ENGINEERING OF SILICON AND GERMANIUM TUNNEL DIODES FOR INTEGRATED CIRCUIT APPLICATION

A Dissertation

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jialin Zhao, B.E., M.S.E.E.

Alan C. Seabaugh, Director

Graduate Program in Electrical Engineering

Notre Dame, Indiana

April 2007
ENGINEERING OF SILICON AND GERMANIUM TUNNEL DIODES FOR INTEGRATED CIRCUIT APPLICATIONS

Abstract
by
Jialin Zhao

In recent years, the tunnel diode has attracted interest from companies and researchers. Integrating tunnel diodes with transistors provide input-out isolation, gain and fan-out ability which the tunnel diode by itself lacks. The sparsity of tunnel diode fabrication processes compatible with transistor processing hinders the wider use of the device. Fabrication processes, which could be applied to integrated tunnel diode/transistor circuits on Si and Ge, are explored in this work.

Silicon tunnel diodes were demonstrated in both vertical and lateral geometries using spin-on diffusants and rapid thermal processing. Silicon tunnel diodes were first formed in the substrate plane through an oxide window process, with peak current densities of approximately 1 µA/µm² and peak-to-valley ratio of approximately 1.3. A self-aligned lateral fabrication process, which forms the junction perpendicular to the substrate plane, has also been successfully developed and yielded backward Si tunnel diodes with peak current densities of 30 nA/µm². To the author’s knowledge, these accomplishments are the first demonstration of lateral Si tunnel diodes using spin-on diffusants and rapid thermal processing. Low current density tunnel diodes can find
applications as zero biased detector.

Germanium tunnel diodes were demonstrated both using a diffusion-based approach and an on-wafer liquid-phase regrowth approach. The diffusion-based approach utilized spin-on diffusants and rapid thermal processing. Germanium tunnel diodes with current densities up to 0.6 nA/µm² and PVR of 1.1 were demonstrated for the first time using this approach. An on-wafer liquid-phase regrowth approach with a silicon nitride microcrucible was developed. Germanium TDs with current densities up to 1.2 mA/µm² were demonstrated. A primary goal of this project, demonstration of a 1 mA/µm² tunnel junction, was fulfilled.
## CONTENTS

FIGURES ........................................................................................................................................... iv

TABLES .............................................................................................................................................. x

ACKNOWLEDGEMENT ........................................................................................................................ xi

CHAPTER 1 INTRODUCTION ................................................................................................................ 1

1.1 Motivation ................................................................................................................................... 1
1.2 Prior art ...................................................................................................................................... 2
1.3 New approaches for tunnel junction formation ........................................................................ 4
1.4 Silicon vs. germanium ........................................................................................................... 6
1.5 Applications ........................................................................................................................... 7

CHAPTER 2 TUNNEL-JUNCTION TRANSPORT AND FORMATION PHYSICS .... 11

2.1 Interband tunneling current-voltage relations ........................................................................ 12
2.2 Tunnel-junction formation ...................................................................................................... 21
   2.2.1 Doping technologies .................................................................................................. 21
   2.2.2 Heavy doping ......................................................................................................... 23
   2.2.3 Abrupt junction formation .................................................................................... 27
2.3 Tunneling current density vs. junction doping density and gradient ...................................... 29
2.4 Switching speed .................................................................................................................... 34
2.5 Curvature coefficient of backward tunnel diode ................................................................ 36

CHAPTER 3 SILICON TUNNEL DIODE FABRICATION ......................................................... 41

3.1 Spin-on diffusants and rapid thermal processing .................................................................. 42
3.2 Diffused-junction tunnel diodes in a vertical geometry ......................................................... 44
   3.2.1 Vertical process flow ............................................................................................... 44
   3.2.2 Demonstration of diffused-junction silicon tunnel diodes ................................... 47
3.3 Lateral diffused-junction tunnel diodes ................................................................. 53
   3.3.1 Lateral process flow and mask set ................................................................. 53
   3.3.2 First demonstration of lateral diffused-junction silicon tunnel diode .......... 59
3.4 Limitations of the diffusion approach .................................................................. 66

CHAPTER 4 GERMANIUM TUNNEL DIODE FABRICATION ........................................ 68
4.1 Prior art vs. new approaches .................................................................................. 69
4.2 Demonstration of diffused-junction Ge TDs ......................................................... 71
   4.2.1 Doping Ge from spin-on diffusants ............................................................... 71
   4.2.2 Diffused Ge tunnel diodes process ................................................................. 77
   4.2.3 Diffused-junction Ge tunnel diodes ............................................................... 81
4.3 Demonstration of on-wafer liquid-phase-epitaxy Ge tunnel diode process ....... 90
   4.3.1 On-wafer liquid-phase-epitaxy approach ...................................................... 90
   4.3.2 Process flow .................................................................................................. 94
   4.3.3 Demonstration of Ge TDs using on-wafer liquid-phase regrowth approach ... 95
4.4 Demonstration of Ge TDs exceeding 1 mA/µm² current density ....................... 101
4.5 Limitations of approach ....................................................................................... 109

CHAPTER 5 CONCLUSIONS ....................................................................................... 110
5.1 Summary of accomplishments ............................................................................. 110
5.2 Achievements relative to prior art ....................................................................... 111
5.3 Recommendation for further study ..................................................................... 112

APPENDIX 1 VERTICAL SILICON TD FABRICATION TRAVELER ......................... 115

APPENDIX 2 LATERAL SILICON TD FABRICATION TRAVELER ......................... 119

APPENDIX 3 GERMANIUM TUNNEL DIODE FABRICATION USING
LIQUID-PHASE REGROWTH APPROACH TRAVELER .......................................... 126

APPENDIX 4 RAPID MELT GROWTH OF TUNNEL JUNCTIONS ....................... 129

REFERENCES .......................................................................................................... 153
FIGURES

Figure 1.1 Schematic current-voltage characteristics of tunnel diode, backward tunnel diode and conventional $p$-$n$ diode. ................................................................. 8

Figure 2.1 Schematic tunnel diode current-voltage curve with corresponding energy band diagrams. Shaded areas represent the filled electron states in the degenerately doped semiconductor. ........................................................................................................... 12

Figure 2.2 Triangular potential energy barrier. ........................................................................................................... 15

Figure 2.3 Tunneling current density vs. electric field for a Si $p$-$n$ junction comparing predicted by Eq. 2.5 with experimental data: $\Delta$ Stork and Isaac [48] and $\Delta$ Fair and Wivell [26]. ....................................................................................................................... 17

Figure 2.4 Schematic current-voltage characteristic and energy band diagram of the tunnel diode biased at $V_P = (V_{dn} + V_{dp})/3$. ..................................................................................................................... 18

Figure 2.5 Peak tunneling current density and internal electric field vs. effective doping density for an abrupt $p$+$n$+ junction ........................................................................................................ 20

Figure 2.6 SIMS profile of B diffusion into Si using SOD and RTP at 900 °C of 1 s. The gradient of diffusion is estimated at the concentration of $7 \times 10^{20}$ cm$^{-3}$, which is the solid solubility of P at 900 °C. ......................................................................................................................... 28

Figure 2.7 (a) Computed band diagram and (b) junction electric field for four doping profiles: an abrupt profile and 1, 2, and 4 nm/decade gradients from $p$ to $n$ with doping density of $1 \times 10^{20}$ cm$^{-3}$. The calculation was made using W. R. Frensley’s BandProf.... 30

Figure 2.8 Junction electric fields vs. position for various doping abruptness and effective doping densities, extracted from the gradient of simulated energy band profiles using W. R. Frensley’s program BandProf........................................................................................................... 31

Figure 2.9 Peak tunneling current density vs. effective doping density with doping abruptness of 0, 1, 2, and 4 nm/decade in Si and Ge. ........................................................................................................ 32
Figure 2.10 Equivalent circuit of a diode detector [34].

Figure 2.11 Zero-bias curvature coefficient vs. effective doping density for abrupt Si and Ge backward TDs. The impurity densities on either side are assumed to be equal. The curves labeled with Si\textsubscript{1} and Ge\textsubscript{1} represent the first term of Eq. (2.11), and the curves labeled with Si\textsubscript{2} and Ge\textsubscript{2} represent the second term of Eq. (2.11).

Figure 3.1 Schematic diagram of the proximity RTP, showing the source wafer with SODs placed on top of the device wafer and separated by quartz spacers.

Figure 3.2 Scale drawing in the vertical direction only, of a vertical Esaki tunnel diode process: (a) following thermal oxidation, and (b) after oxide patterning.

Figure 3.3 Esaki tunnel diode process (continued): (a) after P diffusion, and (b) after P SOD residue cleaning which also thins the oxide.

Figure 3.4 Esaki tunnel diode process (continued): (a) after B diffusion, (b) after B SOD residue cleaning, and following (c) Al metallization.

Figure 3.5 Current–voltage characteristics of Si TDs fabricated using an oxide window process with SODs and proximity RTP.

Figure 3.6 Current-voltage characteristics of Si TDs fabricated using SODs and proximity RTP, with direct contact of Borofilm100 SOD.

Figure 3.7 VASE measurement of SOD residue layer on the silicon surface after proximity RTP and 0, 1, 2 and 3 minutes etch in buffered HF.

Figure 3.8 Current-voltage characteristics of Si TDs fabricated using SODs and proximity RTP, using B diffusion ramp up rate of 60 °C/s.

Figure 3.9 Current-voltage characteristic and curvature coefficient of Si backward TD fabricated by SODs and RTP approach.

Figure 3.10 Scale drawing in the vertical dimension of a lateral TD process: (a) Si substrate with thermally-grown oxide layer, patterned and etched, (b) after P proximity RTP, and (c) following W sputtering.

Figure 3.11 Lateral Si TD process continued: (a) chemical mechanical polishing of tungsten, (b) remove oxide, and (c) boron proximity RTP.
Figure 3.12 Lateral Si TD process continued: (a) PECVD nitride deposition and via patterning, (b) Al evaporation and contact formation, and (c) equivalent circuit of fabricated devices............................................................................................................. 56

Figure 3.13 Schematic layouts of lateral TDs and test structures. Test structures and tunnel diode devices are as indicated............................................................................................................. 57

Figure 3.14 Typical tunnel diode $I_{pnp}$ 8 x 40 layout and the schematic cross-section. ... 58

Figure 3.15 (a) dc and (b) microwave measurement structure layout. ......................... 59

Figure 3.16 SEM pictures of sputtered tungsten film on oxide and silicon before CMP. 60

Figure 3.17 SEM pictures of sputtered tungsten film on silicon after CMP ............... 61

Figure 3.18 SIMS profile of P diffusion into silicon using SOD and RTP at 900 °C of 1 s and 1000 °C of 1 s. The diffusion gradient is given at the concentration of $2 \times 10^{20} \text{cm}^{-3}$, which is the solid solubility of B at 1000 °C. ................................................................. 62

Figure 3.19 Schematic drawings of lateral junction formation process and a SEM picture at the left W corner: (a) after P diffusion, (b) after HF cleaning, W sputtering and CMP, (c) SEM picture of the left W corner corresponding to (b), and (d) after B diffusion. ........ 63

Figure 3.20 $I-V$ curves of a Si lateral tunnel junction at various temperatures between -40 and 80°C. The device is measured from the left contact to the middle contact, as illustrated in the inset ................................................................. 65

Figure 4.1 Scale drawing showing a cross section of the first step in a Ge TD process using SODs and RTP: (a) an $n^+$ P-doped layer is diffused into a $p$-Ge substrate, (b) following a Ga diffusion, and (c) after Al evaporation and etching. ......................... 79

Figure 4.2 (a) Schematic mask layout for Ge TD fabrication using SODs and RTP process, (b) equivalent circuit of measured devices, with device size symbolizing area difference between diode 1 and 2 ....................................................................................................................................... 80

Figure 4.3 Device current-voltage characteristics after three Ga proximity RTP processes at 750, 825 and 900 °C for 1 s, respectively, with (a) current plotted on logarithm scale, and (b) current plotted on linear scale. 750 °C’s result is reduced by 1000x................. 82

Figure 4.4 (a) Germanium TD $I-V$, and (b) good agreement achieved between the extracted conductance and the derivative of dc $I-V$ measurement........................................ 84
Figure 4.5 Real and imaginary part of the impedance of a diffused Ge TD at room temperature, (○) represents the imaginary part and (●) represents the real part; the solid and dashed curves are calculated model fits to the tunnel diode equivalent circuit, inset. 85

Figure 4.6 Inverse junction capacitance squared vs. bias for the diffused Ge TD. 86

Figure 4.7 I-V characteristic and zero-bias curvature coefficient of a Ge TD fabricated by RTP from SODs. 88

Figure 4.8 Germanium-aluminum phase diagram from Okamoto [110]. 91

Figure 4.9 The ratio of the penetration depth, $x$, to the Al deposition thickness, $t_{Al}$, as a function of peak anneal temperature, $T$, and the acceptor doping density as a function of the anneal temperature as determined from the Al-Ge binary phase diagram [110]. 92

Figure 4.10 Phosphorus and Al diffusion in Ge, calculated using $\sqrt{Dt}$. 93

Figure 4.11 Schematic device cross section of the rapid melt growth process drawn to scale in the vertical direction. (a) Germanium is doped $n^+$ by rapid thermal diffusion of phosphorus followed by aluminum deposition, patterning, and (b) etching. (c) At a peak anneal temperature of 600 ºC with a 100 nm Al cap, 110 nm of Ge is consumed. (d) On cooling, a heavily-doped $p^+$ Ge layer of 60 nm is grown. 94

Figure 4.12 Schematic layout for Ge TDs using on-wafer liquid-phase regrowth process. The spacing of square structure to the outside rectangular feature is 2, 4, 10 and 20 µm, as labeled on the top. 95

Figure 4.13 I-V characteristics of Ge tunnel diodes vs. junction area, (a) using a 440 ºC, 5 s anneal, and (b) I-V characteristic and curvature coefficient. 96

Figure 4.14 I-V characteristics of Ge tunnel diodes vs. junction area, (a) using a 700 ºC, spike anneal, and (b) peak current and PVR vs. device area. 97

Figure 4.15 Dependence of Ge TD current-voltage characteristics on rapid-melt-growth anneal condition. P was diffused at 800 ºC for 300 s with direct contact. Al of 100 nm was deposited. The Al-doped regrowth layers were regrown by annealing at 600 or 700 ºC with anneal time of spike, 1 s or 4 s. The 700 ºC, spike annealed sample was damaged during processing and therefore not measured. 98

Figure 4.16 Dependence of current density in the forward, $I_F$, and reverse current, $I_R$, directions as a function of anneal condition for the uncapped Ge diodes of Fig. 4.16. In
the reverse direction, the current magnitude is plotted at a voltage corresponding to minus the forward peak voltage, $V_p$.

Figure 4.17 Top view optical micrographs of Al/Ge mesa structures without a nitride cap, before (a) and after (b) a 440 °C, 5 s regrowth anneal. The Al metal is 100 nm thick.

Figure 4.18 Scale drawing of the device cross section of a vertical Ge TD using on-wafer liquid-phase regrowth with a silicon nitride microcrucible approach.

Figure 4.19 Top view optical micrographs of Al/Ge mesa structures after a 440 °C, 5 s regrowth anneal in a forming gas ambient with (a) a 50 nm oxide, and (b) 250 nm oxide.

Figure 4.20 Top view optical micrographs of Al/Ge mesa structures after a 440 °C, 5 s regrowth anneal in a forming gas ambient with (a) a 50 nm nitride, (b) 50 nm nitride, 10 x 10 and 5 x 5 µm$^2$ devices, and (c) 250 nm nitride. The Al metal is 100 nm thick.

Figure 4.21 Germanium TD peak current density vs. annealing temperatures. Phosphorus was diffused at 800 °C for 300 s with direct contact. Al of 100 nm was deposited. Silicon nitride layer of 50 nm was deposited by PECVD.

Figure 4.22 Current-voltage characteristics of Ge TDs fabricated by the on-wafer liquid-phase regrowth process with 25 nm mesa and 50 nm silicon nitride microcrucible layer. The junction formation regrowth was at 600 °C, 2 s and a ramp up rate of 150 °C/s.

Figure 5.1 Germanium-gallium phase diagram from Okamoto [110].

Figure A.1 Al-Ge phase diagram after Okamoto [11].

Figure A.2 Schematic device cross section of the rapid melt growth process drawn to scale in the vertical direction. (a) Germanium is doped $n^+$ by rapid thermal diffusion of phosphorus followed by aluminum deposition, patterning, and etching. (b) A plasma nitride cap is applied. (c) At the peak anneal temperature (600 °C in this example) an AlGe melt forms with a meltback of 110 nm (d) On cooling, a heavily-doped $p^+$Ge layer is grown; on cooling from a peak temperature of 600 °C, a 60 nm layer is grown.

Figure A.3 The ratio of the penetration depth, $x$, to the aluminum deposition thickness, $t_{Al}$, as a function of peak anneal temperature, $T$, and the concentration of Al, $N_{Al}$ in cm$^{-3}$ as a function of the peak anneal temperature as determined from the Al-Ge binary phase diagram [10].
Figure A.4  Dependence of Ge tunnel diode current-voltage characteristics on rapid-melt-growth anneal condition. The silicon nitride cap was omitted for this set of anneals. Solid lines are used for the 600 °C anneals and dashed lines for the 700 °C anneals; the symbols in the legend and on the curves indicate the hold time at the peak anneal temperature. ................................................................. 147

Figure A.5  Dependence of current density in the forward, $I_F$, and reverse current, $I_R$, directions as a function of anneal condition for the uncapped Ge diodes of Fig. 4. In the reverse direction, the current magnitude is plotted at a voltage corresponding to minus the forward peak voltage, $V_P$. ................................................................. 148

Figure A.6  Peak-to-valley current ratio vs. anneal condition for the uncapped Ge diodes of Fig. A.5. ........................................................................................................ 149

Figure A.7  Optical micrographs showing the Ge wafer surface following rapid melt growth of $p^+$Ge at a peak temperature of 440 °C................................. 150

Figure A.8  Germanium tunnel diode current densities vs. annealing temperature with a 50 nm SiN cap. The solid circles denote measurements at the peak voltage in the forward current direction; the open circles are measurements in the reverse current direction at -0.01 V. The values for the peak-to-valley-current ratio ($PVR$) are given on the forward current curves. ........................................................................................................ 151

Figure A.9  Current-voltage characteristics of Ge tunnel diodes with a 50 nm silicon nitride microcrucible layer and a 150 °C/s heating rate................................. 152
TABLES

TABLE 1.1  Germanium and silicon properties at room temperature ........................................ 6

TABLE 2.1  Summary of the highest measured electrically-active dopant concentration of As, P and B in Si using various approaches ................................................................. 25

TABLE 2.2  Summary of the highest measured electrically-active dopant concentration of B, Al, Ga, P and As in Ge using various approaches ................................................... 25

TABLE 2.3  Record doping gradient for boron diffusion in silicon extracted from SIMS at the solid solubility of the opposing dopant ................................................................. 28

TABLE 2.4  Group IV tunnel diodes properties vs. fabrication approach, selected by highest peak current density or peak-to-valley ratio ................................................................. 35

TABLE 4.1  Summaries of measured doping types, Hall mobilities $\mu_{\text{Hall}}$, sheet carrier concentrations $n_{\text{sheet}}$, calculated diffusion depths $\sqrt{Dt}$ ($t$ is the hold time in RTP), computed carrier densities $n_{\text{Hall}}$ from sheet carrier concentration, $n_{\text{Fistul}}$ from Hall mobility, solid solubilities ($n_{\text{ss}}$) and Ge surface roughness (SR) under different annealing temperatures and times ................................................................. 74

TABLE 4.2  Summaries of measured doping types, mobilities $\mu_{\text{Hall}}$, sheet carrier concentrations $n_{\text{sheet}}$, calculated diffusion depths $\sqrt{Dt}$ ($t$ is the hold time in RTP), computed carrier densities $n_{\text{Hall}}$ from sheet carrier concentration, $n_{\text{Fistul}}$ from Hall mobility, and solid solubilities ($n_{\text{ss}}$) under different annealing temperatures and times ................................................................. 77

TABLE 4.3  Summaries of zero bias curvature coefficient $\gamma$, junction resistance $R_j$, junction capacitance $C_j$, cutoff frequency $f_{\text{cut}}$ and low frequency sensitivity $\beta$ of Si backward diodes in Ref. [35] and this work (section 3.2), and of Ge backward diodes from M-Pulse Microwave Inc. [99] and this work ................................................................. 89

TABLE 4.4  Summaries of the best peak current densities and PVR of Ge TDs fabricated by the on-wafer liquid-phase regrowth approach under various annealing conditions ................................................................................. 100
ACKNOWLEDGEMENT

I would like to thank many people who have supported and helped me to finish my research at Notre Dame and this dissertation.

First, I wish to express my sincere gratitude to my advisor, Dr. Alan C. Seabaugh, for his valuable suggestions and directions concerning my work and life during these years. He gave me endless encouragement and support when I need both in the research work and personal life. He taught me the right attitude towards research and life, which I will benefit for the rest of my life. I am grateful to be his student.

Second, I would like to thank Dr. Bernstein, Dr. Fay, and Dr. Kosel for reading my dissertation and being members of my dissertation defense board. I am very grateful to them for giving me a lot of valuable advice and help on my research during these years.

Finally, I am very grateful to my parents and my girl friend for their encouragement and support. I appreciate my friends and colleagues, Yan Yan, Qingmin Liu, Sajid Kabeer, Bin Wu, Zoltan Racz, Wei Zhao, Surajit Sutar, Dane Wheeler, Qin Zhang, and Notre Dame colleagues, for their help and for the football days and rafting days we had.
1.1 Motivation

In 1957, while investigating internal field emission in heavily-doped Ge $p$-$n$ junctions, Leo Esaki observed, for the first time, the negative differential resistance (NDR) characteristic of the tunnel diode (TD) [1]. In recent years, the TD has attracted interest from companies and researchers [2-9]. The TD, as a two-terminal device, suffers from a lack of isolation between input and output terminals [10]. At the same time, the TD, when used alone, suffers from low voltage gain and fan-out [11]. Integrating TDs with transistors is necessary to provide input-output isolation, higher gain, and fan-out ability. Only discrete devices are now commercially available from five manufacturers: Aeroflex/Metelics, Inc. of Sunnyvale, CA, American Microsemiconductor Inc. of Madison, NY, M-pulse Microwave Inc. of San Jose, CA, US Microwaves, Inc. of Santa Clara, CA, and Solid States Devices, Inc. of La Mirada, CA. The incompatibility of tunnel diode fabrication processes with conventional transistor processing has been an obstacle to the TD’s wider use.

A batch fabrication process would enable the integration of TDs with transistor
technology and allow circuit explorations [2, 3]. Davis and Gibbons [12] developed an alloyed-junction process for Ge TDs in the 1960s. Berger et al. [7, 13] developed a process for molecular beam epitaxy (MBE) Si-based TDs, which can be used to make mixed-signal circuits. The object of this research has been to explore fabrication processes which could be applied to integrated tunnel diode/transistor circuits on Si and Ge. A particular challenge is to achieve high peak current density, exceeding 1 mA/μm², and this has motivated much of the effort. In recent years, two reports of high current density SiGe TDs have proved the feasibility of achieving > 1 mA/μm² tunnel junctions [7, 13]. At the same time, much lower current density devices still have applications as sensors, detectors, and memory.

1.2 Prior art

In the early 1960s, tunnel diodes were fabricated using processes including metal/semiconductor alloying [1, 12, 14], electrical forming [15], solution regrowth [16], and laser annealing [17]. Burrus, even in 1962 [15], demonstrated Ge interband TDs with peak current densities up to 1 mA/μm² using electrical forming, a method in which a current pulse through a Ga-coated probe is used to melt the Ga and form a $p^+$ regrown junction on a heavily $n^+$ doped Ge substrate. Franks fabricated Si interband TDs with peak current densities up to 10 μA/μm² using alloying [14]. Metal/semiconductor alloying, solution regrowth, laser annealing, and electrical forming, are fundamentally liquid-phase epitaxial (LPE) processes. The dopants together with the substrate are heated
and liquified in a near surface layer. Upon cooling, an epitaxial layer incorporating the dopant grows from the liquid. The difference in these approaches is primarily in the heating; alloying utilizes sources like tantalum strip heaters, laser annealing utilizes lasers, and electrical forming utilizes current pulses through attached wires.

Molecular beam epitaxy and ion implantation (II) have also been used for TD fabrication [4-7, 13, 18]. MBE forms the tunnel junction by ultrahigh vacuum elemental evaporation of semiconductors and dopants to achieve precise control of the junction doping concentration and profile. The advantages are high doping concentrations up to $10^{21}$ cm$^{-3}$ [19] and doping abruptness up to 2 nm/dec [20], which are the best among junction formation technologies. With these characteristics, MBE has produced the highest peak current densities up to 2.2 mA/µm$^2$ and peak-to-valley ratio (PVR) up to 6 for Si-based tunnel diodes [6, 7]. MBE fabricated Si/SiGe TDs are the only Si-based TDs that achieve tunneling current density above 1 mA/µm$^2$.

Ion implantation, the most widely utilized commercial doping method, has also been used to produce Si TDs [18]. Ion implantation has the advantage of its compatibility with Si complementary metal-oxide-semiconductor (CMOS) transistor processes. As the most commonly used doping technology in CMOS, ion implantation can incorporate TDs and circuits within Si CMOS without additional equipments or processes. Current densities of approximately 0.1 µA/µm$^2$ (10 A/cm$^2$) and PVR of approximately 2 have been achieved by this method. The lower current density is a result of the lower dopant activation and lesser dopant gradients achieved in II vs. MBE. Koga [18] et al. used
25-30 keV and relatively long annealing times (30 s at 1000 °C) to activate the implants. Higher current density can be expected with lower implantation energies, below 500 eV, and shorter activation anneals.

1.3 New approaches for tunnel junction formation

In this work, a diffusion-based method for junction formation has been explored, using spin-on diffusants (SODs) and rapid thermal processing (RTP) \[8, 21, 22\]. It is shown that device peak current densities achieved by this approaches are up to 2 \(\mu\text{A}/\mu\text{m}^2\), about 20x larger than that achieved by Koga using II \[18\]. The diffusion-based approach has the potential to be developed into a batch process and could find use in sensor/detector applications. Higher current densities are expected by this technique than have been achieved. The reasons for this are discussed in section 3.2.

In the diffusion-based process, phosphorus, boron and gallium-containing SODs have been utilized as the dopant sources. Upon high-temperature anneal, these dopants are released from the SODs to be incorporated into the semiconductor. Tunnel junctions are formed by RTP with annealing times on the order of seconds or less. A desirable attribute of this process is simplicity, while a limitation of the approach is that the diffused junction is not abrupt. Compared to II, more abrupt junctions may be expected using SODs since the dopants in II already have a spatial distribution before they are activated and annealing further distributes the profile.

The halogen lamp sources typically used in RTP are heated and cooled at roughly
100 °C/s rates. At these rates, the substrate is approximately in thermal equilibrium throughout the process, i.e. the wafer temperature is nearly uniform across the wafer volume. With the wafer temperature maintained near equilibrium, the maximum doping concentration obtainable is limited by the solid solubility of the desired impurity at the peak annealing temperature.

In this research, tunnel junction formation using SODs and RTP has been explored on both Ge and Si substrates. For Si, both lateral and vertical junctions were explored. In these processes, the tunnel junction is formed in the (100) surface plane of the wafer. This is called a vertical geometry process. The junction has also been formed perpendicular to the plane of the wafer and this is called a lateral geometry process.

An on-wafer liquid-phase regrowth approach has also been explored for Ge TDs achieving current densities exceeding 1 mA/µm². This approach for high current-density TD fabrication extends the alloying techniques of the 1960s [1, 12, 14]. In this approach, Ge is first heavily $n$-doped ($10^{19}$ to $10^{20}$ cm$^{-3}$) using an SOD film followed by RTP. The TD is then formed by depositing Al which serves both as a $p$ dopant and as a contact. According to the binary Al-Ge phase diagram, a predetermined amount of Ge is dissolved into the Al at the peak annealing temperature and this Ge then epitaxially regrows as a heavily-$p$ Al-doped ($10^{19}$ to $10^{20}$ cm$^{-3}$) layer on the $n+$ doped Ge layer.
1.4 Silicon vs. germanium

Germanium belongs to the same group in the periodic table as Si, and for all its chemical similarities, has several attributes which aid in the formation of higher current density junctions. Selected physical properties of Ge and Si are compared in Table 1.1. The smaller bandgap and lighter effective masses in Ge compared to Si, aid higher current densities while the higher permittivity of Ge over Si acts to reduce the junction internal field and current density. Both Ge and Si are indirect bandgap materials, but Ge has a second conduction band minimum at the zone center which can be populated at the doping densities used for TDs [23]. Direct tunneling provides higher current densities than indirect tunneling processes, which requires phonons to conserve momentum.

### TABLE 1.1

**GERMANIUM AND SILICON PROPERTIES AT ROOM TEMPERATURE**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Symbols</th>
<th>Ge</th>
<th>Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandgap (eV)</td>
<td>$E_g$</td>
<td>0.66 [24]</td>
<td>1.12 [24]</td>
</tr>
<tr>
<td>Dielectric permittivity</td>
<td>$\varepsilon_a$</td>
<td>16 [24]</td>
<td>11.8 [24]</td>
</tr>
<tr>
<td>Reduced tunneling effective mass</td>
<td>$m_{\text{eff}}$</td>
<td>0.08$m_0$ [25]</td>
<td>0.33$m_0$ [26]</td>
</tr>
<tr>
<td>Density of states conduction band (cm$^{-3}$)</td>
<td>$N_C$</td>
<td>$1.04 \times 10^{19}$ [24]</td>
<td>$2.8 \times 10^{19}$ [24]</td>
</tr>
<tr>
<td></td>
<td>$N_V$</td>
<td>$0.6 \times 10^{19}$ [24]</td>
<td>$1.04 \times 10^{19}$ [24]</td>
</tr>
<tr>
<td>Melting point ($^\circ$ C)</td>
<td>—</td>
<td>937 [24]</td>
<td>1415 [24]</td>
</tr>
<tr>
<td>Degenerate impurity density conduction band (cm$^{-3}$)</td>
<td>—</td>
<td>$8 \times 10^{18}$</td>
<td>$21 \times 10^{18}$</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>$4.5 \times 10^{18}$</td>
<td>$8 \times 10^{18}$</td>
</tr>
</tbody>
</table>
From the point of view of doping, Ge, with its smaller effective density-of-states compared with Si, becomes degenerate at a lower doping density, i.e. fewer impurities are needed to achieve degeneracy. At 300 K for Ge, donor and acceptor doping concentrations of approximately $8 \times 10^{18} \text{ cm}^{-3}$ and $4.5 \times 10^{18} \text{ cm}^{-3}$, respectively, are sufficient to align the conduction band and valence band edges with the Fermi level, which is defined as the degenerate impurity density in Table 1.1. The degenerate impurity densities for Ge are approximately 1/2 to 3/5 less than the degenerate donor and acceptor densities, respectively, in Si.

The melting point of Ge is 937 °C, about 1/3 lower than that of Si. Along with the lower melting point, annealing and diffusion temperatures tend to be significantly lower in Ge relative to Si. The thermal budget (temperature-time product) required for Ge TD fabrication will therefore be less than that for Si.

1.5 Applications

The current-voltage ($I-V$) characteristics of a TD, a backward TD, and a conventional $p-n$ diode are shown schematically in Fig. 1.1. The TD features a multi-valued $I-V$ characteristic and a negative differential resistance (NDR) region. The backward TD is a TD in which the forward tunneling current is selected to be at the degenerate concentration. Compared with the $p-n$ diode, the backward TDs has a reverse current higher than the forward current at low bias, showing a rectifying direction that is “backward” compared to that of a $p-n$ junction.
The NDR and high switching speed of TDs have been explored in pulse forming, triggering and sampling [27], analog-to-digital conversion (ADC) [28, 29], logic [30] and oscillator [31] applications. Tunneling-based diodes have also been used in memory [2, 3]. Tunnel diodes with current densities lower than 10 fA/µm² are desired for low power dissipation in tunnel-based SRAM [2].

The backward TD is of particular interest for millimeter-wave detection due to its natural rectification at zero bias [32, 33]. The rectifying characteristic of the backward TD at zero bias eliminates bias control circuits or local oscillators in detector applications [32]. Compared with $p$-$n$ diodes or Schottky diodes, backward TDs are more sensitive microwave detectors with lower temperature dependence and noise [34]. Silicon [35] and Ge [36] backward TDs have been demonstrated for microwave detection. III-V based backward TDs have also been explored for detection into the W-band [32, 33].
Three-terminal tunnel transistors based on the field-effect gating of interband tunnel current may offer a new application for tunnel junctions, in low-subthreshold-swing transistors [37]. The subthreshold swing is the gate voltage required to change the drain current by one decade and in field-effect transistors is limited to \( \ln 10 \left( \frac{kT}{q} \right) \) or 60 mV/decade at room temperature. This swing sets the minimum power supply voltages and the minimum power dissipation of a technology. Recently, several field-effect transistors based on the gating of interband tunneling current have been proposed which can achieve subthreshold swing below 60 mV/decade [37].

The remainder of the thesis is organized as follows. In chapter 2, TD physics and engineering are presented. Analytic expressions are derived to relate the tunnel current density to the junction doping concentration and abruptness. The physical requirements needed to achieve 1 mA/\( \mu \text{m}^2 \) current densities are discussed. Tunnel junction formation technologies are described and the best doping concentration and doping abruptness reported to date are summarized. For backward TD detectors, the curvature coefficient of the current-voltage curve at zero bias is a direct measure of the detection ability. Therefore the relation between the curvature coefficient and the junction doping is outlined and quantified. Approaches to achieve high detection sensitivity are discussed.

Experimental results for Si TDs formed using SODs and RTP are discussed in chapter 3. Silicon TDs, with current densities up to 1.5 \( \mu \text{A/\mu m}^2 \), have been achieved using an oxide window process. A lateral process has also been explored and developed for Si tunnel junction fabrication with current densities of 30 nA/\( \mu \text{m}^2 \) demonstrated using
a self-aligned tungsten chemical mechanical polishing (CMP) process.

Germanium TD fabrication is described in chapter 4. A diffused-junction based approach with SODs and RTP achieves a peak current density of 0.63 nA/µm² and an LPE-based on-wafer liquid-phase regrowth approach has produced diodes with peak current densities up to 1.2 mA/µm². To the author’s knowledge, this is comparable to the highest reported current density Ge TDs [15]. The diffused-junction Ge TDs demonstrated a zero-bias curvature coefficient of 50 V⁻¹. To the author’s knowledge, this is the only reported diffused-junction Ge tunnel device. Chapter 5 summarizes the achievements and proposes future study.
CHAPTER 2

TUNNEL-JUNCTION TRANSPORT AND FORMATION PHYSICS

This chapter outlines the physics of $p^+n^+$ tunnel diode current density and develops analytical expression used to engineer the device for both low and high current density applications. It is shown that the tunneling current density is strongly dependent on the junction electric field in the depletion region, which arises from the ionized fixed impurities. Higher impurities, lower semiconductor band gap $E_g$, and lower carrier effective mass lead to higher tunnel current density. The limits of semiconductor doping impose bounds on the maximum current density that can be achieved. The steepness of the doping gradient for high current density junctions set a further physical constraint. In this chapter, the physical limits constraining the formation of heavily doped semiconductors are also discussed. The tunnel current density is computed for Si and Ge.

For low current density junctions used in imager application, the curvature of the current-voltage ($I-V$) characteristic near zero bias provides a measure of the performance of the device. An analytical expression for this figure of merit is used to establish device development guidelines for Si and Ge.
2.1 Interband tunneling current-voltage relations

The Esaki interband tunneling diode is a bipolar device based on the single barrier interband tunneling of electrons and holes across the energy band gap of a heavily-doped $p$-$n$ junction. The heavy doping forms degenerate electron and hole gases in close proximity separated by a thin depletion region. Current flow is governed by quantum-mechanical tunneling through the thin depletion region, which ranges between 2 and 3 nanometers in high current density devices and to 15 or 20 nanometers in low current density devices.

![Figure 2.1 Schematic tunnel diode current-voltage curve with corresponding energy band diagrams. Shaded areas represent the filled electron states in the degenerately doped semiconductor.](image)

A schematic tunnel diode current-voltage characteristic is shown in Fig. 2.1 along with energy band diagrams corresponding to each of the labeled points on the $I$-$V$
characteristic. In reverse bias, (a), current flows by tunneling of valence band electrons on the \( p \) side to empty states in the conduction band of the \( n \) side. The tunneling probability increases with reverse bias as the junction electric field increases. At zero bias, (b), the net tunneling current between valence and conduction bands is zero. As bias is applied in the forward direction, electrons in the degenerate \( n \) region tunnel into empty states in the degenerate \( p \) region. When the overlap of electron states on the \( n \) side and empty states on the \( p \) side is maximized, the tunneling current reaches a peak, (c). Further increase in the forward bias decreases the overlap and the tunnel current decreases. The current reaches a valley, (d), when the conduction band minimum of the \( n \) side and the valence band maximum of the \( p \) region are equal. For this bias condition, there are no available states for electrons or holes to tunnel across the forbidden gap. One might expect the tunnel current in this circumstance to be zero; however, in practice, defect states in the forbidden gap can provide intermediate states for interband tunneling and the current produced by impurity-assisted or other scattering processes is known as excess current [38]. Beyond the valley current bias region, the current consists of both excess and thermionic emission current with the latter component dominating at large forward bias, (e).

Tunneling is a quantum-mechanical phenomenon describing the transport of particles through potential energy barriers. This transport can occur even if the energy of the particle is low relative to the barrier potential. Tunneling is a probabilistic effect which increases as the thickness of the barrier decreases toward the Bloch wavelength of
the particle. Direct tunneling is seen in semiconductors like GaAs and GaSb, where the band extrema lie at the Brillouin zone center, i.e. $\Gamma = 0$. For semiconductors like Si and Ge, the conduction band minima occur in the [100] and [111] transport directions, respectively, while the valance band maxima are at $\Gamma = 0$, therefore for these semiconductors, tunneling is indirect and must be assisted by phonon or impurity scattering in order for momentum to be conserved. For Ge, there exists a $\Gamma$-point conduction band approximately 0.14 eV above the $L$-conduction band minimum, so in Ge, both direct and indirect tunneling processes can be expected for $n$-doping concentrations exceeding approximately $1 \times 10^{20}$ cm$^{-3}$.

Tunneling physics can be understood from a consideration of the direct tunneling transport [39, 40] because the essential physics are contained in this simple case. Indirect tunneling theory has been developed [41-43], but is not needed to understand the work presented here.

A triangular potential energy barrier is often used to approximate the direct interband tunneling of an electron through the forbidden band gap [44, 45]. The tunneling probability, $T_t$, given by the Wentzel-Kramers-Brillouin (WKB) approximation for a triangular potential barrier, is [46]

$$T_t = \exp\left[-2 \int_{-x_1}^{x_2} |k(x)|\, dx\right]$$

(2.1)

where $k(x)$ is the wave vector in the barrier and $x_1$ and $x_2$ are the classical turning points, i.e. the points where the electron would classically be reflected off the barrier. In the
tunnel diode, the turning points correspond to the places where the Fermi level crosses the band edges. The WKB approximation is valid when the potential energy variation is slow with respect to the wavelength of the charged carrier.

For a triangular barrier, Fig. 2.2, the wave vector is given by

\[ k(x) = \sqrt{\frac{2m^*}{\hbar} \left( \frac{E_g}{2} - q\xi x \right)} \]

where \( E_g \) is the bandgap energy, \( q \) is the fundamental electron charge, \( \xi \) is the internal electric field, \( x \) is the spatial coordinate, and \( m^* \) is the effective mass of the charge carrier times \( m_0 \), the electron rest mass. Boundary conditions are given in Fig. 2.2. For this potential energy profile, the tunneling probability can be shown to be \[46\]

\[ T_t = \exp \left( -\frac{4\sqrt{2m^*E_g}}{3q\hbar \xi} \right) \]

where, for an abrupt junction, the internal electric field is given by
\[ \xi = \left( \frac{2q(V_{bi} + V_a)N_{eff}}{\varepsilon_s} \right)^{1/2}, \]  

(2.4)

where \( V_{bi} = (kT/q)\ln(N_A N_D/n_i^2) \) is the built-in potential, \( V_a \) is the applied voltage (negative for forward bias and positive for reverse), \( n_i \) is the intrinsic carrier density, \( \varepsilon_s \) is the dielectric constant, \( N_{eff} = N_A N_D/(N_A + N_D) \) is the effective doping density, \( N_A \) and \( N_D \) are the acceptor and donor densities, respectively, and \( h \) is Planck’s constant divided by \( 2\pi \). As seen from these equations, the tunneling probability, and therefore the tunneling current can be increased with smaller bandgap, smaller effective mass, lower dielectric constant, and higher doping concentration. Note that larger bandgap produces higher electric fields through higher \( V_{bi} \) but the overall dependence of the exponential argument is to decreased tunneling probability with energy band gap.

It can be shown that the tunneling current density can be written [47]

\[ J_I = \frac{\sqrt{2m^*q^3 \xi}}{4\pi^2 h^2 E_g^{1/2}} \left| V_a \right| \exp\left( -\frac{4\sqrt{m^*E_g^{3/2}}}{3q \xi h} \right). \]  

(2.5)

This relation provides a close fit to measured tunneling current in vertical Si \( p-n \) junctions under reverse bias over eight orders of magnitude [47], as shown in Fig. 2.3. These Si junctions were fabricated using ion implantation. The electric fields were evaluated either by measuring the doping gradients and concentrations using secondary ion mass spectrometry (SIMS) [26] or from capacitance-voltage (C-V) measurements [48]. \^ With its better fit to experimental data, Eq. (2.5) is chosen over Moll’s equation.

\^ The currents were measured at -1 V. Moll’s tunneling current density equation [49] differs by a factor of 2 in effective mass from Eq. (2.5).
Figure 2.3 Tunneling current density vs. electric field for a Si $p$-$n$ junction comparing predicted by Eq. 2.5 with experimental data: △ Stork and Isaac [48] and ▲ Fair and Wivell [26].

The tunneling current density vs. junction electric field for a Si TD is calculated using parameters shown in Fig. 2.3. The built-in potential is assumed to be equal to $E_g/q$. Note the reduced tunneling effective mass extracted by fitting experimental data is $0.33m_0$. This is almost twice what is expected; the tunneling effective mass is given by $m^* = 2m_t^*m_{lh}^*/(m_t^* + m_{lh}^*)$ [50] equals to $0.17m_0$, where $m_t^* = 0.19m_0$ is the electron transverse effective mass and $m_{lh}^* = 0.16m_0$ is the light hole effective mass [24]. For tunneling current calculation in this dissertation, the value of $0.33m_0$ is used since it yields a prediction close to experiments.
The relation between tunneling current density and electrical field in Eq. (2.5) can also be used to approximate the tunnel current in the forward bias direction [51]. To compute the peak tunneling current density and electric field vs. doping concentration, the applied bias $V_a$ in Eq. 2.5 is chosen at the forward peak voltage $V_p$ of the tunnel diode as defined in Fig. 2.4, where $V_p = (V_{dn}+V_{dp})/3$ [51], $V_{dn} = (E_{FN} - E_{C})/q$ is the Fermi potential above the conduction band in the $n$ region and $V_{dp} = (E_{V} - E_{FP})/q$ is the Fermi potential below the valence band in the $p$ region, respectively, Fig. 2.4. The value of $V_{dn}$ and $V_{dp}$ can be evaluated using the Joyce-Dixon approximation given by [52]

$$V_{dn} = \frac{kT}{q} \left[ \ln \left( \frac{N_D}{N_C} \right) + 0.3535 \left( \frac{N_D}{N_C} \right) - 0.00495 \left( \frac{N_D}{N_C} \right)^2 \right]$$

$$V_{dp} = \frac{kT}{q} \left[ \ln \left( \frac{N_A}{N_V} \right) + 0.3535 \left( \frac{N_A}{N_V} \right) - 0.00495 \left( \frac{N_A}{N_V} \right)^2 \right],$$

(2.6)

where $N_C$ and $N_V$ are the effective density of states in the conduction band and valence band at 300 K, respectively.

![Figure 2.4 Schematic current-voltage characteristic and energy band diagram of the tunnel diode biased at $V_p = (V_{dn}+V_{dp})/3$.](image)
The internal electric field and peak tunneling current densities vs. doping concentrations are then calculated and plotted for abrupt Si and Ge TDs in Fig. 2.5. Furukawa [25] has fabricated alloyed Ge TDs, and reported a reduced tunneling effective mass of \(0.08m_0\) for Ge at room temperature. In his work, the junction doping densities were measured; the TD peak current density was then plotted against the junction doping density and the tunneling effective mass was extracted from the slope. Using a similar extraction method, Breitschwerdt [53] has reported a tunneling effective mass of \(0.07m_0\) for alloyed Ge TDs under reverse bias at room temperature, and Meyerhofer et al. [16] reported a tunneling effective mass of \(0.06m_0\) for solution-grown Ge TDs at 4.2 °K. All of the values reported above are extracted based on parabolic barrier tunneling equations. These reported values are close to the theoretical value of \(0.06m_0\), given by \(m^* = 2[1/(3m_l^*)+2/(3m_t^*)+1/m_{lh}^*]^{-1}\) [50], where \(m_t^* = 0.082m_0\), \(m_l^* = 1.6m_0\) is the electron longitudinal effective mass, and \(m_{lh}^* = 0.04m_0\) [50]. In the Ge tunneling current density computation, the value from Furukawa, \(0.08m_0\), is used since it is extracted from a peak current density vs. doping density experiment. The effective doping concentrations of the \(p\) and \(n\) sides are varied from \(0.5 \times 10^{19}\) to \(2.5 \times 10^{20} \text{cm}^{-3}\). The lower bound is roughly the minimum condition for Ge degeneracy with equal \(n\) and \(p\) dopants. The upper bound of \(2.5 \times 10^{20} \text{cm}^{-3}\) is corresponding to half of the highest reported solid solubility of dopant in Ge. Doping concentration limits will be discussed in Section 2.2.2. A second calculation for Ge using tunneling effective mass of \(0.12m_0\), twice the theoretical value, is also conducted. The peak tunneling current with tunneling mass of \(0.12m_0\) is about 3 to
5x lower than that using 0.08\(m_0\), Fig. 2.5. Following discussion is based on the 0.08\(m_0\) calculation.

![Graph showing Electric field vs. Effective doping concentration](image)

**Figure 2.5** Peak tunneling current density and internal electric field vs. effective doping density for an abrupt \(p^+n^+\) junction.

In Fig. 2.5, both the internal electric field and peak tunneling current density increase with effective doping concentration. The required electric field to achieve 1 mA/\(\mu\)m\(^2\) current density is approximately 5 MV/cm for Si and approximately 1.2 MV/cm for Ge. To achieve 1 mA/\(\mu\)m\(^2\) current density requires an effective doping density of approximately \(7 \times 10^{19}\) cm\(^{-3}\) for Si and \(2 \times 10^{19}\) cm\(^{-3}\) for Ge, i.e. Ge, with its smaller
bandgap and tunneling effective mass, requires less doping concentration than Si to achieve the same tunneling current.

This analysis and discussion leads to two questions. First, what doping concentration is achievable in Si and Ge? And second, since in practice an atomically abrupt doping profile is only approached and since a non-abrupt doping profile will decrease the internal electric field and tunneling current, what doping gradient is achievable in Si and Ge?

In the following section, these two questions will be discussed. The junction doping technologies are briefly reviewed first. The physics limiting doping concentration are outlined and the state-of-the-art in doping concentrations and doping gradients are reviewed. Peak current density vs. junction doping concentration and gradient is quantified to guide the experimental work in chapters 3 and 4.

2.2 Tunnel-junction formation

2.2.1 Doping technologies

The most common approach for doping Si is II [54]. For heavy doping of tunnel junctions, MBE [4-7] and solid-phase diffusion from SODs [8] have received attention.

In II, the dopants to be implanted are ionized, accelerated, and impacted into a substrate with precise control of the total dose and dopant distribution. Rapid thermal processing follows to electrically activate the implanted dopants. RTP utilizes lamps (tungsten halogen lamps are widely used) to heat the wafer rapidly to a high temperature,
tens to hundreds of degrees per second to temperatures of order 1000 °C with hold times controllable to approximately 1 s. Short thermal processes are typically used to minimize diffusion. To further minimize the dopant diffusion, flash lamp annealing (FLA) [56] and flash-assisted RTP (fRTP) [57] reduce the anneal process to milliseconds and allow higher peak temperatures to be targeted to increase electrical activation [56, 57]. The lamps (xenon flash lamps [56] and argon arc lamps [58]), because of the low thermal mass [57], can achieve annealing times in the range of milliseconds. As an alternative to RTP, FLA and fRTP, laser annealing is being studied as a means of activating the dopants and repairing the lattice damage from ion implantation [59]. By irradiating with laser pulses, tens of nanoseconds in duration, the implanted Si surface can be melted with a liquefied thickness determined by the energy of the laser beam. The dopants then redistribute across the liquid layers and abrupt junctions reform by liquid-phase epitaxial regrowth. During the regrowth, a large fraction of the dopants, up to 69% of implanted B ions [60], will occupy the substitution sites and become electrically activated.

Molecular beam epitaxy is a widely used approach for the formation of nearly abrupt junctions, with the ability to change both doping and composition with monolayer precision. In MBE elemental materials are evaporated onto a heated substrate in ultrahigh vacuum (~10^{-10} Torr). Junctions formed by depositing heavily-\textit{p}-doped Si or SiGe onto heavily-\textit{n}-doped Si or SiGe are sufficiently abrupt to form TDs [4-7, 13].

Solid-phase diffusion from SODs [61-62] has also been used for doping silicon. SOD can be directly deposited on the Si surface by spinning or spraying. Upon rapid
thermal annealing, the SOD decomposes and releases impurities for diffusion [62]. Spike anneals are used to maximize the doping abruptness.

2.2.2 Heavy doping

The solid solubility sets the doping concentration limits for thermal equilibrium processes; thermal nonequilibrium processes can overcome this limit. Thermal equilibrium doping processes are defined as processes in which temperature and pressure gradients go to zero in times short with respect to the process time. Equilibrium carrier concentrations in Si have been investigated [63-66] using thermal equilibrium processes such as RTP and furnace annealing. The achieved dopant concentrations are limited by solid solubilities [63]. Activated dopant concentrations are lower than the physical dopant concentrations [63]. It is found that precipitation and clustering of As [65, 67] and P [66], or dopant complex formation in As [68] can deactivate part of the dopants.

The $n$ type dopants with the highest solid solubility in Si are As and P, with solid solubilities for both elements exceeding $1 \times 10^{21}$ cm$^{-3}$ in Si [69]. Boron is the only $p$-type dopant which achieves a solid solubility above $1 \times 10^{20}$ cm$^{-3}$ [69]. The $n$-type dopants with the highest solid solubility in Ge are As and P, with the solid solubilities of $8 \times 10^{19}$ cm$^{-3}$ and $2 \times 10^{20}$ cm$^{-3}$, respectively [69]. Aluminum and Ga are the $p$-type dopants with the highest solid solubilities for Ge, both of which achieve $5 \times 10^{20}$ cm$^{-3}$ [69]. These values were obtained from Hall effect measurement, chemical analysis of pulled crystals, and phase diagrams [69]. Solid solubility is temperature dependent.
In a process like laser annealing [70-72], the process time is so short that
temperature cannot reach equilibrium in this period; in process such as low temperature
epitaxial layer growth [19, 73], the growth is also out of equilibrium. These techniques
are able to achieve activated dopant concentrations about one order above those
achievable by thermal equilibrium processes in Si for As [70], P [71] and B [19, 72]. It
has been reported that in laser annealing [74, 75], dopants can occupy metastable
substitutional sites and move to the substitutional sites during the liquid-phase epitaxial
regrowth of Si.

Theories for concentration above the solid solubility have been developed [76, 77].
One theory is that the Columbic repulsions between ionized dopants limit the attainable
maximum doping concentrations [76]. Schubert et al. [76] found experimental results for
Be doping in GaAs to be in agreement with this theory. Luo et al. [77] proposed that
hydrogen in the growth ambient reduces B clusters and increase the B concentration
during growth and achieved good agreement between experimental [19] and theoretical
results.

In Table 2.1, the highest measured (by differential Hall or spreading resistance
method) electrically-active dopant concentrations of As, P, and B achieved in Si using
various techniques are summarized. In Si, the highest measured electrically-active
concentration is achieved for P at 5 \times 10^{21} \text{ cm}^{-3} by laser annealing [71]. The highest
active concentration of \textit{p}-type dopants is 1.3 \times 10^{21} \text{ cm}^{-3} by gas source MBE [19].
Considering the compensation between \textit{n} and \textit{p} dopants when forming the junction,
diffusing arsenic or phosphorus into a heavily-boron-doped Si layer is favorable, since As or P can achieve a higher active concentration than B.

TABLE 2.1

SUMMARY OF THE HIGHEST MEASURED ELECTRICALLY-ACTIVATED DOPANT CONCENTRATION OF As, P AND B in Si USING VARIOUS APPROACHES.

<table>
<thead>
<tr>
<th>Process</th>
<th>As active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium II followed by RTP or furnace annealing</td>
<td>$4.8 \times 10^{20}$ 1200 °C [65]</td>
</tr>
<tr>
<td>Thermal nonequilibrium II followed by laser annealing</td>
<td>$4 \times 10^{21}$ 1414 °C [70]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>P active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium II followed by RTP or furnace annealing SOD+RTP</td>
<td>$5.7 \times 10^{20}$ 1000 °C [65] $9 \times 10^{20}$ 1000 °C [78]</td>
</tr>
<tr>
<td>Thermal nonequilibrium II followed by laser annealing</td>
<td>$5 \times 10^{21}$ 1414 °C [71]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>B active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium II followed by RTP or furnace annealing</td>
<td>$2 \times 10^{20}$ 1000°C [79]</td>
</tr>
<tr>
<td>Thermal nonequilibrium laser annealing</td>
<td>$1 \times 10^{21}$ 1414 °C [72]</td>
</tr>
<tr>
<td>solid source MBE</td>
<td>$6 \times 10^{20}$ 1414 °C [73]</td>
</tr>
<tr>
<td>gas source MBE</td>
<td>$1.3 \times 10^{21}$ 1414 °C [19]</td>
</tr>
</tbody>
</table>

In Table 2.2, the highest measured electrically-active dopant concentrations of B, Al, Ga, P and As achieved in Ge are summarized. The highest measured electrically-activate concentration in Ge by II of P is $5 \times 10^{19}$ cm$^{-3}$ measured using sheet
resistance measurement [55, 81]. The physical P concentration reaches $1 \times 10^{21}$ cm$^{-3}$ in II, and the deactivation of P is possibly due to P clusters [81]. The $n$-type doping density in Ge using SOD/RTP achieves approximately $2 \times 10^{20}$ cm$^{-3}$, as discussed in section 4.2. The highest active concentration of $p$-type dopant achieved for both Al and Ga is $5 \times 10^{20}$ cm$^{-3}$ [82]. In the case of Ge, diffusing Al or Ga into a heavily P-doped Ge layer is preferable since Al or Ga can achieve a higher active density than P.

### TABLE 2.2

<table>
<thead>
<tr>
<th>Process</th>
<th>P active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium</td>
<td></td>
</tr>
<tr>
<td>II followed by RTP annealing</td>
<td>$5 \times 10^{19}$ 600 °C [55, 81]</td>
</tr>
<tr>
<td>SOD followed by RTP annealing</td>
<td>$2 \times 10^{20}$ 800 °C this work</td>
</tr>
<tr>
<td>Thermal nonequilibrium</td>
<td></td>
</tr>
<tr>
<td>II followed by RTP annealing</td>
<td>$1 \times 10^{21}$ 938 °C [83]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>As active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium</td>
<td></td>
</tr>
<tr>
<td>II followed by RTP annealing</td>
<td>$3.6 \times 10^{19}$ 600 °C [55]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>Al and Ga active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal nonequilibrium</td>
<td></td>
</tr>
<tr>
<td>Crystal growth</td>
<td>$5 \times 10^{20}$ 938 °C [82]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>B active concentration cm$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal equilibrium</td>
<td></td>
</tr>
<tr>
<td>II followed by RTP annealing</td>
<td>$2.6 \times 10^{19}$ 650 °C [84]</td>
</tr>
<tr>
<td>Thermal nonequilibrium</td>
<td></td>
</tr>
<tr>
<td>II followed by solid phase regrowth</td>
<td>$4.7 \times 10^{20}$ 400 °C [84]</td>
</tr>
</tbody>
</table>
2.2.3 Abrupt junction formation

In this section, the state-of-the-art in doping gradients using various approaches is reviewed. To compare the doping abruptness of different doping technologies, a new performance measure is defined, applicable to Esaki TD formation. As discussed in section 2.1, the tunnel current density depends exponentially on the maximum electric field in the junction which depends on the doping densities and the impurity gradients at the junction. Thus, when the dopant profile is anything but abrupt, the important dopant slope will be the slope at the center of the junction. When one considers a SIMS measurement of a doping profile, the relevant dopant slope is defined as the slope at the solid solubility density of the opposing dopant type. This provides a best case estimate of the maximum electric field in the tunnel junction for that diffusion profile.

As an example of a doping profile slope, Fig. 2.6 shows a measured SIMS profile of B in Si diffused from an SOD source by RTP at 900 °C for 1 s. The solid solubility of P at 900 °C is approximately $7 \times 10^{20}$ cm$^{-3}$ [69]. Therefore, the gradient of the B diffusion is estimated to be 5 nm/decade at the concentration of $7 \times 10^{20}$ cm$^{-3}$. Using this definition for impurity gradient, the best gradients for B diffusion in Si are summarized from published SIMS data in Table 2.3. Nonequilibrium processes, like MBE and laser annealing achieve the best doping gradients of 2 and 2.5 nm/dec, respectively. Thermal equilibrium processes using RTP achieve doping gradients of 3.5 to 5 nm/dec. Considering that nonequilibrium processes can achieve doping densities about one order
of magnitude higher than equilibrium processes, as discussed in section 2.2.2, nonequilibrium processes should be superior for tunnel junction formation.

Figure 2.6 SIMS profile of B diffusion into Si using SOD and RTP at 900 °C of 1 s. The gradient of diffusion is estimated at the concentration of $7 \times 10^{20}$ cm$^{-3}$, which is the solid solubility of P at 900 °C.

TABLE 2.3

RECORD DOPING GRADIENT FOR B DIFFUSION IN SILICON EXTRACTED FROM SIMS AT THE SOLID SOLUBILITY OF THE OPPOSING DOPANT.

<table>
<thead>
<tr>
<th>Process</th>
<th>gradient nm/dec</th>
<th>Experimental conditions</th>
<th>[Ref.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ion implantation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with RTP</td>
<td>3.5</td>
<td>0.2 keV B+ ion implant, 1100 °C 1 s</td>
<td>[85]</td>
</tr>
<tr>
<td>with FLA</td>
<td>5</td>
<td>1.5 keV BF2+ ion implant, 1100 °C 1.5 ms</td>
<td>[86]</td>
</tr>
<tr>
<td>with laser</td>
<td>2.5</td>
<td>0.25-1 keV B+ ion implant, 600 mJ/cm²</td>
<td>[87]</td>
</tr>
<tr>
<td>MBE</td>
<td>2</td>
<td>MBE growth at 320 °C</td>
<td>[20]</td>
</tr>
<tr>
<td>SOD with RTP</td>
<td>5</td>
<td>SOD, 900 °C 1 s</td>
<td>[8]</td>
</tr>
</tbody>
</table>
Though Ge TDs were widely fabricated in the 1960s, the metrologies available at that time could not provide precise information on the doping profile and gradient. Some recent measurement of II of P in Ge demonstrated a doping gradient of approximately 3 nm/dec by laser annealing [83] and approximately 13 nm/dec for B by II and RTP [81]. For the B doping, the implantation energy is relatively large (6 keV) [81] compared to that for Si (0.2 keV) [85]. Boron diffusion in Ge is known to be very slow, i.e. no measurable B diffusion is detected for furnace annealing at 800 ºC for 3 hours [88]. Therefore, with the lowering of implantation energy, the B doping gradient in Ge could be improved further.

2.3 Tunneling current density vs. junction doping density and gradient

In this section, the relation between peak tunneling current density, junction doping density, and junction abruptness are quantified. The effects of doping density and junction abruptness on the junction electric field are calculated. The peak tunneling current density is then calculated using the procedures discussed in section 2.1.

The effect of doping slope on the electric field of a Si $p$-$n$ diode was computed from solution of Poisson equation for various impurity profiles. The simulated energy band diagram of an atomically-abrupt doping profile (0 nm/decade) is compared with the impurity gradients of 1, 2 and 4 nm/decade, Fig. 2.7(a), from $N_A - N_D = 1 \times 10^{20}$ cm$^{-3}$ to $N_A - N_D = -1 \times 10^{20}$ cm$^{-3}$ for a forward peak bias of $V_P = (V_{dn} + V_{dp})/3$. The doping gradients of 2 and 4 nm/dec are comparable to what has been achieved by MBE and RTP in Table 2.3.
The electric field is the derivative of the simulated potential energy, $\zeta = -dE/dx$, plotted in Fig. 2.7(b). The tunneling distance increases from 4.4 nm for the abrupt profile to 4.5, 4.7, and 5.1 nm for the 1, 2 and 4 nm/decade profiles, respectively. The peak junction electric fields are reduced from 4.22 to 3.83, 3.51, and 3.07 MV/cm, respectively. From Fig. 2.5, the electric field 4.22 MV/cm of an abrupt junction corresponds to a tunneling current of approximately 0.1 mA/µm².

Figure 2.7 (a) Computed band diagram and (b) junction electric field for four doping profiles: an abrupt profile and 1, 2, and 4 nm/decade gradients from $p$ to $n$ with doping density of $1 \times 10^{20}$ cm$^{-3}$. The calculation was made using W. R. Frensley’s BandProf.
The effect of doping concentration and profile abruptness on electric field is then calculated using the same method and is plotted in Fig. 2.8. The impact of varying the effective doping density (from $2.5 \times 10^{19} \text{ cm}^{-3}$ to $2 \times 10^{20} \text{ cm}^{-3}$) and doping gradient (from abrupt to 4 nm/decade) is computed against junction electric field. Doping gradient becomes an increasingly important factor as doping density increases above $1 \times 10^{20} \text{ cm}^{-3}$.

The electric field increases by a factor of approximately 1.5x by improving the gradient from 4 nm/dec to 1 nm/dec at an effective doping density of $2 \times 10^{20} \text{ cm}^{-3}$, compared with the 1.2x increase at an effective doping density of $2.5 \times 10^{19} \text{ cm}^{-3}$. Therefore more improvement in current density from improvements in doping gradient can be expected for junction doping concentrations above $1 \times 10^{20} \text{ cm}^{-3}$.

![Figure 2.8 Junction electric fields vs. position for various doping abruptness and effective doping densities, extracted from the gradient of simulated energy band profiles using W. R. Frensley's program BandProf.](image-url)
The peak tunneling current density vs. junction doping concentration and abruptness is next computed. Similar to Fig. 2.5, the built-in potential is assumed to be equal to $E_g/q$. The reduced tunneling effective masses used are $0.33m_0$ for Si and $0.08m_0$ for Ge. The effective doping concentration range is the same as that in Fig. 2.5, i.e. from $5 \times 10^{18}$ cm$^{-3}$ to $2.5 \times 10^{20}$ cm$^{-3}$. The peak tunneling current density vs. junction doping concentration and abruptness for Si and Ge TDs are plotted in Fig. 2.9.

Figure 2.9 Peak tunneling current density vs. effective doping density with doping abruptness of 0, 1, 2, and 4 nm/decade in Si and Ge.
These calculations show that Ge, with its smaller bandgap and reduced tunneling effective mass compared to Si, can achieve approximately three to four orders of magnitude higher tunneling current at the same effective doping concentration and junction abruptness. As an example, an effective doping density of $3 \times 10^{19} \text{ cm}^{-3}$ with a doping gradient of 4 nm/dec is sufficient to achieve 1 mA/µm$^2$ current density in Ge, but merely leads to approximately $10^{-4}$ mA/µm$^2$ in Si.

The effect of improving the doping abruptness on the peak tunneling current appears to be roughly the same for different doping densities. The peak current density increases by a factor of approximately 4, 2.5, and 2.5x as doping abruptness is changed from 4 to 2 nm/dec, from 2 to 1 nm/dec, and from 1 nm/dec to atomically abrupt profile, respectively in Si and increases by a factor of approximately 4, 2.3 and 2.3x correspondingly in Ge.

The requirements of doping concentration and abruptness to achieve high tunneling current density devices, 1 mA/µm$^2$, are estimated based on Fig. 2.9. Using the experimentally determined highest doping densities of $9 \times 10^{20} \text{ cm}^{-3}$ for P [78] and $2 \times 10^{20} \text{ cm}^{-3}$ for B using thermal equilibrium processes, the highest effective doping concentration is calculated to be $1.6 \times 10^{20} \text{ cm}^{-3}$. From Fig. 2.9, a junction gradient of approximately 4 nm/dec will result in 1 mA/µm$^2$ peak current density. In the case of Ge, the highest active doping concentrations are $2 \times 10^{20} \text{ cm}^{-3}$ for P (this work) and $5 \times 10^{20} \text{ cm}^{-3}$ for Al/Ga [82]. Based on these as maximum $n$ and $p$ densities, the effective doping
density is estimated to be $1.4 \times 10^{20} \text{ cm}^{-3}$. A junction gradient of approximately 40 nm/dec will achieve 1 mA/µm$^2$ peak current density, simulated using BandProf.

2.4 Switching speed

The switching speed of a tunnel diode is determined by the speed at which the junction capacitance can be charged or discharged by the current. According to Diamond [89], the switching time $\tau$ of a tunnel diode is not smaller than

$$\tau \geq \frac{V_P - V_V}{I_P \left(\frac{1}{C_j}(1 - \frac{1}{PVR})\right)}$$

(2.7)

where $V_P$ is the peak voltage, $V_V$ is the valley voltage, $I_P$ is the peak current, $I_V$ is the valley current, $C_j$ is the junction capacitance, and PVR is defined as $PVR = I_P/I_V$. From this equation, it is clear that the switching speed can be enhanced by increasing $I_P/C_j$, called the speed index ratio, and PVR. However, an increase of PVR from 2 to infinity will increase the speed by only a factor of 2. The speed index is the key factor to increase in order to produce a high speed tunnel diode.

Best performance TDs fabricated to date using various approaches are summarized in Table 2.4, showing the best speed index or highest PVR. For Si-based TDs, MBE has produced the best peak current density and speed index, 2.2 mA/µm$^2$ and 36 mV/ps, respectively [7]. The best PVR, 6, is achieved by MBE using Si/SiGe [6]. Silicon and SiGe TDs using RTP and SODs achieve peak current density of approximately 2 µA/µm$^2$ and PVR of approximately 2. Silicon TDs using III achieve peak
current density of approximately 0.1 µA/µm² and PVR of approximately 2 [18].

Germanium TD, with the junction formed by an electrical pulse forming approach has reported current density up to 1 mA/µm² [15]. Germanium TDs with peak current density of 1.2 mA/µm² are achieved using on-wafer liquid-phase epitaxial (LPE) regrowth discussed in section 4.4. The best reported Ge TD PVR is 16 by LPE regrowth [90].

### TABLE 2.4

**GROUP IV TUNNEL DIODES PROPERTIES vs. FABRICATION APPROACH, SELECTED BY HIGHEST PEAK CURRENT DENSITY OR PEAK-TO-VALLEY RATIO.**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Approach</th>
<th>Type</th>
<th>$J_p$ mA/µm²</th>
<th>Speed Index mV/µs</th>
<th>PVR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franks et al.</td>
<td>1965</td>
<td>LPE</td>
<td>Si</td>
<td>0.01</td>
<td>1.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Chung et al.</td>
<td>2006</td>
<td>MBE</td>
<td>Si/SiGe</td>
<td>2.2</td>
<td>36</td>
<td>1.5</td>
</tr>
<tr>
<td>Eberl [6]</td>
<td>2001</td>
<td>MBE</td>
<td>Si/SiGe</td>
<td>0.33</td>
<td>16.5*</td>
<td>6</td>
</tr>
<tr>
<td>Dashiell et al.</td>
<td>2000</td>
<td>MBE</td>
<td>Si</td>
<td>0.47</td>
<td>23.5*</td>
<td>1.3</td>
</tr>
<tr>
<td>Dashiell et al.</td>
<td>2002</td>
<td>MBE</td>
<td>Si</td>
<td>0.16</td>
<td>7.1</td>
<td>-</td>
</tr>
<tr>
<td>Yan et al.</td>
<td>2004</td>
<td>RTP</td>
<td>Si</td>
<td>0.002</td>
<td>0.2</td>
<td>2.1</td>
</tr>
<tr>
<td>Wernersson et al.</td>
<td>2005</td>
<td>RTP/CVD</td>
<td>Si/SiGe</td>
<td>0.002</td>
<td>0.1*</td>
<td>2.6</td>
</tr>
<tr>
<td>Koga et al. [18]</td>
<td>2001</td>
<td>Implantation</td>
<td>Si</td>
<td>0.0001</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Glicksman et al [90]</td>
<td>1965</td>
<td>LPE</td>
<td>Ge</td>
<td>0.2</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>This work</td>
<td>2005</td>
<td>LPE</td>
<td>Ge</td>
<td>1.2</td>
<td>-</td>
<td>1.5</td>
</tr>
</tbody>
</table>

* Computed using a tunnel diode capacitance of 20 fF/µm² obtained from simulation of an abrupt tunnel diode with symmetric $n$ and $p$ doping density of $10^{20}$ cm⁻³ using BandProf by W. R. Frensley.
2.5 Curvature coefficient of backward tunnel diode

In addition to the high current density TDs discussed above, low current density TDs, such as backward TDs, are also of interest in microwave detection applications such as radiometry and millimeter wave imaging [32-36]. In these detection applications, the microwave signal to be detected is incident on the backward tunnel diode detector with power $P_i$. The subsequent average increase in the output current of the diode, $I_r$, is a measure of the incident signal. The ratio $I_r/P_i$ is defined as the detector sensitivity $\beta$ [34].

The curvature coefficient $\gamma$, which is defined as the ratio of the second derivative to the first derivative of the $I$-$V$ characteristic, $\gamma = \left| \frac{(d^2I/dV^2)}{(dI/dV)} \right|$ [91], quantifies the non-linearity of the backward TD. The detector sensitivity is directly proportional to the curvature coefficient [34]. In the following section, the relation between detection sensitivity and curvature coefficient will be discussed, and an analytical expression is developed. Silicon and Ge backward TD curvature coefficients are computed vs. junction doping concentration to establish device development guidelines.

A backward TD detector equivalent circuit is shown in Fig. 2.10 [34]. The voltage source $V_b$ represents the dc bias voltage applied across the diode. The RF source $E_b$ represents the incident microwave signal voltage. The RF source impedance is $R_b$, the diode series resistance is $R_s$, the diode series inductance is $L_s$ which arises from the bond wire, the junction resistance is $R_j$, and the junction capacitance is $C_j$. The dc current is $I_b$ and the r.m.s. average increase of diode current due to the incident microwave signal is $I_r$. The current flow direction is labeled as the direction of the arrow in Fig. 2.10.
The detection sensitivity is given by [34, 92, 93]

\[
\beta = \frac{\gamma R_j (1 - |\Gamma|^2)}{2(1 + R_s/R_j)^2 (1 + (f/f_{ci})^2)}
\]  

(2.8)

where \(f\) is the frequency, \(f_{ci} = \sqrt{(1 + R_s/R_j)/(R_s R_j/2\pi C_j)}\) is the cut-off frequency of the diode. The reflection coefficient representing the mismatch between the diode and the RF source is given by

\[
\Gamma = \frac{R_s + j\omega L_s + 1/(1/R_j + j\omega C_j) - R_b}{R_s + j\omega L_s + 1/(1/R_j + j\omega C_j) + R_b}.
\]  

(2.9)

Increasing the curvature coefficient will proportionally increase the detection sensitivity.

For a \(p-n\) junction or a Schottky diode, the current voltage characteristic is given by \(I = I_0 [\exp(-qV_a/\eta kT) - 1]\), where \(I_0\) is the saturation current and \(\eta\) is the ideality factor. The curvature coefficient for these diodes, \(q/\eta kT\), is independent of bias and varies inversely with temperature. For an ideal diode with ideality factor of one, the maximum curvature coefficient at room temperature is approximately 39 V \(^{-1}\).

The curvature coefficient of a backward TD can be higher than the \(p-n\) or Schottky diode. The backward TD current density, \(J'_b\), at small biases, \(V_a \ll 2kT/q\), is given by Karlovsky [91]
\[ J_i^* = AZ \frac{qV_a}{kT}(qV_{dn} + qV_{dp} - qV_a)^2, \quad (2.10) \]

where \( A = (qm^*)/(72h^3) \) is a constant, and \( Z = \exp[(\sqrt{2m^*E_g^{3/2}}/(3qh\xi))] \) is the tunneling probability. The derivation of Eq. (2.5), assumes a bias voltage \( V_a \rangle qh\xi/2\sqrt{2m^*E_g^{1/2}} \) [49], which is about 60 mV for a junction electric field of 4 MV/cm. For a bias voltage \( V_a \langle 2kT/q \), about 52 mV at room temperature, Karlovsky [91] simplifies the Fermi functions to yield Eq. (2.10). The curvature coefficient at zero bias for the backward diode is [91]

\[ \gamma = \frac{4}{V_{dn} + V_{dp}} + \frac{2}{\hbar} \sqrt{\frac{\varepsilon m^*}{N_{eff}}}. \quad (2.11) \]

By decreasing doping concentrations, the Fermi potential decreases toward zero and the curvature coefficient should tend to infinity. However, other current components, such as impurity-assisted tunneling current or thermal current, will decrease the curvature coefficient. Compared with the curvature coefficient of a \( p-n \) diode or Schottky diode, which decreases linearly with increasing temperature, backward TD curvature coefficient is dependent on temperature only through effective mass and dopant activation changes which change relatively weakly with temperature.

The relations between zero-bias curvature coefficient and effective doping concentrations for Si and Ge backward TDs are computed and plotted in Fig. 2.11. In this computation, the reduced tunneling effective masses for Si and Ge are the experimental values of 0.33\( m_0 \) and 0.08\( m_0 \). Solid curves are the curvature coefficients while the dashed curves are the first and second terms in Eq. (2.11). At low effective doping concentration
of $10^{19}$ cm$^{-3}$, the first term is much larger than the second term. When the doping concentration increases toward $10^{20}$ cm$^{-3}$, the first and second terms are comparable. By lowering the effective doping concentration, the Fermi potential decreases and therefore the curvature coefficient increases. Curvature coefficient higher than 39 V$^{-1}$ is readily achieved for Ge at effective doping density lower than approximately $1.8 \times 10^{19}$ cm$^{-3}$ and for Si, at effective doping density lower than approximately $4 \times 10^{19}$ cm$^{-3}$.

![Graph showing curvature coefficient vs. effective doping density for Si and Ge](image)

Figure 2.11 Zero-bias curvature coefficient vs. effective doping density for abrupt Si and Ge backward TDs. The impurity densities on either side are assumed to be equal. The curves labeled with Si$_1$ and Ge$_1$ represent the first term of Eq. (2.11), and the curves labeled with Si$_2$ and Ge$_2$ represent the second term of Eq. (2.11).
Germanium backward TD can achieve the same zero bias curvature coefficient at a lower doping concentration compared with Si, Fig. 2.11. According to Fig. 2.11, Ge backward TDs can achieve a curvature coefficient of approximately 60 V$^{-1}$ at effective doping concentration of $10^{19}$ cm$^{-3}$, compared to about $3 \times 10^{19}$ cm$^{-3}$ for Si. Theoretically, all semiconductors can achieve the same zero-bias curvature coefficient as long as the Fermi potentials $V_{dn} + V_{dp}$ are the same. Nevertheless, the best reported zero-bias curvature coefficient for a Ge backward TD is 70 V$^{-1}$ [91], which is about twice that for Si, 31 V$^{-1}$ [35]. It is possible that higher doping concentrations in Si backward TDs generate higher impurity assisted tunneling current and therefore reduce the zero-bias curvature coefficient. Since Ge backward TDs can achieve the same zero-bias curvature coefficient at lower concentration, Ge is likely preferred for backward TD detector application than Si. As will be discussed in chapter 3 and 4, Ge backward TD with zero-bias curvature coefficient of 54 V$^{-1}$ and Si backward TD with zero-bias curvature coefficient of 34 V$^{-1}$ are achieved. These diodes show lower zero-bias curvature coefficient than expected from the effective junction doping density, which might suggest current components other than tunneling current exist and decrease the zero-bias curvature coefficient.
CHAPTER 3

SILICON TUNNEL DIODE FABRICATION

Two diffusion geometries for forming Si tunnel junctions using SODs and RTP have been explored in this research. Tunnel diodes with peak current densities of order 1 \( \mu A/\mu m^2 \) and peak-to-valley ratio of approximately 1.3 have been demonstrated using a vertical junction geometry process. In this vertical process, Si was doped to heavily \( n \)-type using Emulsitone’s SOD Phosphorosilicafilm \( 1 \times 10^{21} \) using RTP. Tunnel junctions were then formed by rapid thermal diffusion of B using Emulsitone’s SOD Borofilm100 into the heavily-\( n \)-doped region. Backward TDs were demonstrated with zero-bias curvature of 34 \( V^{-1} \).

Lateral backward tunnel junctions were also demonstrated with a current density of 30 nA/\( \mu m^2 \). In this process, the dopants were diffused using proximity RTP and the same SODs. The tunnel junctions were formed by meeting P and B dopants laterally. A self-aligned tungsten CMP enables the first demonstration of lateral Si TDs using SODs and RTP.
3.1 Spin-on diffusants and rapid thermal processing

The details of the spin-on diffusants and rapid thermal processing approach are described as following. The doping sources used in this work are Borofilm 100 and Phosphorosilicafilm $1 \times 10^{21}$ from Emulsitone. Phosphorosilicafilm $1 \times 10^{21}$ from Emulsitone was selected because it possesses the highest P concentration, up to $1 \times 10^{21}$ cm$^{-3}$, among all the phosphorus-containing spin-on diffusants found on the market.

Spin-on diffusants dope the Si substrate by reactions between Si and the dopant oxides released from the SOD upon rapid thermal annealing [62]. Borofilm 100 consists of water, boric acid, and a carbon polymer as the supporting material [94]. The B concentration in the SOD is $5 \times 10^{20}$ cm$^{-3}$. Borofilm 100 decomposes above 500 °C and leaves $B_2O_3$ as a solid-phase dopant source. Phosphorosilicafilm $1 \times 10^{21}$ incorporates P at a concentration of $1 \times 10^{21}$ cm$^{-3}$, using $SiO_2$ as the supporting material and ethyl alcohol as the solvent [94]. Decomposition at high temperature about 500 to 600 °C results in the formation of $P_2O_5$ as the dopant source. Dopants are incorporated into Si by the reactions $2B_2O_3 + 3Si \rightarrow 4B + 3SiO_2$ and $2P_2O_5 + 5Si \rightarrow 4P + 5SiO_2$ [62, 94].

One method for using SODs with RTP is by spinning the SODs directly onto a Si wafer. Upon anneal, the SOD first forms a dopant-containing oxide with Si on the wafer surface. The dopant diffuses into Si from the oxide with thermal annealing. Another type of rapid thermal doping method, called proximity RTP [61, 62], has been developed, in which a dopant source and a wafer are stacked in close proximity in a rapid thermal processor. In this process, the SOD is spun on an RCA-cleaned Si wafer, called the
source wafer. A second wafer, called the device wafer, is cleaned at the same time. The source wafer is then placed on top of the device wafer, with the coated surface facing the surface to be diffused, and separated by 0.18-0.25 mm quartz spacers. Tungsten halogen lamps in the rapid thermal processor heat the wafers from both sides simultaneously and the temperature is measured by a thermocouple touching the backside of the device wafer. Figure 3.1 shows a schematic cross-section of the reactor with the arrangement of source wafer on top of the device wafer separated by quartz spacers. In the proximity RTP process, the SOD is vaporized during anneal and transports across the gap to the device wafer where it decomposes and incorporates into the Si.

Figure 3.1 Schematic diagram of the proximity RTP, showing the source wafer with SODs placed on top of the device wafer and separated by quartz spacers.
The advantage of proximity RTP over direct spinning is a cleaner Si surface after dopant diffusion. Wang et al. [8] showed that using Borosilicafilm 5257, a residue layer forms at the wafer surface which is insoluble in HF. The residue film is approximately 5 nm using proximity RTP. If the Borosilicafilm 5257 is applied directly to the wafer, an insulating residue layer of approximately 80 nm was formed. Although no further residue layer thickness measurements were performed on other type of SODs, this process is employed for most of the later SOD and RTP experiments of this research.

3.2 Diffused-junction tunnel diodes in a vertical geometry

3.2.1 Vertical process flow

Silicon TDs have been fabricated by diffusing B into a P-doped region through an oxide-defined window. This approach requires two masks, and allows $I-V$ to be measured from front-to-back. A detailed process traveler is provided in Appendix 1.

The process flow for the oxide window approach is outlined in Fig. 3.2. Phosphorus-doped (1-5 mΩ·cm), 100 mm diameter, (100) prime Si device wafers were RCA-cleaned and hydrogen-terminated in buffered HF. Source Si wafers, 100 mm, were similarly cleaned and spin coated with Phosphorosilicafilm $1\times10^{21}$. The device wafer was then oxidized to grow a 600 nm oxide as shown in Fig. 3.2(a). After lithography, the oxide window was patterned then etched using 10:1 buffered HF, Fig. 3.2(b).
Phosphorus was diffused using proximity RTP at 900 °C for 1 s with a heating rate of 30 °C/s. After the proximity RTP, the source wafer was removed and the device wafer was annealed again at 900 °C for 90 s, Fig. 3.3(a). The purpose of this 90 s anneal is to lower P surface concentration and reduce B compensation. Backward TDs were fabricated without this 90 s diffusion [95]. The device wafer was next dipped in 10:1 buffered HF 2.5 minutes to clean the SOD residue prior to loading for B diffusion as shown in Fig. 3.3(b). During this etch the oxide is thinned by approximately 150 nm. A hydrophilic surface after P diffusion turns into a hydrophobic surface after 2.5 minutes cleaning in buffered HF, indicating that the SOD residue has been removed.
The B dopant source Emulsitone Borofilm100, is used in the same way as the P source. A single proximity RTP at 900 °C for 1 s with a heating rate of 30 °C/s was used, Fig. 3.4(a). A 10:1 buffered HF etch for 6 minutes was used to clean the residue and the oxide was further thinned down as shown in Fig. 3.4(b). Aluminum was applied by a blanket electron-beam evaporation, and then lithography and wet chemical etching in Cyantec Al-12 (HNO₃, HPO₃) was used to define the device contacts. The final device cross-section is shown in Fig. 3.4(c). The device $I-V$ characteristic is measured from front-to-back. A parasitic resistor (between Al and $n^+\ Si$ substrate) and a parasitic diode (between $p^+\ B$-doped Si to $n^+\ Si$ substrate) are measured. Since the TD has an area about 300 to 600x of the parasitic devices, the TD current will be much larger than those from the parasitic devices.

Figure 3.4 Esaki tunnel diode process (continued): (a) after B diffusion, (b) after B SOD residue cleaning, and following (c) Al metallization.
3.2.2 Demonstration of diffused-junction silicon tunnel diodes

Using the process flow of section 3.2.1, Si TDs were constructed as shown in Fig. 3.5. The \( I-V \) characteristic measured from top metallization to the wafer backside is shown in the inset of Fig. 3.5. The highest current density, obtained for the process condition indicated in the figure, is 0.26 \( \mu\text{A/}\mu\text{m}^2 \) (26 A/cm\(^2\)) with a PVR of 1.25.

![Graph](image)

Figure 3.5 Current–voltage characteristics of Si TDs fabricated process with SODs and proximity RTP.

In a 540 x 108 \( \mu\text{m}^2 \) area, five devices are measured with ± 5% deviation in peak voltage and ± 6% deviation in peak current. The non-uniformity could be a result of differing parasitic device areas due to process variations. It could also indicate doping
non-uniformity. From Fig. 2.9, the effective doping concentration is estimated to be $2.5 \times 10^{19}$ cm$^{-3}$, and varies by ± 0.4%. From Fig. 2.9, a $2\times$ doping density increase, from 2.5 to $5 \times 10^{19}$, is expected to increase the peak current density $100\times$, from approximately 5 to 560 A/cm$^2$. The peak current density is a sensitive indicator of doping non-uniformity, since the doping gradient should be similar for adjacent devices.

Either improving the effective dopant density or the doping gradient can increase the current density. Silicon TDs were fabricated with direct contact of boron SOD, expecting to improve dopant densities compared to proximity-diffused diodes. A control wafer was similarly processed using proximity RTP. A second wafer was processed using the same P diffusion, while the B diffusion was conducted with Borofilm100 SOD in direct contact with the device wafer. The wafer using direct contact of boron SOD produced the TDs shown in Fig. 3.6, while the control wafer did not produce TDs with negative resistance. The best measured peak current density is 0.78 µA/µm$^2$. The high peak voltage results from a resistance of approximately 70 Ω in series with the TD, possibly from the contact resistance due to the incomplete removal of the B residue [8]. The non uniformity is significantly greater than that in proximity RTP. The greater non-uniformity indicates either greater variation in parasitic devices or greater doping non-uniformity. The greater variation in parasitic devices may arise from difficulty in removing B residue layer in the direct contact experiment. The greater non uniformity in doping might be due to a stress effect, since the SOD, in direct contact with the Si, may introduce stress during the anneal.
Variable-angle-spectroscopic-ellipsometer (VASE) measurements were performed to quantify the spin-on diffusant residue-layer thickness. A Si wafer, 10-18 $\Omega\cdot$cm, 100 mm diameter, was RCA-cleaned and hydrogen-terminated in buffered HF. The native oxide was measured to be approximately 5 Angstroms after the buffered HF treatment. It was then proximity diffused at 900 $^\circ$C for 1 s with Borofilm 100. A 0.5 x 4 cm$^2$ rectangular sample was cleaved along the radius of the 4 inch wafer. A VASE measurement was conducted on the sample after 0, 1, 2, and 3 minutes buffered HF etch respectively, on four spots about 1 cm apart, as shown in the inset of Fig. 3.7. A silicon dioxide on silicon model was used to calculate the residue layer thickness. The measured
layer thicknesses are shown in Fig. 3.7. While the SiO₂ model is almost certainly incorrect, this measurement can be used to evaluate the effectiveness of the HF etch.

![Figure 3.7 VASE measurement of SOD residue layer on the silicon surface after proximity RTP and 0, 1, 2 and 3 minutes etch in buffered HF.](image)

Figure 3.7 VASE measurement of SOD residue layer on the silicon surface after proximity RTP and 0, 1, 2 and 3 minutes etch in buffered HF.

An SOD residue layer about 6-16 nm thick, was measured right after the RTP. After 2 minutes etching, an apparent layer of approximately 2 nm thickness remained, which was not removed by further etching. A similar HF-insoluble layer, possibly silicon boride, has been reported to be 10-30 nm thick, when using a liquid boron tribromide (BBr₃) source in RTP with Si [96]. Though this layer can be removed by low temperature oxidation at 820-920 ºC for 10 minutes [96], this temperature-time budget will degrade
the doping gradient. This oxidation process, though, might improve the dopant activation. Further Hall effect and resistivity measurement might be useful to evaluate this oxidation process.

Silicon TDs have also been fabricated with B diffusion heating rate of 60 °C/s, Fig. 3.8, to reduce the diffusion time, and improve the doping gradient. Peak current density up to 1.55 µA/µm² was achieved. The Borofilm 100 SOD is again in direct contact. In a 960 x 108 µm² area, 8 devices were measured with ± 10% deviation in peak voltage and ± 12% deviation in peak current, Fig. 3.8. The high peak voltage and non-uniformity might be caused by similar reasons as discussed for Fig. 3.6.

![Graph](image)

Figure 3.8 Current-voltage characteristics of Si TDs fabricated using SODs and proximity RTP, using B diffusion ramp up rate of 60 °C /s.
The highest peak current density of 1.55 µA/µm² achieved in this process is about 15x larger than that achieved in II [18]. From SIMS measurement, Fig. 2.6, a doping gradient of approximately 5 nm/decade is expected in the fabricated TDs. The better peak current density in this process could be the result of a better doping gradient from the SOD and RTP process. The highest peak current density on the order of 1 µA/µm², indicates that the effective junction doping density is $4 \times 10^{19}$ cm⁻³ for a 4 nm/dec junction, from Fig. 2.9. Either low activation of dopants or strong compensation between B and P causes the low junction effective doping concentration. Though the tunneling current density is low, these devices are applicable to detector applications.

Silicon backward TDs with curvature coefficient up to 34 V⁻¹, which is comparable to that achieved by MBE approach [35], were achieved using this diffusion-based approach. The Si wafer was B proximity diffused at 1000 °C for 90 s, followed by P proximity diffusion at 900 °C for 1 s with a heating rate of 30 °C/s. The initial design goal of this process was to achieve high doping concentrations of B and P both above $10^{20}$ cm⁻³, since B solid solubility is about $2 \times 10^{20}$ cm⁻³ at 1000 °C and P solid solubility is about $6 \times 10^{20}$ cm⁻³ at 900 °C [69]. Backward TDs resulted.

The $I-V$ characteristic and curvature coefficient for a 150 µm diameter Si TD are shown in Fig. 3.9. The curvature coefficient at zero bias is about 34 V⁻¹ with junction resistance of 740 MΩ·µm². The MBE approach has demonstrated Si backward TDs with curvature coefficient of 31 V⁻¹ and junction resistance of 140 MΩ·µm² [35]. The SODs and RTP approach provides an alternative to achieve Si backward TDs.
3.3 Lateral diffused-junction tunnel diodes

3.3.1 Lateral process flow and mask set

In the lateral formation process, the P and B dopants are both diffused into lightly $n$-type Si instead of into already heavily-doped regions. By precise placement of the junction, compensation of one dopant by another may be minimized. This process features CMP to form a self-aligned contact. Tunnel junctions are formed using SODs and proximity RTP. Lateral TDs are formed in this process, as is a parasitic vertical diode. The vertical diode could be eliminated by processing on properly thinned silicon-on-insulator (SOI) substrate.

The detailed process flow is outlined in Fig. 3.10 to Fig. 3.12. A Si wafer, P-doped,
10-18 $\Omega$·cm, and 100 mm in diameter, is RCA-cleaned and hydrogen-terminated in HF. A thermally-grown field oxide is patterned, Fig. 3.10(a), using reactive ion etching (RIE) followed by buffered 10:1 HF. Phosphorus is diffused into the Si using Emulsitone Phosphorosilicafilm $1 \times 10^{21}$ and proximity RTP, Fig. 3.10(b). The SOD residue is cleaned using buffered HF. A 300 nm thick layer of W is then sputtered over the wafer, Fig. 3.10(c).

**Figure 3.10** Scale drawing in the vertical dimension of a lateral TD process: (a) Si substrate with thermally-grown oxide layer, patterned and etched, (b) after P proximity RTP, and (c) following W sputtering.
A CMP planarizes the surface and leaves the W inside the oxide window, Fig. 3.11(a). With the W protecting the P-diffused area, the oxide is etched off in buffered HF, exposing the lightly P-doped Si, Fig. 3.11(b), for the following B diffusion process. Boron is diffused into the exposed Si using Emulsitone Borofilm 100 and proximity RTP, followed by a buffered HF etch to clean the SOD residue, Fig. 3.11(c).

Figure 3.11 Lateral Si TD process continued: (a) chemical mechanical polishing of tungsten, (b) remove oxide, and (c) boron proximity RTP

After the deposition of silicon nitride of thickness 300 nm by plasma-enhanced chemical vapor deposition (PECVD), the vias are patterned and etched in buffered HF, Fig. 3.12(a). Finally, Al is applied by a blanket electron-beam evaporation, lithography,
and wet chemical etching to define the contacts, Fig. 3.12(b). In total, four masks are used. There are three contacts for this structure allowing connection of TDs separately or in series using the second diode in its reverse tunneling direction. The equivalent circuit of the final structure is two tunnel diodes sharing a common contact, Fig. 3.12(c); note there are parasitic diodes associated with current flow in the $p^+ n$ parasitic vertical diodes.

![Figure 3.12 Lateral Si TD process continued: (a) PECVD nitride deposition and via patterning, (b) Al evaporation and contact formation, and (c) equivalent circuit of fabricated devices.](image)

A four-layer mask set was designed and masks made for patterning using a GCA6300 g-line 5:1 reduction stepper with a 1 micron resolution. The mask set contains
devices, process test structures, and circuits. The lay-out was created using the commercial CAD software Mentor Graphics. The cell occupies a square area, 3.1 mm x 3.1 mm, which is then duplicated in a 5 x 5 matrix. The CAD layout of the tunnel diode test cell is shown in Fig. 3.13. The tunnel diode device structures are designed based on rules that allow 0.5 µm, 1 µm and 2 µm layer-to-layer alignment errors.

Figure 3.13 Schematic layouts of lateral TDs and test structures. Test structures and tunnel diode devices are as indicated. Note the white polygons are just for eye guidance. They are not part of the layout design.
A typical TD structure, $lpnp_{8x40}$, is shown in Fig. 3.14(a). A cut line is displayed in the middle of the device layout, with the cross sectional view at the cut line shown beneath. In the structure $lpnp_{8x40}$, $I$ stands for the $I$ shape of the middle contact, which is designed to minimize series resistance. $Pnp$ stands for the doping type under the three contacts, $p$, $n$, and $p$ respectively, as labeled. In the numbering, the first digit, 8, stands for the width of the emitter layer in microns and the second digit 40 stands for the length of emitter layer also in microns.

![Figure 3.14 Typical tunnel diode $lpnp \times 40$ layout and the schematic cross-section.](image)

The dc measurement pad configuration is shown in Fig. 3.15(a). A microwave 50 $\Omega$ ground-signal-ground coplanar pad structure has also been designed for high frequency
characterization of the S-parameters, Fig. 3.15(b). The microwave pad is designed to fit a coplanar probe with a 100 µm pitch. The open-circuited and short-circuited test structures are designed to measure the pad characteristic, which can then be de-embedded from the tunnel diode S-parameter measurements.

Figure 3.15 (a) dc and (b) microwave measurement structure layout.

3.3.2 First demonstration of lateral diffused-junction silicon tunnel diode

Diffused-junction Si TDs were demonstrated for the first time using SODs and proximity RTP in a lateral geometry. Tunnel devices with current density of approximately 30 nA/µm² were achieved. A detailed process traveler is provided in Appendix 3.

A Si wafer was thermally-oxidized to form an approximately 350 nm field oxide. Oxide window was etched using CF₄/O₂ RIE for 6.5 minutes (etch rate approximately 45
nm/min) down to a depth of approximately 50 nm and finished by 1 minute wet etch in buffered HF (etch rate approximately 60 nm/min) giving an overetch of approximately 10 nm. A W film of approximately 400 nm was sputtered conformal onto the oxide and the exposed Si. With its high melting point (3422 °C) and high etch resistance to buffered HF, W can sustain the high junction formation temperature (~1000 °C) and buffered HF treatment used in this fabrication. The Si wafer was cleaved and the cross-section was measured using an SEM.

The cross-section of an 8 µm wide oxide window, with W on top is shown in Fig. 3.16. The lighter regions are W while the dark regions are oxide and Si. The oxide and Si edge is visible at the cleaved opening edge. The W film covers both the oxide and exposed silicon areas. The oxide sidewall slope is about 75°, which is the result of the anisotropic RIE, followed by isotropic oxide etching in buffered HF.

Figure 3.16 SEM pictures of sputtered tungsten film on oxide and silicon before CMP.
The CMP process for W was developed on a Logitech chemical de-layering and planarization (CDP) system. The CMP process used a table rotation speed of 50 rpm, a carrier head speed of 50 rpm (rotated in the same direction to assure uniformity), a head pressure of $4 \times 10^4$ Pa and a slurry flow rate of 75 ml/minute. The slurry is a mixture of Cabot SEMI-SPERSE W2000 slurry (900 ml), DI water (900 ml) and fresh 30% $\text{H}_2\text{O}_2$ (100 ml). The mixed slurry solution was hand-shaken 5 minutes to homogenize the mixture. To characterize the W polishing rate, a 400 nm W film was sputtered onto a 100 mm Si wafer. During the sputtering, an Al clip blocked one portion of the wafer edge and formed a W step. The wafer was then polished and the W step thickness was measured at different polishing time. Tungsten polishing rate was measured to be 95 nm/minute.

![SEM pictures of sputtered tungsten film on silicon after CMP.](image)

A second Si wafer was similarly processed, followed by a CMP process described above and then cleaved. The cross-section of a 4 µm wide oxide window is shown in Fig.
3.17. Tungsten on the oxide was selectively polished off and only the W in the window was left; the missing W pieces in Fig. 3.17 were likely broken out during cleaving.

On forming the lateral tunnel junction, a first experiment was performed to assess the effects of RTP temperature, time and HF cleaning time. Phosphorus was diffused at 1000 °C for 1 or 5 s using proximity RTP, followed by a cleaning in buffered HF for 60 or 75 s. Variations in buffered HF etch produce different lateral etches of the oxide window, and different placements of the junction with respect to the W edge. Boron was diffusion for 1s at 900 or 1000 °C, using SODs and proximity RTP to form the junction. The temperatures used in these diffusions were chosen to achieve high dopant concentration above $10^{20}$ cm$^{-3}$, while the diffusion times were chosen by considering the dopant diffusion depth from SIMS (Fig. 2.6 and Fig. 3.18) and phosphorus SOD diffusion experiments [97].

![Figure 3.18 SIMS profile of P diffusion into silicon using SOD and RTP at 900 °C of 1 s and 1000 °C of 1 s. The diffusion gradient is given at the concentration of $2 \times 10^{20}$ cm$^{-3}$, which is the solid solubility of B at 1000 °C.](image)
A process, P diffused at 1000 °C for 5 s, followed by a HF etch of 75 s and B diffusion at 900 °C for 1 s, produced backward TDs. A SEM picture of the left corner of W, and a schematic drawing of the lateral junction formed at that corner are shown in Fig. 3.19.

Figure 3.19 Schematic drawings of lateral junction formation process and a SEM picture at the left W corner: (a) after P diffusion, (b) after HF cleaning, W sputtering and CMP, (c) SEM picture of the left W corner corresponding to (b), and (d) after B diffusion.
In Fig. 3.19(a), phosphorus is first diffused into the opened oxide window. From SIMS measurement, the P diffusion depth is about 60-70 nm at a doping concentration of $10^{20}$ cm$^{-3}$. The P diffusion is assumed to be isotropic, and the lateral diffusion of P underneath oxide is the same as the vertical diffusion distance, Fig. 3.19(a). The top layer of Si is oxidized by phosphorus SOD as described in Section 3.1. After a 75 s HF cleaning process, this oxidized layer is removed and leaves a step in the heavily doped Si. At the same time, the field oxide is etched by about 75 nm, both vertically and laterally, Fig. 3.19(b). After W sputtering and CMP, the schematic drawing of the left W corner is shown in Fig. 3.19(b) and a SEM picture at this corner is shown in Fig. 3.19(c). The step in Si due to the SOD oxidation is obvious in Fig. 3.19(c). The field oxide is then removed and B is diffused using SOD and RTP. The diffusion of B is assumed to be isotropic too, and the diffusion length is about 5 nm at a doping concentration of $10^{20}$ cm$^{-3}$ from SIMS data. The junction is formed underneath W, perpendicular to the wafer surface, Fig. 3.19(d). A TEM picture of the interface between W and Si, no shown, indicates no tungsten silicide is formed after P and B diffusion, therefore the tunnel junction is not shorted.

The current-voltage characteristics of a lateral TD from middle contact to left contact are shown in Fig. 3.20. In order to confirm the current is tunneling current instead of thermal currents from parasitic vertical diodes, temperature dependence of the current-voltage characteristics was measured on a Cascade 11861 wafer prober with a Temptronic TP03000A thermal chuck system. The current-voltage curves of the same device from –20 to 80 °C, Fig. 3.20, show a monotonic increase in current with
temperature. The small peak-current change in the forward direction and overall small change in the reverse direction vs. temperature are consistent with a tunneling current. Since the substrate is lightly doped, the parasitic diode is a $p$-$n$ diode. The tunneling current can only be produced laterally between the $p^+$ and $n^+$ diffused Si. Tunneling current vs. temperature characteristic in Fig. 3.20 confirms a lateral TD is formed.

![Figure 3.20](image)

**Figure 3.20** $I$-$V$ curves of a Si lateral tunnel junction at various temperatures between -40 and 80°C. The device is measured from the left contact to the middle contact, as illustrated in the inset.

It is a first demonstration of Si TDs in a lateral geometry using SOD and RTP approach. The lateral tunnel junction area is computed to be 0.6 $\mu$m$^2$, by multiplying the
device perimeter of 120 µm and a junction diffusion depth of 5 nm estimated from SIMS. The peak tunneling current density is calculated to be approximately 30 nA/µm².

The P and B doping densities were computed from the sheet resistance measurements on transfer length measurement (TLM) structures. Sheet resistances of approximately 200 Ω/□ and 530 Ω/□ were measured on the n+ and p+ doped regions, respectively. For a 1000 ºC, 5 s diffusion, the P diffusion depth is approximately 60 nm from SIMS data in Fig. 3.18. The B diffusion depth is estimated about 5 nm at 1 x 10²⁰ cm⁻³, the B solid solubility at 900 ºC [69], from Fig. 2.6. The resistivities of n+ and p+ regions were computed to be 1.2 mΩ·cm and 0.3 mΩ·cm, using ρ = Rₛ·h, where ρ is the resistivity, Rₛ is the sheet resistance and h is the diffusion depth. The corresponding dopant densities are about 1 x 10²⁰ cm⁻³ for P and 4 x 10²⁰ cm⁻³ for B. Due to the limitation of solid solubility, it is thought the doping densities are 2 x 10²⁰ cm⁻³ for P and 1 x 10²⁰ cm⁻³ for B in these lateral tunnel junctions.

3.4 Limitations of the diffusion approach

Peak tunneling current density of order 1 µA/µm² is achieved in a vertical geometry TD. The junction abruptness is about 5 nm/dec, Fig. 2.6 and the effective junction doping concentration is about 4 x 10¹⁹ cm⁻³, Fig. 2.9. Based on a junction abruptness of 4 nm/dec, a junction effective doping concentration of about 1 x 10²⁰ cm⁻³, i.e. B doping of 2 x 10²⁰ cm⁻³ and P doping of 2 x 10²⁰ cm⁻³, is required to achieve 1 mA/µm². Due to the thermal equilibrium property of RTP, activated carrier concentration
greater than $1 \times 10^{20}$ cm$^{-3}$ requires temperature above 1000 °C. However, the junction abruptness degrades from 7 nm/dec at 900 °C to 13 nm/dec at 1000 °C for phosphorus SOD, Fig. 3.18. It is not likely that SODs and RTP process will meet high junction doping density above $1 \times 10^{20}$ cm$^{-3}$ and junction abruptness better than 4 nm/dec, to achieve 1 mA/µm$^2$. Nonequilibrium processes, which have achieved high dopant density above $1 \times 10^{20}$ cm$^{-3}$ and junction abruptness of 2 nm/dec, will mostly like produce 1 mA/µm$^2$ tunnel junctions in Si.
CHAPTER 4

GERMANIUM TUNNEL DIODE FABRICATION

Germanium has been investigated to form tunnel junctions since the 1960s using LPE based processes [1, 12, 15-17, 90, 91]. These junctions have been fabricated by alloying $p$-type dopants, Ga, Al or In, into a heavily $n$-doped bulk Ge substrate, or $n$-type dopants, P or As, into a heavily $p$-doped bulk Ge substrate. Both discrete devices [1, 15-17, 90, 91] and batch-fabrication process [12] for planar Ge Esaki diodes have been developed. The highest peak current density reported is up to 1 mA/µm² using Ga as the $p$-type dopant and electrical forming to create the junction [15].

In this study, both RTP from SODs and rapid melt growth approaches are used separately and in combination to form Ge TDs. Phosphorus and Ga containing SODs were used to form the $n$ and $p$ dopant regions respectively. Germanium backward TDs were successfully demonstrated using this process. To the author’s knowledge, this is the first report of Ge TDs by direct diffusion from SODs.

A rapid melt growth approach has been demonstrated for Ge TD fabrication using on-wafer liquid-phase growth of $p$-type Al-doped Ge on a heavily-doped $n$-type Ge. The LPE process uses RTP to produce the heating/cooling rates needed to form an abrupt
junction. This approach is a further development of the metal/semiconductor alloying approach of the 1960s [1, 12, 15, 16] and the recent reports of rapid melt growth of Ge by Liu et al. [98]. The n+ side of the junction is formed using P SOD and RTP. The tunnel junction is formed by depositing an Al layer, dissolving a predetermined amount Ge into the Al layer, and then regrowing a heavily-p Al-doped Ge layer.

Using this approach, a wide range of peak current densities can be achieved, controlled by the annealing temperature and time. Both backward and high current density tunnel junctions have been demonstrated. The backward Ge TDs show zero-bias curvature coefficient of approximately 20 and junction resistance of 13.5 kΩ·µm², both of which are comparable to those of commercial Ge backward diodes from M-Pulse Microwave, San Jose, CA [99]. Germanium tunnel junctions with current densities up to 1.2 mA/µm² have also been achieved. To the author’s knowledge, this is comparable to the highest reported peak current density of 1 mA/µm² for Ge TDs, fabricated using an electrical pulse forming approach [15].

4.1 Prior art vs. new approaches

A variety of LPE doping approaches have been employed to fabricate Ge TDs including metal/semiconductor alloying [1, 12, 91], electrical forming [15], laser alloying [17], and solution regrowth [16, 90]. These approaches all sought to form an abrupt junction by melting or nearly melting the surface layer and then regrowing or diffusing the heavily-doped semiconductor into the near surface layer.
Phosphorus and As were used as the \(n\)-type dopants while Ga and Al were often used as the \(p\)-type dopants [1, 12, 15-17] in prior Ge tunnel junctions. Gallium and Al, as the \(p\)-type dopants with the highest solid solubilities in Ge, can achieve doping concentrations up to \(5 \times 10^{20} \text{ cm}^{-3}\) at about 700 °C [69]. Phosphorus and As are the \(n\)-type dopants with the highest solubilities in Ge, which can achieve doping concentration of \(2 \times 10^{20} \text{ cm}^{-3}\) and \(8 \times 10^{19} \text{ cm}^{-3}\) at 550 and 800 °C, respectively [69].

Indium (eutectic temperature with Ge is 156 °C) [1] and Sn (eutectic temperature with Ge is 231 °C) [12, 15] have been used as the carrier metals to bring in the dopants during epitaxial Ge regrowth. With these carrier metals, doped Ge layers can be regrown at a low temperature (close to the metal/Ge eutectic temperature) and therefore the dopant diffusion during regrowth is minimized. Gallium (eutectic temperature with Ge is 29 °C) has also been used as the carrier metal and dopant at the same time [15]. Gallium is normally in a liquid form at room temperature, and has been deposited by plating [15].

Germanium TDs with peak tunneling current densities up to 1 mA/µm² have been reported by Burrus [15] using electrical forming. For detector applications, zero-bias curvature coefficients up to 70 V⁻¹ [91] have been reported on Ge backward TDs, and this potential further motivates the investigation of tunnel diodes in Ge. For backward TD fabrication, both diffusion-based and on-wafer liquid-phase approaches are investigated. These two approaches are briefly described below and more details are provided in the next section.

The difference between the diffusion-based approach and prior LPE based
approaches is the junction formation process. The difference between the on-wafer liquid-phase regrowth approach and prior LPE approaches is that the heavily-doped-\( n \) side of the tunnel junction is formed using SOD and RTP, in place of a heavily \( n \)-doped bulk substrate [1, 12, 15-17]. In addition, a microcrucible layer has later been utilized. It improves the surface morphology by confining the melt [100].

4.2 Demonstration of diffused-junction Ge TDs

4.2.1 Doping Ge from spin-on diffusants

Unless otherwise stated, the doping sources used in this work were Ga250 from Filmtronics and Phosphorosilicafilm \( 1 \times 10^{21} \) from Emulsitone. The diffusant Ga250 contains 2.5 weight percent Ga with SiO\(_2\) as a supporting matrix dissolved in an organic solvent [101], similar to Phosphorosilicafilm \( 1 \times 10^{21} \) described in section 3.1.2.

A Ga-doped \( p \)-type, 0.18-0.19 \( \Omega \cdot \text{cm} \), (100) Ge substrate, was first doped heavily \( n \)-type using phosphorus SOD and RTP. The wafer was first degreased in acetone and methanol for 5 minutes each. The native oxide was removed by dipping the wafer in 10:1 buffered HF for 25 s and rinsed in deionized (DI) water for 60 s. The wafer was then spin coated with phosphorus SOD. The volatile organics in the SOD was removed by baking on a hot plate at 150 °C for 10 minutes in air. The sample was next loaded into a Modular Process Technology RTP600 rapid thermal processor for P diffusion.

To investigate the P doping concentration dependence on annealing temperature, four \( 1 \times 1 \text{ cm}^2 \) samples were cleaved from the P SOD-coated wafer and rapid thermal
annealed at four different temperatures in a nitrogen ambient. The highest anneal temperature tested, 850 °C, is about 80 °C lower than the melting point of Ge. An annealing time of 300 s, which is the maximum time allowed by the RTP at this temperature, was used to provide the greatest P incorporation within the limits of the RTP. Annealing temperatures of 500, 600, and 700 °C were also tested for comparison. Annealing times up to 15 minutes were used at these lower temperatures, which is the maximum time allowed at 700 °C.

The activated P concentrations in the diffused samples were determined using Van der Pauw resistivity and Hall effect measurements. Indium contacts were soldered onto the four corners of the sample, and the contacts were confirmed to be ohmic using a Tektronix 370 curve tracer. The Van der Pauw resistivity and Hall effect measurements were then performed at three current densities of 0.05, 0.5 and 5 mA, respectively. The measured sheet carrier concentrations and Hall mobilities are within 5% at these three currents.

The annealing temperatures, times, doping types, measured Hall mobilities and sheet carrier concentrations vs. different temperatures and times are listed in Table 4.1. Annealing processes above 700 °C can dope Ge to heavily n-type using SODs. As an example, an annealing at 850 °C for 300 s converts the Ge surface region from p-type into heavily n-type. In contrast, after an annealing process under 600 °C, the Ge wafer surface region is still p-type. The possible reason is that P in the Phosphorosilicafilm $1 \times 10^{21}$ SOD is bonded to oxygen atoms in a $P_4O_{10}$ or $P_2O_5$ form [61] before diffusion. Certain amount
of thermal energy is required to break the P-O bonds, which will then release P for diffusion. Annealing below or at 600 °C might not be enough to break the P-O bonds and therefore no diffusion of P occurs.

Phosphorus carrier density can be estimated by dividing measured sheet carrier density by diffusion depth. The diffusion depth is calculated by $\sqrt{Dt}$ using the reported P diffusivities of $4.5 \times 10^{-13}$ and $2.1 \times 10^{-11}$ cm$^2$/s at 700 and 850 °C [102], respectively. The carrier density, $n_{Hall}$, is computed by dividing the measured sheet carrier concentration by the calculated diffusion depth. The solid solubilities, calculated diffusion depths, computed carrier densities $n_{Hall}$ from sheet carrier concentrations are also listed in Table 4.1. The computed P carrier density is $25 \times 10^{20}$ cm$^{-3}$ at 700 °C and $8.9 \times 10^{20}$ cm$^{-3}$ at 850 °C. Both these two values are much higher than the maximum solid solubility, $2 \times 10^{20}$ cm$^{-3}$, of P in Ge. The estimated diffusion depth may be too small and causes this discrepancy.

Phosphorus carrier density may also be estimated from measured Hall mobility. Fistul et al. [103], from Hall effect measurements on heavily As-doped single crystal Ge with carrier densities up to $4 \times 10^{19}$ cm$^{-3}$, has come up with an empirical formula

$$\mu = 1.52 \cdot 10^{10} \cdot n_{Fistul}^{-0.4}, \quad (4.1)$$

where $n_{Fistul}$ is the electron carrier concentration and $\mu$ is the electron Hall mobility. Using Eq. (4.1), the computed carrier density from Hall mobility is listed in Table 4.1. The computed carrier density, $10 \times 10^{20}$ cm$^{-3}$, is higher than the solid solubility of $2 \times 10^{20}$ cm$^{-3}$. Since Eq. (4.1) is only valid for a limited carrier concentration range up to $4 \times 10^{19}$ cm$^{-3}$. 

73
cm⁻³ (Hall mobility 220 cm²/V-s), it may not predict an accurate carrier density at a much lower Hall mobility of 45 to 50 cm²/V-s.

**TABLE 4.1**

SUMMARIES OF MEASURED DOPING TYPES, HALL MOBILITIES \( \mu_{\text{Hall}} \), SHEET CARRIER CONCENTRATIONS \( n_{\text{sheet}} \), CALCULATED DIFFUSION DEPTHS \( \sqrt{D_t} \), COMPUTED CARRIER DENSITIES \( n_{\text{Hall}} \) FROM SHEET CARRIER CONCENTRATION, \( n_{\text{Fistul}} \) FROM HALL MOBILITY, SOLID SOLUBILITIES \( n_{\text{ss}} \) [69] AND Ge SURFACE ROUGHNESS (SR) UNDER DIFFERENT ANNEALING TEMPERATURES AND TIMES.

<table>
<thead>
<tr>
<th>Temp. (°C)</th>
<th>Time (s)</th>
<th>Doping type</th>
<th>( \mu_{\text{Hall}} ) (cm²/V-s)</th>
<th>( n_{\text{sheet}} ) (10¹⁵ cm⁻²)</th>
<th>( \sqrt{D_t} ) (µm)</th>
<th>( n_{\text{Hall}} ) (10²⁰ cm⁻³)</th>
<th>( n_{\text{Fistul}} ) (10²⁰ cm⁻³)</th>
<th>( n_{\text{ss}} ) (10²⁰ cm⁻³)</th>
<th>SR (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>300</td>
<td>p</td>
<td>1459</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>600</td>
<td>900</td>
<td>p</td>
<td>930</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>700</td>
<td>900</td>
<td>n</td>
<td>45</td>
<td>50</td>
<td>0.2</td>
<td>25</td>
<td>10</td>
<td>1.2</td>
<td>12</td>
</tr>
<tr>
<td>850</td>
<td>300</td>
<td>n</td>
<td>50</td>
<td>71</td>
<td>0.79</td>
<td>8.9</td>
<td>10</td>
<td>0.8</td>
<td>-</td>
</tr>
</tbody>
</table>

* The Ge wafer in direct contact with P SOD. Carrier density is not calculated for 500 and 600 °C experiments since the doping type remains p-type as the substrate.

Both methods can only provide a crude P carrier density estimation. Since mobility decreases with the increase of doping concentration due to more ionized-impurity scattering, the P carrier densities of samples diffused at 700 and 800 °C, Table 4.1, are very likely to be higher than \( 4 \times 10^{19} \) cm⁻³, with Hall mobilities lower than 220 cm²/V-s. In addition, the measured electron Hall mobilities are relatively constant, about 50 cm²/V-s, in these two samples. A similar constant electron mobility has been
reported by Thurber et al. [104] in Si with phosphorus doping above $10^{20}$ cm$^{-3}$. It is possible that the P carrier density in the SOD and RTP doped Ge surface is above $10^{20}$ cm$^{-3}$, up to the solid solubility of $2 \times 10^{20}$ cm$^{-3}$.

The Ge surface is roughened during P diffusion from SODs, Table 4.1. Prior to annealing, the Ge surface is specular, afterwards the surface turns misty. The surface roughness is measured using an alpha-step surface profiler and the measured differences between the maximum and minimum profile heights are listed in Table 4.1. Roughness from 12 to 18 nm have been measured uncorrelated with annealing temperatures and times.

It has been reported that desorption of GeO$_x$ occurs at temperatures from 430 to 500 ºC [105, 106]. On a free Ge surface, approximately 100 nm of Ge can be removed by RTP at 600 ºC for 60 s in a nitrogen ambient [105], by a reaction of Ge + O $\rightarrow$ GeO$_x$. A possible reason for the rough surface is oxidation of the Ge surface and subsequent desorption of the Ge oxide during RTP, with the phosphorus SOD in direct contact with the Ge surface. Two oxygen sources, the residue oxygen in the nitrogen gas used in RTP and the oxygen in the SOD in the form of P$_4$O$_{10}$ or P$_2$O$_5$, can contribute to the Ge surface oxidation. The phosphorus SOD may block the absorption of ambient oxygen onto the Ge surface. However, the Ge surface could also be locally oxidized by the SOD by a reaction of Ge + P$_2$O$_5$ $\rightarrow$ GeO$_x$ + P during the diffusion, similar to the reaction between SOD and Si in section 3.1.2. Desorption of GeO$_x$ at temperatures above 500 ºC then leads to surface roughness.
Tunnel diodes have depletion widths typically less than 5 nm, therefore a surface roughness on the order of 10 nm can lead to doping non-uniformity in the junction. In order to reduce the surface roughness, the oxygen available for Ge oxidation could be lowered by adding a hydrogen containing ambient in a proximity RTP approach to get the SOD out of contact with the Ge surface. Adding hydrogen in the ambient should lower the residue oxygen by reaction between hydrogen and oxygen to form water. Proximity RTP avoids the local reaction between SOD and Ge. A Ge wafer, which has been proximity annealed with Filmtronics P507 SOD (5 x 10^{20} \text{ cm}^{-3} \text{ P concentration}) in a forming gas ambient (90 N_2:10 H_2), has been reported to visually maintain a specular surface after diffusion [107]. However, no quantitative measurement of surface roughness has been made [107].

To reduce surface roughness, two 1 x 1 cm^2 samples were cleaved from a P SOD-coated Ge wafer. The control sample was proximity-RTP annealed at 850 °C and the other at 750 °C, both for 300 s in a forming gas ambient (95 N_2:5 H_2). Similar to the report by Posthuma [107], the Ge wafers after P diffusion preserved specular surfaces. The P carrier densities were determined by Van der Pauw resistivity and Hall effect measurements using the same procedure described for Table 4.1.

Compared with Table 4.1, a higher electron Hall mobility, i.e., a lower P carrier density is measured. In both annealing processes, the Ge surfaces were converted from \( p \)-type into heavily \( n \)-type. The control sample, after proximity RTP at 850 °C for 300 s, shows a lower sheet carrier of concentration of \( 10^{16} \text{ cm}^{-2} \) and a higher electron Hall
mobility of 153 cm$^2$/V-s, compared with the sample in Table 4.1. The computed P carrier concentration is approximately $1.3 \times 10^{20}$ cm$^{-3}$ using sheet carrier concentration, and is approximately $1 \times 10^{20}$ cm$^{-3}$ using Eq. (4.1), which agrees with the solid solubility of approximately $0.8 \times 10^{20}$ cm$^{-3}$ at this temperature. Spitzer et al. [108] has reported electron Hall mobility of approximately 290 cm$^2$/V-s in a pulled single crystal Ge with P concentration of $4.1 \times 10^{19}$ cm$^{-3}$, which agrees with the measured Hall mobility and estimated P carrier density in the 750 ºC sample. The better agreements achieved in Table 4.2 than in Table 4.1 indicate that Fistul’s equation provides good estimation with doping density up to $5 \times 10^{19}$ cm$^{-3}$ but will likely fail when the doping density is above $10^{20}$ cm$^{-3}$. The lower carrier density in the proximity diffused sample is probably due to the limited dopant source available [61, 62].

### TABLE 4.2

SUMMARIES OF MEASURED DOPING TYPES, MOBILITIES $\mu_{\text{Hall}}$, SHEET CARRIER CONCENTRATION $n_{\text{sheet}}$, CALCULATED DIFFUSION DEPTHS $\sqrt{Dt}$, COMPUTED CARRIER DENSITIES $n_{\text{Hall}}$ FROM SHEET CARRIER CONCENTRATION, $n_{\text{Fistul}}$ FROM HALL MOBILITY, AND SOLID SOLUBILITIES $n_{\text{ss}}$ UNDER DIFFERENT ANNEALING TEMPERATURES AND TIMES.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Time (s)</th>
<th>Doping type</th>
<th>$\mu_{\text{Hall}}$ (cm$^2$/V-s)</th>
<th>$n_{\text{sheet}}$ (10$^{15}$ cm$^{-2}$)</th>
<th>$\sqrt{Dt}$ (µm)</th>
<th>$n_{\text{Hall}}$ (10$^{20}$ cm$^{-3}$)</th>
<th>$n_{\text{Fistul}}$ (10$^{20}$ cm$^{-3}$)</th>
<th>$n_{\text{ss}}$ (10$^{20}$ cm$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>750</td>
<td>300</td>
<td>n</td>
<td>210</td>
<td>7.3</td>
<td>0.23</td>
<td>3.1</td>
<td>0.45</td>
<td>1</td>
</tr>
<tr>
<td>850</td>
<td>300</td>
<td>n</td>
<td>153</td>
<td>10</td>
<td>0.79</td>
<td>1.3</td>
<td>1</td>
<td>0.8</td>
</tr>
</tbody>
</table>

* The Ge wafers were doped using proximity RTP and P SOD.
4.2.2 Diffused Ge tunnel diodes process

The Ge process is described as follows. Chui et al. [55, 109] have degreased Ge in organic solution PRS1000 and hydrogen-terminate it with 50:1 buffered HF in Ge MOS transistor fabrication. In this work, a Ga-doped (0.18-0.19 $\Omega \cdot cm$), 50 mm diameter, (100) Ge wafer was degreased in acetone and methanol at room temperature for 5 minutes. The native oxide was removed by two cycles of dipping in 10:1 buffered HF for 25 s and DI water for 60 s.

Phosphorus was diffused into Ge using proximity RTP at 850 ºC for 300 s with a heating rate of 30 ºC/s in a forming gas ambient, Fig. 4.1. Unless otherwise stated, the source wafer was an RCA-cleaned Si wafer and following RTP, and an etch in 10:1 buffered HF for 5 minutes was used to remove the SOD remnants. A junction depth, $\sqrt{Di}$, of approximately 700 nm is estimated from the P diffusivity of $2.1 \times 10^{-11}$ cm$^2$/s at 850 ºC [102] for the P diffusion, Fig. 4.1(a).

The Ga SOD source was then directly spun on the Ge wafer surface and baked at 150 ºC for 10 minutes in air to remove the SOD solvents. A rapid thermal diffusion, at 825 ºC for 1 s with a heating rate of 30 ºC/s, in a forming gas ambient, was used to diffuse Ga into the heavily n-doped region and form the p-n junction, Fig. 4.1(b). After buffered HF cleaning, Al was applied by blanket electron-beam evaporation, and lithography and wet-chemically-etched in Cyantec Al-12 to define the device contacts. The device mesa was etched with 30% H$_2$O$_2$ using the Al contact as an etch mask, Fig. 4.1(c), to isolate the tunnel junction.
A one-layer photo mask with the test structure shown in Fig. 4.2(a), repeated across a 4 x 4 inch plate, was used to form tunnel devices. Square emitters, with side length from 5 to 40 µm as labeled on the left side of the layout, are surrounded by a large rectangular pad with width of 84 and length of 384 µm. The separation between metals is 2 µm in Fig. 4.2(a). Structures with 4 and 10 µm spacing are also included which have higher yield. Both $I$-$V$ and $C$-$V$ measurements are carried out by putting probe 1, Fig. 4.2, on the square mesa and probe 2 on the large rectangular mesa to measure two TDs in series. The equivalent circuit of the measured devices is shown in Fig. 4.2(b), where the
tunnel diode is represented by the symbol $\text{□}$ $\text{□}$. The physical size difference of the two diodes is schematically represented by the differing symbol sizes. When the small square mesa device, diode 1, is forward biased, a positive voltage is applied on probe 1 and ground is connected to probe 2, the large area device. In this polarity, diode 2, is reverse biased and conducts with low voltage drop in the Zener tunneling direction. The area of diode 2 is larger by a factor of approximately 17x than the 40 x 40 µm² device and 30x greater than the 30 x 30 µm² device to reduce its voltage drop by a similar factor relative to the small area device under test. A process traveler is provided in Appendix 3.

Figure 4.2 (a) Schematic mask layout for Ge TD fabrication using SODs and RTP process, (b) equivalent circuit of measured devices, with device size symbolizing area difference between diode 1 and 2.
4.2.3 Diffused-junction Ge tunnel diodes

In this section, the successful demonstration of diffused-junction Ge TDs will be shown. In the experiments, a Ge wafer was first P-doped using proximity RTP at 850 °C for 300 s in a forming gas ambient. The P carrier concentration is expected to be $1 \times 10^{20}$ cm$^{-3}$, Table 4.1. After spinning on Ga SOD, the Ge wafer was subdivided into three samples. These samples were then individually diffused with Ga using RTP.

On the incorporation of Ga to form the $p^+$ surface layer, an experiment was performed to assess the effect of RTP anneal temperature. Anneal temperatures of 750, 825 and 900 °C are chosen based on diffusion depth and solid solubility estimations. The diffusivity of Ga in Ge is approximately $1.7 \times 10^{-14}$, $1.8 \times 10^{-13}$, and $1.3 \times 10^{-12}$ cm$^2$/s at 750, 825 and 900 °C, respectively [102]. The diffusion depth for 1 second diffusion at these three temperatures, calculated from $\sqrt{Dt}$, is approximately 1.3, 4.2 and 11.4 nm, respectively. The corresponding solid solubility of Ga in Ge at these three temperatures is approximately $5 \times 10^{20}$, $4 \times 10^{20}$, and $2 \times 10^{20}$ cm$^{-3}$ [69]. Among the three conditions, the 750 °C sample is expected to achieve the highest tunneling current with the highest solid solubility and the best Ga diffusion slope due to the lowest anneal temperature. The current-voltage characteristics of Ge diffused-junction diodes fabricated after three Ga diffusion processes are shown in Fig. 4.3.
Figure 4.3 Device current-voltage characteristics after three Ga proximity RTP processes at 750, 825 and 900 °C for 1 s, respectively, with (a) current plotted on logarithm scale, and (b) current plotted on linear scale. 750 °C’s result is reduced by 1000x.

In Fig. 4.3(a), the $I-V$ characteristics of a 40 x 40 µm$^2$ device in each process are shown with the current plotted on a logarithm scale, and in Fig. 4.3(b), the same device $I-V$ characteristics are plotted with the current on a linear scale, which shows the NDR region of tunnel diodes more distinctly.

The 750 °C, 1 s Ga diffusion process resulted in an ohmic $I-V$ characteristic instead of the best TD as expected. At this temperature, assuming Ga concentration is at the solid solubility of $5 \times 10^{20}$ cm$^{-3}$ and the P concentration is $5 \times 10^{19}$ cm$^{-3}$ from Table 4.1, from one-sided abrupt junction approximation, a 2.5 nm junction depth is needed to support the depletion region. Ga diffusion depth of 1.3 nm at 750 °C is not sufficient to form the $p$-$n$ junction. Without the 3 nm or so junction depth to support the field, a resistor like $I-V$ is therefore not surprising as indicated for the 750 °C diffusion.
Negative differential resistance is observed for both polarities of the \( I-V \) characteristic of the 825 °C annealed sample, Fig. 4.3, at bias points of about -0.1 V and 0.05 V. At the 0.05 V bias point, the small area mesa tunnel junction (Fig. 4.2), is forward biased and the large area rectangular tunnel junction is reversed biased. The NDR observed at 0.05 V is the characteristic of the small area mesa tunnel junction. Under -0.1 V bias, the large area tunnel junction is forward biased, and exhibits NDR. The ratio of tunneling currents in the two bias polarities, 17.3, is in agreement with the area ratio of approximately 16.1.

Backward TDs without NDR regions were achieved by the 900 °C, 1 s Ga diffusion process, Fig. 4.3(a). From the 900 °C \( I-V \) characteristic, the voltage shift from the ideal diode 60 mV/dec slope is about 0.5 V at a current density of 1.5 mA, Fig. 4.3(a). A series resistance of approximately 33 Ω is calculated. The extra series resistance could come from spreading resistance and the contact resistance. At 900 °C, the Ga diffusion depth of 11.4 nm increases the tunneling distance and therefore decreases the tunneling current. The diffusion depth of 4 nm at 825 °C is sufficient to support the junction field. Compared with junctions diffused at 900 °C, the Ga diffusion profile could be more abrupt with the lower diffusion temperature. Higher tunneling current is therefore expected and measured.

Capacitance-voltage measurement was used to estimate the junction doping concentration as well as the junction width for the 825 °C, 1 s process. Device impedance measurements were performed on an Agilent 4294A precision impedance analyzer. The
conductance extracted from the impedance measurement, which models a diode by a parallel capacitance and conductance, is first compared with the conductance from the dc $I$-$V$ characteristic, i.e. the derivative $dI/dV$. Discrepancy was observed between conductance from impedance measurement and dc current derivative, $dI/dV$, for devices with NDR in the negative bias region, possibly due to oscillation caused by NDR or a simple parallel resistance/capacitance model is not valid for this kind of devices. Therefore a $30 \times 30 \mu m^2$ Ge TD from the $825 \, ^\circ C$, 1 s process, with no NDR in the reverse bias region, is chosen and its dc $I$-$V$ characteristic is shown in Fig. 4.4(a). At 1 MHz, the device impedance was measured with dc bias from -0.2 to 0.2 V. Good agreement is achieved, Fig. 4.4(b), as would be expected if the parallel resistance/capacitance model provides a good approximation to the device impedance.

![Figure 4.4](image)

Figure 4.4 (a) Germanium TD $I$-$V$, and (b) good agreement achieved between the extracted conductance and the derivative of dc $I$-$V$ measurement.
The junction capacitance $C_j$ and junction width can be extracted from impedance measurement. The impedance was measured at bias points from 0 to 300 mV with a step of 50 mV from 10 kHz to 10 MHz. The measured real and imaginary parts of the impedance are fitted using the TD model shown in the inset in Fig. 4.5. As an example, the measured impedance of the diffused Ge TD in Fig. 4.4 at zero bias is shown in Fig. 4.5. Excellent agreement is achieved between measured data (• and ◦) and fits (solid lines). From the extracted capacitance, a depletion region of approximately 16 nm at zero bias is calculated.

![Figure 4.5](image-url)

Figure 4.5 Real and imaginary part of the impedance of a diffused Ge TD at room temperature, (◦) represents the imaginary part and (•) represents the real part; the solid and dashed curves are calculated model fits to the tunnel diode equivalent circuit, inset.
The junction doping concentration can be estimated from the slope of the $1/C_j^2$ vs. bias curve, Fig. 4.6, i.e. $d(1/C_j^2)/dV = 2/(q\varepsilon_s N_{\text{eff}})$, where $d(1/C_j^2)/dV$ is the slope of the $1/C_j^2$ vs. bias curve. A junction effective doping concentration $N_{\text{eff}}$ of $4.5 \times 10^{18}$ cm$^{-3}$ is calculated. Since the substrate electron carrier concentration $N_D$ is about $1 \times 10^{20}$ cm$^{-3}$, Table 4.2, the acceptor concentration $N_A$ is computed to be $5 \times 10^{18}$ cm$^{-3}$, using $N_{\text{eff}} = N_D N_A/(N_D+N_A)$. An abrupt junction with these doping concentrations has a depletion width of approximately 17 nm, in good agreement with the $C-V$ data. The built-in voltage is computed to be about 0.6 V from the intersection of the curve on the horizontal axis.

![Graph](image)

**Figure 4.6** Inverse junction capacitance squared vs. bias for the diffused Ge TD.
The acceptor carrier concentration estimated above, $5 \times 10^{18}$ cm$^{-3}$, is after Ga being compensated by P. The Ga concentration, without P compensation, would be approximately $1 \times 10^{20}$ cm$^{-3}$, which is lower than $4 \times 10^{20}$ cm$^{-3}$, the solid solubility of Ga in Ge at 825 ºC. A possible reason for this much lower acceptor concentration is that Ga has diffused into Ge at a high concentration but has not activated. Another possible reason is that the Ga available is limited by the dopant supply from the Ga SOD.

Based on these three diffusions of Ga into P-doped Ge, the way to improve diffused-based tunnel diodes toward achieving 1 mA/cm$^2$ peak current density using the SODs and RTP does not look promising. Increasing the Ga diffusion temperature higher than 825 ºC could increase the hole carrier concentration but the tunneling current density will likely degrade unless a shorter diffusion time is utilized. For lamp diffusions, that time is already set 1 s, and could be decreased by only 1 s to a spike anneal. Decreasing the Ga diffusion temperature should improve the gradient, but this cannot be decreased significantly since the dopant density must support the depletion region. Non-thermal-equilibrium processes, such as flash lamp annealing, can achieve high activated Ga concentrations and abrupt junction at the same time, and could be a method to improve the tunneling current density.

Though the tunneling currents are low, the diffused Ge tunnel junctions might be used in microwave detector application. In Fig. 4.7, the $I$-$V$ characteristic of a diffused Ge tunnel junction is shown from -0.1 to 0.02 V. A zero-bias curvature coefficient of 54 V$^{-1}$ and zero-bias junction resistance of 36 MΩ·µm$^2$ have been achieved.
Figure 4.7 $I$-$V$ characteristic and zero-bias curvature coefficient of a Ge TD fabricated by RTP from SODs.

To the author’s knowledge, this is the first Ge tunnel junction demonstrating zero-bias curvature coefficient above 50 using a diffusion approach. For detector application, large curvature coefficient and small junction resistance is desired, as discussed in section 2.8. The zero-bias curvature coefficient $\gamma$, junction resistance $R_j$, junction capacitance $C_j$, and cutoff frequency $f_{ct}$ of Si backward diodes in Ref. [35] and this work (section 3.2), and of Ge backward diodes from M-Pulse Microwave Inc. [99] and this work are listed in Table 4.3. The low frequency sensitivity $\beta$ is calculated using Eq. (2.8), assuming $f_{ct} \ll f_{ct}$, $R_s = 1 \Omega \ll R_j$ and a free space source impedance $Z = 377 \Omega$. 

88
Compared with Si backward diodes, Ge backward diode in this work is superior with higher cutoff frequency and higher low frequency sensitivity. Compared with commercial Ge backward diodes [99], Ge backward diode in this work has higher low frequency sensitivity, but a much lower cutoff frequency.

### TABLE 4.3

**SUMMARIES OF ZERO-BIAS CURVATURE COEFFICIENT $\gamma$, JUNCTION RESISTANCE $R_j$, JUNCTION CAPACITANCE $C_j$, CUTTOFF FREQUENCY $f_{ci}$, AND LOW FREQUENCY SENSITIVITY $\beta$ OF Si BACKWARD DIODES IN REF. [35] AND THIS WORK (SECTION 3.2), AND OF COMMERCIAL Ge BACKWARD DIODES [99] AND THIS WORK.**

<table>
<thead>
<tr>
<th>Type</th>
<th>Approach</th>
<th>$\gamma$ (V$^{-1}$)</th>
<th>$R_j$ (M$\Omega \cdot \mu$m$^2$)</th>
<th>$C_j$ (fF/\mu m$^2$)</th>
<th>$f_{ci}$ (GHz)</th>
<th>$\beta$ ($10^4$ V/W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si [35]</td>
<td>MBE</td>
<td>31</td>
<td>140</td>
<td>9</td>
<td>1.5</td>
<td>2.3</td>
</tr>
<tr>
<td>Si [this work]</td>
<td>SOD/RTP</td>
<td>34</td>
<td>740</td>
<td>20 *</td>
<td>0.3</td>
<td>2.5</td>
</tr>
<tr>
<td>Ge [99]</td>
<td>Alloy</td>
<td>16</td>
<td>0.018</td>
<td>12 *</td>
<td>98</td>
<td>1.2</td>
</tr>
<tr>
<td>Ge [this work]</td>
<td>SOD/RTP</td>
<td>54</td>
<td>36</td>
<td>8.9</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

* Obtained from simulations of an abrupt Si TD with symmetric $n$ and $p$ doping density of $10^{20}$ cm$^{-3}$, and an abrupt Ge TD with symmetric $n$ and $p$ doping density of $2 \times 10^{19}$ cm$^{-3}$ using BandProf by W. R. Frensley.
4.3 Demonstration of on-wafer liquid-phase-epitaxy Ge tunnel diode process

4.3.1 On-wafer liquid-phase-epitaxy approach

An on-wafer liquid-phase epitaxy regrowth process was developed to fabricate Ge TDs. A Ge wafer was first doped heavily n-type (above $5 \times 10^{19}$ cm$^{-3}$) using Emulsitone’s Phosphorosilicafilm $1 \times 10^{21}$ SOD and RTP. Aluminum was thermally evaporated onto the Ge wafer followed by RTP to form a melt from which the $p^+$ side of the Ge junction is grown. In this process, the Al acts both to dissolve and regrow the Ge, and it acts as the acceptor dopant. The process and experimental results are accepted for publication in the Journal of the Electrochemical Society in a paper entitled “Rapid melt growth of germanium tunnel junctions” (see Appendix 4).

The tunnel junction formation process can be understood using the Al-Ge binary phase diagram shown in Fig. 4.8. Upon heating Al in contact with Ge to the eutectic temperature $T_E = 420$ °C, Al and Ge would be expected to react to form a liquid with a Ge weight percentage of 51.6%. Above the eutectic temperature, the weight percentage of Ge in the Al increases along the liquidus line until a peak temperature $T_P$ is reached. Ge will diffuse into the Al which will become molten at an atomic percentage of $C_1$. As the temperature decreases, the Al-Ge melt composition will follow the liquidus curve, and the excess Ge freezes back out onto the substrate. The initial epitaxial layer to freeze will be Ge with an Al atomic percentage equal to $1-C_2$, where $C_2$ is the atomic percentage of Ge in the AlGe alloy. This weight percentage of Al is 0.28% at 600 °C, increasing to 0.4% as the temperature is lowered to the eutectic temperature, 420 °C.
At the peak anneal temperature the ratio of the penetration depth, $x$, of the melt below the original Ge surface (defined in the inset of Fig. 4.9) to the deposited Al thickness, $t_{Al}$, can be determined from the Al-Ge binary phase diagram of Okamoto [110] according to the relation,

$$\frac{x}{t_{Al}} = \frac{\rho_{Al}}{\rho_{Ge} \left(1 - W_{Ge}\right)} W_{Ge},$$  \hspace{1cm} (4.2)

where $\rho_{Al}$ and $\rho_{Ge}$ are the densities in g/cm$^3$ for Al and Ge respectively, and $W_{Ge}$ is the weight percent of Ge in the Al melt, given by the point on the liquidus line at $T_P$. This relation is plotted in Fig. 4.9. As an example, with an Al film of 100 nm and anneal
temperature of 600 °C, approximately 113 nm Ge layer is melted. When cooled down below 420 °C, approximately 52 nm Ge is consumed to form the AlGe eutectic and the rest 61 nm excess Ge will regrow epitaxially with Al concentration of approximately $2.2 \times 10^{20}$ cm$^{-3}$ on the Ge substrate.

The regrown layer thickness, which is controlled by the Al thickness and annealing temperature, is sufficiently thick to contain the depletion region. From Poisson solution for this structure, it is estimated that 4 to 5 nm of regrown layer is required. Calculation in Fig. 4.9 indicates that annealing temperature higher than 440 °C is required to form a regrown layer thickness than 5 nm, assuming the deposited Al is 100 nm.

Figure 4.9 The ratio of the penetration depth, $x$, to the Al deposition thickness, $t_{Al}$, as a function of peak anneal temperature, $T$, and the acceptor doping density as a function of the anneal temperature as determined from the Al-Ge binary phase diagram [110].
In addition to the regrown layer, the diffusion depths of Al into the $n^+$ Ge substrate and P into the regrown $p^+$ layer during the regrowth process, will broaden the depletion layer width and reduce the tunneling current [12]. The diffusivities of P in Ge are experimentally measured to be about $2.8 \times 10^{-16}$, $1.7 \times 10^{-14}$ and $4.5 \times 10^{-13}$ cm$^2$/s at 500, 600 and 700 ºC [102], respectively, while Al diffusivities are $1.42 \times 10^{-19}$, $3.49 \times 10^{-17}$ and $2.75 \times 10^{-15}$ cm$^2$/s at these three temperatures [102]. The diffusion depths for P and Al, calculated from $\sqrt{Dt}$, are plotted in Fig. 4.10, where $t$ is the diffusion time. Since Al has a diffusion coefficient more than 100x less than P, the junction abruptness could be mainly determined by P diffusion, if there were not a co-doping interaction between Al and P.

![Figure 4.10 Phosphorus and Al diffusion in Ge, calculated using $\sqrt{Dt}$](image)

Figure 4.10 Phosphorus and Al diffusion in Ge, calculated using $\sqrt{Dt}$.  

93
4.3.2 Process flow

A Ge wafer was prepared in the same way as described in section 4.2.2 using P-2 and proximity RTP for the donor doping. An Al metal film of 100 nm thick, as the p-type dopant, was applied by blanket electron beam evaporation, and lithography and wet chemically etched in Cyantec Al-12. A cross-sectional diagram for the structure after the evaporation and patterning of Al is shown in Fig. 4.11(a). Germanium mesas were defined using 30% H₂O₂ wet etching using the Al as an etch mask, Fig. 4.11(b). Tunnel junctions were formed by epitaxially regrowing a heavily Al-doped Ge layer on the heavily n-doped Ge substrate using RTP in a forming gas ambient, Fig. 4.11(c) and (d).

![Diagram](image)

Figure 4.11 Schematic device cross section of the rapid melt growth process drawn to scale in the vertical direction. (a) Germanium is doped n⁺ by rapid thermal diffusion of phosphorus followed by aluminum deposition, patterning, and (b) etching. (c) At a peak anneal temperature of 600 °C with a 100 nm Al cap, 110 nm of Ge is consumed. (d) On cooling, a heavily-doped p⁺ Ge layer of 60 nm is grown.

A one-layer photo mask was designed with the test structure shown in Fig. 4.12 repeated across a 4 x 4 inch plate. The large rectangular mesa area is approximately
150000 μm². Both $I-V$ and $C-V$ measurements was carried out by putting Probe 1 on the square mesa and Probe 2 on the rectangular mesa to measure two tunnel diodes in series.

Figure 4.12 Schematic layout for Ge TDs using on-wafer liquid-phase regrowth process. The spacing of square structure to the outside rectangular feature is 2, 4, 10 and 20 μm, as labeled on the top.

4.3.3 Demonstration of Ge TDs using on-wafer liquid-phase regrowth approach

Germanium TDs with peak current densities from 1 to 2.7 μA/μm² were obtained, when the regrowth temperature was varied from 440 to 700 °C. Germanium TDs fabricated by on-wafer liquid-phase regrowth with a 440 °C, 5 s anneal are shown in Fig. 4.13. Two diodes from the 440 °C, 5 s anneal are shown on the left, with peak current density of about 2 μA/μm² and PVR about 1.1. The devices sizes, measured under optical microscope, are 5.5 x 5.5 μm² and 3.6 x 3.6 μm² instead of nominal 5 x 5 μm² and 3 x 3 μm², possibly due to process variations. The Ge TDs demonstrated zero bias curvature
coefficients of approximately 19 V\(^{-1}\), comparable to the 15.9 V\(^{-1}\) curvature coefficient reported for commercial Ge diodes from M-pulse Microwave [99]. The junction resistance of 13.5 kΩ·µm\(^2\) is also comparable to the reported 18 kΩ·µm\(^2\) from M-pulse Microwave. The I-V characteristic and curvature coefficient for a 3 x 3 µm\(^2\) Ge tunnel device fabricated using a 440 °C, 5 s anneal are shown in Fig. 4.13 (b).

![Figure 4.13 I-V characteristics of Ge tunnel diodes vs. junction area, (a) using a 440°C, 5 s anneal, and (b) I-V characteristic and curvature coefficient.](image)

From a 700 °C, spike anneal process, four Ge TDs in the same die, about 25 µm apart, are shown in Fig. 4.14. NDR regions are observed in all tested devices. Peak current density of 2.7 µA/µm\(^2\) was achieved. Compared with diffused Ge TDs, Fig. 4.3, a 500x increase in the peak current density is achieved. The peak tunneling currents scale with junction areas, Fig. 4.14(b). The PVR shows a decreasing trend with size. It is
possible some current component, such as leakage current, which does not scale with junction area, is present in these junctions, and reduces the PVR for the smaller junction.

The peak current density of the on-wafer liquid-phase regrown Ge tunnel diodes are still three orders of magnitude lower than the 1 mA/µm² target. To further improve the tunneling current density, P SOD was directly spun onto a RCA-cleaned Ge wafer and P was diffused at 800 °C for 300 s in a forming gas ambient using RTP. Compared with proximity RTP process, a smaller electron Hall mobility is measured (Table 4.1 and 4.2), which indicates a higher P carrier concentration than that achieved in proximity RTP. The Ge wafer was then cleaned in buffered HF for 5 minutes and 100 nm Al film was deposited. The tunnel junctions were formed by on-wafer liquid-phase regrowth at 600 or

Figure 4.14 $I$-$V$ characteristics of Ge tunnel diodes vs. junction area, (a) using a 700°C, spike anneal, and (b) peak current and PVR vs. device area.
700 °C with annealing time of 0 (spike), 1 and 4 s at a ramp rate of 100 °C/s.

Germanium TDs with peak current densities up to 0.16 mA/µm² were demonstrated. The $I-V$ characteristics of the highest measured peak current density devices in each experiment are shown in Fig. 4.15. The characteristic NDR of tunnel junction is clearly observed for all anneal conditions and reverse Zener tunneling characteristic is also apparent. The series resistance (less than 10 Ω) is responsible for the curvature at high currents; this series resistance is reasonable given the sheet resistance of the phosphorus diffusion (9 Ω/□) and the device geometry.

![Graph showing $I-V$ characteristics for different anneal conditions](image)

Figure 4.15 Dependence of Ge TD current-voltage characteristics on rapid-melt-growth anneal condition. P was diffused at 800 °C for 300 s with direct contact. Al of 100 nm was deposited. The Al-doped regrowth layers were regrown by annealing at 600 or 700 °C with anneal time of spike, 1 s or 4 s. The 700 °C, spike annealed sample was damaged during processing and therefore not measured.
The peak forward current density is plotted as a function of anneal condition in Fig. 4.16. The peak current density decreases monotonically with temperature and time. This change is consistent both with the expected decrease in the junction doping with temperature and the expected broadening of the dopant slope with temperature due to interdiffusion of P. For example, the P diffusion depth, calculated from $\sqrt{Dt}$, is about 1 nm with an annealing cycle at 600 °C for 1 s, and it increases to about 5.6 nm and 11.2 nm with annealing cycle at 700 °C for 1 s and 4 s, respectively. Both of these changes decrease the tunneling probability and tunnel current density which is what is observed. The reverse Zener tunneling current density corresponding to a reverse voltage of minus the forward peak voltage is also plotted in Fig. 4.16. At this voltage, the reverse current density should exceed the forward current density because of the greater overlap of states available for tunneling in the reverse direction over the forward direction, and this is what is observed.

For the highest current density diode in this set, the 600 °C spike anneal, a current density of 0.16 mA/µm² is obtained at a PVR of 1.1. Table 4.3 summarizes the highest peak current densities and PVRs of these devices. Compared to devices in Fig. 4.13, a 50x increase in the peak current density is achieved. The possible reason is that the P concentration has been improved. The PVR in the rapid-melt-growth junctions for these conditions are notably lower than has been achieved previously on alloy Ge diodes [90]. The low PVR is an indication of a high density of defect-states at the junction leading to a high excess current in the valley region of the $I-V$ characteristics. The junction
roughness likely also plays a role in the high valley current by introducing more edge leakage current.

Figure 4.16 Dependence of current density in the forward, $I_F$, and reverse current, $I_R$, directions as a function of anneal condition for the uncapped Ge diodes of Fig. 4.16. In the reverse direction, the current magnitude is plotted at a voltage corresponding to minus the forward peak voltage, $V_P$.

TABLE 4.4

SUMMARIES OF THE BEST PEAK CURRENT DENSITIES AND PVR OF Ge TDs FABRICATED BY THE ON-WAVER LIQUID-PHASE REGROWTH APPROACH UNDER VARIOUS ANNEALING CONDITIONS.

<table>
<thead>
<tr>
<th>Al anneal</th>
<th>$J_P$ (mA/µm²)</th>
<th>PVR</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 °C spike</td>
<td>0.16</td>
<td>1.1</td>
</tr>
<tr>
<td>600 °C 1 s</td>
<td>0.1</td>
<td>1.4</td>
</tr>
<tr>
<td>600 °C 4 s</td>
<td>0.08</td>
<td>1.3</td>
</tr>
<tr>
<td>700 °C 1 s</td>
<td>0.03</td>
<td>1.2</td>
</tr>
<tr>
<td>700 °C 4 s</td>
<td>0.004</td>
<td>1.4</td>
</tr>
</tbody>
</table>
4.4 Demonstration of Ge TDs exceeding 1 mA/µm² current density

During characterization of on-wafer liquid-phase regrown Ge TDs, it was observed that the Al/Ge mesa surfaces were roughened following anneal. The surface morphologies of the Al/Ge mesa structures before the regrowth annealing are shown in Fig. 4.17(a). The top view of one die of various area devices and two 5 x 5 µm² devices are shown. The bright color regions are Al while the darker part is Ge. The corresponding surface morphologies after on-wafer liquid-phase regrowth are shown in Fig. 4.17(b). A height variation from 10 to 600 nm was measured by scanning an alpha step profiler across a 40 x 40 µm² square structure.

Figure 4.17 Top view optical micrographs of Al/Ge mesa structures without a nitride cap, before (a) and after (b) a 440 °C, 5 s regrowth anneal. The Al metal is 100 nm thick.
The fact that the dark part surrounding the $5 \times 5 \, \mu m^2$ feature is Ge with no Al deposit suggests that the dark regions in the metallized areas are the Ge phase of the eutectic mixture, while the white areas are the Al phase. The large areas of Al and of Ge in both the $5 \times 5 \, \mu m^2$ feature and the larger metallized areas, suggesting that the melt formed droplets instead of wetting the Ge. This might be the reason of the rough junction surface after annealing.

A low-pressure chemical vapor deposited silicon oxide layer has been used as a microcrucible in the growth of crystalline Ge layer on silicon [98]. The silicon oxide layer works as a microcrucible that holds the liquid Ge and prevents it from flowing during the Ge melt back and regrowth process. With the similarity between the Ge regrowth process and the Al-Ge on-wafer liquid-phase regrowth, the microcrucible idea is explored on the Ge tunnel diodes formation.

A cap layer is thereby added into the on-wafer liquid-phase regrowth process. The Ge wafer is heavily-$n$ doped using P SOD and RTP. After 100 nm Al is deposited and patterned, Ge mesas of approximately 30 nm were etched by 30 s wet etch in 30% $H_2O_2$. The mesa height is about 130 nm. After the Ge mesa is etched, a dielectric layer, either silicon nitride or silicon oxide, was deposited by plasma-enhanced chemical vapor deposition (PECVD) and covered the Al/Ge mesa structure. Tunnel junctions were then formed using the liquid-phase regrowth in this encapsulated geometry by RTP in a forming gas ambient. A scale drawing of the device cross section is shown in Fig. 4.18. A silicon nitride layer of 50 nm is indicated.
Silicon nitride and oxide layers of various thicknesses, all deposited by PECVD at a substrate temperature of 250 ºC, were tested as the microcrucible. Four Ge samples were prepared as described in last paragraph and deposited with different thickness of silicon nitride or oxide. It was not clear what cap layer thickness should be chosen. A thickness of 50 nm, which is about half of the Al/Ge mesa height and a thickness of 250 nm, which is about twice the Al/Ge mesa height, were deposited for investigation. The Al/Ge mesas were then annealed at 440 ºC for 5 s in RTP in a forming gas ambient. The surfaces are then observed under optical microscope.

Both 50 and 250 nm silicon oxide cap layers yielded surface morphologies similar to those with no caps. The 5 x 5 µm² features after a 440 ºC, 5 s anneal, with 50 nm or 250 nm silicon oxide cap, are show in Fig. 4.19. It has been reported that reaction between Al and PECVD SiO₂ occurred at 300 ºC in an Al/SiO₂/Si structure, through a
reaction of $4\text{Al} + 3\text{SiO}_2 \rightarrow 2\text{Al}_2\text{O}_3 + 3\text{Si}$ [111]. A possible reason for the rough surface could be that Al reacts with Si in the silicon oxide layer at 440 °C and the metal forms droplets instead of wetting the Ge. The tunnel diodes fabricated in these processes have peak current densities of 1 µA/µm², similar to Fig. 4.13(a).

![Figure 4.19 Top view optical micrographs of Al/Ge mesa structures after a 440 °C, 5 s regrowth anneal in a forming gas ambient with (a) a 50 nm oxide, and (b) 250 nm oxide.](image)

A 50 nm silicon nitride cap layer improves the morphology of the Al/Ge mesa structures after the regrowth process, Fig. 4.20. The top view of one die of various area devices, Fig. 4.20(a), one 10 x 10 µm² and one 5 x 5 µm² features, Fig. 4.20(b), with 50 nm silicon nitride cap after a 440 °C, 5 s RTP anneal, are shown. Compared to the uncapped regrowth in Fig. 4.17, a nearly continuous Al or Al/Ge covered mesa area was obtained. It has been reported that reaction between Al and low pressure chemical-vapor deposited (LPCVD) silicon nitride occurs at a higher temperature, 550 °C, than that with oxide [112]. Because of the higher reaction temperature between Al and silicon nitride, it
is possible that less reaction occurs between Al and nitride. The nitride cap holds the Al in place, prevents the liquid from islanding, thereby improving the surface morphology.

Figure 4.20 Top view optical micrographs of Al/Ge mesa structures after a 440 °C, 5 s regrowth anneal in a forming gas ambient with (a) a 50 nm nitride, (b) 50 nm nitride, 10 x 10 and 5 x 5 µm² devices, and (c) 250 nm nitride. The Al metal is 100 nm thick.

In Fig. 4.20(c), 5 x 5 µm² and 3 x 3 µm² features with a 250 nm nitride cap layer after the 440 °C, 5 s RTP anneal is shown. Cracks are observed after anneal and the surface was roughened similar to the no cap situation. It appears that the cracked nitride cap does not hold the Al melt in place. Droplets form and Al/Ge surface is roughened.

Addition of a 50 nm silicon nitride capping layer clearly resulted in much more uniform coverage of the substrate, and the dark lines in the metallization have a typical eutectic morphology. As the size of the capped features was reduced (as small as 3 x 3 µm²), the uniformity of the Al-Ge film was observed to improve and the peak current density increased (as will be shown). A 50 nm nitride layer is thereby employed for later Ge TD fabrication.
With the anneal hold time set at 1 s, a sequence of rapid melt growths was performed on another set of wafers fabricated with the same phosphorus-diffusion condition, but this time with a 50 nm silicon nitride cap. The dependence of forward peak current and reverse current vs. anneal temperature from 440 to 725 °C for these are shown in Fig. 4.21.

![Figure 4.21 Germanium TD peak current density vs. annealing temperatures. Phosphorus was diffused at 800 °C for 300 s with direct contact. Al of 100 nm was deposited. Silicon nitride layer of 50 nm was deposited by PECVD.](image)

In this case a low constant reverse voltage is plotted which similarly tracks the forward-current tunneling-dependence on temperature. A maximum in the tunneling current is found at a temperature of 600 °C. This current-density dependence on
temperature is consistent with an improving electrical activation as temperature increases from 440 to 600 °C. The decreasing tunnel current at temperatures over 600 °C likely results from a degraded dopant slope, due to phosphorus diffusion into the regrown layer. Phosphorus has a diffusion length of approximately 6 nm in 1 s at 700 °C and 1 nm in 1 s at 600 °C in agreement with a degraded dopant slope and lower current density for anneals at 700 °C and above. The Al diffusion coefficient is an order of magnitude lower than phosphorus at these temperatures. Degradation in the peak current would also result from a decrease in junction doping which, as shown in Fig. 4.9, decreases as peak anneal temperature is increased. Germanium tunnel diodes with a peak current density of 0.52 mA/µm² and PVR of 1.5 are achieved under these conditions. More than three times the current density and slightly better PVR are achieved using the silicon nitride cap.

Germanium TDs with current densities up to 1.2 mA/µm² were successfully obtained using a ramp rate of 150 °C/s, an anneal temperature of 600 °C, and a hold time of 2 s. Three devices with nominal areas of 10 x 10, 5 x 5 and 3 x 3 µm² in the same die, about 30 µm apart, were measured. The device current voltage characteristics are shown in Fig. 4.22, showing an area dependence of the current density. The areas of these devices were all measured under a SEM. As the device size was reduced the current density increased. The highest peak current density, 1.2 mA/µm², is achieved at the smallest junction area with reverse current density of 3.8 mA/µm² also achieved in this junction.
Figure 4.22 Current-voltage characteristics of Ge TDs fabricated by the on-wafer liquid-phase regrowth process with 25 nm mesa and 50 nm silicon nitride microcrucible layer. The junction formation regrowth was at 600 °C, 2 s and a ramp up rate of 150 °C/s.

The improvement in current density with reduced junction area could be the result of an improvement in junction morphology or an increase in leakage current. It was observed that small area device generally exhibited more uniform morphology. This could be due to the grain structures of the evaporated Al film. Fewer Al grains are expected on a small area Al/Ge mesa, than on a large area one. With fewer grain structures, the liquid-phase regrowth process could be expected to be more uniform and result in higher tunneling current.

The PVR of Ge devices in Fig. 4.22 is about 1.1, which is notably lower than prior art value of 16 [90]. The low PVR is an indication of a high density of defect-states.
at the junction leading to a high excess current in the valley region of the $I-V$ characteristics [38]. Due to the imperfection of crystal regrowth, crystal defects will form and introduce energy states in the bandgap and enhance the excess tunneling current [38].

A transmission electron microscopy (TEM) on the junction would assist the understanding of those defects. Low temperature RTP (400 °C to 600 °C) in a forming gas ambient has been used to anneal out defects in MBE-grown SiGe TDs thereby improve PVR [7]. Forming gas anneal might be able to improve the PVR in on-wafer liquid-phase regrown Ge TDs. Buckingham et al. [113] annealed alloy Ge TDs in a pure hydrogen ambient at 310 and 340 °C and reported that both peak current and valley current decreased upon annealing and the PVR was approximately constant after annealing.

To the author’s knowledge, this is the highest or comparable to the highest current density Ge TDs reported to date. The goal of achieving 1 mA/µm$^2$ tunneling current density devices is successfully achieved.

4.5 Limitations of approach

The low PVR of Ge TDs using on-wafer liquid-phase regrowth will lead to longer switch time (Eq. 2.8) and limit the device applications in high speed circuit. The regrowth process could be improved to achieve better junction growth by using slower cooling process.
5.1 Summary of accomplishments

Silicon TDs were demonstrated in both vertical and lateral orientations using SODs and RTP. Silicon TDs were formed in the substrate plane through an oxide window process, with peak current densities of approximately 1 μA/μm² and peak-to-valley ratio of approximately 1.3. A self-aligned lateral silicon TD fabrication process, which forms the junction perpendicular to the substrate plane, has also been successfully developed and yielded backward Si TDs with peak current densities of 30 nA/μm². These studies constitute the demonstration of Si tunnel devices using spin-on diffusants and rapid thermal processing approach.

Germanium TDs were demonstrated both using a diffusion-based approach and an on-wafer liquid-phase regrowth approach. The diffusion-based approach utilized SODs and RTP and yielded devices with peak tunneling current densities up to 0.63 nA/μm² and PVR of 1.1 for the first time. An on-wafer liquid-phase regrowth approach with a silicon nitride microcrucible was developed. Germanium TDs with current densities up to 1.2
mA/μm² were demonstrated. A primary goal of this project, demonstration of a 1 mA/μm² tunnel junction, was fulfilled.

5.2 Achievements relative to prior art

The SODs and RTP approach for silicon TD fabrication achieved TDs with peak tunneling current density about 10x larger than that achieved by Koga and Toriumi [18] by II. The diffusion-based approach for Si tunnel devices was shown to produce backward diodes with zero bias curvature coefficient of 34 V⁻¹, which is comparable to those fabricated by the MBE approach [35] for detector applications.

Silicon TDs were also demonstrated in a lateral geometry using spin-on diffusants and rapid thermal processing. The device tunneling current density is about 3 A/cm² and zero bias curvature coefficient of 22. To the author’s knowledge, this is the first silicon tunnel junction formed perpendicular to the wafer surface using diffusion-based approach.

Germanium TDs with zero bias curvature coefficient of 54 V⁻¹ were demonstrated using a diffusion-based approach. This value is lower than the highest reported zero bias curvature coefficient of 70 V⁻¹ [91], but is superior to what have been achieved in Si tunnel devices by MBE [35] and in this work, where the best zero bias curvature coefficient is 31 V⁻¹ [35] and 34 V⁻¹, respectively.

Germanium TDs with peak current densities up to 1.2 mA/μm² were achieved using an on-wafer liquid-phase regrowth approach with silicon nitride microcrucible. To
the author’s knowledge, this is the highest or comparable to the highest current density ever achieved in Ge tunnel devices. The PVR however is significantly less than has been previously achieved with a PVR of 1.5 compared to 16 [90].

5.3 Recommendation for further study

Optimization of the heating and cooling curves, and reduction in the device area could be experimented to improve the junction uniformity. As discussed in Fig. 4.22, tunneling current density increases with the decrease of junction area. Anneal procedure could be experimented to improve the junction growth. A TEM investigation of the junction will help understand the defects formed during liquid-phase regrowth. All of the above techniques will help improve the peak tunneling current density.

An attribute of the Al on-wafer liquid-phase regrowth process is a large amount of Ge dissolved and consumed by Al, i.e. 100 nm Al will dissolve 113 nm Ge, in which 52 nm will form eutectic with Al and 61 nm will regrow back, at 600 °C, Fig. 4.9. Assume the P concentration is at $2 \times 10^{20}$ cm$^{-3}$ uniformly across this 113 nm thickness. After Al and Ge forms the melt of 213 nm (assume the thickness of the melt is equal to the addition of melted Al and Ge), P distributes uniformly across this melt and the P concentration is about $1 \times 10^{20}$ cm$^{-3}$. During epitaxial regrowth, these P dopants will offset one fourth of the Al acceptor doping, which is $4 \times 10^{20}$ cm$^{-3}$ (solid solubility), at this temperature. If the dissolved Ge could be reduced to 10 nm, the P concentration in the melt would be about $2 \times 10^{19}$ cm$^{-3}$, only one twentieth of the Al concentration. In addition,
the melt dissolves the most heavily doped $n^+$ Ge, since the donor diffusion profile has maximum concentration at the surface. This will reduce the $n$-type doping level at the junction and reduce the tunneling current.

One way to reduce the Ge dissolved depth and dopant compensation is using Ga instead of Al. The Ga-Ge phase diagram is shown in Fig. 5.1. With a Ga film of 100 nm and anneal temperature of 600 °C, approximately 33 nm Ge will be dissolved, compared to 113 nm of Ge under the same metal thickness and anneal condition.

![Figure 5.1 Germanium-gallium phase diagram from Okamoto [110].](image)

In the point of view of processing, gallium can achieve solid solubility up to $5 \times 10^{20}$ cm$^{-3}$, which is higher than $4 \times 10^{20}$ cm$^{-3}$ of Al at 600 - 700 °C [69]. The diffusion
coefficient of Ga and Al are comparable, i.e. $3.49 \times 10^{-17}$ and $2.75 \times 10^{-15}$ cm$^2$/s of Al at 600 and 700 °C, respectively, and $1.13 \times 10^{-17}$ and $1.03 \times 10^{-15}$ cm$^2$/s of Ga [102], at corresponding temperatures. Burrus reported Ge TDs with PVR of 3 using electrical forming and Ga metal [15]. It is possible that Ga-Ge has a better regrowth process than Al-Ge, which reduces the junction defects and improves PVR.
APPENDIX 1

VERTICAL SILICON TUNNEL DIODE FABRICATION TRAVELLER

Preparation
- Scribe sample numbers onto the wafer backside avoiding scratch lines parallel with the wafer flats (cleave planes)
- Clean and label wafer trays
- Remove spin-on diffusants (SOD) from refrigerator 24 hours before use

Field Oxidation
- Clean quartz spacers along with wafers in Teflon basket
- RCA1 (DI: NH₄OH: H₂O₂ = 5-10:1:1) bath, 70 °C, 15 min
- Rinse 2 min in DI
- RCA2 (DI: HCl: H₂O₂ = 5-10:1:1) bath, 70 °C, 15 min
- Rinse 2 min in DI
- Etch off surface oxide in buffered HF [1HF: 10H₂O] 25 s
- Rinse in DI water and blow dry
- Inspect
- Oxidation: Purge the tube using dry oxygen, set for 90 min., ramp to 1100 °C, turn on vapor control, turn on oxygen tank
- Load wafers
- Dry oxidation 1100°C, 15 min  actual temp. _____
- Wet oxidation 1100°C, 90 min  actual temp. _____
- Dry oxidation 1100°C, 10 min.  actual temp. _____
- Expect 0.6 µm
- Measure oxide thickness using Gaertner Ellipsometer

Process using ICFAB mask
P-Well Photolithography and Oxide Etch
- Spin on HMDS using program B  1000 RPM 10 s, 4000 RPM 30 s
- Spin on Shipley 1813 photoresist using program B, bake @ 90 C for 60s
• Expose using Mask ____P-Well__, use program IC1 on the stepped, exposed time 0.5s
• Measured intensity ______________ mJ/cm²/s exposed time _________s
• Develop in AZ 327 MIF developer, 80 s
• Rinse in DI
• Hot plate post bake 120 C, 10 min( this step is to harden the PR, might be omitted)
• Take one wafer as a test wafer. Etch approximately 800 s in buffered HF (etch rate ~ 0.8-1.1 nm/s.
• Record etching using alpha step
• Record etching time ____________. Record etching rate ______.
• Record oxide thickness using Alpha step
• Complete remaining wafers using etch time determined on the first wafer
• Soak in Acetone 90 s , IPA 30 s, Rinse with DI
• Repeat etching as necessary to make sure oxide window is open
• Inspect under a microscope and step profile. When the step profile thickness decreases then the oxide window is open.
• Record etching time ____________. Record etching rate ______.
• Record oxide thickness using Alpha step
• Complete remaining wafers using etch time determined on the first wafer
• Soak in Acetone 90 s , IPA 30 s, Rinse with DI
• Repeat etching as necessary to make sure oxide window is open
• Remove SOD from refrigerator
• Allow SOD to warm to room temperature overnight
Wafer Clean-up (if process is delayed overnight)
• Rinse 2 min in DI
• Etch off surface oxide in buffered HF 25 s

Spin on Emulsitone Phosphosilicate SOD
• Drip on SOD phosphosilicate on source wafers
• Spin 1000 RPM 10 s, 4000 rpm 30 s
• NO BAKE NEEDED

Rapid Thermal Diffusion of Phosphorus
• Purge the chamber using PURGE program
• Verify the RTA program by running on a test wafer, program JZ900DIF.rpd

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Temp (C)</th>
<th>gas1</th>
<th>N2</th>
<th>gas3</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>30</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>30</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>30</td>
<td>910</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hold</td>
<td>1</td>
<td>910</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
• Load substrate, place three spacers at 120 degree increments at the edge of the wafer. Run anneal program

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source Wafer</th>
<th>Measured Temperature</th>
<th>Time (s)</th>
<th>Comments</th>
</tr>
</thead>
</table>

Drive for 90s using program JZ900DRI.RPD. For drive-in process, cover the surface of device wafer with a clean wafer, separated by spacers, so that SOD is confined

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Temp (C)</th>
<th>Gas1</th>
<th>Gas3</th>
<th>Gas2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>30</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>30</td>
<td>10</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>30</td>
<td>910</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Hold</td>
<td>90</td>
<td>910</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>

• Inspect under microscope, record observations
• Consider that an etch to remove the remnants of the SOD silicate film from the proximity diffused substrate may be needed.
  Measure I-V before etch
  Measure I-V after etch
• Etch in buffered HF to remove the residue of SOD. Plan to do 2 minutes. Record etching time
• Record oxide thickness using Alpha step
• Remove SOD from refrigerator
• Allow SOD to warm to room temperature overnight

Wafer Clean-up (if process is delayed overnight)
• Rinse 2 min in DI
• Etch off surface oxide in buffered HF [1HF:10H2O] 25 s

Spin-On Emulsitone Borofilm 100
• Drip SOD on source wafers
• Spin program: 1000 rpm 10 s, 4000 rpm 30 s
• Inspect – draw maps of uniformity
• Pre-bake on hot plate: @200 C for 20 min

Rapid Thermal Diffusion of Boron
• Purge the chamber using PURGE program
• Verify the RTA program by running on a test wafer, program JZ900DIF.rpd

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Temp (C)</th>
<th>Gas1</th>
<th>Gas2</th>
<th>Gas3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>30</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>30</td>
<td>10</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>30</td>
<td>910</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Hold</td>
<td>1</td>
<td>910</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>
• Load substrate, place three spacers at 120 degree increments at the edge of the wafer
  Run anneal program

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source Wafer</th>
<th>Measured Temperature</th>
<th>Time (s)</th>
<th>Comments</th>
</tr>
</thead>
</table>

• Observe under microscope, record observation
• Before etch measure the thickness using alpha step
  Etch in buffered HF to remove the residue of SOD, plan for 10 minutes
• Record oxide thickness using Alpha step

**Metal Lithography**

• Bake @ 200 C for 1hr+ to remove moisture
• Deposit 200 nm Al on the surface
• Spin on HMDS using program B
  1000 RPM 10 s, 4000 RPM 30 s
• Spin on Shipley 1813 photoresist using program B, bake @ 90 C for 60 s
• Expose using Mask ___Metal___, use program IC1 in stepped, expose 0.5 s
• Measured intensity ________________mJ/cm²/s exposed time ________s
• Develop in AZ 327 MIF developer, 80 s
• Rinse in DI
• Etch in Al etchant for 3 minutes, should be fine
• Soak in Acetone 90 s, IPA 30 s, Rinse with DI
APPENDIX 2

LATERAL SILICON TUNNEL DIODE FABRICATION TRAVELLER

Field Oxidation Preparation

- Scribe sample numbers onto the wafer backside avoiding scratch lines parallel with the wafer flats (cleave planes)
- Clean and label wafer trays
- To save time, prepare the oxidation furnace together with RCA cleaning
- Oxidation furnace preparation: Purge the tube using High Nitrogen, set for 90 min; ramp to 1000 °C using PC control; the panel will show the furnace ramp from 400 °C to 1000 °C; turn on bubbler control panel, set to 95 °C;
- Set RCA 1 and RCA 2 bath to 70 °C
- Clean quartz spacers along with wafers in Teflon basket
- RCA1 (DI: NH₄OH: H₂O₂=40-50:1:1) bath, 70 °C, 15 min
- Rinse 2 min in DI
- RCA2 (DI:HCL: H₂O₂=40-50:1:1) bath, 70 °C, 15 min
- Rinse 2 min in DI
- Etch off surface oxide in HF [1HF: 10H₂O] 25 s
- Rinse in DI water and blow dry
- Inspect
- After the furnace reaches 1000 °C, load wafers. Leave the wafer at the furnace mouth for 2 minutes, load the wafers slowly for 5 minutes into the furnace until the outside part of quartz rod is ~70 cm (from the furnace surface). Turn on the oxygen tank. Begin oxidation.
- Dry oxidation 1000 °C, 30 min actual temp. _____
- Wet oxidation 1000 °C, 75 min actual temp. _____
- Dry oxidation 1000 °C, 30 min. actual temp. _____
- After the last dry oxidation, leave the wafers in the tube for at least 20 minutes. This is the time for high nitrogen to push out the oxygen before. Unload the wafers in high nitrogen flowing situation. Unload slowly for 5 minutes.
- Expecting 3757A oxide. Measure oxide thickness using VASE system
<table>
<thead>
<tr>
<th>Wafer No.</th>
<th>Point 1 Thickness (Å)/MSE</th>
<th>Point 2 Thickness (Å)/MSE</th>
<th>Point 3 Thickness (Å)/MSE</th>
<th>Comments</th>
</tr>
</thead>
</table>

### Process using TD mask set

**Emitter layer Photolithography and Oxide Etch**

- Spin on HMDS using program J: 1000 RPM 10 s, 4000 RPM 30 s
- Spin on Shipley 1813 photoresist using program J, bake @ 90 °C for 60 s
- Expose using Mask __Emitter__, use program TD1 on the stepper, exposed time 0.6s. changed from 0.54 s to 0.6 s by Mike
- Develop in AZ 327 MIF developer, ____________ (70 s or 80 s)
- Rinse in DI twice and blow dry
- Dry etch first to get straight sidewall.
  - O2 Clean for 20 minutes, 300 W, 80 mT, 30 sccm to clean the RIT chamber
  - Recipe: CF4/O2 26 sccm/4 sccm 40 mT 200 W 6.5 minutes
  - Etch rate 50 nm/min for quarter piece wafer, the etch rate for whole 4 inch is expected to be slower
  - Clean the chamber using program “O2Clean” for 20 minutes. Season the chamber use the dry etch recipe for 10 minutes
  - Process real wafers using Manual mode, which gives more steady plasma

<table>
<thead>
<tr>
<th>Wafer No</th>
<th>DC bias (V)</th>
<th>Time</th>
<th>Comments</th>
</tr>
</thead>
</table>

At this step, program “Flimeasure” can be used to measure the oxide left inside the window. Filmmeasure has a square on the TV screen, ~ 50 x 50um, can measure the film thickness inside the feature.

<table>
<thead>
<tr>
<th>Wafer No</th>
<th>Oxide after RIE(Å)</th>
<th>Oxide after RIE</th>
<th>Oxide after RIE</th>
</tr>
</thead>
</table>

Bake at 120 °C for 2 minutes.
- Etch in buffered HF for ____________ (1 min).
- Record etching using alpha step

<table>
<thead>
<tr>
<th>Wafer No</th>
<th>Etch time (s)</th>
<th>Oxide mid (Angstrom)</th>
<th>Oxide left</th>
<th>Oxide right</th>
</tr>
</thead>
</table>

Complete remaining wafers using etch time determined on the first wafer
- Soak in Acetone 90 s, IPA 30 s to remove the photoresist. Rinse with DI and blow dry

120
• Repeat etching as necessary to make sure oxide window is open
• Remove SOD from refrigerator
• Allow SOD to warm to room temperature. Normally several hours. Overnight is not necessary.

**Wafer Clean-up (if process is delayed overnight)**
• Etch off surface oxide in buffered HF \([1\text{HF}:10\text{H}_2\text{O}]\) 25 s
• Rinse 2 min in DI and blow dry

**Spin on Emulsitone Phosphosilicate SOD**
• RCA clean p- dummy wafers. Note: after RCA, don’t use HF dip. Leave the native oxide will make the spreading of SOD better
• Drip on SOD phosphosilicate on source wafers (with native oxide on)
• Spin 1000 RPM 10 s, 4000 rpm 30 s. (First experiment showed thickness of SOD is around 3000Å)
• Bake @ 200 C for 20 minutes.

**Rapid Thermal Diffusion of Phosphorus**
• Always purge the chamber using PURGE program before any anneal. This is to purge anything left in the chamber from previous anneals.

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>Temp (C)</th>
<th>Gas 1 N(_2) slpm</th>
<th>Gas 2 O(_2) slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>60</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

• Change the normally used thermocouple to ceramic covered thermocouple
• Verify the RTA program by running on a test wafer, program JZ1000CD.RPD

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>Temp (C)</th>
<th>Gas 1 N(_2) slpm</th>
<th>Gas 2 O(_2) slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>90</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>20</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>13</td>
<td>650</td>
<td>2</td>
</tr>
<tr>
<td>Hold</td>
<td>30</td>
<td>650</td>
<td>2</td>
</tr>
<tr>
<td>Ramp</td>
<td>3</td>
<td>1100</td>
<td>2</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

• Run test run before real process. JZ1000CD.030, peak temperature 1028C
• Load substrate, place three spacers at 120 degree increments at the edge of the wafer

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source</th>
<th>Measured Temp</th>
<th>Time (s)</th>
<th>Comment</th>
</tr>
</thead>
</table>

• Inspect under microscope, record observations. Take pictures.
  No black dots can be observed.
• Considering an etch to remove the remnants of the SOD silicate film from the proximity diffused substrate may be needed.
• Etch in buffered HF to remove the residue of SOD. Plan to do 2 minutes. Record etching time
• Record oxide thickness using Alpha step

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Etch time (s)</th>
<th>Oxide before</th>
<th>Oxide after</th>
<th>Etch rate</th>
</tr>
</thead>
</table>

The structure E40 is supposed to use after Al contact. However it might be interrupted at this step to measure the phosphorus doping level by putting down TLM pattern. The other method would be a separate experiment using Hall measurement.

**Tungsten sputtering and CMP**

• Sputtering W supposed thickness ~300nm. (20 mT, 74%×200=148 sccm)
• (7 mins supposed to give 300nm W if the wafer is put under the target without rotating the table 1150W applied forward power)

Since the PE2400 is not working well now, the deposition can only be done at forward/reflected power 500W/50W. Increasing forward power will increase the reflected power above 50W, which is not allowed. The tungsten is sputtered under forward/reflected 500W/50W, 20 mT, 74%×200=148sccm, table spacing 2.5 inch, target bias 500 V, with table rotating. Sputtered for 40 minutes.
• A pilot piece wafer is sputtered at the same time. It’s patterned using ISO mask, etch in room temperature 30% H2O2 for 2 mins, already etch off all the W. measured using α-step, gives ~1512 Å.
• CMP: compose the slurry W2000:D1:30% H2O2=900ml:900ml:100ml. hand shake for 5 to 10 minutes. Regenerate/recondition the pad for 10 minutes with table speed 50 rpm. Polish: slurry 75 ml/min, table speed 50 rpm, carrier head speed 50 rpm. The dent is measured to be 1.17mm; black insert is measured to be 0.57mm, the blue sheath is 0.09mm. the wafer is ~0.6mm.

**Isolation layer Photolithography and Oxide Etch**

• Spin on HMDS using program J: 1000 RPM 10 s, 4000 RPM 30 s
• Spin on Shipley 1813 photoresist using program J, bake @ 90 C for 60s
• Expose using Mask ___ISO__, use program TD1 on the stepper, exposed time 0.54s. The key offset in TD1 is x= -1.55mm y=1.55mm
• Measured intensity _______________ mJ/cm²/s exposed time _________s

122
• Develop in AZ 327 MIF developer, (70 s)
• Rinse in DI for 2 minutes and blow dry
• Take one wafer as a test wafer. Etch approximately 360 s in buffered HF. Rinse with DI twice and blow dry. (etch rate ~ 1.0-1.2 nm/s. This needs further work to verify.)
• Record etching using alpha step

<table>
<thead>
<tr>
<th>Wafer No</th>
<th>Etch time (s)</th>
<th>Oxide mid (Angstrom)</th>
<th>Oxide left</th>
<th>Oxide right</th>
</tr>
</thead>
</table>

• Complete remaining wafers using etch time determined on the first wafer
• Soak in acetone 90 s, IPA 30 s to remove the photoresist. Rinse with DI and blow dry. Repeat etching as necessary to make sure oxide window is open
• Remove SOD from refrigerator
• Allow SOD to warm to room temperature. Normally several hours. Overnight is not necessary.

**Wafer Clean-up (if process is delayed overnight)**
• Etch off surface oxide in buffered HF [1HF:10H2O] 25 s
• Rinse 2 min in DI and blow dry

**Spin-On Emulsitone Borofilm 100**
• RCA clean p- dummy wafers as source wafers. Note: don’t use HF dip afterwards. Removing the native oxide will make the spreading of SOD difficult.
• Drip SOD on source wafers
• Spin program: 1000 rpm 10 s, 4000 rpm 30 s
• Inspect – draw maps of uniformity
• Pre-bake on hot plate: @200 C for 20 min. Rapid Thermal Diffusion of Boron
• Purge the chamber using PURGE program before any anneal. This is to purge anything left in the chamber from previous anneals.

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>Temp (°C)</th>
<th>Gas 1 N2 slpm</th>
<th>Gas 2 O2 slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>60</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

• Change the normally used thermocouple to ceramic covered thermocouple
• Verify the RTA program by running on a test wafer, program JZ900CD.rpd

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>Temp (°C)</th>
<th>Gas 1 N2 slpm</th>
<th>Gas 2 O2 slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>90</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Idle</td>
<td>20</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>Ramp</td>
<td>12</td>
<td>600</td>
<td>2</td>
</tr>
<tr>
<td>Hold</td>
<td>30</td>
<td>600</td>
<td>2</td>
</tr>
<tr>
<td>Ramp</td>
<td>2</td>
<td>900</td>
<td>2</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

123
Load substrate, place three spacers at 120 degree increments at the edge of the wafer
Run anneal program

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source</th>
<th>Measured Temp</th>
<th>Time (s)</th>
<th>Comment</th>
</tr>
</thead>
</table>

• Observe under microscope, record observation
• Measure the diode structures npi diodes at this step. If there were tunnel diodes, we will be able to measure two back to back tunnel diodes.
• (Put wafers in RCA cleaning bath. This will help removal of SOD residual. RCA1 & 2 for 15 minutes each, might be good to remove the W as well)
• Etch in buffered HF to remove the residue of SOD, plan for 2 minutes
• Record oxide thickness using Alpha step

Via layer lithography
• Deposit silicon oxide using PECVD, 2500 Å
  900 mT, 25 W, 250 C. ~500 Å/min deposit for 5 mins. The etch rate of PECVD oxide in BHF is ~33.8 A/s. Recipe SiO2dep.prc
• Spin on HMDS using program J: 1000 RPM 10s, 4000 RPM 30 s
• Spin on Shipley 1813 photoresist using program J, bake @ 90 C for 60 s
• Expose using Mask __ VIA ___, use program TD1 in stepper, expose 0.54 s
• Measured intensity ___________ mJ/cm²/s exposed time _________ s
• Develop in AZ 327 MIF developer, _____________ (70 s).
  Rinse in DI. Blow dry. Baked @120 C for 2 mins
• Etch out via in buffered HF, 90 s
  (Etch out via using RIE etching, using program __________
  power ____________ etchant_________ pressure_________)
• Soak in Acetone 90 s, IPA 30 s, Rinse with DI and blow dry.
• Observe the smallest feature 1x1 um.

Metal layer lithography
• Bake @ 200 C for 1hr+ to remove moisture
• Deposit 200 nm Al on the surface
• Spin on HMDS using program J: 1000 RPM 10s, 4000 RPM 30 s
• Spin on Shipley 1813 photoresist using program J, bake @ 90 C for 60 s
• Expose using Mask __ Bondpad ___, use program TD2 in stepper, expose 0.54 s
  Key offset in TD2 is x=-1.55mm y=-1.55mm
• How are exposures being calibrated on the stepper?  Measured intensity ___________ mJ/cm²/s exposed time _________ s
• Develop in AZ 327 MIF developer, _____________ (70 s). Rinse in DI and blow dry. Baked @ 120 C for 2 mins
• Etch in Al etchant for 4 minutes, should be fine.
• Soak in Acetone 90 s, IPA 30 s, Rinse with DI and blow dry.
Ready for Measurement

Measurement of testing structure

- Tungsten line. 46 squares, 138 squares and 230 squares, respectively.
  Measure the resistance of three tungsten lines.
  - 46 squares
  - 138 squares
  - 230 squares

- Via test structure.
  Measure the resistance of three via test structures. If all the vias were open, we should measure a short.
  - 30 1x1 um via line
  - 50 1x1 um via line
  - 150 1x1 um via line

- TLM test structure.
  E40 is tungsten on n+ silicon. C40 is p+ layer.
  Measure the resistance of E40 and C40

<table>
<thead>
<tr>
<th></th>
<th>2um</th>
<th>4um</th>
<th>8um</th>
<th>16um</th>
<th>32um</th>
</tr>
</thead>
<tbody>
<tr>
<td>E40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2um</th>
<th>4um</th>
<th>8um</th>
<th>16um</th>
<th>32um</th>
</tr>
</thead>
<tbody>
<tr>
<td>C40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Diode test structures E40A (2um alignment tolerance) and C40A (1um alignment tolerance)
GERMANIUM TUNNEL DIODE FABRICATION USING LIQUID-PHASE REGROWTH APPROACH TRAVELLER

Preparation
- Degrease the Ge wafer in acetone and methanol for 5 minutes each
- Etch off surface oxide in buffered 10:1 HF 25 s
- Rinse in DI water and blow dry
- Repeat BOE dip and DI rinse for 2 more times
- Inspect. To minimize the Ge oxide formation, do the HF dip just before RTP anneal
- Remove SOD from refrigerator
- Allow SOD to warm to room temperature. Normally several hours. Overnight is not necessary.

Wafer Clean-up (if process is delayed overnight)
- Etch off surface oxide in buffered HF [1HF:10H₂O] 25 s
- Rinse 2 min in DI and blow dry

Spin-On Phosphorosilicate Film
- Drip SOD directly on Ge wafers
- Spin program: 1000 rpm 10 s, 4000 rpm 30 s
- Inspect – draw maps of uniformity
- Pre-bake on hot plate: @100 C to 150 C for 20 min. Set the temperature to 80 C first, then raise the temperature to 150 C, this will minimize the film crack during bake

Rapid Thermal Diffusion of Phosphorus
- Always purge the chamber using PURGE program before any anneal. This is to purge anything left in the chamber from previous anneals.

<table>
<thead>
<tr>
<th>Gas 1 N₂ slpm</th>
<th>Gas 2 O₂ slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>Temp (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>60</td>
</tr>
</tbody>
</table>
Run O2 ambient 600 C for 5 minutes to burn any organic in the chamber, then run high temperature flush 850 C for 3 s before testing real process recipe.

- Verify the RTA program by running on a test wafer, program JZ800FOR.rpd

<table>
<thead>
<tr>
<th></th>
<th>Time (s)</th>
<th>Temp (°C)</th>
<th>Gas 1 forming gas slpm</th>
<th>Gas 2 O2 slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>90</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>26</td>
<td>800</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Hold</td>
<td>300</td>
<td>800</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>120</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

- Load substrate, place three spacers at 120 degree increments at the edge of the wafer

Run anneal program

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source</th>
<th>Measured Temp</th>
<th>Time (s)</th>
<th>Comment</th>
</tr>
</thead>
</table>

- Observe the surface of germanium
- Hall resistivity measurement if needed, to measure the phosphorus density.

Al source deposition and anneal

- Etch the Ge wafer in BOE for 5 minutes. Rinse in DI twice. Blow dry.
- Bake at 100 C for 2 minutes to further dry the surface. Put in FC-1800 #2 immediately
- Deposit Al of ~ 100 nm, rate ____________
- Pattern the Al: spin on HMDS/AZ5214 1000 rpm 10 s, 3000 rpm 30s using speedway spinner. Expose in Kaul Suss using jzdiode2 mask. Energy 120 mJ/cm². Measured power __________, expose time __________. Develop in AZ327 developer __40__ s. Observe
- Etch in Al etchant for 2-3 minutes, until can see the silver color Al get etched off and exposed the grey color of germanium. Measure mesa.
- Etch in 30% H2O2 for 30 min, Ge etched ~ 25 nm.
- Cap the Al with SiNx. Substrate temp. 250 °C, deposit rate ~ 150 Å/min. Recipe SiNx. Deposit 3 min, ~ 50 nm.
- Anneal in RTP: Always purge the chamber using PURGE program before any anneal. This is to purge anything left in the chamber from previous anneals.

<table>
<thead>
<tr>
<th></th>
<th>Time(s)</th>
<th>Temp (°C)</th>
<th>Gas 1 N2 slpm</th>
<th>Gas 2 O2 slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>60</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

Run O2 ambient 600 C for 5 minutes to burn any organic in the chamber, then run high temperature flush 850 C for 3 s before testing real process recipe.
• Verify the RTA program by running on a test wafer, program JZ600FO2.rpd

<table>
<thead>
<tr>
<th></th>
<th>Time(s)</th>
<th>Temp (C)</th>
<th>Gas 1 N2/H2 95/5</th>
<th>Gas 2 Ar/H2 slpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle</td>
<td>90</td>
<td>20</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Ramp</td>
<td>4</td>
<td>600</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Hold</td>
<td>2</td>
<td>600</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Idle</td>
<td>90</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

• Load substrate, place three spacers at 120 degree increments at the edge of the wafer
  Run anneal program

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Source</th>
<th>Measured Temp</th>
<th>Time (s)</th>
<th>Comment</th>
</tr>
</thead>
</table>

• Take optical pictures.
RAPID MELT GROWTH OF GERMANIUM TUNNEL JUNCTIONS

Jialin Zhao, Alan C. Seabaugh, and Thomas H. Kosel

University of Notre Dame
Department of Electrical Engineering
266 Fitzpatrick Hall, Notre Dame, IN 46556-5637
(574) 631-4473, seabaugh.1@nd.edu

Abstract

A rapid melt growth process for forming interband n+p+ Esaki tunnel junctions on Ge is shown. The process uses a phosphorus spin-on diffusant and rapid thermal annealing to form the n+ side of the junction, while for the p+ side, a deposited aluminum film serves both as an acceptor dopant source and a melt for epitaxial regrowth of p+ Ge. The current density in these junctions depends strongly on the peak temperature of the Al-Ge melt and ranges between a few µA/µm² to over 1 mA/µm². The use of a silicon nitride microcrucible improves the surface morphology by confining the melt. Record forward peak current density of 1.2 mA/µm² and reverse interband tunnel current density of 3.8 mA/µm² is achieved at a peak temperature of approximately 600 °C. ²

² Keywords: germanium, tunnel diode, epitaxy
Interest in germanium technology has been rekindled in recent years for MOSFETs [1], photodetectors [2], and nanowires [3]. This paper discusses a method for forming shallow, abrupt, heavily-doped, \( p \)-type layers in Ge with particular application to the formation of Esaki \( p^+n^+ \) tunnel junctions. The rapid melt growth technique described here is also relevant to Ge \( p^+ \) contact formation in devices where implantation cannot be used due to thermal limits. The ability to form high-current-density tunnel junctions is a central requirement in low-power-dissipation field-effect tunnel-transistors [4] and in circuits combining transistors and tunnel diodes [5].

To achieve high current-density in tunnel junctions requires formation of the highest possible \( p \) and \( n \) doping densities with the maximum doping abruptness (steepest possible slope of the doping concentration at the junction). High doping abruptness and doping density are needed to maximize interband tunneling current. In Ge, the \( p \) dopant with the highest solid solubility is Al, exceeding \( 2 \times 10^{20} \text{ cm}^{-3} \) in the temperature range from 600-700 \(^\circ\)C [6]. Interdiffusion of the dopant elements across the interface should be minimized to achieve the most abrupt doping profile. Hence, in this work a rapid-melt-growth technique has been explored, to minimize the time available for interdiffusion of the dopants across the junction.

The use of rapid thermal growth techniques for recrystallizing deposited Ge from crystalline Si seeds has shown that both lateral and vertical crystallization can be
achieved with high growth velocities (in m/s) driven by rapid thermal processors [7-9]. In this work, the rapid melt growth technique is extended by the introduction of Al over the Ge to act as both an acceptor dopant and as a melt to dissolve and regrow an abrupt $p^+$ on $n^+$ junction. This approach stems from early Ge bipolar transistor manufacturing, where pellets of In or Pb were placed on Ge wafers and used for melt-back and regrowth of $pnp$ Ge transistor junctions [10].

The In-Ge and Pb-Ge alloy systems have simple eutectic binary phase diagrams with relatively low eutectic temperatures. Annealing above the eutectic temperature first dissolved some Ge into the liquid phase, and then solidified a doped Ge layer during cooling, which formed the collector or emitter of the transistor. The remainder of the liquid then formed a solid eutectic mixture which acted as a soft metal alloy contact on the surface. Soft metals were used to reduce mechanical stresses when contact wires were attached. In these alloy junction processes, heating and cooling rates were typically around 0.5 °C/s with a peak annealing temperature of 550 °C [10]. In the present work, heating and cooling rates are increased by more than two orders of magnitude to 100–150 °C/s degrees Celsius per second, to achieve high dopant abruptness. At these high rates, the regrowth time is reduced from approximately 1000 s to around 4, decreasing the dopant diffusion length ($\sqrt{Dt}$) during solidification by approximately 15 times.
**Approach**

The metallurgical principles of our process can be described with the aid of the simple eutectic Al-Ge phase diagram [11] of Fig. 1. Upon heating Al in contact with Ge to the eutectic temperature $T_E = 420 \, ^\circ\text{C}$, Al and Ge would be expected to react to form a liquid with a Ge weight percentage of 51.6%. Above the eutectic temperature, the weight percentage of Ge in the Al increases along the liquidus line until a peak temperature is reached. Upon cooling from the peak anneal temperature, Ge solidifies from the melt, and (as in earlier work [10]) this should be expected to occur epitaxially on the pre-existing Ge surface. The regrown Ge epilayer is degenerately $p^+$ doped, since its Al content should follow the solubility line separating the L+Ge and solid Ge regions; solubility increases as the alloy cools. Especially during rapid cooling, it should be a good approximation to assume no Al diffusion in the solid Ge, so the first regrown Ge (which defines the junction) would have a lower Al content than the last. Precipitation of Ge during cooling drives the liquid composition of the melt along the liquidus line, reaching the eutectic composition at $T_E$, when a eutectic mixture of Al-rich and Ge-rich phases solidify. The temperature at which the eutectic reaction occurs will likely be lower during rapid cooling, but the $p^+n^+$ junction should not be affected by this. The liquid composition attained at the maximum temperature of the process may also differ from that given by the phase diagram.
A schematic description of the tunnel junction process is shown in Fig. 2. Germanium
$p$-type (100) Czochralski substrates were used, with resistivity of 0.18-0.19 $\Omega \cdot \text{cm}$. Wafers
were degreased in acetone and methanol followed by removal of the native oxide in a
10:1 buffered HF etch for 25 seconds. The wafers were then spin coated with an
Emulsitone phosphosilicate containing P at a concentration of $1 \times 10^{21}$ cm$^{-3}$, and
subsequently baked to remove the organics at 150 °C for 10 minutes in air. Samples were
next loaded into a Modular Process Technology RTP600 rapid thermal processor.
Phosphorus was diffused at 800 °C for 300 s in a forming gas (95N$_2$:5H$_2$, 10 slpm)
ambient under atmospheric pressure. From Hall-effect measurements, this process was
found to yield an electron density of approximately $2 \times 10^{20}$ cm$^{-3}$. A buffered HF clean
of 300 s removed the spin-on-diffusant residues prior to electron-beam evaporation of
100 nm of Al. The Al was patterned into square mesas, ranging in size from $3 \times 3$ to $20 \times
20 \mu$m$^2$, by photolithography and etching in Cyantec’s Al-12 etchant (71 wt% H$_3$PO$_4$: 10
wt% CH$_3$COOH : 2 wt% HNO$_3$). The Ge was next etched to a depth of approximately 40
nm in 30% H$_2$O$_2$ using the Al metal as a mask, as shown in Fig. 2(b).

Rapid melt growth (100 - 150 °C/s) was performed in a forming gas ambient, both with
and without a 50 nm thick silicon nitride (SiN) capping layer formed by plasma-enhanced
chemical vapor deposition, Fig. 2(b). On heating the wafer above the Al-Ge eutectic
temperature of 420 °C, the Al consumes an amount of Ge which depends on temperature.
At the peak anneal temperature the ratio of the penetration depth, $x$, of the melt below the
original Ge surface (defined in the inset of Fig. 3) to the deposited Al thickness, $t_{Al}$, can be determined from the Al-Ge binary phase diagram of Okamoto [11] according to the relation,

$$\frac{x}{t_{Al}} = \frac{\rho_{Al}}{\rho_{Ge}} \frac{W_{Ge}}{1-W_{Ge}}.$$

Here, $\rho_{Al}$ and $\rho_{Ge}$ are the densities in g/cm$^3$ for Al and Ge respectively, and $W_{Ge}$ is the weight percent of Ge in the Al melt, given by the point on the liquidus line at $T_P$. This relation is plotted in Fig. 3. For a peak anneal temperature of 600 °C, Ge is consumed to a depth of approximately 110 nm. Equation (1) does not account for lateral melting of the Ge, but for contact dimensions much greater than the melt-back depth, this formula provides a reasonable estimate.

From the peak temperature, cooling rates were approximately 40 °C/s for 10 s through the eutectic temperature of 420 °C, then cooling to under 100 °C in 2 minutes. On cooling, a heavily-Al-doped Ge layer is regrown until the eutectic temperature is reached. As computed from the phase diagram, the doping density, $N_{Al}$ in cm$^{-3}$, decreases linearly with increasing temperature from $4.7 \times 10^{20}$ cm$^{-3}$ at 420 °C to $2.1 \times 10^{20}$ cm$^{-3}$ at 725 °C, Fig. 3. As the temperature decreases below 420 °C, Al-rich and Ge-rich phases are nucleated in roughly equal proportions, forming a eutectic mixture labeled AlGe in Fig. 2(d).
Results

The rapid-melt-growth process was first explored without the presence of a SiN cap during the anneal. Both anneal peak temperature and anneal time at the peak temperature were varied. The anneals were performed on a subdivided single wafer to ensure comparisons of junctions with the same phosphorus diffusion. The current voltage, $I$-$V$, characteristics for 5 different anneal conditions are shown in Fig. 4. The measurements are made between small and large area contacts on the top surface of the wafer as indicated in the inset of Fig. 4. The contacts are in series, and since the voltage drop across the larger device is negligible due to its area, the voltage across the smaller device can be assumed to be equal to the applied voltage. The characteristic negative differential resistance of the tunnel junction is clearly observed for all anneal conditions and the reverse Zener tunneling characteristic is also apparent in each device. The series resistance (less than 10 $\Omega$) is responsible for the curvature at high currents; this series resistance is reasonable given the sheet resistance of the phosphorus diffusion (9 $\Omega/\square$) and the device geometry.

The peak forward current density, $I_F$, measured at the peak voltage, $V_P$, is plotted as a function of anneal condition in Fig. 5. The peak current density decreases monotonically with increasing temperature and time. This change is consistent both with the expected decrease in the junction doping with increasing temperature, Fig. 3, and the expected broadening of the dopant slope with temperature due to interdiffusion of phosphorus and
aluminum at the junction. Both of these changes decrease the tunneling probability and tunnel current density, which is what is observed. Also plotted in Fig. 5 is the reverse Zener tunneling current density corresponding to a reverse voltage with an absolute magnitude equal to the forward peak voltage. At this voltage, the reverse current density should exceed the forward current density because of the greater overlap of available states for tunneling in the reverse direction over the forward direction, and this is what is observed.

Figure 6 plots the dependence of peak-to-valley current ratio (PVR) for these diodes as a function of anneal condition. For the highest current density diode in this set, the 600 °C spike anneal, a current density of 0.16 mA/µm² is obtained at a PVR of 1.06. The PVR in the rapid-melt-growth junctions for these conditions are notably lower than has been achieved previously on alloy Ge diodes, PVR of 8-9 for a current density of 0.16 mA/µm² [12]. The low PVR is an indication of a high density of defect-states at the junction leading to a high excess current in the valley region of the I-V characteristics [13]. The junction roughness likely also plays a role in the high valley current.

A 50 nm SiN cap was added to act as a microcrucible [5, 6] and hold the Al-Ge melt during the alloy process. The addition of the cap prior to annealing was found to improve the morphology of the contact following anneal by holding the Al in place and suppressing islanding. The optical micrographs in Fig. 7 show a comparison of $5 \times 5 \mu m^2$
devices after annealing (a) with and (b) without a SiN capping layer. The left and right thirds of each micrograph are regions that were originally covered with a continuous Al film prior to annealing, while the middle region contains a $5 \times 5 \mu m^2$ deposited Al feature surrounded by a darker Ge region. In (a), the capping layer covered both the Al and Ge areas. The Al deposition was 100 nm thick and the anneal conditions were 440 °C, 5 s, with a 100 °C/s heating rate. The fact that the dark part of the central third of the images is Ge with no Al deposit suggests that the dark regions in the metallized areas are the Ge phase of the eutectic mixture, while the white areas are the Al phase. The uncapped specimen in (b) shows large areas of Al and of Ge in both the $5 \times 5 \mu m^2$ feature and the larger metallized areas at the left and right, suggesting that the melt formed droplets instead of wetting the Ge. Addition of the capping layer in (a) clearly resulted in much more uniform coverage of the substrate, and the dark lines in the metallization have a typical eutectic morphology. As the size of the capped features was reduced (as small as $3 \times 3 \mu m^2$), the uniformity of the Al-Ge film was observed to improve and the peak current density increased (as will be shown).

With the anneal hold time set at 1 s and the same heating rate of 100 °C/s, a sequence of rapid melt growths was performed on another set of wafers fabricated with the same phosphorus-diffusion conditions, but this time with a 50 nm SiN cap. The dependence of forward peak current and reverse current vs. anneal temperature from 440 to 725 °C for these are shown in Fig. 8. In this case a low constant reverse voltage is plotted which
similarly tracks the forward-current tunneling-dependence on temperature. A maximum in the tunneling current is found at a temperature of 600 °C. This current-density dependence on temperature is consistent with an improving electrical activation as temperature increases from 440 to 600 °C. The decreasing tunnel current at temperatures over 600 °C likely results from a degraded dopant slope, due to phosphorus diffusion into the regrown layer. Phosphorus has a diffusion length of approximately 6 nm in 1 s at 700 °C and 1 nm in 1 s at 600 °C in agreement with a degraded dopant slope and lower current density for anneals at 700 °C and above. The Al diffusion coefficient is an order of magnitude lower than phosphorus at these temperatures. Degradation in the peak current would also result from a decrease in junction doping which, as shown in Fig. 3, decreases as peak anneal temperature is increased. Germanium tunnel diodes with a peak current density of 0.52 mA/μm² and PVR of 1.5 are achieved under these conditions. More than three times the current density and slightly better PVR are achieved using the SiN cap.

Germanium tunnel diodes with current densities up to 1.2 mA/μm² were obtained using a ramp rate of 150 °C/s, an anneal condition of 600 °C, a hold time of 2 s, and a SiN cap. The device current voltage characteristics are shown in Fig. 9, showing an area dependence of the current density. The areas of these devices were all measured under a scanning electron microscope. As the device size was reduced the current density increased. The highest peak current density, 1.2 mA/μm², is achieved at the smallest
junction area with reverse current density of 3.8 mA/µm² also achieved in this junction. The improvement in current density with reduced junction area could result from an improvement in junction morphology or an increase in edge leakage.

Conclusions

Germanium tunnel diodes with peak current densities up to 1.2 mA/µm² were fabricated using a rapid-melt-growth technique for growing heavily-Al-doped $p^+$ layers on $n^+$ phosphorus-diffused Ge for tunnel junctions. The use of a silicon nitride cap greatly improved the surface morphology. A wide parameter space exists for raising the peak current density and reducing the valley current including optimization of the melt-back depth, optimization of the heating and cooling curves, and reduction in the device area to improve the junction uniformity.

Acknowledgements

This work was supported by the Office of Naval Research under contract N00014-02-1-0924.
References


Figure Captions

Figure A1. Al-Ge phase diagram after Okamoto [11].

Figure A2. Schematic device cross section of the rapid melt growth process drawn to scale in the vertical direction. (a) Germanium is doped $n^+$ by rapid thermal diffusion of phosphorus followed by aluminum deposition, patterning, and etching. (b) A plasma nitride cap is applied. (c) At the peak anneal temperature (600 °C in this example) an AlGe melt forms with a meltback of 110 nm (d) On cooling, a heavily-doped $p^+$Ge layer is grown; on cooling from a peak temperature of 600 °C, a 60 nm layer is grown.

Figure A3. The ratio of the penetration depth, $x$, to the aluminum deposition thickness, $t_{Al}$, as a function of peak anneal temperature, and the concentration of Al, $N_{Al}$ in cm$^{-3}$, as a function of the peak anneal temperature as determined from the Al-Ge binary phase diagram [10].

Figure A4. Dependence of Ge tunnel diode current-voltage characteristics on rapid-melt-growth anneal condition. The silicon nitride cap was omitted for this set of anneals. Solid lines are used for the 600 °C anneals and dashed lines for the 700 °C
anneals; the symbols in the legend and on the curves indicate the hold time at the peak anneal temperature.

Figure A5. Dependence of current density in the forward, $I_F$, and reverse current, $I_R$, directions as a function of anneal condition for the uncapped Ge diodes of Fig. 4. In the reverse direction, the current magnitude is plotted at a voltage corresponding to minus the forward peak voltage, $V_P$.

Figure A6. Peak-to-valley current ratio vs. anneal condition for the uncapped Ge diodes of Fig. 5.

Figure A7. Optical micrographs showing the Ge wafer surface following rapid melt growth of $p^+$Ge at a peak temperature of 440 °C.

Figure A8. Germanium tunnel diode current densities vs. annealing temperature with a 50 nm SiN cap. The solid circles denote measurements at the peak voltage in the forward current direction; the open circles are measurements in the reverse current direction at -0.01 V. The values for the peak-to-valley current ratio ($PVR$) are given on the forward current curves.

Figure A9. Current-voltage characteristics of Ge tunnel diodes with a 50 nm silicon nitride microcrucible layer and a 150 °C/s heating rate.
Figure A1

Schematic device cross section of the rapid melt growth process drawn to scale in the vertical direction. (a) Germanium is doped $n^+$ by rapid thermal diffusion of phosphorus followed by aluminum deposition, patterning, and etching. (b) A plasma nitride cap is applied. (c) At the peak anneal temperature (600 °C in this example) an AlGe melt forms with a meltback of 110 nm (d) On cooling, a heavily-doped $p^+$Ge layer is grown; on cooling from a peak temperature of 600 °C, a 60 nm layer is grown.
Figure A3

The ratio of the penetration depth, $x$, to the aluminum deposition thickness, $t_{Al}$, as a function of peak anneal temperature, $T$, and the concentration of Al, $N_{Al}$ in cm$^{-3}$ as a function of the peak anneal temperature as determined from the Al-Ge binary phase diagram [10].
Dependence of Ge tunnel diode current-voltage characteristics on rapid-melt-growth anneal condition. The silicon nitride cap was omitted for this set of anneals. Solid lines are used for the 600 °C anneals and dashed lines for the 700 °C anneals; the symbols in the legend and on the curves indicate the hold time at the peak anneal temperature.
Figure A5

Dependence of current density in the forward, $I_F$, and reverse current, $I_R$, directions as a function of anneal condition for the uncapped Ge diodes of Fig. 4. In the reverse direction, the current magnitude is plotted at a voltage corresponding to minus the forward peak voltage, $V_P$. 
Figure A6

Peak-to-valley current ratio vs. anneal condition for the uncapped Ge diodes of Fig. A5.
Figure A7

Optical micrographs showing the Ge wafer surface following rapid melt growth of \( p+\text{Ge} \) at a peak temperature of 440 °C.
Germanium tunnel diode current densities vs. annealing temperature with a 50 nm SiN cap. The solid circles denote measurements at the peak voltage in the forward current direction; the open circles are measurements in the reverse current direction at -0.01 V. The values for the peak-to-valley-current ratio ($PVR$) are given on the forward current curves.
Current-voltage characteristics of Ge tunnel diodes with a 50 nm silicon nitride microcrucible layer and a 150 °C/s heating rate.

Figure A9
REFERENCES


156


