Conventional computers are approaching fabrication limits where statistical variation in the placement of single dopants greatly impacts device performance. Additionally, practical modelling of more than a handful of atoms while accounting for quantum properties can only be performed with a quantum computer. Fabricating a solid state Si:P Kane architecture will require deterministic doping; the ability to accurately place dopants within a substrate.

We have implemented an apparatus for creating arrays of single P donor atoms within Si. We model the fabrication

<table>
<thead>
<tr>
<th>1. REPORT DATE (DD-MM-YYYY)</th>
<th>2. REPORT TYPE</th>
<th>3. DATES COVERED (From - To)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-08-2015</td>
<td>Ph.D. Dissertation</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE AND SUBTITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>The fabrication of arrays of single ions in silicon via ion implantation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5a. CONTRACT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>W911NF-13-1-0024</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5b. GRANT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5c. PROGRAM ELEMENT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5d. PROJECT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5e. TASK NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5f. WORK UNIT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. AUTHORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jessica van Donkelaar</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. PERFORMING ORGANIZATION NAMES AND ADDRESSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of New South Wales</td>
</tr>
<tr>
<td>Research Office</td>
</tr>
<tr>
<td>University of New South Wales</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. PERFORMING ORGANIZATION REPORT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. Army Research Office</td>
</tr>
<tr>
<td>P.O. Box 12211</td>
</tr>
<tr>
<td>Research Triangle Park, NC 27709-2211</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10. SPONSOR/MONITOR'S ACRONYM(S) ARO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>63162-PH-QC.99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>12. DISTRIBUTION AVAILABILITY STATEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release; distribution is unlimited.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional computers are approaching fabrication limits where statistical variation in the placement of single dopants greatly impacts device performance. Additionally, practical modelling of more than a handful of atoms while accounting for quantum properties can only be performed with a quantum computer. Fabricating a solid state Si:P Kane architecture will require deterministic doping; the ability to accurately place dopants within a sub-strate.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15. SUBJECT TERMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>quantum computing, single ion implantation, single P donor atoms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. SECURITY CLASSIFICATION OF:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. REPORT UU b. ABSTRACT UU c. THIS PAGE UU</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. LIMITATION OF ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>UU</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18. NUMBER OF PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19a. NAME OF RESPONSIBLE PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Dzurak</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19b. TELEPHONE NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>612-938-5631</td>
</tr>
</tbody>
</table>
ABSTRACT

Conventional computers are approaching fabrication limits where statistical variation in the placement of single dopants greatly impacts device performance. Additionally, practical modelling of more than a handful of atoms while accounting for quantum properties can only be performed with a quantum computer. Fabricating a solid state Si:P Kane architecture will require deterministic doping; the ability to accurately place dopants within a substrate.

We have implemented an apparatus for creating arrays of single P donor atoms within Si. We model the fabrication of a three-donor device for demonstrating the coherent transfer by adiabatic passage transfer (CTAP) protocol and show that it can be done with deterministic ion implantation through a 20 nm nanostencil stepped 30 nm between implants to produce a yield of 1-in-7 well placed arrays.

Single ion implantation is an established method of device fabrication that involves detecting single 14 keV P+ ions entering a substrate via the charge induced as the ion collides with the lattice. We have tested various designs for the single ion detector to evaluate the sources of good performance and achieved 98±4% counting efficiency with the classic design in high purity silicon.

We have fabricated 200 nm think Si3N4 nanostencils with slot and hole shaped apertures as small as 20 nm using the focussed ion beam (FIB) technique of back filling larger apertures with Pt. These adequately mask a 14 keV P+ ion to define the implant site.

We have built apparatus that employs a piezo-positioner to move the nanostencil above the surface of a single ion detector. This system is LN2 cooled within a high vacuum to implement deterministic doping. The fine alignment of the nanostencil to the implant site is performed using the ion impact signals as feedback over a patterned area of the detector and we have demonstrated this concept to a precision of sub-100 nm.
Author/s:
VAN DONKELAAR, JESSICA

Title:
The fabrication of arrays of single ions in silicon via ion implantation

Date:
2014

Persistent Link:
http://hdl.handle.net/11343/39643

File Description:
The fabrication of arrays of single ions in silicon via ion implantation
The fabrication of arrays of single ions in silicon via ion implantation

Jessica van Donkelaar

Submitted in total fulfillment of the requirements of the degree of Doctor of Philosophy
February 2014

School of Physics, Faculty of Science
The University of Melbourne
Abstract

Conventional computers are approaching fabrication limits where statistical variation in the placement of single dopants greatly impacts device performance. Additionally, practical modelling of more than a handful of atoms while accounting for quantum properties can only be performed with a quantum computer. Fabricating a solid state Si:P Kane architecture will require deterministic doping; the ability to accurately place dopants within a substrate.

We have implemented an apparatus for creating arrays of single P donor atoms within Si. We model the fabrication of a three-donor device for demonstrating the coherent transfer by adiabatic passage transfer (CTAP) protocol and show that it can be done with deterministic ion implantation through a 20nm nanostencil stepped 30nm between implants to produce a yield of 1-in-7 well placed arrays.

Single ion implantation is an established method of device fabrication that involves detecting single 14keV P+ ions entering a substrate via the charge induced as the ion collides with the lattice. We have tested various designs for the single ion detector to evaluate the sources of good performance and achieved 98±4% counting efficiency with the classic design in high purity silicon.

We have fabricated 200nm think Si₃N₄ nanostencils with slot and hole shaped apertures as small as 20nm using the focussed ion beam (FIB) technique of
back filling larger apertures with Pt. These adequately mask a 14 keV P$^+$ ion to define the implant site.

We have built apparatus that employs a piezo-positioner to move the nanostencil above the surface of a single ion detector. This system is LN$_2$ cooled within a high vacuum to implement deterministic doping. The fine alignment of the nanostencil to the implant site is performed using the ion impact signals as feedback over a patterned area of the detector and we have demonstrated this concept to a precision of sub-100 nm.
Declaration

This is to certify that:

1. The thesis comprises only my original work towards the PhD except where indicated.

2. Due acknowledgment has been made in the text to all other material used.

3. The thesis is less than 100,000 words in length, exclusive of tables, bibliographies, appendices and footnotes.

__________________________________________
Jessica van Donkelaar
Preface

The Si substrate detectors we employ for single ion implantation are fabricated at the Semiconductor Nanofabrication Facility (SNF) at the University of New South Wales (UNSW) as a collaboration between nodes of the Centre for Quantum Computer Technology (CQCT). The detectors produced during the course of this work in particular were fabricated primarily by Nai Shyan Lai and Fay Hudson.

Post doctoral fellow Andrew Alves collaborated in designing the apparatus and worked with me to get the MeV IBIC data of section 3.0.9 and the shadow mask implants of section 5.2.
Acknowledgments

I would like to thank the many people who provided both technical assistance and moral support during my PhD project. My sincere thanks go to Roland Szymanski for showing me the inner workings of the pelletron and to Alberto Cimmino similarly laying bare the colutron. I am grateful to Sergey Rubanov for the training of the FIB/SEM system and assistance with the TEM. Thank you Andy Greentree for CTAP advice, to Changyi Yang for all his invaluable help in the lab and to Andrew Alves for working with me.

Thank you to my supervisor David Jamieson for his enthusiasm and attentive guidance over the years.

And finally, thanks go to my family and friends for their support, especially my parents Meryl & Patrick van Donkelaar and my husband David Perry.
# Contents

1 Introduction 1
   1.1 History of the transistor; physics and fabrication . . . . . . . . . . 3
      1.1.1 Current state of the transistor art 9
   1.2 Quantum Computing . . . . . . . . . . . . . . . . . . . . . . . 12
      1.2.1 QC schemes . . . . . . . . . . . . . . . . . . . . . . . . 14
      1.2.2 Coherent Transport . . . . . . . . . . . . . . . . . . . . 25
   1.3 Deterministic Doping . . . . . . . . . . . . . . . . . . . . . . . 28

2 Modelling Ion Implantation 39
   2.1 Ion-solid interaction . . . . . . . . . . . . . . . . . . . . . . . . 39
      2.1.1 Electronic stopping . . . . . . . . . . . . . . . . . . . . . . 44
      2.1.2 Nuclear stopping . . . . . . . . . . . . . . . . . . . . . . . 47
      2.1.3 Binary Monte-Carlo simulation . . . . . . . . . . . . . . . . . 50
   2.2 Coherent tunneling by adiabatic passage . . . . . . . . . . . . . . . . 55
   2.3 Modelling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 59

3 Single ion implantation 70
   3.0.1 The detector device . . . . . . . . . . . . . . . . . . . . . . . . 75
   3.0.2 Fabrication . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79
   3.0.3 Operation . . . . . . . . . . . . . . . . . . . . . . . . . . . . .81
   3.0.4 Pulse height . . . . . . . . . . . . . . . . . . . . . . . . . . . . 84
   3.0.5 Noise . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 86
List of Figures

1.1 Band diagram for crystalline Si. .......................... 4
1.2 The progression of transistor structure. .................. 10
1.3 Spin qubit representations. ................................. 13
1.4 Depiction of NMR spin qubits (a) molecules suspended in liquid
and (b) Si solid state. ........................................... 15
1.5 Depiction of (a) trapped ion and (b) trapped atom qubit devices. 17
1.6 Images of superconducting (a) charge, (b) flux and (c) phase
qubit devices. ...................................................... 19
1.7 Images of physically defined quantum dot structures made with
(a) Si:P, (b) cavity QED and (c) heterostructures. ............ 20
1.8 Images of electrostatic quantum dot qubit devices in (a) GaAs/AlGaAs,
(b) Si/SiGe and c) Si MOS. ........................................ 21
1.9 Spin transport techniques including (a) CNOT swapping, (b)
CTAP and (c) spin bus. ............................................ 26
1.10 The bottom-up STM process of placing P atoms on to Si. ... 29
1.11 Methods of deterministic ion implantation. ................ 31
1.12 The deterministic doping apparatus. ......................... 34
1.13 A schematic of the deterministic doping process. .......... 35
1.14 Comparison of the electrical characteristics of ordered donor
arrays to random dopant placement on a 100 nm scale (a) and
10 nm scale (b). .................................................... 37
1.15 The dopant structures made possible by deterministic doping. . 38
2.1 Comparison of SRIM modelling to Geant4 modelling in the low energy heavy ion regime. 42
2.2 Schematic of the products resulting from ion-solid interaction. 44
2.3 Stopping force comparison of 0.5 MeV He\(^+\) and 14 keV P\(^+\). 51
2.4 The three donor proof-of-concept device for demonstrating CTAP. 55
2.5 Depiction of various systems which utilise the (a) STIRAP and (b) CTAP transport mechanisms. 57
2.6 The (a) alternating and (b) straddling CTAP schemes for long range CTAP transport. 58
2.7 Considerations for fabricating the CTAP device with ion implantation (a)(b) spacings and (c)(d) ion straggle. 60
2.8 The CTAP tunnel matrix elements of two donor placement scenarios that demonstrates the effect of fabrication limitations. 62
2.9 Modelling results that give the fraction of well placed donors for CTAP when fabricated by ion implantation. 67

3.1 A schematic of the deterministic doping process. 71
3.2 (a) Schematic of the Colutron 14kV ion accelerator, (b) the ion source, (c) Fischer slits and (d) the target chamber. 72
3.3 The band bending of p-n junctions that determines charge carrier flow. 76
3.4 A schematic of the PiN detector device from (a) the side and (b) above. 80
3.5 The basic preamplifier circuit. 82
3.6 Leakage current under reverse bias. 83
3.7 Signal sorting to obtain a pulse height spectrum. 84
3.9 Design schematics of the detectors studied in this chapter (a) DP, (b) DP-CS-B and (c) DP-CS-C. 95
3.10 Reverse bias breakdown (a) theory and (b)(c) measured in our PiN detectors. 99
3.11 Comparison of dark current of a good, stable detector and a poor detector. .................................................. 100
3.12 Schematic of the Pelletron 5MU accelerator. ......................... 103
3.13 IBIC CCE image of a single ion detector obtained by Changyi Yang in 2006. .............................. 104
3.14 IBIC CCE map of a DP-CS-B detector. .................................. 104
3.15 IBIC CCE map of a DP-CS-C detector. .................................. 105

4.1 Depiction of lithography methods. ................................. 111
4.2 Comparison of the depth and physical distribution of UV photons, high energy protons and low energy heavy ions used in lithography. ........................................... 113
4.3 The scattering of 10 keV electrons. ................................. 114
4.4 Schematic of the FEI Nova 600 Nanolab dual FIB/SEM system used for the fabrication of nanostencils. ................. 118
4.5 Schematic of electron and ion focussing within an SEM and FIB column respectively. .................................. 119
4.6 Depiction of the Pt deposition process as in [1]. ................. 122
4.7 Deposition profiles obtained at different current to precursor ratios. ......................................................... 123
4.8 Various SEM images of issues encountered during the FIB fabrication process. ............................................. 125
4.9 (a) the current feedback during the initial FIB milling and (b) images showing Pt deposition narrowing the initial aperture. . . 127
4.10 TEM images of the FIB milled slots before Pt deposition. .... 129
4.11 TEM images of 20 nm wide (a) slot and (b) hole apertures .... 130

5.1 Comparison of shadow mask techniques for metal evaporation and ion implantation. .................................. 134
5.2 Depiction of ion scattering through the mask regions surrounding a nanostencil aperture. ........................................ 137
5.3 Images of (a) alignment markers on the sample and (b) the nanostencil used for the experiment. .......................... 138
5.4 (a) AFM and (b) KFM images of the implanted sample. .... 140
5.5 A comparison of (a) AFM (b) SEM and (c) KFM scans of the same samples. .................................................. 141
5.6 The deterministic doping apparatus for the fabrication of single ion arrays. ................................................. 144
5.7 Photograph and schematic of the Transax micropositioner stage used to mount the apparatus in the target chamber. ....... 146
5.8 Close-up photograph of the camera in position to align the nanostencil to the target implant site. ......................... 147
5.9 Photographs from the camera showing the alignment process. ... 148
5.10 Schematic of the axes in the apparatus requiring alignment. ... 149
5.11 Squaring of the apparatus axes. ........................................ 149
5.12 Experimental ion induced charge signals from ions passing through PMMA and directly into the detector via an EBL fabricated trench. .................................................. 151
5.13 nanoIBIC image of the surface compared to AFM scans of the same sample. .................................................. 151
5.14 Depiction of a position alignment process with a resolution better than the width of the nanoaperture used. ............... 152
5.15 (a)(b)(c) Possible sacrificial alignment patterns in different scenarios. .................................................. 154
5.16 The coefficients of thermal expansion of the main materials used in our apparatus. ...................................... 156
5.17 (a) nanoIBIC and (b) microscope image of an S301 detector. .. 158
5.18 nanoIBIC trace that shows the detector surface. ............... 159
5.19 Proposed design for future detectors.
List of Tables

2.1 The percentage of three donor devices which satisfy control or coherence restrictions within our model. . . . . . . . . . . . . . 68

3.1 Performance summary for the original single ion system. . . . . 96

3.2 Performance summary for the new deterministic doping system. 96
Chapter 1

Introduction

The most important material for the semiconductor industry, Si, owes much of its success to the impurities contained within it. In particular, our ability to place certain dopants within specific regions allows us to precisely control the electrical properties of the material and create devices such as the transistor, ubiquitous in modern technology. In fact, the scale of a single transistor within an integrated circuit has been steadily decreasing since its inception, providing improved processing power and lower monetary cost [2]. The length of a transistor channel is approaching the scale where the positions of single dopant atoms can drastically affect the properties of the device [3, 4]. This implies there is a limit to the gains acquired by shrinking transistors and that we will require alternative means of improving computer performance, such as novel materials or device structure [5–7]. Further to this, exact simulations with full reckoning of quantum interaction in real systems are intractable on classical computers [8,9] but possible with the introduction of a new paradigm in computing, the quantum computer. While classical information processing is based on bits represented by zeroes and ones with Boolean logic, quantum logic based on the entanglement between $|0\rangle$ and $|1\rangle$ qubit states allows quantum parallelism which outstrips the limits of classical computing with only
30-40 qubits [10]. Although there are many areas that could possibly benefit from quantum computation [11] it has been calculated explicitly that brute force searching can require as few as \( \sqrt{N} \) steps compared to \( N \) for a classical computer [12] and a factorising quantum algorithm that is exponentially faster than any classical algorithm has been found [13]. There are many physical systems that provide necessary qualities for implementing a quantum computer and a few in particular have surpassed critical milestones on the route towards full scale fabrication including photonic systems [14], trapped ion schemes [15], superconducting circuits [16] and solid state spin systems [17]. Most however are now limited by scalability; even where a full scale architecture has been envisioned, our ability to build it has not yet been established. Strategies which employ the vast repertoire of techniques already in place for silicon IC fabrication, would seem to be at an advantage. This includes the approach devised by Bruce Kane to utilise the spin of a P atom in silicon, known as the Si:P system [18]. In QC research, transistors small enough to be effected by single atoms are utilised to observe the quantum properties of single P [19] or As donors [20]. Many of the seminal results published on single to few atom silicon devices were obtained from dopants that were incorporated by chance [20] or during timed blanket implants [21]. In order to advance our research further than current successes, we require the ability to precisely place single P atoms into a silicon substrate to create multi donor devices. This thesis presents a solution to this obstacle in the form of deterministic ion implantation. The ability to fabricate arrays of precisely placed single atoms within a semiconductor will not only allow us to fabricate devices to demonstrate coherent transport of quantum information [22] but could also prove useful in classical electronics [23] or novel device production.

In order to fully appreciate the context of deterministic doping via ion implantation, we must turn to the history of semiconductor physics [24, 25], the transistor [26–28], ion implantation [29] and the main driver of this industry,
the computer [30]. We have drawn on these literature reviews to provide the following summary of the early development that occurred to bring about the invention of semiconductor electronics and computers that are now ubiquitous in scientific research. Each story is intermingled with the others and credit has historically gone to the first person to the patent office or the first to publish results of a working device, despite being the culmination of years of effort from many different people. Vast amounts of money, time and effort have been invested by many people and since the inception of the transistor until now. In this introductory chapter the people credited for the invention and early development stages of semiconductor technology have an implied et al after each name. A critical understanding of solid state physics evolved along side the manufacturing processes involved, but the emphasis here is on the history of semiconductor manufacture for simplicity.

1.1 History of the transistor; physics and fabrication

The advent of semiconductor physics came after quantum mechanics was first established. With the discovery of the electron (1897) came the discovery of the photon (1905), the Bohr structure of the atom (1913) and eventually the Pauli exclusion principle (1925), Schrodinger wave theory (1926), Fermi-Dirac statistics (1926) and the Heisenberg uncertainty principle (1927). Quantum mechanics was required to explain the electrical properties of a semiconductor which appeared erratic to scientists who had no real measure of the purity of the samples or knowledge of p- and n-type distinction. In 1928 Bloch applied quantum mechanics to metallic solids and determined that electrons travelling through a periodic lattice are only scattered by imperfections in this periodicity, such as temperature fluctuations, impurities and at the ends of the
Figure 1.1: The band gap of Si is indirect at 1.1 eV, from [31].

structure [32], providing the origins of electrical resistance. At this time in history, the uncontrolled doping of semiconductors worked against the favour of Si. The very existence of an intrinsic semiconductor was once in question and Gudden suggested in 1930 that no substance could be semiconductor without impurities, but would become insulating when pure [33]. However, formal band structure theory was developed by Wilson in 1931 [34] by extending Bloch’s theory of quantum solids to encompass the idea of a band gap, including the first explicit description of donors and acceptors [35]. The strange behaviour of semi-conductors compared to metals could now be understood. While the resistance of conductors increases with temperature due to increased thermal scattering, the increase in conductance of semiconductors with temperature could be explained with Fermi-Dirac statistics of valence electrons, which distributed the energy of more electrons above the band gap with increasing temperature. Although it was decades before the theory was refined enough to produce the fine multi-valley structure of the bandgap seen in figure
1.1, Wilson’s work was enough to convince some scientists that such materials were worth pursuing. Before this enlightenment, vacuum tube components were considered to be the future of electronics. This included the vacuum tube diode produced by Fleming in 1904 [36] and the thermionic triode which acted as an amplifier, invented in 1907 by de Forest [37].

Perhaps in consequence, Ohl worked at the Bell Laboratories, a department for research and development for a telephone manufacturing company, to produce semiconductor replacements for poorly performing vacuum tube radio receivers. His initial results were largely unreproducible until he engaged Scaff and Theuerer from the Bell Laboratories metallurgical department for the purpose of purifying silicon. The earliest methods to purify silicon were simple melting techniques since doped regions were known to melt at lower temperatures than intrinsic silicon [38]. When Scaff cooled some Si ingots slowly to avoid cracking dislocations in 1939, he also unwittingly separated p and n dopants via diffusion along the length of the material. One of the samples produced for Ohl from such a slab straddled the p-n junction and Ohl noted its tendency to increase in conductance in the presence of light, which we now know is due to the photoelectric generation of carriers within the naturally occurring depletion width. Here was a current rectifying replacement for the vacuum tube diode! Like much of the industrial research undertaken during the second World War [27], his discovery was kept confidential until 1946 [39]. While Scaff and his team continued to characterise doped silicon [40, 41], the alteration of the electrical characteristics of germanium by the presence of impurities was written into a patent by Woodyard working at Sperry Gyroscope Company in 1944 [42] which is considered to constitute the first formally explicit description of intentional semiconductor doping.

Meanwhile, the first programmable computing machines which were considered ‘Turing complete’ [43,44] were built, such as the Z3 built from telephone relays
by Zuse et al in Berlin during 1941 and the Colossus series, secret machines built from a few thousand vacuum tube components by the British government in 1943 onwards to crack communication encryption during the war [45]. A Turing machine is essentially a black box that can be used to solve any given problem, which requires the ability to program. While reprogramming of these machines was possible, it was difficult and required physical rewiring. Flexible memory was only used for data writing. After the war, vacuum tube computers based on the Von Neumann [46] architectures were built that allowed memory to be used as program storage as well. Relay and vacuum tube based computers with the Harvard architecture were built at the same time which read program instructions from tape and executed them in the memory which meant that unlike the Von Neumann architecture, it had physically separate pathways for the input of instruction and data writing. Current computer architecture retains elements from both these designs: programs are stored as software alongside data, but they are executed within a distinct central processing unit (CPU) with a cache for efficient data flow so that multiple processes can occur at once.

Also working at Bell Labs, Bardeen and Brattain built the first semiconductor transistor in 1947 [47]. It was a point-contact transistor formed by mounting a slab of germanium on a large metal contact and pressing two point contacts close together on the surface, as shown in figure 1.2. Voltage on one point contact modulated the current between the remaining surface point contact and the bottom contact. It was unreliable and not well understood but it brought sharp attention to the possibilities inherent in semiconductor technology. Schockley figured out how Ohl’s p-n junction worked and published the theory in 1949 [48]. Based on this he was able to fabricate a Ge bipolar junction transistor in 1950 [49] which resembles the transistors in use today, depicted in figure 1.2. The three were jointly awarded the 1956 Nobel Prize in Physics for their research in semiconductors, and in particular their discovery of the
transistor. (The reference [50] provides Schockley’s perspective of the story). While semiconductor transistors were immediately applied to computer architecture in laboratories, such discrete transistor computers were outperformed by vacuum tubes which continued to dominate industrial computer fabrication while transistor research continued.

One of the major impediments to reliable semiconductor device function was lifted with the development of better techniques for purifying Ge and then Si [51, 52]. Now known as the Czochralski process, the procedure involved using a seed crystal to extract a single large crystal from molten material. Impurities could be introduced during the growth or removed by applying the process repeatedly, since most of the dopant impurities would be relegated to the bottom of the ingot which could be removed. The availability of pure material greatly assisted semiconductor research and by 1954, the first silicon transistor was developed by Tanenbaum et al, also at the Bell Laboratories [53] and in 1957 Sah and Schockley amended the original p-n junction theory with the addition of carrier generation and recombination via trapping impurities within the depletion width to account for the seemingly anomalous behaviour of Si compared to Ge [54, 55].

Like most semiconductor devices up to this point, the dopants in the first silicon transistor were incorporated during the crystal growth [56, 57]. Though dopant incorporation via ion implantation was already recognised [58–60], the apparatus required for the ion source was costly and of questionable quality. By 1955, Tanenbaum and Fuller had invented the diffusion method of doping [61] where dopants are introduced to solid substrates via gases at high temperatures, at which the dopants are naturally mobile within the material and diffuse down into the substrate. It was taken up immediately and used in the invention of integrated circuitry which spanned the next decade, largely at the Farchild Camera and Instrument Corporation [62]. Kilby demonstrated
the first working integrated circuit (IC) while working at Texas Instruments in 1958 (for which he was awarded the Nobel prize in 2000) [63]. The UV lithography fabrication process, which is still in use today, was developed by Jack Last and Robert Noyce [64] for the reproducible manufacture of transistors at Fairchild. Hoerni, also at Fairchild, decided to use oxide to protect otherwise vulnerable devices which he patented in 1959 [65, 66]. At the same time, Kahng working at Bell Labs with Atalla, created the first working MOSFET (metal-oxide-semiconductor field effect transistor) which worked on the same use of SiO$_2$ and was patented and publicly presented in 1960 [67, 68]. Hoerni and independently conceived of the planar MOS structure in figure 1.2 that we recognise today and this made it possible for Noyce to develop integrated circuits with evaporated Al interconnects [69, 70] which he patented in 1961 and perfected with Last over the next few years [64]. The development of reliable lithographic UV fabrication technology, whether for IC or batch single transistor manufacture, marked the turning point in semiconductor technology. Oxide protected components were far more reliable than exposed junction transistors which enabled them to be sold commercially as replacements bulky vacuum tube switches.

Ion implantation superseded diffusion in the early 70s after the establishment of MOS [71] technologies which allowed batch lithography based fabrication of a higher standard than before. Ion implantation for semiconductor fabrication purposes was funded by NASA during the 1960s for space sensors [72], and technology that improved the stability and purity of ion beams was adapted from general nuclear physics accelerator ion beam technology. With these improvements the physical distribution and dopant concentrations that were achievable with ion implantation were more accurate than masked diffusion techniques could produce. Thus implantation fabricated devices performed better and could be built with higher density which swayed manufacturers to invest in ion implantation equipment. Si eventually superseded Ge, since
there were no dielectrics available for Ge that were as clean as the Si/SiO₂ interface. Si also has a higher melting point than Ge, which enabled the fabrication of high power devices. With the production of CMOS (complementary MOS) patented in 1967, high density low power circuits could be made. Now large scale integrated circuit fabrication was possible and with improvements afforded by the use of clean rooms and designs based on decades of vacuum tube computer research, the first working supercomputer made from silicon ICs was built by Cray et al in 1975 [73]. The first commercially sold Cray computer weighed more than 5 Tonnes, cost more than $8 million (1977) and had a maximum processing speed of 133 Mflops (floating-point operations per second), which is exceeded by the average commercially available personal computer today. The best supercomputer of 2011 exceeds 10 Pflops [74]. This billion-fold improvement was made possible by continuous down scaling of the IC circuit elements [75, 76] and the onset of massively parallel computing with thousands of CPUs in the early 1990s [77, 78], whether this is distributed ‘grid’ computing [79] or the centralised cluster computing approach. Ion implantation has retained dominance in IC fabrication [29, 80, 81], in great part due to the progressive physical models for ion-solid interaction [82–84] based on better experimental data. The physical models allowed manufacturers to select the energy for the appropriate depth and the angle of implantation for crystal substrates to avoid channeling [85]. There can be almost three dozen different ion implantation steps in a typical Complimentary-MOS (CMOS) circuit today, the relatively complex structure of which is seen in figure 1.2.

1.1.1 Current state of the transistor art

Right before the invention of MOS technology, Feynman gave a lecture at an American Physical Society meeting at Caltech in 1959 that proved to be a Jule’s Verne like premonition entitled ‘There’s plenty of room at the bottom’.
Figure 1.2: The evolution of transistor structure from the very first point contact and bipolar junction transistors to the advent of MOS structure and modern Complimentary-MOS with complex features for good device performance such as shallow trench isolation (STI).

Among other things he estimated that facial recognition, a typical human ability we take for granted, would require hardware the size of the Pentagon. He suggested the miniaturisation of all types of machine, starting with a motor, and he noted the existence of a quantum limit to this downscaling. He pronounced that man’s ultimate triumph over matter would be the ability to build structures atom by atom. At the time, the ultimate CMOS structure of electric components was not yet fixed and the IC scaling would not occur for another 15 years or so [2] and we are reaching the quantum limit of electronics scaling just now, 50 years later [7].

Other predictions made by Feynmann are in progress including the use of
artificial neural networks [86], atom-by-atom fabrication [87] and the keen interest in nanotechnology in general [88–91]. While current artificial facial recognition technology exists, even state of the art systems are very crude - they generally require people to face the camera and remove confusing things like hair, glasses or smiles [92,93]. If we continue the comparison of computers to the human brain, the IBM Blue Gene/P supercomputer managed to simulate an artificial neural network [94] with 1.6 billion neurons and 9 trillions which represents approximately 1% of the human cortex only [95]. During 2007-2008 this particular computer featured within the top ten on both the Top500 [74] and the Green500 [96] lists, which measure the top speed and efficiency of supercomputers respectively [97]. Since then research scientists have simulated a neural network with a size equivalent to the brain of a rat [98] on the IBM Blue Gene/L and the brain of a cat [99] IBM Blue Gene/P. The functionality of these networks, let alone the size, are a far cry from reaching the complexity or capability of human brain function [86].

Otherwise supercomputers are mostly used for weather/climate modelling [100] or protein folding for biomedical applications [101]. The computer was originally conceived as a universal operator that was capable of providing exact solutions [43, 44]. But since classical computers are not based in quantum physics but Boolean logic, they are inherently incapable of providing exact solutions for the simulation of quantum mechanical systems [8, 9]. This may seem like philosophical question of definition but it is simple to demonstrate the inefficacy of classical computing in a specific case without any explicit understanding of universal computing. The local memory of one of the thousands of CPUs in a supercomputer cluster is about 17 GB (that is, $2^{34}$ bytes) [74]. Such a computer could record the states of 34 entangled spins, maximum, (which does not allow for any computation of the evolution of the system). A quantum computer could perform the same task with only 34 entangled qubits, since n entangled spins provide $2^n$ entangled states [10]. Very recently, a physical
system has been demonstrated with up to 300 atoms in prepared entangle-
ment [102]; were this system the basis of a quantum computer, it would offer
a $10^{80}$ improvement on a contemporary state-of-the-art RAM! There is no fea-
sible way for conventional high performance computing to approach quantum
systems with this complexity.

The enticements of quantum information processing have prompted much ef-
fort in research labs around the world to find viable physical systems for the
realisation of quantum simulation in the past two decades [17, 103]. Here we
describe the requirements that need to be satisfied for physical implementa-
tion, and the main systems which are on their way to demonstrating these
properties.

### 1.2 Quantum Computing

The requirements for the physical realisation of a quantum computer were
originally set out by DiVincenzo in 2000 [104] and have been continuously re-
vised by researchers, particularly with the hindsight provided by a decade of
research in the QC field [11, 105]. Here we will only consider discrete quantum
computing strategies that consist of qubits, though we note that continuous
variable computing is possible in principle [106–108] where variables are entan-
gled via the Heisenberg uncertainty principle, and exploiting the quadrature
amplitudes of electromagnetic radiation is particularly suited to quantum key
distribution [109–111].

The majority of quantum computer strategies are based on the ability to en-
tangle $|0\rangle$ and $|1\rangle$ states to create superpositions within the Bloch sphere shown
in figure 1.3, which gives the quantum computer an edge over a classical bi-
nary computer system. For this to be possible, the proposed $|0\rangle$ and $|1\rangle$ states
of a physical system must be distinct in energy in order to be preparable,
1.2.0

Figure 1.3: A spin qubit can be initialised to the state \( |1 \rangle \) or \( |0 \rangle \), or an arbitrary superposition of both. It is the entanglement between spins that enables universal quantum logic operations where \( n \) qubits provide \( 2^n \) entangled states.

and must be reasonably long lived in order to undergo operations. For spin qubits [112, 113], the state distinctions may be the Zeeman split spin-up spin-down states of a single spin or the singlet-triplet splitting of two coupled spins. The lifetime of a spin qubit depends on its interaction with the environment. Referred to as the relaxation time, \( T_1 \) measures the time before the higher energy state relaxes to the ground state which constitutes a loss in energy and classical information. \( T_2 \) measures the coherence of an arbitrary superposition on the Bloch sphere before the phase information is lost. Though theoretically \( T_2 \) can be as long as twice \( T_1 \), in practice \( T_2 \) is much shorter than \( T_1 \) and so the literature is dominated by measurements of \( T_2 \) coherence times in comparison with operation times [17, 114], which is one measure of the viability of a qubit. The general consensus is is that operation speed should be \( 10^4 \) times faster than decoherence in order to apply error correction protocols. Optical proposals that proffer the polarisation of photons as distinct qubit states have the advantage here since photons do not suffer appreciable decoherence.

The second indication of the feasibility of a qubit candidate is whether it can
be employed for universal logic operations. By 1995 it had been determined that any logic operation is composed of two basic operations, an arbitrary qubit rotation and a two-qubit CNOT gate [115]. Therefore demonstration of these two functions are key to implementing a quantum computer, including the experimental practicalities involved.

The remaining criteria for physical realisation arise from the need for a practical quantum computer architecture [104]. In the literature this is conveyed by the term ‘scalable’, meaning that a small set of qubits can be scaled to the many which are required for any useful simulation, including some form of error correction [116–118] to make the system fault tolerant. The scalability of a system often hinges on a coherent transport mechanism for conveying quantum information between different sections of an architecture, but is fundamentally dependent on the ability to literally build the system.

### 1.2.1 QC schemes

We will now summarise the main strategies for physically implementing a quantum computer, mentioning their relative strengths and weaknesses along with research progress to date.

**NMR manipulation of spins**

The control of nuclear spin states goes back to the discovery of nuclear magnetic resonance by Isidor Rabi in 1938 [121]. The application of resonant radio frequency (rf) fields, induces transitions between Zeeman split nuclear spin states. The resonance at which the spin states flip is dependent on the magnetic field present via the gyromagnetic ratio. The gyromagnetic ratio of different elements differ slightly due to the different number of proton magnetic moments present in the nucleus and the effect of the electron spin cloud, enabling
magnetic resonant imaging (MRI) of selective atoms in living tissue [122]. It is no surprise then, that one of the first quantum computing schemes was devised on the NMR control of molecular spins. In 1997, the nuclear spin states of molecules suspended in liquid were proposed as qubits [123, 124], and both Shor [119, 125] and Grover’s [126] algorithms were demonstrated soon after. Each nuclear spin in a molecule is a potential qubit and each of the $10^{18}$ molecules forms an identical processing unit. The ensemble is controlled globally by a radio frequency (RF) field coil using standard NMR techniques since the resonant frequency of each nucleus differs due to variations in the local charge distributions within the molecule and each nucleus can be individually addressed. Though coherence times are more than adequate [127] because the nuclear spins are shielded by the local electron clouds, only a few dozen qubits could ever be exploited in such a system because of temperature fluctuations [128]. Liquid NMR schemes have been abandoned except as a proving ground [129] where they have been very useful for testing the viability of spin logic operations [130]. Instead, researchers in NMR QC have turned to solid state schemes in order to facilitate a large scale architecture. Si has been the most widely exploited semiconductor in the least century owing to its relative
abundance and resistance to high temperatures within transistor channels so it is no surprise that several QC schemes hinge on the sophisticated fabrication techniques available for constructing Si devices. In addition, the theoretical understanding of Si is well advanced. Common intrinsic Si contains 92.23% $^{28}\text{Si}$ and 3.1% $^{30}\text{Si}$ each with a nuclear spin of 0, while the remaining 4.67% is $^{29}\text{Si}$ with a nuclear spin $\frac{1}{2}$. In most silicon QC schemes, $^{29}\text{Si}$ is a hindrance that hastens decoherence [131, 132] but the solid state NMR strategy exploits the nuclear spin of $^{29}\text{Si}$ atoms as a qubit [120, 133] which have demonstrated $T_2$ times exceeding minutes [134]. Atoms are placed in rows on a stepped $^{28}\text{Si}$ structure to form collections of qubits analogous to the suspended molecules in liquid NMR. A nearby magnet provides a gradient of field along the row that causes the resonance of each atom to differ for universal logic operations. In this scheme it may be necessary to employ the sensitive magnetic resonance force microscopy [135] which utilises the sensitivity of the magnetic force microscopy technique by putting the sample on a flexible structure that can move when the spins are aligned at resonance, just as the cantilever of a magnetised atomic force microscopy probe deflects over magnetised samples. A single spin can be measured with MRFM [136], but while the Si NMR scheme provides an ensemble of identical spin processors [114], such structures are yet to demonstrate any qubit operations.

**Trapped ions and atoms**

The trapped ion scheme was first proposed in 1995 [139] to take advantage of previous research undertaken in the 1980s, on manipulating laser cooled atoms in traps [140–142]. The system suggested optical transitions (ground state to metastable exited state) of laser cooled ions as qubits, held in place by electric fields and the coupling between particles produced by inducing collective harmonic motion with further application of lasers. A CNOT operation was
Figure 1.5: (a) Ions are trapped within an electric field and lasers induce harmonic oscillator potentials, from [137]. (b) Neutral atoms on the other hand are trapped directly into an optical lattice, distributed according to temperature, from [138].

...
herent coupling of trapped atoms to photons \[153\] can serve as a readout mechanism \[154\] or herald the long distance teleportation of atomic states \[155\], essential for building quantum networks.

**Optical Strategies**

The photon is a natural qubit candidate since it has few degrees of freedom and thus is relatively free of the decoherence that is suffered by spin systems. However, this also means that photons are weakly interacting without the non-linearities afforded by QED cavities \[156, 157\]. Even so, all optical quantum computing is possible in principle \[158\] with single photon sources and detectors. Demonstrations of quantum algorithms have shown promise \[14, 159–161\], particularly the optical demonstrations on specially designed waveguide chips \[162\] that suggest a route to scale up. The real strength of photon based systems however, is in secure long range quantum communications including quantum key distribution \[163, 164\].

For this reason, much of the optical QC literature involves interfacing with other qubit schemes either directly \[153, 165\] or indirectly via QED cavity structures \[166–168\]. QED cavity QC systems are widely researched as well \[169, 170\].

**Superconducting Qubits**

Superconducting qubits \[16\] are based on the use of Josephson junctions within circuits and come in three types, charge \[171\], flux \[172\] and phase \[174\]. Qubits and qubit interaction are manipulated by microwave pulses delivered via selective on-chip electrodes \[175, 176\]. Though coherence times of several \(\mu s\) are 10-100 times longer than initialisation, operation and readout times, which has allowed simple quantum logic demonstrations \[177, 178\] the sources
1.2.1

Figure 1.6: Superconducting qubit circuits including (a) a charge qubit circuit (aka a cooper-pair box) from [171], (b) a flux qubit circuit from [172] and (c) a circuit that was used to entangle 3 phase qubits from [173].

of decoherence are not well understood and will need to be addressed in order to scale the system. Current progress in qubit operations has been supported by the integration of QED cavity spin bus architectures [179] as long range mediation across superconducting circuits.

Donors and dots: spin qubits

The trapped ion scheme and optical qubit scheme are both trying to implement solid state, on chip systems in order to mesh with current fabrication technologies. QC schemes that have operated with this advantage from the beginning include the use of donors [18,180] or quantum dots [181–183] within
solid state architectures that deliver spin based qubits via the Zeeman split states of a single spin or the singlet-triplet spin splitting of a biased double dot [184]. A donor within a semiconductor lattice provides quantised energy levels within a sharp atomic potential well for electrons to occupy, such as the valence electron level of a P atom within a Si lattice [185, 185, 186]. A quantum dot is an artificially constructed potential well, also with quantised energy levels available for electrons which depend mostly on its size [187]. It is considered analogous to electrons bound to a donor nucleus, though a quantum dot can accommodate many 10-100s of electrons and a donor will hold maybe one more than is natural for the atom.

Metallic quantum dots [188] which were among the first to be recognised, are not inherently useful for quantum computing since they are usually highly disordered and are difficult or impossible to tune down to the single few electron regime [189–192]. Another form of physically defined quantum dots can be formed by imbedding self-assembled, nano-scaled single crystals in materials with a higher bandgap [187, 193] to eliminate any surface states, such as the InAs/GaAs system that is popular because it provides optically addressable electron spin [194] or charge [195] qubits. Physically defined quantum dot structures are seen in figure 1.7.

![Quantum dot structures](image)

**Figure 1.7:** Quantum dot structures, physically defined. In (a) the dot is defined by a region of heavy P doping from [192], (b) with QED cavity coupling from [167] and in (c) the dot is defined by buried heterostructures from [196].

Electrostatically defined quantum dots that are formed in 2-dimensional elec-
tron gases (2DEG) are favourable because they are highly tunable rather than having couplings largely predetermined during fabrication. The literature on electrostatically defined quantum dot QC demonstrations initially contained predominately GaAs structures [197–199] owing to the high mobility of the material and readily available GaAs/AlGaAs heterostructures, but Si/SiGe dot technology is emerging and improving upon the GaAs results [200, 200, 201]. Such devices can be seen in figure 1.8(a) and (b) where conduction is observed via quantum point contacts (QPC) [202] or a resonant circuit [203]. Another way of electrostatically defining dots are with metal-oxide-semiconductor (MOS) structures [204–206] where the conduction through the dot is observed directly via source/drain leads seen in figure 1.8 (c).

![Figure 1.8](image.png)

**Figure 1.8**: Quantum dot structures, in (a) GaAs/AlGaAs from [197], (b) Si/SiGe from [207], (c) Si MOS from [208], electrostatically defined by the surface electrodes shown.

Double [201, 208, 209] and triple dot [198, 199, 210] devices have been demonstrated, controlled to single electron occupation [206]. Spin dependent transport allows the determination of spin in a single shot readout process for both GaAs dots [197, 211, 212] and Si dots [213, 214]. GaAs coherence times are shorter than in Si due to the spin moment 3/2 of GaAs nuclei [215]. The fluctuations in the magnetic forces exerted on the qubit charge cause by this background spin field can be alleviated by the application of small magnetic fields [215, 216] but spin diffusion can still occur [217] and GaAs single spin coherence times remain in the µs [212]. The analogous Si/SiGe systems have
demonstrated $T_1 = 2.8 + 0.3 \text{s}$ [213] and though they exhibit $T_2$ at $3 \mu\text{s}$ [218], certain triplet states can be longer lived [214], which is also seen in GaAs [219].

Donor schemes have the advantage of even longer coherence times due to the sharp potential that confines the electron in all three dimensions. Donors under consideration include P [18,180,220–222], Bi [223–228], As [20,229,230], Sb [231,232] or even Er, due to its optical properties [227,233,234]. The nuclear spin of zero of $^{28}\text{Si}$ makes it a nice clean medium for long coherence times as long as the presence $^{29}\text{Si}$ is sufficiently small [132], meaning that enriching will be necessary as well as purification to avoid spin diffusion via other spins [217]. $T_1$ can be hours even at room temperature [235,236] in bulk Si and ms-s in devices [213,237] while $T_2$ was recently measured exceeding 10s at 1.8 K in enriched $^{28}\text{Si}$ [238]. Nuclear spin coherence is more difficult to measure due to its smaller magnetic moment and utilising the electron in the measurement shortens $T_2$ [239], compared to the coherence of ionised P donors [240,241]. The decoherence lengths of Bi in Si are similar to P times since they are affected by the same mechanisms [223], despite the nuclear spin of Bi being $3/2$. The spin $1/2$ of both the P nucleus and donor electron make it the simplest donor system and scalable Si:P architectures have been envisioned [18,180,220].

While the original Kane scheme presented the nuclear spin as the qubit since the state is longer lived due to its smaller magnetic moment, the electron spin is easier and faster to control due to its larger magnetic moment. The current strategy is to exploit the P donor electron as a spin qubit while the longer lived nuclear spin can be exploited as a quantum memory via microwave transferral [242], an avenue unavailable to 2DEG quantum dots or photons alone. There are however donor-dot architectures [183,229] and photon-atom mechanisms [243,244] that can provide atomic memory to these systems. The study of donor states within silicon is well advanced, having been established in the 1950s in parallel with semiconductor fabrication technology [185,186,245], particularly for P which is the main n-type donor used in the electronics
industry. Certain P donor states in Si can be optically pumped [246] but state manipulations required in QC devices are electrostatically controlled [247,248]. Si:P devices can be fabricated via top-down ion implantation [249] or bottom-up STM [250] methods which produce similar results (cf [251] and [192]). Single qubits have been realised, in both the donor electron spin [252] and the nuclear spin [241] of $^{31}\text{P}$ in Si. The qubits are initialised and controlled with radio-frequency pulses delivered by sophisticated on-chip transmission lines [253] and the single shot readout of a single electron spin is achieved [21] with the same spin to charge conversion used in dot measurements. The potentials are tuned so that the Zeeman splitting prevents an electron in the lower spin state from tunneling to the electrostatic dot formed by the measurement SET. The spin of the nuclear qubit is measured the same way by first coupling it to the electron spin.

While the wavefunction overlap of adjacent quantum dots can be large for relatively easy entanglement, P donors have yet to be individually specifically entangled. However naturally formed cluster states have been observed in photoluminescence [254] and EDMR measurements [240] that could be exploited in a global cluster scheme where an array of donors is addressed rather than individual spins [222]. Entanglement has also been deliberately induced in an ensemble of donors [255] but the observation of controlled exchange between two individual donors is yet to be realised. All of the single or few atom experiments have been performed on devices with serendipitous placement of donors, whether they were implanted directly [21, 251] or expected from device lead diffusion during fabrication [20, 256, 257]. In order to demonstrate operations with more complexity we will require deterministic doping to accurately and knowingly place single dopants within semiconductors.

Other dopant:semiconductor systems are possible [7], the most successful of which is the use of the nitrogen vacancy centre (NV$^-$) in diamond [258–260]
which has many of the same qualities of the Si:P scheme with the added convenience of faster optical manipulation and room temperature operation. While diamond devices can be more difficult to engineer than Si due to its hardness and NV\(^{-}\) centres are much more complicated to produce [261] than the simple implantation of P in Si, the optical properties of the NV\(^{-}\) centre are better suited to long range transport or communications [165]. The carbon atom has zero nuclear spin and the diamond lattice provides a very large bandgap for long coherence times, with \(T_1\) surpassing seconds and \(T_2 = 2\) ms, in 99.7% pure diamond at room temperature [262]. Memory storage has been demonstrated with a spin transfer from the electron to the nitrogen nucleus of the NV\(^{-}\) centre [263] analogous to the Si:P system. Room temperature spin readout [264, 265] and entanglement [266] have been demonstrated, controlled and observed optically. Actual qubit operations require many carefully placed NV\(^{-}\) centres, just as further experimental progress within the Si:P system now requires deterministic doping.

Hybrid architectures are emerging to take advantage of the relative strengths of donor and dot systems. One such proposal suggests shuttling electron qubits from dots to donors [229, 267] to combine the ease of fabrication and operation of quantum dots with donor memory storage. Mixed donor schemes also exist such as a S:Bi system where the Bi atoms are optically addressed via Er donors [227] and the all-silicon solid state NMR scheme includes the use of P donors to mediate efficient initialisation of the \(^{29}\)Si nuclear spin [114]. What the strategies have in common - single, well placed donors - requires deterministic doping to fabricate. One of the most pressing reasons for introducing deterministic doping now is for the demonstration of transport mechanisms in order to show that fault-tolerant donor architectures are scalable.
1.2.2 Coherent Transport

Transport is often the key to demonstrating the scalability of QC systems. Having made great progress in fundamental qubit demonstrations in the QC strategies we have just outlined, researchers are now seeking coherent means of transport for quantum information, within and between various sections of information processing architectures, necessary for large scale operation. The term ‘quantum wire’ is sometimes used to describe such a pathway but this can be misleading as it was originally used to define a conventional, one-dimensional wire structure so small that electron conducting states become quantised within it; an effect which we frequently see in the current spectroscopy of QC MOS devices [20,251] due to quantisation in the nm scaled electrodes. Proposals for charge qubit computation with mobile electrons in electron waveguide wires have been made [268–270] analogous to photon waveguide systems and although greater electron-electron interaction might make for easier entanglement than in exclusively optical systems, it also puts electrons at a disadvantage by causing fast decoherence. The coherence of delocalised electrons with all three motional degrees of freedom, is much shorter than electrons confined in 1 or 2 dimensions [271–274] but electron waveguides in high mobility GaAs are possible [275]. In general however, conventional conduction is unavailable for spin information processing architectures. While coherent conduction band transport is potentially possible in Si [276], the ‘hot electron’ temperatures required (>60 K) are incompatible with dot and donor systems which operate at 4 K and mK respectively. Further to this, in such a conventional wire, there is no reason to expect that the electron emerging from the final end is the exact same electron that went into it. For the transfer of quantum information, we need a guarantee that the final electron, if not actually the exact same initial electron, is entangled with the initial electron in such a way as to have the exact same spin state [277]. Photons are a natural candi-
date for long range transport of coherent information and offer the promise of secure communication with quantum key distribution (QKD) [163, 164]. For this reason, research groups working on different QC schemes have published results on coupling the state of their qubit of choice to a photon either directly [153, 165] or indirectly via QED cavities [166–168]. The flying photon qubit can then couple back to a solid state system, which makes it a mediators for long distance teleportation of spin states between trapped atoms [155]. For long range transport within an architecture, it would be more convenient to use solid state solutions. For transport within dot or donor systems such as Si:P there are several options. Information is coherently transferred from one donor to the next during a swap operation [278] and while repeated swap operations could transfer the electron down a chain of donors (depicted in figure 1.9(a)) this is a relatively slow, and thus short range solution. In a spin bus architecture, long range qubit exchange is affected by an intervening chain of interacting spins [279–281] which can be exploited for rapid transport [282] (seen in figure 1.9(b)).

![Diagram](image)

**Figure 1.9:** (a) [277] CNOT swapping along a chain of donors (b) coherent tunneling by adiabatic passage [22] (c) spin bus chains for the mediation of exchange between quantum dots [280]

For compatibility with electron spin Si:P donor architecture [180], there is the
coherent transfer by adiabatic passage (CTAP) mechanism [22]. At the same time the CTAP mechanism was published, a similar proposal was made for ion trap systems [283], which has now been realised in a three well system [284]. CTAP transports the qubit electron along a chain of donors though unlike transport methods which pass along the quantum state via photon or spin entanglement, there is no such mediation here; the electron is physically transported through the material along a null state of the donor system. This is also the case in the digital Kane approach [220] where electrons would be transported along a H terminated Si interface in a process akin to SWAP shuttling. The adiabatic techniques described in [183] for the transport of information within a quantum dot architecture, also physically transport electrons within a travelling wave potential created by adiabatically evolving voltage pulses from gate electrodes. This is effectively creating mobile quantum dots where the coupling between multiple adjacent dots moving in parallel can be preserved. CTAP involves adiabatic, counter intuitive pulse sequences to tunnel an electron which is localised around the first donor, to the last donor in the chain without significant interaction with any of the intervening donors, for preservation of the spin state coherence. It forms a solid state analogy to the well known stimulated Raman adiabatic passage (STIRAP) technique which allows the transfer of excitations across an otherwise forbidden transition via a dark state [285, 286] for efficient optical pumping. CTAP has an advantage over the adiabatic travelling wave technique of [183] when the straddling [22] mechanism is used for long range transport because the required gate density is greatly reduced which limits the scalability of other transport techniques [277]. While STIRAP is affected by the application of radiation, the CTAP transfer of electrons in Si:P is modulated by gate voltages [22] and the transfer of trapped ions is modulated by the proximity of the adjacent wells [283]. The CTAP mechanism has been demonstrated in optical [287] and ion trap systems [284] but a solid state demonstration with a device shown in figure 1.9(c)
is required to demonstrate the scalability of the Si:P system.

Reasonably strict spacing is required between adjacent donors for a wave-function overlap that allows exchange to be manipulated with feasible surface gate voltages [247, 288]. This spacing is the same for a two donor device needed to demonstrate entanglement during a CNOT operation and the yield of such devices could be improved with deterministic doping when compared to timed implants through a single aperture. Deterministic doping would also be required to obtain the correct fixed exchange between adjacent spin bus dopants [280] and between NV$^-$ centres in diamond [289].

1.3 Deterministic Doping

Donor architectures are among the most challenging solid state systems to study, due in great part to the near atomic precision required in placing the donors and alignment of the control gates. Larger systems that use electrostatically defined quantum dots or superconducting circuits have relaxed electrode scales, fewer alignment steps and are easier to observe electrically. Despite the initial fabrication and other experimental hurdles, the SiP strategy has produced results demonstrating single P donor electron spin rotations [252], spin-to-charge readout [21] and shows great promise as a scalable system [180].

In order to establish that logic operations are possible, we are now required to exhibit a two-qubit CNOT operation [115] which requires controlled entanglement between two adjacent donors.

The importance of deterministic doping is clearly noted in the International Technology Roadmap for Semiconductors (ITRS) [290] which deals with the progression of classical electronics. The position of very few atoms has been shown to significantly effect the behaviour of a transistor [23, 291, 292] which dissuades continued down sizing of classical transistors without more accu-
rate ion implantation. For researchers attempting to demonstrate donor based quantum computing mechanisms, the interest is in increasing the yield of results from devices which would otherwise be made with less certain dopant placement methods such as timed P implants [251] or via predicted diffusion of As from the highly doped leads into the finFET channel [20].

Figure 1.10: P donors can be accurately placed on a Si surface by removing atoms from the H-terminated (001) surface with an STM probe tip and introducing phosphene gas under specific conditions.

The most accurate method for placing a phosphorous donor in silicon involves the use of STM probe tips for nm precision patterning [250], developed by Michelle Simmon’s group at the University of New South Wales and is depicted in figure 1.10. It begins with an epitaxially grown (001) $^{28}$Si substrate that is H-terminated. A very fine STM probe tip is brought to the surface to scratch off hydrogen atoms and where three atoms have been removed, Si dimers are available to bind with the phosphorous of incoming molecules of PH$_3$ gas. The number of P ions incorporated in the target region is inferred from scans of the exposed dimers before the injection of phosphene gas which means the location of each donor is known to within three specific lattice sites and has the accuracy of an atomic STM scan. A subsequent in situ anneal incorporates the phosphorous atom into the lattice. Highly doped source/drain leads and in-plane gates can be constructed in the same way to make devices
that exist on the surface plane of the Si substrate \[192, 293\]. This fabrication method and the devices produced with it have been thoroughly characterised and STM Si:P few atom devices produce similar results to Si:P implanted devices (cf \[251\] and \[192\]). Currently, it has only been optimised for Si:P, though other materials are of course possible \[7, 294\]. This approach and others like it are known as bottom-up methods.

The other approach to deterministic doping is top-down, where dopant atoms are implanted into devices by directing accelerated ions at masked targets. The depth at which the ion comes to rest depends on the species of ion, its acceleration energy and properties of the target substrate. An ion entering a substrate will be slowed as it ionises and damages its surroundings in collisions with the lattice and comes to rest when all its kinetic energy has been given to the substrate. Thus, in general, more energetic ions travel further into the substrate as do lighter atoms. Strongly ionised atoms will reach a shallower depth than singly ionised atoms because the extra charge lends them larger interaction cross sections. The physics of ion-solid interaction will be discussed in chapter 2.1. Unlike the STM bottom up methods, many species of semiconductor dopant can be accelerated and implanted to a usable depth. In order for ion implantation to be deterministic, the ions need to be accurately placed as well as counted so that the exact number of ions entering the semiconductor substrate is known. An accelerated ion incident on a substrate will cause electrons (and in some cases atoms) to be ejected from the surface, create electron-hole pairs as it collides with the lattice and leave behind it a trail of damage. These by-products can be exploited to detect a single ion entering a substrate. The secondary emission from a substrate caused by ions with the low energies required in quantum device applications, is generally too small to be a reliable indicator \[295, 296\] though the emission can be enhanced by using highly charged ions or carefully biased substrates \[297\]. In theory secondary electron detection could be used for both bulk and SOI samples.
Figure 1.11: An accelerated ion entering a substrate will cause secondary electron emission and ionisation within the material; each can be exploited for detection. Ion beams for implantation can be focussed or collimated.

The ionization of the substrate by even a single ion, is large enough to detect if we specially fabricate the substrate into a PiN diode [249]. In conjunction with a JFET charge-to-voltage converter, both cooled below 120 K with LN₂, the noise of the system is sufficiently small so that the ionisation produced by a single ion can be observed. This method is applicable to bulk Si samples only but can be applied to any accelerated ion species.

The competing method of detection exploits thin layer silicon on insulator (SOI) channels rather than bulk Si. The small Si channel is electrostatically affected by the damage induced in the channel (or its immediate surroundings) by the passage of an ion [298–300]. This is exhibited as a permanent change in the source drain current. This method is appealing because it can be used at room temperature, unlike the e-h pair detection method which requires lengthy liquid nitrogen cooling. However, the SOI channel needs to be small for the signal to be significant, so the useful area is much smaller when compared to ionization detectors which can encompass more than \((45 \mu m)^2\) active area. While the ionisation detection signal is well understood and predictable, the mechanisms involved in the SOI method have not been clearly characterised.
For the positioning of the donors, the ion beam can be collimated with an aperture or focussed with magnetic lensing, seen in figure 1.11. Schenkel’s team employs a scanned AFM cantilever with a hole drilled through the tip which can be used for both collimation and position registration when used in a scanning microscopy capacity. The resolution of this method is limited by the straggler of the ion species, the resolution of the AFM system and the size of the collimator. Reports typically contain apertures of 80 nm [299] or tapered holes with 30 nm exits [301] but holes as small as 4 nm have been constructed and would be required for the probe tip to remain small for good imaging resolution [296]. The placement accuracy of focussed ion beam implantation currently stands at 60 nm [23] which is enough to produce MOS channel modification [23, 292, 299] with deterministic dopant placement. Though such low energy ion beams can be focussed to a few nm, it is in positioning the beam that the accuracy becomes limited. A limited number of species can form a liquid metal, focussed ion beam source (LMIS) for low energy implants. The collimation method can handle various species without alteration, as long as the aperture is thicker than the range of the particular ion. Either method is compatible with both ionization or SOI channel detection.

A new ion implantation method that promises to be more accurate than these is the use of a magneto-optical ion trap to produce a cold ion source [302], currently under development by Jan Meijer et al for the creation of NV$^-$ centres in diamond. Unlike the previous schemes, the ion counting stage occurs when an ion is registered in the trap, not when or after it hits the substrate which allows deterministic doping of diamond since the electrical characteristics of diamond otherwise preclude it from the ionisation method [261]. Preliminary experiments are promising [303,304] demonstrating 97% efficient deterministic loading of 2 mK single ions which are ejected from the trap at only 80 eV. Researchers are now working on increasing the brightness of the source [305] and integrating it with focussing optics [303] and acceleration apparatus for the
creation of -NV centres in diamond [261]. This deterministic doping technique is essentially a focused ion method that aims to exceed the resolution of the work by Shinada et al by developing a new cold ion source and optical system. The ion trap source is also able to accommodate ions that cannot form a typical LMIS.

The bottom-up STM method, being the most accurate, is clearly suited to large scale fabrication where the error tolerances are increasingly small. However, the slow throughput of this direct write method limits its usefulness in quick prototyping and proof of concept devices and flexibility in dopant species has not yet been established. Additionally, the method currently only produces single plane devices [192, 306, 307] which limits the possible gate density. Ion implantation methods, though inherently less accurate due to ion straggle, are capable of implanting many ion species into various target materials and are much faster due to compatibility with standard lithographic techniques for making surface device electrodes.

In order to continue experimenting on proof-of-concept, counted atom devices while other techniques are developed, we have implemented a deterministic implantation strategy that exploits existing technology. Our deterministic doping method consists of three main components that form the chapters of this thesis; a single ion detector, a high aspect ratio nanostencil and the ability to position the stencil accurately above the detector. A schematic of the deterministic doping apparatus is shown in figure 1.12.

Low energy ions are collimated by a positionable aperture which can be stepped across the implant site where the dopants entering the Si substrate are counted by the ionization detection system. In this way single atoms can be implanted precisely for the fabrication of donor arrays. The stack of three piezo positioners are capable of nm precision movement and control the xyz position of a nanoscaled aperture which we deploy over an ionisation detector. The detector
and the charge sensitive JFET behind it are cooled to cryogenic temperatures with LN$_2$ cooled N$_2$. The concept is shown in figure 1.13.

The nanoaperture in the Si$_3$N$_4$ membrane is aligned to the implant site with a microscope ex situ to an accuracy of 5 µm. With the detector cooled, active, in the target chamber and under vacuum, a series of EBL defined slots in the PMMA covering the active region can be used to align the aperture to the EBL defined implant window with nm precision. The nanostencil can be stepped across to the next implant site and the implant repeated to form a one dimensional array of single donors. For a two dimensional array, a circular aperture is used.
In chapter 3 we explain the detection of single 14 keV P$^+$ ion implantation and based on rigorous testing we present remarks concerning the main contributors to detector performance. In chapter 4 we describe the fabrication of a nanoscaled aperture via focussed ion beam techniques. Chapter 5 demonstrates the alignment of a nanostencil to a single ion detector for the implementation of deterministic doping.

The potential of the deterministic doping apparatus goes beyond the fabrication of single atom arrays. The ion beam modification of materials has a long history and a wide range of applications besides semiconductor doping [308,309]. Low energy ion implantation can be used to nucleate the formation of semiconductor nanowires [310], quantum dots [311], or nanorods [312]. Implanting metals into semiconductors can create buried metal nanocrystals for plasmonics [313] or conductive thin films [314]. While high energy ions can be required to modify the refractive index of diamond for optical appli-
cations [315], low energy, heavy or highly charged ions can induce the graphiti-
tisation of diamond [316], or conversely the crystallisation of graphite into
diamond [317]. In general only one ion dose is delivered such to samples with-
out intervening mask changes. The ability to implant precise ion fluences and
a varied distribution of dose in one step could facilitate further nanotechnol-
ogy research whether for novel nanoelectronic device design or medicinal drug
delivery [318].

Other semiconductor applications include quantum-well intermixing induced
by the implantation of 18-180 keV P or As into InP based laser diodes [319].
The implantation step carried out after the growth of the heterostructure,
softens the junctions by intermixing the heterostructure species and blueshifts
the laser luminescence. Graded junctions are useful in MOS semiconductor
fabrication as well [320,321]; the threshold voltage is determined by the channel
doping near the source junction in particular [320] and the magnitude of the
current through the channel depends on the doping at the drain side [291].
These results were obtained from devices with a one-step gradient in dopant
concentration. Deterministic doping would allow experiments on channels with
a dopant concentration of controllable gradient for engineered band bending.

The difference in performance of channels with ordered and disordered arrays
of dopants has been clearly demonstrated at a pitch of 100 nm at room tem-
perature [23], seen in figure 1.14(a) and a 10 nm pitch at mK shown in figure
1.14(b) [322]. The ordered array provides less variation on the Voltage thresh-
hold of the MOSFET channel at a relatively large scale. On a mesoscopic scale,
we can see that the ordered array provides an even potential across the channel
which requires a certain bias to be traversible. The disordered array however
provides ready potentials for percolating currents to move through the effec-
tively higher concentration of donors at a lower bias voltage. Creating ordered
and disordered arrays could be used to characterise a deterministic doping
Figure 1.14: (a) shows the difference in threshold variation between ordered (O) and disordered (D) channels in a transistor at a pitch of 100 nm, from [23]. (b) shows the threshold voltage required in a 10 nm pitch ordered array (O), compared to the immediately ohmic behaviour of a disordered array (D), from [322].

System since below a certain pitch, we would expect little to no difference between the ordered and disordered array. When device size is very small, the particular configuration of single dopants is expected to have an effect on device performance because of the anisotropy of both the semiconductor and the dimensions of the structure. For example, the deterministic doping apparatus could be used to produce the arrays seen in figure 1.15. The rectangles have dopants with the same pitch and the same average cm$^{-2}$ concentration, but alternate distribution. On a large scale the differences may not be noticeable but there could be a marked difference on the few atom scale, particularly if the current becomes quantised through the various donor levels.
Figure 1.15: The structures made possible by deterministic doping. In addition to single ion arrays of varying pitch and orientation, specific dopant concentration gradients can be fabricated.

However, we now concentrate on the specific goal of fabricating a three donor array which can be used to demonstrate the CTAP protocol. Chapter 2 begins with the accepted physical models for the interaction of ions with matter during low energy ion implantation which we will use to model the fabrication of three donor arrays. Using hydrogenic modelling for the CTAP protocol, we estimate the effectiveness of deterministic ion implantation by the fraction of arrays that can be reasonably used to demonstrate CTAP.
Chapter 2

Modelling Ion Implantation

In order to assess the potential of deterministic doping, we require a model for ion implantation. The models and the Monte-Carlo simulations that take advantage of them are well advanced. In this chapter we present the accepted physical description of the process and the results of such a model when employed by Monte-Carlo simulators in a binary approximation. The following section contains a review of CTAP. Within the hydrogenic approximation, we estimate the donor positions required for a timely CTAP operation. With this, and device limits on donor positions, we use the ion implantation modelling to estimate the effectiveness of deterministic doping in fabricating a three donor array for the demonstration of CTAP.

2.1 Ion-solid interaction

The interaction of accelerated particles with matter has been studied closely for more than a century, even before Geiger and Marsden first directed an alpha source of a few MeV at a substrate of gold [323] and found that an atom consists of a small positive nucleus with even smaller electrons orbiting around it [324]. Since then, scientists have been purpose building accelerators, achieving higher
energies to peer deeper into the structure of matter [325]. There is no doubt that the creation of exotic particles is a fascinating science, not least due to the immense machines used to facilitate such minuscule events and this means that the physics of high energy collisions and ion-solid interaction has been studied in great detail including relativistic effects, particle creation/annihilation and particle tracking through millions of individual detector cells [326]. A high level of understanding of the physics that occurs within the detectors is necessary in the pursuit of further refinement of the standard model, like the widely publicised search for the Higgs boson [327]. Much of the initial technology used for ion implantation was adapted from high energy particle accelerator physics [72] but the physical models of ion-solid interaction are very different. The physical models for the very low energy scale, where the work in this thesis has been carried out, are less complicated because they are without relativistic effects or nuclear reactions. However, because the initial ion energy is so low, a single collision holds greater significance and this makes it difficult to obtain both data and accurate approximations for modelling.

Though we can simulate the trajectory of a single ion through some material, it is more instructive to observe the trends seen when we consider the behaviour of many ions. As such, the literature on low energy ion-solid interactions contains mostly phenomenological descriptions of the physics, tailored to suit specific purposes. For instance, a semiconductor manufacturer might like to know what energy a particular ion should be given in order to reach a desired depth in a specific material. The best way to use the various physical models developed is in Monte-Carlo simulations.

One of the most widely used tools used to evaluate properties of ion implantation is the freely available [328] ‘Stopping and Range of Ions in Matter’ (SRIM) package [329] which is based on ZBL theory for ion-target interaction [330], informed by extensive experimental databases. The strength of the program lies
in its comprehensive range of ion-target interactions which includes all atomic elements with energies from 10 keV to 2 GeV in over 150 target compounds.

Another widely used simulation system is the Geant4 [331–333] toolkit, also freely available on the internet [334]. It is a collaborative work developed at CERN; the major publication [331] has over 70 authors. It has been designed to cover a vast breadth of physical interactions with matter. Its accuracy and applicability in the low, sub-0.1 MeV range has been a recent development [335, 336] since it was initially developed to simulate high energy interactions with detection systems. It has since broadened so far as to cover ion/particle interactions with biological matter [337]. Its versatility comes from being a package of physical process subroutines that can be selected and applied to any target geometry, unlike SRIM which uses layered, infinite plane targets only. The Geant4 toolkit offers the choice of using code and data based on the Ziegler, ICRU or its own Penelope model.

The other prominent authority on the stopping power of ions in solids in the literature is the 1993 Report 49 of the International Commission on Radiation Units and measurements (ICRU) [338]. This report covers stopping tables for electrons, protons and alpha particles in 73 materials and covers the methods used to generate them. The implementation of this report is manifest in programs estar, pstar and astar [339] which generate stopping tables for electrons, protons and alphas respectively. But this project covers the implantation of heavy ions (Z > 2) so only geant4 and SRIM are now considered.

A database of compared stopping tables for H⁺, He²⁺ and 37 heavy ions between lithium and uranium in a useful range of target materials is available at [340] maintained by H. Paul. Extensive comparisons of available models and stopping tables have been made by Paul et al. [340–342]. In 2003 [341] SRIM and Paul’s own MSTAR were judged to generate stopping tables that showed the best fit to experimental data overall in the energy range 1 keV –
Figure 2.1: A comparison of SRIM modelling to Geant4 modelling in the low energy heavy ion regime. Reprinted from [335], Copyright (2005), with permission from Elsevier.

1 GeV. This study included Geant4.

The three have been compared in [343] and found to give similar results when the entire package range is taken account. For low energy ions in particular, for 10 – 20 keV B implanted into Si, the Geant4 range was a few nm more than that calculated by SRIM while the straggle a few nanometres less [335], as seen in figure 2.1. In the case of As into Si with the same energy range, the range was again a few nanometres more while the straggle was about was a few nm more than the SRIM calculations. This difference is in accord with the discrepancy Klaus Wittmaack uncovered regarding the number of atoms SRIM2000 calculates as sputtered per incident ion [344]. Where the incident mass is much larger than the target (As:Si), the estimated number of sputtered ions was fewer than experiments suggest. While projectile atoms with masses smaller than the target particle mass (B:Si) produced more sputtered ions than expected. If we consider the affect on the division of energy between the sputtered ions, the range and the straggle then the differences seem to stem from the same discrepancy in the SRIM model when compared to Geant4. Since we are employing P which is roughly equal in mass to Si, we do not expect to encounter a significant difference in the results between the two programs. Wittmaack also states that SRIM2000 electronic stopping powers for low en-
ergy ions are consistently underestimated, which leads to an overestimation of the range only. This is in conflict with our own experimental results concerning electronic stopping powers, which will be discussed in section 3.0.4. Since it is the straggle that most informs our calculations in chapter 2, this is a minor concern.

SRIM results provide the closest fit to experimental data when the entire energy range is considered so it is generally considered to be the best, despite shortcomings in the heavy ion regime [345]. The recent 2011 update to SRIM has attempted to address remaining concerns [346]. The main changes to the program concern the accuracy of the simulation in compounds based on new experimental evidence and the accuracy of heavy ions modelling in the >1 MeV/amu energy range and the results for 14 keV P into Si are unchanged.

We use the package as a guide to our experiment, which will be presented in the following section 2.3. A full description of the program can be found in [347], which we have summarised here. Most of the following theory underlies all ion implantation simulations which differ mainly in execution. In particular, the ZBL theory for the interaction potentials of nuclear collisions is the most widely applied. We will make special mention of the failings of the SRIM code as they apply to our work, and how we adjust accordingly.

A phosphorous ion that has been accelerated to 14 keV will embed itself ∼25 nm into a silicon substrate (20 nm below the interface of a 5 nm oxide). Collisions between the incident ion and particles of the target material rob the accelerated ion of its kinetic energy in a piecemeal process throughout this distance. In the sub-100 keV regime we are considering here, there are no relativistic effects or nuclear reactions occurring. Only the valence electrons of the target material can change state; the inner shell electrons of both the incident ion and the atoms in the target lattice are too tightly bound to be accessible at these low energies.
Figure 2.2: Where the electrons of the ion and the material collide the result is ionisation. An ion can remove atoms from the surface or displace them within the lattice if it imparts enough energy during a nuclear collision.

When an ion enters a substrate, target material atoms can be ejected if the impact causes enough vibration to break the surface binding energy. This is known as sputtering and is exploited in applications such as focused ion beam (FIB) fabrication that is discussed in greater detail in chapter 4. Sputtering yields calculated by SRIM predict 1.85 O atoms and 0.56 Si atoms per 14 keV P⁺ ion hitting a SiO₂ surface, which is negligible damage. The main forces at work which slow the ion are known as electronic stopping and nuclear stopping.

2.1.1 Electronic stopping

The electrons of the target material cause electronic stopping via both elastic and inelastic collisions with the incident ion. Inelastic collisions, where the substrate electron is ionised by the impact, dominate and elastic collisions are typically ignored in most ion-solid interaction descriptions.
The description of electronic stopping can be separated into five distinct regimes: light ions (H and He), low energy heavy ions (<25 keV/amu), intermediate energy heavy ions (25 - 250 keV/amu), high energy heavy ions (>250 keV/amu), and relativistic ions which will not be covered here. The electronic stopping cross-section, $S_e$, of light ions is well described by weighting the interaction function $I(\mu, \rho)$ of the ion with an electron density $\rho$ by the probability of the ion encountering that electron density in the target material:

$$S_e = \int I(\mu, \rho) (Z_1)^2 \rho \, dx^3$$

(2.1)

where $I$ is the interaction function (units eV.m$^2$) of the ion charge $Z_1$, travelling at velocity $\mu$ into a plasma of electron density $\rho$. This methodology was initially developed by Lindhard [348]. The approach is most accurate when the electron density $\rho$ is defined by the Hartree Fock description of an atom which accounts for shell densities and solid state confinement. The Hartree-Fock electron density was introduced by Wilson, Haagmark and Beirsack in 1977 [349, 350] and improves on the previous work by Lindhard who used the semi-classical Thomas-Fermi model [351]. Other classical models previously used include Lens-Jensen [352, 353], the Moliere [354] and the Bohr potential [355]. The solution to equation 2.1 assumes that the impinging hydrogen ion remains a lone proton and the charge $Z_1$ is therefore constant. Without discussing the specific results which can be found in [356]; this description agrees well with experiment for all energies of protons with only a few corrections for specific materials. This theory can be scaled to singly charged heavy ions using Brandt-Kitigawa theory [357]. In the regime where the ion energy is above 250 keV/amu, the electrons of the incident ion are travelling much faster than the electrons in the target material and the electrons in the inner levels of the incident heavy atom have no means of promotion. This means the charge state of the incident ion can be assumed constant, as in the proton case. The stopping power simply scales to a heavier ion travelling at the same
speed in the same medium. We employ 0.5-2 MeV He ions in chapter 3.0.7 for performing analysis on the single ion detectors, which fall into this category.

Within the intermediate regime where the speed of the incident electrons and the target electrons are comparable (25 - 250 keV/amu), exchanges of energy cause ionisation of both the material and the incident ion which invalidates the Lindhard description. First order perturbation theory is introduced in the Brandt-Kitagawa theory for accurate modelling of this regime. We do not use any such ions in this work so we leave the discussion elsewhere [347, 357].

For low energy heavy ions, the relatively slow moving electrons of the incident ion are less likely to promote electrons in the target material. This occurs below 25 keV/amu, which is the case for a 14 keV $^{31}\text{P}^+$ ion which travels at less than 0.5 keV/amu. Using experimental observations, it has been determined that the electronic stopping force scales with velocity rather than being dependent on the electron density of the material. There are exceptions to this, including Si which is our target material. Being a semiconductor, the band gap of Si has an effect on the interaction function and experiment has determined that the electronic stopping forces of Ge, Si and some other semiconductors, scales with velocity $v$, proportional to $v^{0.7}$. For a more sophisticated description of situation, there exists some density functional theory that gives good agreement with experimental data [358, 359], but it does so only above an atomic mass of 57, which does not help us with 14 keV P$^+$. SRIM relies on the experimental results for modelling electronic stopping in the regime below 57 amu. In these experiments, ions are directed through thin films and detected via charge collection methods to evaluate the stopping forces of different materials. Very low energy, heavy ions travel a very short distance into most materials and produce very little charge which means experiments are difficult. Besides, for these low energy heavy ions, the range and distribution is influenced heavily by the nuclear stopping.
2.1.2 Nuclear stopping

*Nuclear stopping* dominates the energy loss when an ion does not have energy to promote electrons in the material nor the ion. This occurs at the end of range of the light, high energy ions and over the entire range of heavy and low energy ions. The incident ion experiences elastic, Coulomb screened collisions with atoms in the material. These collisions can displace atoms within the lattice and the recoil of the atom will incite a cascade of further displacements and ionisation until the energy transferred in the initial collision has dissipated.

The nuclear stopping $S_n(e)$ is defined as the integration of the transferred energy $E_t$, over all interatomic distances $p$:

$$S_n(e) = \int_0^{p_{\text{max}}} E_t \, p \, dp$$

This transferred energy can be determined from simple conservation of momentum $J_c$ and energy $E_c$ principles, expressed in centre-of-mass, polar coordinates:

$$E_c = \frac{1}{2} M_c \left[ \left( \frac{dr}{dt} \right)^2 + r^2 \left( \frac{d\Theta}{dt} \right)^2 \right] + V(r)$$

$$J_c = M_c r^2 \left( \frac{d\Theta}{dt} \right)^2 = M_c V_0 p$$

Which can be used to find [347]:

$$E_t = \frac{4E_0 M_1 M_2}{(M_1 + M_2)^2} \sin^2 \frac{\Theta}{2}$$

$$\Theta = \pi - 2 \int_{r_{\text{min}}}^{r_{\text{max}}} \frac{p \, dr}{r^2 \left[ 1 - \frac{V(r)}{E_c} - \frac{p^2}{r^2} \right]}$$

So that the nuclear stopping becomes:

$$S_n(e) = 2\pi \gamma E \int_0^{p_{\text{max}}} \sin^2 \frac{\Theta}{2} \, p \, dp$$
2.1.2

given here in centre of mass, polar coordinates. The impact parameter $p$ is defined as the closest distance that the atoms approach each other. The choice of interatomic potential here is crucial to the accuracy of the model. While the interatomic potential can be found experimentally, [360], we require an analytical approach to determine potentials for all atom target combinations. The interatomic potential $V(r)$ has a number of terms contained within it:

$$V = V_{nn} + V_{en} + V_{ee} + V_k + V_a$$

(2.8)

where $V_{nn}$ is the potential arising between nuclei, $V_{en}$ is the potential energy between each nuclei and the other electron cloud distribution and $V_{ee}$ is between the respective electron distributions.

$$V_{nn} = Z_1 Z_2 e^2 / r_{12}$$

(2.9)

$$V_{ee} = e^2 \int \left[ \Psi_2(r_2) + Q_2(r_2) \right] \rho_1 \, dx^3$$

(2.10)

$$V_{en} = -Z_1 e^2 \left[ \int_{r_{12}}^{\infty} \frac{4\pi r^2 \rho_2}{r} dr + \frac{1}{r_{12}} \int_0^{r_{12}} (4\pi r^2 \rho_2) dr \right]$$

(2.11)

$$V_k = \kappa_k \int \left[ (\rho_1 + \rho_2)^{5/3} - (\rho_1^{5/3} + \rho_2^{5/3}) \right] \, dx^3$$

(2.12)

$$V_a = -\kappa_a \int \left[ (\rho_1 + \rho_2)^{4/3} - (\rho_1^{4/3} + \rho_2^{4/3}) \right] \, dx^3$$

(2.13)

$V_{nn}$, $V_{ne}$ and $V_{ee}$ are straightforward; they are derived simply from the electrostatic energy that arises from the Coulomb force between two charges. $V_k$ is a term for the increase in energy due to the Pauli exclusion principle existing where the two electron clouds overlap. We assume that the electron distribution remains constant, there is no ionisation or significant distortion of either electron cloud. It is derived by taking the Fermi energy of an electron and integrating over the electron distribution within the overlap space. $V_a$ describes the exchange energy arising in the overlap, in the Slater approximation.
Each term in the total interatomic potential requires an accurate description of the electron distribution densities, $\rho$. For a realistic result, both the target atom and the incident ion are considered to have a solid-state Hartree-Fock electron arrangement that accounts for shell structure and truncates the distribution within the interatomic spacing of the target material. For computational ease, a universal interatomic potential equation has been used from the original TRIM up until version SRIM2003 which conforms to all atomic ion-target potentials calculated with the previous equations within 5%. In the latest version SRIM2011 however, individual interatomic potentials for all elements and compounds are now included, in keeping with the processing memory available to the average 2011 user [346].

An impinging ion transfers a variable energy $E_t$ (from equation 2.5) with each collision, until it has less than the material specific displacement energy, $E_d$, remaining, when it will stop and the last of its energy will be released as phonons. $E_d$ is usually larger than the binding energy $E_b$. When more than the displacement energy, $E_d$, is imparted to the target atom, it will be displaced from its lattice site and move in the material, subject to the stopping forces of a free ion at the transferred energy $E_t$, minus $E_b$, the binding energy. This means that where the initial ion energy is much higher than this displacement energy, a cascade of collisions between so-named recoil atoms will cause further (nuclear) displacements and (electronic) ionisation. Where the transferred energy is less than $E_d$, the material atom will not leave the lattice site but release the energy as phonons as it recoils back into place.

The displaced substrate atoms create vacancy and related defects in the lattice [361] that can only be healed by heating the substrate to high temperatures [362]. For single ion implantation where the displaced atoms are relatively few, a rapid thermal anneal of 950° for 5 s repairs the lattice with minimal diffusion of the dopant atoms. Much longer anneal times are need for higher fluence
implantations, such as those required in device fabrication where the damage is extensive. Annealing is also required to incorporate the dopant atom into the host lattice so that it becomes electrically active in the potential landscape of the material.

In the nuclear stopping formalism there is no distinction between the energy regimes that categorise the electronic stopping, but it bears reminding that nuclear stopping only becomes significant where the energy has been reduced so far as to make electronic stopping unlikely - hence the unchanging electron distribution conditions on $\rho$ in the electron cloud overlap equations for $V_a$ and $V_k$. This means that high energy ions cause lattice displacements only at the end of their range, while ions such as 14keV P$^+$ will be subject to nuclear stopping from the moment they enter the substrate. Meanwhile, it is the ionisation created within the substrate that we exploit to detect an incident ion.

2.1.3 Binary Monte-Carlo simulation

In this work we employ two ion beams that are in distinct energy regimes. For characterisation of charge collection, a 500 keV He$^+$ microprobe is used which is an energetic light ion (250 keV/amu) (Such an experiment will feature in section 3.0.7). For the construction of proof of concept devices, we use a low energy, heavy 14keV P$^+$ ion beam (0.5 keV/amu). Shown in figure 2.3 is the ionisation caused by the two species in Si as calculated by SRIM. It has been sectioned into the ionisation caused by the original ion and the ionisation induced by the recoiling atoms of the target material. The ionisation is directly related to the electronic stopping. The nuclear stopping is shown for comparison. Note that in figure 2.3(b) the electronic stopping dominates and the nuclear stopping only occurs near the end of range when the ions have slowed significantly. The 14 keV ions on the other hand undergo considerable
nuclear stopping over the entire range of the ion. Note also that while the 14keV P\(^+\) ion travels around 25 nm into the Si, a 500 keV He ion reaches a depth of 2 um.

![Figure 2.3](image.png)

**Figure 2.3:** The stopping differences in the two light, high energy regime (0.5 MeV He\(^+\) > 250 keV/amu) and the low energy heavy regime (14 keV P\(^+\) < 25 keV/amu) are contrasted.

While the path of a single ion is simulated by SRIM based on the physics we have just presented, it is more useful to assess the statistical average of a simulation of a large number of ions. For example, the range of an ion has a Gaussian spread about the mean depth. This deviation is known as straggle and exists in the lateral dimension also. SRIM uses the binary-collision approximation (BC) [363] in a semi-empirical Monte-Carlo simulation. Many ion tracks are simulated until clear statistics emerge, this is the Monte-Carlo approach. One statistical result of an SRIM simulation is straggle, quoted as the square root of the variance in the physical distribution of implanted ions. Straggle exists around both the mean depth of the ion and in the lateral dimension. The depth and distribution of implanted ions is of great importance to IC manufacturers, for example, who may use dozens of implantation steps in the fabrication of one chip [81]. Perhaps for this reason, the SRIM code
functions by calculating the ion track first and then distributing the ionisation along the ion path. Each ion track is determined by a series of two-body collisions based on the energy transfer in equation 2.5, each one within a volume of material corresponding to the inverse of the atomic density of the material. This is the volume you would expect to find only one target atom and each collision event has no specific positional data which suits an amorphous target substrate. Minor collisions that deflect the ion less than 0.1° are ignored or grouped into a single larger collision. The electron stopping losses are distributed over the length of the ion path between collisions according to the stopping cross section $S_e(E)$ of equation 2.1. The remaining energy is presumably lost as phonons. Random numbers are introduced into the scattering angle $\Theta$ and the transferred energy $T$ in order to simulate many random ion tracks which then demonstrate implantation statistics. The range of the variables $\Theta$, $T$ and the impact parameter $p$, are chosen so that the results conform to experimental observations and basic physical limitations. This requires SRIM to take a different approach when simulating low energy ions because with fewer collisions along a much shorter track, each single collision becomes more important. Thus the minimum threshold for ignoring collisions is set well below the displacement energy $E_d$ and the range of the allowed impact parameter $p$ is increased because the atoms are travelling so slow. The ion track through the substrate is thus compiled from the many sequential collisions. Electron stopping induced ionisation is distributed over the path length between collisions for higher energy ions and for electron stopping below 25 keV/amu where the slow moving particles are less likely to promote electrons, the empirical model of SRIM relates the electronic stopping power to the impact parameter formulated by Oen and Robinson [364]. The SRIM simulation results, such as those presented in figure 2.3 conform to experimental data within 5% over the entire energy and ion species range in general. The behaviour of low energy ions specifically however, has been seen to diverge sig-
nificantly from the ionisation predictions of SRIM [365]. Additionally, SRIM simulates ion implantation into amorphous substrates and implantation into crystalline targets can be significantly different where channeling occurs [366]. We will treat the two discrepancies separately here.

Firstly, we would like to be able to accurately predict the ionisation induced in the substrate because this is what we measure to detect the passage of a single ion. We would like to be able to compare the predictions with results to ensure that our detectors are responding to 100% of the charge created for the maximum signal. Each 3.6 eV lost to ionisation has the potential to create one electron-hole pair and SRIM predicts that one 14 keV P+ ion will produce on average 5.0 keV of ionisation in the substrate that we can observe. Any discrepancy has traditionally been attributed to charge lost in the detector due to recombination with substrate defects. With the emergence of high purity ionisation detectors with very thin passivation layers, several reports suggest that SRIM overestimates the ionisation expected from low energy ions [365, 367–369] because experimental values are always lower, even with these very high efficiency detectors. While some authors have used their data to improve the equation relating the electron stopping power to the impact parameter, [370,371], others have found it more instructive to provide an updated partitioning factor between nuclear $f_n$ and electronic $f_e$ losses [365, 369, 372]. Both empirical methods provide more accurate results than SRIM in the low energy regime where nuclear stopping dominates. A physical explanation for the problem has been offered by Funsten in [365] as being the exclusion of a lower limit to the electronic stopping in SRIM. A minimum energy (3.63 eV) is required to promote an electron across the indirect bandgap of silicon but such a limit is not accounted for in the SRIM simulation. This causes SRIM to overestimate the ionisation produced where the velocity of the ions is less than the threshold, as is typical during recoil atom collisions in low energy ion simulations. We will discuss the implications for our own experiment in section
3.0.4. This shortcoming of SRIM does not have any bearing on the nuclear stopping portion of the simulation since they are treated independently. Thus the physical distribution of low energy ions as calculated by the ZBL theory is considered accurate and is backed up by experimental data. However, this is for amorphous targets only and the distribution of implanted ions can be vastly different in crystalline targets where channelling occurs [366]. Since we employ (100) single crystal substrates in our work, this is of some concern.

Channeling occurs at specific crystallographic axes that effectively present a material lower in density than an amorphous structure to an incident ion. This can drastically increase the mean depth and distribution of implanted ions [84]. Channelling is minimised by implanting at an off-axis angle or (at low energies) through an oxide layer that will randomise the direction of ions before they enter the substrate [373,374]. The channelling of low energy (sub-50 keV) ions is not sufficient to effect the mean depth of the implanted distribution [374] but there is always some channeling which effectively adds a tail to the amorphous distribution that is predicted by SRIM. From the available data, the minimum fraction of ions found outside the amorphous distribution is 18.5% for 5 keV B [373] and we expect the fraction to be similar or less for 14 keV P [374]. Effectively, channeling increases the range of these few ions, beyond the depth of the remaining 80%. A program called crystal-TRIM exists to simulate the ion implantation into crystalline materials [85] but it is very computationally intensive to keep track of the target atoms. Therefore we will still employ SRIM in our modelling for quick estimates, and it will be necessary to employ the experimental evidence found in the literature to estimate the effect that channeling will have on our model.
2.2 Coherent tunneling by adiabatic passage

In order to determine exactly what we require of a deterministic doping apparatus, we need to specify an intended use. The multi-dopant device we would like to fabricate is a proof-of-principle device for a solid state demonstration of the coherent tunneling by adiabatic passage (CTAP) mechanism [22]. In this chapter we will revise the workings of the CTAP mechanism and discuss the limitations placed on the physical Si:P system required to provide a CTAP pathway. We will show with Monte-Carlo modelling of implanted 14 keV ions that it is feasible to produce the array of donors required for such a proof-of-principle device, with deterministic ion implantation.

CTAP is a protocol for the transfer of a particle along a chain of potential wells and is compatible with Si:P QC architectures [180]. It functions as a ‘wire’ for physically transferring quantum information from one area to another within a quantum information processor. The simplest solid state demonstration requires three well placed donors as shown in fig 2.4. The electron localised around the first atom is transferred directly to the third atom with a counter intuitive pulse sequence that exploits the null eigenstate of the three site sys-

![Diagram of CTAP](image)

**Figure 2.4:** Three donor, one charge system with surface gate control for realising CTAP. The symmetry gates $S_1$ and $S_3$ maintain the degeneracy of the end of chain states, whilst the variation in the tunnel rates is effected by the barrier gates $B_{12}$ and $B_{23}$.
tem. The transfer is affected by applying appropriate voltages and voltage pulse sequences to surface gates situated above (S gates) and between (B gates) the donors. A similar concept was independently introduced for a trapped atom transfer between three wells, where the proximity of the wells is modulated to affect the transfer. Both proposals are analogous to the well known stimulated Raman adiabatic passage (STIRAP) protocol that transfers excitations between energy levels via a dark state using electromagnetic radiation. A schematic of the STIRAP concept is shown in figure 1(a). The final and intermediate levels are coupled with a Stokes laser pulse before coupling the initial and intermediary levels with a pumping pulse to form a trapped state. We obtain 100% population transfer from the initial state to the final state regardless of whether this transition is otherwise forbidden. The route is known as a dark state because there is no interaction of the population from the intermediary level. CTAP is distinguished from STIRAP in being particle transport through matter rather than the manipulation of excitations via photon stimulation. The words ‘stimulation’ and ‘Raman’ are optical nomenclature while CTAP is a solid state protocol without the use of electromagnetic radiation.

STIRAP across a spatial degree of freedom has been suggested in for the transfer of a hole charge across two quantum dots via an exciton state. This left/right charge qubit is proposed as a long lived memory storage qubit. The term STIRAP still applies because photons are used to affect the transfer. In the case of cooper-pair transport in superconducting nanocircuits, STIRAP can be used to control the state within one cooper pair box, and CTAP can be applied to transfer a single cooper-pair from one cooper-pair box to another via a continuum of states in the adjoining lead, regardless of whether the lead is superconducting or not. Schematics of the STIRAP and CTAP systems are shown in figures 2.5a and 2.5b. It is possible to find a CTAP pathway in many systems including quantum dots, donors, spin
Figure 2.5: Various systems utilising the (a) STIRAP and (b) CTAP transport mechanisms. STIRAP is affected by the application of electromagnetic radiation ((i) atomic levels, (ii) quantum dots and (iii) a cooper-pair box) whilst CTAP is affected by non photonic means such as the application of electric fields ((i) donors, (ii) spin chains and (v) cooper-pair boxes) or proximity modulation ((iv) waveguides and (iii) optical traps). chains [379], optical traps [283], superconducting cooper-pair boxes [380] and waveguides [381]. Optical traps alone can accomodate ions, neutral atoms [283, 382] or even BECs [383, 384]. CTAP was indeed demonstrated with photons in waveguides where the state modulation was produced by the separation of the three waveguides as shown in fig 2.5b [287]. The proposal made by Eckert etal [283] has also been realised with trapped ions in an optical lattice, with or without a second species in the intermediate trap [284], which demonstrates the robustness of CTAP to fabrication variation in the central well in particular.

Devices of similar complexity to the proof-of-concept device of fig 2.4 have been fabricated with conventional UV and electron beam lithography tech-
niques [205, 206]. The experimental technique for manipulating donors with the application of electric fields is also well advanced where readout of electron transfer [251] and spin state [21] can be achieved with the use of SETs [19] for donors or quantum point contacts for electrostatically defined dots [385]. The experimental control of these complex devices is also well advanced. Triple dot systems have experimentally demonstrated coulomb blockade with controlled tuning of electron numbers in GaAs [198, 199, 386] and gated carbon nanotubes [387].

The CTAP protocol is robust to fabrication irregularities, such as those introduced by the straggle of implanted ions and as long as the transition is adiabatic, the pulse shape is not critical. The CTAP protocol can be extended to much longer, odd numbered chains with a $\sqrt{n}$ scaling of the transfer time without compromising fidelity [388]. Long range transport can be affected via application of the counterintuitive pulses to alternate tunnel couplings [389] as shown in figure 2.6(a) or by straddling the intervening donors and applying the counterintuitive voltage pulses only to the tunnel couplings at either end of the chain [22], seen in 2.6(b). The straddling scheme allows a great reduction in gate density compared to direct swapping techniques or the adiabatic travelling wave electron transport in [183] which limits many transport architectures [277] and increases the effects of cross-talk [390]. As well as being

![Diagram of CTAP schemes](image)

**Figure 2.6:** The (a) alternating and (b) straddling CTAP schemes for long range CTAP transport.
the simplest to fabricate, successful demonstration of the three site protocol is a necessary milestone towards demonstrating the advantages of long range CTAP transport. We will now model the three atom device as fabricated by deterministic single ion implantation.

2.3 Modelling

So that transport can be effected in timescales short compared with the expected decoherence times, the donors need to be relatively close. Ideally the device should comprise three donor atoms spaced 20-30 nm apart, 20 nm below the surface. Here we model departures from the ideal configuration imposed by the limitations of the implantation of ions through nano-scale apertures. An ion entering a substrate will endure direction changing collisions before it comes to rest. We employ SRIM [329] Monte-Carlo simulations to model the straggling of implanted ions into the substrate. SRIM is the most widely used program for modelling the range and distribution of all ion species from 10 keV to 100 MeV, in a range of materials. The limitations of the SRIM simulation will be discussed in greater detail in the next chapter, section 2.1 which requires detailed context. The SRIM data in figures 2.5 (c) shows the probability distribution of 100,000 14 keV P\(^+\) ions implanted through circular apertures 10 nm in diameter centred on the desired atom locations. A 14 keV P\(^+\) ion will travel around 23 nm into the substrate before stopping but can straggle away from this median position as it undergoes collisions with the Si lattice, giving the range distribution a standard deviation of 10.5 nm. A less energetic 7 keV ion will penetrate 14 nm below the surface and has 40% less straggle but also provides half the ionisation available for detection. We have included the 7 keV results with a view to improve the single ion detector system, which will be discussed in section 3.0.1.
Figure 2.7: (a)(b) For the greatest chance of producing a working device, 5 nm B gates will be required. They would be laid in the space left by the S gates. spacings of 20 nm and 30 nm were used as the extrema of the nominal range. (c) top view and (d) side view of SRIM simulations (see text) showing the spatial probability distribution of 100,000 14 keV phosphorus ions implanted into silicon through three apertures 10 nm in diameter spaced 30 nm apart, patterned in 180 nm of PMMA. There is a 5 nm oxide present on the Si surface.

To model devices containing atoms implanted in this way we begin by writing down the Hamiltonian of the three donor one electron problem in the three-state approximation where we only keep the lowest state of the electron localised around the donor.

We will use the basis $|1\rangle, |2\rangle, |3\rangle$ to represent the three donor sites with on site energies initially set to $E_i = 0$ and tunnel matrix elements $\Omega_{12}(t)$, $\Omega_{23}(t)$, and $\Omega_{13}$ as a measure of the interaction between the donors. Although we will later treat the next-nearest neighbour tunneling as being a potential noise limit, for the discussion that follows it will be assumed to be zero. The CTAP protocol requires tuning of the energies of the sites so that they are equal. This is achieved by the surface gates $S_1, S_2$ and $S_3$ in figure 2.5. The adiabatic pathway is effected by modulating the barrier gates $B_{12}$ and $B_{23}$. In practice,
all gates will affect all of the relevant parameters, and so some cross compensation will be required \([390, 391]\) which can be determined by studying the three-dimensional charging diagram of the three-site system \([199, 392]\). The Hamiltonian is:

$$H(t) = \Omega_{12}(t)|2\rangle\langle 1| + \Omega_{23}(t)|3\rangle\langle 2| + \Omega_{13}(t)|3\rangle\langle 1| + h.c., \quad (2.14)$$

and with \(\Omega_{13}(t) = 0\), the eigenstates are

$$|D_0\rangle = \frac{\Omega_{23}|1\rangle - \Omega_{12}|3\rangle}{\sqrt{\Omega_{12}^2 + \Omega_{23}^2}}, \quad (2.15)$$

$$|D_{\pm}\rangle = \frac{\Omega_{12}|1\rangle \pm \sqrt{\Omega_{12}^2 + \Omega_{23}^2}|2\rangle + \Omega_{23}|3\rangle}{\sqrt{2(\Omega_{12}^2 + \Omega_{23}^2)}}, \quad (2.16)$$

with energies

$$E_0 = 0, \quad (2.17)$$

$$E_{\pm} = \pm \sqrt{\Omega_{12}^2 + \Omega_{23}^2}. \quad (2.18)$$

The coherent transfer by adiabatic passage of an electron can now be understood quite simply. The idea is to keep the system in the state \(|D_0\rangle\) while varying the tunnel matrix elements. For example, the system will be in the desired initial state \(|D_0(t = 0)\rangle = |1\rangle\) when \(\Omega_{12}\) is large while \(\Omega_{23}\) is small and in order to produce the desired final state at time \(t = t_{\text{max}}\), then the tunnel matrix elements must be smoothly transitioned to large \(\Omega_{23}\) and small \(\Omega_{12}\) to obtain \(|D_0(t = t_{\text{max}})\rangle = |3\rangle\).

To effect CTAP, there is a large amount of flexibility to choose the pulsing scheme for the tunnel matrix elements. In STIRAP protocols, Gaussian or Gaussian-like pulses are most commonly employed because of the necessity to turn on the excitation (laser pulse) before varying it \([285]\). However in the solid state, where non-zero tunnel matrix elements arise solely due to donor proximity, such pulsing is not required, and so we advocate the use of the pulses that vary between their extrema at \(t = 0\) and \(t = t_{\text{max}}\). Such pulses are explicitly stated below and illustrated in figure 2.8 and are chosen to be sinusoidal for
Figure 2.8: (a) Two of the many possible scenarios for three implanted donors: where they are equally spaced and where the right spacing is half the left. (b) Tunnel matrix elements as a function of time. The solid lines correspond to the case that $W_{12} = W_{23}$ for the sinusoidal variation as defined in the text, whereas the dashed line corresponds to the case that $W_{23} = W_{12}/2$. (c) Eigenenergies of the states as a function of time with the sinusoidal variation, again solid lines are for $W_{12} = W_{23}$ and the dashed to $W_{23} = W_{12}/2$. (d) the value of the adiabaticity parameter $A$ throughout the process. Note that it is maximised at $t = t_{\text{max}}/2$ irrespective of the values of $W_{12}$ and $W_{23}$, although there are minor differences in the value of $A$. As expected, the process is slightly less adiabatic with the smaller values of $W$.

convenience. In Refs. [383, 388, 393] error function pulses were employed which have some advantages in terms of smoothness of evolution and in nonlinear systems avoid certain complications due to eigenstate degeneracy at the ends of the protocol. It should be noted that the tunnel matrix elements between sites vary exponentially with applied gate voltage [247] and so although the tunnel matrix elements are varied sinusoidally, the form of the control biases will be different [394].

If the CTAP can be considered coherent as long as it’s performed adiabatically to ensure the system remains in the null eigenstate and there is negligible population of state $|2\rangle$. Without a change in energy there can be no change of
the spin state. The theory applied here presents the CTAP pathway as a dark state and this finite state formalism is the same applied to STIRAP. In the realistic spatial setting however, the electron has a wave function that extends out into the material so perhaps application of Shrodinger wave formalism would be more pertinent [394]. With the SW theory, even with perfect adiabaticity, there exists a finite population in state $|2\rangle$ during the protocol, further distinguishing CTAP form STIRAP. However, the theory presented here still captures all of the essential physics of the CTAP protocol, a fact confirmed by recent analyses of the CTAP in the triple square well case [394, 395] and atomistic simulations of the ideal case [391].

To determine whether the system remains in the target state, we invoke the adiabaticity criterion, and due to symmetry we define (without loss of generality) the adiabaticity parameter to be between $|D_0\rangle$ and $|D_+\rangle$. The adiabaticity parameter can be defined:

$$\mathcal{A} \equiv \frac{\langle D_+ | \frac{\partial H}{\partial t} | D_0 \rangle}{|E_+ - E_0|^2},$$

(2.19)

and for adiabatic evolution we require $\mathcal{A} \ll 1$. Note that this is equivalent to other means of quantifying the adiabaticity, e.g. that given in Ref. [285]. Adiabaticity does not translate to a direct measure of fidelity, it is rather a measure of when the assumption of adiabatic evolution is justified. Choosing as the form for the surface gate controlled tunnel matrix elements:

$$\Omega_{12}(t) = W_{12} \sin^2 \left(\frac{\pi t}{2t_{\text{max}}}, \right),$$

(2.20)

$$\Omega_{23}(t) = W_{23} \cos^2 \left(\frac{\pi t}{2t_{\text{max}}}, \right),$$

(2.21)

gives

$$\mathcal{A} = \frac{\pi W_{12} W_{23} \sin \left(\frac{\pi t}{t_{\text{max}}}, \right)}{\sqrt{2t_{\text{max}}} (W_{12}^2 + W_{23}^2)^{3/2}}.$$  

(2.22)

Adiabaticity is a maximum when $t = t_{\text{max}}/2$ irrespective of the relative values of $W_{12}$ and $W_{23}$. This is significant in the design of robust sequences and in
affording a quick estimate of the required timescale for CTAP operation on devices fabricated subject to the statistical variations of ion implantation. To illustrate the energies, tunnel matrix elements and adiabaticity, in figure 2.8 we present characteristic plots for the case that $W_{12} = W_{23}$ and when $W_{23} = W_{12}/2$.

In a realistic experiment, we will set the time for CTAP given a certain desired adiabaticity. So it is more important to rearrange (2.22) to determine the value for $t_{\text{max}}$ that keeps the maximum value of $\mathcal{A}$ at or below some threshold, i.e.

we determine:

$$ t_{\text{max}} = \frac{\pi W_{12} W_{23}}{\sqrt{2} \mathcal{A} (W_{12}^2 + W_{23}^2)^{3/2}} $$  

(2.23)

where we have dropped the sine term because we are evaluating $t_{\text{max}}$ with respect to the maximum value of $\mathcal{A}$. Equation (2.23) is particularly useful in gaining insight into the practicalities of CTAP. The most obvious impediment to very long time scales will be the limits placed by decoherence, which will set a maximum length of time over which the protocol can be conducted [22, 396, 397].

Until now we have neglected the next nearest neighbour tunneling, i.e. $\Omega_{13}$. Formally there is no CTAP pathway for non-zero $\Omega_{13}$ as the Hamiltonian (2.14) has no null (or dark) state, however we can place a good bound on whether neglecting $\Omega_{13}$ will be valid by comparing the period for oscillation on the $|1\rangle - |3\rangle$ transition directly, with the total time for CTAP. We introduce $J \equiv \Omega_{13} t_{\text{max}}$ and assert that if $J \ll 1$ then we may ignore the effect of $\Omega_{13}$ in our analysis. Although this criterion is helpful for proof of concept devices, it may not suffice to assess applicability for a large-scale quantum computer where more rigorous error control is required [180].

To determine the bare (unperturbed) tunnel matrix elements, we use the hydrogenic approximation [398, 399] applied by Openov [400] as a simplified approach to the study of singly ionized double-donor structures [401, 402]. These
results are applicable to donor separations predominantly along the [100] direction in silicon where the valley degeneracy has less effect [221, 245, 401, 403]. Though this approximation ignores the lattice site variability in the donor exchange coupling, it is the only computationally feasible method for simulating a large number of straggle disordered arrays. This should be contrasted with the more accurate, but substantially more computationally intensive rigorous identification of the CTAP pathway in an ideal three-donor device using the atomistic NEMO3D package [391].

We use:

$$W_{ij} = 4E^* \left( \frac{d_{ij}}{a_B^*} \right) \exp \left( - \frac{d_{ij}}{a_B^*} - 1 \right), \quad (2.24)$$

Where $E^*$ is the effective Hartree, and $d_{ij}$ is the interdonor separation between donor $i$ and $j$. With these constraints we can immediately write down $t_{\text{max}}$ and $\mathcal{J}$ as

$$t_{\text{max}} = \frac{\pi a_B^* d_{12} d_{23} \exp \left( - \frac{d_{12} + d_{23}}{a_B^*} - 2 \right)}{4\sqrt{2} E^* A \left( (F_{12})^2 + (F_{23})^2 \right)^{3/2}}, \quad (2.25)$$

$$\mathcal{J} = \frac{\pi d_{12} d_{23} d_{13} \exp \left( - \frac{d_{12} + d_{23} + d_{13}}{a_B^*} - 3 \right)}{\sqrt{2} A \left( (F_{12})^2 + (F_{23})^2 \right)^{3/2}}, \quad (2.26)$$

$$F_{ij} = d_{ij} \exp \left( - \frac{d_{ij}}{a_B^*} - 1 \right) = \frac{W_{ij} a_B^*}{4E^*}. \quad (2.27)$$

A rigorous calculation of $t_{\text{max}}$ and $\mathcal{J}$ would involve full band structure considerations, but these formulae give an extremely efficient mechanism for determining these important parameters with relatively minimal computational cost. The hydrogenic approximation of equation (2.24) also allows us to determine the fraction of simulated devices that have tunneling rates $W_{ij}$ larger than the valley-orbit splitting (10.8 meV) to the 1s excited states. Taking the simple gate geometry shown in figure 2.7 we can exclude implanted donors that prohibit the placement of a 5 nm gate between them which will prevent control of the tunnelling rates. Donors too close to the surface may hybridise with states at the oxide interface [20, 248, 404, 405].
Whilst this may not necessarily be a catastrophic failure channel, we will exclude ion arrays where one or more of the donors are closer than 5 nm to the interface as the coherence times are expected to be shorter and the hydrogenic approximation assumed in our CTAP analysis is expected to be invalid. Donors deeper than 25 nm are expected to be difficult to control with reasonable surface gate potentials, hence we identify this as another failure mode.

To explore the statistical variations in possible three donor devices we performed SRIM simulations of the straggling for both 14 keV and 7 keV P$^+$ implants where the ion strike location was chosen randomly within one selected aperture in a three aperture array. Each aperture was assumed to be 10 nm in diameter (state-of-the-art for EBL in PMMA), separated linearly by either 20 nm or 30 nm. 14 keV P$^+$ reach an average depth of 23 nm below the substrate surface and it is therefore possible to implant through a typical 5 nm gate oxide employing the demonstrated on-chip detector electrodes [249]. Simulations for the 7 keV P$^+$ ions assume a state-of-the-art 1.2 nm gate oxide and will achieve a higher placement precision because the effect of lateral and longitudinal straggle will be less. Near-future developments of the on-chip detection system to improve sensitivity may allow the deterministic implant of 7 keV P$^+$ ions.

The results of our simulations are shown in figure 2.9 and table 2.1. For each three donor array, in which the donor positions vary according to the statistics and the randomly selected impact site within the corresponding aperture, we have calculated $\Delta S_{AS}$, $t_{max}$ and $J$. For a donor spacing of 20 nm, the straggle of implanted ions is more detrimental to the control requirements than to the coherence restrictions. When the donor implantation spacing is increased to 30 nm the lateral straggle affecting gate control is less important while the fraction of donors with $t_{max} < 10$ ns decreases. One in nine simulated arrays of 14 keV P$^+$ implanted into the 10 nm apertures spaced 20 nm apart will
Figure 2.9: The overlap of success modes in a given three donor array. In the 14 keV, 20 nm spacing Venn diagram, most arrays with a short enough $t_{\text{max}}$, small $\mathcal{J}$, suitable $\Delta_{\text{SAS}}$ or $d_{13} > d_{12}/d_{23}$ (favourable coherence) are also suitably positioned for alignment with the surface gates (well positioned laterally) and/or for suitable gate control (well positioned vertically). Increasing the implant spacing to 30 nm subtly decreases the total number of arrays that have sufficiently short $t_{\text{max}}$ times but greatly increases the number of arrays that are well positioned laterally and this increases the number of successful arrays to 1 in 6.

have suitable configurations while spacing the implantation sites 30 nm apart increases the yield to one in six. Decreasing the energy of the phosphorous ion to 7 keV increases this to better than one in two devices with favourably placed ions.

If we increase the aperture size from 10 nm to 20 nm, the results vary little for the 14 keV case, but drastically for the 7 keV ions. This is due to the lateral straggle of the respective energies. Since the standard deviation of the lateral distribution of 14 keV ions is 11 nm, increasing the diameter by 10 nm makes little difference because though there are ions scattered further away from the right range, there are just about as many scattered back towards the centre.
The percentage of three donor devices which satisfy control or coherence restrictions with our present model including straggling. The success modes are listed for implanted donors that are implanted P<sup>+</sup> ions with energies of either 7 keV or 14 keV at a randomly selected impact site within three 10 nm diameter apertures spaced either 20 nm or 30 nm. An array of three donors needs to satisfy all placement conditions to be considered viable (see Fig 2.9); the percentage of three donor arrays which satisfy all control and coherence requirements is given for each implant scenario.

The 7 keV ions however have a straggle of 6 nm, so increasing the diameter by 10 nm puts a significant fraction of ions away from the center.

To include the suspected effects of channeling to the modelling results we remove 18.5% of the favourably placed donors which in a worst case scenario, removes 55% of the favourably placed arrays with one channeled ion in each triplet. Best case scenario, all three ions are in 18.5% of the pairs. This changes the favourable placement fraction for 14 keV ions placed 30 nm apart between 1-in-11 and 1-in-8. The 7 keV ions however have a much shallower peak depth in the amorphous case we estimate half the channeled ions will still

<table>
<thead>
<tr>
<th></th>
<th>20 nm spacing</th>
<th>30 nm spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>success mode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( J_{\text{max}} &lt; 0.001 )</td>
<td>97% 98%</td>
<td>99% 100%</td>
</tr>
<tr>
<td>( t_{\text{max}} &lt; 10 \text{ ns} )</td>
<td>95% 100%</td>
<td>84% 91%</td>
</tr>
<tr>
<td>( \Delta_{\text{SAS}} &lt; 10.8 \text{ meV} )</td>
<td>95% 95%</td>
<td>99% 100%</td>
</tr>
<tr>
<td>( d_{13} &gt; d_{12}, d_{23} )</td>
<td>83% 87%</td>
<td>96% 100%</td>
</tr>
<tr>
<td>not too shallow &gt; 5 nm</td>
<td>72% 75%</td>
<td>72% 75%</td>
</tr>
<tr>
<td>not too deep &lt; 25 nm</td>
<td>42% 90%</td>
<td>42% 90%</td>
</tr>
<tr>
<td>lateral separation &gt; gate width</td>
<td>34% 63%</td>
<td>62% 90%</td>
</tr>
<tr>
<td>CTAP suitable arrays</td>
<td>11% 41%</td>
<td>17% 56%</td>
</tr>
</tbody>
</table>

**Table 2.1:**
be within a suitable range. Thus P ions spaced 30 nm apart implanted with an energy of 7 keV provide between 2-in-5 and 1-in-2 suitable three donor arrays.

This highlights the need to improve the single ion detection system for the deterministic doping of lower energy ions which will produce more favourable results. Nevertheless, our results suggest that an array yield between 1-in-6 and 1-in-11 is immediately realisable. The odds of getting one ion in each of three apertures without the assurance of the deterministic collimator are 2/9. This means if we were to take our chances with an array of apertures these yields plummet to between 1/27 - 2/99 clearly unfeasible. But by making our single ion implantation deterministic, we have a reasonable chance of success. The suitably placed two-atom implants for simple qubit demonstrations would be even higher.

The overall scaling to the longer CTAP chains required for scalable QC is obviously expected to be far less favorable and STM hydrogen resist lithography methods for producing such CTAP chains is the preferred method. However, with the implementation of deterministic doping, ion implantation has a continuing role in the development of prototype, proof-of-principle devices for the study of Si:P quantum computing.

The following chapter describes the single ion detection process and the current state of the system, with a view to improving the resolution and signal-to-noise ratio in order to facilitate greater accuracy in the placement of lower energy ions.
Chapter 3

Single ion implantation

A deterministic doping method needs to be able to register the incorporation of a single dopant atom into a substrate and to control the location of this incorporation accurately. Our approach is to exploit the interaction of the ion with the substrate, specifically the ionisation caused by the impact of a single ion [249].

In the basement of the physics building in the Parkville Campus of the University of Melbourne, we routinely implant 14 keV $P^+$ ions into silicon. This energy is a compromise between range, straggle and detectability. The average depth of such an ion is around 25 nm from the surface, a useful depth for surface gate control [247] though there is evidence that the addition of 5 nm to the oxide, effectively bringing the donors closer to the oxide interface, still produces suitable devices. Single ion trajectories are forged by random collisions within the material and this results in what we call straggle. The lateral distribution of 14 keV $P^+$ ions covers a range 10 nm either side of the entry point. This straggle is reduced in a lower energy ion but such an ion would also produce fewer electron-hole pairs in the substrate, resulting in a smaller detection signal. The detectors we currently use reliably count single 14 keV ions, a significant proportion of lower energy ions would not be observed above
the noise level. In the future better detectors may be available that will enable us to use lower energy ions but currently 14 keV is optimal.

![Diagram of deterministic doping process](image)

**Figure 3.1:** A schematic of our deterministic doping process. A nanostencil in combination with surface apertures define the implant site while the charge detection capability is utilised for position registration and single ion counting.

This chapter is about the PiN detector, shown in the context of deterministic doping in figure 3.1. We will cover the sources of noise and factors that have an impact on the resolution of the PiN detector in section 3.0.1. The results of our recent attempts to incorporate a channel stopper into the design will be presented in section 3.0.7. For reference, here is a description of the 14 keV P$^+$ source.

**The 14 keV P$^+$ source**

To obtain 14 keV phosphorous ions we use a Colutron source [406]. A schematic of the beamline is displayed in figure 3.2. The system can be considered in three parts, the source chamber, the beam line, and target chamber. Each of
Figure 3.2: (a) Schematic of the Colutron 14kV ion accelerator. (b) The ion source. (c) Fischer slits. (d) The target chamber as seen from above.
these sections has an independent vacuum pump system and can be isolated from the others with gate valves. Simple operation occurs as follows:

A filament in the source chamber heats the immediate vicinity. For a $P^+$ beam, gas consisting of 70% Ar and 30% He is let into the source chamber at a controlled rate. The phosphorous is introduced in solid, red phosphorous form on the end of a glass rod which is positioned within the filament as seen in figure 3.2(b). The heat from the filament atomises the phosphorous and with the right argon pressure ($\sim 10^{-5}$), an arc from the filament ionises the atomised phosphorous straight to a plasma. The source chamber thus has two vacuum gauges, one for monitoring the plasma pressure and one for the general chamber pressure. The rest of the source chamber is kept cool by pipes of Dupont HPC-134a, an ozone safe refrigerant. The source plasma is created between the cathode and anode plates which provide the acceleration field; the beamline end is grounded and the anode plate is biased at 14kV. After focussing the accelerated plasma into an ion beam and some vertical steering, the ions pass through a horizontal Wein velocity filter where the magnet separates incident ions by mass and charge [407]. Here $P^+$ is identified relative to the large $Ar^+$ peak (present due to the plasma inducing gas), measured on the Faraday cup in the beamline. Helium ions, nitrogen, $H_2O^+$ and molecular $P_2$ are also present in the mass spectrum though the nitrogen dissipates as the source warms up. With an appropriate gas bottle, $^{15}N$ can also be obtained but an undiminishing $^{14}N$ peak when using the Ar bottle is indicative of a vacuum leak. There are two additional ways that $P^+$ is identified and distinguished from ions with a similar mass and charge such as $^{15}N_2$; pushing the source rod closer to the filament or increasing the filament current. Both of these actions will cause an increase in the phosphorous plasma and thus the $P^+$ peak current.

After leaving the source chamber, the ion beam is collimated with an aperture 500 $\mu$m wide. The vacuum gauge in the beam line typically reads $10^{-8}$ Torr.
Deflector plates that steer the beam with an electric field are then used to keep the beam aligned with the beam line axis and can also be used for scanning. Fischer slits control the beam current; these are crossed V shaped, cylindrical edged slits that form a variable size aperture as shown in figure 3.2(c). With the Fischer slits inserted so that they cross at their narrowest end, the current can be lowered to a pA range, or 1 ion every 30 seconds in the target area.

The target chamber, shown from above in figure 3.2(d) has viewing windows from four different angles, one of which is used to position the target in the beam. A layer of polymethyl methacrylate (PMMA) will be damaged by ion bombardment and this modification is visible to the naked eye. An optical camera with a magnification lens is used to align the spot with the camera focal point which is subsequently used as the sample position. There is a second Faraday cup that can be wound in and out of the beam, used to measure the beam current. An $^{55}$Fe X-ray source, collimated with a Be window can be swung in from the side of the chamber for calibrating the detectors. The chamber must be let up to atmospheric pressure when changing the target sample so it is isolated from the rest of the beam line by a gate valve. The reason for the cylindrical detector apparatus is now obvious, it has to fit through the circular flange on the target chamber. Circular collimating apertures located before the target chamber determine the final beam area incident on the sample, 100, 250 or 600 $\mu$m. The end result is a 14keV P$^+$ ion beam with a selected current and a specific area that allows us to choose the fluence a sample will be exposed to. As we described in the previous section 2.1.1, an ion entering a substrate will cause ionisation of the material. The rest of this chapter is devoted to an explanation of how we exploit this ionisation to alert us to the passage of a single ion.
3.0.1 The detector device

The e-h pairs created by the incoming ion can be exploited to detect such a passage so long as they can be kept from recombining. In practice this is achieved with an electric field; the electrons move against the electric field and the holes in the opposite direction. We generate a field by fashioning the Si substrate into a PiN diode and reverse biasing to full depletion. The following explanation of the detection process includes p-n junction theory, a description of the device and the factors that contribute to noise and signal strength which determine the confidence level we have in counting every 14 keV ion.

The p-n junction

The PiN diode we employ to detect single 14 keV P$^+$ ions can be understood by examining the physics that governs the simple but useful, p-n junction. When a substrate is doped with an acceptor type species like B, they provide an abundance of states near the bottom of the band gap as shown in figure 3.3(a).

These states are available to electrons which makes the material more conductive than regular intrinsic Si. Another way of envisioning this is to say that the p-type material has an excess of positive charge carriers that move oppositely to electrons, that we call holes. On the other hand doping a Si wafer with a donor type species like P introduces extra electrons close to the conduction band, also making the Si more conductive since the effective gap between these states is much smaller as seen in figure 3.3(b). When the two are brought together, we get a p-n junction. We define the Fermi level as the potential energy which has a 50% chance of being occupied at a given temperature. As temperature increases, so does the Fermi level and vice versa. The dopants in the Si change the effective Fermi level of the material, dragging it
Figure 3.3: Electrons can only travel right to left along a descending conduction band, while holes will only travel upwards left to right. Thus figures (c) and (e) represent equilibrium and reverse bias exaggerated barriers to conduction respectively. Forward biasing lowers the barrier to allow electrons to flow from the n-type side to the p-type side.

up or down to where the majority of available states are. When the p-type material is brought against the n-type material, the Fermi level needs to be consistent across the material as shown in figure 3.3(). In order for the system to reach equilibrium, a number of electrons leave the n-type material and cross into the available states in the p-type side. This leaves the p-type side of the junction negative and the n-type side, now depleted of electrons, positively charged. This means that there is now a ‘depletion region’ where the material has been emptied of the dominant carriers. The extent of the depletion region depends on the concentration of the charge carriers within the material.
since the amount of charge needs to be mirrored either side of the junction to achieve equilibrium as depicted in figure 3.3(a) and (b). This leads to a small built-in voltage $V_{bi}$ over this natural depletion region that exists without any externally applied bias. This built in bias is a barrier that prevents any current flow which is seen in the energy band representation in figure 3.3(c). When applying bias across the p-n junction, it can either overcome the barrier created and allow current to flow through the region, depicted in figure 3.3(d) or it can enhance the effect, increasing the depletion width $\omega$ proportionally to the applied bias $V$ via equation 3.1 which is depicted in band structure form in figure 3.3(e).

$$\omega = \sqrt{\frac{2\epsilon N_a + N_d}{q N_a N_d}} (V_{bi} - V)$$ (3.1)

where $\epsilon$ is the permittivity of silicon, $N_a$ and $N_d$ are the acceptor and donor concentrations and $q$ is the charge of an electron. The latter is called reverse bias as it empties the device of charge carriers and prevents any further current flow. Forward bias however, as long as it is larger than the built in voltage, encourages current to flow constantly through the device.

This leads to the ingenuity of the PiN junction detector; the structure is actually P-n-N since the intrinsic region is very lightly n doped. When the diode is reverse biased, the depletion region extends across the whole 300um wafer. In practice, the current is finite rather than nonexistent because of generation and recombination by defects within the depletion region. Traps exist within the band gap of Si, and will release or accept electrons on a timescale that depends on the energy of the trap level below the conduction band or above the valence band and the temperature of the substrate [54, 408]. When an electron or hole is generated by a trap, the carrier is swept out by the electric field and contributes to the ‘dark’ or ‘leakage’ current. The reverse bias current of our PiN detector is naturally very small due to the high purity of the substrate which provides a low noise background, particularly at cryogenic
temperatures. When an ion induces charge within this region, the reverse bias electric field sweeps the charge back out, inducing a mirror charge signal on the back N contact.

The charge produced by an ion impact can be considered a plasma which decays quickly into free charges. Each charge is subject to the Coulomb forces from all the other present charges in range. This includes the repulsion felt by like charges that leads to diffusion of the induced charge. The rate of diffusion of charge carriers, \( J_{\text{diff}} \), is proportional to the gradient of the charge density (\( \nabla n \) for electrons, \( \nabla p \) for holes) as in equations 3.2-3.4 which clearly show the preference of charge to be evenly distributed.

\[
J_{n,\text{diff}} = q D_n \nabla n \tag{3.2}
\]
\[
J_{p,\text{diff}} = -q D_p \nabla p \tag{3.3}
\]
\[
D_{n,p} = \frac{k_B T}{q\mu_{n,p}} \tag{3.4}
\]

where \( \mu_h \) and \( \mu_n \) are the mobility of holes and electrons respectively in the detector Si substrate and D is the diffusivity of the material, dependent on Boltzmann’s constant \( k_B \) and temperature \( T \). The presence of the electric field over the region adds another element of charge transport - drift. Charge carriers will drift with or against an electric field which separates the holes from the electrons and prevents them from recombining. The electron drift current, \( J_{\text{drift}} \), is given in equation 3.5:

\[
J_{n,\text{drift}} = q \mu_n n E \tag{3.5}
\]
\[
J_{p,\text{drift}} = q \mu_p p E \tag{3.6}
\]

where \( E \) is the electric field strength and \( n \) and \( p \) are the electron and hole charge concentrations respectively. The electron and hole currents of course, are in opposite directions. While drift and diffusion widens the gap between holes and electrons, Coulomb attraction between unlike charges is also active. Charge carrier lifetimes in Si are fairly long due to the indirect bandgap which
means that e-h pairs require phonon assistance to recombine. While the plasma density in our case is thought to be too low to promote recombination at a higher than usual rate [409–411], the proximity of opposite charges in the initial charge cloud may lead to some ambipolar transport. The mobility of holes and electrons can differ, and do in Si, but the pull of the positively charged holes can slow down the negative electrons which will thus move with the mobility of holes:

\[
D = \frac{(n + p) D_n D_p}{n D_n + p D_p}
\]

\[
p << n \quad , \quad D \approx D_p
\]

Since the charge appears at the top surface of the detector, the holes are quickly taken by the top p-type electrodes and normal electron transport occurs. In fact, if we estimate the time a single free electron will need to drift from the top surface to the back contact, it is much quicker than the recombination time of such an electron in the bulk [412]. We will discuss the implications of this in section 3.0.4.

### 3.0.2 Fabrication

A detailed description of the fabrication steps that create our PiN detector can be found in [413] but the structure of the finished detector is presented here in figure 3.4(a) and from above in figure 3.4(b). We should keep in mind that the single ion detector is a means to an end, the end being the manipulation of the single electron spins of the implanted donors. For this reason, many of the features of the device have been optimised for the later construction of control devices in balance with the requirements of detection. The intrinsic substrate is an ultra-high-purity Topsil wafer that is 300 µm thick and very lightly n-type with phosphorous present to a resistivity of around 15 kΩ. The top leads visible in fig 3.4(b) are p+ diffused boron while the entire back surface is n+.
Figure 3.4: (a) A schematic of the detector device as seen from the side showing the PiN structure. (b) A photograph of the detector device as seen from above. Barely visible are the detector leads (yellow) but the device leads used for later construction (red) are clear.

Diffused phosphorous as seen from the side in figure 3.4(a). These regions are electrically contacted via aluminium pads deposited directly onto the silicon and the entire back surface is covered with Al to form the back contact. A gap as large as 45 $\mu$m between the top leads yields as much charge collection as closer electrodes [414] so the electrodes are more than adequate set 15$\mu$m
apart. Between the top surface p+ leads is a region of high quality 5 nm oxide with a nice, low interface trap density \((10^{11} \text{ cm}^{-2}\text{eV}^{-1})\) that forms the implant site. This 5 nm oxide is the minimum thickness needed to separate the substrate from the control device leads. It needs to be present before implantation because the hot growth conditions required would diffuse the implanted donors away from their target positions. Therefore the small loss in energy is tolerated for the sake of positional accuracy. Surrounding the thin oxide region and the Al contacts is a thermally grown 200 nm oxide. For further definition of the implant site within the thin oxide region, the detector surface is covered with a 180 nm layer of polymethyl methacrylate (PMMA) in which 20 nm windows are fabricated by electron beam lithography (EBL). All other features of the device, being much larger, are fabricated with UV lithography defined masks. Each detector occupies a 1 mm\(^2\) region of the wafer and is wire bonded from the Al contacts to one of 28 gold coated pins on a commercially available, 5 cm ceramic chip carrier.

### 3.0.3 Operation

Operating the diode under reverse bias produces very little leakage current which provides a low noise background, while cooling the detector and JFET to LN\(_2\) temperatures minimises the thermal noise. The detectors are typically biased to -20 V which fully depletes the intrinsic region for maximum charge collection. When an accelerated ion enters the substrate, the e-h pairs generated are kept from recombining by the electric field and as they drift in opposite directions, a current is induced in the back contact. This signal is picked up by a JFET based, charge sensitive preamplifier that is LN\(_2\) cooled along with the PiN detector for low noise operation. A simplified circuit diagram for the preamplifier stage is shown in figure 3.5 [412].

The drift of electrons to the back, n-type end of the PiN diode causes an
induced charge on this contact which collects on the feedback capacitor $C_f$ and the JFET gate. The current through the n-type JFET channel is modulated by the charge on the top gate, causing the internal resistance of the JFET to vary and thus the voltage input to the integrating op-amp with the output:

$$V_O = -\frac{q}{C_f}$$  \hspace{1cm} (3.9)

where $C_f$ is the capacitance of the feedback capacitor that forms the integrating amplifier. While the detector is under reverse bias there is a so called ‘dark’ or ‘leakage’ current that continuously feeds the detector-preamplifier circuit. This means the output is a steadily increasing voltage with a gradient proportional to the leakage current through both the detector and the JFET. Because the circuit signal would saturate when the capacitor is fully charged, there is a transistor switch to periodically drain the charge from the capacitor and the JFET gates. The duration of the reset is necessarily dead time so the leakage of the JFET is controlled along with the triggering of the reset switch to provide
a nice long ramp of active time, shown in figure 3.6. A reset time of 20-200 ms is ideal. Too short and the dead time becomes comparable to the live time. This can happen where a device has a high leakage due to poor bonding or proximity to the edge of the chip where the current circumvents the substrate. In case the reset time is too long (as it is in figure 3.6), there is an LED installed in the chamber to produce some extra current from photon generated carriers, without significantly adding to the noise. In the deterministic doping apparatus however, the LED is usually unnecessary as the presence of the piezopositioners keeps the leakage in the system high. While the reset time may be up to 500 ms with the piezopositioners off, it is decreased by an order of magnitude with the power on.

![Plot](image)

**Figure 3.6:** Leakage current under reverse bias provides a steady voltage ramp before the circuit resets at saturation of the feedback capacitor.

The charge induced by the ion impacts then manifests as discrete jumps in the output voltage which are shaped into Gaussian pulses. From equation 3.9 we know the height of a pulse is directly proportional to the amount of charge induced on the detector electrodes which in turn depends on the energy of the incident ion as discussed in section 2.1. The pulses are put through a multi-channel analyser to obtain a pulse-height histogram of energy vs counts like the spectra shown in figure 3.7. Single ion pulses are distinguished from noise events by the symmetry of the ion pulse shape, as seen in figure 3.7. Noise events have an asymmetrical tail.
3.0.4

Figure 3.7: The single ion signal is converted from charge to a voltage by the preamp and then these are sorted into a pulse height spectrum. The height of the pulse is proportional to the energy of the impinging ion/X-ray. The noise, P\(^+\) and X-ray peaks have each been fitted with a Gaussian. Low energy signals from incomplete charge collection in the X-ray spectrum have been subtracted.

3.0.4 Pulse height

While the pulse height obtained is primarily proportional to the energy of the incoming ion, it is also dependent on a few additional factors. The percentage of the ion induced charge that is collected by the detector is called the charge collection efficiency (CCE). We use X-ray sources of known energies, \(^{55}\text{Fe}\): 5.9 keV \(K\alpha\) and \(^{57}\text{Co}\): 6.4 keV \(K\alpha\) and 14.41 keV \(\gamma_1\) to calibrate the detector. With sufficient statistics the \(K\beta\) peak can be visible as a shoulder on the main
$K_{\alpha}$. The mechanism of X-ray attenuation in the detector means that we can assume that the peaks observed have been formed with 100% CCE since an X-ray will give up all its energy at once. There is usually a small amount of Compton scattering visible but this is negligible compared to the main X-ray peak. The calibration is only accurate to 0.2 keV.

When compared to commercial devices, the CCE of our detectors is generally better due to the commercial detectors having much thicker dead layers [415]. The thickness (or rather the thinness) of the oxide present on the active area is crucial in obtaining a good signal so this is continually monitored with ellipsometry from batch to batch. The sensitivity of the signal to the dead layer thickness is the reason for utilising the cleanroom in our work, we need to keep the detector surface clear of any debris. As previously discussed, 5.0 keV of the initial 14 keV energy given to the ion is expected to manifest as e-h pairs in the detector as modelled by SRIM. However, the phosphorous peak has been measured at an energy of $3.5 \pm 0.2$ keV. This implies a CCE of only 70%.

It has been known for a long time that semiconductor detectors exhibit a non-linear response to the decreasing energy of heavy ions, not all of the kinetic energy of an ion will manifest as electron hole pairs in the device and this is referred to as the pulse height defect (PHD) [367]. We know a higher proportion of the energy is lost to nuclear stopping for slower, heavier ions, as we discussed in section 2.1, which results in displacements and phonons within the lattice rather than ionisation. Additionally, lower energy, heavier ions do not travel as far as higher energy, lighter ions and therefore the charge produced within the ‘dead layer’ becomes a significant fraction of the total ionisation. The percentage of the charge that remains trapped in the dead layer can be difficult to ascertain but assuming total loss, there is still a difference in the energy predicted by SRIM and the measured charge [368, 416]. This difference
3.0.5 Noise

The noise arising from the detection system provides a minimum energy threshold for detection. Single pulses above this threshold are detected, pulse heights...
less than the threshold cannot be distinguished from the background noise. The pulse height of a single 14 keV P$^+$ will vary according to the statistics of pair production and the noise mechanics of the detector system. This gives a Gaussian-like distribution of possible pulse heights, the lowest of which should be above the noise threshold for 100% counting efficiency. This section deals with the sources of noise in the system as well as the causes of the spread of the ion peak, and how we minimise their impact for sufficient single ion detection.

The deterministic doping apparatus consists of a detector mount, JFET and preamplifier electronics in order to combine detection with the scanned aperture for precision confinement. In the original design and implementation of the system, the main concern was minimising the noise for a clear single ion signal [249, 420]. We utilise the same hardware as the original design, with the exception of the piezo-positioner stack which introduces additional grounding points and requires slightly different tuning of the preamplifier system. Otherwise, the sources of noise are the same in the deterministic doping (DD) system as they are in the original single ion implantation (SII) system.

In all electronic devices there is an intrinsic time variation in conductance and/or resistivity; we call this noise since it causes deviation from the desired or expected signal. The main offenders in our case and in most DC signal applications, are Johnson (thermal) noise and Shot noise. Traditionally these have been described as distinct types of noise with different sources. The thermal noise $V_n$ in equation 3.10 arises due to the fluctuations in state of the valence electrons which occupy a Fermi distribution of states within the material. It is present even in the absence of a bias, dependent only on the temperature $T$ and resistance $R$ of the detector. The Shot noise given in equation 3.11 however is only present during a current flow. It is the consequence of the discrete nature of electrons, arising from fluctuations in the total number of electrons present in the total charge that contributes to the current $I_l$. 
the Shot noise equation given in 3.11 can be found as a special case of a more general, but fundamentally quantum mechanical, description that has been formulated for mesoscopic devices where the original equation 3.11 no longer applies. Beginning with simple tunneling probability and assuming the current in a given direction is dependent on the density of occupied states \( f(E) \), we can obtain expressions for the fluctuation of the current spectra \( S_{I_T-I_T} \) transmitted through the barrier [421]:

\[
i_n^2 = S_{I_T-I_T} = 2 \frac{q^2}{2\pi\hbar} \int dE T f(1-Tf)
\]

where \( \hbar \) is the Planck constant divided by \( 2\pi \). In this equation \( T \) refers to a transmission coefficient where the transmission and reflection \( R, T + R = 1 \). The Poissonian Shot noise in 3.11 is recovered where the temperature is large for ballistic (\( T=1 \)) conductors or the transparency of the barrier is very low (\( T<<1 \)), which is the case in a device as large as ours at 80 K. Note that the thermal noise is present where the density of states \( f(E) \) was introduced in these equations.

We minimise the thermal noise given in equation 3.10 by dropping the Si substrate and the JFET to LN\(_2\) temperatures, \( \sim80 \) K. Contributions to the Shot noise originating in the JFET are actually minimised at 120 K which is compensated for with a slight warming current delivered from the preamplifier electronics. Capacitance matching is significant in minimising the thermal Johnson noise supplied by the JFET [420] which has a fluctuating transistor channel length during operation.
where the total capacitance $C_{\text{input}}$ includes contributions from the JFET capacitance $C_{\text{JFET}}$ and the detector capacitance $C_D$ in producing a root-mean-square charge noise per unit bandwidth. The term $g_m$ is the transconductance of the JFET, which is 4.5 mS for the MX-40 we employ, and $a_n$ is a unitless constant that relates to the length of the JFET, usually between 2/3 and 1 [422]. Though not included in equations 3.14, stray capacitance resulting from the proximity of metals to the detector also contributes to the input capacitance $C_{\text{input}}$. Since the JFET capacitance is very small, typically 40 pF, the capacitance of the detector needs to be minimised. This is aided by the purity of the wafer, biasing to full depletion, minimising the number of wire bonds that create stray capacitance and limiting the size of the detectors since $C \propto \frac{A}{d}$. Though we note that regardless of the number of wired detectors, the back contact occupies the entire size of the chip which makes the chip size the limiting factor. Even across a single chip, the variability in the noise and gain of the detectors can be more than the difference between separate chips which is evident in figure 3.8. We suspect this arises from the wedge bonding and proximity of the detector to the chip edge, rather than wafer or other fabrication variability, which should be more consistent.

During typical operation, we acquire a spectrum for 60 s to set the noise threshold. We run the beam at a current set to give us one P ion per 30 s in the open region, typically (30x30) nm$^2$. If the detector output is stable and there are no noise counts above a certain threshold in 120 s we then feel safe to start the beam and assume that any signal we observe above this threshold is the result of an ion entering the detector. To further assess the validity of this
procedure, both the signal P⁺ peak and the noise can be fit with a Gaussian approximation. Integrating over the Gaussian above the threshold generally gives < 0.5 an ion count. Using this method, the noise threshold of the best detector currently in our possession is 1.45±0.2 keV in the deterministic doping system.

In addition to these considerations, the grounding of the pre-amp control electronics proved to be very important. With the inclusion of the attocube piezopositioners, the extraneous pins of the chip carrier need to be grounded otherwise they act as antennas and pickup electromagnetic noise radiating from the attocubes and deliver it straight to the detector chip. The nanostencil also needs to be grounded to the same plane as the detector to avoid adding to the stray capacitance of the system which would increase the noise via equation 3.13. The noise floor our PiN detector does not significantly change from 80–120K after which it does tend to increase. In fact, even at at room temperature the noise floor, is 4.7 keV which is low enough to detect 6 keV X-rays and 14 keV H⁺ and He⁺ ions which loose most of their energy to ionisation of the substrate.

When an ion peak in its entirety lies above the noise, we can be confident in counting every ion we implant. Therefore we would like the highest possible CCE and the lowest possible noise. It is also important to minimise the noise in order to maximise the resolution of the detector which means narrowing the ion beam peak which can improve the ability to separate the signal from the noise.

3.0.6 Resolution

The P⁺ beam we get from the Colutron contains a Gaussian spread of energies, centred around 14 keV and though the average energy available for pair production is supposedly 5 keV, this will also vary greatly between ions due to
the different collisional paths taken through the lattice. The process of pair production itself leads to a Gaussian statistical variation in the number of e-h pairs produced, even where the energy available to ionisation from electronic stopping, \( E_i \), is the same:

\[
E_{FWHM} = 2 \sqrt{2 \ln(2)} \frac{w F E_i}{2} \quad (3.15)
\]

where \( w \) is the energy required to produce one e-h pair (3.63 eV in Si at 300 K) and \( F \) is the Fano factor (0.12 in Si at these ion energies [423]). The impact of pair production variation gives \( E_{FWHM} = 0.2 \text{keV} \) but additional widening cause by the variation in \( E_i \) is difficult to estimate and may account for some of the discrepancy between calculated and measured resolution. The electronic noise in the system also impacts on resolution, widening the FWHM of any peak with increasing noise. The achievable resolution is:

\[
\Delta E^2_{\text{tot}} = \left( 2 \sqrt{\ln(2)} \right)^2 \left[ \frac{w^2 ENC^2_{\text{tot}}}{q^2} + w F E_i \right] \quad (3.16)
\]

\[
ENC^2_{\text{tot}} = 4 k_B T \frac{a_N}{g_m} C_{\text{input}}^2 \frac{A_1}{\tau} + 2 k_B T D_d C_d A_2 + q I_l A_3 \tau \quad (3.17)
\]

where the total equivalent charge noise \( ENC_{\text{tot}} \) contains the Johnson noise 3.10, the Shot noise 3.11 and dielectric noise [424] weighted by factors \( A_1 = 0.51, A_2 = 1.2 \) and \( A_3 = 1.04 \) for a semi-Gaussian shaping amplifier [425] and with \( D_d \) and \( C_d \) being the dissipation factor and capacitance of the preamplifier feedback capacitor respectively [426]. The shaping time, \( \tau \) of our system is 10 \( \mu \)s. In figure 3.8 the relationship between noise and resolution is evident. In both the single ion implantation (SII) (a) and the deterministic doping (DD) (b) systems, increasing noise tends to increase the FWHM of the X-ray peak.

The percentage of the \( P^+ \) peak that falls below the threshold is 2.34\pm5\%, leaving us with a counting efficiency of 98\pm5\%. The error threshold takes into account the the variation observed in measurements of many detectors
on the same chip, so this can be considered a generic result for the S301 chip and detectors fabricated and mounted with the same specifications. The next chapter discusses the efficiency of other detector designs and mounting parameters.

### 3.0.7 New detector designs

As we have just demonstrated in figure 3.8, the variation in performance across a single chip can be larger than the variation between detectors on different chips so it is difficult to draw conclusions about the effect of specific changes in the fabrication process. It is possible that small changes that seem to be acceptable, can accumulate over time to alter the detector performance more drastically. With the addition of the piezopositioner in the deterministic doping (DD) system compared to the original single ion implantation (SII) system, the DD system may have a different response to changes in the detector. Metal near the detector surface in the SII system will add to the stray capacitance while in the DD system it is likely to pickup rf noise from the piezopositioners and deliver it to the detector capacitively. Both systems would seem to be susceptible to detector design changes that effect capacitance. It has recently become necessary to upgrade the detector design for improved observation of electron transport via the implanted donors and this section documents the performance results of various detector designs in both the DD and the SII systems.

The purpose of the single ion detector is the fabrication of few donor, proof of principle devices for Si:P quantum information processing. To that end, after the ions have been implanted, the Si chips are sent back to UNSW for processing. They are cleaned of all PMMA and Al before a 5 second, 950°C rapid thermal anneal that returns the displaced atoms to the lattice, including the implanted P donor [362]. Afterwards, aluminium electrodes are laid
Figure 3.8: The FWHM of the X-ray peak against the noise threshold for all recorded detectors tested in (a) the single ion implantation system and (b) the new deterministic doping system. These detectors vary in design, intrinsic wafer stock, chip size and mounting method, all of which are expected to affect performance. The FWHM is given in channels rather than energy since it is the most consistent calibration across all detectors and chips. Note the variation between detectors on a single chip (red and blue markers) can be greater than the variation between detectors on distinct chips. The uncertainties arise from the fitting of Gaussian curves to the data.
down through EBL masks to form devices with 30 nm gate leads over the implanted site. These are connected to the diffused phosphorous leads that are already present in the device at the time of implanting. The process flow is amenable to building various devices for different purposes including qubit control, transport or qubit entanglement, only altered masks are required. The recent success with spin transport devices etc [21] was in part facilitated by improved device design. Before the adjustment, there was a significant amount of leakage obscuring the finer details of the current spectroscopy. A channel of p-type boron doping was placed between the n-type device leads to prevent leakage through a two-dimensional electron gas brought to the interface by fixed oxide charge, an inevitable result of the oxide growth. This channel stopper was introduced into the process flow before the implantation step where it was incorporated into the detector design. Several iterations of the design have been tested, here we present observations on two. It is important that these detectors function as well as previous devices for the success of the deterministic doping scheme. Therefore we have put some efforts into determining where these designs have met expectations and where they have been diminished as detectors.

3.0.8 Performance results

The first design incorporating a channel stopper is shown in figure 3.9(a). The channel stopper is a ring of p-doping, with a lighter concentration than the detector leads. Together with the detector leads, it forms a loop around each of the four device leads which form leads for source, drain or gates in the quantum manipulation devices. The channel stopper has been designed to prevent leakage between adjacent leads. They serve their purpose in this regard, adequately preventing signal obscuring leakage to allow measurements like the current spectroscopy presented in [251]. We tested three detector de-
The detector designs studied in this chapter (a) the original DP only design and (b) DP-CS-B and (c) DP-CS-C with channel stopping guard rings. The results are summarised in tables 3.2 and 3.1 which compares the noise level, X-ray peak position and X-ray peak FWHM in channels since it is the only calibration that holds across various detector chips. The exact calibration on each chip varies with the gain, made evident by the varying X-ray peak position in tables 3.1 and 3.2, which may be due to differing detector capacitance. The changing gain of the detectors makes the measure of noise threshold in channels unreliable for comparing detector performance. Even if we had calibrations for each of the noise thresholds in keV, the counting efficiency also relies on the P\textsuperscript{+} peak widths. Thus, we have chosen here to compare the FWHM of the X-ray peak in channels, since the
electric charge noise is directly responsible for this via equation 3.17 and this resolution contributes to the $P^+$ peak width. Since $P^+$ widths are strongly dependent on the width of the dead layer, we have left out this variable for simplicity: we already know that a thin oxide and clean, high vacuum environment is required to maximise this part of detector performance.

**Table 3.1: Summary of detector performance as measured by the SII system.**

<table>
<thead>
<tr>
<th>detector</th>
<th>X-ray peak (ch)</th>
<th>X-ray FWHM (ch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP (S301)</td>
<td>659</td>
<td>97</td>
</tr>
<tr>
<td>DP (FayA)</td>
<td>635</td>
<td>130</td>
</tr>
<tr>
<td>DP-CS-B1-3</td>
<td>595</td>
<td>165</td>
</tr>
<tr>
<td>DP-CS-B1-9</td>
<td>607</td>
<td>153</td>
</tr>
<tr>
<td>DP-CS-B2</td>
<td>632</td>
<td>110</td>
</tr>
<tr>
<td>DP-CS-C</td>
<td>594</td>
<td>308</td>
</tr>
</tbody>
</table>

**Table 3.2: Summary of detector performance as measured by the DD system.**

<table>
<thead>
<tr>
<th>detector</th>
<th>X-ray peak (ch)</th>
<th>X-ray FWHM (ch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP (S301)</td>
<td>245</td>
<td>54</td>
</tr>
<tr>
<td>DP-CS-B2</td>
<td>235</td>
<td>64</td>
</tr>
<tr>
<td>DP-B2</td>
<td>239</td>
<td>101</td>
</tr>
<tr>
<td>DP-CS-C</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The main observations we can make from these results and the results of figure 3.8 are as follows: S301 performs better than all the subsequently fabricated devices, including the DP detectors with the exact same design. DP-CS-B2 performed better than DP-CS-B1. There was no significant increase in the performance of DP-B compared to DP-CS-B and DP-CS-C performed very poorly.
The design of the detector is not the only difference in the detectors listed. While the fabrication process itself might vary slightly between the different people carrying out the fabrication, we will neglect this for now and focus on more significant factors. The size of the detector designs are roughly the same (1mm$^2$), but the chips have been cut to contain anywhere between 3x3 and 6x6 detectors which means the size of the chip will vary. This is important because the entire back surface of the chip is n+ diffused and coated in Al to form the back contact. The area of the chip can affect capacitance which in turn determines the noise via equation 3.13. There are a differing number of bonded detectors which might also effect stray capacitance and certainly determines the local stray capacitance of any given detector - if one side of the chip carrier has more bonds for example, the capacitance of the detectors on this side might be higher. The method of mounting the detector onto the chip carrier was formerly silver paste and changed to PMMA for easier removal. The sliver paste would presumably increase stray capacitance but it is possible for the PMMA to increase dielectric losses. The chip carrier itself was originally custom made from shapal-m ceramic but now commercially available 28 pin ceramic chip carriers are used which have a closer pin spacing and thus might contribute more stray capacitance. The difference in DP-CS-B1 and DP-CS-B2 is the quality of the substrate. DP-CS-B1 was fabricated in high resistivity wafer stock, not high purity. The X-ray peak data of DP-CS-B1 devices in table 3.1 shows a lower gain and poorer resolution than DP-CS-B2 device. This loss in the measured energy could be ascribed to recombination via the impurities present in the depletion width. It is difficult to exactly quantify the effects of these changes, particularly since the variance in performance across one chip can be greater than the difference between differently mounted chips. Based on the low temperature performance data we have collected we can make speculations, but we can make some firmer conclusions with additional evidence. We will attempt to explain these results based the IV characteristics
of the devices, and the CCE as mapped by ion beam induced charge (IBIC) techniques.

**IV of a PiN**

There is an expected profile to the current response obtained from reverse biasing a P-i-N device, shown in figure 3.10(a). As we know from p-n junction theory, during a forward bias larger than the built in voltage, the current should be large and increase with voltage. In reverse bias however, you would expect the current to increase as the depletion region widens to cover more current generating defects until the device is fully depleted and then the leakage current should saturate. However, reverse bias breakdown can occur via two different mechanisms, Zener [427] or avalanche [428].

In avalanche breakdown, carriers that are generated in the device gain enough momentum as they move through the intrinsic layer to cause further ionisation which multiplies forming an ‘avalanche’ of charge carriers. With increasing temperature, the mobility of the depletion region is reduced meaning the carriers require a higher voltage difference before the same effect is achieved. This means that the avalanche breakdown coefficient is positive with temperature. Zener breakdown however, has a negative temperature coefficient. It is the result of having a very narrow depletion width which occurs over a very heavily doped region. Though movement of a charge may be prevented by a potential barrier, we know that quantum mechanics allows tunneling to occur where the wave function of the carrier is longer than the barrier width. As temperature increases, the energy of the carriers is greater, which increases the probability of tunneling. This means Zener breakdown occurs at lower voltage differences as temperatures increase.

While our PiN detectors exhibit breakdown before full depletion at room temperature, they do not exhibit any breakdown before -25 V at LN$_2$ tempera-
Figure 3.10: (a) While pn junction theory tells us that once a diode is reverse biased to full depletion there is no more increase in current flow, reverse bias breakdown may occur. This can happen before full depletion, as it does in the PiN detector shown in (b) as well as in a circular PiN diode fabricated on the same wafer (c).

This indicates that the breakdown we see is of the Zener variety at room temperature, most likely occurring in the thin p-type layer of the device. The intrinsic layer is very lightly doped with phosphorous, making it n-type which extends the n-type depletion width to 300 $\mu$m, far beyond any likelihood of tunneling. The electric field within our PiN diode has been estimated at $1.5 \times 10^3$ V.cm$^{-1}$ at a reverse bias of -20 V, much less than the usual $1-4 \times 10^5$ V.cm$^{-1}$ required for avalanche breakdown in silicon [429]. At 80–120 K after prolonged ion exposure, a different mechanism causes breakdown in the device. We know it is due to the build up of charge in the oxide because the result can be reversed by exposing the substrate to light which dissipates the charge. While a 200 nm oxide will hold a significant amount of charge, there is a finite limit to this ability. Since the accumulated charge is positive, it draws electrons from the negatively charged p-type depletion layer at the oxide interface which allows current to flow in the device. Reducing the bias also counteracts the effect; presumably at a lower bias the magnitude of the negative charge on the depletion layer is too small to allow the carriers to succumb to Coulomb force attraction over the SiO$_2$/Si interface. The same thing happens when the PMMA breaks down and can no longer hold charge [430],
but in this case the effect is irreversible because it is caused by irradiation damage to the polymer chains.

While the leakage current contributes to the noise as stated in equation 3.11, variation over an order of magnitude or two in the value of the leakage current measured at room temperature is not a reliable indication of the quality of the detector at cryogenic temperatures. Stability in the current is a better indicator because this does not tend to change after cooling the substrate. The steadier the current, the smaller the fluctuation in electron number, the less noisy a detector is likely to be (via Shot noise equation 3.11). Figure 3.11 shows the resetting output of a good detector and a poor detector for comparison. A stable current provides a steadily increasing output with a constant time between resets. A fluctuating current will cause a fluctuating reset time, and sometimes a significantly nonlinear output gradient. While the semi-Gaussian amplifier will attenuate high frequency noise very well, the relatively slow, unstable fluctuations seen in figure 3.11 result in in a poor signal-to-noise-ratio.

![Figure 3.11: Comparison of a good device with a stable dark current and a poor device with unstable leakage and variable reset times.](image)

The DP-CS-B1 detector chip was fabricated in Si wafer stock that was manufactured to be high resistivity but was not high purity. The leakage of these devices was relatively unstable; the length of the signal ramp fluctuated by 15%. In addition, the reset time increased and the noise decreased over time.
This is an indication that defect traps are present in the substrate leading to Shockley-Read-Hall generation currents [431]. As the temperature increases, the Fermi level may sweep up through particular defect levels so that the traps become emptied (holes) or full (electrons), slowing generation/recombination in the depletion width and reducing the Shot noise as in equation 3.11. The same device design fabricated in high purity stock, DP-CS-B2, displayed better reset stability, better gain and better resolution, seen clearly in the values of table 3.1.

3.0.9 MeV IBIC analysis

To characterise the CCE over the area of a detector, we employ ion beam induced charge (IBIC) techniques. With an ion beam which is focussed to about 1µm, we can scan the beam over the detector and compare the signal obtained in one region with another. Since we are exposing the detector to a beam with a known energy, any difference in the charge collected is due to CCE variation arising in the detector.

Experimental details

The beam we use is the 5MU Pelletron microprobe located in the basement of the School of Physics. The maximum energy is 5 MeV and it will accelerate protons (H\(^+\)), He\(^+\) and alpha particles (He\(^{2+}\)). For IBIC experiments we typically use 0.5-2 MeV He\(^+\). These ions behave differently to the slower, heavier 14 keV P\(^+\) created in the Colutron and the Pelletron is accordingly very different.

The Pelletron is depicted in figure 3.12. There is no solid source; a mix of H\(_2\) and He gas is let into the source chamber to the appropriate pressure and ionised into a plasma with an RF field. The ions are kicked into the charging
tank beamline with a few keV from a typical high voltage supply. To achieve the rest of the enormous electric potential, a string of conducting metal pellets on an insulating plastic chain are rotated past charge collectors within a tank filled with an insulating gas to avoid arcing. The machine is effectively a large, more sophisticated Van De Graaf generator. With electrostatic steering and focussing, the emergent beam then goes through a velocity selector that also bends the selected species to the horizontal plane. There are several beam lines in the lab that the bending magnet may be swiveled around to, we use the line designated MP2. Unlike the low energy P⁺, the MeV He⁺ will visibly fluoresce on glass or a phosphor screen. This means that we can insert precollimating or current reducing slits and focus the beam with visual feedback. The beam is focussed with quadrupole magnets that sit right before the target chamber. We used 0.5 MeV He⁺ focussed to ∼1µm. The current produced in the detectors by these ions is an order of magnitude larger than the P⁺ so it can be detected at room temperature without the aid of a JFET preamplifier. While the stopping profile of 0.5 MeV He⁺ is very different to that of 14 keV P⁺ (seen in figure 2.3 in section 2.1.3) and cannot predict the CCE of a 14keV P⁺ ion, it is useful for characterising the general behaviour of the new device structures. We image parts of the detector outside the 5 nm oxide region and check they are functioning as expected.

**IBIC Results**

The IBIC images obtained are shown in figures 3.13 – 3.15. The He⁺ peak shifts according to the CCE at each point; a loss in charge corresponds to a loss in pulse height. The median energy of the ions collected at each point is used to create a map. Figure 3.13 was obtained by Changyi Yang in 2006 from a previous generation DP device that contained no channel stopper and no device leads. The image clearly shows the detector leads and the CCE drop
Figure 3.12: The Pelletron 5MU accelerator. A chain of charging pellets inside the three storey tank provide up to 5MeV of acceleration to H or He ions that are filtered and focussed along the beam line to a μm spot at the target chamber.

Away from these leads that is indicative of the falling electric field strength. A higher resolution scan of the active area is shown on the right of figure 3.13. The maximum CCE occurs in the 5nm SiO₂ window and is better than the best commercially available PiN diodes that have much thicker dead layers (200 nm). This particular experiment was performed with 2 MeV He⁺. An outline of the features has been added for clarity.

In figure 3.14 we can see that IBIC reveals the structure of the new DP-CS-B device through the variation in CCE. The detector leads and the channel stopper are brightly lit indicating that the electric field is strongest here as expected. The wire bonds, the device leads and the fabrication markers can be seen because they absorb some of the energy of the ions that pass through them which leads to an apparent step down in CCE. Even the detector designation OP12 can be seen in the top right corner. The CCE in the 5nm SiO₂ region is at 100% CCE. What is most striking about the IBIC image in
Figure 3.13: This is an IBIC map performed in 2006 with 2 MeV He$^+$ by Changyi Yang. The simple leads of the detector are visible either side of the maximum charge collection area, the 5 nm SiO$_2$ window.

figure 3.14 is that the linked channel stopper in DP-CS-B acts to increase the effective area of the detector. An increase in the front area may increase the capacitance of the detector since $C = \epsilon A/d$. Since the entire back contact is biased anyway, the increase in the top detector lead area may not be the most significant contributor to the increased noise.

Figure 3.14: This is an IBIC map of DP-CS-B, performed with 500 keV He$^+$. The detector leads are visible, as is the channel stopper which is linked to the leads. Other features of the device are visible within the electric field halo.
Never the less, for the second incarnation, DP-CS-C, we chose to physically separate the channel stopper from the detector leads in an attempt to reduce the active area of the detector. The IBIC analysis image figure 3.15 revealed that the detector lead and the channel stopper were capacitively coupled which still imparted more active area to the detector. The image also shows uneven patches of CCE across the device, seen clearly in figure 3.15(b). The proximity of the floating p-type channel stopper means they are capacitively coupled to the detector ring. While the channel stopper outside the detector ring adds to the total lateral area of the electric field, the channel stopper within the detector ring is excluding the electric field, which prevents charge collection within this region. This particular design is unsuitable for single ion implantation and the operation at cryogenic temperatures within the respective detection systems reflects this also.

![IBIC map of DP-CS-C](image)

**Figure 3.15:** This is an IBIC map of DP-CS-C, performed with 500 keV He⁺. The detector lead is the visible ring around the top half of the device. In this case the floating channel stoppers located around the device leads are physically separate from the detector leads but apparently capacitively coupled. The channel stopper outside the detector ring adds to the electric field but the channel stopper within the detector ring is excluding the electric field and causing the CCE to be uneven over the active 5 nm SiO₂ area.
3.0.10 Summary

The variance between detectors on a single chip can be greater between detectors on distinct chips so testing of all detectors on a chip is required for good statistics. High purity substrates are required to avoid a reduced signal due to charge recombination and increased noise from trap-assisted generation currents. Designs like DP-CS-C which introduce floating p+ regions on the surface, are detrimental to detection. The integration of a channel stopping ring into the top p+ detector contact as in the design DP-CS-B should not significantly effect detector performance. The addition of the channel stopper increases the active area of the detector but the back contact covers the entire chip and dominates the detector capacitance.

DP-B, S301, and DP (Fay A) all have the same two lead detector design. The vast difference between the performance of the S301 detector and the other DP detectors is likely due to the mounting of the chip - while the S301, 2T-SM1 detectors listed on figure 3.8(a) were mounted on custom made carriers built from shapal-m ceramic and gold coated pins, all the other devices listed in tables 3.1 and 3.2 were mounted on commercially available, pre-packaged two inch, 28 pin chip carriers. The dielectric properties of shapal-m and the prepackaged carrier ceramic are practically equivalent so this is not likely to be a significant noise factor. However the PMMA used to mount the detector chip on the commercial chip carriers, may introduce some dielectric noise. The dominant contribution of stray capacitance would appear to arise from the distance between adjacent pins. While the S301 pins are clearly separated by > mm and are positioned far from the edge of the carrier, the pins on the commercial carrier sit closer than 0.5 mm and are located on the very edge, closer to the surrounding metal of the apparatus. This increases the stray capacitance of the system, leading to increased noise via equation 3.13.

While the shift from custom made shapal-m carriers to the far cheaper com-
mercially available carriers was previously deemed acceptable, it is clear that there are gains to be had by returning to custom made carriers. This could possibly be implemented without the use of expensive shapal-m ceramic as long as the separation between adjacent pins is large.

In conclusion, the deterministic doping (DD) single ion detection works as well as the original single ion implantation (SII) detection system. The presence of the piezopositioners increases the output current due to induction from stray rf fields, but this does not appear to change the Shot noise level significantly. The current can be tuned within the preamplifier circuit so that the output provides an appropriate live time between resets. The best noise threshold is $1.45 \pm 0.2\,\text{keV}$ with a $14\,\text{keV}$ P$^+$ ion counting efficiency of $98 \pm 5\%$ and was obtained on an older generation device without a channel stopper. The noise in the system is dominated by stray capacitance and increasing the counting efficiency of our PiN detectors will require careful consideration of the chip carrier mounting including size of the chip, distance between adjacent pins and closeness of the pins to the edge. For the lowest possible noise threshold, we suggest mounting a single (high purity) detector on a shapal-m carrier with as few pins as possible. Integration of a channel stopper to the detector architecture should be successful once the chip mounting improvements are made.

It was clear from the results of chapter 2 that to facilitate real improvement in the accuracy of our ion implantation deterministic doping system, we need the ability to detect lower energy ions since straggle is the limiting factor. The problem with lowering the energy is that all ion impact indicators are reduced – the damage used in SOI channels, the secondary electron emission and the induced substrate charge. Detecting very small amounts of charge may be possible with avalanche detection [432] where detectors are intentionally operated at avalanche breakdown voltages so that a small signal is multiplied.
into a detectable one. Lowering the dark currents of such detectors is key to detecting single low energy ion impacts [433].
Chapter 4

Constructing a nanostencil

A stencil is an aperture that is used to define an area that we wish to expose to some form of modification. Stencils are used during various fabrication processes in order to make a process reproducible and fast. Compare an ink pad stamp to a ball point pen; the pen has a sharp point to provide good resolution but filling in the shape that a stamp can produce in one hit would take much longer. Stencils are particularly useful where manufacturing speed and uniformity across each final product is desired over resolution. For the masking of ions in particular, apertures have always played an integral part in accelerator hardware - they are used to collimate ion beams by blocking divergent ions. In the context of this thesis however, the aperture in question is present to mask the beam outside the target implant site and thus functions as a stencil.

A 14 keV P\(^+\) beam will be completely blocked by 200 nm of Si or Si\(_3\)N\(_4\). Denser materials like Pt or amorphous carbon have greater stopping powers and a mere \(\sim 50\) nm is required to mask 14 keV P\(^+\) ions. The following section contains a review of nanofabrication techniques that can be used to make nanostencil type structures in a range of materials. Many of these methods also employ nanostencils which provides a wider context for the history of stencil use in
both a scientific and commercial setting.

## 4.1 Available fabrication methods

Most of the techniques available for fabricating nanoscaled apertures have arisen from the integrated circuit industry. For the last few decades the techniques have been continuously refined to allow the manufacture of smaller and smaller components with reliable accuracy. In fact, industry projections and accounts such as the International Technology Roadmap for Semiconductors (ITRS) have historically predicted and measured progress in nm [290] while the typical end user may only see an increase in RAM bytes or processing speed. We present here a short summary of standard fabrication techniques that can be used to fabricate apertures in standard materials, most of which are used at some point along the IC fabrication process flow. We will also mention recent some recent innovations that are yet to be used in classical components but may prove essential in quantum information processing systems.

Note: The techniques described produce layered, two-dimensional structures such as masks or stencils. For further reading about more advanced techniques with the emphasis on three dimensional solid structures rather than apertures, see [434] for a large range of techniques and [435] for basic nanowire growth.

While lithography in a general sense tends to mean printing with the use of an existing template, in semiconductor fabrication it applies to processes that are used to fabricate polymer masks. A positive resist consists of long polymer chains that will crumble into short sections after exposure to UV radiation, ions or electrons. A negative resist polymer will have its short chains reinforced into longer cross-linked chains during exposure to UV, ion or electron radiation. This will provide a large difference in etch rates for the effective removal of the weaker material. Techniques that involve a mask include UV
4.1.0

Figure 4.1: Lithography can be performed with focussed or masked electrons, UV radiation or ions. Masked processes include contact/proximity lithography and projection lithography where a smaller image of a larger mask is projected onto the surface to get beyond the diffraction limit of proximity lithography.

Lithography, electron projection lithography (EPL), ion projection lithography (IPL) and the recently developed nanoimprint techniques. Maskless methods include electron beam lithography (EBL), ion beam lithography (IBL) and scanning tunneling microscope probe (STM) techniques. These lithography methods create devices one layer at a time, the constructed polymer masks are used for doping specific areas of a substrate or the deposition of metal structures. There are however 3-dimensional fabrication techniques that can print complete structures on demand [436–438]. They are known as two-photon fabrication techniques because it is two-photon absorption that polymerises a resist, cross-linking or breaking polymer chains to produce either a positive or negative 3-dimensional structure respectively, with sub-100 nm features.

Three methods of lithography are shown in figure 4.1. A maskless process requires a focussed beam that will scan over the resist to create each feature in the pattern one after the other. The methods that utilise a mask are obviously faster, since an entire pattern is exposed at once. The resolution
of contact or proximity UV lithography is defined primarily by the mask size and the conditions of resist development [439, 440] while projection lithography where the mask pattern is inserted into the optical system as shown in figure 4.1, depends on the quality of the system as well. Both are ultimately diffraction limited, which makes the resolution proportional to the wavelength of the radiation. This has led to a foray into extreme, EUV [441] but the energy of a photon can only be increased so far before Compton scattering starts to make a significant contribution to the resolution. Charged particles can get beyond the UV diffraction limit but the method of energy transfer to the resist involves substantial scattering. While photons will attenuate without significantly scattering outside the exposure area, ions will scatter to varying degrees, depending on the mass and velocity as described in section 2.1. Two examples are shown in figure 4.2.

Both ions and UV photons produce ionisation within the material and electrons scatter strongly in most materials as seen in figure 4.3 which extends the material exposure beyond the range of the ions and photons seen in figure 4.2. However, it is not simply the trajectories of ions or electrons that count but the dose of energy received by the resist which is illustrated by the contour in figure 4.3. A resist requires a certain dose of electrons within a specific time for that region to be damaged sufficiently in order for it to be developed. Accounting for this there is still not much to be gained in using electron projection lithography (EPL) [442] or for that matter ion projection lithography (IPL) [443] over UV techniques due to the scattering. Focussing the particles for a direct writing process can provide better resolution [444, 445] but the slow throughput of this pixel-by-pixel method has relegated EBL to the creation of masks for UV projection lithography which does the high throughput IC fabrication. Outside the IC fabrication industry however, all of these techniques are used in various prototyping research and development work within the scientific community.
The energy of the UV radiation used in lithography is too low for significant scattering, the dominant absorption mechanism is the photoelectric effect. High energy protons also undergo little scattering before the end of range (50 µm for 2 MeV). Low energy heavy ions scatter more and reach a much shallower depth (50 nm for 30 keV Ga).

Figure 4.2: Not to Scale.

Ion beam techniques provide the greatest versatility, since variation in charge, mass and energy provide ion beams for all occasions [446] but have been held back by the lack of bright ion sources [447]. Direct sputtering can be brought about with heavy ions to directly form holes in any material. By the same token, gaseous precursors containing metals can be dissociated on contact with the heavy ion beam in order to deposit the material in a pattern prescribed by the ion beam. These FIB techniques available in dual SEM/FIB systems [448]. Proton beam writing provides the ability to modify materials at different depth levels by exploiting different MeV energies to create 3-dimensional structures.
Figure 4.3: UV and ions create secondary electrons in substrate material. Electrons scatter heavily because they interact strongly with electrons in the material. A certain dose of energy per unit time is required to alter a resist material so that even though the spread of electrons is large (1 \( \mu m^2 \)), the affected area is small as shown by the thick line (300 nm deep).

[449]. Since such a beam employs a relatively large RF plasma source (as opposed to the point-like, solid, field emission sources for EBL), resolution as fine as electron beam or lower energy, heavy ion beams is difficult to achieve but 30 nm has been demonstrated [450] for the fabrication of sub-100 nm structures [449].

The best resolutions achievable with these techniques (bar MeV proton writing) tend to be at a shallow depth, similar to the feature size. In general, the shorter in depth you wish your feature size to be, the better resolution you can obtain laterally. For FIB milling where the 30 keV Ga\(^+\) ions only travel \(~50\) nm into the surface, the more you sputter, the bigger your hole gets as the removed material and secondary recoil atoms contribute to further sputtering at the surface of the hole [451]. This also occurs in EBL which a major impediment to fine resolution. Secondary electrons caused by high energy collisions travel further than the region of exposure and can similarly weaken the
material, though perhaps not as much as in the exposed region. The overlap between adjacent exposure regions may have received a dose large enough for development to occur, degrading the resolution of a pattern in what is known as the 'proximity effect'. Though etching can be anisotropic in crystalline samples [452], in polymer resist samples it spreads in all directions at once. Resist materials need to have a high contrast, meaning the rate of etching of the unexposed regions compared to the exposed regions is large so that there will be an insignificant amount of etching of the unexposed regions during the development. Other conditions of the developing stage can be crucial, such as temperature [453]. However, where secondary electrons increase the etch rate of the edges to a rate comparable with the fully exposed region, the etching occurs outwards as it travels down, leaving a tapered shape to the feature and a slight curve at the top due to the Gaussian beam profile. In the case of FIB milling, it is the recoiled atoms travelling out of the exposure region that weaken the material which can then be sputtered by further recoils and returning sputtered atoms as the ion beam progresses down the substrate. To avoid this in FIB milling, the irradiation must be short in duration - either the current is high or the substrate is very thin. A higher current produces poor resolution, thus most of the fabrication of the holes and the few slots published in the literature was performed in membranes a few tens of nanometres thick. This is less important for MeV proton techniques because as we know from section 32.1 of this thesis, the energy transfer is relatively small along the length of the range of the ion where there is very little straggler or recoil. Very high aspect ratio structures can be fabricated [454] using this part of the ion track. On a side note, exploiting the end of range damage instead leads to a technique for the fabrication of very thin layers of silicon [455].

To achieve even smaller definition, there are a raft of relatively new STM techniques, that have been reviewed thoroughly in [87]. STM tips may be used to physically indent a surface similarly to nanoimprint lithography [456, 457].
They can be used as a field emission electron source which enables them to perform the same functions as a traditional electron beam: they can directly remove [430, 458, 459] or modify [460, 461] material, etch with the assistance of volatile gases [462–464] or they can even be used to assist deposition [294, 465]. Electrically they can oxidise a surface [305, 466, 467], deposit material from the end of the tip [468, 469], and with the ability to pick it back up again [470, 471] they can move single atoms [472–476] or molecules [477] around for bottom-up building of structures atom by atom. While they provide the highest resolution, limited mainly by the tip size, they are similarly slow-throughput, direct write techniques and can only create very shallow structures.

A popular method of constructing small holes for use as stencils [478] or DNA filters [479] is to use the FIB to mill through a membrane substrate and then deposit material to narrow the opening. This can be done with ion or electron assisted deposition in situ [480–482] or in a separate step [478, 483, 484], depending on the material of choice. The process can be controlled if you can observe it, as you can in a TEM. The TEM current closes the hole, apparently due to surface tension in the oxide but it could also be deposition of contaminants in the vacuum [485]. The TEM method has produced holes as small as 3 nm in a 40 nm substrate [483], while FIB deposition as small as 4 nm in a similar substrate [480]. Deposition can also be assisted with a focussed electron beam though the deposition rate is much slower and the ratio of C:Pt is much higher due to incomplete dissociation of the mostly carbon precursor [486–488].

Our primary goal is a the implantation of a line of single atoms and for this we only require a nano-scaled slot rather than a hole to ease the alignment demands in at least one dimension. A substrate of 200 nm Si₃N₄ rules out direct writing or sputtering as a fabrication means. While the observational control of the TEM closing method makes it the most accurate, the high current required
means that the field of view is small and this disallows an even exposure along
the entire length of a 5 \( \mu \text{m} \) slot. This leaves us with the method of milling a
large hole and then back filling it in the manner of Schenkel et al. [480], with
Pt, also adopted in [479]. This has resulted in holes as small as 4.3 nm [480]
in a 30 nm substrate. Standard with most SEM/FIB systems is the ability
to inject a precursor gas like \( \text{C}_5\text{H}_4\text{CH}_3\text{Pt(CH}_3)_3 \) (methyl cyclo
pentadienyl Pt
(IV) tri methyl) which will adsorb to the substrate surface and dissociate on
contact with a sufficiently energetic ion or electron beam. Deposition requires a
specific beam current, too small and very little deposition will occur, too large
and sputtering will dominate. In fact, the sputtering rate can be balanced
with the deposition rate to directly write features that would otherwise be
unattainable with the ion beam alone [482] in very thin substrates, specifically
11 nm in 45 nm of silicon nitride. The same can be done with electron beams
and injection of an etchant such as \( \text{H}_2\text{O} \) or \( \text{XeF}_2 \) where electron beam induced
selective etching can produce an 18 nm hole in a 10 nm amorphous carbon
substrate [489].

4.2 The FIB system

The FEI Nova 600 Nanolab, a dual FIB/SEM system which is located in the
microscopy suite of the University of Melbourne’s Bio21 facility, was used for
all the nanoaperture fabrication presented in this chapter. It comprises of a
vertical 0.2-30 keV SEM column and a 5-30 keV Ga FIB column angled at 52°
as shown in figure 4.4, directed at the rotating sample holder within a high
vacuum chamber. Metallorganic precursor gases, including \( (\text{CH}_3)\text{Pt(C}_p\text{CH}_3) \)
for Pt deposition, are injected through a needle that can be inserted close to
the sample surface. Available gas chemistry allows the deposition of Pt, W,
SiO2 and C. Gases used to enhance milling include \( \text{I}_2 \) and \( \text{XeF}_2 \) which produce
volatile etching molecules when dissociated. The sample stage is wired to
4.2.1

Figure 4.4: The FEI Nova 600 Nanolab is a dual FIB/SEM system which consists of an SEM column, a FIB column at 52° and a gas injection system, in this case for the deposition of Pt. The sample stage rotates for use with either column at normal incidence.

measure the incident current which can be observed during a milling procedure.

4.2.1 Beam diameter

Schematics of the ion beam and electron columns are shown in figure 4.5. The focal spot will be an image of the original source with additional contributions from chromatic aberration \( d_c \), spherical aberration \( d_s \), and possibly diffraction \( d_d \) as given in equations 4.1 [490]. Thus the total beam diameter, \( d \), at the surface will be:

\[
d = \sqrt{d_g^2 + d_s^2 + d_c^2 + d_d^2}
\]  

(4.1)
Figure 4.5: Electron and ion focusing in an SEM and a FIB column respectively are shown.

\[
d_g = \frac{d_v}{M} \quad (4.2)
\]
\[
d_s = \frac{1}{2} C_s \alpha^3 \quad (4.3)
\]
\[
d_c = C_c \alpha \frac{\Delta V}{V} \quad (4.4)
\]
\[
d_d = 0.6 \frac{\lambda_e}{\alpha} \quad (4.5)
\]
\[
\lambda_e = \frac{1.2}{\sqrt{V}} \quad (4.6)
\]

where \( d_v \) is the source diameter, \( M \) is the demagnification induced by the lens system, \( C_s \) is the spherical aberration coefficient particular to the system, \( C_c \) is the chromatic aberration coefficient particular to the system, \( \alpha \) is the beam convergence angle through the final collimating aperture, \( V \) is the beam energy, and \( d_d \) is the beam diffraction which is limited by \( \lambda_e \), the wavelength of the electron.
In an SEM, electrons are emitted from the point of an electron field emission source at a very small angle with $<1\text{ eV}$ spread in the energy. The Ga$^+$ ions however originate from a liquid metal ion source (LMIS) [491] and emerge at a wider angle with a larger energy spread ($\sim 15\text{ eV}$). While electrons are controlled with magnetic lenses since a magnetic field can exert more power for better focussing, the FIB ions require electrostatic lenses because magnetic fields would separate the ions by mass, velocity and charge leading to a much wider beam spot. Alignment and scanning are each performed by a separate set of magnetic field coils and electrostatic octupoles in the SEM and FIB column respectively. At each crossover where the beam is focussed to a point, an increase in the chromatic abberation $d_C$ as well as the beam spot size occurs due to the Coulomb repulsion felt by the charged particles. While this is tolerated in the SEM, the heavier ions which are already at a disadvantage due to the use of electrostatic lenses and the larger liquid source, are only subjected to one crossover - the final focus on the substrate. Along each beam line there is a selectable aperture that serves to control the beam current which reaches the sample. A smaller aperture will improve resolution that would otherwise be lost to spherical abberation $d_s$ by preventing widely scattered electrons from continuing, effectively reducing $\alpha$. The gain continues until the aperture is so small that significant diffraction occurs, this limit depends on the energy of the electron. Ion diffraction is not possible since the ion de Broglie wavelength is extremely small. The crux of this discussion is that while an SEM electron beam can be focussed to less than a nm [492], the 30 keV Ga$^+$ ion beam is at best focussed to $\sim 10\text{ nm}$.

## 4.2.2 Deposition resolution

This section turns into a discussion of the specific factors effecting the resolution focussed ion and electron beam induced deposition and how they can be
modelled. The results are in support of our choice of the electron beam for induced deposition of Pt in our application.

A 30 keV Ga\(^+\) will sputter 6 Si atoms from a Si substrate [493]. Typical milling and deposition resolution when using the focussed ion beam is 50 nm which is larger than the 10 nm beam width. The feature size achievable is generally convolution of the beam width, secondary electron emission and recoil atom distribution. For simplification, the following discussion of resolution limiting factors ignores any contribution from secondary electrons or recoil atoms. This is acceptable for electron beam work but less apt for ion beam milling where features are severely widened by recoil atoms. Secondary emission is less important in ion beam assisted deposition. The distribution of deposited material \( R(r) \) can be evaluated by multiplying the volume of the decomposed molecule \( V \) with the density of dissociated atoms adsorbed to the surface \( n(r) \) and the dissociation cross section \( \sigma(E) \) and the beam current distribution \( f(E,r) \) [494], ignoring the finer detail of variation with impinging particle energy.

\[
R(r) = V n(r) \int_{0}^{E_{PE}} \sigma(E) f(E,r) \, dE \tag{4.7}
\]

\[
\approx V n(r) \sigma f(r) \tag{4.8}
\]

Following the formulation laid out in [1], we can asses the rate of deposition by considering the incoming precursor flux \( J \), the probability that a precursor will stick to the substrate surface \( s \) and include the effect of the density already present where the maximum \( n(r) \) is one full monolayer, \( n_0 \). The model also includes the diffusion of precursors from the region outside the beam width into the region where the density has been depleted by dissociation, as seen in figure 4.6. \( D \) is the diffusion coefficient and the concentration gradient appears in polar co-ordinates. We introduce the time \( \tau \) before a precursor would spontaneously desorb outside the dissociation region and subtract such precursors from the density distribution. Also removed are the precursors that
are dissociated where the beam profile $f(r,t)$ crosses the adsorbed precursor layer, $\sigma n$, since these are now deposited on the surface.

\[
\frac{\partial n(r,t)}{\partial t} = sJ \left(1 - \frac{n}{n_0}\right) + D \left(\frac{\partial^2 n}{\partial r^2} + \frac{1}{r} \frac{\partial n}{\partial r}\right) - \frac{n}{\tau} - \sigma f(r,t) n \quad (4.9)
\]

Which leads to the description of three regimes: electron limited, precursor limited and diffusion enhanced, the effects of which are shown in figure 4.7. In the electron limited regime, $n = n_0$ always which means that the adsorbed precursor layer is refreshed as fast as the beam can dissociate it and the deposited profile equals the beam current profile - Gaussian with a high central peak. We define the precursor limited regime where $J$ is slower than the beam current and there is no diffusion. The precursors are deposited within the cross section of the beam according to the rate of the flux only so that the width of the deposition is similar to the width of the beam but the profile is flat. The particular current profile of the beam is not relevant. There is a third regime called diffusion enhanced, where there is enough diffusion to keep the adsorbed concentration useful, but not as high as $n_0$. Where the diffusion length $\rho$ is less
than half the beam width, there will be a dip in the centre of the deposition profile because the centre is where the precursor density remains scarce \( n = 0 \). A plot of the profiles obtained within these regimes is displayed in figure 4.7.

\[ \tilde{\rho} = \frac{2\rho}{\text{FWHM}_B} \]

**Figure 4.7:** \( \tilde{\rho} = \frac{2\rho}{\text{FWHM}_B} \) The deposited profiles obtained in the different regimes of current to precursor ratios. Reprinted with permission from [1]. Copyright 2007, American Vacuum Society.

Therefore, FIB assisted deposition of Pt is much wider than FEB assisted deposition. In particular, this is due to the much higher rate of precursor dissociation by massive heavy ions which puts FIB assisted deposition on the precursor limited regime. FEB deposition on the other hand is more likely to be in the electron limited regime since the dissociation rate is much slower owing to the tiny mass of the electron. If the deposition profile has a narrow peak then it also becomes visible on a shorter time frame since a certain thickness is required before the secondary electrons from the Pt make it visible to the SEM. In contrast, if we use the FIB even at minimum practical currents, we run the risk of obtaining a diffusion enhanced double peak in the deposition profile.
4.3 Results

The samples used were commercially available 200 nm SiN$_3$ membranes for 14 keV masking. The SiN$_3$ membranes were coated in a standard ∼50 nm of carbon to prevent charging of the sample and a ∼10 nm layer of sputtered gold to assist in SEM focussing. Incidentally, larger random debris from the surrounding environment also assisted in focussing both the SEM and FIB which allowed high resolution milling, deposition and imaging. Any remaining material in the vacuum can be deposited during imaging use of the SEM so the work was carried out at a vacuum of 10$^{-7}$ Torr or below.

Images depicting various factors that need to be considered during FIB work are shown in figure 4.8. If the sample is insulating it can hold on to the positive charge of the ion beam which will cause the ion beam to shift creating a distorted image like the one shown in 4.8(a). To counter this we deposit a layer of carbon on the surface to allow the charge to leave the sample. Such a carbon layer is fairly dark to the SEM so an additional layer of sputtered gold coat is laid down which helps with visibility and due to its granular nature it allows better focussing of the SEM. In the image 4.8(b) there are obvious scratches in the gold layer from contact with plastic tweezers during mounting. Carbon tape (seen on the left of the same image) helps complete the charge diffusion path to the sample stub. Though some debris on the surface is preferred for even easier focussing of the SEM at a lower magnification, contamination by salt crystals from contact with naked fingers is too much, as seen in figure 4.8(c). As an interesting side note, debris we observed include fibres, dust particles and occasionally strange, seemingly self-assembled structures such as the one pictured in 4.8(d).

As we discussed previously, when milling through a thick substrate the surface side of the aperture is wider than the bottom opening which is evident in figure 4.8(e). Here, where perpendicular slots have been milled from either
side of a 3 µm thick Si substrate, the second slot has pushed material into the cavity of the first slot. The edges of the second slot are more precise because it
was milled from the surface opposite the one pictured. The Si$_3$N$_4$ membranes used are under tension and can only tolerate $\sim 15 \mu$m long apertures before they break, pictured in figure 4.8(f). The incoming ion beam used to mill a slot will cause Ga poisoning of the immediate area [495]. The dark halo seen around the 10 $\mu$m slot in 4.8(g) could be Ga but is also likely to be due to the removal of the gold surface layer, exposing the carbon underneath. Any material in the ambient environment may be deposited onto the surface by the SEM beam during viewing. The effect is particularly pronounced at high magnification and thus current per unit area. What looks like redeposited gold particles can be seen in figures 4.8(h)(i) where the end of the slot is narrowed by increased electron exposure. The precursor gas needle is inserted very close to the sample surface, seen on the left in figure 4.8(j). This results in an asymmetrical deposition rate. Though we are depositing onto the solid surface around an aperture, occasionally the beam will drift and unintentionally dissociate precursors within the gap that are not already adsorbed onto the surface. This causes lumpy, missshapen deposits as seen in 4.8(k) since the dissociated molecules tend to stick to pointier regions. We know we are operating in the precursor limited regime by the formation of Pt wires where the electron beam dwells longer than the main pattern. This is usually at the ends of the pattern where the beam reverses. In figure 4.8(l), the wires were formed because the electron beam drifted in this pattern due to a particular feedback system being turned off.

The nanoaperture procedure was as follows: With the sample at eucentric height, normal to the FIB column, a FIB current of 30 pA was used to mill a 10 $\mu$m long slot through the membrane samples in about 20 seconds. The current monitor is a handy way to ensure we have created an aperture through the sample, a typical current trace is shown in figure 4.9(a). When the current, that is measured from the sample stub, suddenly drops it indicates that the ion beam has passed through the sample. At higher ion beam currents
Figure 4.9: (a) the feedback current measured from the sample stub; the drop after 30s alerts us to the ion beam reaching the back side of the membrane and passing through. (b) consecutive deposits of Pt with the electron beam narrow the initial aperture to the desired width. The bright spots are Pt precipitates that form at the dwell site of each circular deposit.

(>100 pA) the sample is automatically exposed to the electron beam periodically to compensate the positive charge build up caused by the ion beam. Therefore when these higher currents are required, the observed current reads less than the milling current for the duration of the mill and then increases to the set milling current as the ion beam passes through the sample and directly onto the stub. When using currents as low as 30 pA however, there is no compensation and the current observed is higher than the set milling value with the slow release of trapped charge. The current drops to the milling value when the beam passes through the sample and is directly incident on the stub only. The step is not immediate as the curve in figure 4.9(a) shows. Initially the beam opens the aperture unevenly and the current continues to drop as the ion beam widens the bottom of the aperture until the sidewalls are relatively straight and then the current due to the ion beam hitting the stub becomes constant.
The apertures milled in this way are typically >100 nm. To narrow them to the required width we deposit Pt on the side walls of the aperture. While using the electron beam for this results in a higher percentage of carbon in the deposited material, the resolution of the deposit is better which provides greater control over the narrowing process. Lines of Pt are deposited with the SEM beam at 10–15 keV, 0.54-0.98 nA for 8 s intervals. The vacuum gauge is a good indicator of the residual precursor gas in the chamber and we wait for the vacuum to return to <10^{-7} between deposition and subsequent imaging to avoid deposition of the leftover carbon precursor pieces. We deposit lines of Pt alternating sides of the slot as shown in figure 4.9(b) to gradually narrow the slot. Because the Pt precursor comes in from the left, more Pt tends to be deposited on the right sidewall with the same exposure times. We try to compensate by increasing the number of deposits on the left side. For narrowing a hole, the Pt is deposited in concentric rings around the initial circular aperture. We found that depositing Pt over the hole with a raster scan of the electron or ion beam, produced unpredictable aperture shapes. The roundness of the final hole also depends strongly on the shape of the initial hole. The same principle applies for the slots, the slot must be straight and have even, regular sides before Pt deposition. Non-uniform sides produce non-uniform deposits because the Pt is deposited in monolayers which conform to the shape of the surface. While 50 nm slots are achievable by direct milling, there can be residual silicon at the bottom edge, as seen in the TEM image in figure 4.10. This residual material narrows the slot faster at this bottom edge which leads to a very thin membrane of material closing the aperture. The edges of a slot are straight with a width of 100 nm when milled with a similar current but for a longer duration.

While we do our best to ensure that the beam is focussed and the vacuum is high before imaging, there are some factors beyond our control. This includes variation in the thickness of the membranes, the condition of the source which
depends on how far along it is in its typical lifetime and the cleanliness of
the vacuum which depends on previous users. This generally leads us to alter
the beam current, energy and times employed in order to achieve the same
outcome. Narrow slots can be reliably fabricated with an initial width of
100 nm and a final closed width of 30 nm. It takes several depositions to achieve
the desired width so we need to observe the narrowing of the slot after each
deposition step. It is difficult to observe the real width of apertures with the
SEM because the narrowest part of the slot may be further down than the reach
of the electrons. The final width is not known until the membrane is viewed
with a high resolution transmission electron microscope (TEM), sometimes
weeks later, where the high energy electrons penetrate right through the slot
and substrate.

TEMP images of the smallest slot and smallest hole apertures that we have fab-

\begin{figure}
\centering
\includegraphics[width=0.8\textwidth]{figure4.10.png}
\caption{The spikes in intensity seen in the TEM trace of a 50 nm slot (left) indicate scattering from small amounts of material at the bottom edge of the slot. A wider 100 nm slot however milled for a longer time, has straighter more even edges (right).}
\end{figure}
ricated are shown in figure 4.11. The nodules in figure 4.11(b) have occurred where the electron beam dwells momentarily longer as it turns around at the end of the pattern and they show that our procedure is working in the electron limited regime which we desire for good resolution. Each aperture is \( \sim 20 \text{ nm} \) wide.

![Figure 4.11: TEM images of 20 nm wide (a) slot and (b) hole apertures. Deposited material can be seen at the end of the slot in (a). The concentric deposits can be seen in (b) around the bright aperture.](image)

While this is larger than our aim of 10 nm, a 10 nm slot is not inconceivable it is just difficult to observe in situ: the hole closure and the variability in conditions mean we cannot rely on previous successful parameters. Most of our attempts to fabricate 10 nm wide slots resulted in a closed structure. The results of chapter 2 showed us that a 20 nm aperture can produce 1-in-7 well placed arrays of P donors implanted through the aperture stepped 30 nm apart. The slots we have produced are suitable for proof-of-concept experiments that demonstrate the implementation of the deterministic doping system.
4.4 Summary

We successfully fabricated a 20 nm wide slot and circular aperture for use as a mask in the deterministic doping apparatus. This was done by milling large apertures in a $\text{Si}_3\text{N}_4$ substrate and then narrowing them with the deposition of Pt. The key to fabricating uniform aperture structures with a FIB/SEM system is beginning with a large FIB milled aperture that is uniform and depositing filler material with the electron beam on the substrate surface, near the aperture but not within the void. This procedure does not have a specific recipe that earns the same results every time due to natural variations in the FIB equipment. This means that it would be difficult to translate into a large-scale fabrication method but it has proven adequate for our purposes.

The following chapter introduces our nanostencil to the detection apparatus which is necessary to demonstrate deterministic doping.
Chapter 5

Deterministic Doping

The alignment of the stencil to the implant site has been the most challenging aspect of the deterministic doping implementation. The piezo positioners move in accurately predictable steps and the single ion detection is routine but merging the two is the technical novelty that allows us to fabricate multidonor devices. An ion implantation with high fluence is compatible with electron beam lithography (EBL) fabrication of devices because device structures can be laid down directly on top of the visible implant. Therefore high accuracy in the location of the implantation is unnecessary. Single donors however are practically invisible so they need to be precisely positioned relative to specific fabrication markers to be useful in subsequently fabricated devices. This chapter covers the positioning system in place for aligning our nanostencil to the target site for true deterministic doping.

5.1 The use of a nanostencil

Our deterministic doping method will allow us to build multi-atom arrays for complex devices. The main difference between our deterministic doping apparatus and the single ion implantation system is the introduction of a nanostencil
to mask the ion beam. The use of template masks to quickly replicate structures for semiconductor device fabrication was one of the main contributing factors to the rise of silicon electronics. We have discussed in detail the range of fabrication techniques that employ masks in various ways in the previous chapter. Commonly, masks are used in projection systems that magnify the stencil structure when the UV, electrons or ions reach the substrate. The benefits of projection lithography cannot be applied to the evaporation of metals because there is no compatible lens system, the only way to define the area we want deposited metal is by using either a lithographic surface mask or a free-standing stencil. When such a stencil is used to define an area of exposure, the method is known as shadow masking. It applies not only to the evaporation of metals but to ion implantation and is the basis of our deterministic doping system. Shadow mask techniques are useful for applications where standard lithographic chemicals are undesirable or simply ineffective, such as they are on C60 fullerenes [496]. The method is also useful for fragile or thin film substrates that cannot survive the spin coating procedures used in conventional lithography [497]. Essentially, a nanostencil provides a way of directly depositing material or implanting ions into a substrate in the desired pattern, bypassing the lithography process. The method covers feature sizes from <50 nm to >100 μm [496, 498] and can be used to a full wafer scale [499].

While more of the literature concerns the deposition of metals than ion implantation, the results are interchangeable with a few important differences. Because evaporation sources are large and difficult to collimate since the deposition of material tends to close apertures [500], a penumbral broadening of the deposited features occurs, as seen in figure 5.1(a). This is dealt with by minimising the gap between the mask and the substrate. One such method is the use of innovative compliant masks that have been fabricated by Sidler et al [501]. A thin SiN film is suspended on flexible cantilevers such that the forces enacted on the mask by the substrate will bend the cantilevers both towards
Figure 5.1: (a) Metal evaporation uses a large uncollimated source resulting in a penumbral shadow of material that widens the features compared to the mask. (b) Ion beams however are collimated and only suffer from scattering which is a much less significant effect.

and away from the substrate so that the mask closely conforms to the contours of the sample surface. The association of ions and stencils have a history in ion beam lithography [502] as discussed in the previous chapter, and shadow masking for ion implantation is much the same with only the substrate being differently prepared. The beam currents used in ion implantation are many orders of magnitude less than the deluge of material involved in metal evaporation and this makes them much easier to collimate with apertures which are much longer lived. Additionally, ion beams can be focussed and steered by electromagnetic lenses and the result is a highly collimated beam for a system that does not suffer shadow broadening and therefore can be carried out with the mask at large distances from the surface [503, 504]. Evaporated material deposited on a nanostencil tends to close apertures [500] though some report a widening of apertures with use [505], suggesting that the stress between the exposed side and the underside of their 150 nm SiN membrane acts to open the slotted aperture. Ion irradiated apertures tend to close as well [301] though the rate is much slower, particularly in a very high vacuum.

Impinging ons can be scattered within the stencil material and the region im-
mediately surrounding the aperture can contribute forward scattered ions that reach the substrate, as seen in figure 5.1(b). This is alleviated by using high aspect ratio apertures [506] in dense material and is a less significant effect than deposition broadening. The most significant factor limiting ion implantation resolution is straggling, as discussed in chapter 2 which is independent of the nanostencil system.

A nanostencil shadow mask system can be described as static, multi-step or dynamic. Static systems allow no movement of the nanostencil relative to the substrate while in use. Moving stencils (or substrates) offer greater flexibility. Multi-step approaches move the stencil in discrete distances such as between different aperture patterns on the same mask or between target sites. This allows flexibility in the patterns that can be produced [507,508] or allows fabrication on a large scale, such as covering a full wafer with nanomechanical resonators for integration with CMOS [509]. A dynamic nanostencil can be used to trace out the desired pattern during the irradiation of ions or evaporated material for a direct write system. The simplest inception of this would occur with a single hole in the mask. While it is easy to deflect an ion beam with a magnetic field to turn it on and off, evaporated metal can only be shuttered with a physical barrier. This can be done on the nanostencil by exploiting the piezoelectric actuation of a polysilicon cantilever that extends over the aperture and contracts when required [510]. The dynamic techniques outlined in the literature are dominated by the use of apertures constructed within AFM probe cantilevers for both the deposition of metals [496,511–513] and ion implantation [514,515]. With the AFM probe used as positional feedback, the positional accuracy may be a few nm and the resolution of the fabrication is comparable the size of the aperture. However, the small size of cantilevers limits the masking ability of such nanostencils, and usually they only partially covers the total area exposed to the evaporated metals. The majority of static or multi-step shadow mask experiments demonstrated in the literature, were
performed with nanostencils fabricated in SiN chips because they are readily fabricated or commercially available, are relatively mechanically robust, chemically stable and come in thicknesses from microns down to a few tens of nanometers. Grevin et al have managed to combine the convenience of SiN membrane chips with the advantages of AFM feedback [516]. While simply affixing an AFM tip to the chip itself provides an inflexible cantilever that is poorly suited to conventional AFM, attaching the chip to the prong of a tuning fork sensor (named the Qplus arrangement) allows a more sensitive AFM measurement [517]. They have achieved 165 nm resolution with this method.

The deterministic doping system we have implemented uses a different position feedback mechanism. It is based on ion beam induced charge techniques discussed in section 3.0.9. The substrate we implant ions into has been formed into a detector which enables us to observe each and every ion enter the substrate by the charge it produces as it collides with the silicon lattice. The charge induced is proportional to the energy of the ion when it reaches the live part of the detector, which allows us to produce a map of CCE by scanning and ion beam across the surface of the detector. If an ion passes through a dead layer before entering the Si substrate, it will have lost some energy which will be reflected in the observed signal. By engineering a surface pattern in a charge trapping layer like PMMA [430], we can associate the charge signal output with nanostencil position. This chapter describes our proof-of-concept experiment with both 0.5 MeV He$^+$ and 13.5 keV Ar$^+$ (where 14 keV P$^+$ was unavailable at the time of the experiment).

But first we will present timed 14 keV P$^+$ implants that were performed with the nanostencils we fabricated, which show the good masking ability of our nanostencil and demonstrates the multi-step ability of our deterministic doping system.
5.2 Timed shadow mask implants through a multi-step nanostencil

There will always be some ions that have passed through the mask material and scattered into the open aperture, as depicted in figure 5.2. This means some of the initial energy of the ion will have been lost before it reaches the detector substrate. The area that is responsible for scattering is the section immediately around the aperture, depicted in figure 5.2(b).

![Figure 5.2](image)

**Figure 5.2:** While a 14 keV P⁺ ion will stop after 30-40 nm of Pt, ions entering the nanostencil near the edge of the aperture may scatter into the aperture and continue with a reduced energy.

The edges of our nanostencil appear very straight and uniform in the TEM images that appear in the previous chapter 4.3, which bodes well since irregular edges (in shape and thickness) will cause more scattering of the ions. The straggle of a 14 keV P⁺ ion within the relatively dense Pt, is only 5.2 nm, in Si has a straggle of 11 nm. From the same TEM pictures we can see that the Pt extends further than 30 nm from the edges of aperture which means ions can only be forward scattered into the detector if they have entered the Pt within 5.2 nm of the aperture wall. If the slot were made entirely from Si₃N₄, there would be many more scattered ions reaching the substrate because they could find their way into the aperture from further away in the mask substrate. It is clear that the higher the aspect ratio of a stencil, the more
scattered ions are blocked by reentering the mask material at the opposite wall of the aperture. The nanostencils we fabricated have been made in 200 nm membranes to produce an aspect ratio of 10:1 to minimise the percentage of scattered ions that reach the substrate outside the confines of the shadow mask.

While it is possible to perform Monte-Carlo ion scattering simulations [518,519] a precise profile of the slot is required for accurate results in the low energy regime and this is difficult to obtain. Since the 14 keV P$^+$ peak is so near to the noise threshold, we cannot expect to see any scattered ions in the pulse height spectra obtained with the deterministic doping system.

By performing a multi-step, timed 14 keV P$^+$ implantation at high fluence through the nanostencil we are able to assess the masking ability of the nanostencil. A high fluence implantation is visible to SEM and AFM due to the damage created in the substrate that provides a high contrast to the surrounding crystalline material. If there is significant scattering, the modification of the implanted substrate will be broadened compared to the size and shape of the aperture.

Figure 5.3: (a) alignment markers etched through 5 nm of SiO$_2$ on Si. (b) an SEM image of the array of slots used for the timed implantation experiment.
5.2.1 Results

To measure the effects of implants we performed AFM, KFM and SEM imaging on the implanted sample. The results can be seen in figures 5.4 and 5.5. While standard atomic force microscopy maps the physical height of a sample by monitoring the strength of the proximity based van der Waals force, Kelvin force microscopy probes respond to the capacitive force between the probe and the surface that is a result of the potential difference between the two [520]. While the alignment markers are apparent in figure 5.4(a), the KFM scan is performed away from the surface where the van der Waals force is ineffective so only the change in electric potential is visible in figure 5.4(b). The AFM image shows that the implantation caused amorphisation of the c-Si which raises the area since the density of a-Si is lower than c-Si. The difference in potential could be caused by the doping, though at this dose there is experimental evidence that once incorporated the KFM change disappears. This implies that the potential difference is also altered due to the amorphisation of the c-Si. During the experiment specific to figure 5.4, the sequential exposures were spaced closer (50 nm) than the width of the slots (75 nm) which blurs the exposures into one. Each visible shape is the result of a different aperture in the nanostencil, which is pictured in 5.3(b). Though we can not measure the width of the modified region caused by a single exposure in this instance, note that the third exposure shows a weaker electrostatic response. The particular slot that formed this implant was narrower than the others, hence this area received a reduced dose compared to the others and the KFM reflects this.

In figure 5.5 we demonstrate the difference in AFM, KFM and SEM imaging of the same sample. While AFM is only sensitive to surface contours and KFM is sensitive to the surface potential, SEM would appear to be sensitive to both. The dark implanted regions appear such because there is less secondary emission from these regions. This agrees with the KFM result that
the potential, and hence the work function of the material is greater in the irradiated regions. While the etched SiO$_2$ is evident in the AFM images due to the height difference, the markers are distinguished in the SEM images due to material properties. The electrons channel down into the exposed c-Si but charge the SiO$_2$ to create greater secondary electron emission and hence the oxide area appears lighter in figure 5.5(b).

In this particular implant run, the step between implants (200 nm) was greater than the width of the slot (76.7±2.5 nm) so a single exposure can be imaged. The width appears narrower in the SEM (76±12 nm) than in the AFM (104±16 nm). This is primarily due to differences between each technique rather than any physical difference. Surface features in an AFM image are broadened due to the width of the probe tip - the centre of the probe is kept at least one tip radius away from the surface feature by physical width of the probe. The electron beam suffers no such effect. In addition, while the AFM
is highly sensitive to the surface modification on the sample, the electrons of
the SEM penetrates some way into the sample and images the modification to
the material below.

The uncertainty limits on the AFM and SEM measurements reflect on the
resolution of the images. The uncertainty in the TEM measurement is larger
than the resolution and reflects the variability in the slot width. The implant
conforms to the size of the aperture but both the surface contours and the
SEM imaging are not sensitive enough to give us meaningful results. We chose
the fluence to be an order of magnitude more than that required to produce a
visible change in the substrate.

Let us consider straggler. Due to straggler, 20% of the implanted ions would be
outside the original confines defined by the aperture. The area they occupy
is larger than the area inside the aperture limits so that the concentration
outside is only an eighth of the main concentration and this in turn is not
distributed evenly over the area - 16 nm out from the aperture limits. There

Figure 5.5: A comparison of (a) AFM (b) SEM and (c) KFM scans of the same
samples.
should be a visible widening of the implanted structure due to the straggling but there is not, at least outside the uncertainty limits of the measurements. If the scattering profile is similar to the straggling profile, and 40% of ions were scattered, we would not see the 20% that scatter outside the aperture limits.

In any case, for our deterministic doping method, the lateral scattering direction of ions is less important than the altered energy of these ions; widely scattered ions would be stopped by the surface PMMA mask but ions that scatter within the implant zone with a severely reduced energy will produce insufficient charge to be detected. These ions would be detrimental to the later experiments carried out at mK temperatures either by directly pulling the electron away from the target donors or increasing spin diffusion to reduce the coherence time of the system [238]. Because these ions can’t be detected, there is no easy way to measure their prevalence.

A TEM scan of the cross section of an implanted region would reveal little because the contrast of P to Si is very small and therefore we would not obtain better resolution than we have already achieved. It should be possible to irradiate a detector with 14 keV He\(^+\) and see a significant part of the spectrum. The noise threshold of the best S301 detector device at 300 K is 4.7 keV and the SRIM determined ionisation for 14 keV He\(^+\) is about 11 keV. With such a good detector we would expect to see at least the higher energy half of the scattered spectra and use it, along with modelling such as in [519] to extrapolate to 14 keV P\(^+\).

We have minimised scattering through the aperture by fabricating a slot with a high aspect ratio and straight, uniform sides. The results so far show that the implant region is confined to the area of the aperture.
5.2.2 Lifetime of the nanostencil

There has been evidence expressed in the literature that such apertures may close over time under ion exposure [301] and in fact this is the backbone of the fabrication methods outlined in the previous chapter that entail depositing metals dissociated by ion irradiation. While we use a very low current when implanting single ions, the initial positioning of the ions is done with a higher current so closing under exposure concerns us. If it is due to deposition of outgassed elements in the vacuum, we can only minimise it by using very high vacuum. If the closing mechanism is a result of the migration of atoms into the void due to surface tension in the material, as suggested by [483], we have alleviated the tension of the Si$_3$N$_4$ with the deposition of Pt rather than SiO$_2$ which was used in [483]. The apertures in [301] were milled into SiN without any further deposition. For long term use, the characterisation of the lifetime of a stencil is important data to have but for the duration of the proof of concept experiments in this work, we found no significant closing evidence. As described, the experiments in question include TEM imaging and ion exposure. Implanting $\sim 10^{15} \text{cm}^{-2}$ Ar or P into the mask did not modify the Pt/C material near the aperture enough to narrow it but there are surely limits to this durability.

5.3 Deterministic Doping

In the previous section we learned that heavy fluences create implanted regions that are obvious under the scrutiny of a scanning electron microscope. Single ion arrays however will be next to invisible. Ion irradiation can modify the etch rate of SiO$_2$ [521, 522] which might make it possible to locate the entry point of single P$^+$ ions with an STM scan of an etched sample. However, this method of observation is incompatible with lithographic fabrication pro-
cesses that employ SEM imaging to locate alignment markers. This requires us to make our implantation deterministic by precisely locating the nanostencil above the target implant site that has been predefined relative to conventional fabrication markers. We are able to use the ion strike signals we can obtain from our PiN substrate for positional feedback and this section describes that method of alignment.

First, we will review the deterministic doping apparatus we have implemented, which combines the detection system described in chapter 3 with the nanostencil detailed in chapter 4.

Figure 5.6: The deterministic doping apparatus for the fabrication of single ion arrays.
The apparatus

The apparatus is seen in figure 5.6. The nanostencil is positioned and relocated with commercially available, attocube™ nanopositioning stacks. The PiN substrate and a JFET preamplifier, output a voltage pulse proportional to the energy of any ion that enters the substrate. The PiN substrate is mounted to a chip carrier and consists of several detectors wired to different pins. The JFET preamplifier is located behind the chip carrier position within a block of macor® ceramic (for good thermal contact and low dielectric losses) inside a copper housing (for thermal mass and noise shielding). The Cu housing, JFET and the detector chip are reduced to cryogenic temperatures with LN₂ cooled nitrogen gas flowing through the coils that pass through a copper heat sink attached to the top of the Cu housing. This detector housing is thermally isolated from the positioner stage block because the position feedback encoders within the piezo-positioners will not function correctly when so cold. The nanostencil supporting block is electrically isolated from the positioner stack and grounded to the same plane as the detector to prevent the RF noise that would be transferred capacitively from the piezo-positioner stack if the membrane were a floating contact.

The apparatus hangs from a large micropositioner stage mounted onto the target chamber of the 14 keV P⁺ ion accelerator (which was discussed in section 3). The micropositioning stage is shown in figure 5.7 and assists in aligning the apparatus via a microscope ex situ, as seen in figure 5.8. It is also used to position the selected detector in to the path of the ion beam once the apparatus is in the target chamber. During any type of implantation, the ion beam is aligned with the area of interest on the sample by ‘burning a beam spot’ into the corner of the chip. A high current of ions into PMMA will modify the structure visibly and by moving the focal point of the chamber-mounted camera into the resultant beam spot, the focal point becomes the
5.3.1 Alignment of the aperture to the implant site

Because the membrane is optically transparent, we are able to view the detector features through the membrane as in figure 5.9 which allows us to align a

reference point where the ion beam crosses the target plane. This works by assuming the sample surface is relatively flat, or rather, perpendicular to the beam. This is not always the case and so a repeated burning of the beam spot and adjustments to the camera are sometimes required at opposite corners of the chip to ensure accuracy.
Figure 5.8: The magnification lens of the camera is seen in position to align the nanostencil to the target implant site.

specific aperture with the implant site. If the membrane is within 50 µm of the detector surface, we can simultaneously view the nanostencil and the detector within the depth of view of the microscope. Otherwise, we use the z-axis of the transax micropositioner to move the apparatus in and out of the microscope focal plane, and align the features of the two surfaces to the same place in the microscope image, as depicted in the bottom images of figure 5.9.

There is room for misalignment between the plane of the membrane, the plane of the camera, the axes of the piezo-positioners and the axes of the transax micropositioner as shown in figure 5.10 which shows how each angle can be
Figure 5.9: The apertures in the nanostencil can be seen clearly in the camera image and as the membrane is transparent, sample features are visible through the nanostencil membrane. This allows alignment of selected apertures to specific detectors as shown.

determined. In practice, we square the various axes as shown in 5.11. We define the x-axis to be perpendicular to the beamline while the y-axis runs parallel. First we align the camera normal with the y-axis of the transax micropositioner motion: detector features will appear to move off to the left or right when moved out of focus by the micropositioner y-axis if they are not aligned. When the feature blurs but stays in the same spot, then it is moving parallel to the camera normal. The piezopositioner axis is aligned to the detector surface by naked eye. The membrane is aligned to the detector surface by close observation with the microscope and a purpose built tilt mount. This is required not just for simplicity but to keep the gap between the membrane and the substrate constant during long range movement to prevent crashing the membrane into any part of the detector, which may destroy either surface. Alignment of the piezopositioner axis to the others is necessary for the encoder
5.3.1

Figure 5.10: There are many axes in the apparatus that require alignment for accurate use. The angles can be determined with simple measurements and similar triangle principles.

Figure 5.11: For the piezo- and micro-positioner feedback to be accurate, the apparatus axes need to be aligned as shown through close, magnified observation with the camera.

position readout to accurately reflect the movement of the nanostencil across the detector. The alignment of the entire apparatus to the beamline is then critical in replicating the ex situ alignment inside the target chamber. From observation, we estimate a maximum misalignment of $\pm 5^\circ$ as the sum of all possible misalignments. The distance this translates to on the surface of the detector, is determined by the height of the nanostencil above the surface. The current design of the detector with wire bonds spaced 1 mm apart keeps the membrane at least 700 um away from the surface which means a maximum
misalignment of ±60 µm laterally from the target implant site. While we are able to demonstrate proof-of-concept experiments with this, device fabrication in the future will require modified detectors that allow a much closer approach. If we use the width of the active region 10 µm to set a required coarse alignment of ±5 µm, we demand a membrane-detector separation of 50 µm.

This ex situ alignment is only the coarse alignment and a separate mechanism is in place for fine alignment, which is the main subject of this chapter. The fine alignment is based on the collection of ion induced charge and thus occurs inside the high vacuum target chamber with the detector cooled and under reverse bias, in situ. As we have previously discussed, the height of the signal pulse obtained from an ion strike is proportional to the energy of the ion. This means we can distinguish between ions that have entered the detector directly and ions that have lost energy passing through an additional layer of PMMA, as seen in figure 5.12. With good energy resolution we can also distinguish ions with energies in between, such as scattered ions, or ions that have passed through some of the total PMMA thickness (green in figure 5.12). By stepping the nanostencil across a patterned section of PMMA and recording pulse height spectra at each position, we can reconstruct a surface contour. Although we are not observing the CCE of the detector with this technique, it still utilises the ion beam induced charge (IBIC) principles, so we call it nanoIBIC imaging.

5.3.2 MeV nanoIBIC

We have demonstrated this fine alignment concept using MeV He\(^+\) ions and a 50 nm wide slot made in a re-purposed, 3 µm Si AFM cantilever that has the required thickness of material to stop 2 MeV He\(^+\). Data showing the beam crossing the edge of an EBL trench in PMMA in 14 nm steps is shown in figure 5.12. The spectra are each compressed into a single row of pixels and then placed together to form the image in figure 5.13.
5.3.2

Figure 5.12: Experimental pulse height spectra. The ions traversing the PMMA layer enter the detector with less energy than an ion passing directly through the trench. Stepping the ion beam over an edge produces the spectra seen above.

Figure 5.13: Lining up the spectra reproduces the EBL pattern on the surface of the detector where each column of pixels represents a 100 nm step. Though the finest trench cannot be seen by an optical microscope, it is visible in the nanoIBIC pattern.

The colour scale corresponds to the number of counts recorded at that pixel, the x-axis is the position of the nanostencil and the y-axis of the image is pulse height/energy. It is a 1-dimensional nanoIBIC trace of the surface features of
the detector. By direct comparison of the shapes of the surface map to an AFM trace of the pattern, shown also in figure 5.13, the resolution of the nanoIBIC map is about $50 \pm 3 \text{ nm}$. This is exactly what we expect since the width of the slot utilised in the experiment is $47 \pm 3 \text{ nm}$ and the smallest EBL slot is much larger. The technique could be improved with smaller EBL trenches or a smaller nanostencil aperture.

For example, if the nanostencil aperture is larger than the width of the surface trench, alignment can be achieved by observing the relative count rates of low energy ions to high energy ions, as seen in figure 5.14.

**Figure 5.14:** (a) the motion of the nanostencil (b) the relative count rate strengths of low(red) and high(blue) energy ions which can be exploited to indicate reaching the centre point of the trench for accurate alignment. The resolution is comparable to the step size of the scan rather than the width of the slot.

As long as the step size is smaller than the width of the EBL trench, the point where the low energy count is minimum and the direct high energy ion count
rate is at maximum, shows that the aperture is located over the centre of the EBL trench. This improves the resolution of the position feedback beyond the width of the aperture to the same order as the step size of the scan. This is a very important feature of the position finding method. It ensures that the full accuracy of the piezo movement can be exploited, even when the size of the nanostencil aperture is larger than 10 nm.

We need to consider the possibility that the nanostencil slot is not aligned parallel to the trench. In this case we will only get meaningful results if the trenches are spaced far apart so that the nanostencil slot can not cross multiple trenches simultaneously. Then the signal will still show a plateau, the centre of which occurs when the middle of the slot length is over the trench. The longer the slot, the further apart the surface trenches need to be so a shorter length is preferred.

The question of alignment in the other dimension can be resolved in different ways. When using a slot for one-dimensional scanning, we ensure that the alignment trenches are shorter than the nanostencil aperture as seen in figure 5.15. We can then observe the count rate as the slot is moved up/down over the trench and where this is at a maximum plateau, we have ensured that the nanostencil aperture is held over the centre of the trench which is in turn aligned line with the implant site. Even with tapering as shown in figure 5.15(a), this pattern requires us to approach the detector from the right or left with less than $\pm 10\,\mu m$ misalignment in a vertical direction. To allow approach from any direction during a search scan, a concentric type of pattern such as that shown on figure 5.15(b) is required in conjunction with a hole type aperture. Remembering that the nanostencil membrane window is $500\,\mu m^2$, much larger than both the $20\,\mu m^2$ detector window and the $250\,\mu m$ beam diameter, we can have both types of aperture on the same nanostencil spaced far enough apart for independent use. Thus it should be possible to combine elements of
both the 1-d and 2-d alignment patterns on the detector to allow both a multi directional approach and ease of alignment in the vertical dimension for the implant, as in figure 5.15(c).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.15.png}
\caption{(a) A possible pattern to for aligning the nanostencil aperture relative to the implant site, assuming misalignment in one dimension only. (b) A possible pattern for locating the nanostencil aperture relative to the implant site with misalignment in two dimensions. (c) A possible pattern that allows a multi-directional approach in addition to ease of alignment in one dimension in the implant zone.}
\end{figure}

In summary, utilising a tiny aperture or EBL trench, good energy resolution in the detector and small positioner steps, we can achieve fine alignment.
5.3.3 14 keV nanoIBIC

As discussed in the detector chapter 3.0.1, using the S301 detector we obtained 98% counting efficiency with the stencil in proximity to the detector. With a circular aperture we effectively have a nano-scaled ion beam with a diameter matching the aperture size. Labview control of the piezopositioners gives us the ability to raster scan this 14 keV ion beam that we can use to make a 2-dimensional nanoIBIC map, analogous to those we presented in section 3.0.9. Performing the experiment with low energy ions presented some unique challenges.

The detectors currently have wire bonds that prevent the membrane from getting close enough to the detector for good alignment. If any part of the membrane holder touches the wire bonds, the bias is effectively lost and single ion detection is no longer possible. The Si$_3$N$_4$ membrane is under tension and this makes it fragile; touching a wire bond would easily destroy the nanostencil. This meant that the closest we were able to get to the substrate was 700 µm. At this distance, a misalignment of the aperture to the detector site of 5° translates to a lateral misalignment of 60 µm, far further than the detector area of 10x10 µm$^2$. While the in situ detection alignment enables us to scan the substrate until we are aligned with the detector, exposing that much of the substrate to ion bombardment is not ideal for the fabrication of devices that are highly sensitive. In the future we need to redesign the detector packaging to allow the membrane to get closer to the surface.

The second major hurdle in nanostencil alignment is the requirement of cooling the detector. The linear thermal expansion of the materials used in the deterministic doping apparatus are shown in figure 5.16 and have been obtained from [523–526]. The expansion (or contraction) is less within the 80-120 K operating window than at room temperature, but it is still significant because we are working at the nm scale. The encoders of the piezo nanopositioners
Figure 5.16: The coefficients of thermal expansion of the main materials used in our apparatus. The dotted line shows the temperature window we are required to keep the detector and JFET within for low noise operation.

will not function at cryogenic temperatures, so the base of the apparatus that holds the positioner stack is thermally isolated from the detector housing. This means the main shrinkage occurs in the copper housing of the detector/JFET. Since it is anchored to the frame at the top, the direction of the contraction is primarily in the vertical axis which moves the detector up, relative to the membrane. The cooling needs to remain off during single ion detection because the vibrations caused by flowing N\textsubscript{2} create noise in the detector. With the cooling switched off, the copper expands as it warms. From observation, the rate of the detector motion is a few hundred nm per Kelvin. Since the length of the slot is about 7 \(\mu\text{m}\), the vertical motion is not a problem because it is relatively slow and can be roughly correlated to temperature. There is also a negligible sideways motion that is consistent over our working temperature range. In the next incarnation of the apparatus, we could increase the thermal mass of the copper housing to slow the warming and thus slow the detector motion. Since the single ion implants can be carried out within a few seconds, we do not expect the fabrication of a three donor array to be much affected by this movement.
Low energy ion signals are collected differently to those in the room temperature, MeV IBIC experiments. The movement of the piezo positioners introduces an additional dead time to the signal since the movement induces both rf noise during movement and a small amount of echoing acoustic noise when it stops. Fortunately for us, there are different options for affecting positioner motion. During typical use the positioner is thrown repeatedly by application of voltage pulses until the encoder is satisfied it has reached the selected destination. This includes jumping back and forth at the end of range until the closest approach is found. This takes a relatively long time and 300-500 ms of noise is picked up by the detector. The alternative is to throw the mechanism just once with a particular voltage; the step size will be roughly the same each time. This only adds noise for the duration of the jump (10-20 ms) so we can prescribe a lean 40 ms dead time. The step sizes achievable with this method range from 10 nm to 1 \( \mu \)m. For scans with step sizes larger than 1 \( \mu \)m, a 600 ms dead time is used to cover the longer working duration of the piezo positioners.\(^1\)

Because of the movement of the detector as it warms during the scan, we have to use a large scanning area which includes a large portion of dark area outside the detector region. Ion irradiation of the outside of the detector can shift the electrostatics of the detector so we avoid this as much as possible by performing the scan on a counted basis rather than timed. The beam current is set such that we expect 4 ions to pass through the aperture each second. This is low enough that each impact is easily distinguished by the data collection system to ensure single ion incidence. In practice we measure the current of the ion beam and make an estimate based on the known geometry of the beam system and the geometry of the membrane aperture. While only the pulse height

\(^1\)Note that the positioners also have a highly accurate but short 2 \( \mu \)m range of movement that utilises piezo extension directly. With this we can reproducibly obtain step sizes from 5 to 2000 nm which we will require for the fabrication of single ion arrays.
forms the nanoIBIC map, the full pulse (voltage-time) data of each charge event is recorded. The shape of the pulse indicates whether it was triggered by an ion (symmetrical) or noise (generally not symmetrical). This doesn’t account for ions that enter with energy below the noise threshold of 1.45 keV, which is estimated to be between 2 and 7% of the ions for the S301 detector with 98±5% counting efficiency. If no counts are detected within 0.5 s, the piezo is moved on to the next position. Once a set number (3-6) of ions are counted, the positioner is moved on to the next step. We record the voltage pulse of each ion, the encoder position and the time it was obtained.

**Results**

Using this approach, we obtained the nanoIBIC map seen in figure 5.17. At the time of this experiment, 14 keV P\(^+\) was unavailable so we used 13.5 keV Ar\(^+\) which has a similar electronic stopping profile to 14 keV P\(^+\) in order to give us qualitatively similar results. Quantitative analysis of the ion charge spectra is not useful.

![Image of nanoIBIC map and microscope image](image_url)

**Figure 5.17:** (a) The 13.5 keV Ar\(^+\) nanoIBIC image of an S301 detector. (b) microscope image of the same detector showing the active 5 nm SiO\(_2\) region.

As we expect, the 14 keV CCE is consistent across the whole detector area. Because of the relative depth that 0.5-2 MeV He ions reach in the detector, there has always been some question of the relevance of the high energy ion CCE to the low energy performance we require. With a resolution of 200 nm, there is no discernible CCE variance across the S301 thin oxide detector area.
This is important for the position feedback mechanism of our deterministic doping method: we can expect the same CCE within a sacrificial, patterned alignment site as at the implant site situated suitably far away. For minimum impact on the quantum manipulation experiments performed on our implanted arrays, we should locate the alignment area as far away as possible. For this reason, it would be interesting to discover if we could replicate the high energy 2 MeV He$^+$ experiments of [414] with 14 keV P$^+$. This MeV experiment observed 100% CCE with the detector electrodes seperated 45 $\mu$m (or less) which would provide ample space for a sacrificial alignment pattern.

The smooth corner shape of the active, thin oxide area can clearly be seen in the nanoIBIC map. In the line traces seen in figure 5.18 for counts with 200 nm steps, the mean energy trace demonstrates the ability to trace the surface contours that we first performed with 2 MeV He$^+$. In this case there is effectively one large trench only and in practice a detailed pattern will be required for high accuracy alignment.
Though we would like to know the fraction of the detected spectra that results from scattered ions, the 13.5 keV Ar\textsuperscript{+} we used only produces roughly 2.8 keV of ionisation so that most scattered ions are below the detection threshold which is 1.45 keV and we cannot calculate a useful number. In figure 5.17 and in other scans not shown, we can distinguish widely scattered ions that have entered the detector while the aperture was positioned outside the detector area. These are the bright pixels that appear in the silent/dark region outside the detector area in the nanoIBIC map. We could calculate a minimum scattering percentage with the data we have but this would not be a useful number. These widely scattered ions will not effect the implant quality because they are scattered far outside the implant zone and would be stopped by the surface PMMA mask which is later removed. It is the forward scattered ions that fall within the implant site that worry us. Ions below a certain energy will be stopped by the oxide. Ions with energies above this threshold but below the detection threshold are the ions that concern us the most since they will enter our devices undetected. With scattering spectra there is usually a continuum of ions ranging from low energy to the peak energy so that we would observe a skewed peak with significant scattering. Additionally, if there were a large percentage of scattered ions, we would see a change in CCE at the edges of the detector where some collimated ions that enter the PMMA are scattered back towards the open detector. Seeing as there is no such feature in our scans, we continue to assume the scattering is not excessive and effects the fabrication of a three donor array less than straggle does.

5.3.4 Summary

We have demonstrated the step and repeat ability of the deterministic doping apparatus by irradiating a sample through an aperture while stepping the nanostencil across with the piezo positioners. The shadow mask implants
5.3.4

did not show significant scattering beyond the confines of the mask which conforms to the straggle theory presented in chapter 2. We are confident that our nanostencils work as intended.

We performed 14 keV nanoIBIC on a detector, the first time this has been done with low energy ions to such a resolution. It shows once and for all that the CCE is constant across the thin oxide area of the detector, for 14 keV ions as well as the 2 MeV He\(^+\) that were used in all previous IBIC testing. In addition, the nanoIBIC map of the detector we have shown represents the capability required of the apparatus for fine alignment.

Though there are experimental limitations still to overcome, we did not see significant scattering in the spectra and this is promising. The resolution of our scanning is sub-100 nm; we will require further work to demonstrate sub-10 nm mapping.

**Future Work**

For further work to be carried out on the deterministic doping system, the detectors need to be revised. We require high quality detectors with a low noise floor in order to measure the proportion of scattered ions that are implanted. These detectors need to be redesigned such that the wire bonds connecting to the top contact on the substrate do not interfere with getting the nanostencil close to the surface of the implant site. A sacrificial alignment pattern needs to be incorporated into the fabrication process flow such as those described earlier in this chapter. Concept drawings for the alignment pattern, detector design and chip mounting as discussed in chapter 3 are shown in figure 5.19.

High energy experiments with 2 MeV He\(^+\) tell us that the PiN detector retains full CCE with the leads either side of a 45 µm gap [414] and with the deterministic doping apparatus, we could repeat these experiments with 14 keV
ions. This would allow us to determine the maximum space in which we can build a sacrificial alignment pattern on the detector. It is necessary to avoid interfering with the integrity of the implant site.

The resolution of the apparatus should be comparable to the scale of the
alignment pattern features and with appropriate detectors we will approach
the sub-10 nm resolution that we require to fulfill the goal of fabricating a
three-donor array for a demonstration of CTAP.
Conclusions

Conventional semiconductor fabrication techniques are rapidly approaching the atomic scale, necessitating precise positional control of single dopants to guarantee device function. This is known as deterministic doping. We have implemented apparatus for building arrays of single dopants within silicon by masking an ion beam with a movable nanostencil and detecting successful single ion entry via induced charge. An important application of this work is the fabrication of proof-of-concept devices for experiments in quantum computing.

In Chapter 1, we provide the context for deterministic doping as a fabrication method by reviewing the history of silicon in the semiconductor industry, including the development of semiconductor theory and fabrication techniques. These were applied most notably to the fabrication of the computer which has consistently improved in performance as transistor density has increased [75]. However, modelling quantum physics on a useful scale is beyond the Boolean logic of the conventional computer which prompted a new information processing paradigm where quantum logic based on the entanglement between $|0\rangle$ and $|1\rangle$ states, affords a vast increase in processing speed and ability. Such a quantum computer will require only 30-40 qubits to out perform current state-
of-the-art classical super-computers made of clusters of thousands of CPUs, each containing a local memory of 17 GB [10].

The various approaches to building a quantum computer are reviewed in the latter part of Chapter 1. We are pursuing the Si:P architecture proposed by Bruce Kane where qubits are formed from the spin state of the outer electron and nuclei of phosphorus atoms in a silicon substrate [18]. Silicon is an ideal candidate on which to base a solid state quantum information processing architecture because experimental methods and fabrication techniques making use of Si are well advanced. Part of the appeal of Si in conventional technology is the willingness with which it accepts various dopants into its crystal structure which enables the engineering of a wide range of properties and the spin 0 of $^{28}\text{Si}$ provides a clean substrate for quantum technologies.

Both the nucleus and donor electron of P have a simple spin $\frac{1}{2}$ and they have long coherence times in Si at liquid helium temperatures and below. Fully functioning single Si:P qubits have been constructed; the experiments exploit both the electron spin and the longer lived nuclear spin (via entanglement with the electron spin) of single P atoms that were implanted into Si [241, 252]. The qubits are initialised and manipulated with RF pulses delivered by on-chip microwave lines and the spin state is read out with the use of sensitive SETs. It took great refinements in fabrication and experimental skill to produce these results and the stage is now set for the arrival of multi-atom and more specifically, multi-qubit devices. In order to produce these devices, we require the ability to place single dopants accurately within the substrate; this is known as deterministic doping.

Our deterministic doping method involves the use of a nanostencil that masks a wide beam in order to define the implantation site. The aperture is maneuvered over a PiN detector with piezo-electric positioners in order to accurately direct ions that are then detected via the charge induced as they enter the
In Chapter 2 we model the fabrication of a deterministically implanted three-donor device and estimate the proportion of arrays that would suitable for demonstrating the coherent tunnelling by adiabatic passage (CTAP) protocol.

As a functional goal for our apparatus, we chose the 3 donor device which is required to demonstrate the simplest form of coherent tunneling by adiabatic passage (CTAP) [22]. CTAP has been proposed as a means of shuttling quantum information along a chain of P donors within the Si:P architecture [180]. In a protocol akin to the optical STIRAP, voltages applied to surface gates control the wave function overlap between donors and suitably adiabatic pulses allow a coherent transfer of the electron from the first donor to the third without ever having been present in the potential of the second donor. CTAP has already been demonstrated with photons in three waveguides [287] but a solid state demonstration across three donors would cement CTAP as a viable transport option for quantum information processing.

To model the fabrication of devices we used the Monte-Carlo software package SRIM [329] which utilises the ion-solid interaction physics outlined in section 2.1 to track the movement of accelerated ions as they slow within a material.

We then applied hydrogenic modelling [400] to assess the feasibility of performing a CTAP protocol on these three donor arrays that varied from the ideal positioning due to straggle. We found that the CTAP transfer should be possible in 1-in-7 donor arrays fabricated with 14 keV P⁺ ions implanted through a 20 nm nanostencil stepped in 30 nm intervals. This shows that ion implantation is a viable fabrication method for making multi-atom proof-of-concept devices with the introduction of deterministic doping.
In Chapter 3 we review the methodology of single ion detection in order to present the considerations that went into designing the new apparatus. The integration of our deterministic doping apparatus with the detection system demands high performance from the detector due to the capacitively introduced noise arising from the close proximity of the piezo driven stencil.

Single ions are detected entering a PiN doped substrate via a charge pulse at the back contact that mirrors the charge induced as the ion collides with the Si lattice [249]. The detector design was changed due to constraints imposed by the experimental conditions of qubit demonstration. The spin state of a Si:P qubit is read out by observing the transfer of a single electron across the device. Charge trapped in the oxide during growth can cause a two-dimensional electron gas (2DEG) to form at the SiO$_2$ interface and this forms a pathway for extraneous electron flow. To prevent this, a channel stopper was introduced to the device and could only be introduced into the fabrication process flow before the single ion implantation occurs, meaning the detector design has changed. Extensive testing of the performance of the new detector designs provided an opportunity to review the detector system and investigate the sources of noise and reduced charge collection in preparation for the deterministic doping apparatus. We examined and compared the noise threshold of all detectors in our inventory and performed ion beam induced charge (IBIC) analysis on each design.

The new detectors were found to have higher noise and have poor charge collection efficiency compared to their predecessors. With careful testing this was shown to be primarily due to the use of high resistivity but lower purity substrates as well as altered mounting of the chips. These issues trump the design of the detector which should still function properly with a channel stopper. These findings provide a pathway to remedy the poor performance of the current detectors and successfully integrate a channel stopper into the design.
Chiefly, high purity substrates are required to avoid losing signal strength to charge recombination within defects and to avoid increased noise from trap-assisted generation currents. Additionally, the custom made chip carriers that were previously in use should be reintroduced. They had contact pins that were more widely separated which introduced less stray capacitance, hence noise, than the pins on the new prepackaged carriers.

The best available detector had a counting efficiency of 98±5% and this was equivalent in both the original single ion detection system and the new deterministic doping apparatus. This means the detection hardware in the new system is functioning correctly.

Chapter 4 concerns the high aspect ratio nanostencil with a 10 nm aperture that is required for masking the ion beam in our deterministic doping apparatus. Constructing such an aperture pushes the boundaries of nano-fabrication, which we review. By employing focussed ion beam (FIB) technology, we were able to construct both circular and slot shaped apertures 20 nm wide.

We fabricated nanostencils in 200 nm Si$_3$N$_4$ membranes by backfilling a larger aperture with the electron assisted deposition of Pt. The dual FIB/SEM system used for this was readily available and the this method requires fewer steps than conventional lithographic techniques. It is also one of the only methods that allows the formation of high aspect ratio apertures. The literature contains many descriptions of this method for forming circular apertures [480–482] however slot apertures as small as 20 nm in a 200 nm thick membrane have not previously been reported in the literature.

TEM imaging of these structures showed the edges to be uniform with even thickness which limits scattering. The key to control was to keep the vacuum high, work in the electron limited regime [1] and rather than raster scanning across the whole aperture as reported in other literature, we deposited Pt in
lines on the edge of the hole or slot which itself must have neat edges originally.
The width of apertures was typically 30 nm and we suspect that at around this width, surface tension of the material increases the rate of closure which makes it difficult to selectively stop the process at 20 or 10 nm.

Chapter 5 demonstrates the successful integration of the multi-step movement of the nanostencil with single ion detection, using novel low-energy nanoIBIC experiments (14 keV P⁺).

The alignment of the nanostencil to the implant site is done in two stages, *ex situ* with a microscope and inside the chamber with the detection system running. Because the pulse strength induced by an ion is proportional to its energy, an ion which has passed through a dead layer of PMMA on the surface will be distinguishable from an ion entering directly through the thin oxide. We demonstrated this technique by scanning 0.5 MeV He across a patterned detector which produced a one-dimensional IBIC image of the surface features. We also performed a 2-dimensional 14 keV nanoIBIC scan of the detector which has not been done before. Crucially, we observed that charge collection efficiency remained consistent across the detector area in agreement with MeV He testing. This also means that a sacrificial alignment pattern can be placed well clear of the target implantation site, thus avoiding unintended ion bombardment.

In our implementation, we have achieved an accuracy of <100 nm, with scope for further refinement. Future work to improve this precision will require properly functioning detectors with low noise thresholds that incorporate necessary new features such as the alignment pattern and channel stopper and bond pads spaced further apart to accommodate the nanostencil. Experiments in novel detection techniques are also ongoing.
Over the course of this project, experiments on Si:P qubits demonstrated that it is a viable option that is still worth pursuing. Deterministic doping is required for the role of ion implantation to be significant in the fabrication of multi atom Si:P devices. We have implemented such an apparatus and demonstrated key concepts for its use. Further refinement will have it working to fabricate arrays of ions with precision.


[239] D. R. McCamey, C. Boehme, G. W. Morley, and J. van Tol. Electrically
detected spin echoes of donor nuclei in silicon. *Physical Review B*, 85:073201,
2012.

[240] Lukas Dreher, Felix Hoehne, Martin Stutzmann, and Martin S. Brandt. Nu-

Zwanenburg, D. N. Jamieson, A. S. Dzurak, and A. Morello. High-fidelity
readout and control of a nuclear spin qubit in silicon. *Nature*, 496:334 – 338,
2013.

A. Ardavann, T. Schenkel, E. E. Haller, J. W. Ager, and S. A. Lyon. Solid-
state quantum memory using the $^{31}$P nuclear spin. *Nature*, 455:1085 – 1088,
2008.

[243] Alexander I. Lvovsky, Barry C. Sanders, and Wolfgang Tittel. Optical quant-

[244] Holger P. Specht, Christian Nolleke, Andreas Reiserer, Manuel Uphoff, Eden
Figueroa, Stephan Ritter, and Gerhard Rempe. A single-atom quantum mem-


subsecond hyperpolarization of the nuclear and electron spins of phosphorus
in silicon by optical pumping of exciton transitions. *Physical Review Letters*,


[255] Stephanie Simmons, Richard M. Brown, Helge Riemann, Nikolai V. Abrosimov, Peter Becker, Hans-Joachim Pohl, Mike L. W. Tewalt, Kohei M. Itoh,


[496] Percy Zahl, Martin Bammerlin, Gerhard Meyer, and Reto R. Schlittler. All-in-one static and dynamic nanostencil atomic force microscopy/scanning tun-


[513] Haiming Guo, David Martrou, Tomaso Zambelli, Jrme Polesel-Maris, Agns Piednoir, Erik Dujardin, Sbastien Gauthier, Marc A. F. van den Boogaard,


