

7.1 Rapid thermally processed samples. 108

7.2 Measured Raman peak shift, $\Delta\omega$, for oxidised wafers D-F. 109

7.3 Measured Raman peak shift, $\Delta\omega$, for doped wafers A-C. 110

7.4 Measured Raman half widths, Γ, for wafers F, A and C. 113

List of Symbols

A	Cross-sectional area (cm^2)
A_J	Junction area (cm^2)
A_S	Area of the depletion region at the silicon surface (cm^2)
B	Body/substrate
C_s	Surface concentration (cm^{-3})
D	Drain
D_i	Intrinsic diffusion coefficient (cm^2/s)
D_P	Hole diffusion constant (cm^2/s)
E_Y	Horizontal electric field (V/cm)
f_{AI}	Fractional interstitialcy component
g_m	Transconductance (A/V)
G	Gate
I	Current (A)
I_D	Drain current (A)
I_{DST}	Subthreshold current (A)
I_{GEN}	Generation current (A)
I_{REC}	Recombination current (A)
k	Boltzmann's constant (8.62×10^{-5} eV/°K)
L	Channel length
n	Ideality factor
n_i	Intrinsic carrier concentration (cm^{-3})
N_D	Donor impurity concentration (cm^{-3})
q	Charge on an electron (1.602×10^{-19} C)
R_s	Sheet resistance (Ω/sq)
S	Source
S_0	Surface recombination velocity (cm/s)
T	Temperature

V_A	Applied voltage (V)
V_{DB}	Drain-body voltage (V)
V_{DS}	Drain-source voltage (V)
V_{DSSAT}	Pinch-off voltage (V)
V_{GB}	Gate-body voltage (V)
V_{GS}	Gate-source voltage (V)
V_{SB}	Source-body voltage (V)
V_T	Threshold voltage (V)
x_j	Junction Depth
$\Delta\omega$	Raman shift (cm^{-1})
ε	Misfit factor
Γ	Raman half width (cm^{-1})
σ	Stress (Pa)
τ_g	Generation lifetime (s)
τ_0	Effective lifetime of carriers in the depletion region (s)
τ_P	Hole minority carrier lifetime (s)

Acronyms and Abbreviations

APCVD	Atmospheric Pressure Chemical Vapour Deposition
BPSG	Borophosphosilicate Glass
BRT	Back Reflection Topography
BSG	Borosilicate Glass
DIBL	Drain Induced Barrier Lowering
FD	Furnace Diffusion
FTIR	Fourier Transform Infrared Spectroscopy
GILD	Gas Immersion Laser Doping
GOLD	Gate-drain Overlapped LDD
IC	Integrated Circuit
IR	Infrared
ITLDD	Inverse-T Lightly Doped Drain
LDD	Lightly Doped Drain
LOCOS	Local Oxidation of Silicon
LPCVD	Low Pressure Chemical Vapour Deposition
LTO	Low Temperature Oxide
MLDD	Moderate Lightly Doped Drain
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
MRS	Micro Raman Spectroscopy
OD1	Open Diode type 1
OD2	Open Diode type 2
OED	Oxidation Enhanced Diffusion
PSG	Phosphosilicate Glass
RIE	Reactive Ion Etching
RTC	Rapid Thermal Cure
RTD	Rapid Thermal Diffusion
RTO	Rapid Thermal Oxidation

RTP	Rapid Thermal Processor
RVD	Rapid Vapour-phase Diffusion
SD	Surround Diode
SEM	Scanning Electron Microscope
SIA	Semiconductor Industry Association
SIMS	Secondary Ion Mass Spectrometry
SOD	Spin On Dopant
SOG	Spin On Glass
SOI	Silicon On Insulator
ST	Section Topography
SUPREM	Stanford University Process Engineering Model
SXRT	Synchrotron X-ray Topography
TED	Transient Enhanced Diffusion
TEOS	Tetraethylorthosilane
TOPS	Total Overlap with Polysilicon Spacer
ULSI	Ultra Large Scale Integration
VLSI	Very Large Scale Integration

Author's Publications

Refereed Journal Publications

1. **M. Nolan**, T. Perova, R.A. Moore and H.S. Gamble, "Boron diffusion from a spin-on source during rapid thermal processing", *J. Non-Crystalline Solids*, **254**, 89-93, (1999).

2. **M. Nolan**, T. Perova, R.A. Moore, C.J. Moore, K. Berwick and H.S. Gamble, "Micro-Raman study of stress distribution generated in silicon during proximity rapid thermal diffusion", *Materials Science and Engineering B*, 2000 (In press).

3. T.A. Briantseva, Z.N. Lebedeva, D.V. Lioubtchenko, **M. Nolan**, T.S. Perova and R.A. Moore, "Precise chemical analysis development for silicon wafers after rapid thermal processing", *Applied Surface Science*, 2000 (In press).

4. **M. Nolan**, T.S. Perova, R.A. Moore, C.E. Beitia and H.S. Gamble, "Spectroscopic investigations of borosilicate glass and its application as a dopant source", *J. Electrochem. Soc.,* 2000 (In press).

Referred Conference Proceedings

1. **M. Nolan**, T. Perova, R.A. Moore and H.S. Gamble, "Material and surface analysis of p-type junctions in silicon formed by rapid thermal diffusion", *Materials Research Society Proceedings*, 30 November–4 December, Boston, C3, 58, (1998).

2. **M. Nolan**, T. Perova, R.A. Moore and H.S. Gamble, "Boron diffusion from a spin-on source during rapid thermal processing", *Material Research Society Proceeding of the 2nd International Conference on Amorphous and Crystalline Insulating Thin Films,* 12-14 October, Hong Kong, 89, (1998).

3. **M. Nolan**, T. Perova, R.A. Moore, C.J. Moore, K. Berwick, H. Byrne and H.S. Gamble, "Micro-Raman study of stress distribution generated in silicon during proximity rapid thermal diffusion", *Proceedings of the European Material Research Society Spring Meeting,* 1-4 June, Strasbourg, France, p. E-21, (1999).

4. T. Briantseva, Z. Lebedeva, D. Lioubtchenko, **M. Nolan**, T. Perova, R.A. Moore and K. Berwick, "Precise chemical analysis development for Si and GaAs surfaces", *Proceedings of the International Congress on Advanced Materials and Processes,* EUROMAT'99, 27-30 September 1999, Munich, Germany (In press).

5. T. Briantseva, Z. Lebedeva, D. Lioubtchenko, **M. Nolan**, T. Perova, R.A. Moore and H.S. Gamble, "Process-induced modification to the surface of crystalline silicon by precise chemical analysis", *Proceedings of the International Congress on Advanced Materials and Processes,* EUROMAT'99, 27-30 September 1999, Munich, Germany (In press).

6. **M. Nolan**, T.S. Perova, C.E. Beitia, R.A. Moore, J. McGilp, H.S. Gamble, C. Moore and K. Berwick, "Spectroscopic investigations of dopant source borosilicate glass", The Electrochemical Society 197th Meeting, Toronto, Ontario, Canada, May 14-18, 2000 (Accepted).

7. **M. Nolan**, T.S. Perova, C.E. Beitia, R.A. Moore, J. McGilp, H.S. Gamble, C. Moore and K. Berwick, "Spectroscopic investigations of dopant source borosilicate glass and porous silicon for integrated circuits", NATO Advanced Workshop (ARW)/Spring School "Frontiers of Nano-Optoelectronic Systems: Molecular Scale Engineering and Processes" SA (PST.ARW 975678) 5068, Kyiv, Ykraine, 22-26th May, 2000 (Accepted).

8. **M. Nolan**, T.S. Perova, R.A. Moore, H.S. Gamble, D. Lowney, P.J. McNally, T. Tuomi, R. Rantamaki and A.N. Danilewsky, "Micro-Raman and Synchrotron X-Ray Topography of Boron Doped Silicon Wafer using Rapid Thermal Diffusion", The XVIIth International Conference on Raman Spectroscopy, ICORS 2000, Beijing, China, August 20-25, 2000, China (Accepted).

9. **M. Nolan**, T.S. Perova, R.A. Moore and H. S. Gamble, "Characterisation of the borosilicate glass as a dopant source and its application for low-doped drain PMOS devices", The 3rd International conference on Materials for Electronics, 16-17 October 2000, Dublin, Ireland (Oral presentation)(Submitted).

Preface

The IC industry continually strives towards manufacturing higher density integrated circuits, since increasing the density of microchips on the wafer increases the profitability of the process. Increasing the density means using shorter channel lengths in MOS transistors. Reducing the channel lengths did not present a major problem until the 1970s, when channel lengths were reduced to < 2 μm. Beyond this point the short-channel devices began to display characteristics that were not observed in long-channel devices. These deleterious characteristics were termed short-channel effects, and continue to exacerbate engineers to the present day.

Low-doped drain (LDD) extensions in MOS transistors improve short-channel effects. Initially, it was possible to fabricate these shallow junctions using traditional ion implantation. However, ion implantation is limited for ultra-shallow junctions (< 500 Å), primarily because it generates defects in the lattice that enhance dopant diffusion. In particular, producing ultra-shallow boron junctions is a major challenge confronting researchers, since boron tends to channel through silicon resulting in deep junctions. Alternative doping techniques have been widely reported in the literature.

Rapid thermal processing (RTP) is an important technology for ULSI chip manufacturing, and is increasingly replacing conventional furnace processing. RTP is a single-wafer process, and offers the advantage of better wafer-to-wafer uniformity and better control of ambient and peak temperature. The thermal budget is also minimised during RTP since the wafer is rapidly heated to 1000 °C in 10-20 s, typically. RTP is already widely used for annealing implant damage and for forming high quality silicide layers. Rapid thermal diffusion (RTD) is a prospective technique for ultra-shallow junction formation, primarily because anomalous diffusions are avoided due to the minimal thermal budget, and because it does not introduce unwanted defects into the substrate.

This thesis begins by introducing the difficulties that have arisen in MOS transistor device physics and MOS doping technology, in deep submicron device production. RTD of boron from a borosilicate layer is then investigated as a novel technique for fabricating boron junctions. The main aim of this work is to fabricate high quality ultra-shallow boron junctions. *P-n* junction diodes are fabricated and electrically tested to determine the quality of the junctions. The stress that is generated during RTD is also analysed.

Chapter 1

MOS Transistor Device Physics

1.1 Introduction

During the past 50 years, the IC industry has aimed towards making smaller and faster devices. In the 1970's MOSFET devices were fabricated with gate lengths smaller than ~2 μm. Beyond this point the MOSFETs began to exhibit characteristics that had not been observed in larger long-channel devices. These characteristics were termed short-channel effects. Short-channel effects deteriorate as the gate length continues to decrease. To date, MOS devices with gate lengths less than 0.25 μm have been manufactured, and short-channel effects have become a critically important issue. Therefore, an understanding of MOS device physics becomes increasingly important in sub-quarter-micron VLSI and ULSI technology. Device aspects that could once be ignored, such as source-drain doping profiles, are now extremely significant.

Shallow junctions eliminate short-channel effects in MOS devices. The main aim of this work is to fabricate shallow boron junctions for PMOS transistors. Therefore, this chapter tends to put more emphasis on PMOS devices

This chapter begins by discussing MOS capacitors, since an understanding of a MOS capacitor is central to understanding the operation of a MOSFET. Secondly, long-channel MOSFETs are discussed. Thirdly, short-channel MOSFETs and issues of MOSFET device scaling are presented. Finally, lightly-doped-drain (LDD) devices are presented. LDD devices are the most widely used devices for improving the reliability and device performance of short-channel MOSFETs.

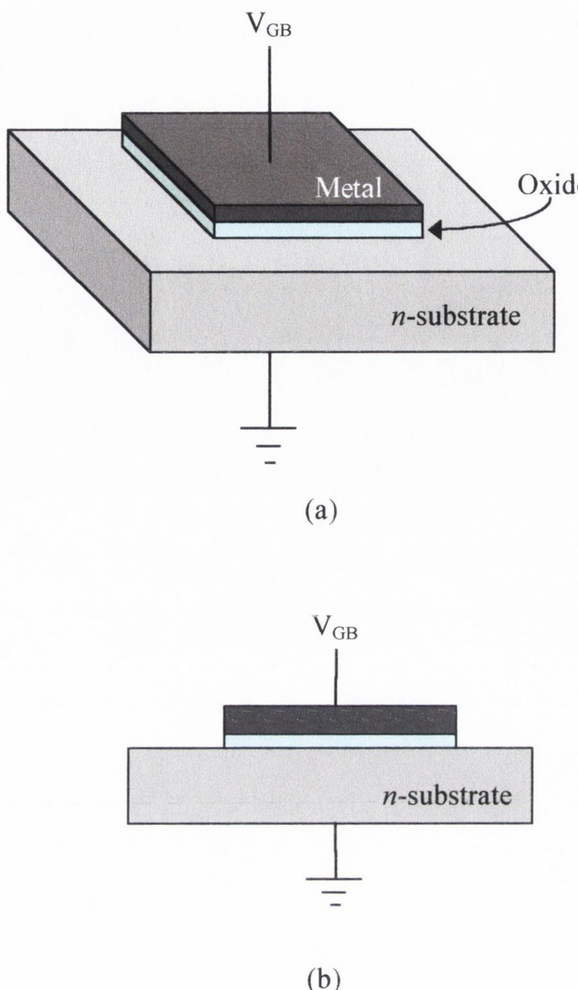


Figure 1.1. (a) Perspective view of a MOS-C, (b) Cross-section of a MOS-C.

1.2 MOS Capacitor

The MOS capacitor (MOS-C) exhibits behaviour that is the foundation of both long-and short-channel MOSFET operation. Such behaviour is easiest to describe in a MOS-C since some of the more complex MOS device characteristics are absent in the MOS-C. Figure 1.1 shows the structure of a MOS-C. It consists of a metal conducting layer (referred to as the gate), an insulating layer (usually silicon dioxide) and a semiconductor substrate, hence the acronym MOS (*metal-oxide-semiconductor*). The semiconductor substrate may be *n*-type or *p*-type. Originally, the gate in the early MOS-C structure was made of aluminium metal, however, polysilicon has replaced aluminium in modern MOS structures, but the original MOS acronym remains unchanged. Figure 1.1 shows

terminals connected to the gate and to the substrate of the MOS-C. An external bias, V_{GB} (gate-body voltage), can be applied across these terminals. Sections 1.2.1-1.2.4 will consider the operation of a *n*-substrate MOS-C, Fig. 1.1.

1.2.1 Accumulation

In the first biasing case a positive bias is applied ($V_{GB}>0$). The positive bias causes electrons (the majority carriers in *n*-type Si) to be attracted to the Si surface adjacent to the oxide. More electrons accumulate near the Si surface than if $V_{GB}=0$. The effect of attracting additional majority carriers to the semiconductor surface is called *accumulation*. The conductivity of the region near the Si surface is increased by the presence of the extra electrons.

1.2.2 Depletion

In the second bias condition, a small negative voltage is applied to the gate of the *n*-type MOS-C ($V_{GB}<0$). This causes electrons in the semiconductor to be repelled from the surface, creating a depletion region, of width d, consisting of positively charged donor ions. This mode of operation of called the *depletion* mode.

1.2.3 Inversion

When the negative voltage that is applied to the gate is increased, minority holes are attracted to the Si surface and the Si surface is changed from n-type to p-type. This layer of holes at the Si surface is called the *inversion layer*.

1.2.4 Strong Inversion

If V_{GB} is made even more negative, the hole density per unit volume at the surface will be equal to the electron density in the bulk. This is known as the point of *onset of strong inversion*. When strong inversion occurs, the depletion region reaches a maximum width of d_{max}. The voltage applied between the gate and the substrate at the onset of strong inversion is called the threshold voltage, V_T. For further increases in negative voltage, the hole concentration in the inversion layer exceeds the bulk equilibrium electron concentration, but there is no further change in the depletion width. This mode of operation is referred to as *strong inversion*.

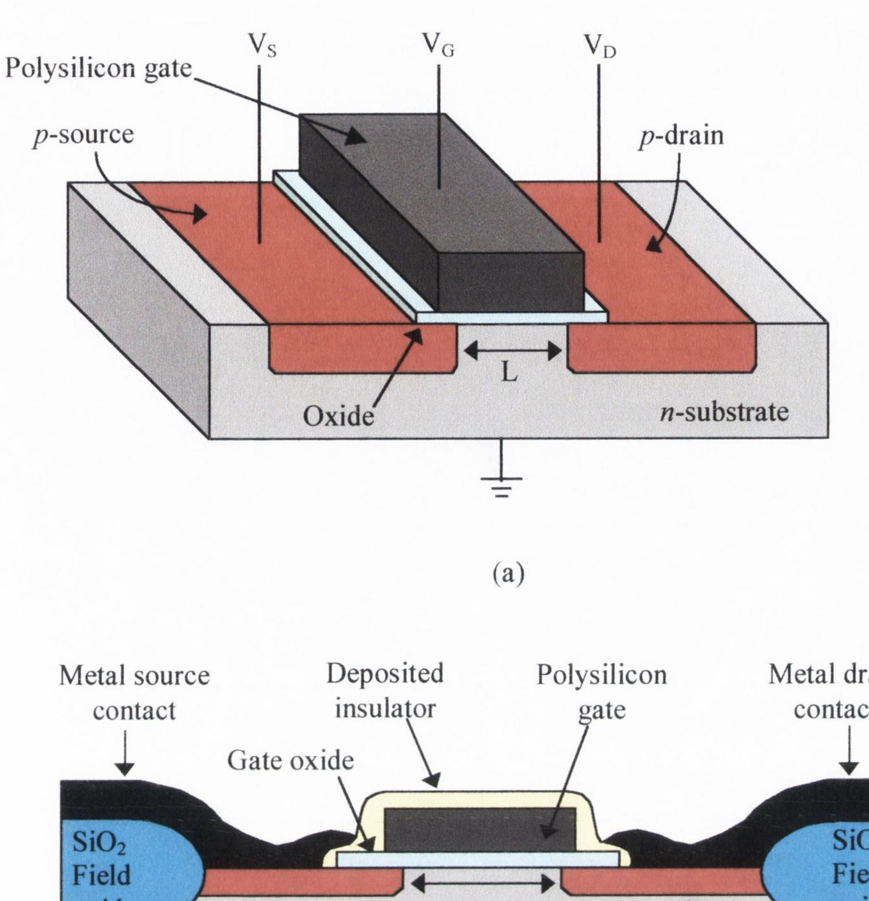


Figure 1.2. (a) Perspective view of a MOSFET, (b) Cross-section of a MOSFET.

1.3 Long-Channel MOSFET

Figure 1.2 shows the cross-section of a *p*-channel MOSFET. The important difference between the MOS-C and the MOSFET is that in the latter a *pn* junction exists at each side of the MOS-C channel region. The MOSFET has four terminals, referred to as the source S, gate G, drain D, and the substrate/body B. The source contact will be used as a reference throughout this section. When no voltage is applied to the gate, the source-to-drain electrodes correspond to two back-to-back pn junctions. The only current that can

flow from source to drain is the reverse leakage current. When a sufficiently large negative voltage is applied to the gate (V_{GB}>V_T), the central MOS structure is inverted forming a surface inversion channel that connects the source and drain regions together. A drain current, I_D, can then flow if a voltage, V_{DS}, is applied between the source and drain terminals.

Now, the I_D variation with V_{DS} will be considered. When $V_{DS} = 0$ V, $I_D = 0$, even if $V_{GS} > V_T$. When V_{DS} is increased slightly I_D starts to flow. When V_{DS} is small (i.e. a few tenths of a volt or less), the surface channel behaves like a simple resistor. Therefore, an I_D proportional to V_{DS} flows into the drain. A plot of I_D against V_{DS}, Figure 1.3a, is used to demonstrate the I_D-V_{DS} behaviour; the response to small values of V_{DS} is shown as the straight line from the point 0 to the point A.

When V_{DS} is increased above a few tenths of a volt, the MOSFET exhibits a new phase of behaviour. Current flowing in the channel gives rise to a significant voltage drop in the channel. Specifically, the channel voltage is V_{SB} at the source end and V_{DB} at the drain end. Since, $V_{SB} = 0$ V and $V_{DB} < 0$ V, the channel voltage equals 0 V at the source end, and increases to $V_{DB} < 0$ V at the channel end. As shown in Figure 1.3c, the voltage increase across the channel causes the depletion region under the gate to widen towards the drain end of the channel. As the depletion width increases, more positively charged donor ions are uncovered. Since V_{GB} is fixed, the total charge in the semiconductor must remain unchanged to maintain balance with the unchanged value of the charge on the gate. Therefore, the mobile carrier concentration in the inversion layer decreases as the width of the depletion region increases. The smaller number of mobile carriers causes the channel conductance to decrease, which in turn is manifested as a smaller slope in the I_D-V_{DS} characteristic as V_{DS} is increased. As the drain voltage is increased further, the depletion region continues to widen and the slope of the I-V characteristic continues to decrease.

Eventually, if V_{DS} is increased sufficiently, the width of the inversion layer is reduced to zero in the vicinity of the drain, Fig 1.3d. The reduction of the inversion layer width to zero due to an increase in V_{DS} is called pinch-off. The pinch-off point on the I_D-V_{DS} characteristic is shown at point B in Figure 1.3a. As the drain voltage increases beyond

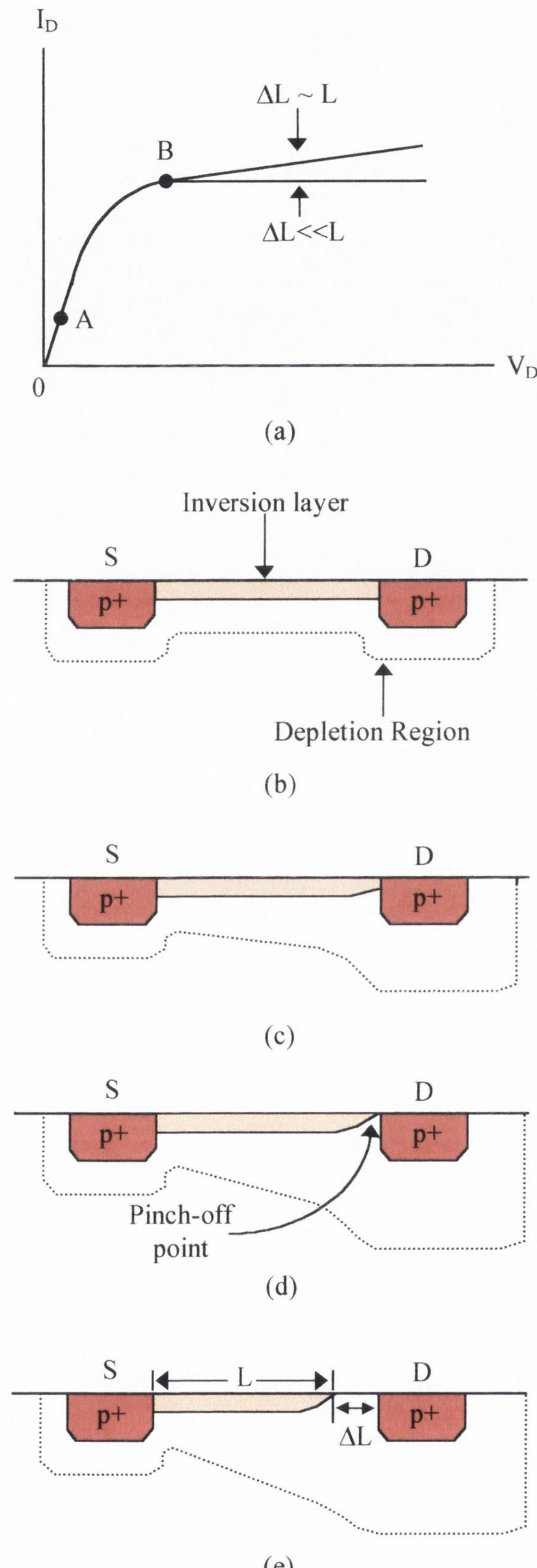


Figure 1.3. (a) I_D-V_{DS} characteristic of a MOSFET being operated at a fixed $V_{GS}>V_T$; (b) $V_{DS}=0$ V; (c) inversion layer narrowing under moderate V_{DS} biasing; (d) pinch-off, $V_{DS}=V_{DSSAT}$; and (e) post pinch-off, $V_{DS}>V_{DSSAT}$.

the pinch-off voltage V_{DSSAT}, the pinch-off section of the channel widens from a point into a segment, ΔL in length, Figure 1.3e. Most of the voltage drop in excess of V_{DSSAT} is dropped across this pinched-off section ΔL. I_D remains approximately constant (saturated) for drain voltages greater than V_{DSSAT} so long as $\Delta L \ll L$. If ΔL becomes comparable to L (which may occur in short-channel MOSFETs), the same voltage drop appears across a shorter channel, and the post-pinch-off current I_{DSAT} in such devices will increase as V_{DS} is increased above V_{DSSAT}, Figure 1.3a.

For each value of $V_{GS} > V_T$, a characteristic of the form shown in Figure 1.3a will be observed. Since the conductance of the channel will increase as V_{GS} is increased, the initial slope of I_D will also become steeper with increasing V_{GS}. Also, since the inversion layer contains more holes, a larger value of V_{DS} is needed to produce a pinch-off, and V_{DSSAT} increases for increasing V_{GS}. Based on these phenomena, Figure 1.4 shows *I-V* characteristics of an ideal long-channel MOSFET.

1.3.1 Subthreshold Currents in Long-Channel MOSFETs

In section 1.2.4, the onset of strong inversion was defined as the point when the hole density per unit volume at the surface is equal to the electron density in the bulk. The voltage applied between the gate and the substrate at the onset of strong inversion was defined as V_T. The mode of operation when the hole density at the surface is smaller than the electron density in the substrate is referred to as weak inversion. Since mobile carriers can exist in the channel region below threshold (i.e., in weak inversion), a small

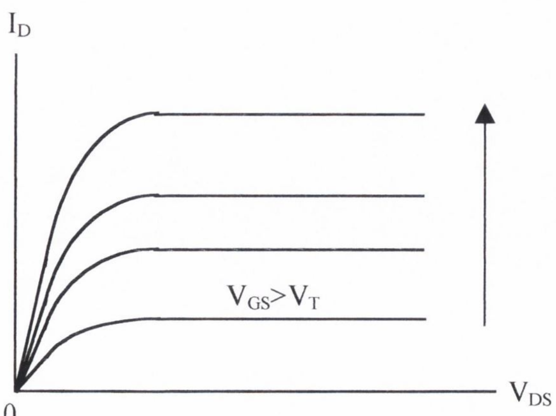


Figure 1.4. I_D-V_{DS} characteristics for a long-channel MOSFET.

drain current can flow in the channel if a voltage is applied to the drain. This small current is called the subthreshold current, I_{DST}. In most applications I_{DST} is far too small to be useful as a drive current. However, it can represent an unwanted leakage current. The common range of V_T in submicron CMOS ICs is 0.6-0.8 V. Therefore, when $V_{GS} = 0$ V the MOSFETs in the circuits may be close to weak inversion. Consequently, it is critical to be able to establish and maintain a uniform and stable V_T in MOSFET devices. If the magnitude of V_T drops even slightly below its designated value, the device may exhibit excessive subthreshold leakage current when $V_{GS} = 0$ V. Factors that may adjust the threshold voltage are discussed in Section 1.4.

1.4 Short-Channel MOSFET

The continuous drive by the IC industry towards higher device densities and faster circuits increases the importance of MOSFET device scaling. As mentioned in the Introduction, in the 1970s MOS transistors with gate lengths less than 2 μm began to exhibit characteristics that were not observed in larger long-channel devices. The control of these short-channel effects is critical in todays sub-quarter-micron technology. This section will discuss these short-channel effects and the resulting device degradation.

1.4.1 Hot-Carrier Effects

If MOSFET device dimensions are reduced and the voltage supply remains constant, the lateral drain applied electric field in the channel increases. A large drain applied electric field exists in submicron short-channel MOSFETs. Mobile carriers injected from the source gain kinetic energy from this and are accelerated towards the drain. Some of the carriers can gain sufficient energy to surmount the Si-SiO_2 barrier and are injected into the gate oxide, where they may be trapped. Those carriers trapped in the gate oxide are detrimental for the device since they may lead to a V_T shift and deterioration in transconductance, g_m ($g_m \equiv \delta I_{DS}/\delta V_{GS}$). As mentioned in Section 1.3.1, if the magnitude of V_T drops even slightly below its designated value, the device may exhibit excessive subthreshold leakage current when $V_{GS} = 0$ V. Other injected carriers flow towards the gate electrode contributing to the gate current. Those carriers flowing to the gate electrode may generate trapping sites in the oxide and at the interface. Figure 1.5 shows

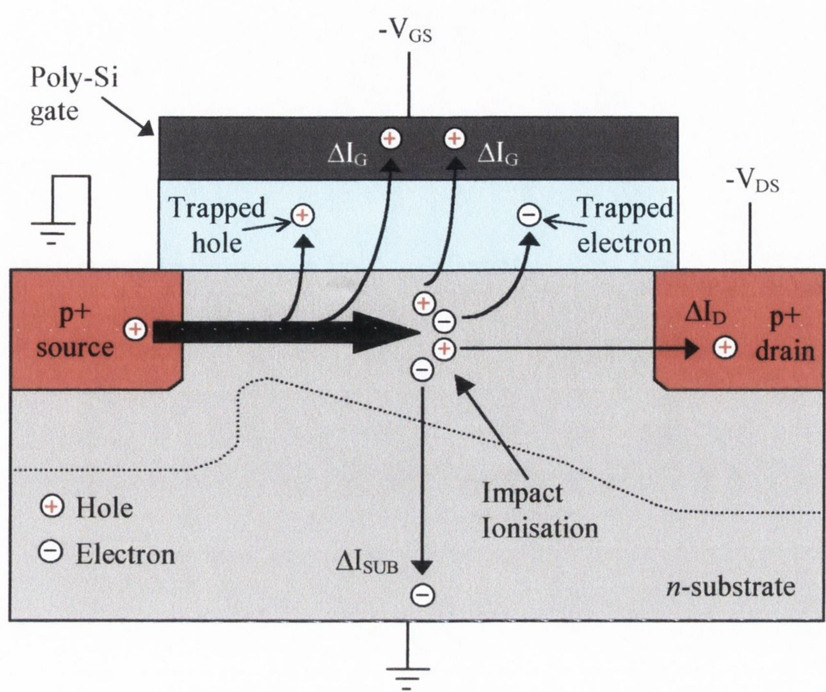


Figure 1.5. Cross-section of a PMOS transistor showing hot-carrier processes.

these effects schematically. Hole injection into the oxide is far less probable than electron injection because the potential energy barrier for holes at the Si-SiO_2 interface is higher.

Those carriers not injected into the gate oxide can continue accelerating towards the drain, causing avalanche multiplication as a result of impact ionisation, Fig. 1.5. Thus, a large number of electron-hole pairs are generated. Some of the electrons and holes are injected into the gate oxide and become trapped there. Some flow to the gate electrode appearing as a gate current. The majority of the generated carriers flow to the drain, contributing to drain current. Those carriers flowing to the substrate become a substrate current. Excessive substrate current may induce latchup in CMOS circuits. Secondly, as some of the carriers are accelerated leaving the drain-substrate depletion region, they may gain enough energy to cause secondary impact ionisation far from the drain region. Some of the electrons generated in this manner may then travel to other nodes on the chip to be collected. This may lead to the reduction of storage time of dynamic circuit nodes in DRAMs.

Hot-carrier effects are observed in NMOS transistors with gate lengths < 2 μm, however, they are only observed in PMOS transistors with gate lengths < 1 μm. The reasons for this are that impact ionisation due to holes is 1-2 orders of magnitude lower than that due to electrons at a given electric field.

There is one particular hot-carrier effect that only occurs in PMOS transistors [1]. Hot electrons that are trapped in the gate oxide cause a build-up of negative charge in the oxide near the drain. This charge eventually inverts the *n*-type Si surface near the drain, resulting in an extension of the p^+ drain region, Fig. 1.6. As a result, the effective channel length, L, shrinks. A reduction in L at low gate voltages is undesirable for short-channel PMOS devices since it gives rise to a poor subthreshold characteristic and a reduction in the source-drain punchthrough voltage (Section 1.4.3). A second effect arising from the build-up of negative charge in the oxide of the PMOS is that the magnitude of V_T is reduced.

In summary, hot-carrier effects cause serious device degradation. Oxide trapped charges cause detrimental shifts in the threshold voltage and deterioration in transconductance. Carriers that flow towards the gate electrode may generate oxide and interface trapping sites as they move through the gate oxide. Carriers that contribute to the substrate current may induce latchup in CMOS circuits. These carriers may also generate electron-hole pairs by secondary impact ionisation far from the drain. The generated electrons may then escape the drain field and cause device deterioration in other parts of the chip.

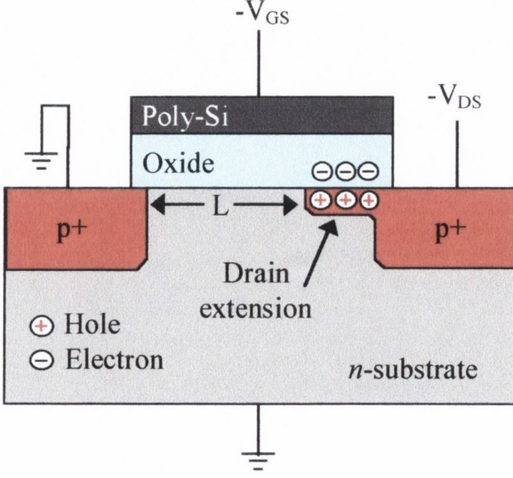
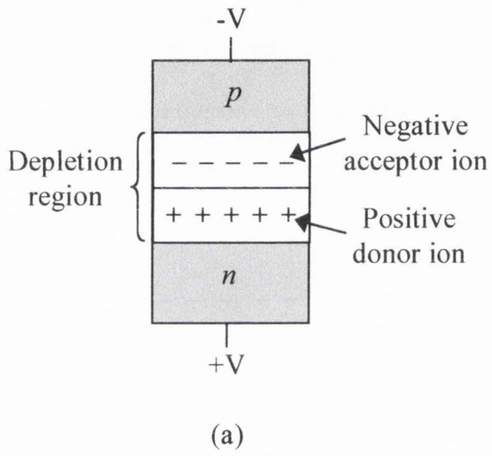


Figure 1.6. Hot electron injection into the gate oxide in a PMOS transistor.

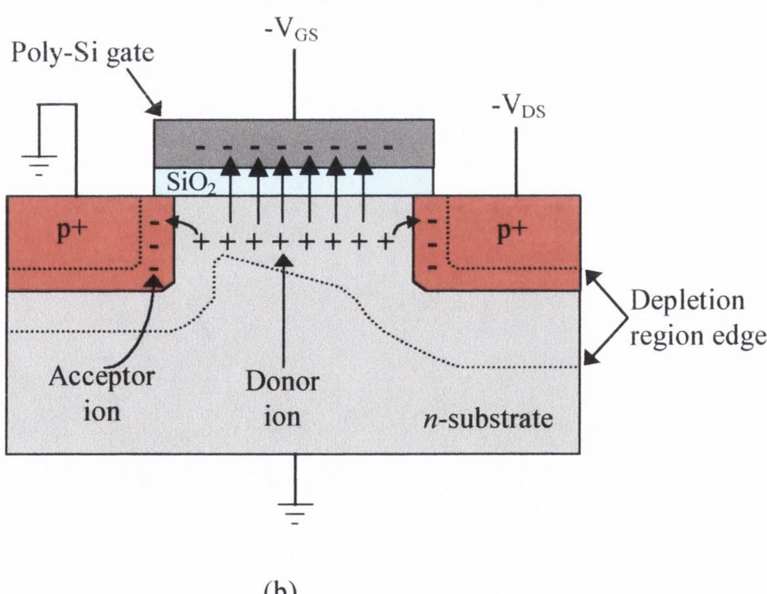
Lightly doped drain (LDD) structures have been used to increase the hot-carrier reliability in submicron MOSFETs. LDD devices will be presented in Section 1.5.

1.4.2 Drain-induced-barrier-lowering (DIBL)

In an ideal MOSFET device the electric field lines emanating from the charge under the gate would terminate on the gate charge. However, in real devices some terminate on space charge in the source and drain depletion regions. Figure 1.7 shows this effect schematically. A *pn* junction depletion region consists of positively charged donor ions



(a)



(b)

Figure 1.7. (a) Reverse biased *pn* junction; (b) Cross section of a PMOS transistor showing electric field lines from the channel charge terminating on acceptor ions in the source and drain depletion regions.

and negatively charged acceptor ions. In a PMOS transistor the electric field lines emanating from the positive donor ions in the channel depletion region can terminate on the negatively charged acceptor ions on the *p*-type side of the source and drain depletion regions instead of on the gate charge, Fig. 1.7. Consequently, less gate charge is required to cause inversion, thus reducing the magnitude of V_T. This effect is called drain-induced-barrier-lowering (DIBL). The fraction of the charge induced by the source and drain becomes significant as the channel length is of the order of the junction depletion widths. Therefore, DIBL critically degrades submicron MOS devices.

LDD devices suppress DIBL in submicron MOSFETs, and will be discussed in Section 1.5.

1.4.3 Punchthrough

In Section 1.3.1, subthreshold currents, I_{DST}, in long-channel MOSFETs were described. In weak inversion ($V_{GS}<V_T$) a subthreshold current may flow at the surface of the channel region. If V_{GS} is fixed and V_{DS} is increased, larger values of I_{DST} are observed in short-channel MOSFETs than in long-channel devices. This larger I_{DST} is partly due to an increase in DIBL. However, another short-channel effect called punchthrough can also give rise to the increase in I_{DST}.

When V_{DS} in increased in MOSFETs with $L<2$ μm the source and drain depletion region edges get closer together and eventually merge. Since the depletion regions of a *pn* junction widen as the reverse bias is increased, all MOSFETs would eventually enter punchthrough if a sufficiently high V_{DS} was applied. However, in MOSFETs with $L>2$ μm, junction breakdown usually occurs before the punchthrough voltage is reached. Consequently, punchthrough is not a limiting factor in long-channel devices. In shorter channel devices, $L<2$ μm, punchthrough does represent a serious limitation.

The next section will discuss changes that have been made to submicron MOS devices in an attempt to suppress punchthrough and the other short-channel effects that have been presented in the previous sections.

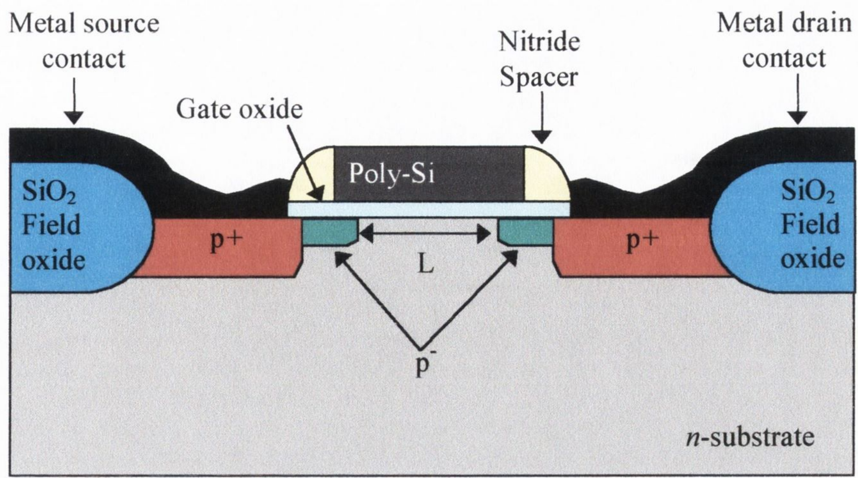


Figure 1.8. Cross-section of a LDD PMOS transistor.

1.5 Lightly Doped Drain (LDD) Devices

The problems that have arisen in submicron MOSFETs have put enormous pressure on process and design engineers to make modifications to the basic MOSFET structure that will improve the reliability and performance of these smaller devices. Use of a lightly doped drain (LDD) is one of the modifications that have been investigated. The remainder of this section will describe the LDD MOSFET structure and explain its impact on device performance. LDD device fabrication will be discussed later in Chapter 8.

Figure 1.8 shows a cross section of a conventional LDD PMOS transistor. A shallow lightly doped region has been added to the source and drain. LDD structures are the most widely used device structures for increasing hot-carrier reliability in submicron MOSFETs. Figure 1.9 shows the electric field profile at the drain of a MOSFET, both with and without a LDD structure [2]. In this case, the LDD provides a 30-40% reduction in the drain-applied electric field. A large proportion of the drain-source voltage drop is across the LDD. As a result, a mobile carrier that is injected from the source into the channel will not gain sufficient energy from the lower drain applied electric field to become a hot-carrier. Therefore, the hot-carrier reliability of the device is increased. Figure 1.10 shows the hot-carrier lifetime improvement gained by using a *p*-channel LDD [1]. It compares conventional 0.95 μm PMOS devices with 0.8 μm LDD

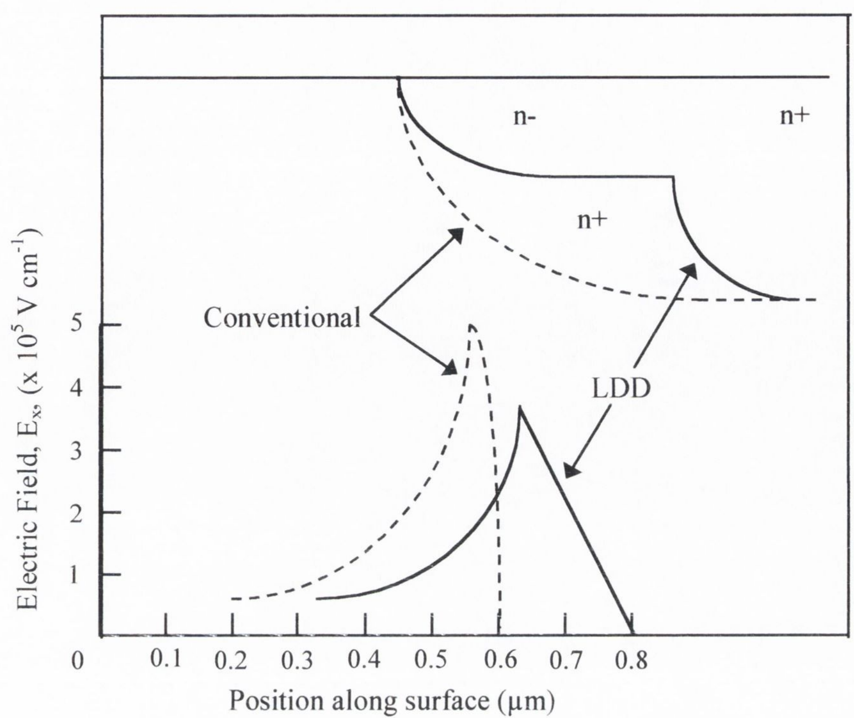


Figure 1.9. Magnitude of the electric field at the Si-SiO_2 interface as a function of distance: L = 1.2 μm, V_{DS} = 8.5 V, V_{GS} = V_T [2].

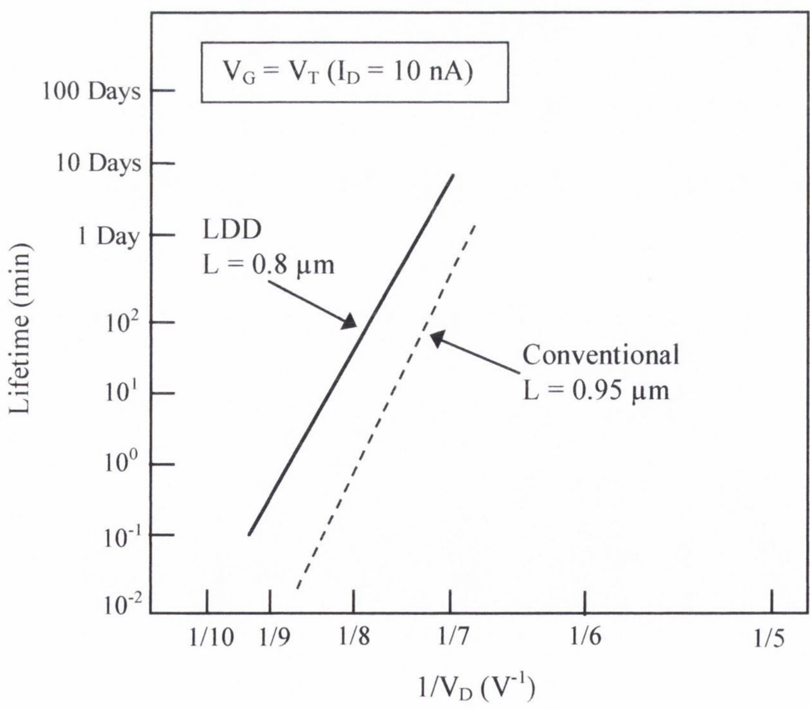


Figure 1.10. Device lifetime versus reciprocal stressing drain voltage for a conventional PMOSFET and a LDD PMOSFET [1].

PMOSFETs. The LDD devices have superior reliability despite the shorter gate lengths, and an extrapolation of the data to $V_{DS} = 5$ V implies that lifetimes in excess of 10 years are possible. Unfortunately, while LDDs do improve the hot-carrier reliability in submicron devices, they also cause an unwanted reduction in the drain current. LDD regions increase the total resistance along the channel, thus, reducing I_D. Several improvements have been made to the conventional LDD device shown in Figure 1.8, including increasing the LDD doping concentration. These improvements will be discussed in Chapter 8.

DIBL critically degrades submicron devices by lowering the threshold voltage, Section 1.4.2. DIBL can be reduced by fabricating shallow drain junctions in the MOS transistor. However, reducing the drain junction depth also degrades the device performance. The reduced junction depth increases the sheet resistance and the contact resistance of the drain. The shallower junction is also more susceptible to aluminium junction spiking. Another way to achieve a shallow junction in a PMOS device is to use a LDD. The junction depth of the shallow LDD reduces unwanted DIBL and also avoids the issues that arise when the entire p^+ drain junction depth is decreased.

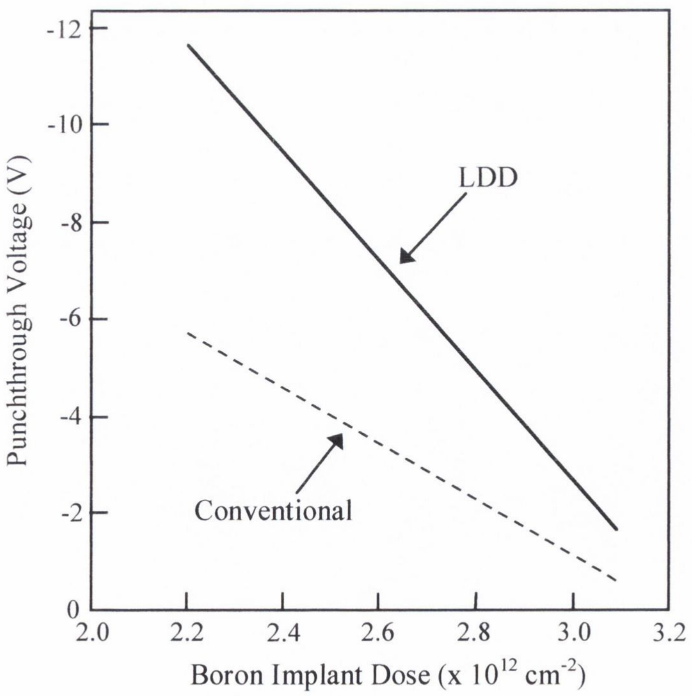


Figure 1.11. Punchthrough voltage as a function of boron implant dose for a 0.4 μm PMOSFET with and without an LDD structure [3].

Table 1.1 Projected LDD junction depths versus gate lengths [4].

Year of first product shipment	1997	1999	2001	2003	2006	2009	2012
Gate Length (μm)	0.25	0.18	0.15	0.13	0.1	0.07	0.05
LDD junction depth (Å)	500-1000	360-720	300-600	260-520	200-400	150-300	100-200
LDD Conc. (cm^{-3})	$1x10^{18}$	$1x10^{19}$	$1x10^{19}$	$1x10^{19}$	$1x10^{20}$	$1x10^{20}$	$1x10^{20}$

Solutions exist *Solutions being persued* *No known solution*

Reducing the source/drain junction depths also suppresses punchthrough in MOS transistors. Shallow junctions reduce the lateral spread of dopants under the gate. Therefore, the lateral spread of the source/drain depletion regions is reduced, making MOSFETs less susceptible to punchthrough. However, the entire length of the source/drain junctions is not made shallow, to avoid junction series resistance and junction leakage problems as mentioned above. So, once again, a LDD structure is a more suitable way to achieve shallow junctions. Figure 1.11 indicates that the punchthrough voltage is increased by 2-3 V when an LDD is incorporated into a PMOS device [3].

In summary, LDDs improve the reliability and device performance of submicron MOSFET devices. Table 1.1 shows the projected junction depths for LDDs for the next few generations of technology [4].

1.6 Conclusions

In submicron MOS technology, devices exhibit characteristics that were not observed in larger long-channel devices. These characteristics are termed short-channel effects. Hot-carrier effects, drain-induced-barrier-lowering (DIBL), and punchthrough are some of the

short-channel effects that have been observed in MOS transistors. These effects degrade device performance and reduce reliability.

Hot-carrier effects cause a detrimental shift in threshold voltage. If the magnitude of V_T drops even slightly below its designated value, the device may exhibit excessive subthreshold leakage current when $V_{GS} = 0$ V. Hot-carrier effects also deteriorate transconductance and change the substrate, drain and gate currents. Excessive hot-carrier generated substrate current may induce latchup in CMOS circuits.

The second short-channel effect that was discussed was DIBL. DIBL also causes an unwanted harmful shift in threshold voltage. Finally, as the gate length is reduced, MOS devices become more susceptible to punchthrough. Punchthrough increases the subthreshold leakage current in short-channel MOSFETs.

Lightly doped drain (LDD) devices have been used to improve device performance and reliability in submicron MOS transistors. They eliminate hot-carrier effects by reducing the drain-applied electric field across the MOS channel. The shallow LDD junction depth reduces DIBL. LDDs also reduce the lateral spread of the source/drain depletion regions, making MOSFETs less susceptible to punchthrough. The disadvantage of introducing LDD into the conventional MOSFET structure is that the LDD causes an unwanted reduction in the drain current. Improvements that have been made to the conventional LDD MOSFET will be discussed later in Chapter 8. Fabrication of *p*-type LDDs is extremely challenging due to the high diffusivity of boron in silicon. This doping technology issue will be presented in the following chapter.

References

[1] M. Koyanagi, A.G. Lewis, R.A. Martin, T.-Y. Huang and J.Y Chen, *IEEE Trans. Electron. Dev.*, **ED-34**, 839, (1987).

[2] S. Ogura, P.J. Tsang, W.W. Walker, D.L. Critchlow and J.F. Shepard, *IEEE Trans. Electron. Dev.*, **27**, 1359, (1980).

[3] A. Schmitz and J.Y. Chen, *IEEE Trans. Electron. Dev.*, **ED-33**, 148, (1986).

[4] *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Association, (1997), p. 74.

Chapter 2
MOSFET Doping Technology

2.1 Introduction

Ion implantation has been the method of choice for doping semiconductor devices for more than 20 years, due to its unquestionable advantage of precise concentration control and reproducibility. However, for ultra-shallow junctions it has limitations related to defect formation in the target crystal, which affect both junction depth and electrical properties of devices. In particular, shallow *p*-type junction formation through boron implantation creates difficulties. In an attempt to avoid these problems, alternative doping techniques that produce shallow defect-free junctions have been investigated. This chapter discusses the difficulties that have arisen when using ion implantation to fabricate shallow junctions. Rapid thermal doping techniques are also discussed as an alternative to ion implantation.

2.2 Ion Implantation

Ion implantation is a process by which energetic impurity atoms can be introduced into a semiconductor substrate. As an energetic ion penetrates a solid target material, it transfers energy by Coulombic interaction with the electrons in the target material (*Electronic Stopping*) and by collisions with the target nuclei (*Nuclear Stopping*), and eventually come to rest. In electronic stopping, the energy transferred to the electrons can lead to excitation of the electrons to higher energy levels, or to ionisation of the electrons. The energy loss due to these electronic interactions gradually slows the ion, eventually bringing it to a stop. In nuclear stopping, the transfer of energy to the atomic nuclei causes a deflection of the projectile ion and also dislodges the target nuclei from

their original sites. Consequently, nuclear stopping results in physical damage to the semiconductor, which may take the form of point or line defects. Often the semiconductor can become amorphous and/or semi-insulating as a result of this process.

After ion implantation an anneal treatment is necessary both to activate the implanted species as well as to restore the implantation damage. For shallow junction implants, rapid thermal processing (RTP) is often used to minimise unwanted diffusion during the thermal anneal. Transient enhanced diffusion (TED) of the dopant, resulting from implantation damage, occurs during the activation anneal and has been reported extensively. TED due to several types of implant damage has been observed. An ultrafast diffusion pulse with a low activation energy has been identified for low damage levels (Si implantation with a dose $< 10^{12}$ cm^{-2} and the tail end of high dose implants) and low annealing temperatures (< 600 °C) [1]. At higher temperatures (670-815 °C) and higher implant doses ($5x10^{12}$ - $1x10^{14}$ cm^{-2}), {113} interstitial clusters have been observed [2]. These {113} defects dissolve during the anneal releasing interstitials. These interstitials are responsible for TED of dopants, such as boron, which diffuse by an interstitial mechanism in silicon [3-5] (the boron interstitial mechanism of diffusion in Si will be discussed in detail in Chapter 5). At even higher implantation doses ($> 1x10^{14}$ cm^{-2}) end-of-range loops are generated [6]. Dislocation loops are much more stable than {113} defects giving rise to slow annealing and an enhancement of the interstitial concentration.

Ultra-low energy (<500 eV) B implants, which cause less physical damage to the Si substrate, have been studied for shallow junction fabrication [7, 8]. However, from a manufacturing point of view, ultra-low energy implants are not practical because commercially available ion implanters were not designed for such low energy implantations.

In single crystal lattices, there are some crystal directions (known as channels) along which the ions do not encounter any target nuclei, and will be channeled, or steered along such open channels of the lattice. As the implanted ions travel down channels, the slowing down is accomplished mainly by electronic stopping, and the ions can penetrate

the lattice several times more deeply than in amorphous targets. Channeling effects are difficult to control and lead to anomalous diffusion tails in the implanted material.

Channeling effects are particularly difficult to control during B implantation. Boron is a very light ion (B^{11}), and when implanted at room temperature does not amorphise the silicon substrate as a heavier ion, like phosphorous (P^{31}), would. Thus long channeling tails are present in the silicon. The crystalline-defect damage caused by boron must be annealed at temperatures > 900 °C [9], and this leads to significant unwanted diffusion of the implanted ions.

Several attempts have been made to eliminate channeling effects. The most widely adopted technique to minimise channeling has been to tilt the wafer surface relative to the incident beam direction. Tilting the wafer reduces the channeling but it does not eliminate the effect completely. When atoms enter the lattice in an apparently random direction, they can be diverted into one of the open channels after entering the lattice. Also, even when the wafer is tilted the Si atoms may still be aligned in a highly symmetrical array of planes, and *planar channeling* [10] can still produce channeling effects.

Other methods that have been investigated to reduce channeling effects include: i) pre-amorphising the lattice by a prior implantation [11-13]; ii) implanting through a surface oxide, to randomise the directions of the ions as they enter the lattice [14, 15]; and iii) using heavy ions for the implantation, for instance, using BF_2 instead of B [16, 17].

Ge^+ and Si^+ preamorphisation have been used to suppress B channeling. However, for low energy implantation, the junction depth is the same with and without preamorphisation [13]. Implantation through a screening oxide layer, proved not only to be ineffective for very shallow junctions but also resulted in boron profile broadening [18].

The molecular species BF_2 has been investigated as an alternative to B implantation. When the BF_2 penetrates the Si, it collides with target nuclei and the BF_2 ion dissociates, releasing a low energy B atom that will penetrate the substrate less than a higher energy

ion. Since BF_2 is a heavy ion (~5 times the mass of a B atom), collision with Si atoms can create an amorphous layer. This amorphous layer reduces channeling effects and results in shallower junction depths. Unfortunately, BF_2 has a very significant disadvantage that prevents it from being used in shallow junction formation; the incorporated F^+ will enhance boron penetration into the gate oxides [19-21]. Fluorine-induced increase in sheet resistance was reported by Sato *et al.* [19], and can be attributed to the increase in the effective diffusion coefficient of B in SiO_2, which enhances boron migration from the Si to the gate oxide. Therefore, a single B implant is essential for *p*-type shallow junction formation in sub-micron devices.

Reducing channeling effects during the B implant and minimising the enhanced diffusion during annealing are currently major challenges confronting researchers in this field. In an attempt to avoid these problems, alternative doping techniques, such as rapid thermal diffusion (RTD), rapid vapour-phase doping (RVD) [22] and gas-immersion laser doping (GILD) [23, 24], have been investigated. The remainder of this chapter will discuss RTD as a doping technique for submicron technology.

2.3 Rapid Thermal Diffusion

Rapid thermal diffusion (RTD) is a prospective technique for ultra-shallow junction formation. RTD offers several attractive features for shallow junction formation. Firstly, RTD is a single-wafer process, and offers the advantage of better wafer-to-wafer uniformity and better control of ambient and peak temperatures. Secondly, the thermal budget is minimised, therefore avoiding unwanted anomalous diffusion. Finally, RTD does not introduce defects into the substrate during the diffusion process.

Rapid heating and rapid cooling are the characteristics in rapid thermal processing. Consequently, one of the main challenges in RTD lies in minimising the development of thermal stress in the wafer. Stress generation during RTD will be discussed later in Chapter 7.

Rapid thermal diffusion using various dopant sources such as planar dopant sources, doped polysilicon, doped silicides and spin-on dopants (SOD) have been investigated. Difficulties have arisen in several of these IC doping processes.

Planar dopant sources have been successfully used in furnace diffusion processes, and they were expected to be cheap and simple solutions for RTD processes even though they required lengthy thermal preparation prior to diffusion. BN disks are used in the case of boron diffusion. These disks are oxidised to form a layer of boron trioxide (B_2O_3) on the surface. The B_2O_3 then acts as a B dopant source. However the release of B_2O_3 from the disk is not efficient during RTD due to the short process duration, thus resulting in poor process reproducibility. Hydrogen injection improves the B_2O_3 transfer, but results in very thick borosilicate glass (BSG) on the wafer that is difficult to remove [25].

Doped polysilicon sources are fabricated by implanting polysilicon layers that are deposited on the silicon substrate. The implantation energy is selected to confine the dopant to the polysilicon layer and avoid implantation damage in the substrate. In conventional furnace diffusion processes a doped polysilicon would be considered a constant dopant source. However, in RTD due to short process duration, many phenomena such as grain boundary diffusion, epitaxial recrystallisation, dopant pile-up at the poly/mono crystalline interface and interface diffusion, affect the diffusion efficiency [26].

Doped refractory metal silicides such as $TiSi_2$ [27] and $CoSi_2$ [27-29] have been used as dopant source in RTD. As in the case of doped polysilicon, the dopant is implanted into the metal silicide layers, and an implantation energy is chosen to confine the dopant to the silicide. Junction leakage due to metal knock-on, formation of metal-dopant compounds and interface segregation are some of the problems that have been encountered using metal silicides as a dopant source in RTD.

Spin-on dopant has also been investigated as a dopant source for RTD. The following chapter will deal with spin-on dopant sources specifically.

2.4 Conclusions

The IC industries continuous drive towards smaller and faster devices puts critical demands on vertical scaling in sub-micron MOSFET devices. At present, ULSI and VLSI technologies depend on ion implantation for doping. However, there are limitations on this technique in the formation of shallow boron p-type junctions. Boron, in particular, being a light ion requires low energy for implantation and the resultant damage is difficult to anneal. At room temperature, B does not have sufficient energy to amorphise the Si substrate so long channeling tails are present. Preamorphisation still results in defects, and does not reduce the junction depths for low implantation energies. The high diffusivity of B in Si and the enhanced diffusion of the channeling tails result in deep junctions in when the B is thermally annealed. All of these difficulties make control of shallow junctions using ion implantation very difficult.

In an attempt to avoid the problems associated with ion implantation, alternative doping techniques have been investigated. Rapid thermal diffusion is a prospective technique for fabricating shallow boron junctions. Employing this technique minimises the thermal budget and offers better control of junction depth. Dopant sources, including planar dopant sources, doped polysilicon and doped silicides have been discussed. However, dopant pile-up at the poly/mono crystalline interface, formation of metal-dopant compounds and junction leakage due to metal knock-on are some of the problems that have been encountered using these various dopant sources.

Spin-on dopant is an alternative dopant source for RTD. The following chapter will discuss the suitability of SOD as a dopant source in RTD for the fabrication of shallow *p*-type junctions.

References

[1] N.E.B. Cowern, G.F.A. van de Walle, P.C. Zalm and D.W.E. Vandenhoult, *Appl. Phys. Lett.*, **65** 2305, (1994).

[2] A. Agarwal, H.-J. Gossman, D.J. Eaglesham, L. Pelaz, S.B. Herner, D.C. Jacobson, T.E. Haynes and R. Simonton, *Materials Science in Semiconductor Processing*, **1**, 17, (1998).

[3] A.E. Michel, *Appl. Phys. Lett.*, **51**, 487, (1987).

[4] N.E.B. Cowern, K.T.F. Janssen and H.F.F. Jos, *J. Appl. Phys.*, **68**, 6191, (1990).

[5] H.-J. Gossman, T.E. Haynes, P.A. Stolk, D.C. Jacobson, G.H. Gilmer, J.M. Poate, H.S. Luftman, T.K. Mogi and M.O. Thompson, *Appl. Phys. Lett.*, **71**, 3862, (1997).

[6] R.B. Fair and Y. Kim, *Electrochem. Soc. Proc. 2nd Symp. Defects in Silicon*, p. 423, (1991).

[7] A. Bousetta, J.A. van der Berg and D.G. Armour, *Appl. Phys. Lett.*, **58**, 1626, (1991)

[8] M.C. Ozturk, J.J. Wortman and R.B. Fair, *Appl. Phys. Lett.*, **52**, 963, (1986).

[9] S. Wolf and R.N. Tauber, in "*Silicon Processing for the VLSI Era, Vol. 1*", Lattice Press, (1986), p. 325.

[10] N.L. Turner, M. Current, T.C. Smith and D. Crane, *Solid State Tech.*, **28**, 163, (1985).

[11] M. Servidori, Z. Sourek and S. Slomi, *J. Appl. Phys.*, **62**, 1723, (1987).

[12] X. Bao and X. Hua, *Phys. Stat. Sol. A*, **123**, K89, (1991).

[13] S. Saito, S. Shishiguchi, K. Hamada and T. Hayashi, *Materials Chemistry and Physics*, **54**, 49, (1998).

[14] A.E. Michel, R.H. Kastl, S.M. Mader, B.J. Masters and J.A. Gardner, *Appl. Phys. Lett.*, **44**, 404, (1984).

[15] L. Kaabi, B. Remaki, C. Gontrand, P.F. Lo and B. Balland, *Nucl. Instrum. Methods B*, **112**, 196, (1996).

[16] G. Queirolo, P. Caprara, L. Meda, C. Guareschi, M. Anderle, G. Ottaviani and A. Armigliato, *J. Electrochem. Soc.*, **134**, 2905, (1987).

[17] D.F. Downey, C.M. Osburn and S.D. Marcus, *Solid State Tech.*, **40**, 71, (1997).

[18] P. Changae, K.M. Klein, Al.F. Tasch, R.B. Simonton and G.E. Lux, *IEDM Tech. Dig.*, p. 67, (1991).

[19] Y. Sato and I. Kawashima, *J. Electrochem. Soc.*, **141**, 1381, (1994).

[20] C.T. Liu, Y. Ma, H. Luftman and S.J. Hillenius, *IEEE Electron Dev. Lett.*, **18**, 212, (1997).

[21] J.M. Sung, C.-Y. Lu, M.L. Chen and S.J. Hillenius, *IEDM Tech. Dig.*, p. 447, (1989).

[22] T. Uchino, T. Shiba, T. Kikuchi, Y. Tamaki, A. Watanabe and Y. Kiyota, *IEEE Trans. Electron Dev.*, **42**, 406, (1995).

[23] K.H. Weiner and T.W. Sigmon, *IEEE Electron Dev. Lett.*, **10**, 260, (1989).

[24] G.K. Giust and T.W. Sigmon, *IEEE Electron Dev. Lett.*, **18**, 394, (1997).

[25] K.T. Kim and C.K Kim, *IEEE Electron Dev. Lett.*, **EDL-8**, 569, (1987).

[26] T.D. Kamins, "Preparation and properties of polycrystalline-silicon films", in *Handbook of Semiconductor Silicon Technology*, W.C.O. O'Mara, R.B. Herring and L.P. Hunt, Eds., Park Ridge, NJ, (1990).

[27] F. La Via and E. Rimini, *IEEE Trans. Electron Dev.*, **44**, 526, (1997).

[28] M.H. Juang, C.T. Lin and H.C. Cheng, *J. Appl. Phys.*, **76**, 1323, (1994).

[29] F. La Via, V. Privitera, C. Spinella and E. Rimini, *Semicond. Sci. Technol.*, **8**, 1196, (1993).

Chapter 3
Proximity Rapid Thermal Diffusion

3.1 Introduction

Rapid thermal diffusion and the difficulties that have arisen with various dopant sources, such as planar dopant sources, doped polysilicon and doped silicides, were discussed in Chapter 2. This chapter investigates spin-on dopants (SOD) as dopant sources for RTD.

Shallow p^+-n junctions have been fabricated by rapid thermal diffusion of boron from spin-on dopants (SOD) [1, 2]. Ling *et al.* [1] were the first to develop this technique, and they used SOD deposited directly onto a silicon substrate as a dopant source in RTD. However, there are several disadvantages to this technique. Deglazing of the thick SOD glass layer results in thinning of the masking oxide layer [3]. This oxide thinning, particularly at the junction edges, causes leakage currents in the device. The SOD also leaves a carbon rich, residual film, which is insoluble in HF, on the surface of the silicon [4]. In an attempt to avoid these problems, proximity RTD was investigated [5-7].

3.2 Proximity Rapid Thermal Diffusion Process

Proximity rapid thermal diffusion (RTD) has been investigated as a technique for fabricating shallow boron junctions in silicon. Spin-on dopant is spun onto a silicon wafer. This dopant source wafer is then stacked in proximity to a bare silicon product wafer, Figure 3.1, and used as a planar dopant source in proximity RTD.

During proximity RTD, boron trioxide (B_2O_3) evaporates from the spin-on dopant, across the separation gap to the product wafer, where it is adsorbed onto the surface. Boron

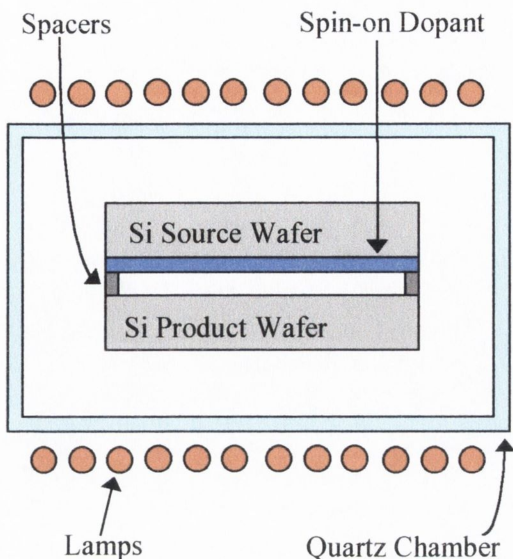


Figure 3.1. Wafer Configuration in Proximity RTD.

diffusion is then accomplished by means of a surface oxidation-reduction reaction between the B_2O_3 and the silicon wafer [8], given by:

$$2B_2O_3 + 3Si \leftrightarrow 4B + 3SiO_2 \quad (3.1)$$

In this reaction, doped SiO_2 is formed on the surface of the product wafer and becomes a dopant source for elemental boron, which diffuses into the silicon substrate. Excessive amounts of B_2O_3 can lead to the formation of silicides and other compounds of boron on the silicon surface. This *boron skin* causes a dark brown stain, is electrically insulating and often results in device failure due to open contacts. In order to avoid the formation of this boron skin, all of the proximity RTDs in this study are performed in an oxidising ambient of 25% O_2: 75% N_2 which promotes the formation of SiO_2.

Proximity RTD is a very suitable technique for shallow junction formation for several reasons; (i) the dopant source is quickly and easily prepared; (ii) there are no defects introduced into the wafers during the diffusion process and (iii) the dopant diffusion is minimised since the wafers are heated to high temperatures for short times.

Figure 3.2. Sitesa rapid thermal processor.

3.2.1 Experimental Procedure

A Sitesa rapid thermal processor was used throughout this study, Fig. 3.2. A schematic representation of the RTP is shown in Figure 3.3. The RTP furnace is composed of a water-cooled, highly reflective stainless steel, double-sided casing. A quartz chamber is located inside the furnace. A single wafer is supported inside the chamber on 3 quartz pins. Twenty four tungsten-halogen lamps surround the quartz chamber and supply radiant heat energy to the wafer in seconds. The intensity of the lamps is controlled by 2 thermocouples. One thermocouple is used to monitor the temperature of the process wafer and the second thermocouple is used to monitor a control wafer section. The control thermocouple provides the feedback loop that regulates the temperature of the lamps.

Initially, the temperature in the chamber was calibrated using a thermocouple wafer. (High temperature ceramic cement encapsulates the thermocouple junction and thermally bonds it to the substrate.) To ensure repeatability, the wafers are always inserted into the chamber with the major flat parallel to the furnace door. At a given temperature, there

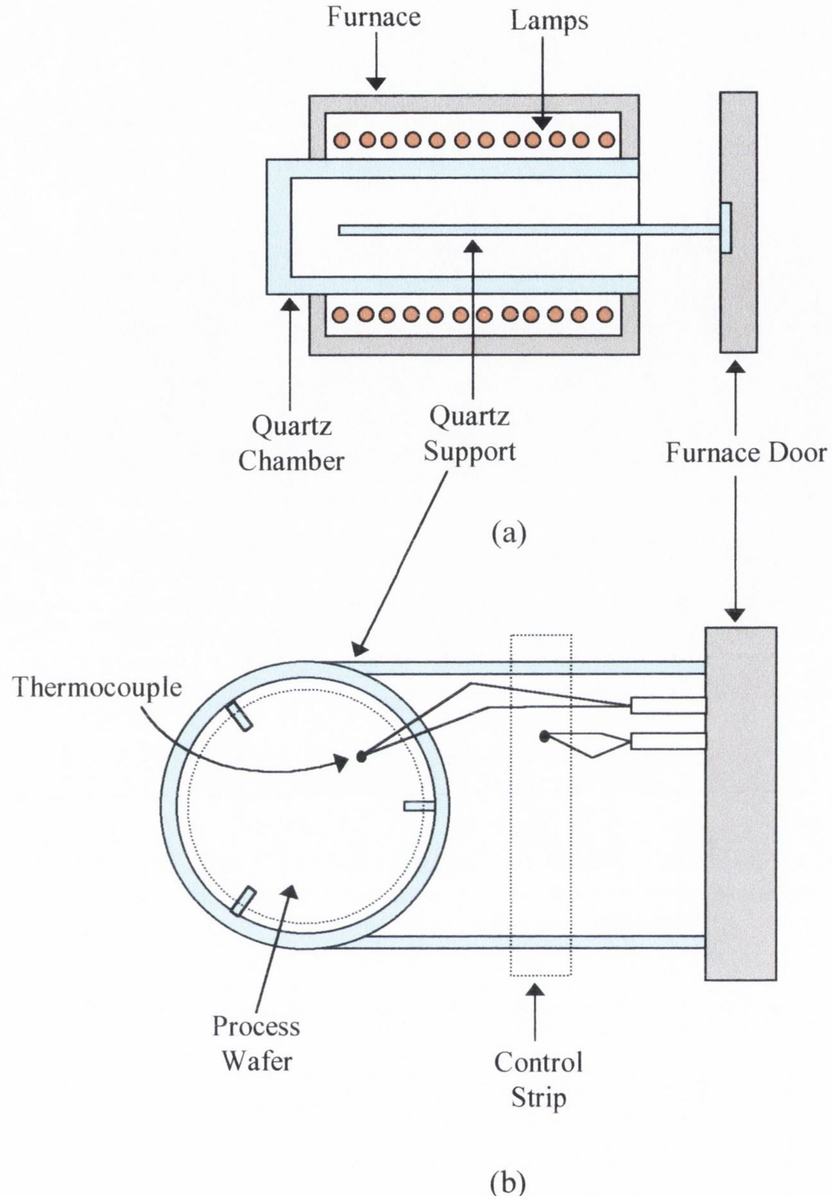


Figure 3.3. Schematic representation of Sitesa RTP. a) Side-view of RTP b) Plan-view of the quartz support.

was a variation across a 4 inch diameter, with the wafer temperature at the furnace door being 90 °C cooler than the wafer temperature at the back of the chamber. This anomalous temperature profile is due to a variation in the power supply to the halogen lamps. The temperature distribution across the wafer was exploited by measuring the temperature dependence of material and electrical properties, such as junction depth and sheet resistance, on a single wafer. A ramp-up rate of 48 °C/s and a ramp-down rate of

60 °C/s were used throughout this study. All of the RTDs were performed in 25% O_2 : 75% N_2.

Czochralski-grown 4 inch *n*-type, <100> oriented, silicon wafers, with a background concentration of 3 x 10^{14} cm^{-3}, were used throughout this work. As-received wafers were cleaned. SOD gel was spun onto the wafers at 6000 rpm for 15 s. The SOD is commercially available from Filmtronics, U.S.A. The SOD contains ethanol as a solvent. Since light organic solvents, such as ethanol, are a source of contamination in proximity RTD, it is necessary to remove these solvents from the dopant layer prior to RTD [5, 6]. The following section will determine the thermal process that is required to extract these contaminants from the SOD layer.

3.3 Removal of Contamination from Spin-on Dopants

Contamination during IC fabrication must be avoided. The main source of contamination in proximity RTD is from the SOD, in the form of organic solvents and water. Carbon, which is the principal element in organic solvents, acts as a nucleating agent for point defect condensation in silicon and generates lattice strain in the silicon substrate [9]. Therefore, the organic contaminants must be evaporated from the SOD before RTD in order to avoid generating defects in the product wafer during the doping process. If water is transferred from the SOD to the product wafer it will cause MOS device degradation. The presence of moisture in gate oxides has been found to increase their susceptibility to hot-carrier degradation by creating interface trapping states [10]. Moisture has already been observed to diffuse from spin-on glass (SOG) to gate oxides [11]. Consequently, it is also important to evaporate water from the SOD prior to RTD.

Fourier transform infrared spectroscopy (FTIR) was used to determine the thermal process that was required to evaporate the moisture and organic solvents from the SOD. The SOD was spun onto bare silicon wafers at 6000 rpm for 15 s. The SOD covered wafers were then baked in an oven at temperatures ranging from 20-400 °C for 15 minutes. Structural and compositional changes in the SOD layers were determined using a Fourier transform Biorad 60A spectrometer. FTIR spectra in the range of

450-4000 cm^{-1} were measured with a resolution of 2 cm^{-1} and 8 cm^{-1}. Figure 3.4 shows the infrared absorption spectra obtained for the source wafers after post-spin baking. Changes can be observed with changing temperature. A band in the 3000-3600 cm^{-1} region (which is the superposition of O-H stretching vibration and B-O-H and Si-OH vibrations) almost disappears after baking at 100 °C. This decrease is due to the evaporation of water present in the wafer and in the SOD layer. The absorption intensity of the Si-OH band in the 3200 cm^{-1} region increases with baking temperature and the wavelength position of the absorption maximum shifts to a smaller wavenumber. These are characteristic changes that occur on densification of borosilicate structures [2, 12-14]. The inset in Figure 3.4 shows that the organic contaminant, CH_3, disappears after baking at 200 °C.

In summary, a 200 °C bake for 15 minutes is adequate to remove the moisture and light organics from the SOD layer. The SOD dopant source is now ready to be used as a planar dopant source in proximity RTD.

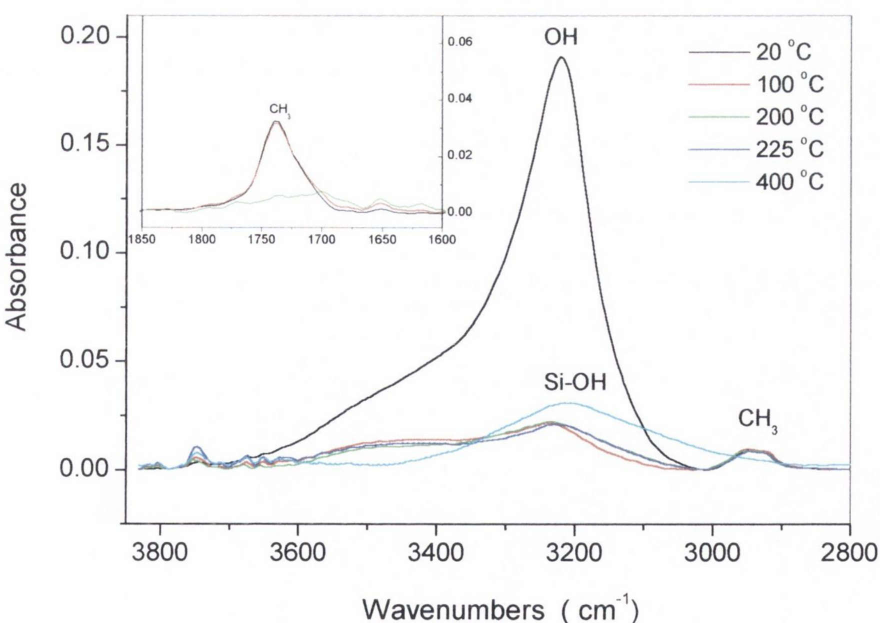


Figure 3.4. FTIR spectra of the source wafers after 15 minutes at various post-spin baking temperatures.

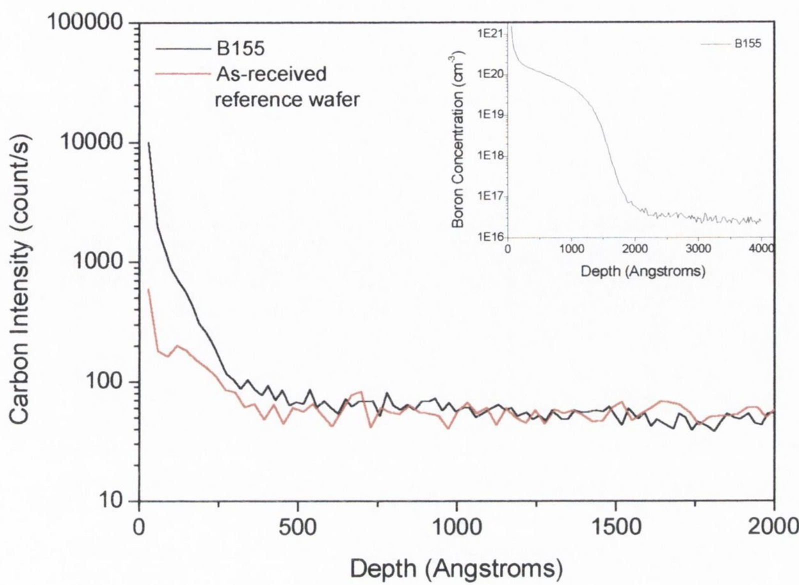


Figure 3.5. SIMS Carbon and Boron profiles of a product wafer that was doped with B155 during a RTD at 1050 °C for 16 s.

3.4 Spin-on Dopants

Two different types of SOD, B155 and B153 (Filmtronics, USA), were investigated as dopant sources in proximity RTD. These particular SODs were chosen because they offered a suitable compromise between physical characteristics and cost efficiency. The following sections discuss the properties and the suitability of these SODs for use in RTD applications.

3.4.1 Spin-on Dopant B155

SOD B155 is a vinyl backbone polymer. Vinyl backbone polymers consist of carbon CH_2=CH- chains [15], so the B155 SOD could consist of CH_2=CH-BO chains. Unfortunately, the exact chemical formula of B155 is not available since it is a patented Filmtronics product. B155 was spun onto a silicon source wafer and was then baked at 200 °C for 15 minutes to remove moisture and light organics. The dopant source wafer was then stacked in proximity to a bare silicon wafer and heated to 1050 °C for 16 s. Figure 3.5 shows the SIMS concentration profiles for boron and carbon in the doped

product wafer. The inset in Figure 3.5 shows that B is transferred to the product wafer. Carbon can be incorporated into Czochralski silicon during crystal growth, therefore, the C profile of a bare as-received silicon reference wafer was also measured. In Figure 3.5 the C concentration of the doped wafer is higher than the reference wafer, which provides evidence that C was transferred from the SOD layer to the product wafer. Since the light organic solvents were evaporated from the SOD before RTD, this C must be from the vinyl polymer itself. Zagozdzon-Wosik *et al.* [4] also reported C contamination from B155 during RTD. The transfer of C from the SOD to the product wafer will generate defects in the wafer, making SOD B155 an unsuitable dopant source in proximity RTD.

3.4.2 Spin-on dopant B153

Spin-on dopant B153 is a borosilicate gel, consisting of a B_2O_3-SiO_2 network [12, 13, 16]. The exact chemical formula of B153 is not exactly known because it is also a patented Filmtronics product. Unlike B155, there is no C in the B153 borosilicate network. Therefore, when the organic solvents have been evaporated, there should not be any C in the SOD gel.

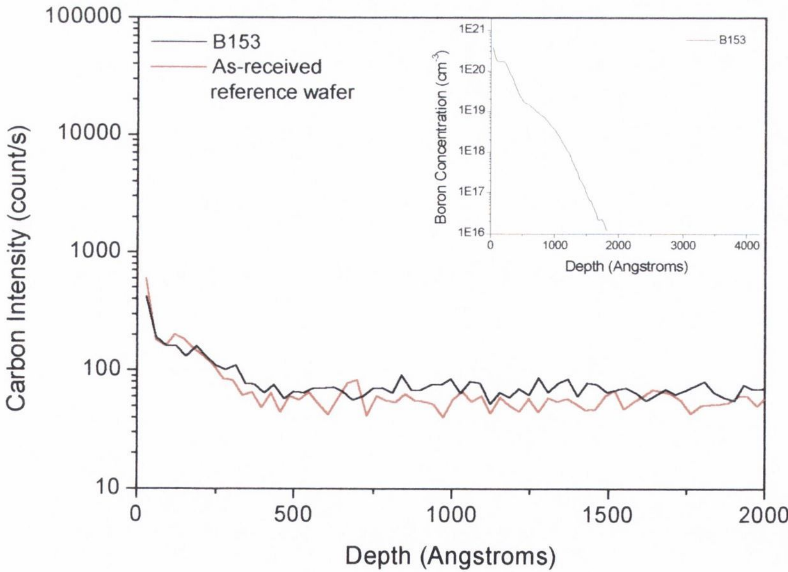


Figure 3.6. SIMS Boron and Carbon profiles of a product wafer that was doped using B153 in RTD 1050 °C for 16 s.

As in the previous section, B153 was spun onto a silicon source wafer and was then baked at 200 °C for 15 minutes to remove moisture and light organics. The dopant source wafer was then stacked in proximity to a bare silicon wafer and heated to 1050 °C for 16 s. Figure 3.6 shows the SIMS B and C concentrations for the doped product wafer. The C profile of the doped wafer is the same as the C profile in the reference wafer, showing that no C was introduced into the product wafer during proximity RTD. The inset in Figure 3.6 shows that B is successfully transferred from the B153 to the product wafer.

In summary, there is no transfer of C from B153 to the product wafer, and the wafer is doped with B after RTD. These characteristics make B153 a very suitable dopant source for proximity RTD.

3.5 B153 Doping Effectiveness

The B153 doping effectiveness was determined by doping 3 bare silicon wafers with a single B153 dopant source wafer. Each RTD was for 16 s at 1050 °C. The sheet resistance of each of the wafers was measured using a Jandel four-point probe, Table 3.1. If B153 was a stable constant dopant source, the 3 wafers would be equally doped and have equal values of sheet resistance, however, Table 3.1 shows that this is not the case. The first doped product wafer has the lowest sheet resistance which means that it has the highest dopant concentration. The 2nd product wafer has a higher sheet resistance, and the third wafer has the highest sheet resistance of the three. These 3 wafers are definitely not equally doped.

Table 3.1. Sheet resistance of 3 wafers that were doped with a single dopant source wafer.

Wafer	Sheet Resistance (Ω/sq)
1st Product	371
2nd Product	1023
3rd Product	4592

There are 2 possible explanations for the variation in sheet resistance in Table 3.1. If B153 was a limited dopant source, then it may have been depleted of B during the RTDs. This would cause the sheet resistance of the wafers to increase as the supply of boron decreased. However, it is known that borosilicate gels, such as B153, are thermodynamically unstable [2, 17]. At temperatures > 800 °C, the composition and structure of the SOD gel changes to that of a borosilicate glass (BSG). The boron supply from the dopant source changes as the structure changes. Therefore, there is a variation in the sheet resistance of the 3 product wafers because the state of the dopant source, and hence the dopant supply, changed during each high temperature RTD.

FTIR and SIMS were used to confirm that the B supply from the SOD during proximity RTD depends on the thermal treatment that the SOD receives before RTD. Firstly, Figure 3.7 shows the FTIR absorbance spectra for the B-O stretching vibration between 1300-1500 cm^{-1} for 5 different product wafers after RTD. Each product wafer was doped using proximity RTDs of 1050 °C for 16 s. The SOD layers were cured at 900 °C for 0-60 s prior to diffusion. The level of B that was transferred from each SOD layer during

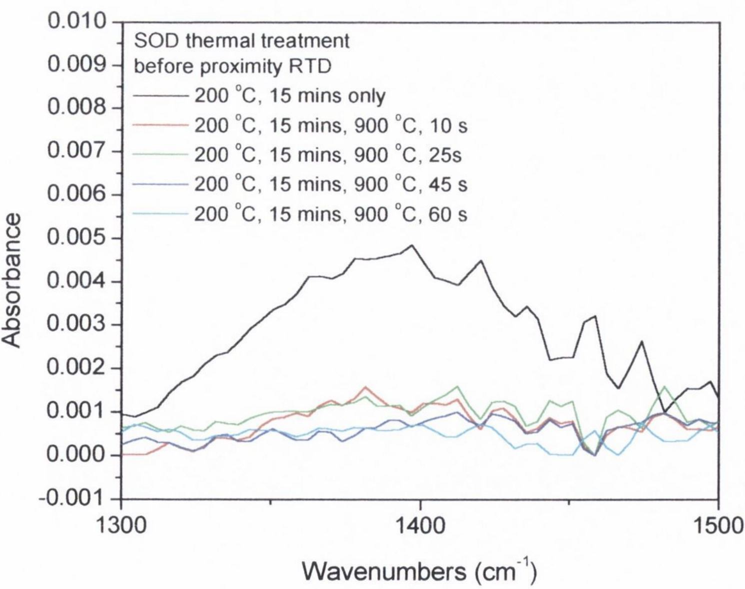


Figure 3.7. FTIR spectra of product wafers that were doped at 1050 °C for 16 s. Each SOD source received a different thermal treatment prior to diffusion.

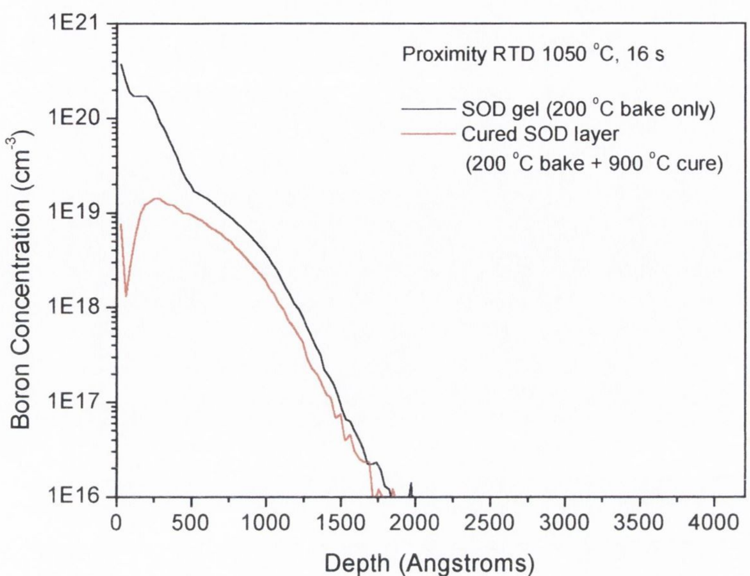


Figure 3.8. SIMS profiles of junctions that were formed by diffusion of boron from a SOD gel and a cured SOD layer.

RTD can be determined by the absorbance intensity of the B-O band in Figure 3.7. Figure 3.7 shows that the level of B that is transferred to the product wafers depends on the thermal treatment that the SOD receives before proximity RTD.

Secondly, one product wafer was doped with a SOD gel layer directly after the 200 °C post-spin bake, and a second product wafer was doped with a SOD layer that had been cured at 900 °C for 45 s. The SIMS dopant profiles of the wafers are shown in Figure 3.8. The wafer that was doped with the SOD gel (200 °C bake only) has a higher surface concentration and a deeper junction depth than the wafer that was doped with the cured (200 °C bake and 900 °C cure) dopant source. The SIMS analysis shows that the B supply from the SOD gel is greater than the dopant supply from the cured layer.

3.6 Conclusions

Spin-on dopants were investigated as dopant sources for RTD. Commercially available SODs were used as dopant sources. These dopants contain ethanol as a solvent. Ethanol is a source of C contamination, which generates lattice strain in silicon. Therefore, it is necessary to evaporate the solvent from the SOD prior to proximity RTD. It is also necessary to remove moisture from the SOD prior to RTD to prevent moisture-induced hot-carrier degradation in gate oxides on the product wafer. FTIR analysis determined that a 200 °C bake for 15 minutes was sufficient to remove the solvents and moisture from the SOD.

Two different types of SOD were investigated, B155 and B153. B155 is a vinyl polymer and SIMS analysis revealed that as well as B, C was also transferred from the B155 SOD to the silicon product wafer during proximity RTD. Carbon acts as a nucleating agent for point defect condensation in silicon and generates lattice strain in the silicon substrate. Therefore, B155 is not a suitable dopant source for high quality shallow junctions in proximity RTD. B153 SOD is a borosilicate gel. SIMS analysis proved that B153 was an efficient B dopant source, and that C was not transferred from this SOD to the product wafer, making it a suitable dopant source for proximity RTD.

Borosilicate gel structures, such as SOD B153, are thermodynamically unstable. At temperatures > 800 °C, the composition and structure of the SOD gel changes to that of a borosilicate glass (BSG). Sheet resistance measurements, FTIR and SIMS analysis showed that the boron supply from the dopant source depends on the thermal treatment that the SOD receives before RTD. The conversion of B153 gel to BSG will be investigated in Chapter 4.

References

[1] E. Ling, P.D. Maguire, H.S. Gamble and B.M. Armstrong, *IEEE Electron Dev. Lett.*, **EDL-8**, 96, (1987).

[2] A. Slaoui, L. Ventura, A. Lachiq, R. Monna and J.C. Muller, *Mat. Res. Soc. Proc.*, **387**, 365, (1995).

[3] A. Usami, M. Ando, M. Tsunekane and T. Wada, *IEEE Trans. Electron. Dev.*, **39**, 105, (1992).

[4] W. Zagozdzon-Wosik, P.B. Grabiec and G. Lux, *IEEE Trans. Electron. Dev.*, **41**, 2281, (1994).

[5] W. Zagozdzon-Wosik, J.C. Wolfe and C.W. Teng, *IEEE Electron Dev. Lett.*, **12**, 264, (1991).

[6] M. Nolan, T. Perova, R.A. Moore and H.S. Gamble, *J. Non-Crystal. Solids*, **254**, 89, (1999).

[7] M. Nolan, T. Perova, R.A. Moore and H.S. Gamble, *Mat. Res. Soc. Proc.*, 30th November – 4th December, Boston, **C3**, 58, (1998).

[8] S.K. Ghandhi, in *VLSI Fabrication Principles*, Wiley, New York, (1994), p. 211.

[9] K.V. Ravi, in *Imperfections and Impurities in Semiconductor Silicon*, Wiley, New York, (1981) p. 77.

[10] S. Wolf in "*Silicon Processing for the VLSI Era*", Vol. 3, Lattice Press, (1995), p. 647.

[11] N. Lifshitz and G. Smolinsky, *IEEE Electron Dev. Lett.*, **12**, 140, (1991).

[12] W. Kern, *RCA Review*, **32**, 429, (1971).

[13] J. Wong, *J. Electron. Mater.*, **5**, 113, (1976).

[14] W.A. Pliskin, in *Semiconductor Silicon*, H.R. Huff and R.R. Burgess, Eds., Electrochemical Society, Pennington, NJ, (1973), p. 506.

[15] H. Remy, in *Treatise on Inorganic Chemistry*, Elsevier, (1956), p. 437.

[16] M.A. Villegas and J.M. Fernández Navarro, *J. Mater. Sc.*, **23**, 2464, (1988).

[17] R.M. Almeida and C.G. Pantano, *J. Appl. Phys.*, **68**, 4225, (1990).

Chapter 4

Stabilisation of B153 Spin-on Dopant Source

4.1 Introduction

In the preceding chapter, it was established that spin-on dopant B153 is a suitable dopant source for proximity RTD. B153 SOD is in the form of a borosilicate gel. On heating, the composition and structure of the SOD gel changes to that of a borosilicate glass (BSG). The boron supply from the dopant source depends on the thermal treatment that the SOD receives prior to RTD, as described in the previous chapter. Therefore, in order to use one dopant source to dope several wafers repeatably, it is important that the borosilicate gel is initially converted into a stable BSG dopant source. This chapter shows how FTIR spectroscopy and spectroscopic ellipsometry were used to determine, and subsequently optimise, the conversion of the gel to BSG.

In the past decade, a number of investigations have been performed to understand how different thermal treatments influence the properties of SOD and spin-on glass (SOG) films. Information on different properties, such as doping efficiency, density, refractive index and thickness have been found for phosphosilicate glass (PSG) films [1-5] and for spin-on glass films [6-8]. However, very few studies have been done on spin-on dopant BSG films [9-12]. Detailed investigations were performed on BSG [13-18] and borophosphosilicate (BPSG) [19-27] films, obtained by low-pressure and atmospheric pressure chemical vapour deposition (LPCVD and APCVD). However, as was noted by Becker *et al.* [21], the properties of these films strongly depend on the preparation technique. This chapter presents the detailed FTIR and ellipsometric analysis of BSG films spin-coated on silicon wafers. As deposited, the SOD layer had a thickness of 120 nm. This layer is very thin in comparison to the BSG layers that were used in previous works. It should also be noted that there is no thin capping silicon dioxide layer

on the surface of the BSG films studied here, since the concentration of boron is not high and the films do not absorb moisture on exposure to the atmosphere after the curing process. The measurements were repeated six months after the initial measurements were made, and there was no change in the FTIR spectra recorded.

4.2 Experimental Procedure

4.2.1 Proximity rapid thermal diffusion

Czochralski-grown 4 inch n-type, <100> oriented, 9-15 Ω cm resistivity silicon wafers were used throughout this study. As-received wafers were cleaned using H_2SO_4:H_2O_2 followed by a HF dip. SOD gel was spun onto the wafers at 6000 rpm for 15 s. The SOD contains ethanol as a solvent. Therefore, following application of the SOD, it was necessary to bake the wafers at 200 °C to evaporate moisture and light organics from the dopant layer, as described in the previous chapter. The rapid evaporation of the solvent during film deposition (through the spinning) leads to the formation of a porous gel film. The film structure in this case [6] strongly depends on the molecular weight distribution of the oligomers present in the solution. Thus, even after baking at 200 °C, silanol (Si-OH) and molecular water remain in the gel network to some extent. Fully dense films are only produced after heat treatment at temperatures between 800 and 1200 °C [6,25].

Dopant source wafers were cured in a Sitesa rapid thermal processor (RTP) at different temperatures in the range 800-1000 °C for 2-60 s (Table 4.1), to convert the SOD gel to a BSG layer prior to RTD. FTIR and spectroscopic ellipsometry were used to determine, and subsequently optimise, the conversion of the gel to a BSG layer.

Following the stabilisation of the dopant source, the source wafer was then stacked in proximity to a silicon product wafer on 0.5 mm silicon spacers. Silicon product wafers were doped using the optimised dopant source. All rapid thermal heat treatments were performed in 25% O_2:75% N_2.

Table 4.1. Rapid thermal treatment to convert SOD gel to BSG

Sample	Curing Temperature (°C)	Curing Time (s)
1	800	10
2	800	25
3	800	45
4	800	60
5	850	10
6	850	25
7	850	45
8	850	60
9	900	10
10	900	25
11	900	45
12	900	60
13	950	5
14	950	10
15	950	25
16	950	45
17	1000	2
18	1000	5
19	1000	10
20	1000	25
21	1000	45

4.2.2 FTIR Measurements

The measurements on BSG films were performed in the spectral range from 4500-500 cm^{-1} with a Fourier transform BioRad 60A spectrometer using a Globar source, a KBr beam splitter, and a mercury cadmium telluride (MCT) detector. The spectra were collected with 8 cm^{-1} resolution, and 64 scans were averaged for each spectrum to improve the signal-to-noise ratio. A bare silicon substrate was used as a reference for all of the samples, and the analysis of the transmission spectra was performed neglecting reflectance. A clear and detailed analysis of the different films requires consideration of spectra both in the high-frequency region (from 3000-2000 cm^{-1}) and in the low-

frequency region (from 1600-400 cm^{-1}). The spectra are the superposition of the spectra of the BSG and the silicon wafer and it is necessary to subtract the spectra of the bare wafer from the measured spectra in order to obtain the spectra for the BSG. Hence, all spectra discussed below are difference spectra. It should be noted that the influence of the substrate absorption becomes especially important if heat-treated samples are analysed, as was shown by Becker *et al.* [21]. However, we believe that, due to the extremely short thermal treatment in the RTP reactor (the maximum time used was 60 s), there will be no change to the substrate absorption.

4.2.3 Spectroscopic Ellipsometry Measurements

The refractive index and the thickness of the sample were obtained by means of spectroscopic ellipsometric measurements. This technique has been widely used for the characterisation and study of physical properties of thin films [28-30].

The samples were measured at room temperature with a rotating polarizer spectroscopic ellipsometer (SOPRA GESP5) in the visible to near ultraviolet wavelength range (225-880 nm). For each spectrum 250 points were measured, giving a wavelength resolution of 3 nm. The data were collected using the current tracking mode for the position of the analyser in order to improve the accuracy of the measurements. The ellipsometer has the capability of measuring variable angles of incidence and, in order to improve the accuracy in determining the physical parameters of the samples, each spectrum was measured at two different angles of incidence (65 and 75 degrees). Several different models were used to fit the data measured (Tan Ψ and Cos Δ) simultaneously at the two angles of incidence. A three-layer model (air, oxide layer, and silicon), using a Cauchy law fit for the oxide layer, was chosen based on the accuracy of the results obtained. This chapter presents the results obtained from this model and compares them to results obtained from FTIR.

4.2.4 SIMS and Sheet Resistance Measurements

The sheet resistance of the wafers was measured using a Jandel four-point probe. The boron concentration profiles of the product wafers were measured using a CAMECA IMS 3F secondary ion microscope in the National Microelectronics Research Centre (NMRC), Ireland.

4.3 Results and Discussion

4.3.1 FTIR

IR spectroscopy is known as a fast and non-destructive method of determining a number of important properties of dielectric films. Those properties are: boron and phosphorous concentration, absorbed water content, chemical bonds, thickness and density [13-15, 40]. The ratio of the B-O band peak intensity at a frequency near 1370 cm^{-1} to the Si-O-Si absorbance maximum near 1070 cm^{-1} is widely used to determine the boron content for glass films of thickness up to 2.5 μm [1-18].

FTIR was used to investigate the structure and properties of the SOD layer after the rapid thermal curing (RTC) process step. The wafers were annealed at 800, 850, 900, 950 and 1000 °C for 2-60 seconds, Table 4.1. The FTIR spectra for the wafers treated at 800, 900 and 1000 °C are shown in Figures 4.1-4.3. These spectra clearly show the change in structure with variation in temperature and time. In particular, as the heating time is increased, the Si-O-Si asymmetric stretching vibration band at ~1075 cm^{-1} shifts to a higher frequency and the absorption intensity increases (Figure 4.4). This shows that the silicon dioxide thickness has increased and the SOD layer has densified [8, 13-15].

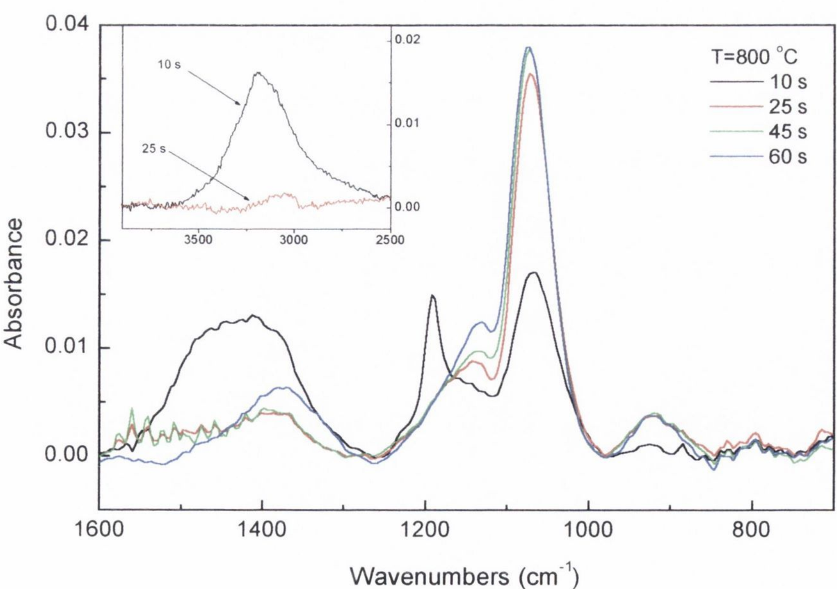


Figure 4.1. IR absorbance spectra of SOD layers which were annealed at 800 °C for various times.

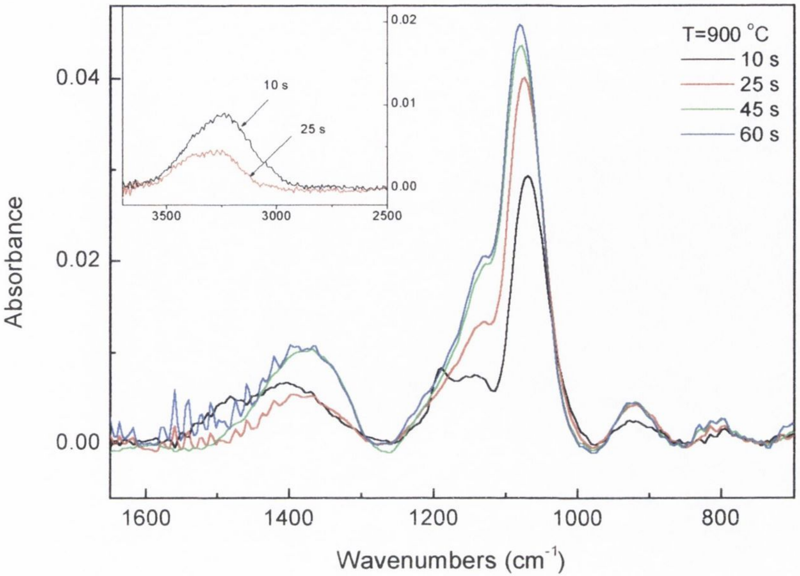


Figure 4.2. IR absorbance spectra of SOD layers which were annealed at 900 °C for various times.

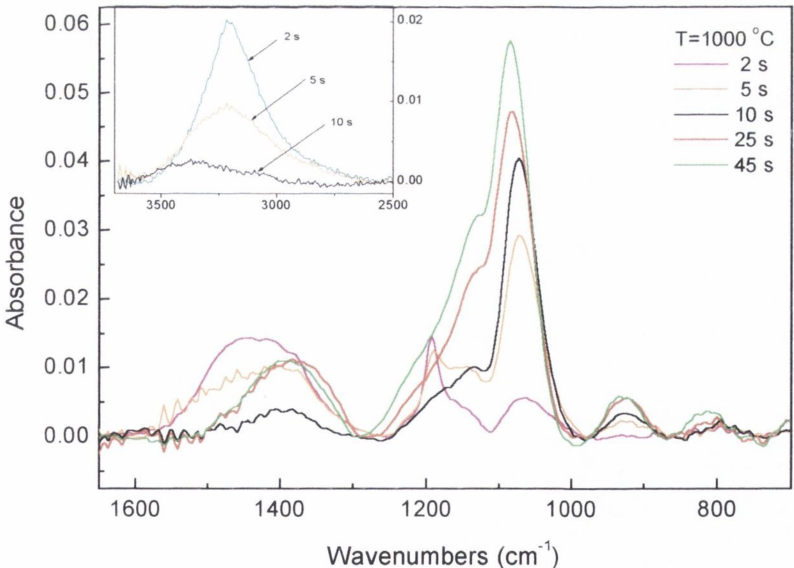


Figure 4.3. IR absorbance spectra of SOD layers which were annealed at 1000 °C for various times.

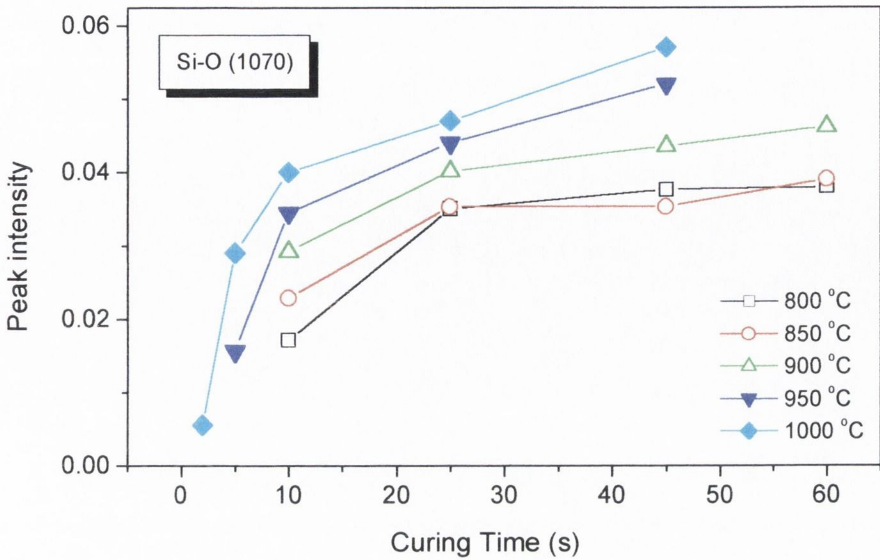


Figure 4.4. The dependence of the peak intensity of Si-O-Si stretching vibrations band on RTC temperature and time.

For a given temperature, the B-O stretching vibration band in the region 1370-1440 cm^{-1} shifts to a lower frequency as the heating time increases (Figure 4.5). This is a characteristic change, which occurs on densification of borosilicate glass [13, 14]. The absorption intensity and the contour of this band also changes significantly as the heating time is increased. These changes were observed for all of the RTC temperatures investigated. The intensity of the broad band that appears at ~1430 cm^{-1} (peak position of B-O stretching vibrations [13-15]) decreases rapidly with time and the peak shifts to a lower frequency. Then the intensity of this band increases again and reaches a saturation value with a peak position at ~1370 cm^{-1}. This complicated behaviour can be explained by the influence of a two different processes; i) the decrease in film thickness, as a result of the layer shrinking with temperature; ii) the formation of the borosilicate network through the bridging oxygen atoms.

The dramatic decrease in the absorption intensity of the OH stretching vibration band (shown as insets in Figs. 4.1-4.3) verifies the progressive loss of the OH groups from the

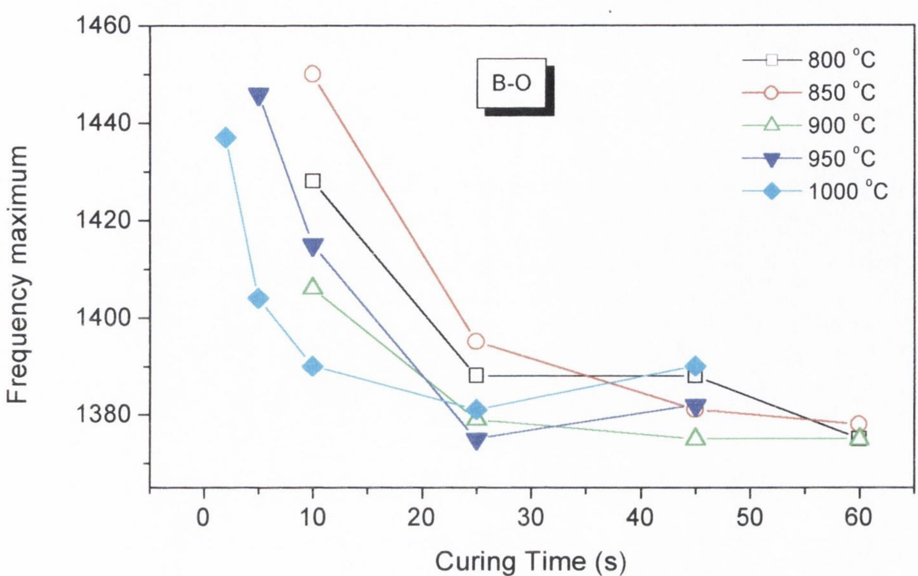


Figure 4.5. The dependence of the peak position of B-O stretching vibrations band on RTC temperature and time.

SOD. This result indicates that during thermal treatment, hydrolysis and condensation reactions are continuing, releasing water, and producing Si-O-Si, B-O-B and Si-O-B bonds [27, 31-33] at the expense of the Si-OH, B-OH. This is also consistent with the growth of the Si-O-Si and Si-O-B vibrational bands at 1075 and 910 cm^{-1}, with increasing temperature. The rate of loss depends on the RTC process temperature. Enhancement of band intensity with temperature is explained by incorporation of the B-OH groups [31, 32] into the borosilicate phase, during which Si-O-B bonds are progressively formed. The observed minimum in the intensity of B-O band at certain temperatures and time occurs during the transformation of tetraborate to borate and then to a borosilicate glass. At the intermediate stage, some boron may exist in ionic forms that are not IR active. The minimum observed for B-O peak intensity may be explained by the presence of IR-inactive intermediate products that are formed during reconstruction of the layer rather than by boron loss from the layer (Figs. 4.1-4.3). However, the shift of the B-O band at ~1370 cm^{-1} to the high frequency side seen in Fig. 4.3 after 45 s treatment at 1000 °C is due to boron diffusion inwards into the underlying silicon, and outwards into the atmosphere.

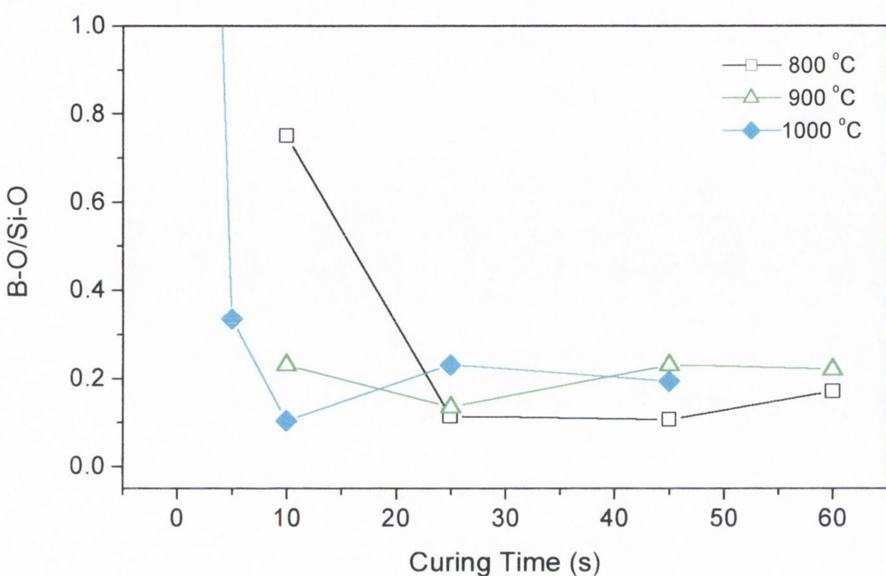


Figure 4.6. The dependence of the B-O/Si-O infrared peak ratio on curing time for different temperatures.

The band at 910 cm^{-1} can also be used to determine when the BSG forms. This band belongs to the bending vibration of B-O-Si units, and is a characteristic band, which appears on the formation of BSG [13-15]. The band intensity increases as the RTC time increases, Figs. 4.1-4.3. Figure 4.3, in particular, shows the increase in intensity and a stabilisation of the peak at 25 s.

Analysis of the ratio of two vibrational bands B-O/Si-O (B-O at 1370 cm^{-1} and Si-O-Si at 1075 cm^{-1}) allows the concentration of the boron in the dopant layer to be determined [10-12]. Figure 4.6 shows the dependence of the B-O/Si-O ratio on RTC time. For 800 °C, there is very little change in the ratio between 25 and 45 s, which signifies a very slow rate of conversion of the gel to a glass. For 900 °C, the ratio reaches a maximum value of 0.2 after 45 s cure and remains the same after a 60 s cure. The 1000 °C profile reaches a maximum at a shorter time of 25 s.

4.3.2 Spectroscopic Ellipsometry

The refractive index and the thickness of the SOD layers have been determined from spectroscopic ellipsometry measurements. Figure 4.7 shows the refractive index values

as a function of the curing time for different temperatures. All of the samples have a high value of refractive index for short curing times, ≤ 10 s. These high initial values are due to OH groups in the SOD layer, the existence of which are clearly shown in Figures 4.1-4.3 by the presence of the 3000-3600 cm^{-1} band.

There is a significant decrease in refractive indices for curing times > 10 s, following the loss of the OH groups. This observation is also consistent with the FTIR results, Figs. 4.1-4.3, obtained for the OH band in the 3000-3600 cm^{-1} region. The refractive indices decrease to values between 1.47 and 1.55, which are typical values for BSG [35].

Initially, an increase in the refractive index could be expected due to the densification of the gel to a BSG structure [7]. However, the decrease caused by the evaporation of OH (0.25-0.15) is one order of magnitude bigger than the expected decrease on BSG densification (0.025) [7, 8].

Oxide growth at the underlying Si/BSG interface might also be expected, since the thermal process is performed in an oxidising ambient. Sol-gel layers can retain a structure of interconnected pores, even after high temperature thermal treatment. We

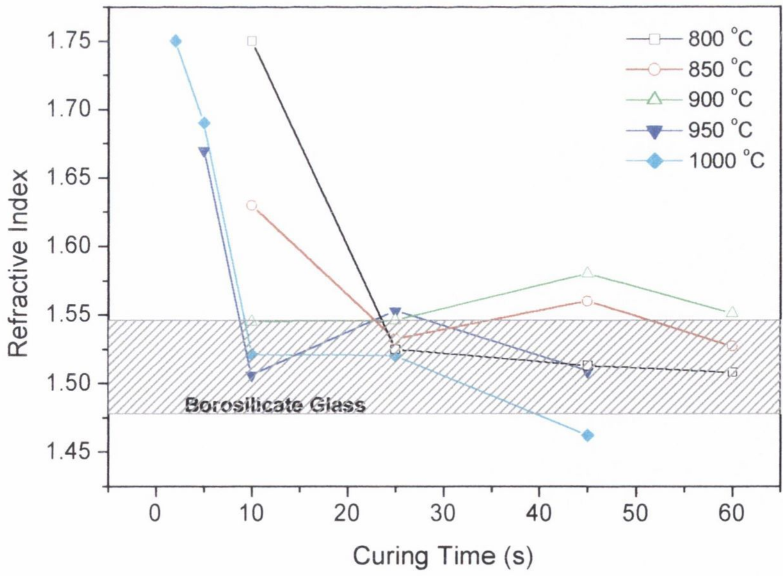


Figure 4.7. The dependence of the refractive index of SOD layers on RTC temperature and time.

have estimated the porosity of the BSG layer to be about 3% [36]. Hence the growth of SiO_2 beneath the BSG layer, via the interconnected pores of the BSG layer, should also be considered. However, it is very difficult to observe changes in the global refractive index due to any densification or oxidation process while the OH groups influence the results.

For curing times > 10 s, the refractive index continues to vary with temperature and time. There are several factors that must be taken into consideration: (i) the refractive index will decrease if the underlying Si oxidises, because of the decrease in the overall B_2O_3/SiO_2 ratio; (ii) the refractive index will increase as the layer becomes more dense; (iii) any diffusion of boron from the BSG layer will also cause a decrease in the refractive index.

At 800 °C, there is no significant change in refractive index for curing times > 20 s. This is in good agreement with the B-O/Si-O dependence at 800 °C measured by FTIR (Fig. 4.6) and further verifies that any change is very slow at this temperature. The refractive index for 850 and 900 °C increases by a small amount and reaches a maximum value at 45 s. This refractive index behaviour is consistent with densification of the BSG layer, evidence for which also comes from the characteristic decrease in B-O frequency (1400 cm^{-1} region in Figs. 4.1-4.3). Next, the refractive index for the 850 °C and 900 °C samples decreases by 0.03 between the 45 and 60 s cures. Oxidation of the underlying Si or boron out-diffusion from the BSG layer would produce such a decrease. The film thickness determined from the spectroscopic ellipsometry calculations (Figure 4.8), is in excellent agreement with the FTIR results in Figure 4.4, and shows that the film thickness increases with curing time. In contrast, there was no difference in the sheet resistance of the silicon source wafer before and after the 850 °C and 900 °C RTC steps, confirming that the boron did not diffuse into the silicon source wafers, in significant quantities, at these process temperatures. Therefore, the decrease of 0.03 in refractive index must arise from the oxidation of the underlying silicon.

At 950 °C, the refractive index shows the same trend as the 850 °C and 900 °C plots, except that the changes occur more rapidly. The increase in reaction rate at higher temperatures is revealed in Figures 4.4 and 4.8 by the more rapid increase in oxidation

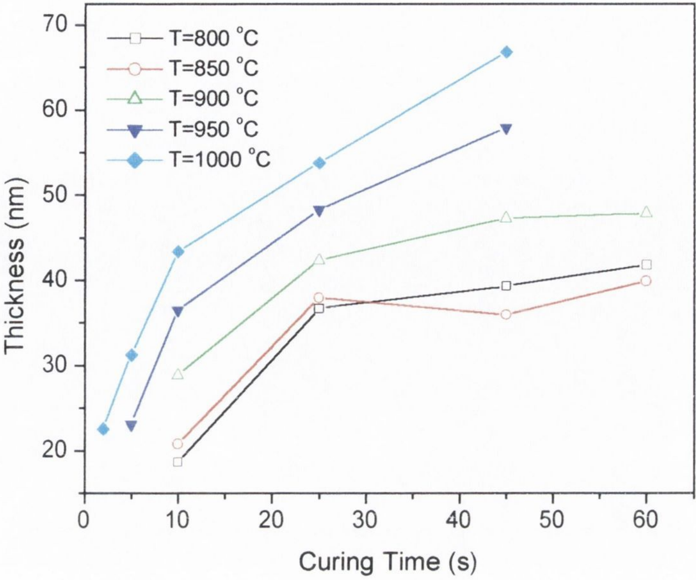


Figure 4.8. Variation in thickness of SOD layers (ellipsometry measurements) with changing RTC temperature and time.

and layer thickness. The refractive index is highest at 950 °C after a 25 s cure. This value is not as high as the maximum value measured for 900 °C, but it is possible a higher, unmeasured value exists between 20 s and 30 s. After processing for 45 s, the refractive index decreases by over 0.03 from the 25 s value. In this case, however, both oxide layer growth and boron out-diffusion may be contributing to the decrease. In Figure 4.5, the maximum B-O frequency position for 950 °C increases from 1370 cm^{-1} to 1390 cm^{-1} between the 25 and 45 s cures. This change in frequency is consistent with a decrease of boron content in the BSG layer. Sheet resistance measurements also confirm that boron diffuses from the oxide layer into the underlying silicon. These trends are more pronounced for RTC at 1000 °C but due to the increase in the conversion rate at high temperatures, additional data points would be required to determine the detailed refractive index behaviour.

The most favourable RTC process step would be at a temperature that does not result in boron loss from the BSG layer, since the layer will be used as a boron dopant source, but that also converts the SOD gel to BSG in the shortest time possible. It is possible to determine the optimum conditions by combining the FTIR, spectroscopic ellipsometry

and sheet resistance results. We know from sheet resistance measurements, and from spectroscopic investigations (see Figures 4.5-4.7), that boron diffuses from the BSG layer at temperatures greater than 900 °C. As the rate of conversion increases with RTC temperature, a 900 °C curing step will be more time efficient than thermal treatments at 800 °C or 850 °C. Figures 4.5-4.7 show that the BSG layer is at its most stable after a 45 s cure at 900 °C. The optimum RTC conditions for the conversion of SOD gel to BSG are thus curing at 900 °C for 45 s.

Having determined the optimum RTC conditions, a SOD layer was converted to a BSG layer by curing at 900 °C for 45 s. The efficiency of the resulting BSG layer as a dopant source was determined by using a single dopant source wafer to dope two bare silicon wafers in proximity RTD. Each RTD was at 1050 °C for 1 s. Figure 4.9 shows the SIMS dopant profile for the two wafers that were doped with the same dopant source wafer. The doping profiles of the two wafers are the same, having a surface concentration of $5x10^{18}$ cm^{-3} and a depth of ~306 Å at 10^{18} cm^{-3}. The proximity doped wafers will be analysed in more detail in Chapter 5.

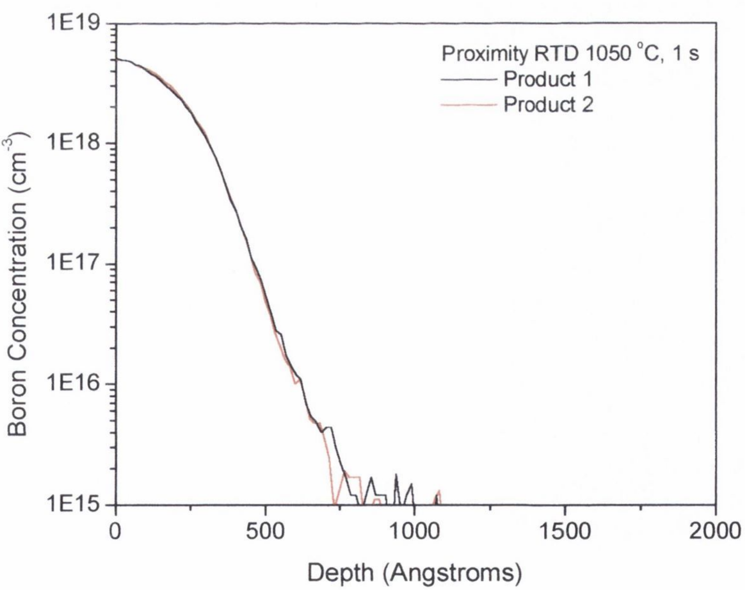


Figure 4.9. SIMS B concentration profiles of two silicon wafers that were doped by a single dopant source in proximity RTD.

4.4 B_2O_3 Transfer to the Product Wafer

The remainder of this chapter reports the step coverage and ease of removal of the B_2O_3 layer that is transferred from the B153 dopant source to the product wafer. In each case, the B153 SOD was cured at 900 °C for 45 s prior to proximity RTD.

4.4.1 B_2O_3 Step Coverage on the Product Wafer

B_2O_3 step coverage on the product wafer is very important in the fabrication of LDD PMOS devices. The shallow boron junction must align with the edge of the polysilicon gate. Therefore, the B_2O_3 dopant source must flow into the corner of the gate, to ensure that the boron diffuses and aligns with the edge of the gate. It is well known that B_2O_3 has low viscosity, in particular, it is added to PSG to decrease the temperature of viscous deformation and improve step coverage [26, 38, 39]. For this application, B_2O_3 should flow into the gate corner and provide uniform step coverage. Figure 4.10a shows a SEM micrograph of a plasma-etched polysilicon gate. Figure 4.10b shows the patterned gate covered in B_2O_3 after proximity RTD. The B_2O_3 reflows into the gate corner, providing adequate step coverage and ensuring that the boron diffused junction will align with the gate edge. This is a significant result, as it highlights the suitability of proximity RTD with SOD for LDD device fabrication.

4.4.2 B_2O_3 Layer Removal

Having fabricated a shallow boron junction on the product wafer, it is important that the B_2O_3 layer is removed to prevent further unwanted diffusion in subsequent process steps. The layer was easily removed following 5 s in 10:1 HF. FTIR measurements confirmed that all of the B_2O_3 had been removed.

4.5 Conclusions

The thermal conversion of a SOD borosilicate gel layer to a stable BSG dopant source layer has been confirmed by combining the results of FTIR, spectroscopic ellipsometry and sheet resistance measurements. The films were cured in a rapid thermal processor in the temperature range 800-1000 °C for periods of 2-60 s in an atmosphere of

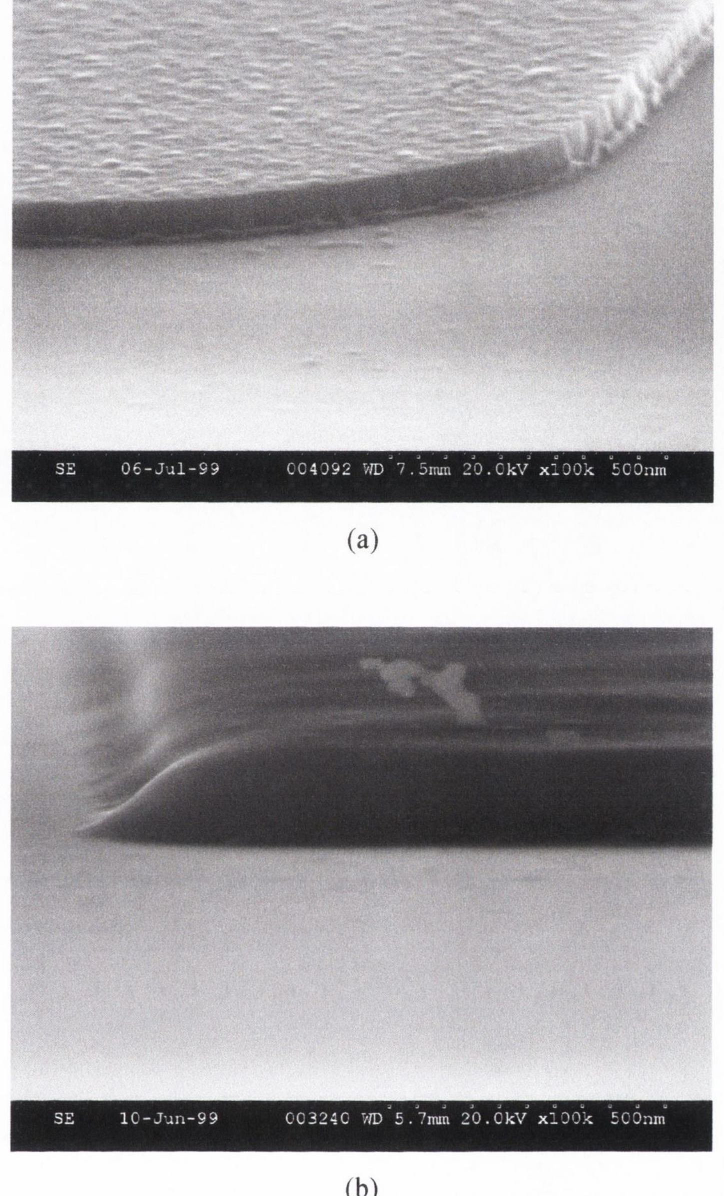


Figure 4.10. SEM micrographs (x100k) (a) Plasma-etched polysilicon gate structure; (b) Polysilicon gate covered in a conformal B_2O_3 layer after proximity RTD.

25% O_2:75% N_2. The changes in the properties of the layer can be summarised as follows:

1. Thickness – the thickness of the layer increases with increasing temperature and time. This increase in thickness is attributed to oxide growth at the Si-BSG interface during heat treatment in an oxidising ambient.

2. Refractive Index – initially, the refractive index decreases following the loss of OH groups from the layer. The refractive index then increases as the BSG layer densifies, and subsequently decreases again as the underlying silicon oxidises. For the 950 °C and 1000 °C treatments, the refractive index decreases again due to boron diffusion from the BSG layer.

Heating at 900 °C for 45 s was the optimum rapid thermal curing process step required to convert the gel to a BSG, since there was no detectable boron loss from the BSG dopant source at this temperature.

The optimised BSG dopant source was used to uniformly dope two silicon product wafers. The wafers had a boron surface concentration of $5x10^{18}$ cm^{-3} and a depth of ~306 Å at 10^{18} cm^{-3}.

The B_2O_3 reflows during proximity RTD providing excellent step coverage. It is important that the dopant reflows into the gate corners to ensure that the shallow boron junction aligns with the polysilicon gate in an LDD PMOS transistor. The B_2O_3 is easily removed in HF after RTD to prevent unwanted diffusions in subsequent process steps.

References

[1] W. Zagozdzon-Wosik, P. Grabiec, F. Romero-Borja, L.T. Wood and G. Lux, *Mat. Res. Symp. Proc.,* **303**, 297, (1993).

[2] S. Montandon, N. Zagozdzon-Wosik, Jia Li, W.T. Taferner and B. Bensaoula, *Electrochem. Soc. Proc.*, **95-20**, 552, (1955).

[3] W. Zagozdzon-Wosik, P. Grabiec and G. Lux, *J. Appl. Phys.*, **75**, 337, (1994).

[4] W. Zagozdzon-Wosik, P. Grabiec and G. Lux, *IEEE Electron Devices*, **41**, 2281, (1994).

[5] P. Grabiec, W. Zagozdzon-Wosik and G. Lux, *J. Appl. Phys.,* **78**, 204, (1995).

[6] R.M. Almeida and C.G. Pantano, *J. Appl. Phys.*, **68**, 4225, (1990).

[7] L. Ventura, B. Hartiti, A. Slaoui. J.C. Muller and P. Siffert, *Mat. Res. Soc. Symp. Proc.,* **284**, 197, (1993).

[8] A. Slaoui, L. Ventura, A. Lachig, R. Monna and J.C. Muller, *Mat. Res. Soc. Symp. Proc.*, **387**, 365, (1995).

[9] D.M. Haaland and C.J. Brinker, *Mat. Res. Soc. Symp. Proc.,* **32**, 267, (1984).

[10] M. Rastogi, W. Zagozdzon-Wosik, F. Romero-Borja, J.M. Haddleson, R. Beavers, P. Grabliec and L.T. Wood, *Mat. Res. Soc. Symp. Proc.,* **342**, 369, (1994).

[11] S. Mone, W. Zagozdzon-Wosik and M. Rastogi, *Mat. Res. Soc. Symp. Proc.,* **405**, 351, (1996).

[12] D.J. Taylor, D.Z. Dent, D.N. Braski and B.D. Fabes, *J. Mater. Res.*, **11**, 1870, (1996).

[13] W. Kern, *RCA Review*, **32**, 429, (1971).

[14] J. Wong, *J. Electron. Materials*, **5**, 113, (1976).

[15] W.A. Pliskin, in *Semiconductor Silicon*, H.R. Huff and R.R. Burgess, Eds., Electrochemical Society, Pennington, NJ, (1973), p. 506.

[16] A.S. Tenney, *J. Electrochem. Soc.,* **118**, 1658, (1971).

[17] E.A. Taft, *J. Electrochem. Soc.*, **118**, 1985, (1971).

[18] A.S. Tenney and J. Wong, *J. Chem. Phys.*, **56**, 5516, (1972).

[19] W. Kern and G.L. Schnabe, *RCA Review*, **43**, 423, (1982).

[20] W. Kern, W.A. Kurylo and C.J. Tino, *RCA Review*, **46**, 117, (1985).

[21] F.S. Becker, D. Pawlik, H. Shäfer and G. Staudigl, *J. Vac. Sci. Tech.,* **B4**, 732, (1986).

[22] J.E. Franke, T.M. Niemczyk and D. Haaland, *Spectrochim. Acta*, **50A**, 1687, (1994).

[23] R.A. Carpio and J. Taylor, *SPIE*, **2638**, 38, (1995).

[24] T.W. Dyer, *J. Electrochem. Soc.*, **145**, 3950, (1998).

[25] S. Rojas, R. Comarasca, L. Zanotti, A. Borghesi, S. Sassella, G. Ottaviani, L. Moro and P. Lazzeri, *J. Vac. Sci. Technol.,* **B10**, 633, (1992).

[26] L.D. Madsen, A.C. de Wilton and J.S. Mercier, *Chemtronics*, **5**, 35, (1991).

[27] D.M. Haaland and C.J. Brinker, *Mat. Res. Soc. Symp. Proc.,* **32**, 267, (1984).

[28] I. Susuki, M. Ejima, K. Watanabe, Y. Xiong and T. Saitoh, *Thin Solid Films*, **313-314**, 214, (1998).

[29] B. Drevillon, *Thin Solid Films*, **313-314**, 625, (1998).

[30] S. Bruynooghe, F. Bertin, A. Chabli, J.-C. Blanchard and M. Couchaud, *Thin Solid Films*, **313-314**, 722, (1998).

[31] M. Nogami and Y. Moriay, *J. Non-Crystaline Solids*, **48**, 359, (1982).

[32] A.D. Irwin, J.S. Holmgren, T.W. Zerda and J. Jonas, *J. Non-Crystaline Solids*, **89**, 191, (1987).

[33] M.A. Villegas and J.M. Fernández Navarro, *J. Mater. Sci.,* **23**, 2464, (1988).

[34] R.M. Levin, *J. Electrochem. Soc.,* **129**, 1765, (1982).

[35] T. S. Izumitani, *Optical Glasses*, American Institute of Physics, USA, (1986), p. 17.

[36] M. Nolan, T.S. Perova, R.A. Moore, C.E. Beitia, J.F. M[c]Gilp and H.S. Gamble, to be published.

[37] *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Association, (1997), p.74.

[38] W. Kern and R.K. Smeltzer, *Solid State Tech.*, **28**, 171, (1985).

[39] J.E. Tong, K. Schertenleib and R.A. Carpio, *Solid State Tech.*, **27**, 161, (1984).

Chapter 5

Boron Diffusion & Properties of Proximity RTD Doped Wafers

5.1 Introduction

Considerable progress has been made in the past thirty years in understanding dopant diffusion in silicon. Specific diffusion mechanisms have been analysed. Anomalous diffusion effects, such as oxidation-enhanced diffusion and transient-enhanced diffusion, have been investigated in detail. However, in the case of boron diffusion, a satisfactory understanding has not been reached due to the complexity of the diffusion process. Exact values for diffusion parameters, such as fractional interstitialcy, have not been agreed upon to date.

Much of the effort in understanding dopant diffusion in the past decade has focused on ion implantation and damage annealing because of the industrial importance of these processes. More recently, alternative diffusion processes, such as RTD, have been investigated. This study concentrates on RTD of boron from borosilicate glass. Very little work has been reported on RTD from BSG in the literature, and the doping mechanism is not fully understood.

Firstly, this chapter presents the basic diffusion mechanisms in silicon. Some of the effects that alter dopant diffusivities are then introduced in an attempt to demonstrate the complexity of boron diffusion. Next, an analysis of bare and oxidised wafers that were doped using proximity RTD is presented. Sheet resistance, SIMS boron concentration depth profiles and diffusion coefficients are shown. Possible doping mechanisms are proposed, although the exact process is not fully understood. Finally, the shallow

junctions fabricated using proximity RTD are compared to other fabrication techniques that have been presented in the literature.

5.2 Boron Diffusion

5.2.1 Diffusion Mechanisms

Diffusion describes the process by which atoms move in a crystal lattice. There are several mechanisms by which an atom can diffuse through a lattice: i) Interstitial Diffusion; ii) Substitutional Diffusion; iii) Interstitial-Substitutional Diffusion and iv) Interstitialcy Diffusion. These mechanisms occur in all semiconductor materials, however, in this section the diffusion processes are described specifically for diffusion through a silicon lattice.

i) Interstitial Diffusion

This mechanism of diffusion is illustrated in Figure 5.1, where an impurity atom moves through the crystal lattice by jumping from one interstitial site to an adjacent interstitial site. This process is relatively fast, because of the large number of interstitial sites in a semiconductor. Impurities such as sodium and lithium move through silicon by this mechanism.

ii) Substitutional Diffusion

Figure 5.2 depicts substitutional diffusion. An impurity atom diffuses through the crystal by jumping from one lattice site to the next, thus substituting for the original silicon host atom. However, it is necessary that this adjacent site is vacant in order to allow substitutional diffusion to occur. Since the concentration of vacancies is quite low, substitutional diffusion occurs at a much slower rate than interstitial diffusion. All *n*-type impurities, except phosphorous, diffuse through silicon by this mechanism.

iii) Interstitial-Substitutional Diffusion

In this case, impurity atoms occupy substitutional as well as interstitial sites. However, they only move at a significant rate when in interstitial sites. A substitutional impurity atom can become an interstitial, leaving behind a vacancy, Figure 5.3. This mechanism

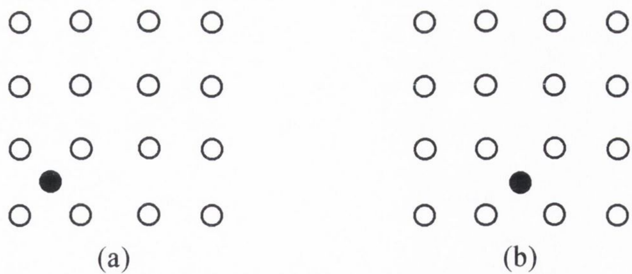


Figure 5.1. Diffusion by the interstitial mechanism.

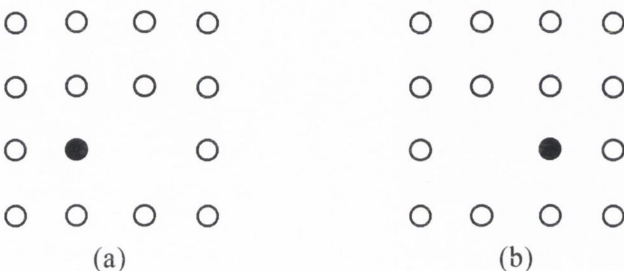


Figure 5.2. Diffusion by the substitutional mechanism.

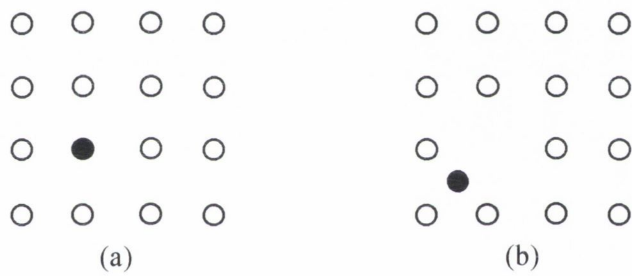


Figure 5.3. Interstitial-substitutional diffusion by the dissociative mechanism.

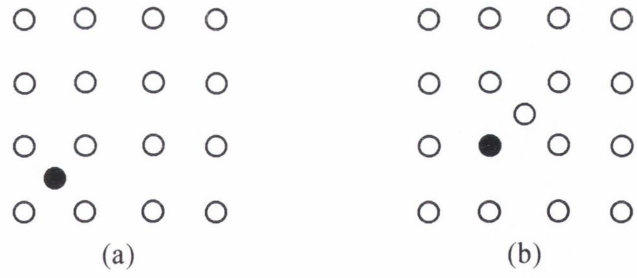


Figure 5.4. Interstitial-substitutional diffusion by the kick-out mechanism.

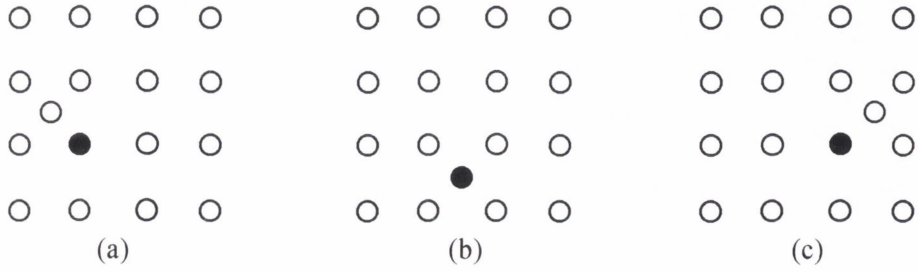


Figure 5.5. Interstitialcy diffusion process.

is called the *dissociative* mechanism. Copper and nickel move in silicon by the dissociative mechanism. The *kick-out* mechanism is an alternative pathway for interstitial-substitutional diffusion, Figure 5.4. A rapidly moving interstitial diffuser can move into a substitutional site by displacing a silicon atom that is already in place, resulting in a silicon interstitial. Gold and platinum diffuse through silicon by the kick-out mechanism.

iv) Interstitialcy Diffusion

This is a modified version of substitutional diffusion. Figure 5.5 shows a silicon interstitial pushing a substitutional impurity atom into an interstitial site. These interstitial impurities can now diffuse to adjacent substitutional sites and create new silicon interstitials. Thus, the interstitial position of the impurity atoms is a transitional state, in moving from one substitutional site to another. Interstitialcy diffusion is faster than substitutional diffusion. It is generally accepted that the diffusion of boron and phosphorous is dominated by this process.

Some distortion of the silicon lattice must occur in accommodating either native point defects or dopant impurity atoms. Silicon has a tetrahedral radius of 1.18 Å. Boron has a tetrahedral radius of 0.88 Å, corresponding to a misfit of $\varepsilon = 0.254$. Since boron is smaller than silicon, the presence of boron in the silicon lattice contracts the lattice and induces strain. Strain-induced defects in boron-doped silicon will be discussed later in Chapter 7.

5.2.2 Fractional Interstitialcy Component

One of the most controversial single issues relating to diffusion in silicon since the 1960s has been the identification of which mechanism is responsible for dopant diffusion; interstitial type or vacancy type. As a consequence, a parameter called the fractional interstitialcy component, f_{AI}, was introduced, where f_{AI} is:

$$f_{AI} = \frac{D^*_{AI}}{D^*_{AI} + D^*_{AV}} \tag{5.1}$$

D^*_{AI} is the diffusivity of dopant A with interstitials under equilibrium conditions and D^*_{AV} is the diffusivity of dopant A with vacancies under equilibrium conditions. Simply, f_{AI} is the fraction of dopant diffusion that occurs through the interstitial mechanism. In general, if f_{AI} is less that 0.5 for a particular dopant, the vacancy mechanism plays a larger role in the diffusion process than the interstitial mechanism, and conversely, if $f_{AI} > 0.5$ the dominant diffusion mechanism is the interstitial mechanism. There is general agreement that the fractional interstitialcy component for boron is $f^B_{AI} > 0.5$, however, a range of very different values ($f^B_{AI} = 0.25$ [1], $= 0.3$ [2], $= 0.8$ [3], $= 0.94$ [4, 5]) is to be found in the literature. To date, no agreement has been reached for an exact value of f^B_{AI}.

5.2.3 Oxidation-Enhanced Diffusion

In the 1970's experiments showed that non-local effects such as surface oxidation could dramatically alter dopant diffusivities [6]. Understanding these effects helped to clarify the nature of diffusion processes in silicon, however, the debate continues to this day concerning the precise reaction mechanisms. It was observed that boron diffusivity was enhanced during silicon oxidation [7]. It was also observed that stacking faults grow during oxidation, by addition of silicon atoms to the dislocation ends. Hu [8] was the first to present a model of B diffusion during oxidation that combined the observations of stacking fault growth and B diffusion characteristics. The explanation for both of these observations was that silicon interstitials are injected into the silicon lattice during surface oxidation [9, 10]. Hu assumed that B diffused by the interstitialcy mechanism, Fig. 5.5. The oxidation-induced flux of Si interstitials interchange with B atoms and create B interstitials. These B interstitials then migrate through the lattice until they interact with a vacancy. They then continue to diffuse via a vacancy mechanism, until interaction with another Si interstitial occurs. Thus, the introduction of Si interstitials during oxidation enhances B diffusion.

Silicon nitridation is the exact complement of oxidation. During the nitridation of silicon, vacancies are injected into the lattice; as a result stacking faults shrink and boron diffusion is retarded [11, 12]. The oxidation-enhanced diffusion (OED) and the

nitridation-retarded diffusion results indicate that boron diffusion has a substantial interstitial component of diffusion.

5.2.4 Transient-Enhanced Diffusion

Transient-enhanced diffusion (TED) adds another complication to understanding boron diffusion in silicon. TED was introduced briefly in Chapter 2. Implanted boron atoms in silicon exhibit TED during implant annealing. The enhanced diffusion arises from the excess interstitials that are generated by the implant, for instance, {113} interstitial clusters are generated during implantation, these defects dissolve during the subsequent thermal anneal releasing interstitials, which subsequently enhance B diffusion [13, 14]. Although TED is minimal under rapid thermal annealing conditions, it becomes more pronounced and critical with the continuing trend of reducing the thermal budget in manufacturing. TED now threatens to impose severe limitations on the minimum device dimensions attainable in next generation devices. Another ambiguous feature of TED is the fact that the peak portion of the implanted boron profiles remains static and electrically inactive upon annealing when an apparent critical dopant concentration less than solid solubility is reached [15]. This feature may be due to interstitial-induced boron clustering, though further investigation is necessary to verify this phenomenon [13, 16]. During the past decade, substantial experimental efforts have been devoted to untangling the mechanism of TED with the objective of improving simulation programs designed to predict dopant diffusion during processing [13, 14, 16, 17]. However, a satisfactory state of understanding has not been reached yet due to the inherent complexity of the process.

While TED is not an issue in this proximity RTD study, since ion implantation was not used, it could become an issue when complete PMOS devices are fabricated. During the fabrication of Disposable LDD devices (Chapter 8), the deep source/drain junctions are formed using ion implantation prior to LDD formation. Therefore, defects that are generated during ion implantation may cause anomalous diffusion during subsequent proximity RTD.

5.2.5 Other Anomalous Diffusion Effects

It is already evident that boron diffusion is an extremely complex issue. Unfortunately, as technology advances, the level of complexity also increases. Since the introduction of rapid thermal processing in the 1980s, further anomalous diffusions have been observed. Enhanced diffusion during rapid thermal annealing of ion implanted silicon was observed [18, 19]. The enhanced diffusion results from annealing implant damage, as discussed in the previous section. Enhanced boron diffusion from doped polysilicon into silicon during rapid thermal processing, and the difficulties encountered modelling the diffusion has also been reported [20]. In Ref. [20] four different process simulators failed to model the experimental boron diffusion profiles.

5.3 Enhanced Diffusion of Boron from Borosilicate Glass during RTD

Miyake reported enhanced diffusion of boron in silicon during doping from borosilicate glass (BSG) [21, 22]. Enhanced diffusion was observed for high-concentration BSG, and the enhancement of the diffusion coefficient increased with increasing boron concentration in BSG. Enhanced diffusion was observed during rapid thermal diffusion (RTD) and during furnace diffusion (FD) which suggests that the enhanced diffusion is inherent to the BSG/Si structure. In a BSG/Si system, the boron is transferred across the BSG/Si interface by:

$$2B_2O_3 + 3Si \rightarrow 4B + 3SiO_2 \quad (5.2)$$

This reaction indicates that oxidation of Si takes place at the BSG/Si interface. In a process analogous to OED, silicon interstitials are also generated that enhance boron diffusion in the BSG/Si system.

Miyake determined that diffusion by FD was less than by RTD. He also observed that boron-enhanced diffusion did not occur during RTD when the BSG layer was not present. Therefore, RTD itself does not cause enhanced diffusion, it promotes the enhanced diffusion that is inherent to the BSG/Si structure. The reason boron diffusion

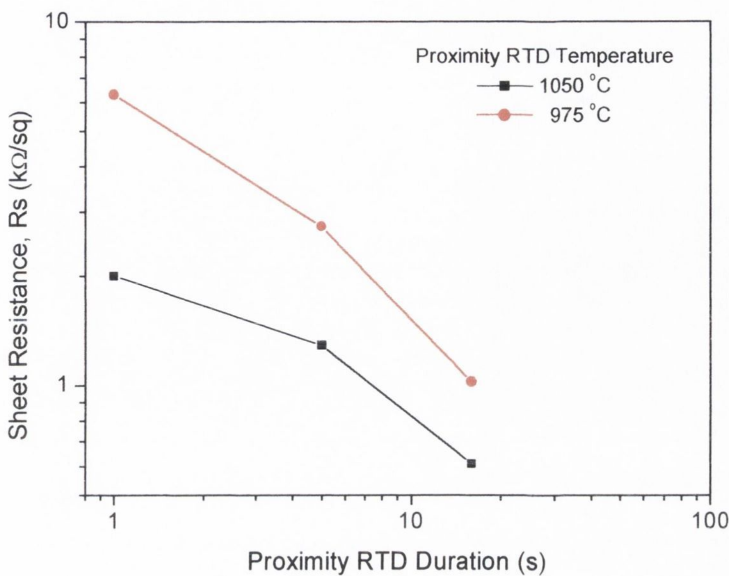


Figure 5.6. Sheet resistance of doped product wafers as a function of RTD duration at 975 °C and 1050 °C.

is further enhanced during RTD is not understood. Ishikawa *et al.* [23] reported boron-enhanced diffusion during IR heating, which implies that the enhanced diffusion in RTD may be related to IR heating. The mechanism of enhanced diffusion related to IR heating is also not clear at present.

5.4 Properties of Proximity RTD Doped Product Wafers

5.4.1 Sheet Resistance Measurements and SIMS Analysis

Proximity RTD was used to dope several bare silicon wafers. The B153 spin-on dopant was converted to stable BSG before RTD (Chapter 4). Figure 5.6 shows measured sheet resistance as a function of RTD duration at 975 and 1050 °C. The sheet resistance, *Rs*, decreases with increasing RTD temperature and duration, as expected.

Figure 5.7 shows SIMS boron concentration profiles for silicon wafers that were doped at 1050 °C. Before the doping mechanism is discussed, it is important to draw attention to

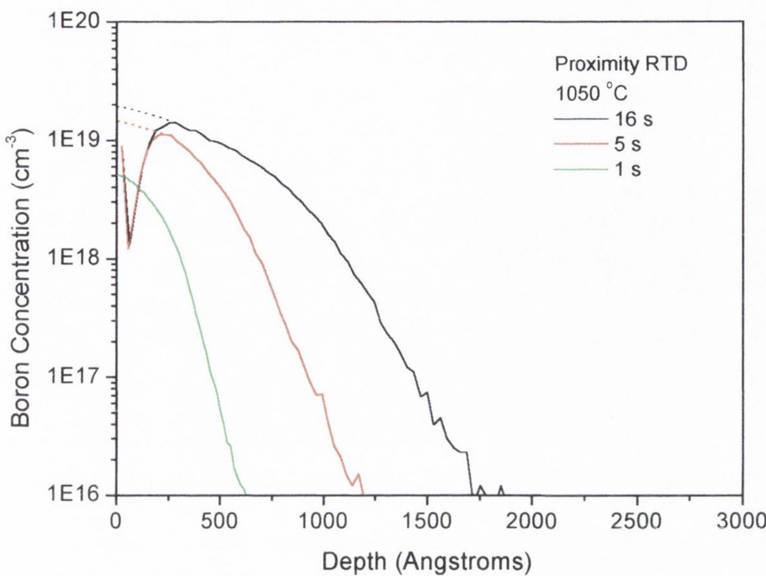


Figure 5.7. Boron concentration profiles of bare silicon wafers that were doped during proximity RTD at 1050 °C for various durations.

a specific feature of SIMS analysis. At the start of any SIMS analysis, it is necessary to stabilise the primary beam sputtering process before any data can be collected. This set-up procedure involves sputtering the sample surface, and results in a loss of information at the surface. The apparent decrease in boron concentration near the silicon surface for two of the samples (5 and 16 s RTD) in Figure 5.7 is caused by this artefact of SIMS. To avoid this loss of important information from the surface of the 1 s sample, a 1000 Å layer of polysilicon was deposited on the doped product wafers prior to SIMS analysis. Subsequently, the sputtering process stabilised in the polysilicon layer instead of in the silicon substrate, and avoided any loss of information from the doped silicon surface. The SIMS profiles for the 5 and 16 s samples were extrapolated to the silicon surface (dotted lines in Figure 5.7) to allow boron surface concentration to be determined.

The surface boron concentration, *Cs*, and the profile depths that were measured in Figure 5.7 are shown in Figure 5.8 as a function of RTD duration. The profile depths were determined by taking the 10^{18} cm^{-3} B concentration level as a reference level. The depth increases with increasing RTD duration. It can be seen that *Cs* also increases initially with RTD duration, and then tends to become saturated for longer diffusion

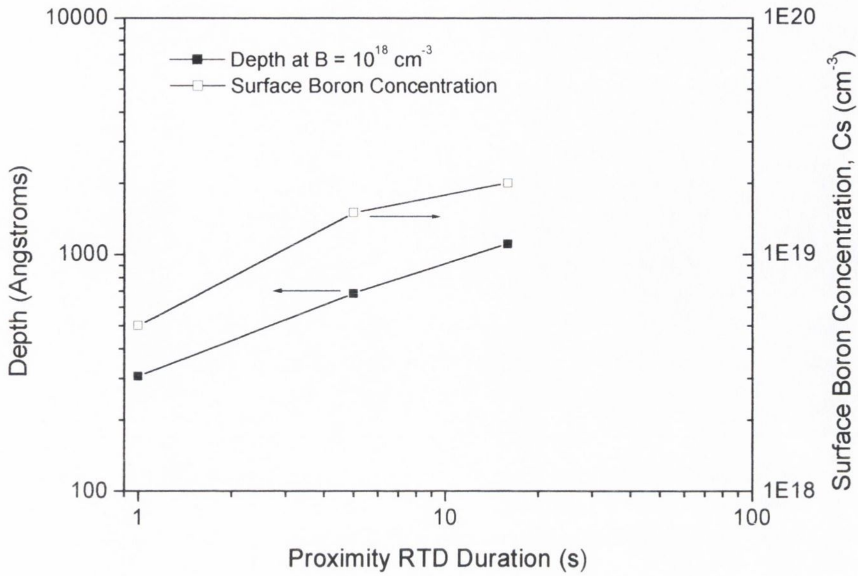


Figure 5.8. Surface boron concentration and depth as a function of RTD duration.

times. The following explanation will clarify this observation. The proximity RTD doping mechanism involves two processes that occur simultaneously. Firstly, the dopant oxide, B_2O_3, must be transported from the BSG to the surface of the product wafer, and secondly, the B_2O_3 must react with the silicon substrate to produce SiO_2 and B. Initially, the silicon is not effectively doped with boron since the B_2O_3 must be transferred to the wafer. The doping efficiency improves with time as the supply of boron increases at the silicon surface. Figure 5.6 shows that *Rs* decreases with RTD temperature, which demonstrates that doping efficiency also improves with temperature. At 1050 °C, higher vapour pressure of B_2O_3 improves mass transport in the gas phase, thus improving doping efficiency, for shorter diffusion times in particular.

Currently, in MOS technology, the gate oxide is left in place during LDD formation. A screen oxide is also grown on top of the gate oxide before LDD doping to protect the gate oxide against damage during implantation, and subsequent spacer removal (Chapter 8). Proximity RTD does not introduce any damage into the silicon lattice, however, until spacer removal technology is improved to avoid oxide damage, it is still necessary to dope the LDD extensions through an oxide layer, typically 120 Å thick. A

120 Å oxide was grown on a bare silicon wafer during a 5 minute dry oxidation at 900 °C in a mixture of 50% O_2:50% N_2. This oxidised wafer was doped during a 5 s proximity RTD at 1050 °C. The SIMS concentration profile is shown in Figure 5.9. The profile shows a depth of 308 Å at 10^{18} cm^{-3}, and a surface concentration of $1.2x10^{20}$ cm^{-3}, which corresponds to the solid solubility of boron in silicon at 1050 °C [24]. The depth is less than for an identical diffusion into a bare silicon wafer (Fig. 5.7), so the oxide layer does act as a barrier for B diffusion. However, the surface concentration is higher and the depth is greater than might be expected. B_2O_3 must diffuse through the SiO_2 to the Si interface, where it reacts to form SiO_2 and B. The introduction of impurities in SiO_2 radically changes its properties. Boron is a substitutional impurity (or *network former*) in SiO_2, and replaces the Si atom, weakening the SiO_2 network. In this case, the weakened bond structure in the 120 Å oxide layer may allow B_2O_3 to enter the SiO_2 more easily, and also to diffuse through it more rapidly, resulting in a deeper than expected B profile. As boron is formed at the silicon interface, it can either diffuse into the silicon lattice or into the SiO_2 layer. Boron segregates in SiO_2 (segregation coefficient $m < 1$) and because it has a slow diffusivity in SiO_2, it will remain trapped in the oxide. A combination of this B segregation with the continuous supply of B_2O_3 from the B153

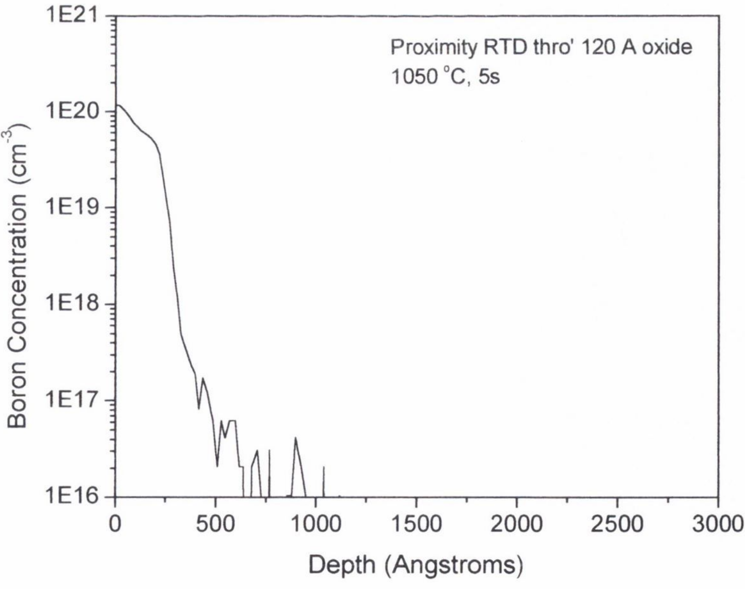


Figure 5.9. SIMS concentration profile of an oxidised wafer that was doped wafer during proximity RTD at 1050 °C for 5 s.

BSG would result in a very highly doped oxide layer on the surface of the product wafer. This build-up of boron in the oxide could cause a reduction in B depth in the Si, but it would also provide a very high concentration dopant supply to the silicon. This could be the reason that the wafer in Figure 5.9 is doped to solid solubility. Very little work has been reported in the literature about boron RTD from BSG, in particular RTD though thin oxides. The doping mechanism related to RTD of B from BSG through thin oxides is not clear at present, and further study is required to determine the mechanism.

Since the 120 Å oxide changes the boron dopant profile significantly, the effect of doping through a native oxide was investigated. All of the wafers in Figure 5.6 had a native oxide on the surface before proximity RTD. A wafer that was HF passivated was doped for 5 s at 1050 °C, however, the dopant profile was exactly the same as for a wafer that had a native oxide present. In conclusion, the native oxide does not influence the dopant profile during proximity RTD.

SUPREM (Stanford University Process Engineering Model) was used in an attempt to model the diffusion profiles shown in Figures 5.7 and 5.9. Unfortunately, the profiles have not been modelled successfully to date using this particular simulator. In general, new technologies have always been developed before a modelling capability for these technologies. SUPREM uses various parameters, including fractional interstitialcy, f_{AI}^{B}, (Section 5.2.2), to model boron diffusion. It is clear from the literature that the precise boron diffusion mechanism is not understood and that no exact value of f_{AI}^{B} has been agreed upon to date. General issues, such as the value of f_{AI}^{B}, and issues specific to proximity RTD must be fully understood before a satisfactory diffusion model can be developed.

5.4.2 Diffusion Coefficients

Boron diffusion coefficients were evaluated from the SIMS depth profiles in Figures 5.7 and 5.9 using Boltzmann-Matano analysis [25-27]. The Boltzmann-Matano analysis allows the diffusion coefficient to be calculated from an experimentally determined concentration profile if the surface concentration is constant. In Figure 5.6 Boltzmann-Matano analysis cannot be applied to the 1 and 5 s diffusions since *Cs* is not constant.

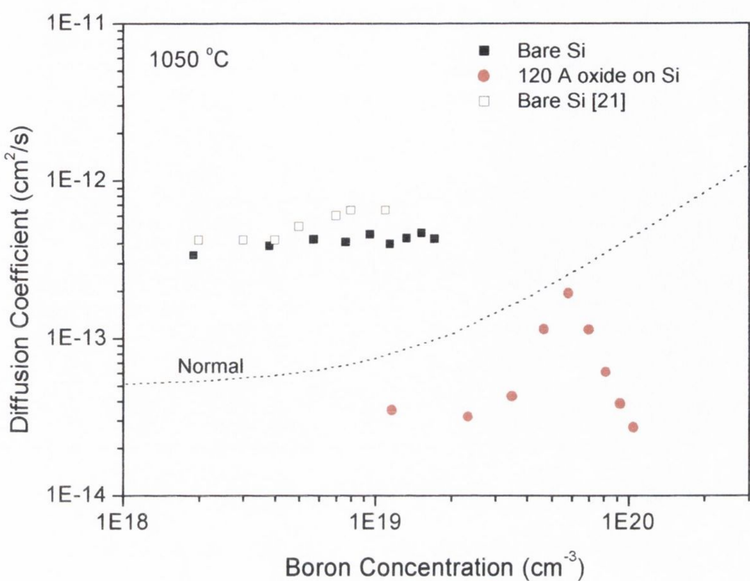


Figure 5.10. Dependence of boron diffusion coefficient on boron concentration in Si, evaluated from 16 s proximity RTD into bare Si at 1050 °C, and from a 5 s RTD through a 120 Å oxide on Si at 1050 °C. The dotted curve represents the normal diffusion coefficient.

However, if we assume that *Cs* has saturated and reached a constant value for the 16 s diffusion then the diffusion coefficient can be evaluated. Similarly, we will assume that the surface concentration of profiles that are formed by diffusion through 120 Å oxides is fixed at the solid solubility value of 1.2×10^{20} cm^{-3}. Figure 5.10 shows the dependence of boron diffusion coefficient on boron concentration, evaluated from the 16 s diffusion profile in Figure 5.7 and from the 5 s profile in Figure 5.9. The dotted curve represents the boron concentration dependence of the normal diffusion coefficient, and is represented by the following equation [28]:

$$D = D_i (N/2n_i)\left\{1 + \sqrt{1 + 4n_i^2/N^2}\right\} \quad (5.3)$$

where D_i is the intrinsic diffusion coefficient, n_i is the intrinsic carrier concentration and N is boron concentration. At 1050 °C, $D_i = 4.9 \times 10^{-14}$ cm^2/s [28] and $n_i = 1.07 \times 10^{19}$ cm^{-3} [29]. The normal boron diffusion coefficient is constant for low boron concentration (the intrinsic region) and increases with boron concentration in the extrinsic region, i.e., for boron concentration roughly larger than n_i. The increase in diffusion coefficient with an

increase in boron concentration is attributed to an increase in vacancy concentration [28, 30]. The behaviour of the diffusion coefficients evaluated in this work for RTD into bare Si (black squares in Fig. 5.10) resembles the boron concentration dependence of the normal diffusion coefficient. However, the diffusion coefficient is larger than the normal diffusion coefficient. This proves that enhanced diffusion takes place during proximity RTD from B153 BSG into bare Si. Miyake also observed enhanced diffusion during RTD from atmospheric pressure chemical vapour deposited (APCVD) BSG. Miyakes results are represented by the open squares in Figure 5.10, and are comparable with the diffusion coefficients evaluated in this work. Two reactions occur at the Si surface during proximity RTD of bare Si; i) $Si + O_2 = SiO_2$, since the RTD is performed in an oxidising ambient of 25% O_2 : 75% N_2, and ii) $B_2O_3 + Si = B + SiO_2$. Therefore, the enhanced diffusion may be caused by a combination of OED (Section 5.2.3) and BSG-enhanced-diffusion (Section 5.3), both of which enhance B diffusion by injecting Si interstitials into the silicon lattice during the formation of SiO_2.

Figure 5.10 also shows the diffusion coefficients for proximity RTD from B153 BSG through a 120 Å oxide on silicon. This behaviour does not resemble the boron concentration dependence of the diffusion coefficients into bare silicon. The diffusion coefficients are ≤ the normal diffusion coefficients. This proves that there is no enhanced diffusion during proximity RTD of B through a thin oxide layer. Si interstitials are still generated at the Si interface by the formation of SiO_2. However, the generated Si interstitials are reabsorbed into the SiO_2, where they are further oxidised to SiO_2 [31, 32]. Dunham *et al.* [32] estimated that the diffusion flux of interstitials into the Si substrate was negligible. Therefore, the level of oxidation-induced interstitials in the bulk is insufficient to enhance B diffusion. It is not clear at present why the diffusion coefficient for B concentration < 5×10^{19} cm^{-3} are smaller than the normal diffusion coefficients.

Initially, in Figure 5.10, the diffusion coefficient for diffusion through an oxide is approximately constant. Then, for boron concentration > 3×10^{19} cm^{-3}, the diffusion coefficient increases with increasing B concentration. The OED generation rate of Si interstitials at the Si/SiO_2 interface is known to increase under extrinsic conditions [30] ($N > n_i$, $n_i = 1.17 \times 10^{19}$ cm^{-3} at 1050 °C). If the generation rate increases significantly, then the flux of interstitials into the Si bulk may increase to a level that is sufficient to

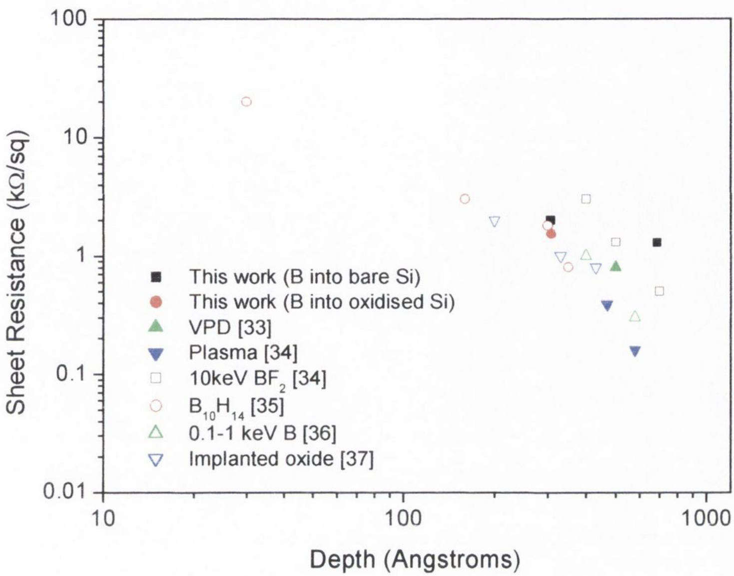


Figure 5.11. Boron depth at 10^{18} cm^{-3} versus sheet resistance for various shallow boron junction formation techniques.

increase B diffusion. This could explain why the diffusion coefficient increases for higher B concentrations.

Next however, the diffusion coefficients decrease unexpectedly with further increases in B concentration. When the solid solubility of an impurity in silicon is exceeded, precipitation of the impurity will occur. In this case, the solid solubility of B in Si ($1.2x10^{20}$ cm^{-3} at 1050 °C) has not been exceeded. However, excess interstitials may induce clustering of substitutional B at concentrations below the solid solubility limit of B in Si. This observation has been reported in ion implanted Si, where excess interstitials are generated by implantation damage [15-17]. If this sub-solid-solubility clustering does occur, then the amount of substitutional B that is available for diffusion is reduced at high B and Si interstitial concentrations, thus retarding B diffusion, and causing the diffusion coefficient to decrease at high boron concentrations.

5.4.3 Comparison with other Fabrication Techniques

Finally, a comparison of proximity RTD from B153 BSG with other shallow junction fabrication techniques is presented in Figure 5.11. The sheet resistance and B profile

depth combinations achieved with proximity RTD are comparable with the best results obtained with other techniques.

5.5 Conclusions

Bare wafers were doped by proximity RTD of boron from B153 BSG layers. The sheet resistance decreases with increasing RTD temperature and duration. The surface concentration increases with RTD duration, and eventually saturates for longer diffusion times. Initially, the silicon is not effectively doped due to a delay in transferring B_2O_3 from the BSG to the surface of the silicon wafer. The doping efficiency improves with time as the dopant supply increases at the silicon surface. At higher RTD temperatures, higher vapour pressure of B_2O_3 improves mass transport in the gas phase, thus improving doping efficiency. Some wafers were HF-passivated to remove the native oxide from the silicon wafer prior to RTD. It was determined that the native oxide did not influence the boron dopant profile. Boron profiles with depths of 686 Å at 10^{18} cm^{-3} and surface concentration of $1.5x10^{19}$ cm^{-3} were formed by a 5 s RTD at 1050 °C.

A silicon sample with 120 Å oxide layer on the surface was doped for 5 s at 1050°C by proximity RTD. The B profile had a depth of 308 Å at 10^{18} cm^{-3}, and a surface concentration of $1.2x10^{20}$ cm^{-3}. It was postulated that boron weakened the oxide bond structure, allowing the B_2O_3 to enter the SiO_2 more easily, and diffuse through it rapidly. It was also suggested that boron segregation in the oxide caused the silicon surface to be doped to solid solubility.

Boltzmann-Matano analysis was used to determine that the boron diffusion was enhanced during RTD into bare silicon wafers due to the injection of silicon interstitials into the lattice. Enhanced diffusion was not observed during RTD into the oxidised wafer, since the oxide layer traps the generated interstitials. Retarded diffusion at high boron concentration may be caused by interstitial-induced boron clustering.

The sheet resistance and B profile depth combinations achieved with proximity RTD are comparable with the best results obtained with other techniques.

For 0.1 μm technology, the Semiconductor Industry Association [38] project that depths < 400 Å with $Cs = 1x10^{20}\ cm^{-3}$ will be required to prevent short-channel effects. The boron profile formed in the oxidised sample during proximity RTD satisfies these requirements. The next chapter will integrate this shallow junction into *pn* junction diodes for device characterisation.

References

[1] D. Tsoukalas and P. Chenevier, *J. Appl. Phys.*, **66**, 1858, (1989).

[2] D.A. Antoniadis and I. Moskowitz, *J. Appl. Phys.*, **53**, 9214, (1982).

[3] P.A. Packan and J.D. Plummer, *J. Appl. Phys.*, **68**, 4327, (1990).

[4] P. Fahey, G. Barbuscia, M. Moslehi and R.W. Dutton, *Appl. Phys. Lett.*, **46**, 784, (1985).

[5] D.J. Roth and J.D. Plummer, *J. Electrochem. Soc.*, **141**, 1074, (1994).

[6] G. Masetti, S. Solmi and G. Soncini, *Solid State Electron.*, **16**, 1419, (1973)

[7] R.B. Fair, *J. Electrochem. Soc.*, **122**, 800, (1975).

[8] S.M. Hu, *J. Appl. Phys.*, **45**, 1567, (1974).

[9] T.Y. Tan and U Gosele, *Appl. Phys. Lett.*, **39**, 86 (1981).

[10] W.A. Tiller, *J. Electrochem. Soc.*, **128**, 689, (1981).

[11] Y. Hayafuji, K. Kajiwara, S. Usui, *J. Appl. Phys.*, **53**, 8639, (1982).

[12] S. Mizuo and H. Higuchi, *J. Electrochem. Soc.*, **130**, 1942, (1983).

[13] P.A. Stolk, H.-J. Gossman, D.J. Eaglesham, D.C. Jacobson, H.S. Luftman and J.M. Poate, *Mat. Res. Soc. Symp. Proc.*, **354**, 307, (1995).

[14] A. Agarwal, H.-J. Gossman, D.J. Eaglesham, L. Pelaz, S.B. Herner, D.C. Jacobson, T.E. Haynes and R. Simonton, *Materials in Semiconductor Processing*, **1**, 17, (1998).

[15] P.A. Stolk, H.-J. Gossman, D.J. Eaglesham, D.C. Jacobson, C.S. Rafferty, G.H. Gilmer, M. Jaraiz and J.M. Poate, J. Appl. Phys, 81, 6031, (1997).

[16] M. Orlowski, *Tech, Digest. IEDM*, 505, (1997).

[17] E.J.H. Collart, G. de Cock, A.J. Murrell and M.A. Foad, *Mat. Res. Soc. Symp. Proc.*,**525**, 227, (1998).

[18] R.B. Fair, *J. Vac. Sci. Technol.*, **A4**, 926, 1986.

[19] R.B. Fair in "*Rapid Thermal Processing*", Ed. R.B. Fair, Academic Press, (1993), p.169.

[20] A. Sultan, S. Batra, G.E. Lux and S. Banerjee, *Materials Science and Eng.*, **B32**, 25, (1995).

[21] M. Miyake, *J. Electrochem. Soc.*, **138**, 3031, (1991).

[22] M. Miyake, *J. Electrochem. Soc.*, **145**, 2534, (1998).

[23] Y. Ishikawa, K. Yamauchi and I. Nakamichi, *Jpn. J. Appl. Phys.*, **28**, L1319, (1989).

[24] R.B. Fair, *J. Electrochem. Soc.*, **137**, 667, (1990)

[25] L. Boltzmann, *Ann. Physik.*, 53, 960, (1894).

[26] C. Matano, *Japan. Phys.*, 8, 109, (1933).

[27] *Atomic Diffusion in III-V Semiconductors*, B. Tuck, Hilger, Bristol, (1988), p. 30.

[28] R.B. Fair, *J. Electrochem. Soc.*, **122**, 800, (1975).

[29] F.J. Morin and J.P. Maita, *Phys. Rev.*, **96**, 28, (1954).

[30] M. Miyake, *J. Appl. Phys.*, **57**, 1861, (1985).

[31] P.M Fahey, P.B. Griffin and J.D. Plummer, *Rev. Mod. Phys.*, **61**, 289, (1989).

[32] S.T. Dunham and J.D. Plummer, *J. Appl. Phys.*, **59**, 2551, (1986).

[33] Y. Kiyota, T. Nakamura and T. Inada, *IEICE Trans. Electron.*, **E77-C**, 362, (1994).

[34] M. Takase, K. Yamashita, A. Hori and B. Mizuno, *IEDM Tech. Dig.*, 475, (1997).

[35] K. Goto, J. Matsuo, Y. Tada, T. Tanaka, Y. Momiyama, T. Sugii and I. Yamada, *IEDM Tech. Dig.*, 471, (1991).

[36] E.J.H. Collart, K. Weemers, D.J. Gravesteijn and J.G.M. van Berkum, *J. Vac. Sci. Technol.*, **B16**, 280, (1998).

[37] J. Schmitz, M. van Gestel, P.A. Stolk, Y.V. Ponomarev, F. Roozeboom, J.G.M. van Berkum, P.C. Zalm and P.H. Woerlee, *IEDM Tech. Dig.*, (1997).

[38] *National Technology Roadmap for Semiconductors*, Semiconductor Industry Association, (1997), p.74.

Chapter 6

Electrical Characterisation of *p-n* Junction Diodes

6.1 Introduction

In the preceding chapters, proximity RTD was developed as a doping technique. Chapters 3 and 4 concentrated on the dopant sources and Chapter 5 focused on the doped product wafer. In this chapter, proximity RTD is integrated into a PMOS process and *p-n* junction diodes are fabricated.

The current-voltage (*IV*) characteristics of the *p-n* junction diodes are discussed. When a forward bias is applied to a *p-n* junction, the current increases rapidly as the voltage increases. However, when a reverse bias is applied, virtually no current flows initially. As the reverse bias is increased, the current remains very small until a critical voltage is reached at which point the current suddenly increases. This sudden increase in current is referred to as the junction breakdown.

The ideal diode equation describes certain real diodes; however, several conditions exist where the ideal diode fails to adequately represent physical devices. Forward bias deviations occur due to recombination in the depletion region, high level injection and series resistance in the bulk and contact regions. Reverse bias deviations occur due to generation of carriers in the depletion region and subsequent junction breakdown. These deviations from ideality will be discussed in this chapter, and the quality of the junctions fabricated using proximity RTD will be determined.

6.2 Fabrication of *p-n* Junction Diodes

In order to investigate the properties of the shallow boron junctions, 3 different types of *p-n* junction diodes were designed and fabricated, Figure 6.1. The Surround Diode (SD), Fig. 6.1a, consists of a shallow boron junction, surrounded by a deeper boron junction. This diode was designed in an attempt to examine the planar shallow boron junction while also avoiding issues concerned with shallow boron junction edge effects (Section 6.5.1).

Open Diode type 2 (OD2), Fig. 6.1c, was designed to complement the SDs, and enable the shallow boron junction edge effects to be analysed. Ideally, the metal contact should be made to the shallow junctions in these diodes, however, it is extremely difficult to prevent aluminium spiking through very shallow junctions during metallisation. Therefore, it was necessary to form a deeper boron junction at the centre of each diode to allow aluminium contacts to be made to the devices.

Open Diodes type 1 (OD1), Fig. 6.1b, were designed in an attempt to determine the behaviour of shallow boron junctions when they are integrated into a PMOS transistor. The shape of the OD diode is comparable with a PMOS transistor that is missing a gate structure.

A four-mask process was used to produce the SD, OD2 and OD1 diodes. A detailed mask set for the diodes is given in Appendix A. Czochralski-grown 4 inch *n*-type, <100> oriented silicon wafers, with a background concentration of $3x10^{14}$ cm^{-3} were used throughout this study. The substrate was wet oxidised at 1150 °C to produce a masking oxide layer thickness of approximately 0.4 μm. This oxide was patterned and a *p*-type diffusion from a boron nitride disk was carried out at 1000 °C for 25 minutes, followed by a 60 minute drive-in at 1000 °C, to form *deep p-type, p^d*, junctions. These deep *p*-type junctions have a depth of 1.55 μm at 10^{18} cm^{-3}. A second mask was used to open the windows for the very shallow proximity RTD *p*-type regions. In order to minimise process variations, wafers D4 and D5 were then dry oxidised at 900 °C for 5 minutes to produce a 120 Å thick oxide layer in the diffusion windows. This thin oxide layer was subsequently removed from wafer D5, to enable comparisons between junctions formed

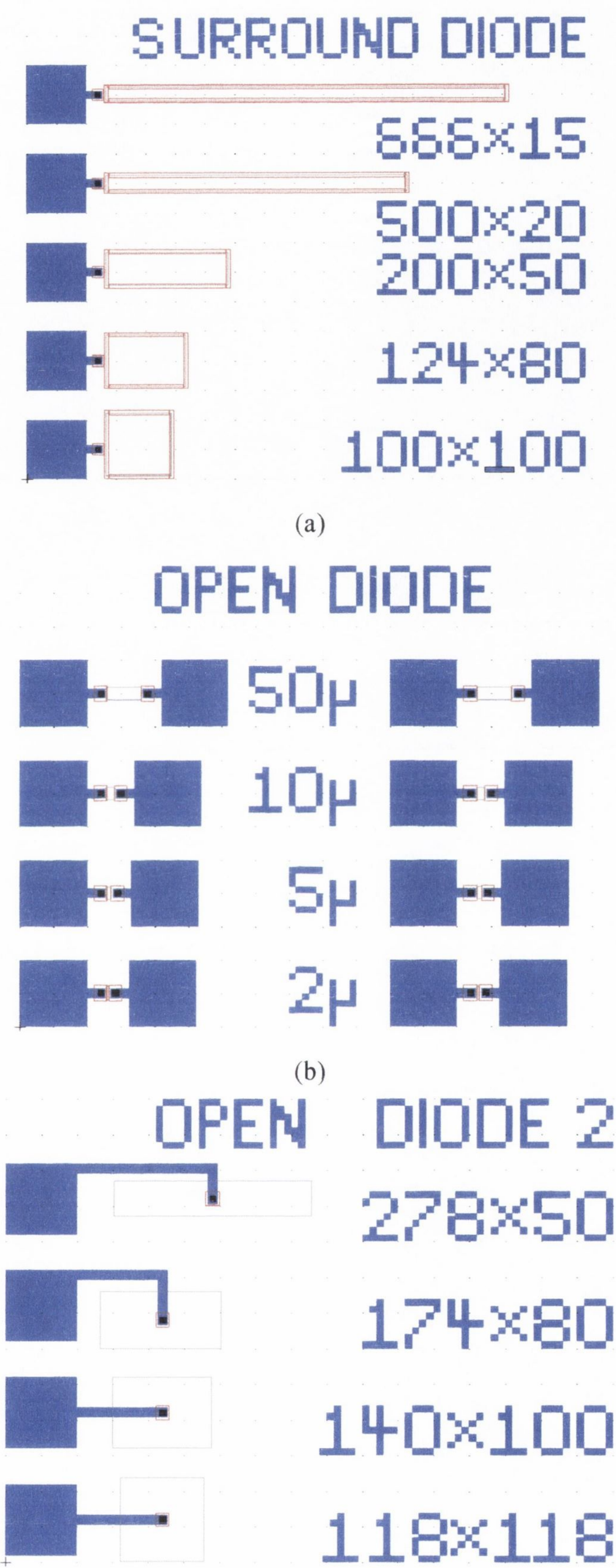


Figure 6.1. Mask Layout for Surround Diodes (SD), Open Diodes type 1 (OD1), and Open Diodes type 2 (OD2).

by RTD of boron into bare silicon (wafer D5) and RTD through a 120 Å oxide layer (wafer D4). Proximity RTD was then performed at 1050 °C for 5 s to form very *shallow p-type*, p^s, junctions. The shallow *p*-type junction, p^s, in wafers D5 and D4 have depths of 686 Å and 308 Å at $10^{18}\,cm^{-3}$, respectively (Chapter 5). An LPCVD oxide was deposited from a TEOS source at 720 °C. The device manufacture was completed by opening contact windows in this oxide layer to the deep *p*-type, p^d, junctions, followed by aluminium metallisation. The devices were annealed in forming gas *prior* to aluminium metallisation to avoid aluminium spiking through the shallow junctions. The process is described in more detail in Appendix B.

6.3 The Ideal Diode Equation

The ideal diode equation for an abrupt one-sided p^+-n junction is given as:

$$I = I_D\left(e^{qV_A/kT} - 1\right) \qquad (6.1)$$

$$I_D = A_J q \frac{n_i^2}{N_D}\sqrt{\frac{D_p}{\tau_p}} \qquad (6.1a)$$

The ideal diode equation only takes the diffusion current into consideration, and assumes that no generation-recombination occurs in the depletion region. However, in order to accurately describe a real diode, the generation-recombination current must be taken into consideration.

Under reverse bias conditions, carrier concentrations in the depletion region fall far below their equilibrium concentrations. Electron-hole pairs are then generated in the depletion region. The current due to generation, I_{GEN}, in the depletion region is:

$$I_{GEN} = A_J q n_i \frac{W}{\tau_g} \qquad (6.2)$$

The total reverse current for a p^+-n junction, for $V>3kT/q$, can be approximated by the sum of the Eqns. 6.1 and 6.2:

$$I = A_J q \frac{n_i^2}{N_D}\sqrt{\frac{D_p}{\tau_p}} + A_J q n_i \frac{W}{\tau_g} \tag{6.3}$$

Under forward bias conditions, carrier concentrations in the depletion region exceed their equilibrium concentrations, and electron-hole pairs recombine. The recombination current, I_{REC}, is given by:

$$I_{REC} = I_R e^{qV_A/2kT} \tag{6.4}$$

$$I_R = \frac{1}{2} A_J n_i q \left\{ \frac{W}{\tau_0} + S_o \frac{A_S}{A_J} \right\} \tag{6.4a}$$

The first term in the brackets (Eq. 6.4a) relates to recombination in the bulk, while the second term in the brackets relates to surface recombination. The total forward current for a p^+-n junction, for $V>3kT/q$, can be approximated by the sum of the Eqns. 6.1 and 6.4:

$$I = I_D e^{qV_A/kT} + I_R e^{qV_A/2kt} \tag{6.5}$$

or as a single expression:

$$I = I_o e^{qV_A/nkT} \tag{6.6}$$

where $I_o = I_D$ when the diode current is dominated by the ideal diffusion process, and $I_o = I_R$ when the current is dominated by the depletion region recombination process. The parameter n, known as the *ideality factor*, has a value of 1 for the diffusion current and is 2 for the recombination current. When the two currents are comparable, n lies

between 1 and 2. The ideality factor, *n*, and the current I_0 can be determined from the slope and intercept of a forward log *I* versus V_A plot.

Minority carriers that are injected into either side of a *p-n* junction diffuse into the neutral material and recombine with the majority carriers. The minority carrier current decreases exponentially with distance into the neutral region. The ideal diode equation assumes that the width of the *p*- and *n*-type regions is very long compared to the minority carrier diffusion lengths [1]. Thus several diffusion lengths away from the junction, the minority carrier current is small and most of the current is by the majority carriers.

This assumption is limited for very shallow junctions. When the junction depth is comparable to the minority carrier diffusion length, the minority carrier current will not decay to a negligible value, and will represent a significant percentage of the total current. The ideal diode equation does not take this effect into consideration. SD, OD1 and OD2 have very shallow *p*-type regions, and the minority carrier current may not decay exponentially to zero. These diodes are very novel devices and may not be accurately modelled by the ideal diode equation. Since, these types of device have not been reported in the literature to date, there is no other model available for discussion. Therefore, the ideal diode equation will be used in the following sections to get an idea of the behaviour of the SD, OD1 and OD2 diodes.

6.4 Forward IV Characteristics

The forward *IV* characteristics of all the *p-n* junction diodes on wafers D4 and D5 were measured using a HP 4155A Parameter Analyser. The forward current was measured as the voltage was increased in 0.01 mV increments in the range 0-1 V. Table 6.1 shows the ideality factor, *n*, and I_0 for each of the diodes.

6.4.1 Surround Diodes (SD)

Figure 6.2 shows the forward *IV* characteristics for all of the surround diodes on wafer D5 from 0.2-0.3 V. It can be observed that the current varies between each diode. A

Table 6.1. Ideality factor, *n*, and I_o for wafers D5 and D4.

Diode	Wafer D5		Wafer D4	
	Ideality, *n*	I_0 (A)	Ideality, *n*	I_0 (A)
SD 666μm x 15μm	1.219	8.69e-13	1.124	3.53e-13
SD 500μm x 20μm	1.210	7.11e-13	1.117	2.64e-13
SD 200μm x 50μm	1.188	4.66e-13	1.132	1.98e-13
SD 124μm x 80μm	1.182	3.97e-13	1.122	1.82e-13
SD 100μm x 100μm	1.194	4.23e-13	1.121	1.90e-13
OD 50μm	1.102	1.18e-13	1.083	6.90e-14
OD 10μm	1.097	1.05e-13	1.082	6.45e-14
OD 5μm	1.089	8.46e-14	1.075	6.42e-14
OD 2μm	1.084	7.56e-14	1.075	6.41e-14
OD2 278μm x 50μm	1.074	1.92e-13	1.051	1.68e-13
OD2 174μm x 80μm	1.073	1.92e-13	1.049	1.63e-13
OD2 140μm x 100μm	1.075	1.94e-13	1.052	1.69e-13
OD2 118μm x 118μm	1.072	1.89e-13	1.051	1.69e-13

graph of current (at V = 0.25 V) against area is shown in Figure 6.3 for wafers D5 and D4. While the area of each diode is the same there is a significant increase in current when the width of the diode is reduced to less than 50 μm. A similar trend in current is observed for both diffusion into bare silicon (Wafer D5) and diffusion through the thin oxide (Wafer D4). These results provide evidence of current crowding in the devices. A

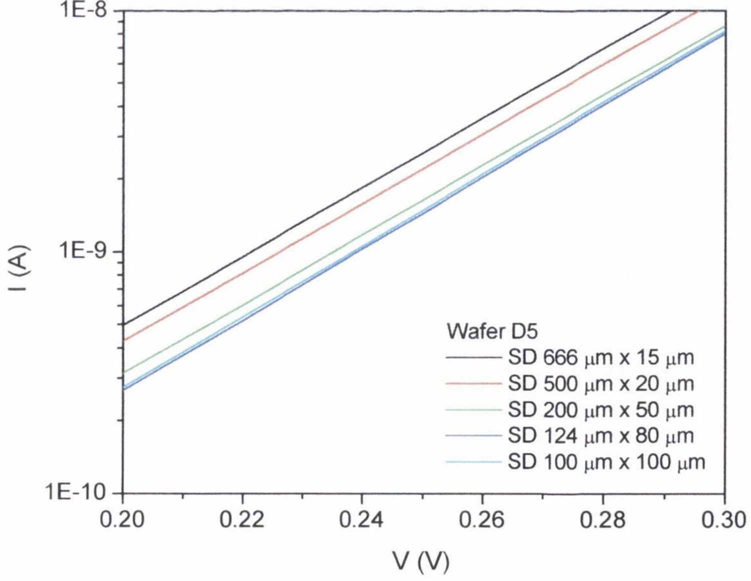
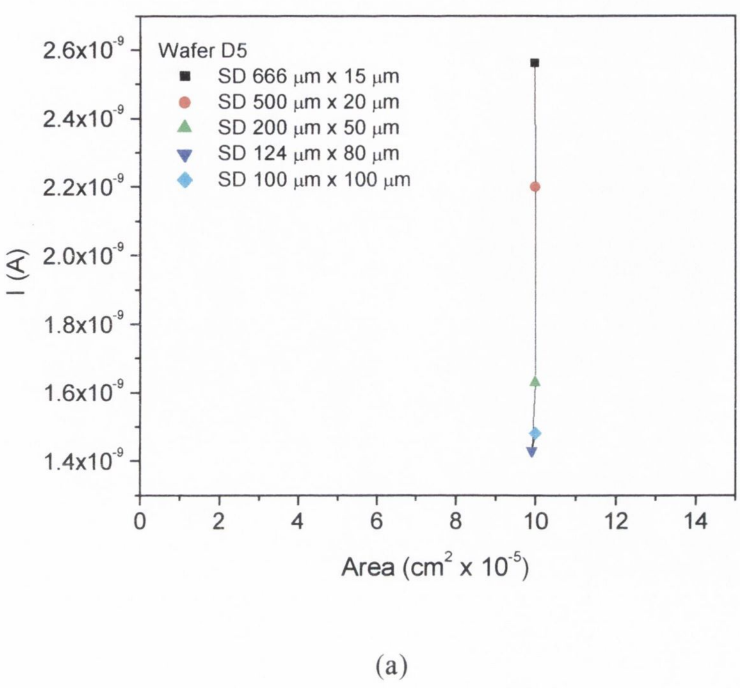
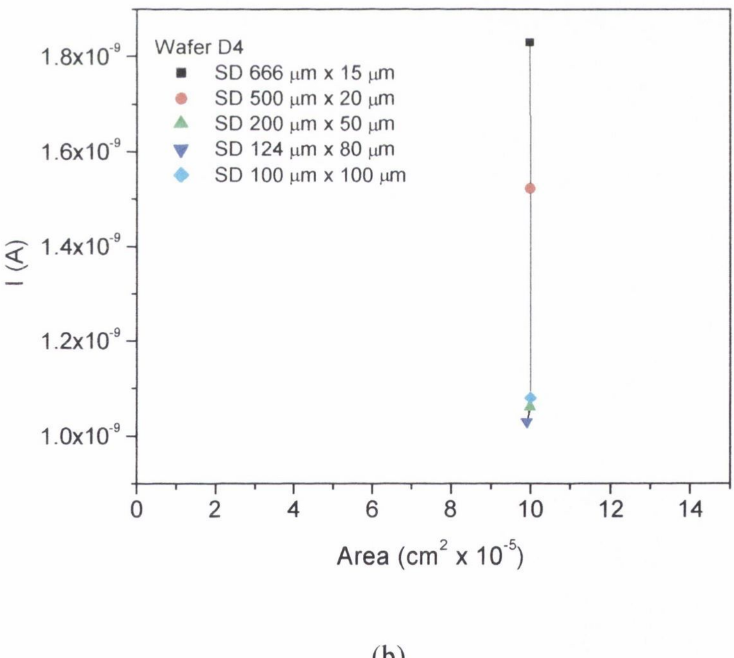


Figure 6.2. Forward IV characteristics of surround diodes on wafer D5.



(a)



(b)

Figure 6.3. Current v's total p-type area of surround diodes on (a) wafer D5 and (b) wafer D4.

cross-section of a surround diode is shown in Figure 6.4a. The current density through the p^s region will decrease away from the p^d contact due to a voltage drop across the lateral series resistance. Therefore, the effective area of the diode is reduced from its

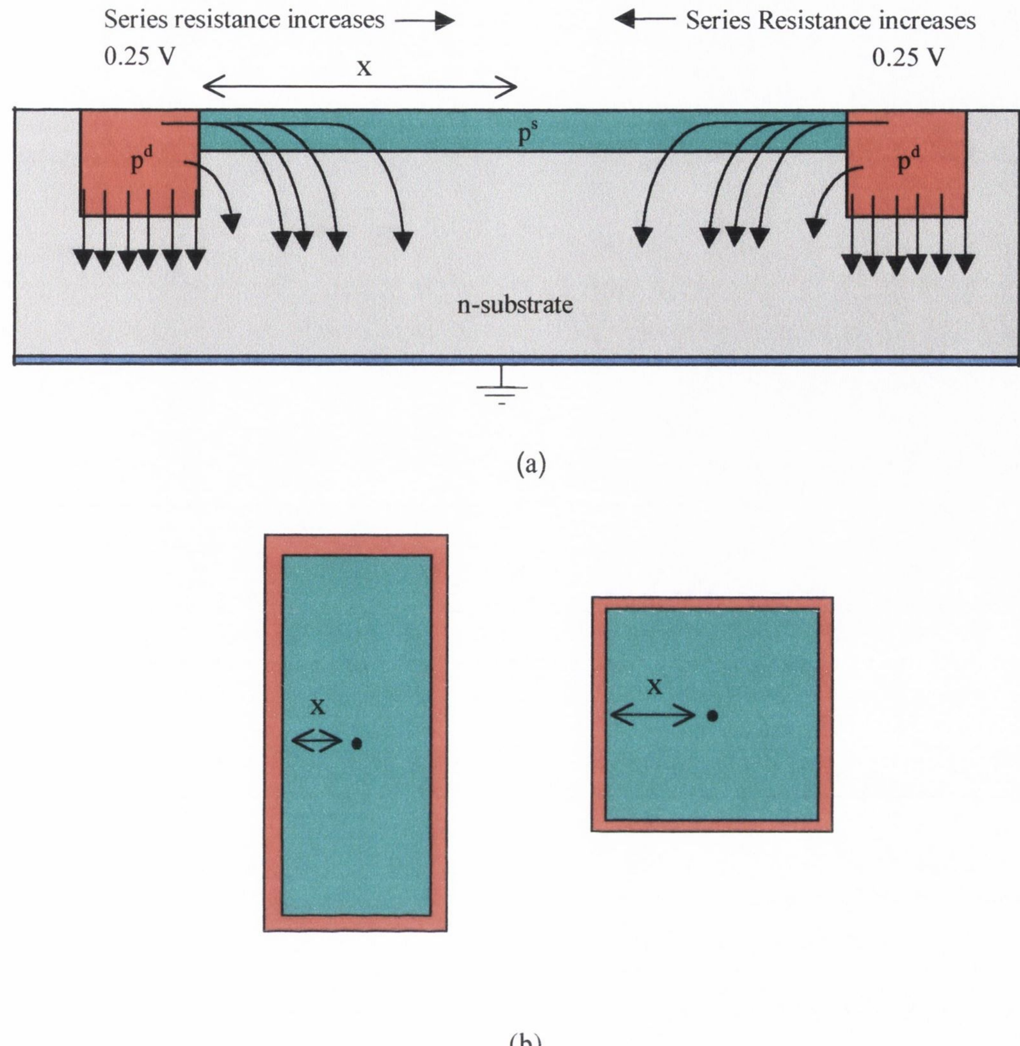


Figure 6.4. (a) Current crowding in a surround diode, (b) As width x increases, the current crowding increases in the diode.

geometric area. For a constant geometric area, the current will decrease as the width increases, assuming that the voltage drop along the p^d region is negligible, Figure 6.4b. Surround diode SD 666 µm x 15 µm has the narrowest width, therefore less current crowding will occur across the p^s region, allowing a higher current to flow through the diode. Figure 6.3 shows that SD 666 µm x 15 µm does have the highest current flowing through it. The current also decreases in the other diodes as the width increases, and reaches a relatively constant minimum value for SD's with widths > 50 µm in wafer D5 (Fig. 6.3a) and widths > 20 µm in wafer D4 (Fig. 6.3b). Since the p^s junction depths are shallower on wafer D4 current crowding is more severe, which limits the current flowing across the p^s region to a higher extent than in wafer D5.

Table 6.1 shows the ideality factor, *n*, and I_0 for each of the SD diodes. I_0 and *n* will increase ($n > 1$) if there is recombination in the depletion region. The recombination current, I_{REC}, varies with junction area, A_J, depletion region width, *W*, area of the depletion width at the silicon surface, A_S, and surface recombination velocity, S_0 (Eq. 6.4a). Since, the doping levels of all of the SD's are equal, there will be no variation in *W*. As already discussed, the effective junction area, A_J, of the diodes decreases due to current crowding. Subsequently, the decrease in I_0 in both wafers between SD 666 μm x 15 μm and SD 100 μm x 100 μm is due to a decrease in the effective area A_J. The I_0 values of the SD diodes on wafer D4 are smaller than the I_0 values of the SD diodes on wafer D5. As mentioned above, the p^s junction depths are shallower on wafer D4 and current crowding is more severe; resulting in a significant reduction in the I_0 values of all of the SD diodes, even that of SD 666 μm x 15 μm.

The area of the depletion width at the silicon surface, A_S, is proportional to the perimeter length of the diode. The ideality factor between SD 666 μm x 15μm and SD 100 μm x 100 μm on wafer D5 decreases as the perimeter length and hence, A_S decreases. (The decrease in perimeter also contributes to a decrease in I_0 in both wafers, although it is not the main influencing factor.) The same trend is not observed for the SD's on wafer D4. Thus, the surface recombination velocity, S_0, is dominant for wafer D5 but not for wafer D4. Different values of fixed charge in the oxide surrounding the diodes would give different depletion widths at the oxide-silicon interface. This could be due to slightly different processing between the two wafers.

6.4.2 Open Diodes Type 1 (OD1)

The *IV* characteristics of each of the OD1 diodes were measured. The p^d area is the same for all of the diodes. The p^s area varies between the diodes, with OD1 50 μm having the maximum area and OD1 2 μm having the minimum area. Figure 6.5 shows the dependence of the current flowing through the diodes on diode area. The increase in current between OD1 2 μm, OD1 5 μm and OD1 10 μm is almost directly proportional to the increase in area between the diodes. However, there is a relatively small increase in current between OD1 10 μm and OD1 50 μm even though the area of OD1 50 μm is

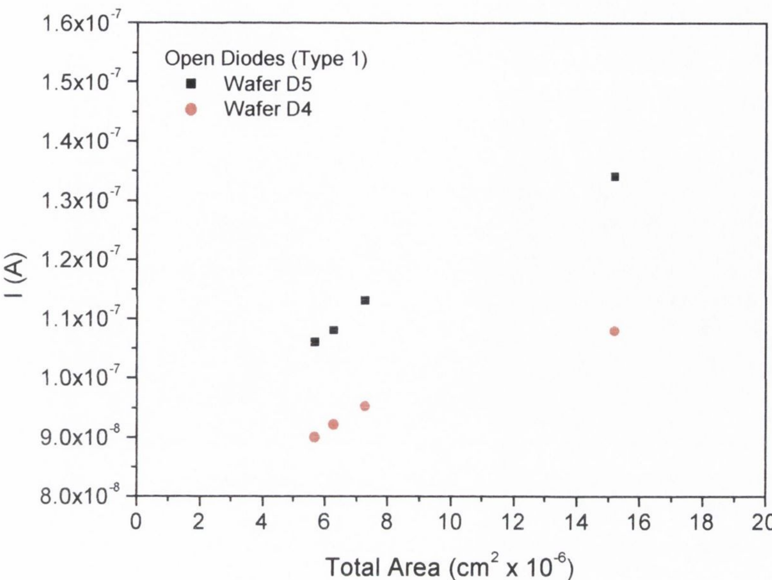


Figure 6.5. Current v's total area of OD1 diodes in wafers D4 and D5.

double the area of OD1 10 μm. The current flowing through OD1 50 μm is limited by current crowding in the device.

As discussed in the previous section, I_0 and n increase as I_{REC} increases. The I_0 values decrease in the OD1 diodes with decreasing junction area, A_J. This trend is observed in both wafers; with I_0 values on wafer D4 being less than I_0 values on wafer D5 due to shallower junctions and more severe current crowding on wafer D4. Also, as in the previous section, n decreases as the diode perimeter length, and hence A_S, decreases.

6.4.3 Open Diode Type 2 (OD2)

Figure 6.6 shows that the current flowing through each of the OD2 diodes in wafer D5 is equal. The same result was observed for all of the OD2 diodes in wafer D4.

In Section 6.4.1 it was observed, that due to current crowding, the current decreased to a minimum in SD's with widths > 50 μm ($2x$ in Figure 6.4) in wafer D5. In other words, the current density in the p^s region is negligible at approximately 25 μm (x) from the p^d contact. Similarly, in OD2 diodes, the current density in the p^s region is negligible at

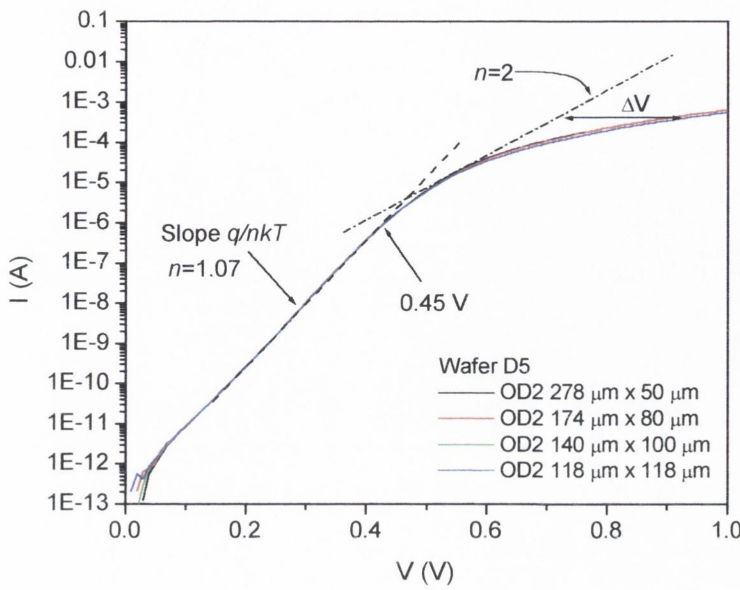


Figure 6.6. Forward bias IV characteristics for OD2 diodes on wafer D5.

approximately 25 μm from the p^d contact. Hence, each OD2 diode will have an effective junction area around the p^d contact, with a radius of approximately 25 μm. Therefore, since each diode has the same effective junction area, and doping level, the forward *IV* characteristics will be almost identical, as observed in Figure 6.6.

Table 6.1 shows that the ideality factor and I_0 are essentially equal for each OD2 diode on wafer D4 and D5, with D4 having smaller I_0 values because of more severe current crowding. It is interesting to note that the OD2 diodes have the smallest ideality factors, when compared to all of the other SD and OD1 diodes. This may be because the OD2 diodes are not electrically active at the edge of the p^s region, which would mean that there would not be any surface recombination at the oxide/depletion interface.

6.4.4 High Level Injection and Series Resistance

The forward *IV* characteristics of all of the diodes on wafers D4 and D5 have one feature in common; the slope of log *I* versus *V* changes at approximately $V = 0.45$ V. Figure 6.6 shows the change in slope for the OD2 diodes on wafer D5. Firstly, the slope changes from $n = 1.07$ to $n = 2$. This change in slope is due to the onset of high level injection.

(i) High Level Injection

A *p-n* junction operates under high level injection when the injected minority carrier concentration on the low doped side of the junction becomes comparable to the majority carrier concentration. The net result of high level injection is an increase in recombination that changes the ideality factor to $n \cong 2$.

All of the diodes have p^+-*n* junctions, therefore the high level injection is on the *n*-side of the junction. The onset of high level injection can be determined when $p_n = N_D$, (where p_n is the concentration of holes in *n*-type material, similarly, p_p is the concentration of holes in *p*-material),

$$p_n = \frac{n_i^2}{N_D} e^{qV_A/kT} \tag{6.7}$$

At room temperature, in this case, the onset of high level injection occurs at $V = 0.5$ V. Subsequently, the slope will change at this voltage. This calculation approximately agrees with Figure 6.6, where the slope of the forward *IV* changes at $V = 0.45$ V.

Secondly, the slope of Figure 6.6 changes at $V = 0.6$ V from $n = 2$ to $n > 2$. This change in slope is due to series resistance across the device.

(ii) Series Resistance, R_s

In the derivation of the ideal diode equation, it is assumed that all of the applied voltage appears across the depletion region, and the bulk *n* and *p* regions are field free. No consideration is given to the possible voltage drop across the ohmic contacts. For most devices these are good assumptions at lower current levels. However, at large current levels the bulk resistance can produce a significant voltage drop and the applied voltage is greater than the voltage across the depletion region. Also, the metal-silicon contacts can behave like small resistors, adding to the voltage drop across the bulk regions.

The series resistance can be calculated by measuring the voltage difference (ΔV) between the ideal forward *IV* plot which includes high level injection ($n = 2$) and the measured

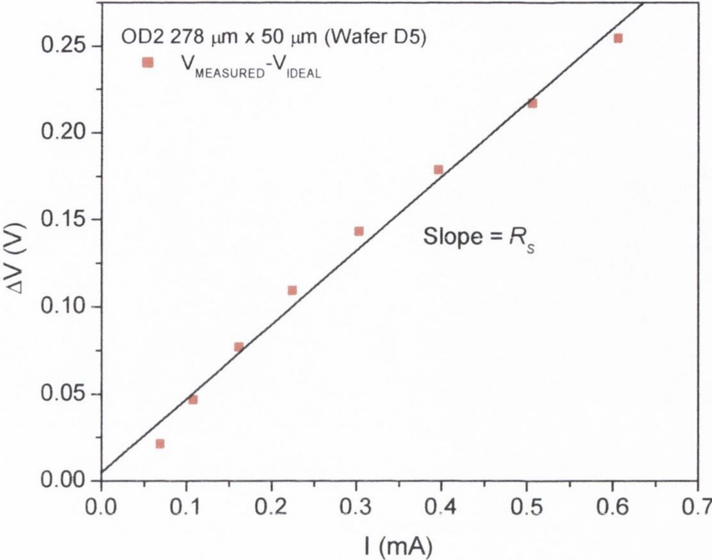


Figure 6.7. ΔV versus I for OD2 278 μm x 50 μm. The slope of the graph is the series resistance, R_s, of the diode.

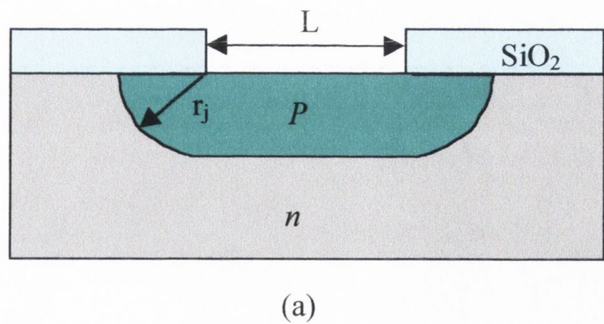
forward *IV*. If ΔV is plotted against current, R_s can be determined from the slope, Figure 6.7. A value of $R_s = 425\ \Omega$ is deduced from the slope of Figure 6.7.

6.5 Reverse IV Characteristics

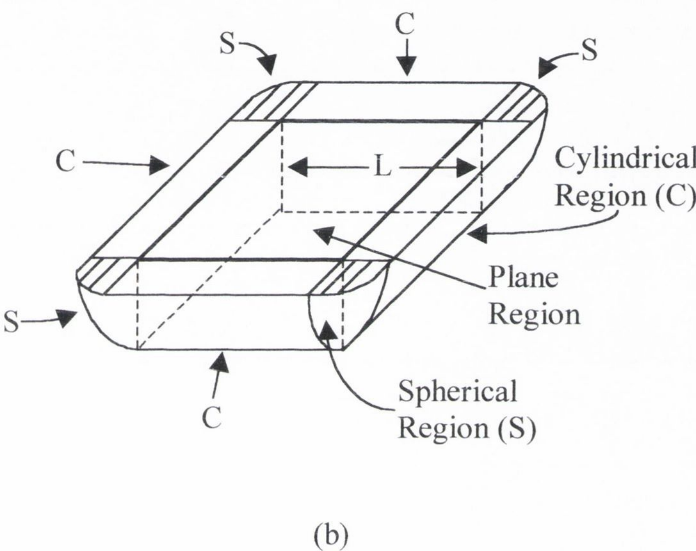
The reverse *IV* characteristics of all the *p-n* junction diodes on wafers D4 and D5 were measured using a HP 4155A Parameter Analyser. The reverse current was measured as the voltage was increased in 0.01 mV increments until the breakdown voltage was reached. All of the reverse *IV* characteristics were measured in total darkness to prevent photogenerated current.

6.5.1 Influence of Junction Curvature on Breakdown Voltage

When a *p-n* junction is formed by diffusion through a window in the SiO_2 layer, the impurities diffuse vertically into the silicon and laterally beneath the Si-SiO_2 interface, creating a junction with curved boundaries as shown in Figure 6.8a. Depending on the



(a)



(b)

Figure 6.8. (a) Planar diffusion process that forms junction curvature near the edge of the diffusion mask, where r_j is the radius of curvature. (b) Formation of the cylindrical and spherical regions by diffusion through a rectangular mask.

shape of the window in the SiO_2, the curved boundaries will be either spherical of cylindrical in shape, Figure 6.8b. In both cases, the electrical field intensity will be higher than in the plane region of the junction, causing avalanche breakdown to occur there at unexpectedly low voltages. The influence of the junction curvature, r_j, on the breakdown voltage of one-sided abrupt junctions in silicon is shown in Figure 6.9. The reduction in breakdown voltage is especially severe for a shallow junction with a small radius of curvature. It becomes less serious as the junction depth increases.

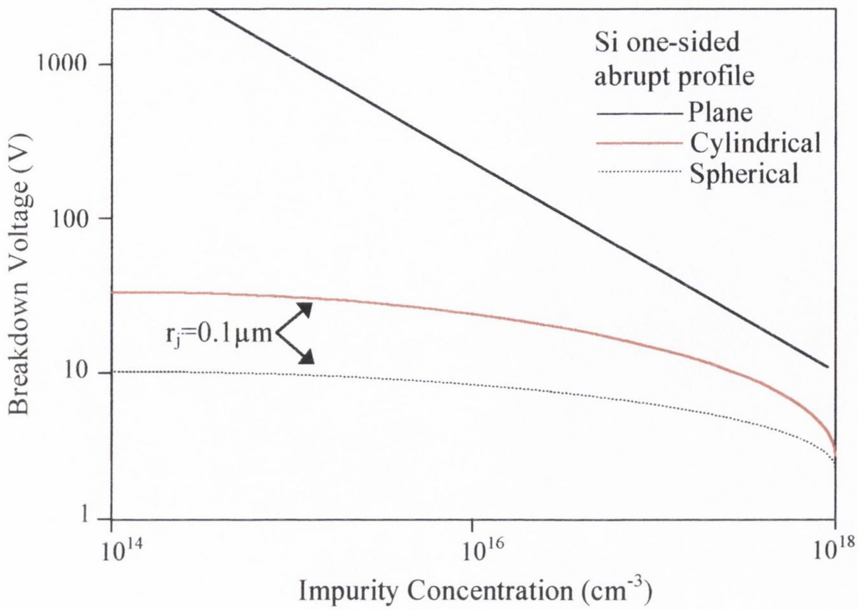


Figure 6.9. Breakdown voltage versus impurity concentration for one-sided abrupt doping profile with cylindrical and spherical junction geometries, where rj is the radius of curvature as indicated in Fig. 6.14 [2].

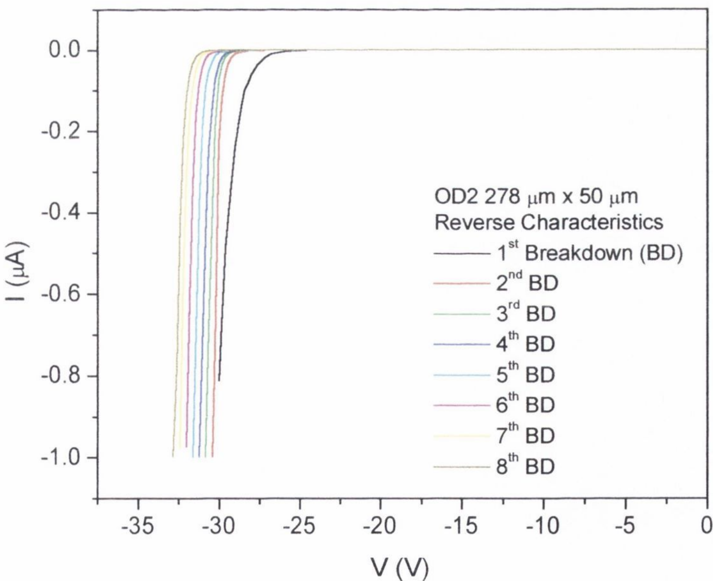


Figure 6.10. Breakdown voltages of OD2 278 μm x 50 μm.

6.5.2 Open Diodes Type 2 (OD2)

All of the OD2 diodes on wafer D5 display almost identical reverse *IV* characteristics. Figure 6.10 shows the reverse IV characteristics of OD2 278 μm x 50 μm on wafer D5. The reverse current increases gradually with bias until a critical voltage is reached at which the current increases abruptly with voltage. This critical voltage is called the breakdown voltage, $V_{BD.}$ Initially, the diode had a breakdown voltage of V_{BD} = -30 V. Following this initial breakdown, a reverse bias was applied to the junction again. The breakdown voltage increased to a higher value of V_{BD} = -30.4 V. Figure 6.10 shows that as the reverse breakdown is repeated, the breakdown voltage increases. This observation can be explained in the following way.

Initially, at thermal equilibrium, there is an accumulation of electrons at the oxide-substrate interface due to the trapped positive charges in the oxide, Fig. 6.11a. This accumulation of electrons results in a decrease in the depletion width, from W to W_1, at the oxide-substrate interface. If a sufficiently large reverse voltage is applied to the *p-n* junction, avalanche breakdown will occur. Following the application of this voltage, a large number of electron-hole pairs are produced in the junction depletion region by avalanche multiplication. Some of the electrons generated in the avalanching depletion region acquire sufficient energy to surmount the Si-SiO_2 energy barrier and are trapped in the oxide layer, as shown by label (1), Fig. 6.11b. Fringing electric fields may also cause electrons to be injected directly from the p-region into the oxide layer, label (2) in Figure 6.11b. This negative charge in the oxide causes strong inversion in the channel. The subsequent increase of the depletion width, from W_1 to W_2, reduces the electric field and increases the breakdown voltage of the diode, Fig. 6.11c. This process is called *p-type walkout*. If avalanche breakdown is continually repeated, the depletion width at the oxide-substrate interface will eventually saturate to the bulk depletion width W, Fig. 6.11d. P-type walkout was observed for all of the OD2, OD1 and SD diodes on wafers D4 and D5.

P-type walkout occurs at the oxide-substrate interface at the outer edge of the *p*-region. Hence, observation of *p*-type walkout is proof that the junction breaks down at the *p-n* substrate interface, which means that the entire p^s-region of the OD2 diodes is

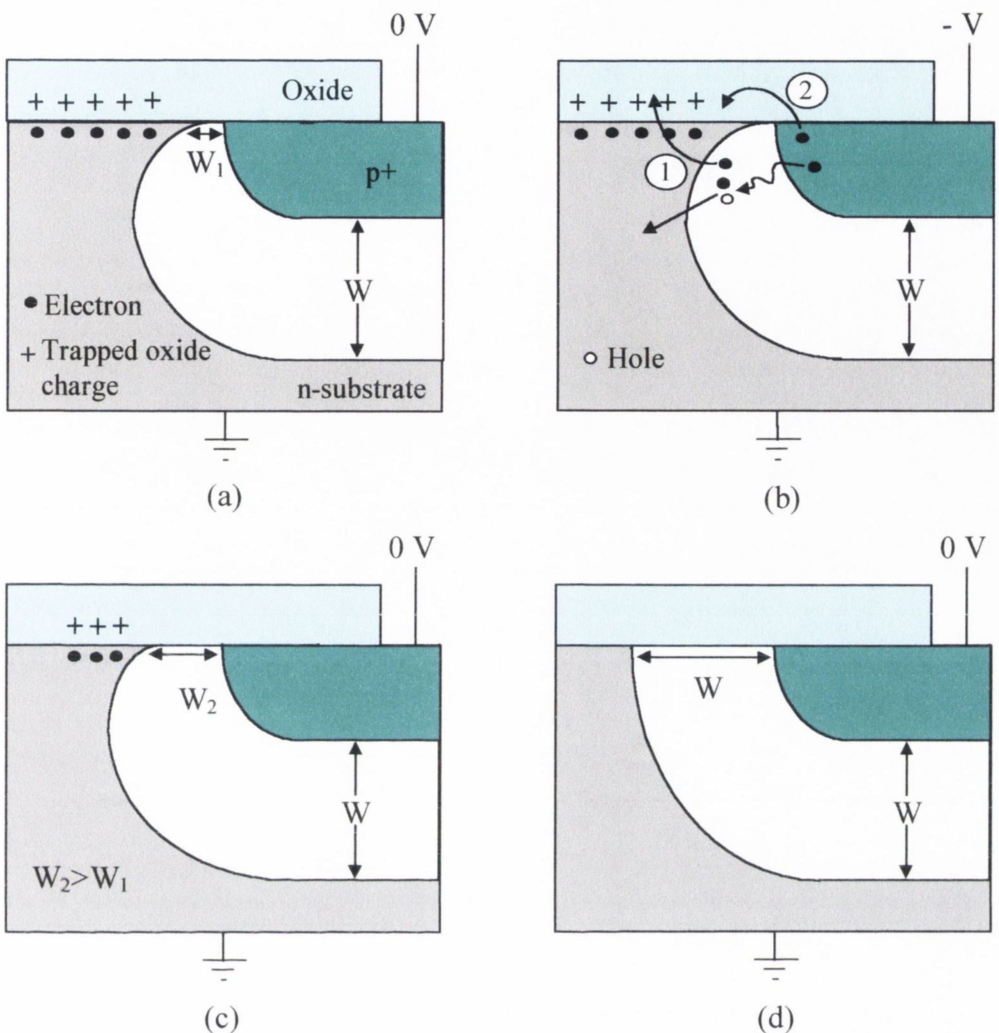


Figure 6.11. (a) Accumulation of electrons at the oxide-substrate interface reduces depletion width to W_1, $W_1 \ll W$ (b) Avalanche multiplication process injects an electron into the oxide (c) Injected electron causes depletion width to change to W_2, $W_2 > W_1$ (d) If avalanche breakdown is repeated, depletion width at oxide-substrate interface increases an saturates to W.

electrically active under reverse bias. Specifically, the OD2 diodes breakdown at the p^s corners, since the electric field intensity is highest at these spherical regions.

It is interesting to note that the first breakdown in Figure 6.10 has a different, more rounded shape than the other subsequent breakdown curves. The difference in breakdown may be due to the change in depletion width with *p*-type walkout. Alternatively, a variation in the fixed charge density in the oxide surrounding the diode could result in a variation in the depletion width around the diode. Following avalanche breakdown, the injected electrons would smooth out further variations in the fixed oxide

charge. However, it is more likely that either one of these possibilities would result in a gradual change in the shape of the breakdown curve instead of an abrupt change between the shape of the first breakdown and all of the other curves, as observed in Figure 6.10. A third possible explanation for the rounded shape of the first breakdown curve is that Zener breakdown may have occurred. Zener breakdown involves valence electrons tunnelling through the energy bandgap from the valence band to the conduction band. Since the probability of transmission of an electron through an energy barrier is a strong function of the thickness of the barrier, tunnelling is only significant in highly doped p^+-n^+ material in which the depletion region is very narrow and the electric fields are very high. In the case of Zener breakdown, the current rises slowly with voltage in the breakdown region, and a sharp breakdown point (or knee-point) is not observed. Such a characteristic is called a *soft* breakdown. The OD2 diodes have p^+-n junctions. However, a high fixed charge density in the oxide could result in a depletion width that is sufficiently thin to allow tunnelling to occur. If p-type walkout occurs during the breakdown, the depletion width will increase. As the depletion region width increases the probability of tunnelling decreases rapidly, and avalanche breakdown then becomes more likely than Zener breakdown. Thus, the first soft breakdown in Figure 6.10 may be due to tunnelling, and the subsequent sharp breakdowns may be avalanche breakdowns.

The breakdown voltage of the OD2 diodes on wafer D4 is V_{BD} = -19 V. The diodes on wafer D4 have a lower breakdown voltage because the p^s junctions are shallower on wafer D4 than on wafer D5 (Section 6.5.1).

The influence of the junction curvature, r_j, on the breakdown voltage of one-sided abrupt junctions in silicon was discussed in Section 6.5.1. The reduction in breakdown voltage is especially severe for a shallow junction with a small radius of curvature. It becomes less serious as the junction depth increases. Figure 6.12 shows that the breakdown voltage of OD2 278 μm x 50 μm increases as the junction depth, and hence the radius of curvature r_j, increases.

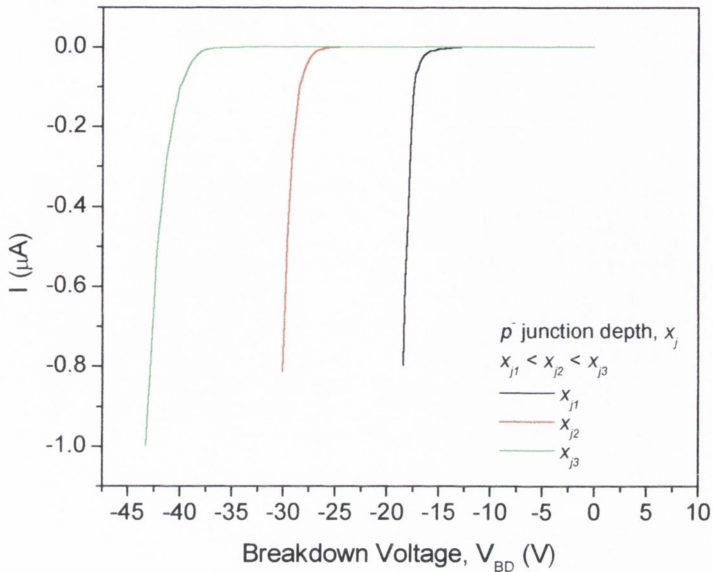


Figure 6.12. Breakdown voltages for OD2 278 μm x 50 μm on wafer D5.

6.5.3 Surround Diodes (SD)

The p^s region in the surround diodes (SD) is surrounded by a p^d region. The breakdown voltage of the p^d corners will determine the breakdown voltage of the SD's, since the electric field intensity is highest at a spherical junction region. Each SD diode would be expected to have the same breakdown voltage. However, SD 666 μm x 15 μm and SD 500 x 20 μm on both wafers have a breakdown voltage of V_{BD} = -48 V. SD 200 x 50 μm has a slightly higher breakdown voltage of V_{BD} = -50 V, and both SD 124 x 80 μm and SD 100 x 100 μm have a breakdown voltage of V_{BD} = -59 V. In order to explain why the breakdown voltage varies between SD diodes, the depletion region width must be taken into consideration. The depletion region width, W, is given by:

$$W = \sqrt{\frac{2\varepsilon_S V}{N_D q}} \tag{6.8}$$

The depletion region width of SD 666 μm x 15 μm and SD 500 x 20 μm at V_{BD} is W = 16.6 μm. Since both of these diodes are narrow, the depletion regions overlap at

either end of the diodes. As a result, the shape of the depletion region at the narrow end of the diode is completely spherical (Fig. 6.13a), resulting in a high electric field intensity and a lower than expected breakdown voltage.

The depletion region width at V_{BD} for SD 124 x 80 µm and SD 100 x 100 µm is W = 18.3 µm. The depletion regions do not overlap at the end of these diodes. The less spherical shape of the depletion region (Fig. 6.13b) will have a lower field intensity than the narrowest diodes, giving a higher breakdown voltage of V_{BD} = -59 V.

SD 200 x 50 µm is an intermediate case, since the breakdown voltage is slightly higher than the two narrowest diodes and lower than the 2 widest diodes. The depletion regions (W = 16.9 µm at V_{BD}) do not overlap, but the cylindrical region is only 20 µm long. The depletion region acquires a roughly spherical shape, with a field intensity that is higher than a diode in which the junction corners are further apart.

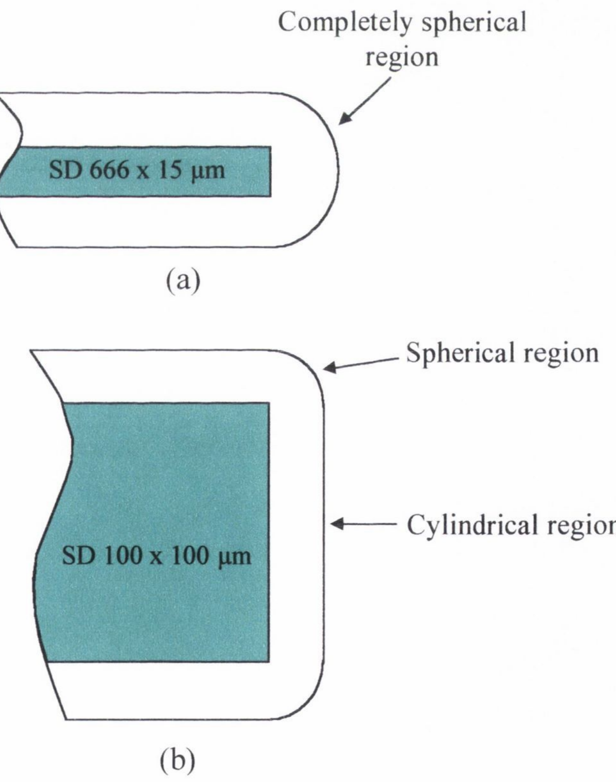


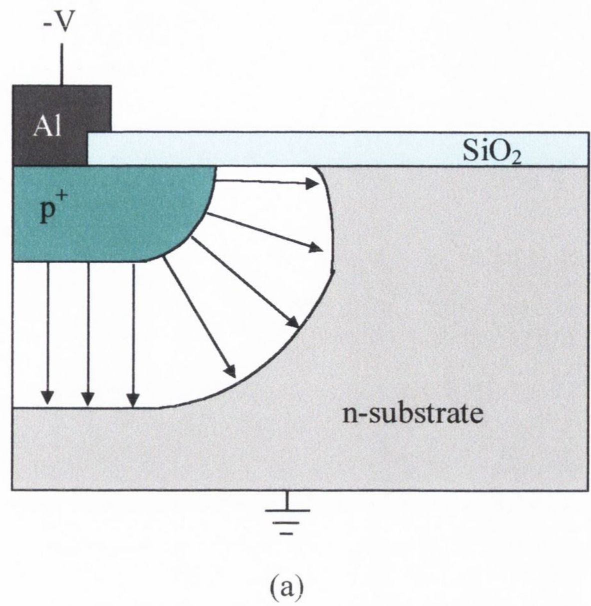
Figure 6.13. Shape of the depletion region around (a) SD 666 x 15 µm and (b) SD 100 x 100 µm.

6.5.4 Open Diodes Type 1 (OD1)

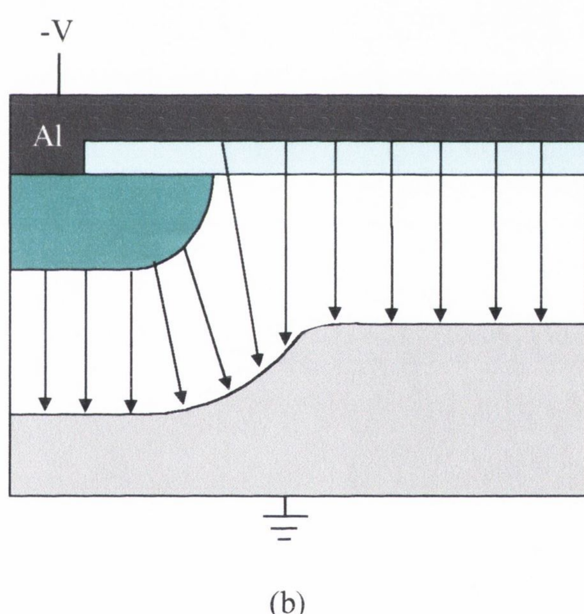
The OD1 diodes on wafers D4 and D5 have a breakdown voltage of V_{BD} = -86 V. It is expected that these diodes will breakdown at the p^d corners, in which case the V_{BD} should be similar to the V_{BD} of the SD diodes, which also breakdown at the p^d corners. However, the OD1 diodes have an unexpectedly high breakdown voltage of V_{BD} = -86 V. One possible explanation for the unexpectedly high breakdown voltages in these diodes is that the aluminium metal contact is acting as a *field-plate*, and is influencing the surface of the *p-n* junction. The ionic charge on the field-plate induces an image charge in the semiconductor and thereby leads to the formation of a surface depletion region. Normally, in planar technology the thermally grown silicon dioxide layer reduces any surface effects. However, the oxide layer in the OD1 diodes is 0.4 μm thick, and may not be sufficiently thick to eliminate the surface effects.

Figure 6.14 shows the effect of a field-plate voltage on the shape of the depletion region. Figure 6.14a shows the shape of the depletion region in a diode that has no field-plate. Under many conditions the electric field near the semiconductor surface will be higher than in the bulk, for instance, the electric field will be higher at the surface if there is trapped charge in the oxide layer. Therefore, breakdown will occur near the surface at a lower reverse bias than that corresponding to the breakdown voltage in the bulk. Figure 6.14b shows the effect of a field-plate voltage on the shape of the depletion region. A surface depletion region of approximately the same width as the junction depletion region is formed and, therefore, the electric field intensity at the surface will be approximately the same as the intensity in the bulk, resulting in high breakdown voltages.

Alternatively, if the junction breakdown voltage due to the field-plate is very high, the OD1 diodes may breakdown at the cylindrical region of the junction, that is not influenced by the field-plate, before reaching the high field-plate breakdown voltage. In either case, the high breakdown voltages of the OD1 diodes are as a result of surface effects induced by a junction field-plate.



(a)



(b)

Figure 6.14. Influence of field-plate voltage on the shape of the depletion region. (a) *p-n* junction with no field-plate (b) *p-n* junction with a field-plate.

6.6 Conclusions

The junction diodes that were fabricated displayed excellent characteristics. The typical forward turn on voltage for all of the diodes was ~0.6 V. The diodes have a near ideal forward slope and the measured ideality factors varied between 1.05 and 1.2, depending

on the diode type. In particular, the OD2 diodes on wafers D4 and D5 have near ideal forward characteristics, with ideality factors of 1.05 and 1.07 respectively, with good exponential increase in current flow over approximately 6 orders of magnitude. These characteristics indicate little influence of generation/recombination effects, and show that the silicon is of very high quality.

Under reverse bias, the diodes have sharp avalanche breakdown with $V_{BD} > -19$ V, depending on the type of diode. The leakage currents observed were very low, rising to only 50 pA as the diodes approached breakdown. These low breakdown voltages and very low leakage currents gives further evidence of the quality of the substrate.

The quality of these *IV* characteristics indicates that the silicon substrate is of high quality showing no evidence of stress induced defects. Therefore, it can be concluded from these results that the proximity RTD process step does not cause any significant degradation of the devices, and is a suitable doping technique for CMOS technology.

References

[1] J.L. Moll, in *Physics of Semiconductors*, McGraw-Hill, N.Y., (1964), Chapter 7.

[2] S.M.Sze and G. Gibbons, *Solid State Electron.*, **9**, 831 (1966).

Chapter 7

Stress Generated in Silicon during Rapid Thermal Processing

7.1 Introduction

Rapid thermal processing (RTP) is an important technology for ULSI chip manufacturing, and is increasingly replacing conventional furnace processing. RTP is already widely used in the IC industry for annealing implant damage and for forming high quality silicide layers. Its low thermal budget allows the formation of high quality shallow junctions, as has been demonstrated in the preceding chapters. However, one of the main challenges in rapid thermal processing lies in minimising the development of thermal stress in the wafer [1]. Stress-induced defects influence physical processes, such as diffusion, and directly lead to degradation of electrical parameters, such as ideality factor, leakage currents and breakdown voltages.

Several efforts have been made to model thermal stress during RTP. Lord [1] modelled the wafer temperature and stress distribution for unpatterned silicon wafers during RTP, using a simple two-dimensional reactor scale model and assuming the temperature profiles to be axisymmetric. Erofeev *et al.* [2] modelled the three-dimensional temperature and stress distributions of wafers. In [3] the effect of metallisation patterns on thermal stress during RTP processing in silicon wafers have been predicted. Bentini *et al.* [4] evaluated the topographic distribution of the slip lines in thermally stressed <100> silicon. The largest deformation in the silicon crystal is experienced in the <110> direction. When the thermally induced resolved stress is larger than the yield strength of the silicon, then slip on the {111} planes in the <110> direction occurs. There are 12 positions on the wafer where stress concentration occurs and where the elastic limit is expected to be exceeded first. Figure 7.1 illustrates schematically the distribution of

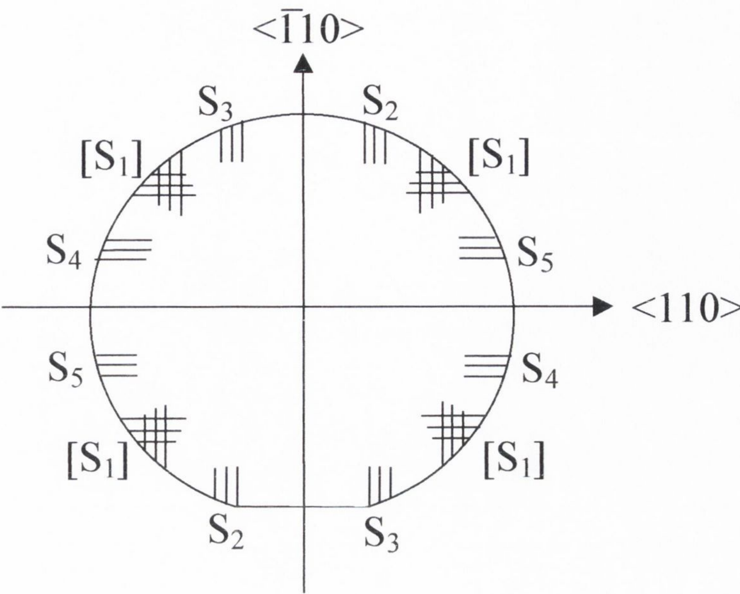


Figure 7.1. Locations of the maxima of the thermally induced resolved stresses and the corresponding slip line directions on a <100> silicon wafer.

these positions. The thermally induced stress component S_1, produces glide which results in steps on the wafer edge, with no visible effect on the wafer surface. The stress components S_2, S_3, S_4 and S_5 give rise to slip lines, which should be visible on the polished surface of the wafer.

This chapter investigates the thermal stress that in generated during proximity rapid thermal oxidation and diffusion using micro-Raman spectroscopy, synchrotron X-ray topography and optical microscopy. Since the introduction of impurity atoms can also lead to the generation of stress within the crystal, boron-induced stress will also be determined.

7.2 Micro-Raman Spectroscopy (MRS)

It has only recently become evident that micro-Raman spectroscopy is a valuable technique for the study of the local mechanical stress in devices and structures used in microelectronics [5-8]. In particular, it has been shown that this technique can be used for the analysis of defects introduced during metallisation in LOCOS structures and

trenches and also in SOI (Si-On-Insulator) materials. To date, the technique has been principally applied to measurements of the stress and strain generated in silicon wafers during thermal processes in furnaces [5]. However, MRS has very recently been used to study the stress that is generated in silicon wafers during rapid thermal diffusion [8, 9].

7.2.1 Experimental Procedure

The micro-Raman measurements were performed in backscattering mode on a Renishaw Raman microscope system 2000. The 514.5 nm line of Ar^+ laser at a power of 25 mW was used as the excitation source. The incident light was polarised parallel to the width of the line, the scattered light was not analysed. Two objective lenses (40x, 4x) located outside the microscope were used in conjunction with a pinhole, with a diameter of 10 μm, to expand the laser spot size. The light was then focused onto the sample using a 50x objective lens mounted on the Olympus BH-2 optical microscope. The size of the laser spot was ~ 2 μm. A Peltier cooled CCD detector with 1 cm^{-1} resolution was used to analyse the spectrum.

The sample was placed underneath the objective lens on an XY translation stage. Raman spectra were recorded at several points across the sample, the positions of which are shown in Figure 7.2. The exciting line peak was measured at each point. In addition, the phonon vibration peak at ~ 520 cm^{-1} was registered at each point, in order to avoid inaccuracy in the peak position due to drift of the calibration frequency. The spectral line was fitted to a Lorentzian curve, allowing the frequency shift to be measured with an accuracy of 0.02 cm^{-1} [6].

7.2.2 Extracting Information from Raman Spectra

The crystalline silicon Raman spectrum in a phonon region mainly consists of a narrow peak around 520 cm^{-1} with a half width of about 4.2 cm^{-1}. The spectrum arises from scattering by long-wavelength transverse optical phonons [10]. With a state-of-the-art spectrometer it is possible to identify a shift in the Raman band of the order of ~0.01 cm^{-1} [5]. Background baseline removal, followed by a line fit using a Lorentzian function allows three components of the Raman spectrum to be determined: intensity; half width and position. These variations are related to the composition, defect density,

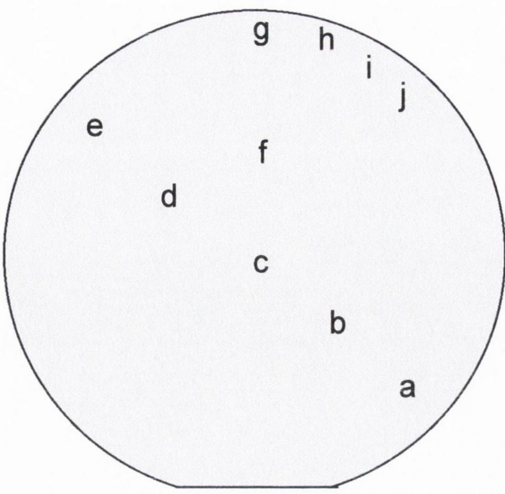


Figure 7.2. Raman spectra were measured at points a-j across each wafer.

and magnitude of stress, σ, respectively. A relationship exists between the stress, σ (in Pa), and the Raman shift, Δω (in cm^{-1}) [11]:

$$\Delta\omega = -2\times10^{-9}\sigma\,, \tag{7.1}$$

where $\Delta\omega = (\omega_{stress} - \omega_{ref})$ (in cm^{-1}), ω_{stress} is the peak frequency of the phonon band of silicon under the stress and ω_{ref} is the peak frequency of the phonon band of the stress-free silicon wafer. A positive or negative shift in the Raman peak position corresponds to compressive or tensile stress, respectively [5], assuming uniaxial stress only, i.e., within the plane of the wafer.

7.3 Synchrotron X-Ray Topography (SXRT)

Synchrotron X-Ray Topography (SXRT) is a non-destructive technique, which can provide a map of the strain/defect distribution in crystals [12]. This technique is based on the difference in reflecting power between perfect and distorted parts of a crystal. The strain induced by local imperfections and stress in the crystal substrate perturb the lattice for several microns from the core of the imperfection. These strained regions modify the

diffraction of the X-rays and give rise to an image of the defect or strained region by means of a local deviation from the ideal Bragg reflection observed in the perfect crystalline regions. SXRT has been used to study melt-growth of alloy systems [13], process-induced defects in Si and GaAs [14, 15], including strain in epitaxial layers [16] and shallow trench isolation CMOS structures [17], and recently strain/defect distribution in Si samples following rapid thermal processing [18].

7.3.1 Experimental Procedure

RTD and RTO Si samples were analysed using two experimental arrangements; White Beam Back Reflection Topography (BRT) and White Beam Topography in Section Topography Mode (ST).

i) White Beam Back Reflection Topography (BRT) is illustrated in Figure 7.3. In this case, the substrate is placed perpendicular to the large area white light X-ray beam. The various lattice planes reflect according to Bragg's Law. Each reflection from a crystal plane is itself a high-resolution topograph and is recorded on a photographic plate.

ii) White Beam Topography in Section Topography Mode (ST) is shown in Figure 7.4. Here, the incident X-ray beam is collimated into a narrow ribbon by a slit typically 10-15 μm wide. Again, a set of reflections from various planes is produced, but in this case each image gives a cross-sectional topograph with direct information about defect depth in the crystal.

The X-ray measurements were performed at HASYLAB-DESY, Hamburg, Germany (Hamburger Synchrotronstrahlungslabor am Deutschen Elektronen-Synchrotron), utilising the continuous spectrum of synchrotron radiation from the DORIS storage ring bending magnet. The ring operated at a positron energy of 4.45 GeV and at typical currents of 80-150 mA. The aforementioned Bragg pattern of topographs was recorded on Kodak SO-181 High-Resolution Professional X-ray film having an emulsion grain size of about 0.05 μm. The distance from the sample to the film was 40 mm in both back reflection and section transmission configurations. However, in section transmission the sample was tilted at 18° to facilitate the recording of certain reflections.

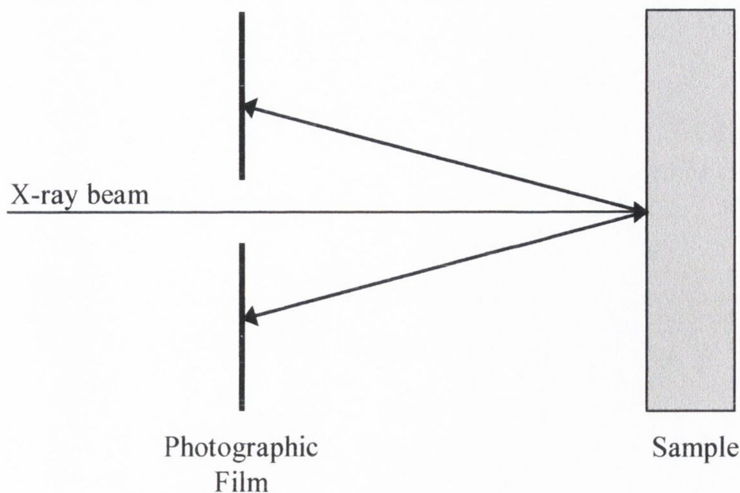


Figure 7.3. Schematic diagram of White Light Back Reflection Topography.

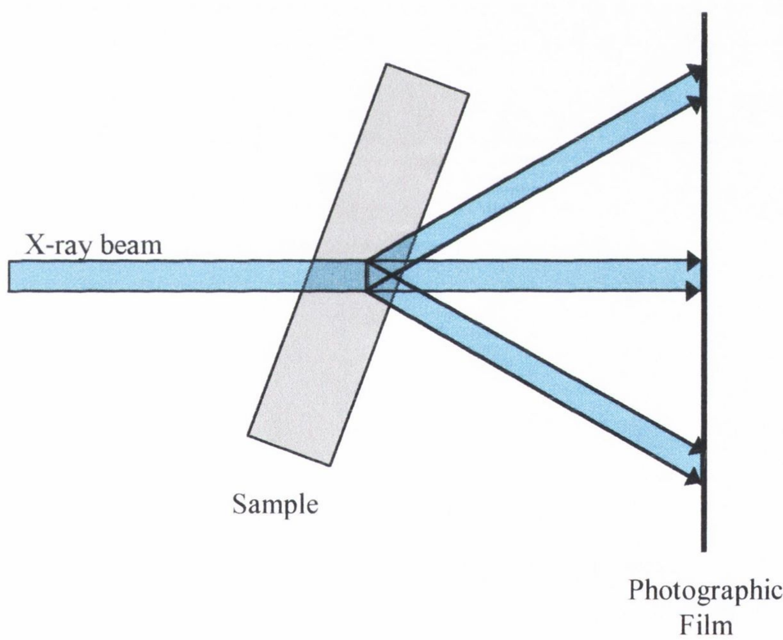


Figure 7.4. Schematic diagram of White Beam Topography in Section Topography Mode.

7.4 Stress Generated during Proximity Rapid Thermal Processing

In Section 7.4.1, the thermal stress that is generated during rapid thermal diffusion (RTD) and rapid thermal oxidation (RTO) will be determined. Section 7.4.2 will concentrate specifically on the strain that boron introduces to the silicon lattice. RTD and RTO samples were prepared as follows:

Table 7.1. Rapid thermally processed samples.

Wafer	Rapid Thermal Process	Temperature (°C)	Duration (s)
A	Diffusion	1050	5
B	Diffusion	1050	16
C	Diffusion	1050	166
D	Oxidation	1050	5
E	Oxidation	1050	16
F	Oxidation	1050	166

(i) RTD Process: SOD B153 was spun onto the source wafers at 6000 rpm for 15 s. The SOD wafers were then baked at 200 °C for 15 minutes to evaporate moisture and light organics from the SOD layer (Chapter 3). Next, the SOD layer was cured at 900 °C for 45 s to convert it to a BSG dopant source (Chapter 4). The dopant source was then stacked in proximity to a bare Si product wafer (Figure 3.1, Chapter 3). RTD was performed at 1050 °C for 5-166 s, as shown in Table 7.1.

(ii) RTO Process: RTO samples were fabricated by stacking a bare Si dummy wafer in proximity to a bare Si product wafer in the RTP. By doing this, the product wafer will still experience the same temperature profile as during RTD, thereby minimising temperature variations between the doped and oxidised wafers. RTO was performed at 1050 °C for 5-166 s, Table 7.1.

All of the rapid thermal processing was performed in an oxidising ambient of 25% O_2: 75% N_2. The RTP ramp rate was 48 °C/s for all processes.

7.4.1 Thermal Stress

Wafers A-F were inspected using an optical microscope. Slip lines were only observed in wafers B, C and F. The slip lines observed extended from the edge of the wafers in 8 positions ($S_2 \ldots S_5$), as predicted by Bentini *et al.*, Figure 7.1. The length of the slip lines in wafers B, C and F were 0.5 cm, 2.5 cm and 0.03 cm, respectively.

Table 7.2 shows the measured Raman peak shift, $\Delta\omega$, for the oxidised wafers D-F. Raman spectra were measured at positions a-j across the wafer (Figure 7.2). The positive

Table 7.2. Measured Raman peak shift, Δω, for oxidised wafers D-F.

Position on the wafer	Raman shift, Δω (cm^{-1})		
	Wafer D	Wafer E	Wafer F
a	0.12	0.2	0.59
b	0.09	0.2	0.29
c	0.09	0.35	0.65
d	0.01	0.20	0.44
e	0.18	0.19	0.27
f	0.03	0.20	0.07
g	0.06	0.15	0.004
h	0.04	0.06	0.38
i	0.23	0.24	0.64
j	0.03	0.10	0.31

Raman shift in Table 7.2 indicates that a compressive stress was generated in each of the wafers. For simplicity, the values of stress at the centre of each wafer are plotted in Figure 7.5. It can be seen that the stress increases with RTO duration.

Table 7.3 shows the measured Raman peak shift for the doped wafers A-C. A compressive stress was generated across wafer A and B. However, a different stress

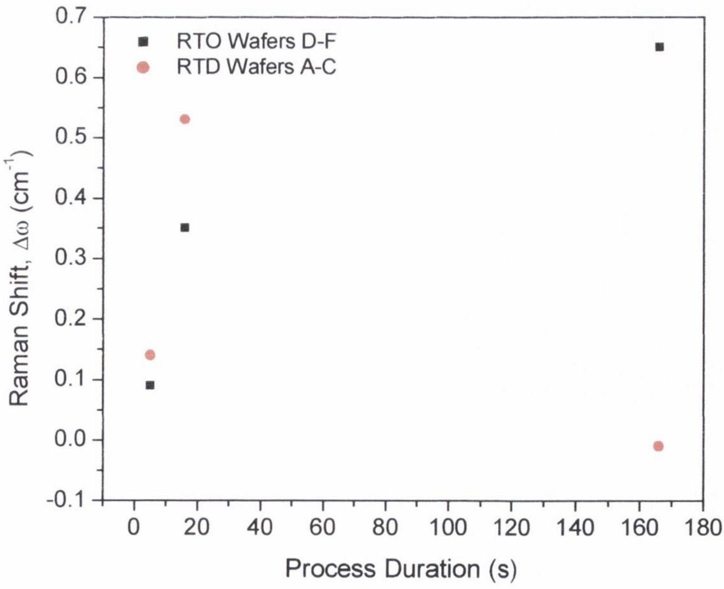


Figure 7.5. Raman shift, Δω, as a function of process duration for RTO and RTD at 1050 °C.

Table 7.3. Measured Raman peak shift, Δω, for doped wafers A-C.

Position on the wafer	Raman shift, Δω (cm^{-1})		
	Wafer A	**Wafer B**	**Wafer C**
a	0.15	0.17	0.32
b	0.16	0.29	0.17
c	0.14	0.53	-0.01
d	0.15	0.19	0.04
e	0.14	0.35	0.15
f	0.16	0.23	0.08
g	0.18	0.41	0.23
h	0.19	0.46	0.24
i	0.34	1.04	0.44
j	0.15	0.36	0.14

distribution was observed in wafer C. The induced stress was tensile at the centre of the wafer, and compressive at the edges of the wafer. The exact reason for this is not clear at present. The values of stress at the centre of wafers A-C are also plotted in Figure 7.5. Figure 7.5 shows that the stress in wafer B is greater than the stress in wafer A, due to an increase in RTD time at 1050 °C. However, the magnitude of stress in wafer B is larger than the stress in wafer C. The induced stress in both wafers exceeded the yield strength of silicon causing the crystal to slip. The formation of slip dislocation lines increased with increased time at the peak temperature, resulting in a greater stress relaxation in wafer C compared to wafer B.

Figure 7.5 also shows that the induced stress in the doped wafers is greater than the induced stress in the oxidised wafers for short times (5 and 16 s). This suggests that the presence of boron in the silicon lattice increases the generation of stress. However, the stress measured in the oxidised wafer appears to be greater than the stress in the doped wafer after 166 s. Since boron increases the generation of stress in silicon, the induced stress in a doped wafer will exceed the yield strength of silicon quicker than in an undoped wafer. Furthermore, slip lines will be generated in a doped wafer quicker than in an undoped wafer. Optical microscopy has already revealed that slip lines were generated to a greater extent in doped wafer C than in undoped wafer F. Therefore, there is greater stress relaxation in wafer C than in wafer F. This explains why the measured stress in wafer F was greater than in wafer C.

Lord [1] predicts that the rapid ramp rates are not the cause of the most damaging stresses in the wafers; it is the temperature non-uniformity at the peak temperature that tends to stress the wafers beyond the yield stress. Due to the strong temperature dependence of the silicon critical yield stress [19], the extent of slip line formation depends on both the peak wafer temperature and the time spent at the peak temperature. Our observations of slip lines and Raman measurements agree with this previous work, and show that for a given peak temperature, slip line formation increases with increasing time at peak temperature.

In Tables 7.2 and 7.3, point i corresponds to stress component S_2 in Figure 7.1. The Raman peak shift, and therefore the induced stress, at point i in all of the wafers is greater than the peak shift at points g, h and j. This observation also agrees with the predictions of Bentini *et al.*

7.4.2 Boron-Induced Strain

Some distortion of the silicon lattice must occur in accommodating dopant impurity atoms. Silicon has a tetrahedral radius of 1.18 Å. Boron has a tetrahedral radius of 0.88 Å, corresponding to a misfit of $\varepsilon = 0.254$. Since boron is smaller than silicon, the presence of boron in the silicon lattice contracts the lattice and induces strain.

SXRT was used to analyse doped and undoped wafers after rapid thermal processing at 1050 °C (wafers A, C and F). A bare as-received Si wafer was used as a reference wafer. Figure 7.6a shows a section transmission topograph for the reference wafer. The dark and light bands that are observed in this topograph are called *Pendellösung Fringes*. These fringes are indicative of extremely high quality, crystalline samples. Figure 7.6b shows a topograph for an oxidised wafer (wafer F). The Pendellösung fringes present in this image are slightly less sharp than in Figure 7.6b. However, the sample is very nearly as high in quality as the reference wafer. A degradation in crystallinity, caused by boron doping and the high temperature anneal was deduced in Figure 7.6c for wafer A. This is evident from the reduction in the definition of the Pendellösung fringes. Strain features are also present near the upper doped surface. Figure 7.6d shows a topograph of a heavily boron doped wafer (wafer C) after a prolonged RTD. Slightly distorted Pendellösung fringes can be seen in the left-hand side of the topograph suggesting

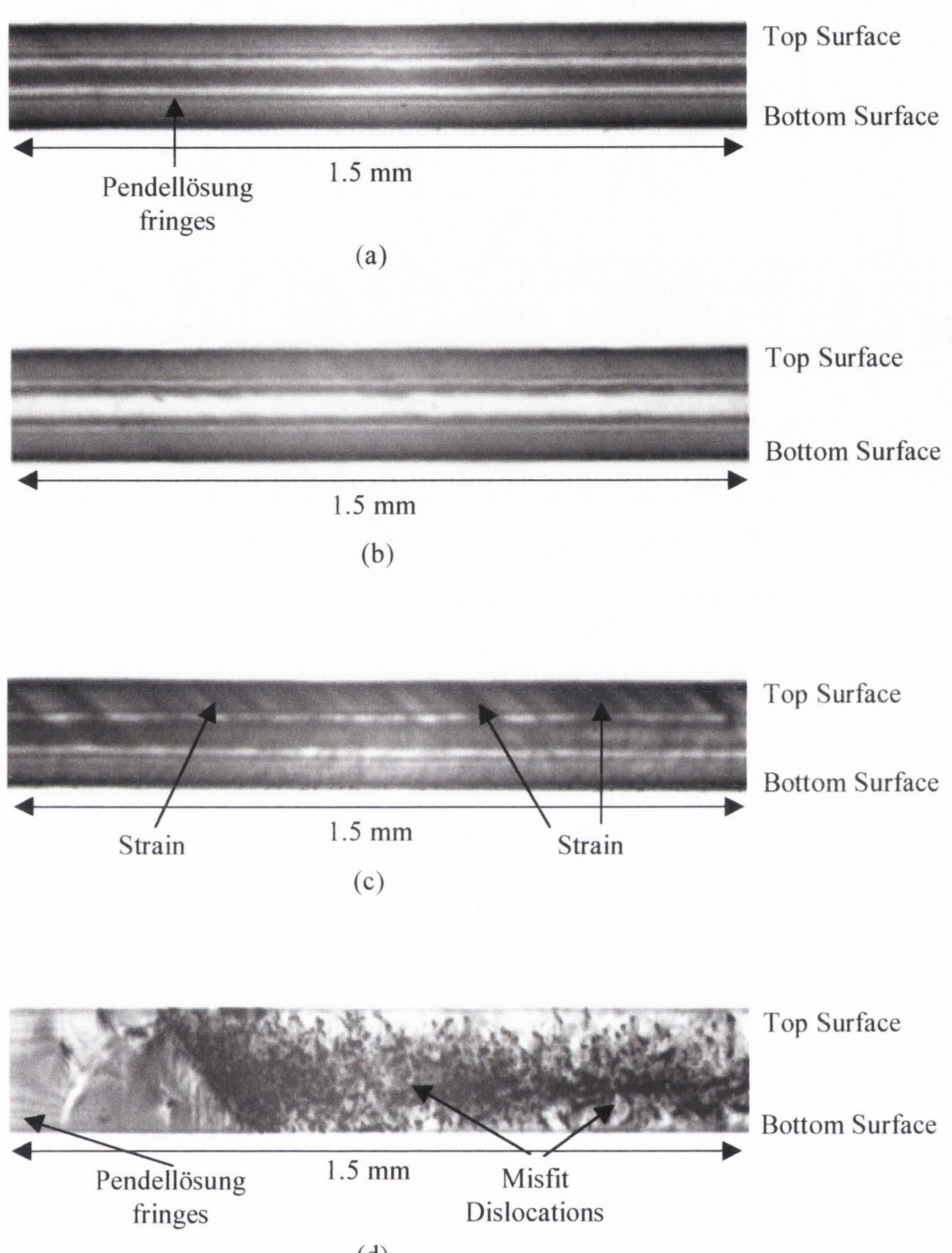


Figure 7.6. SXRT section topographs: (a) Bare as-received Si reference wafer; (b) Wafer F: RTO for 166 s at 1050 °C; (c) Wafer A: RTD for 5 s at 1050 °C; (d) Wafer C: RTD for 166 s at 1050 °C.

reasonable substrate quality. This region of the sample is not doped. However, the boron-doped region (right-hand side) possesses a dense dislocation array. These dislocations are also evident in Figure 7.7, which is a large area reflection topograph of

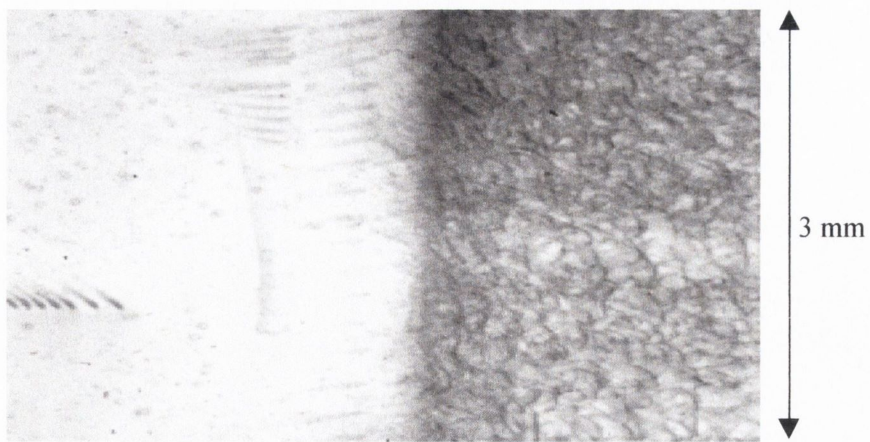


Figure 7.7. SXRT back reflection topograph of wafer F.

the same wafer. These images confirm the observation that heavy boron doping induces misfit dislocation.

These SXRT results are confirmed from the analysis of the Raman half widths. Raman half width is related to the defect density in a material. The half width for a bare as-received reference wafer is 4.18. As shown in Table 7.4, there is very little change in the values of half width obtained for wafers A and F. However, for the case of wafer C there is a noticeable increase in half width, which proves that a significant structural disorder exists in this wafer.

Table 7.4. Measured Raman half widths, Γ, for wafers F, A and C.

Position on the wafer	**Raman half width, Γ (cm^{-1})**		
	Wafer F	**Wafer A**	**Wafer C**
a	4.19	4.20	4.65
b	4.14	4.18	4.72
c	4.18	4.19	5.21
d	4.22	4.19	5.26
e	4.09	4.23	5.14
f	4.18	4.18	5.32
g	4.17	4.25	4.86
h	4.19	4.20	4.95
i	4.18	4.24	5.06
j	4.17	4.25	5.17

7.5 Conclusions

Thermal stress and boron-induced strain that is generated during RTO and RTD has been investigated. In conclusion, stress generation, and the corresponding formation of slip lines, increases with the time spent at 1050 °C. The introduction of boron atoms into silicon deteriorates the stress that is induced in the lattice. SXRT clearly shows a dense array of misfit dislocations that are formed in heavily doped silicon after prolonged RTD. These results were confirmed by micro-Raman spectroscopy. Surface strain was observed in wafers that were doped for 5 s at 1050 °C, however, there was no evidence of any dislocations.

Chapter 6 described the electrical characteristics of diodes that were formed by RTD at 1050 °C for 5 s. These diodes displayed excellent electrical characteristics with near ideal forward characteristics, low leakage currents and sharp breakdown voltages. Therefore, even though strain was observed in wafers after RTD at 1050 °C for 5 s, the electrical characteristics show no evidence of any serious stress-induced defects. In conclusion, the strain that is induced after 5 s at1050 °C is minimal and does not degrade the device performance, proving that proximity RTD is a suitable doping technique for CMOS technology.

References

[1] H.A. Lord, *IEEE Trans. Semiconductor Manufacturing*, **1**, 105, (1988).

[2] A.F. Erofeev, T.M. Makhviladze, A.V. Panjukhin, O.S. Volchek and O. Adetutu, *Proc. 4th Int. Rapid Thermal Processing Conf.*, Boise, Idaho, Sept. 11-13, (1996), p. 342.

[3] J.P. Hebb and K.F. Jensen, *IEEE Trans. Semiconduct. Manufacturing.*, **11**, 101, (1998).

[4] G. Bentini, L. Correra and C. Donolato, *J. Appl. Phys.*, **56**, 2922, (1984).

[5] I. de Wolf, *Semicond. Sci. Technol.*, **11**, 139, (1996).

[6] K. Ajito, J.P.H. Sukamoto, L.A. Nagahara, K. Hashimoto and A. Fujishima, *J. Vac. Sci. Technol.*, **A13**, 1234, (1995).

[7] S. Webster, D.N. Batchelder and D.A. Smith, *Appl. Phys. Lett.*, **72**, 1478, (1998).

[8] M. Nolan, T. Perova, R.A. Moore, C.J. Moore, K. Berwick and H.S. Gamble, *Materials Science and Eng. B*, (2000) (In press).

[9] M. Nolan, T. Perova, R.A. Moore, C.J. Moore, K. Berwick, H. Byrne and H.S. Gamble, *European Mat. Res. Soc. Symp. Proc.*, E-21, (1999).

[10] M. Cardona, in *Light Scattering in Solids II*, M. Cardona and G. Güntherodt, Eds., Springer, Heidelberg, (1982), p. 19.

[11] T. Englert, G. Abstreiter and J. Pontchara, *Solid-State Electron.*, **23**, 31, (1980).

[12] D.K. Bowen and B.K. Tanner, in *High Resolution X-ray Diffractometry and Topography*, Taylor & Francis Ltd., (1998), p. 221.

[13] G. Grange, J. Gastaldi, C. Jourdan and B. Billia, *J. Crystal Growth*, **151**, 192, (1995).

[14] T. Tuomi, M. Tuominen, E. Prieur, J. Partanen, J. Lahtinen and J. Laakkonen, *J. Electrochem. Soc.*, **142**, 1699, (1995).

[15] P.J. McNally, P.A.F. Herbert, T. Tuomi, M. Karilahti and J.A. Higgins, *J. Appl. Phys.*, **79**, 8294, (1996).

[16] P.J. McNally, A.N. Danilewsky, J.W. Curley, A. Reader, R. Rantamaki, T. Tuomi, M. Bolt and M. Taskinen, *Microelectronic Eng.*, **45**, 47, (1999).

[17] P.J. McNally, J.W. Curley, M. Bolt, A. Reader, T. Tuomi, R. Rantamaki, A.N. Danilewsky and I. DeWolf, *J. Mat. Sci: Materials in Microelectronics*, **10**, 351, (1999).

[18] M. Nolan, T.S. Perova, R.A. Moore, H.S. Gamble, D. Lowney, P. J. McNally, T. Tuomi, R. Rantamaki and A.N. Danilewsky, The XVIIth International Conference on Raman Spectroscopy, ICORS 2000, Bejing, China, August 20-25, (2000). (Accepted)

[19] I. Matsuba, K. Mokuya and T. Aoshima, *IEDM Tech. Dig.*, 530, (1986).

Chapter 8

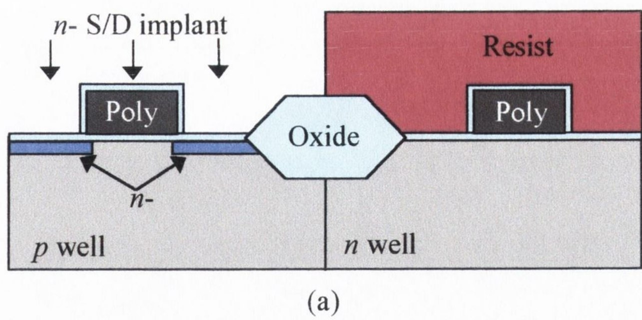
LDD and Proximity RTD Process Integration

8.1 Introduction

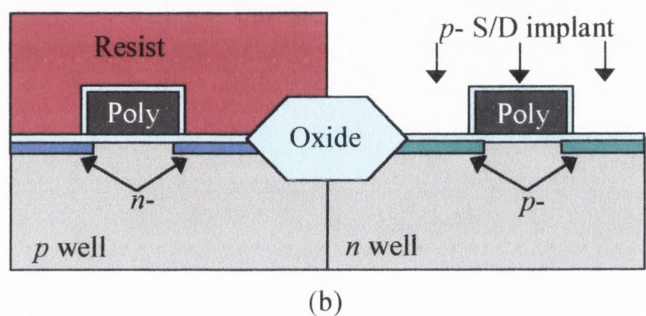
Proximity RTD has been investigated as a technique for fabricating shallow *p*-type junctions for LDD devices. The technique was successfully integrated into a CMOS process to produce high quality *p-n* junctions diodes. However, it is beyond the scope of this work to fabricate LDD MOS transistors. In order to complete the study, this chapter will describe the fabrication process and the limitations of conventional LDD devices, followed by a proposed method of integrating proximity RTD into an improved LDD process.

8.2 Conventional LDD Device Fabrication

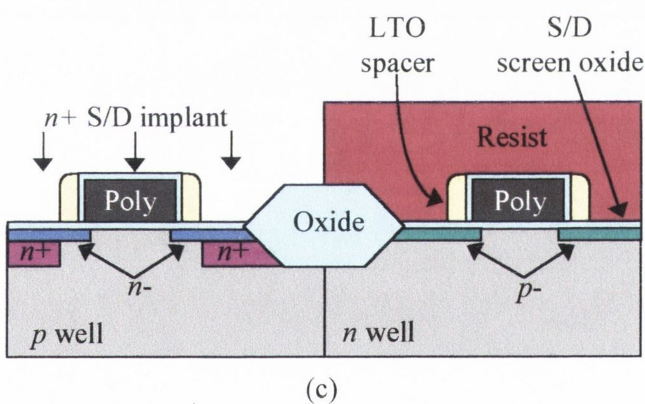
The conventional LDD spacer process is illustrated in Figure 8.1 [1, 2]. The poly gate is patterned with an RIE etch process designed to produce a poly feature with vertical sidewalls. The poly gate etch process may form microtrenches in the gate oxide at the gate edge, thus causing all of the oxide there to be removed [3]. Therefore, following the poly etch, the patterned poly is oxidised to improve the gate oxide edge integrity. Next, a 1×10^{13} cm^{-2} phosphorous *n*- implantation and a 1×10^{13} cm^{-2} boron *p*- implantation are carried out to form the NMOS and PMOS LDD regions, Fig 8.1a and b. Permanent oxide sidewall spacers are formed by LPCVD and RIE processing of a conformal low temperature oxide (LTO) layer. Again, the RIE processing may damage or remove the oxide layer over the source/drain regions. This oxide layer protects the underlying silicon from implant damage, therefore, it is important that it is reoxidised prior to the



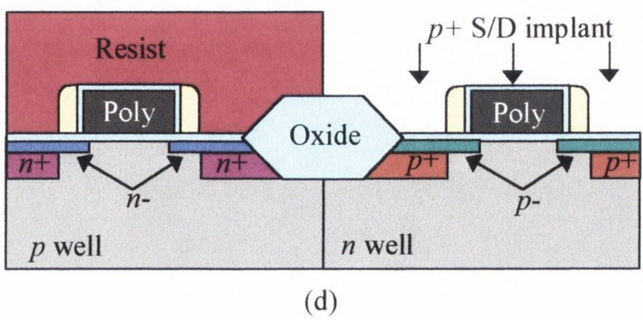
(a)



(b)



(c)



(d)

Figure 8.1. Conventional LDD spacer process. (a) Selective n- implantation; (b) Selective p-implantation; (c) LTO spacer deposition and RIE, followed by a source/drain screen oxidation. Next, a selective n+ implantation is performed; (d) Selective p+ implantation.

n^+/p^+ implantation. This reoxidised layer is referred to as a screen oxide. The source/drain regions are then formed by selective arsenic n^+ and boron p^+ implantations, Fig. 8.1c and d.

8.3 Limitations of Conventional LDD Devices

The advantages of LDD devices were discussed in Chapter 1. In brief, LDDs improve the reliability and device performance of submicron MOSFETs by suppressing short-channel effects, such as hot-carrier effects, drain-induced-barrier-lowering and punchthrough. However, the early LDD devices exhibited some disadvantages. One of the primary disadvantages lies in the conventional LDD process sequence; the LDD is formed *before* the n^+/p^+ source/drain junctions. Therefore, the LDD junctions are exposed to the high temperature source/drain thermal anneals. This results in an unwanted deepening of the LDD junction depths. The LDD also diffuses laterally under the gate during the S/D anneal causing a further reduction in the effective channel length. The Semiconductor Industry Association (SIA) have predicted that the LDD junction depth will continue to decrease with each new generation of CMOS technology (Table 1.1, Chapter 1). Therefore, greater control of the LDD junction depth is essential and the doping sequence of the conventional LDD process becomes intolerable.

Additional concerns arise when salicided source/drain junctions are integrated with deep submicron channel length CMOS device processing. The salicide (self-aligned silicide) process has become widely used in submicron MOS technology as a method of reducing the source/drain series resistance [4, 5]. Due to the consumption of silicon in the salicide process [4], deep source/drain junctions are required to minimise junction leakage. Such a constraint conflicts with the requirement for shallow junctions in submicron devices for improved short-channel behaviour. Deeper source/drain junctions result in deeper and wider LDDs and the problems mentioned above are aggravated even further. Improved LDD structures will be discussed in Section 8.4.

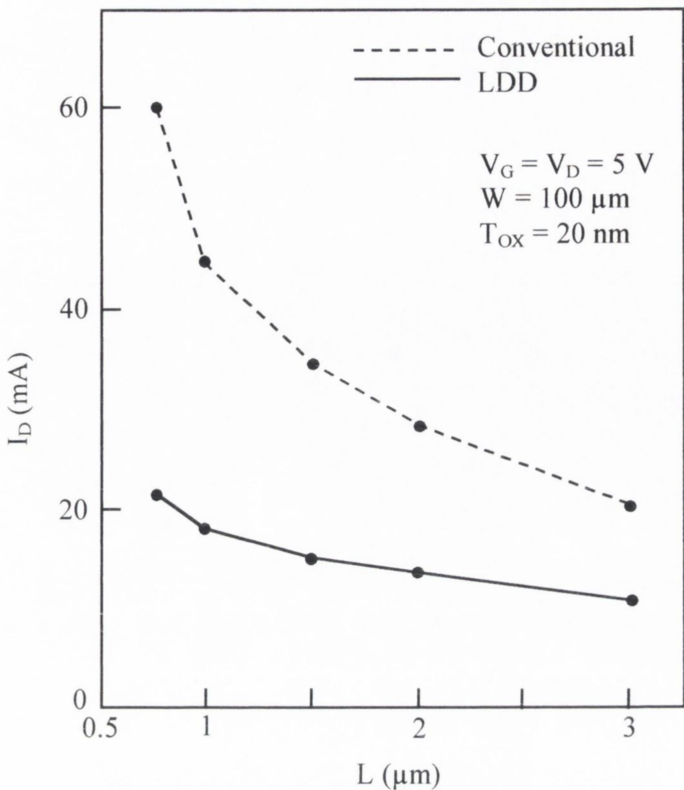


Figure 8.2. Drain current in conventional and LDD MOSFETs [6].

The introduction of a LDD also causes device degradation by reducing the drain current, I_D. Figure 8.2 compares the drain current in a conventional MOSFET with that in a LDD MOSFET [6]. The LDD MOSFET exhibits smaller I_D values. Adding LDD regions to the MOSFET is equivalent to adding two parasitic resistors to the total source/drain resistance. Hence, a higher total resistance results in a lower value of I_D.

In early LDD MOSFETs, the criterion for selecting the optimum device structure for hot-carrier resistance was based on minimising the horizontal electrical field, E_Y. Although this assumption was correct, it was not necessarily the case that devices with smaller E_Y exhibited better hot-carrier reliability. The location of E_{YMAX} influences the hot-carrier reliability of the device, and its location is dependent on the doping concentration of the LDD region. E_{YMAX} moves closer to the edge of the p^+ region of the drain if the p^- concentration is low. Consequently, since the damage due to hot-carrier injection is localised in the gate oxide above the location of E_{YMAX}, this would place the damage in the spacer region of the channel. This segment of the channel is outside the edge of the

gate, therefore, it will tend to be less heavily inverted compared to the rest of the channel. The accumulation of injected hot-carriers in the oxide above this location will reduce the level of inversion even more, thus causing the parasitic resistance to increase and I_D to decrease. This hot-carrier degradation is called structure-dependent degradation effect, and it was observed in conventional LDD structures.

One method of moving the location of E_{YMAX} is to increase the LDD doping concentration. This method will be discussed in Section 8.4. An alternative method to avoid the problem is to fully overlap the LDD region with the gate. These device structures with fully overlapped LDD regions, such as inverse-T LDD (ITLDD) [7], total overlap with polysilicon spacer (TOPS) [8] and gate-drain overlapped LDD (GOLD) [9], require extremely complex manufacturing processes and are very difficult to fabricate. For this reason, the remainder of this chapter will concentrate on partially overlapped LDDs since they are easier to fabricate than fully overlapped devices, but still improve hot-carrier reliability.

8.4 Improvements to the Conventional LDD Structure and Proposed Integration of Proximity RTD

A Disposable Polysilicon LDD spacer technology was introduced by Parrillo *et al.* in 1991 [1]. The new Disposable Polysilicon spacer process, as illustrated in Figure 8.3, has demonstrated improved short-channel effects in both NMOS and PMOS transistors. After the 3500 Å undoped polysilicon gates are etched and reoxidised, a 1500 Å layer of undoped polysilicon is deposited by LPCVD, Fig. 8.3a. Using resist to cover the PMOS areas, the NMOS spacers are formed by RIE. After a $5x10^{15}$ cm^{-2} arsenic n^+ implant is performed, the NMOS spacers are removed by dry etching, the resist is stripped, and the arsenic implant is thermally annealed, Fig. 8.3b. This key thermal cycle provides deep n^+ arsenic regions to accommodate the silicon consumption during the later salicidation step for low leakage salicided junction requirement. In addition, the arsenic implant damage is sufficiently annealed, to prevent defect enhanced diffusion in the subsequent LDD implant. With the poly layer covering the PMOS areas, a blanket phosphorus $5x10^{13}$ cm^{-2} *n-* implant is performed, Fig. 8.3c. (It is important to note that this *n-*

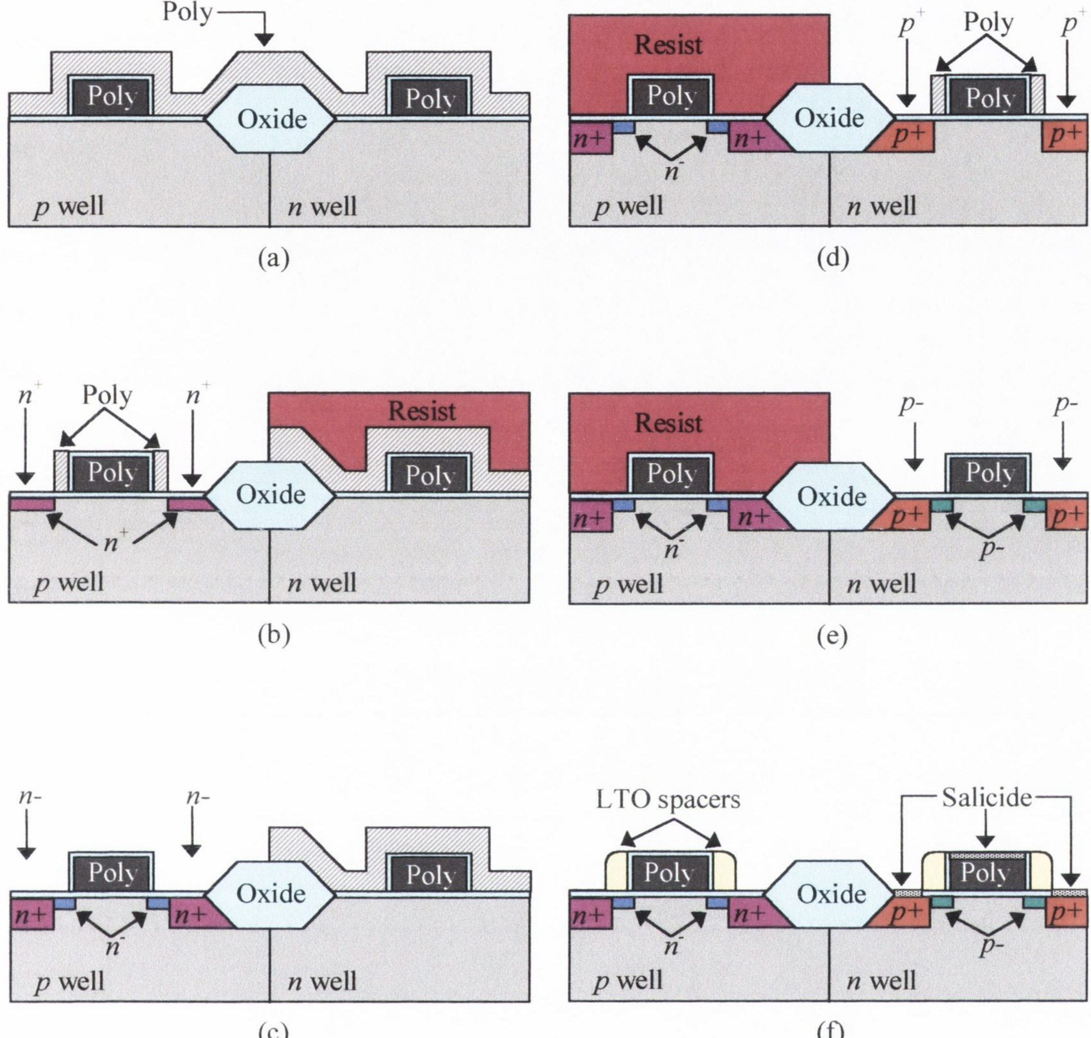


Figure 8.3. Disposable spacer process. (a) Undoped polysilicon is deposited as the spacer material. (b) Using resist to cover PMOS areas, the NMOS is formed and followed by an arsenic n+ implantation. The resist is stripped and the arsenic profile is annealed. (c) A blanket phosphorus n- implant is performed which uses the nonetched polysilicon spacer material to protect the PMOS areas. (d) Using the second patterning step, the NMOS areas are covered with resist and the PMOS poly spacers are formed, followed by the p+ source/drain implant. (e) The spacers are removed and the p- implant is performed. (f) The resist is stripped and the permanent LTO spacers are formed to define the salicided source/drain/gate regions. Conventional processing follows with contact and aluminium metallisation definition. [1]

implant dose is greater than the *n-* implant dose in the conventional LDD (*n-* = 1×10^{13} cm^{-2}), Section 8.2; the explanation will be given towards the end of this section.) Using the second resist mask, the PMOS poly spacers are formed by RIE, followed by a 5×10^{15} cm^{-2} boron p^+ implantation and thermal anneal, Fig. 8.3d. The poly

spacers are then removed by dry etching and a $1x10^{14}$ cm^{-2} boron *p-* implantation is performed, Fig. 8.3e. (The *p-* implant dose is also greater than the conventional LDD dose.) A permanent 2000 Å CVD oxide spacer is then formed prior to $TiSi_2$ salicidation to provide an appropriate offset distance to avoid bridging between the source/drain and the gate regions, without compromising the LDD design, Fig. 8.3f.

Proximity RTD could be integrated into the Disposable Polysilicon spacer process. Instead of using a *p-* implantation in Figure 8.3e, proximity RTD could be used to fabricate the shallow junctions. Chapter 5 has already shown that boron junctions with a surface concentration of $1.2x10^{20}$ cm^{-3} and a depth of 308 Å at 10^{18} cm^{-3} can be produced by proximity RTD. These junctions satisfy the SIA requirements for 0.1 μm technology (Chapter 1, Table 1.1). The B_2O_3 glass layer that is transferred to the product wafer during RTD would have to be removed after the doping process, to prevent unwanted anomalous diffusions during subsequent thermal processes. Also, if Proximity RTD was used to fabricate the *p-* junctions in Figure 8.3e, the resist could not be used to mask the NMOS, since it could not be heated in the RTP. SiN could be used instead of the resist. The SiN could be removed with phosphoric acid heated to 180 °C. This process would be adequate for demonstration purposes, however it would not be a suitable manufacturing process due to the hazardous nature of phosphoric acid. It is beyond the scope of this work to investigate how Proximity RTD could be fully integrated into a manufacturable Disposable Polysilicon spacer process.

The Disposable LDD process has several advantages compared to the conventional LDD process. Firstly, the source/drain and LDD process sequence has been reversed. Therefore, the LDD junctions are not exposed to the source/drain thermal anneal. Since the LDD implants only receive low temperature annealing of the back-end process, shallow LDD junctions depths are maintained. Excellent NMOS and PMOS short-channel characteristics have been obtained using this process with no deleterious lateral dopant diffusion being observed [1]. Secondly, the Disposable LDD process allows independent adjustment of the LDD and salicide spacers to optimise LDD design while avoiding salicide bridging of source/drain to gate regions.

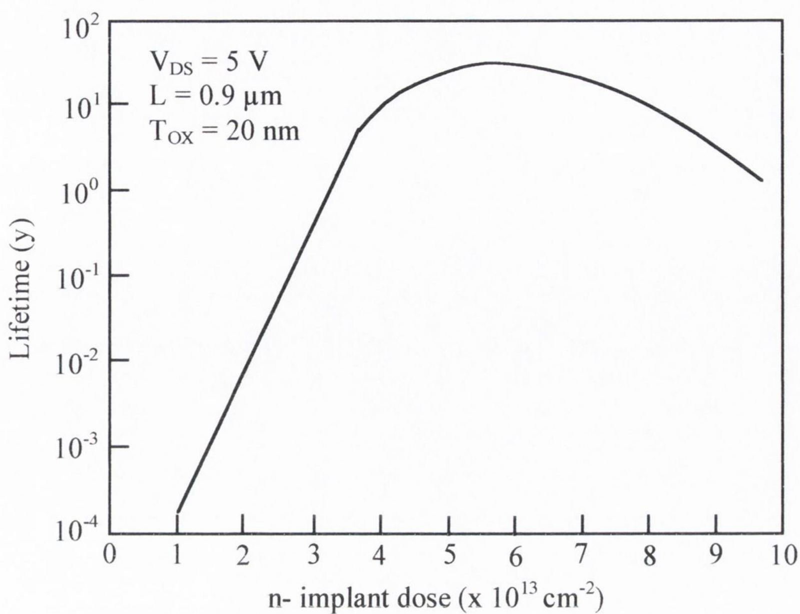


Figure 8.4. The dependence of device lifetime on the LDD n- implant dose. The maximum lifetime point is at 5-6x10^{13} cm^{-2} [10].

Finally, the LDD doping concentration is higher in the Disposable LDD process than in the conventional LDD process. The advantage of altering the doping concentration will be explained in the following discussion. The reduction in I_D and the influence of the location of E_{YMAX} in LDD devices were discussed in Section 8.3. One solution to these problems is to increase the LDD doping concentration. These devices have been called moderate lightly doped drain (M-LDD). The increased LDD doping implies that E_Y is not reduced as much as in conventional LDD devices. However, the sheet resistance of the M-LDD is lower than that of the conventional lightly doped LDD, thus improving the drain current. Secondly, the increased doping concentration moves the location of E_{YMAX} under the gate edge, such that the damaged region in the oxide also lies beneath the gate edge, instead of underneath the spacer as occurs in the lightly doped LDD. The gate field counteracts the damage in the oxide, resulting in less depletion than before. Also, the heavier doping makes it harder to deplete the LDD region. These combined effects made the device more resistant to hot-carrier degradation. Figure 8.4 shows the lifetime of an LDD as a function of LDD *n-* concentration [10]. At optimum doping (5-6x10^{13} cm^{-2}), the M-LDD exhibits 4-5 orders of magnitude improvement in lifetimes over conventional LDDs (*n-* = 1x10^{13} cm^{-2}). The advantages of the M-LDD are so great that it is frequently incorporated into MOS devices, including the Disposable Polysilicon LDD spacer

process. The SIA predicts that the LDD doping concentration will continue to increase with each new technology generation (Table 1.1, Chapter 1), in an attempt to improve hot-carrier resistance.

8.5 Conclusions

LDD MOSFET structures have been used extensively for improving short-channel effects. However, even though the LDD suppresses short-channel effects, conventional LDD structures cause other device degradation. Firstly, the LDD junctions receive the n^+/p^+ source/drain thermal anneal which increases the LDD junction depth. This becomes a critical issue in deep submicron salicided junctions. Secondly, the drain current in very lightly doped LDD devices is reduced due to an increase in series resistance and unfavourable positioning of the maximum lateral electric field, E_{YMAX}.

Several improvements have been made to the conventional LDD MOSFETs. By reversing the doping sequence, the Disposable Polysilicon spacer process provides deeper source/drain junctions for lower salicided junction leakage while simultaneously providing shallow LDD junctions for improved short-channel effects. Additionally, the process allows the LDD and the salicide spacer technology to be optimised independently. Increased LDD doping (M-LDD) improves the drain current. The heavier doped M-LDD also exhibits increased hot-carrier reliability, by eliminating the structural degradation effect that leads to more rapid hot-carrier degradation in lightly doped LDD devices. Proximity RTD could be integrated into the improved Disposable spacer technology.

References

[1] L.C. Parrillo, J.R. Pfiester, J.-H. Lin, E.O. Travis, R.D. Sivan and C.D. Gunderson, *IEEE Trans. Electron. Dev.*, **38**, 39, (1991).

[2] S. Ogura, P.J. Tsang, W.W. Walker, D.L. Critchlow and J.F. Shepard, *IEEE Trans. Electron. Dev.*, **27**, 1359, (1980).

[3] T.J. Dalton, J.C. Arnold, H.H. Sawin, S. Swan and D. Corliss, *J. Electrochem. Soc.*, **140**, 2395, (1993).

[4] Y. Taur, J.Y.-C. Sun, D. Moy, L.K. Wang, B. Davari, S.P. Klepner and C.-Y. Ting, *IEEE Trans. Electron. Dev.*, **34**, 575, (1987).

[5] H. Ronkainen, G. Drozdy and S. Franssila, *IEEE Electron Dev. Lett.*, **12**, 125, (1991).

[6] F. Hsu and K.-Y. Chiu, *IEEE Electron Dev. Lett.*, **5**, 162, (1984).

[7] J.R. Pfiester, F.K. Baker, R.D. Sivan, N. Crain, J.-H Lin, C. Seelbach, C. Gunderson and D. Denning, *IEEE Electron Dev. Lett.*, **11**, 253, (1990).

[8] J.E. Moon, T. Garfinkel, J. Chung, M. Wong, P.K. Ko and C. Hu, *IEEE Electron Dev. Lett.*, **11**, 221, (1990).

[9] R. Izawa, T. Kure and E. Takeda, *Tech. Dig. IEDM*, 38, (1987).

[10] M. Kinugawa et al., *Symp. VLSI Tech.*, 116, (1985).

Chapter 9

Conclusions and Suggestions for Future Work

9.1 Conclusions

This work has investigated proximity RTD as a technique for fabricating high-quality, shallow, boron-doped *p*-type junctions for LDD MOSFET devices. Spin-on dopants were investigated as dopant sources for RTD.

Spin-on dopants contain organic solvents and moisture. Carbon generates lattice strain in silicon and moisture induces hot-carrier degradation in gate oxides. Therefore, it is necessary to remove both of these contaminants prior to proximity RTD. It was determined using FTIR that a 200 °C bake for 15 minutes efficiently removed the solvents and moisture from the SOD. Having removed the contaminants, the commercially available B153 SOD was found to be a suitable dopant source, since it transferred boron efficiently to the product wafer during proximity RTD without causing any carbon contamination.

SOD B153 is a borosilicate gel. Borosilicate gels are thermodynamically unstable. Upon heating, the composition and structure of the borosilicate gel changes to that of a borosilicate glass. Sheet resistance measurements, FTIR and SIMS analysis showed that the boron supply from the SOD changed as it was converted from a gel into a glass layer. Therefore, in order to use one B153 dopant source to dope several wafers repeatably, it was necessary to initially convert the borosilicate SOD gel into a thermodynamically stable BSG dopant source.

The B153 SOD was cured in a RTP to convert it from a gel to a BSG layer. FTIR and spectroscopic ellipsometry were used to determine the structural and compositional changes that transpired during the thermal treatment. FTIR showed that for a given temperature, the B-O stretching vibration band shifts to a lower frequency as the heating time is increased. This is a characteristic change that occurs on densification of a borosilicate glass. The B-O-Si bending vibration band, which is a characteristic band that appears on the formation of BSG, was also observed using FTIR. The spectroscopic ellipsometry measurements were in excellent agreement with the FTIR results. The refractive index increased with time and reached a maximum value as the BSG layer was formed. Sheet resistance measurements, FTIR and spectroscopic ellipsometry demonstrated that a RTC process step of 900 °C for 45 s converts the SOD gel to a BSG layer efficiently without any loss of boron from the layer. A single BSG dopant source wafer was successfully used to dope two wafers equally.

B_2O_3 is transferred to the product wafer during proximity RTD. SEM studies show that the B_2O_3 reflows during proximity RTD providing excellent step coverage. This is a very important characteristic, since the boron diffused junction must align with the gate edge.

Bare and oxidised (120 Å oxide) silicon wafers were doped using a BSG source during proximity RTD. The doping efficiency was observed to improve with temperature and time as the dopant supply at the silicon surface increased. Boron profiles with a surface concentration of $1.5x10^{19}$ cm^{-3} and a depth of 686 Å at 10^{18} cm^{-3} were fabricated in the bare silicon wafers. Boron profiles with a surface concentration of $1.2x10^{20}$ cm^{-3} and a depth of 308 Å at 10^{18} cm^{-3} were fabricated in the oxidised wafer. Source/drain junctions with depths < 400 Å and surface concentrations of 10^{20} cm^{-3} are required for 0.1 μm technology. The profile in the oxidised wafer satisfies these requirements, thereby establishing the suitability of proximity RTD for ultra-shallow LDD device fabrication.

Boltzmann-Matano analysis showed that boron diffusion was enhanced for RTD into bare Si, due to the injection of silicon interstitials into the lattice during RTD. Enhanced diffusion was not observed for RTD into the oxidised wafer, since the oxide layer traps

the generated interstitials. It was postulated that retarded diffusion at high boron concentrations might be caused by interstitial-induced boron clustering.

The sheet resistance and boron profile depth combinations obtained with proximity RTD were comparable to the best results obtained with other doping techniques.

P-n junction diodes were fabricated using proximity RTD (5 s at 1050 °C). The diodes displayed excellent *IV* characteristics. The typical turn-on voltage for all of the diodes was ~0.6 V. The ideality factors varied between 1.05 and 1.2, depending on the diode type. Under reverse bias, the diodes had sharp avalanche breakdown with $V_{BD} > -19$ V. The leakage current observed was very low, rising only to 50 pA as the diodes approached breakdown. These characteristics indicate that the silicon is of very high quality, and was not degraded by the proximity RTD process.

Micro-Raman spectroscopy, synchrotron X-ray topography and optical microscopy were used to determine the thermal stress and the boron-induced strain that is introduced into silicon wafers during proximity RTD. Thermal stress increases with time spent at the peak temperature. Slip lines were observed in oxidised wafers after 166 s at 1050 °C, and after 16 s at 1050 °C in boron doped wafers. Boron aggravates the stress that is induced in the wafer. A dense array of misfit dislocations was observed in heavily doped boron wafers. Surface strain was observed in wafers after 5 s at 1050 °C, however there was no evidence of any dislocations. Similarly, the electrical characteristics of the *p-n* junction diodes that were subjected to proximity RTD of 1050 °C for 5 s showed no evidence of any stress-induced defects.

9.2 Future Work

Proximity RTD has successfully been used to fabricate high-quality shallow boron junctions that satisfy requirements for 0.1 μm technology. This highlights the suitability of proximity RTD for deep submicron technology. However, several areas still require further work.

In particular, the boron diffusion mechanism is not completely understood, especially RTD through thin oxides. This lack of understanding is further complicated due to the difficulty in fabricating uniform high-quality thin oxides. Variations in the oxide thickness will be magnified in the junction depth. Therefore, before detailed investigation into RTD through thin oxides can be considered, the fabrication of thin oxides must be optimised.

Once the boron diffusion mechanism is fully understood, it would be extremely useful to develop a program that can accurately model the process.

As wafer size continues to increase, the thermal uniformity and induced stress across the wafer continue to be a critical issue. Therefore, if proximity RTD is to be used to dope $\geq$ 8 inch wafers, it will be necessary to do a more detailed stress analysis on silicon wafers after RTD and RTO. In particular, it would be interesting to determine when the boron doping profile begins to cause defects that will degrade device performance.

Appendix A

Mask Set for *p-n* Junction Diodes

A four-mask process was used to produce the *p-n* junction diodes that were described in Chapter 6. The mask sets for the Surround Diodes, Open Diodes type 1 and Open Diodes type 2 are shown in Figures A.1, A.2 and A.3, respectively. Each of these figures has 50 micron grid points included. Each mask is highlighted in a different colour:

- Mask 1 (red): opens the deep boron-diffused windows.
- Mask 2 (green): opens the shallow boron-diffused windows.
- Mask 3 (black): opens the contact windows.
- Mask 4 (blue): patterns the aluminium metallisation layer.

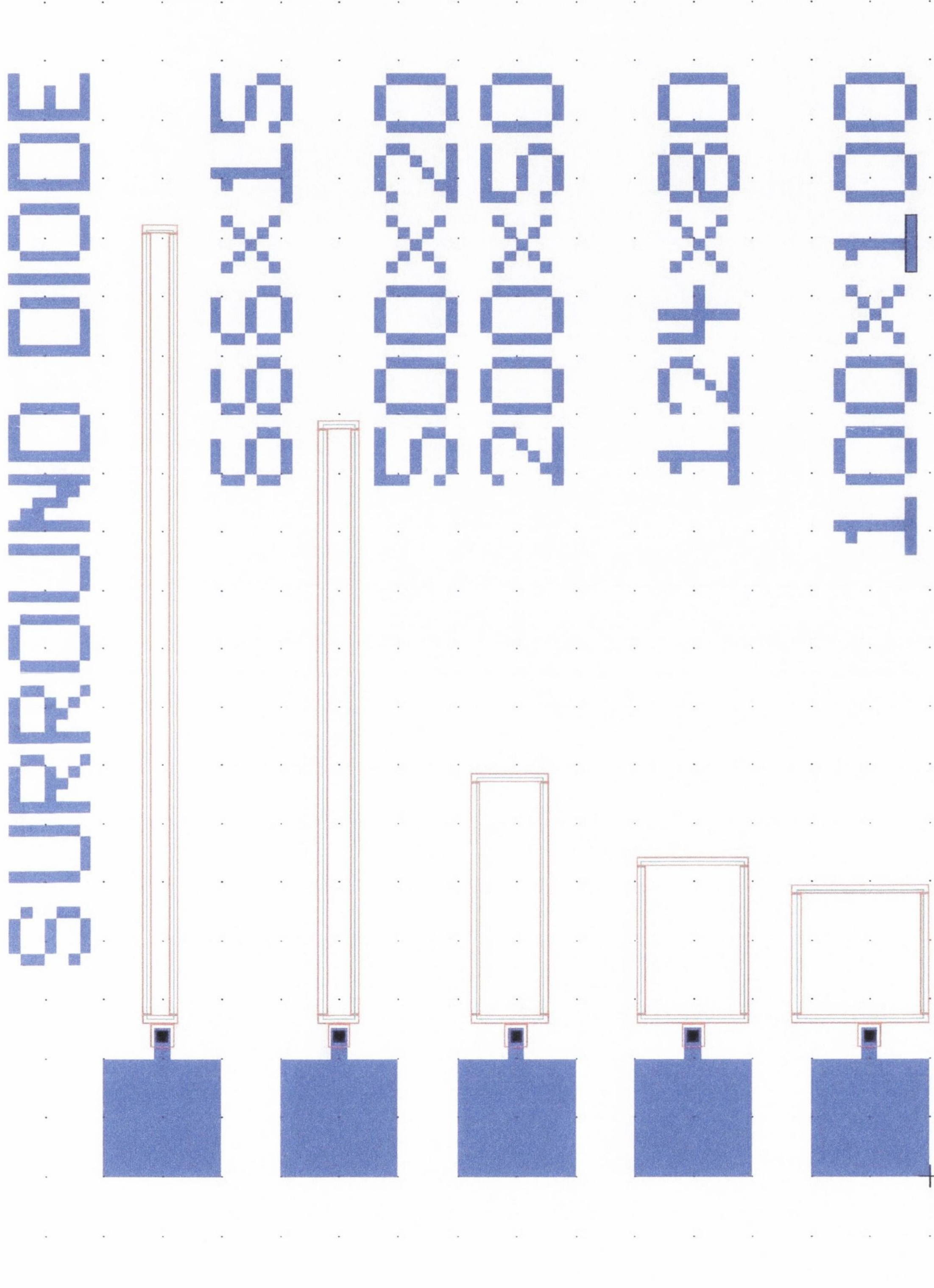


Figure A.1. Mask set for Surround Diodes.

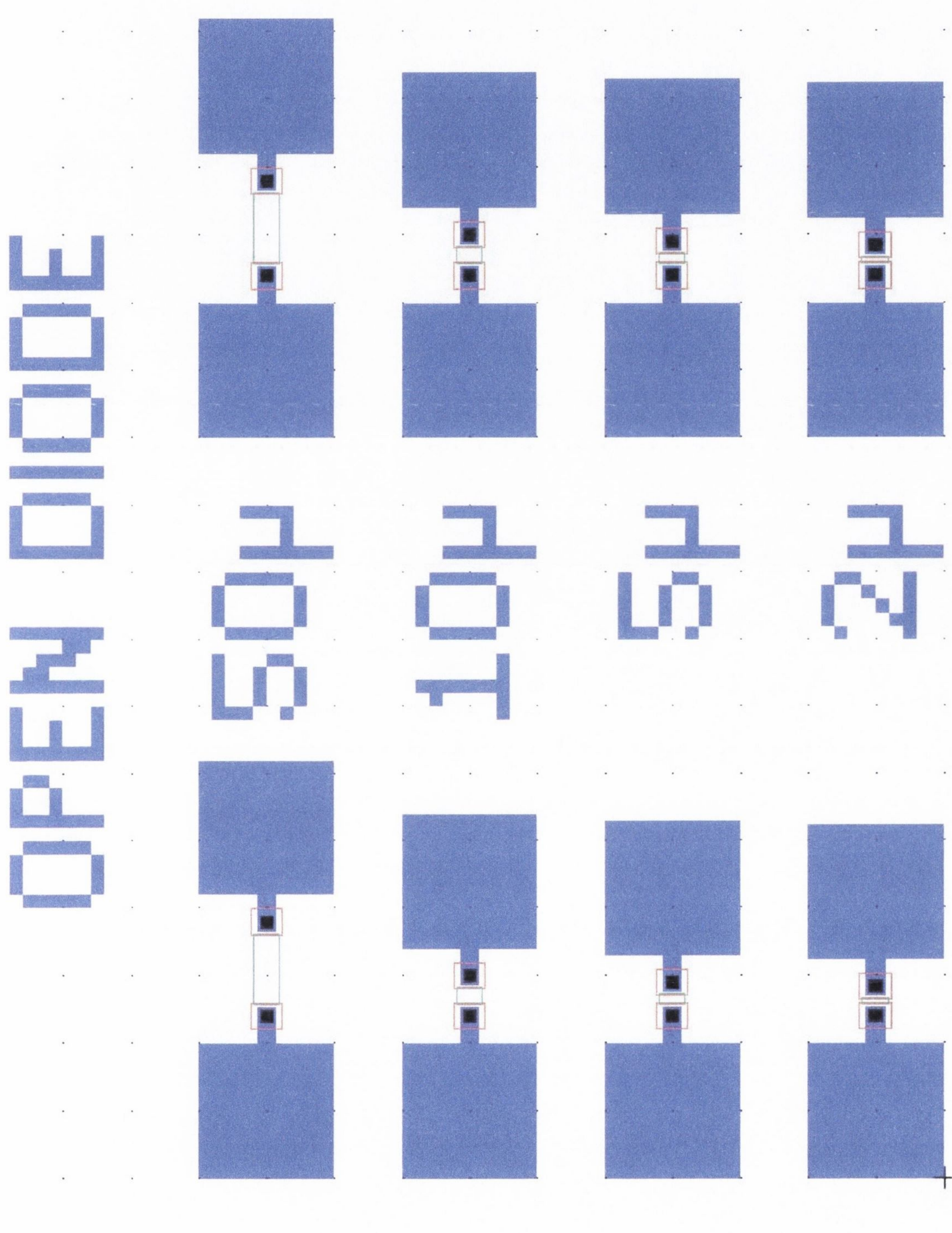


Figure A.2. Mask set for Open Diodes type 1.

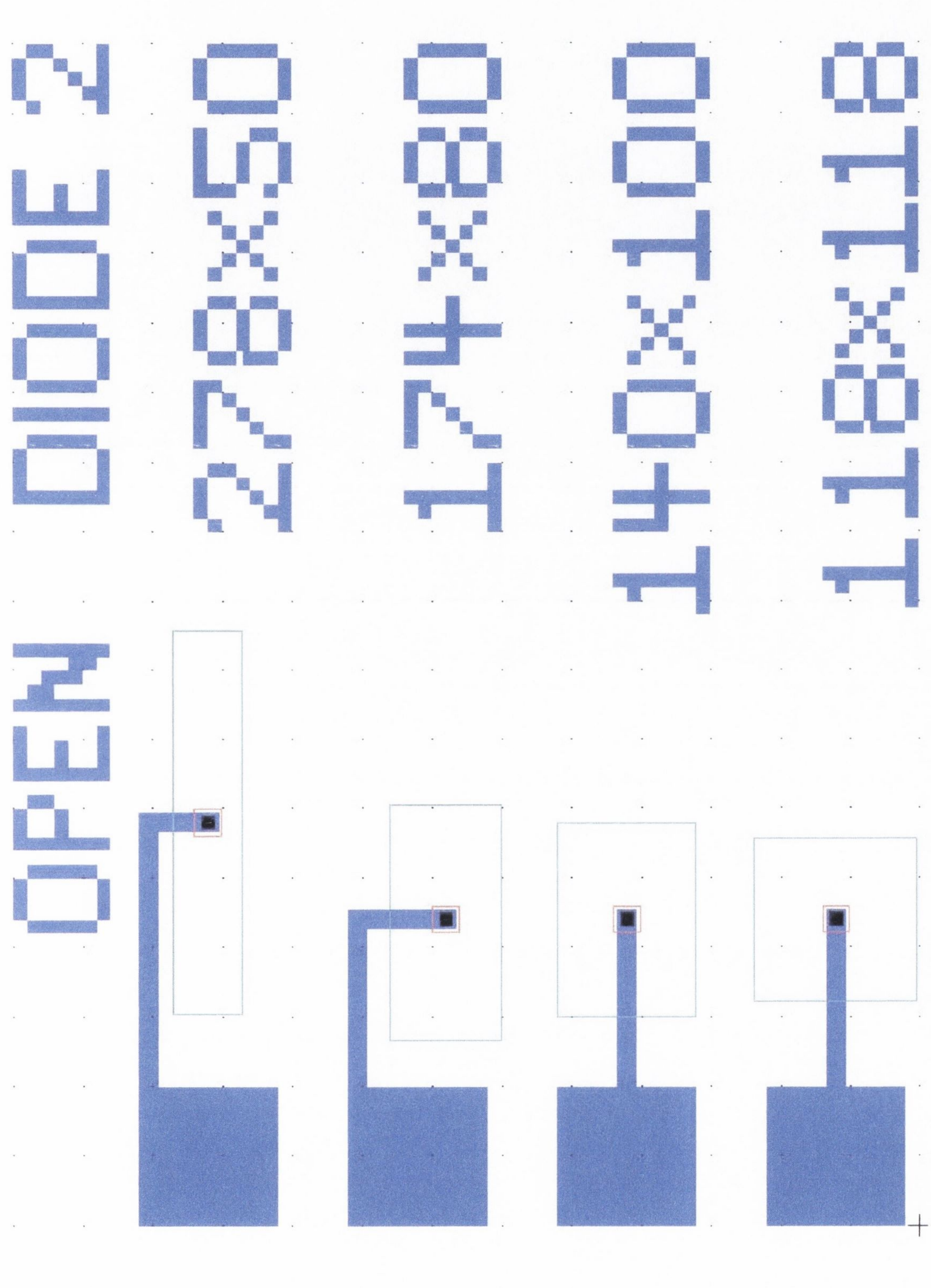


Figure A.3. Mask set for Open Diodes type 2.

Appendix B

Process Flow for *p-n* Junction Diodes

1. Czochralski grown 4-inch *n*-type, <100> oriented, 9-16 Ω cm resistivity silicon wafers were used throughout this study. The as-received wafers were labelled D4 and D5. The wafers were cleaned before any processing was performed.
2. The substrate was wet oxidised at 1150 °C for 12 mins to produce a masking oxide layer thickness of 0.4 μm.
3. Mask 1 opened the deep boron-diffused windows. CF_4/CHF_4 RIE was used to transfer the pattern to the oxide. The photoresist was removed using fuming nitric acid, and the wafers were cleaned before proceeding to the next process step.
4. Predeposition: a boron nitride disk was used as the dopant source during predeposition,
    - Ramp up: 800-1000 °C in 5% O_2:95% N_2
    - Predeposition: 1000 °C for 15 mins in 5% O_2:95% N_2, followed by 10 mins in 100% O_2
    - Ramp-down: 1000-800 °C in 5% O_2:95% N_2
5. Drive-in:
    - Ramp-up: 800-1000 °C in 5% O_2:95% N_2
    - Drive-in: 1000 °C for 1 min in 5% O_2:95% N_2, 49 mins in 100% N_2, and 10 mins in 100% O_2
    - Ramp-down: 1000-800 °C in 5% O_2:95% N_2
6. Wafers D4 and D5 were etched in 10:1 H_2O:HF for 2 mins 10 s to remove the boron glass from the diffusion windows. The wafers were then cleaned. The wafers were dehydrated at 400 °C for 20 mins before the next photolithography process step.
7. Mask 2 opened the shallow boron-diffused windows. CF_4/CHF_4 RIE was used to transfer the pattern to the oxide. The photoresist was removed using fuming nitric acid, and the wafers were cleaned before proceeding to the next process step.

8. Wafers D4 and D5 were dry oxidised at 900 °C for 5 mins in 50% O_2:50% N_2 to grow a 120 Å thick oxide over the shallow boron-diffused windows. Wafer D5 was then etched in 10:1 H_2O:HF to remove this thin oxide from the diffusion window. Both wafers were dry oxidised to minimise process variations.

9. Proximity RTD in 25% O_2:75% N_2:

- Ramp-up: 0-1050 °C at 48 °C/s
- RTD: 1050 °C for 5 s
- Ramp-down: 1050-800 °C at 60 °C/s

10. Wafers D4 and D5 were etched in 10:1 H_2O:HF for 70 s and 60 s respectively, to remove the boron glass from the diffusion windows. The wafers were cleaned.

11. A 0.2 μm LPCVD oxide was deposited from a TEOS source at 720 °C for 20 mins.

12. Mask 3 opened contact windows. Buffered HF was used to transfer the pattern to the oxide layer. The photoresist was removed using fuming nitric acid, and the wafers were cleaned before proceeding to the next process step.

13. The wafers were annealed in forming gas at 450 °C for 20 mins to passivate the dangling bonds. Normally, this step is performed after the aluminium metallisation, however, in this process, the forming gas anneal is performed before the metallisation to avoid aluminium spiking through the shallow *p*-type junctions.

14. The wafers were given a 30 s dip in 50:1 H_2O:HF prior to metallisation. Approximately 0.2 μm layer of aluminium was evaporated onto the wafer followed by 0.5 μm layer of sputtered aluminium to improve step coverage.

15. Mask 4 patterned the aluminium metallisation layer.

16. Finally, HF was used to remove the oxide from the backside of the wafer, and 0.1 μm layer of aluminium was evaporated onto the backside to improve the backside contact during electrical testing.