Nanoelectromechanical Relays for Low Power Integrated Circuits

Peter Gleeson
Department of Chemistry
Trinity College Dublin

This dissertation is submitted for the degree of

Doctor of Philosophy

Prof. John J. Boland Group
September 2019
Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university, and it is entirely my own work. I agree to deposit this thesis in the University’s open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Dublin Library conditions of use and acknowledgement.

Peter Gleeson
Acknowledgements

I thank my supervisor, Prof John Boland, for his advice and guidance throughout not only this work, but the many projects I have been engaged in over the years. The work in this thesis involved a large team working in Ireland and various locations in the US. I would like to thank my Intel colleagues, Dr Kevin Lin for his work on analytical and finite element modelling, Dr Patrick Theofanis for his work on molecular dynamics simulations and Dr Chytra Pawashe for his numerical modelling. I thank my Trinity College colleague, Dr Plamen Stamenov, for his generosity with his time and skill in conductive AFM measurements, and I thank Dr Alexis Potie, working between the Boland and Cross groups in Trinity College, for his nanorelay contact performance mapping study.

As an Intel researcher in residence, I thank the AMBER programme team at CRANN for hosting me in the centre, and providing me with such a rich opportunity during my many years here. I acknowledge my former manager, Leonard Hobbs, who encouraged me to undertake a PhD, and Bernie Capraro, Research Manager for Silicon Technology at Intel Ireland, for his ongoing encouragement and support. Thanks also to the Intel research group members, Jennifer, Roger and Matt, and to all former members, particularly Chris, for his friendship down the years, and whose fingerprints can be found on the many “SketchUp” diagrams contained herein. Armin, Riley and Jen-
nifer also left their mark, for which I am grateful.

It has been a real pleasure to do this PhD surrounded by so many talented people in the CRANN Institute, particularly present and former members of the Boland group. Indeed many former colleagues have become personal friends.

I am grateful to have had parents who inspired me to work towards this PhD. My only regret is that they did not live to see its completion. This thesis is dedicated to them.
Abstract

The semiconductor industry roadmap was described by Intel cofounder, Gordon Moore, fifty years ago when he predicted that the number of transistor devices on a silicon integrated circuit would double with every new generation of the technology on a two year cycle [1]. That this exponential growth has been maintained for several decades to the point where a state of the art Si integrated circuit now contains billions of transistors is remarkable, and has transformed the world we live in. The solid state device technologies on which these products have been built now feature dimensions on the order of 10nm, and this aggressive scaling brings new challenges. With growing demand for data volumes and computing power, much of it mobile, leakage currents and power dissipation are key challenges for the industry. While the progress of the technology to date has been evolutionary, with new materials, device designs and circuit topologies, the fundamental nature of the problems confronting the industry now require revolutionary approaches, with completely new devices and technologies being developed to maintain Moore’s law.

The future may well look to the past for inspiration in this regard. In the 1930s, computing was carried out using mechanical relay devices by Alan Turing and others [2]. If mechanical devices can be scaled to be compatible with current state of the art integrated circuits, they may well offer solutions
to the leakage and power performance issues of state of the art solid state technology. In this work, the challenge of incorporating nano scale relays (nanorelays) to complement or replace solid state devices is explored. The specific challenge of integrating these devices at the 22nm technology node is explored using validated models and experimental investigation.
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>PN junction diode</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>Bipolar Junction Transistor</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>MOS transistor in OFF state</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>MOS transistor in ON state</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>MOS transistor with electron energy distribution function</td>
<td>18</td>
</tr>
<tr>
<td>2.4</td>
<td>Transistor transfer characteristic</td>
<td>19</td>
</tr>
<tr>
<td>2.5</td>
<td>Scaling and leakage in CMOS</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>CMOS circuit configuration</td>
<td>22</td>
</tr>
<tr>
<td>2.7</td>
<td>Cantilever beam (CB) type nanorelay device</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Nanorelay schematic</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>Balance of forces in nanorelay</td>
<td>32</td>
</tr>
<tr>
<td>3.3</td>
<td>Hysteresis in nanorelay operating characteristic</td>
<td>33</td>
</tr>
<tr>
<td>3.4</td>
<td>Multi terminal relay device</td>
<td>35</td>
</tr>
<tr>
<td>3.5</td>
<td>Schematic of four terminal relay</td>
<td>38</td>
</tr>
<tr>
<td>3.6</td>
<td>Schematic of relay with 4nm nanogap</td>
<td>38</td>
</tr>
<tr>
<td>3.7</td>
<td>Nanorelay with modified beam</td>
<td>39</td>
</tr>
<tr>
<td>3.8</td>
<td>Nanorelay with composite beam</td>
<td>40</td>
</tr>
<tr>
<td>3.9</td>
<td>Nanorelay with large area actuation paddle</td>
<td>41</td>
</tr>
<tr>
<td>3.10</td>
<td>Vertical cantilever beam device</td>
<td>44</td>
</tr>
<tr>
<td>Number</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.11</td>
<td>Integrated circuit incorporating nanorelays</td>
<td>45</td>
</tr>
<tr>
<td>3.12</td>
<td>Nanorelay operating characteristic</td>
<td>49</td>
</tr>
<tr>
<td>4.1</td>
<td>Schematic and micrographs of nanorelay array</td>
<td>56</td>
</tr>
<tr>
<td>4.2</td>
<td>Fabrication process flow</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Zeiss Gemini electron optical column</td>
<td>59</td>
</tr>
<tr>
<td>4.4</td>
<td>Micrograph of sub 30nm wide beam</td>
<td>62</td>
</tr>
<tr>
<td>4.5</td>
<td>Origin of DC bias in plasma etch process</td>
<td>63</td>
</tr>
<tr>
<td>4.6</td>
<td>Si etch profile</td>
<td>64</td>
</tr>
<tr>
<td>4.7</td>
<td>Etching / polymerisation process space</td>
<td>66</td>
</tr>
<tr>
<td>4.8</td>
<td>Nanocantilever collapse during drying process</td>
<td>68</td>
</tr>
<tr>
<td>4.9</td>
<td>Schematic of Atomic Force Microscope</td>
<td>70</td>
</tr>
<tr>
<td>4.10</td>
<td>Schematic of stiffness measurement experiment</td>
<td>71</td>
</tr>
<tr>
<td>4.11</td>
<td>Force displacement characteristic</td>
<td>72</td>
</tr>
<tr>
<td>4.12</td>
<td>F-d map showing measurement locations on beam</td>
<td>74</td>
</tr>
<tr>
<td>4.13</td>
<td>Stiffness versus beam length simulation and measurement</td>
<td>75</td>
</tr>
<tr>
<td>4.14</td>
<td>Cross sectional images of beams</td>
<td>76</td>
</tr>
<tr>
<td>4.15</td>
<td>Moment of inertia for ideal and realistic beams</td>
<td>79</td>
</tr>
<tr>
<td>4.16</td>
<td>FEA model of realistic beams</td>
<td>81</td>
</tr>
<tr>
<td>4.17</td>
<td>Beam stiffness pre and post oxide strip</td>
<td>82</td>
</tr>
<tr>
<td>5.1</td>
<td>Trigate transistor at 22nm node</td>
<td>88</td>
</tr>
<tr>
<td>5.2</td>
<td>Transfer characteristic at 22nm node</td>
<td>89</td>
</tr>
<tr>
<td>5.3</td>
<td>Device used for numerical model</td>
<td>91</td>
</tr>
<tr>
<td>5.4</td>
<td>Dimensional dependence of pull-out force</td>
<td>92</td>
</tr>
<tr>
<td>5.5</td>
<td>Dimensional dependence of resonance frequency</td>
<td>94</td>
</tr>
<tr>
<td>5.6</td>
<td>CMOS / nanorelay four input AND gate</td>
<td>96</td>
</tr>
<tr>
<td>5.7</td>
<td>Schematic of 2/3/4 terminal nanorelay</td>
<td>97</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8</td>
<td>UNCD and NCD AFM probe chips</td>
<td>101</td>
</tr>
<tr>
<td>5.9</td>
<td>Low noise DC measurement circuit</td>
<td>101</td>
</tr>
<tr>
<td>5.10</td>
<td>AFM head with preamplifier</td>
<td>102</td>
</tr>
<tr>
<td>5.11</td>
<td>Substrate mount with AFM probe chip holder</td>
<td>103</td>
</tr>
<tr>
<td>5.12</td>
<td>F-d curve: UNCD - Si</td>
<td>105</td>
</tr>
<tr>
<td>5.13</td>
<td>Damage to Si through repeated F-d measurement</td>
<td>106</td>
</tr>
<tr>
<td>5.14</td>
<td>F-d curve: NCD - Si</td>
<td>107</td>
</tr>
<tr>
<td>5.15</td>
<td>Simulated &amp; measured adhesion pressure for various materials</td>
<td>110</td>
</tr>
<tr>
<td>5.16</td>
<td>Relay design space</td>
<td>111</td>
</tr>
<tr>
<td>6.1</td>
<td>Nanorelay failure mechanisms</td>
<td>119</td>
</tr>
<tr>
<td>6.2</td>
<td>Encapsulated device in BEOL</td>
<td>120</td>
</tr>
<tr>
<td>6.3</td>
<td>F-d characteristic showing repulsive behaviour</td>
<td>125</td>
</tr>
<tr>
<td>6.4</td>
<td>Cantilever beam with mechanical stop</td>
<td>126</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Truth table: AND ........................................ 4
1.2 Truth table: OR ........................................ 4
1.3 Truth table: NOT ....................................... 4

4.1 EBL process parameters ................................. 62
4.2 ICP plasma etch process parameters ................. 67

5.1 Critical dimensions at the 22nm technology node. 90
5.2 Nanorelay device specification ....................... 99
## Abbreviations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Polarisability</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Dielectric Permittivity</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Surface Energy</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck’s Constant</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Electron Mean Free Path</td>
</tr>
<tr>
<td>$\mu$</td>
<td>mobility</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Poisson Ratio</td>
</tr>
<tr>
<td>$\omega(r)$</td>
<td>van der Waals interaction</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Density</td>
</tr>
<tr>
<td>$AFM$</td>
<td>Atomic Force Microscope</td>
</tr>
<tr>
<td>$BEOL$</td>
<td>Back End Of Line</td>
</tr>
<tr>
<td>$BOE$</td>
<td>Buffered Oxide Etch</td>
</tr>
<tr>
<td>$BOX$</td>
<td>Buried Oxide</td>
</tr>
<tr>
<td>$c \sim AFM$</td>
<td>conductive - Atomic Force Microscope</td>
</tr>
</tbody>
</table>
CB  Cantilever Beam

CMOS  Complementary Metal Oxide Semiconductor

CRT  Cathode Ray Tube

CVD  Chemical Vapour Deposition

DAC  Digital to Analogue Convertor

DI  DeIonised

DLC  Diamond Like Carbon

EBL  Electron Beam Lithography

EsB  Energy selective Backscatter

FEA  Finite Element Analysis

FEG  Field Emission Gun

FIB  Focused Ion Beam

FinFET  Fin Field Effect Transistor

FPGA  Field Programmable Gate Array

FSD  Full Scale Deflection

GND  Ground

HDP  High Density Plasma

HOPG  Highly Ordered Pyrolytic Graphite

HVM  High Volume Manufacturing
ABBREVIATIONS

ICP  Inductively Coupled Plasma

ICT  Information and Communications Technology

IPA  IsoPropyl Alcohol

ITRS International Roadmap for Semiconductors

LMIS Liquid Metal Ion Source

MD  Molecular Dynamics

MEMS Micro ElectroMechanical Systems

MOS Metal Oxide Semiconductor

MOSFET Metal Oxide Semiconductor Field Effect Transistor

NCD NanoCrystalline Diamond

NEMS Nano ElectroMechanical Systems

NMOS N type Metal Oxide Semiconductor

OFN Oxygen Free Nitrogen

PMOS P type Metal Oxide Semiconductor

PolySi PolySilicon

PPE Personal Protective Equipment

RF Radio Frequency

RIE Reactive Ion Etch

SE Secondary Electron
ABBREVIATIONS

SEM Scanning Electron Microscope
Si Silicon
SOI Silicon On Insulator
SRAM Static Random Access Memory
STM Scanning Tunnelling Microscope
STP Standard Temperature and Pressure
TEM Transmission Electron Microscope
TEOS LR Tetra Ethyl Ortho Silicate Low Rate
TEOS TetraEthylOrthoSilicate
TMD Transition Metal Dichalcogenides
TTL Transistor Transistor Logic
UNCD Ultra NanoCrystalline Diamond
# Contents

1 Introduction ........................................ 1
   1.1 Overview ........................................ 1
   1.2 Digital Computing ................................ 2
   1.3 Device Physics .................................. 4
       1.3.1 The semiconductor pn junction diode .... 5
       1.3.2 The bipolar transistor ..................... 9
       1.3.3 The MOS transistor ......................... 10
       1.3.4 The nanorelay ............................... 11

2 Motivation .......................................... 15
   2.1 Introduction .................................... 15
   2.2 Metal Oxide Semiconductor (MOS) Technology . 15
   2.3 CMOS - State of the Art and Challenges .......... 20
   2.4 Nanorelays for Reduced Leakage .................. 23
   2.5 Thesis Overview .................................. 25

3 Theory ............................................. 29
   3.1 Mechanics of Cantilever Beam Nanorelays .......... 29
   3.2 NEMS Devices: State of the Art .................. 34
   3.3 Circuits incorporating NEMS: State of the Art .... 42
3.4 NEMS / Nanorelay Technology: Challenges .......................... 47
  3.4.1 Miniaturisation .................................................. 47
  3.4.2 Adhesion .......................................................... 48
  3.4.3 Contact reliability ................................................. 51

4 Silicon Nanocantilevers .................................................. 53
  4.1 Introduction ......................................................... 53
  4.2 Review of Beam Mechanics ......................................... 54
  4.3 Design & Fabrication of Nanocantilever Array .................. 55
  4.4 Processing .......................................................... 57
    4.4.1 Application of (electron beam) resist ....................... 57
    4.4.2 Exposure and development .................................... 57
    4.4.3 Dry plasma etch .............................................. 62
    4.4.4 Wet release etch ............................................. 66
    4.4.5 Resist strip ................................................... 68
  4.5 Experimental Measurement of Stiffness ............................ 69
  4.6 Experimental Results and Analysis ................................ 73
  4.7 Conclusions ........................................................ 84

5 Nanorelays at the 22nm node .......................................... 85
  5.1 Introduction ........................................................ 85
  5.2 The 22nm Technology Node ........................................ 88
  5.3 Device Specifications ............................................... 90
    5.3.1 Mechanical specifications .................................... 90
    5.3.2 Electrical specifications ..................................... 96
  5.4 Experimental Approach ............................................. 100
  5.5 Results ............................................................ 104
    5.5.1 The (U)NCD - Si dimple - source material system .......... 104
CONTENTS

5.5.2 Molecular dynamics simulations .................................. 108
5.5.3 Performance mapping of candidate material combina-
tions ............................................................................. 109
5.6 Conclusions ................................................................... 113

6 Possible Future Directions .................................................. 115
  6.1 Introduction ................................................................... 115
  6.2 The Adhesion and Electrical Challenge .............................. 116
  6.3 Device Fabrication .......................................................... 117
  6.4 Reliability ...................................................................... 118
  6.5 Repulsive Behaviour in Liquids ........................................... 120
  6.6 Novel Device Design ...................................................... 126
  6.7 Alternative Materials ..................................................... 127

7 Summary and Conclusions .................................................... 131

Appendix .............................................................................. 137

A Processing Details .............................................................. 137
  A.1 Application of Photoresist ................................................. 137
  A.2 Electron Beam Lithography .............................................. 137
    A.2.1 Electron beam exposure ........................................... 137
    A.2.2 Development .......................................................... 138
  A.3 Plasma Etch ................................................................. 138
    A.3.1 Silicon plasma etch .................................................. 138
    A.3.2 Silicon dioxide plasma etch ...................................... 139
  A.4 Wet Release Etch .......................................................... 139
  A.5 Removal of Electron Beam Resist ...................................... 139
    A.5.1 Barrel asher process ................................................. 139
Chapter 1

Introduction

1.1 Overview

Information and communications technology (ICT) has revolutionised life across the globe. In an increasingly connected world, diverse facets of life such as transport, healthcare, education and the workplace rely on computing power. The ubiquity of this technology is evident, for example, in the proliferation of internet-enabled smartphones used by billions of people all around the world. This has been enabled by Moore’s Law which has led to extremely powerful processors containing billions of transistors. The move from PC and laptop based computing to mobile computing has given rise to cloud computing, where the compute function is carried out in large server farms with data transmitted between these server farms (the “cloud”), and processors in smartphones or embedded controllers in cars or smart televisions, for example.

All these devices, telecommunications networks and servers consume power. In a research paper, Fehske et al [3] quantify the carbon footprint of mobile technology in the context of its ecological and economic significance, and pre-
dict that emissions will rise to 235 MTo by 2020. Indeed, taking in the wider scope of ICT, Fettweiss et al [4] note that power consumption is increasing by 16-20% per annum. While the entire system is complex (the deployment of ICT in many sectors of industry has led to major gains in efficiency), it is clear that the power being consumed by ICT is a problem in the context of climate change.

In parallel with Moore’s Law, efficiency gains have been made according to Koomey’s law [5] which describes how technological advances have resulted in a doubling of the number of computations per joule every 1.5 years (though this has slowed down more recently). However, Complementary Metal Oxide Semiconductor (CMOS) technology, on which processors have been built for the past few decades, is now facing a fundamental limit in efficiency. To retain Moore’s Law as the guiding principle of the industry, new power efficient approaches are required. Taking inspiration from the early days of computing, relay technology, scaled to the nano regime, may well hold the solution.

1.2 Digital Computing

Information and communications technology driven by advances in computing has revolutionised society over many decades. In the field of computing, information, whether voice, text or video, is represented in digital form, consisting of binary digits or bits, with a value of one or zero. While the decimal system is ubiquitous in everyday life, the binary system, with just two values, lends itself to computing where data can be processed in binary form (for example, the number eight in decimal, “8”, can be represented in binary as “1000”).
Taking voice as an example, the analogue signal from a sensor (microphone) is sampled, and the amplitude of the signal at each sample is measured and recorded in binary form. In this way, the signal can be represented by a series of digits having either of just two values (zero or one). According to the sampling theorem [6], an analogue signal can be recovered from a sampled version of the signal if the sampling frequency is at least double the highest frequency component of the original signal. Hence an analogue signal can be represented by a digital signal by sufficiently fast sampling and subsequent analogue to digital conversion. In digital electronics, each binary digit or bit is represented by charge, where the presence of charge (voltage) represents a “1”, and the absence of charge represents a “0”. This binary signal is memorised or processed by utilising transistors as switches where they can be switched to the power line to represent a “1” or to ground (GND) to represent a “0”. In addition to the memory function, transistors can be connected together to perform numeric algebra / calculation based on this binary system. Such transistor circuits can also be used to implement Boolean logic developed by George Boole [7], [8], which utilises two states, “True” and “False”. There are three primary operators, AND, OR and NOT. These functions are described in so called truth tables below. This use of binary coding to perform logical functions was developed by mathematician and electrical engineer C. E. Shannon, and is the basis for all digital computing [9].

The complexity of computing, which underpins so many aspects of daily life, requires billions of transistors connected together on a single integrated circuit. This integration of electronic devices gave the industry leader, Intel, its name (from integrated electronics). The first integrated circuit was demonstrated by Jack Kilby at Texas Instruments in 1958 [10], and Robert
CHAPTER 1. INTRODUCTION

Table 1.1: Truth table for AND: output is true if inputs A AND B are true.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>A.B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1.2: Truth table for OR: output is true if inputs A OR B are true.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>A+B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1.3: Truth table for NOT: output is inverse or opposite (NOT) of input.

<table>
<thead>
<tr>
<th>A</th>
<th>(\overline{A})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Noyce [11], then of Fairchild Semiconductor, demonstrated the first Si integrated circuit using planar technology, and both he and Gordon Moore went on to found Intel [12]. At this time, Gordon Moore’s prediction of scaling became the business model for the industry, and has set a rapid pace which has been maintained to this day. Moore’s Law, as it is known, describes how the transistor density on the Si integrated circuit doubles with every new technology generation on a 2-3 year cycle [1]. Having followed this model for five decades, integrated circuits now contain billions of transistors on a single silicon chip [13].

1.3 Device Physics

Metal semiconductor junctions have been investigated since the nineteenth century. The rectifying nature of the junction was discovered by Braun in 1874 [14]. By the 1930s, these rectifying diodes were used in radio detectors. This was given some impetus by the development of radar, but it was the
1.3. DEVICE PHYSICS

subsequent understanding of semiconductor device physics that ultimately led to the development of transistors (first on polycrystalline Ge, and later on crystalline Si), and the integrated circuit.

A description of a simple two terminal diode is given below, and this is followed by a description of bipolar and MOS transistors. The challenges for CMOS technology as the industry continues along the roadmap described by Moore’s Law, is then reviewed, and nanorelay technology is presented as a potential solution.

1.3.1 The semiconductor pn junction diode

In this section, the basic physics of a p-n junction diode is explained. Fig. 1.1 shows a diode formed by connecting p type and n type Si together. The n type Si is doped with phosphorus which has five valence electrons. Four of these electrons are used to form bonds with neighbouring Si atoms, while the fifth electron is liberated and moves freely through the material. The positively ionised donors together with the electrons (in red) can be seen in Fig. 1.1. Similarly, the p type Si is doped with boron which has three valence electrons which bond with neighbouring Si atoms, thereby creating a hole (the negatively charged ionised acceptors together with the liberated holes can also be seen in Fig. 1.1. Taking a typical solar cell as an example, the concentration of electrons in the n type Si is equal to the dopant concentration, \( N_D = 10^{18}\text{cm}^{-3} \), while the concentration of holes is \( 10^2\text{cm}^{-3} \). On the p type side of the junction, the hole concentration is \( N_A = 10^{16}\text{cm}^{-3} \), while the concentration of electrons is \( 10^4\text{cm}^{-3} \) (again taking typical values for a solar cell). The concentration gradient of electrons drops from \( 10^{18}\text{cm}^{-3} \) in the n type side of the junction to \( 10^4\text{cm}^{-3} \) on the p type side, giving rise to a diffusion current described by Eqn. 1.1. The hole concentration gradient
Figure 1.1: (a) PN junction diode. Note the negatively charged ionised donors in the p type Si and positively charged donors in the n type Si. In the depletion region, there are no carriers. The electric field, shown in (b), is zero outside the central depletion region, and there is an electric field and built in potential caused by the recombination of minority carriers in this region. At equilibrium, the built in potential causes a drift current which compensates the diffusion current of electrons from the n type Si to the p type Si, and holes from the p type Si to the n type Si. The applied bias reduces this built in potential in forward bias, and current increases exponentially. In reverse bias, the applied voltage acts to increase the built in potential thus reducing this current.
1.3. DEVICE PHYSICS

(from $10^{16}\text{cm}^{-3}$ on the p type side to $10^2\text{cm}^{-3}$ on the n type side) also gives rise to a diffusion current given by Eqn. 1.2. The minority carriers recombine with majority carriers on either side of the junction, that is to say, minority carriers diffusing to the p side (electrons) combine with majority carriers (holes) and are no longer liberated. Conversely, on the n type side, the holes which are the minority carriers combine with electrons which are the majority carriers, as shown in Fig. 1.1. These areas immediately adjacent to the metallurgical junction are therefore depleted of charge, and this region is referred to as a depletion region or space charge region. The redistribution of charge to either side of the region gives rise to an electric field, and the field causes a drift current in the opposite direction to the diffusion current given by Eqn. 1.3 for electrons and Eqn. 1.4 for holes. At thermal equilibrium, the drift and diffusion currents are equal, and the resulting net current is zero.

\[ J_{e(\text{diff})} = qD_e \frac{dn}{dx} \]  
\[ J_{h(\text{diff})} = -qD_h \frac{dp}{dx} \]  
\[ J_{e(\text{drift})} = nq\mu_e E \]  
\[ J_{h(\text{drift})} = pq\mu_h E \]

where $D_e$ and $D_h$ are the electron and hole diffusion coefficients, $q$ is the electronic charge, $n$ and $p$ are the electron and hole concentrations and $\frac{dn}{dx}$ and $\frac{dp}{dx}$ are the electron and hole concentration gradients. Looking at the bandstructure in Fig. 1.1, the electric field in the depletion region that gives rise to the drift current is labelled as the built in voltage, $V_{bi}$. In a pn junction diode, there are electrodes on the p type Si (anode) and the n type Si (cathode). When an applied positive bias is applied to the anode with respect to the cathode, the superposition of the applied voltage and the built
in voltage causes a reduction in the built in voltage. At this point the balance
between the diffusion and drift currents at thermal equilibrium is disturbed,
with the drift current no longer able to compensate the diffusion current.
This causes an increase in minority carrier concentration on either side of
the barrier. The quasi neutral regions (outside the depletion region) now lose
their neutrality as the ionised donors are not compensated by the majority
carriers (lost to recombination). Neutrality is restored by the external circuit
which provides a measurable current to balance the diffusion current that is
not compensated by the drift current, and the recombination current. As the
applied bias is increased, the depletion region will shrink and the electric field
will decrease leading to an exponential increase in the (forward bias) current
flowing through the diode. Under reverse bias (i.e. with a negative bias
applied to the p type Si with respect to the n type Si), the applied voltage is
dropped across the depletion region (as the quasi neutral regions outside the
depletion region are conductive). The sum of $V_{bi}$ and the applied bias now
increases the barrier, and gives rise to a potential barrier which minimises
diffusion of majority carriers across the junction. While this voltage does
favour the drift current based on minority carriers across the depletion region,
recall that the minority carrier concentration is low (a hole concentration of
$10^2\text{cm}^{-3}$ in the n type Si and an electron concentration of $10^4\text{cm}^{-3}$ in the p
type Si). The resulting current is, therefore, much smaller than the forward
bias current, and rectifying behaviour is seen, where the diode conducts under
(sufficiently high) forward bias, while negligible current flows under reverse
bias conditions.
1.3. DEVICE PHYSICS

Figure 1.2: Schematic and circuit symbol for an npn transistor. There are three terminals, the emitter connected to the n++ region, the base connected to a thin lightly doped p type region, and a collector connected to an n+ doped region.

1.3.2 The bipolar transistor

The bipolar transistor was invented by Bardeen, Brattain and Shockley in 1947 [15], [16]. The name comes from “transfer” and “resistor”, and the device acts to transfer a signal from low to high resistance. To understand how the device operates, consider the “npn” transistor in Fig. 1.2. The npn bipolar transistor consists of three regions, an n type emitter, a thin p type base, and larger n type collector. Though there are various configurations depending on the circuit / use, the base emitter voltage can be thought of as the input which controls the collector emitter current (the output). The voltage on the base puts the base emitter junction in to forward bias, giving an influx of electrons in to the lightly doped base region, where they are minority carriers. These electrons diffuse from the base emitter junction to the (reverse biased) collector base junction, where they drift in to the
CHAPTER 1. INTRODUCTION

collector region under the electric field arising from the collector base reverse bias. The base region is made thin so that the minority carriers can diffuse in less time than the minority carrier lifetime. Note also in Fig. 1.2 that the doping in the base is kept low in order to minimise carrier loss through recombination. The device can operate as an amplifier (when the emitter base junction is forward biased and the collector base junction is reverse biased as described, or as a logic switch. In the latter case, the transistor is considered to be “ON” when the device is in saturation (both junctions forward biased), or “OFF” when the device is in the cut off region with both junctions reverse biased. Early integrated circuits were built using these bipolar transistors in transistor transistor logic (TTL) [17]. In particular the 7400 series developed by Texas Instruments [18] became an industry standard, with compatible products offered by all the major industry players of the time. In comparison to CMOS technology, built on MOS devices, TTL logic consumed significantly more power, particularly when not switching.

1.3.3 The MOS transistor

Unlike the bipolar transistor, the metal oxide semiconductor (MOS) transistor conducts current using either electrons (NMOS) or holes (PMOS). Like the bipolar transistor, it is a three terminal device, with current flow between the source and drain controlled by a third gate terminal. The gate terminal is electrically isolated from the source and drain terminals, and the channel region between them, and device operation is based on modulation of the conductivity of the channel by an electric field arising from a voltage signal on the gate terminal. This principle was first disclosed by Julius Edgar Lilienfeld in his patents from the 1930s [19], [20]. The MOS transistor is, therefore, a field effect device. Detailed operation of the MOS device is
described in Chapter 2. The key advantage of the MOS transistor over the bipolar transistor can be appreciated by comparing its performance at circuit level. As noted above, TTL logic is not highly energy efficient. When MOS transistors are used in integrated circuits, the most common implementation is CMOS or complementary MOS. While a more detailed description follows in Chapter 2, it is sufficient to note that in the CMOS configuration, there is essentially no current flowing when the circuit is not switching from one logic state to the other, leading to a dramatic improvement in energy efficiency. While CMOS has been the standard technology for several decades, even while aggressively scaling from the micro regime to the nano regime, it has recently started to face a fundamental energy crisis, to be described in Chapter 2.

1.3.4 The nanorelay

The power consumed by the MOS transistors in a CMOS circuit consists of both dynamic power, which scales with the supply voltage, $V_{DD}$, when the transistor is switched ON, and static power when the transistor is switched OFF. To reduce the dynamic power consumption, the supply voltage, $V_{DD}$, can be reduced. However, to maintain the drive current, $I_{ON}$, and thus circuit speed, while scaling the supply voltage, the voltage at which the device switches on or threshold voltage, $V_{TH}$, must also be reduced by the same amount given the dependence of $I_{ON}$ on $V_{DD} - V_{TH}$ (known as the gate overdrive) according to Nathanael [21]:

$$I_{ON} = \mu_{ef} C_{ox} \frac{W}{L} (V_{DD} - V_{TH})^2$$  \hspace{1cm} (1.5)

where $\mu_{ef}$ is the effective mobility in the channel, $C_{ox}$ is the gate oxide capacitance per unit area, and $W$ and $L$ are the gate width and length.
respectively. It is important to maximise this drive current, to avoid a reduction in circuit speed due to slower charging and discharging of the parasitic load capacitance of subsequent logic gates. However, scaling the threshold voltage causes an exponential increase in the leakage current, Eqn. 1.6 [21], due to non abrupt switching which will be explained in detail in Chapter 2.

\[ I_{OFF} \propto 10^{-\frac{V_{TH}}{SS}} \] (1.6)

with SS, the subthreshold swing, given by

\[ SS = \ln(10) \frac{kT}{q} \left( 1 + \frac{C_{dep}}{C_{ox}} \right) \] (1.7)

where \( \frac{kT}{q} \) is the thermal voltage and \( 1 + \frac{C_{dep}}{C_{ox}} \) is a non ideality factor which accounts for the capacitive voltage divider effect of the capacitance arising from the depletion region in the semiconductor channel, \( C_{dep} \), and the gate oxide capacitance, \( C_{ox} \).

Nanorelays offer a solution to this fundamental energy crisis for CMOS. The nanorelay device utilises an electrostatic force to draw a movable beam across an air gap and in to contact with an opposing contact, thus closing the switch. When the electrostatic force (i.e. control voltage) is released, the spring restoring force causes the movable beam to move back to the neutral position and out of contact. In this way, the switching characteristic is abrupt, thus resolving the CMOS leakage issue. Furthermore, once the beam is out of contact, the air gap (which is designed to be greater than 2nm to prevent tunnelling) essentially eliminates OFF state leakage. From an energy efficiency perspective, therefore, nanorelay technology is compelling. However, there are a number of challenges to resolve as nanorelays are scaled to state of the art CMOS dimensions.
The application of microtechnology developed by the semiconductor industry to the fabrication of devices which feature both electrical and mechanical functionality is not new. Indeed, the first accelerometer was developed in Stanford University in 1979 by Roylance et al [22], and microelectromechanical devices (MEMS) now find widespread use as accelerometer based airbag sensors in the automotive industry. To compete with CMOS, however, devices would need to be aggressively scaled to the nano regime, and this brings many challenges. If these challenges can be satisfactorily addressed, it is likely that nanorelays will complement rather than replace CMOS technology, particularly in applications such as power gating [23], where energy efficiency is critical.

A number of these challenges were addressed in this work. Understanding the mechanical behaviour of nanoscale cantilever beams, and how beam stiffness scales is important, as both the pull-in voltage and the ability of the beam to overcome stiction force is determined by the beam stiffness. In this work, the impact on stiffness of scaling single crystal Si cantilever beams to the nanoscale was investigated. Arrays of nanoscale cantilever beams were fabricated with systematically varying lengths and widths. Detailed measurement of physical dimensions and composition were made, and used to validate analytical models. Due to the departure from ideal rectangular shapes with well defined dimensions [24], [25], finite element analysis (FEA) was undertaken, and these models were validated by experiment.

In addition, the challenge of intercepting the industry roadmap at the 22nm technology node was used to define realistic specifications for nanorelays at these dimensions [26]. Experimental work was carried out to investigate a number of material systems in terms of meeting both mechanical and electrical target specifications. It was found that while mechanical (pull-
out / adhesion) and conductivity specifications could be achieved separately, there was no relay design space identified where both specifications could be met simultaneously. A novel idea based on overcoming adhesion force by immersing the system in a carefully chosen medium was proposed, together with preliminary experimental work. This and other potential solutions are presented as future work.
Chapter 2

Motivation

2.1 Introduction

In 1968, Gordon Moore, one of the cofounders of Intel, predicted that the number of transistors in a given area of Si would double with each new technology node, on a two year cycle [1]. After 50 years of relentless miniaturisation, his prediction still holds. Indeed, it has become the business model for the semiconductor industry, where new products with ever increasing functionality can be delivered through increased device density, thus generating profits to fund a virtuous cycle of innovation and growth.

2.2 Metal Oxide Semiconductor (MOS) Technology

For at least three decades, the workhorse device of the semiconductor industry has been the metal oxide semiconductor system, shown in Fig. 2.1. This is a three terminal device, with current flowing from source to drain controlled by bias applied to the gate. To understand the operation of the
device, consider the MOS transistor in Fig. 2.1. The current in this device is carried by electrons (hence NMOS). The device is built in a p type well, with n+ implants in the source and drain regions. The p type region between the n type source and drain is called the channel, and this is electrically isolated from the metal gate above by the gate oxide. This gate stack gives the device its name - Metal Oxide Semiconductor (MOS). With positive bias on the drain and the source at ground, the band structure is shown in Fig. 2.1. Note that there is an energy barrier for electrons at the source to channel junction. This reverse biased junction prevents current flow. With the application of positive bias to the gate as shown in Fig. 2.2, the barrier to current 

**Figure 2.1:** Metal Oxide Semiconductor (MOS) transistor schematic (top) showing source terminal connected to n+ region, gate terminal over channel region separated by SiO₂, and drain terminal connected to drain n+ region and the associated bandstructure (below). Note how the built in voltage described in Chapter 1, \( V_{bi} \), between the source and channel, acts to minimise electron flow from the source to the channel.
flow is reduced. The source to channel barrier is first driven to flatband condition, and ultimately an inversion layer is formed where the channel region becomes n type (with the same absolute carrier concentration as in the p type well). This is referred to as the threshold voltage, and the transistor is now considered to be in the ON state with current flowing from source to drain. In reality, the transition from the OFF state ($V_{GS} < V_{TH}$) to the ON state ($V_{GS} \geq V_{TH}$) is not abrupt.

The carriers in the source region are described by the Boltzmann approximation of the Fermi probability function [27], [28] given by

$$f(E) = e^{\frac{E_f - E}{kT}}$$

(2.1)

where $f(E)$ is the probability of occupancy by an electron, $E$ is the electron energy, $E_f$ is the Fermi energy, $k$ is the Boltzmann constant and $T$ is absolute temperature. Note that there is a high energy tail, where some electrons have sufficient energy to be thermionically emitted over the barrier as the barrier
height is reduced by $V_{GS}$, even before the threshold voltage is reached. This can be seen in Fig. 2.3, where the high energy tail is clearly higher in energy than the barrier. Electrons in this high energy tail give rise to a leakage current preventing an abrupt switching characteristic. Fig. 2.4 shows a plot of source drain current on a log scale against the gate voltage on a linear scale where this subthreshold current is clear. The inverse slope of this plot

---

**Figure 2.3:** Electron energy distribution function showing high energy tail (labelled thermal tail) with energy exceeding the source - gate energy barrier. Electrons in this high energy tail give rise to subthreshold current (adapted from Tehrani, nanoHub.org).
in the subthreshold region is given by

\[
\left( \frac{kT}{q} \right) \ln(10) = > 60 \text{mV/decade}
\]  \hspace{1cm} (2.2)

At room temperature, the slope is 60mV / decade. In other words, to reduce the source drain current by one order of magnitude, the gate voltage must be reduced by at least 60mV. Through innovation such as high permittivity gate dielectric materials and non planar fin field effect transistor devices (FinFETs) \[29\], capacitive coupling of the gate to the channel has been enhanced to reduce the subthreshold slope, but the slope is fundamentally limited to 60mV/dec. This limitation on the subthreshold slope represents a major constraint for the industry for a number of reasons. To optimise performance, the drive current, \(I_{\text{ON}}\), (source drain current with maximum gate signal ap-

![Figure 2.4: Drain current as a function of gate voltage showing subthreshold slope.](image-url)
CHAPTER 2. MOTIVATION

plied) should be maximised, and the OFF (leakage) current, $I_{OFF}$, should be minimised. Maximising $I_{ON}$ is important as this current charges / discharges the parasitic capacitance of the following gate(s), and thus determines the speed performance of the circuit. To turn on a transistor in a circuit, the full $V_{DD}$ signal is applied to the gate. Although the transistor is in the ON state once the threshold voltage is reached, this “gate overdrive” ($V_{DD} - V_{TH}$) increases $I_{ON}$ as the carrier concentration in the channel increases beyond inversion as the gate voltage is increased, leading to higher $I_{ON}$, see Fig. 2.4. It might therefore be concluded that to maximise $I_{ON}$, the threshold voltage should be reduced as much as possible to maximise $V_{DD} - V_{TH}$. However, for reasons of power conservation, $I_{OFF}$ must be minimised. As outlined above, the transistor does not switch off abruptly below $V_{TH}$. In the best case scenario, the rate of current reduction is limited to 60mV/dec. Therefore, as $V_{TH}$ is decreased to maximise the drive current, the leakage current increases exponentially as shown in Fig. 2.5

2.3 CMOS - State of the Art and Challenges

The significance of this leakage current on circuit performance stems from the use of MOS transistors in a complementary manner in Complementary Metal Oxide Semiconductor (CMOS) technology, which was invented by Frank Wanlass in 1963 [30]. An example is shown in Fig. 2.6. The basis for the binary representation and processing of data in computing was described in Chapter 1, and this computing has been based on CMOS technology for several decades. The reason why can be seen in the basic CMOS circuit, which consists of two MOS transistors connected in series across the power rail. The upper transistor is a PMOS device, where the current is
To improve energy performance, $V_{DD}$ and $V_{TH}$ can be scaled down, but the $>60$ mV/dec subthreshold swing causes rapid increase in $I_{OFF}$.

Figure 2.5: CMOS transistor characteristic. This transistor is turned ON when the gate signal is low, and OFF when the gate signal is high. Conversely, the lower transistor is an NMOS device, and is turned ON when the gate signal is high, and OFF when the gate signal is low. The output from this circuit is either logic 1 ($\sim V_{DD}$) or logic 0 (GND), as described in Chapter 1. In the case of a logic 1 output, the PMOS transistor is ON, thus connecting the output to $V_{DD}$, whereas in the case of a logic 0 output, the NMOS transistor is ON thus connecting the output to GND. With the circuit in either state, one of the transistors will be in the OFF state, thus preventing current flow in the
ideal case of zero OFF state leakage. This is the key to the energy efficiency of CMOS technology, and is what precipitated the move from NMOS and bipolar technologies in the mid 1980s. Up to that point, integrated circuit complexity and power consumption had been increasing, and power dissipation was becoming a challenge. The move to CMOS allowed more energy efficient designs. However, due to the finite subthreshold slope (in addition to gate leakage), there is a finite OFF state leakage current. With today’s extremely high device densities (with 100 million transistors on 1mm\(^2\) area of Si for example), this leakage current is now a serious impediment to future scaling [31] as the industry once again encounters the issue of dissipating the power consumed by the integrated circuit. This power dissipation issue is compounded by the recent expansion in battery powered mobile computing in laptops and cellular ’phones, where the power supply is limited by the battery capacity. While the problem can be mitigated by power gating, where idle blocks of the microprocessor are temporarily powered down, the

Figure 2.6: CMOS circuit configuration: no connection across the power rail when circuit is stable (not switching).
fundamental energy efficiency of the MOS transistor must be addressed if Moore’s law is to continue. This is a major issue for the industry, and it is this challenge that provides the high level motivation for the work outlined in this thesis.

2.4 Nanorelays for Reduced Leakage

A number of alternative devices that do not rely on thermionic emission have been suggested to address the power efficiency crisis of MOS transistors. One such candidate is Tunnelling Field Effect Transistors (TFETs), where current flow is based on tunnelling of carriers through a barrier rather than thermionic emission over a barrier. The switching mechanism is based on modulating the ability of the carriers to tunnel through to the channel, and subthreshold slopes of less than 60mV/dec have been demonstrated by Ionescu et al [32], for example. However, these transistors are hampered by limited $I_{\text{ON}}$ current, with implications for circuit speed as noted above. The prospects and challenges for this work are reviewed by Avci et al in [33].

Mechanical relay type switches are a promising category of device to replace (or, more likely, complement) CMOS technology. As a candidate device, mechanical relays, scaled appropriately, offer compelling power performance based on abrupt switching and essentially zero OFF state leakage. A cantilever beam style relay is shown in Fig. 2.7. The relay is a three terminal device analogous to a MOS transistor, with a source, gate and drain. When a voltage is applied to the gate, the electrostatic force pulls the cantilever down to contact the opposing (source) electrode, thus closing the switch. As electrostatic force is ambipolar, the device relies on the stiffness or spring restoring force of the beam to break the contact and return the beam to
Figure 2.7: (a) A nanorelay with a suspended beam of length $L$, and thickness $t$, separated by an air gap $g$, from the substrate. Note the dimple at the end of the beam which gives a reduced separation of $g_d$. (b) A bias applied between the beam (Drain) and Gate terminals causes the capacitor comprising the beam and gate with air dielectric, to charge up. The electrostatic force then causes the beam to bend bringing the Drain in to contact with the Source at the dimple contact.

Although these advantages are compelling in view of the power crisis afflicting CMOS, a number of challenges remain to be resolved before nanorelay technology can be adopted in commercial products. As noted in Chapter 1, to be applicable in state of the art CMOS technology, these nanoelectromechanical (NEMS) nanorelay devices need to be aggressively scaled to the nano regime. NEMS devices in the literature are typically micron sized in at least one dimension. While there may be applications such as power gating (switching power to blocks of circuitry only when the block is active) where the device energy efficiency outweighs the penalty associated with utilising relatively large areas of Si real estate on the chip, more widespread adoption would require more aggressive scaling. This miniaturisation brings further
problems. The electrostatic force available to pull the active element (cantilever beam) into contact with the source scales with the active device area. Further, the requirement to operate at CMOS compatible voltage (~1V) limits the force, and this problem is exacerbated by the inverse beam stiffness dependence on length, as will be shown in Chapter 3. On the other hand, once the switch is in the closed position, stiction forces which are known to dominate at the nanoscale, must be overcome to allow the beam to be restored to the neutral position when the gate voltage is removed. Contrary to the case for low voltage operation, this requires that the beam be sufficiently stiff. The nature of the contact interface is also critical. In addition to the mechanical properties just discussed, the contact must be capable of carrying sufficient current, and must withstand repetitive impact without significant wear. For example in [34], Peschott et al point out that to guarantee functionality for ten years based on 100MHz operating frequency with a duty factor of 1%, the device must withstand 3x10^{14} ON / OFF switching cycles. In Chapter 5, it is shown that for this operating frequency, a current of 100nA per nanorelay is required in a typical circuit configuration.

2.5 Thesis Overview

The thesis is organised as follows. The theory of NEMS devices is presented in Chapter 3. Recent developments in the field, and the state of the art as it relates to both device technology and the integration of these devices into CMOS circuitry is then reviewed. The chapter ends with a discussion of the challenges facing the technology, as NEMS devices are scaled from micron dimensions to nanoscale dimensions. Specific challenges relating to stiction, electrical performance and reliability are addressed.
Chapter 4 begins with a review of cantilever beam mechanics. The stiffness of theoretical beams is described, and the need to build and validate models of the stiffness of these beams is addressed. This is followed by a description of the fabrication of arrays of nanoscale cantilever beams. The dimensions of these beams are varied in a systematic and controlled manner. The stiffness of these beams is then measured using force spectroscopy in an atomic force microscope (AFM). This technique is described in detail, along with the results from the fabricated array. The need for validated analytical and finite element models to accurately predict the mechanical behaviour of these nanoscale beams is then presented, and the model results are compared to the results of experimentally measured fabricated nanoscale beam arrays. The models are modified to account for practical nonidealities found in the fabricated beams.

In Chapter 5, the objective of intercepting the industry roadmap at the 22nm node is discussed. Taking this technology node as a basis, reasonable dimensions for a cantilever beam device are determined, and this leads to adhesion force requirements, while electrical specifications are derived from practical circuit implementations. Using an AFM tip and substrate to replicate the dimple and source of a nanorelay, it is possible to measure the adhesion performance of a range of material combinations by performing force-displacement measurements and measuring the adhesion force. By applying electrical bias across the tip-substrate contact, the electrical performance of various dimple-source contact materials can be simultaneously tested. These results are then compared to the aforementioned specification to determine if the particular candidate materials support nanorelay operation.

Chapter 6 addresses future work, including fabrication challenges, and device reliability. Possible solutions such as novel device design and materials
are reviewed. Preliminary experimental work demonstrating the elimination of stiction is presented as a potential route to reliable device operation.

The work is summarised, and conclusions are presented in Chapter 7.
Chapter 3

Theory

3.1 Mechanics of Cantilever Beam Nanorelays

Consider the cantilever beam nanorelay design shown in Fig. 3.1. The device is based on a cantilever beam which acts as the active element. A bias is applied to the gate terminal, which, together with the cantilever beam, acts as a capacitor. The electrostatic force arising from the charge on this capacitor causes the cantilever to accelerate towards an opposing source contact, thus closing the switch. Note the contact dimple on the end of the cantilever, which defines the contact interface, and reduces switching time by minimising beam travel. As the gate voltage is reduced, the combination of electrostatic force and van der Waals (vdW) force ensures the switch remains in the closed position until the pull-out or release voltage is reached, at which point, the beam spring restoring force overcomes these forces, and the beam moves back to the neutral position. The device can also be designed to remain in the closed position when the gate voltage is removed, allowing nonvolatile
memory functionality. The dimple gap, $g_d$, is a key design parameter which determines whether or not the device will display this hysteretic behaviour as explained below. Note that the mechanical spring restoring force acting on the beam is a linear function of displacement,

$$F = -kz = -k(g_0 - g)$$  \hspace{1cm} (3.1)$$

where $F$ is the applied force, $k$ is the stiffness and $z$ is the displacement or distance the beam has travelled towards the opposing contact, $g_0$ is the as fabricated actuation gap thickness and $g$ is the actual gap thickness. By modelling the beam and gate electrode as a parallel plate capacitor of capacitance, $C$, it can be seen that the electrostatic force experienced by the cantilever, on the other hand, is a superlinear function of displacement,

$$F_{elec} = \frac{1}{2}qE = \frac{1}{2} \left( \frac{CV^2}{g} \right) = \frac{1}{2} \left( \frac{\epsilon_0WL}{g} \right) \left( \frac{V^2}{g} \right) = \frac{\epsilon_0WLV^2}{2g^2}$$  \hspace{1cm} (3.2)$$

where $q$ is the electronic charge, $E$ is the electric field, $W$ and $L$ are the width and length of the beam, $V$ is the applied voltage and $g$ is the actuation gap distance. At equilibrium, the electrostatic force is balanced by the spring restoring force.

$$k(g_0 - g) = \frac{\epsilon_0WLV^2}{2g^2}$$  \hspace{1cm} (3.3)$$
3.1. MECHANICS OF CANTILEVER BEAM NANORELAYS

As the displacement of the beam from the neutral position increases due to the applied gate voltage, there comes a point where the electrostatic force will be unconditionally larger than the spring restoring force. At this point, the beam will abruptly jump to contact. This phenomenon, described by Rebeiz [35], is known as “pull-in”, and Kaajakari [36] has shown that this critical displacement occurs when the beam has moved by $\frac{g_0}{3}$. This phenomenon occurs at the pull-in voltage given by Yaung [37] as

$$V_{pi} = \sqrt{\left(\frac{8}{27}\right)\left(\frac{k g_0^3}{\epsilon_0 (W L)}\right)}$$  \hspace{1cm} (3.4)

As noted above, nanorelays frequently utilise a contact dimple which can be seen in Fig. 3.1, where the dimple gap, $g_d$, is shown. This dimple defines the patterned contact area and can be designed independently of the as fabricated actuation gap. If the dimple gap is less than $\frac{g_0}{3}$, then the pull-in phenomenon will not occur, because the dimple will contact the source electrode before the beam displacement reaches $\frac{g_0}{3}$. The mode of operation (whether pull-in or non pull-in) has implications for the device operating characteristic. For a device operating in pull-in mode ($g_d > \frac{1}{3}g_0$), the actuation gap, which reduces to $g_0 - g_d$ on contact, leaves the beam closer to the gate electrode, so the electrostatic force is larger for the same $V_{pi}$. To release the beam, the voltage must be reduced below $V_{pi}$ leading to the aforementioned hysteresis. The spring restoring force now has to overcome both surface adhesion force, and the electrostatic force that exists when contact is made. In non pull-in mode, the spring restoring force must overcome the surface adhesion force only, leading to reduced hysteresis [38]. In all cases, the spring restoring force must overcome the adhesion force, to allow the switch to pull out. Fig. 3.2 shows the electrostatic and mechanical forces that apply at $V_{pi}$ for both a pull-in mode and non pull-in mode device.
CHAPTER 3. THEORY

Figure 3.2: Mechanical restoring force ($F_k$) and electrostatic force ($F_e$) for (a) a pull-in mode nanorelay and (b) a non pull-in mode nanorelay. When operating in pull-in mode, the beam is closer to the gate electrode when the device is ON (labelled $g_0$-$g_d$ on the $Z$ axis), leading to higher electrostatic force. The net downward force ($F_e$-$F_k$) holds the beam in the contact position until the electrostatic force is reduced below the restoring force by reducing the gate voltage. This gives rise to the hysteresis visible in Fig. 3.3. In the non pull-in mode device, the net force ($F_e$-$F_k$) is zero when in the contact position (again labelled $g_0$-$g_d$ on the $Z$ axis), and the beam returns to neutral position as soon as the gate voltage is reduced, neglecting adhesion force (adapted from [37]).

In the case of the pull-in mode device, the electrostatic force (solid line) is greater than the linear force (dashed line) when the relay is ON. In order to turn the device OFF the gate voltage must be reduced below $V_{pi}$ to reduce the electrostatic force to less than the opposing spring restoring force. This gives rise to the hysteretic characteristic shown in Fig. 3.3. In the case of the non pull-in device, the electrostatic force does not exceed the spring restoring
3.1. MECHANICS OF CANTILEVER BEAM NANORELAYS

Figure 3.3: Nanorelay hysteretic operation: the nanorelay turns on at $V_{pi}$ but does not turn off until the gate voltage is reduced to $V_{PO}$

force but balances the spring restoring force when the device is turned ON. Thus the spring restoring force immediately overcomes the electrostatic force as the gate voltage is reduced from $V_{pi}$. When a relay operating in non pull-in mode ($g_d < \frac{1}{3}g_0$) turns ON, the actual gap, $g$, is given by $g_0 - g_d$, and the turn on voltage (denoted $V_{pi}$ by convention) is given by

$$V_{pi} = \sqrt{\frac{2k_{eff}g_d(g_0 - g_d)^2}{\epsilon_0 A_{eff}}}$$

(3.5)

where $k_{eff}$ is the effective spring constant of the actuation beam, and $A_{eff}$ is the overlap area between the actuation beam and the fixed gate electrode. When contact is established, the force is described by

$$\frac{\epsilon_0 (WL)^2}{2g^2} + F_{adh} = k(g_0 - g)$$

(3.6)

where $F_{adh}$ is the surface adhesion force. At the release point where $g = g_0 - g_d$, the release voltage is given by

$$V_{rl} = \sqrt{\frac{2(k_{eff}g_d - F_{adh})(g_0 - g_d)^2}{\epsilon_0 A_{eff}}}$$

(3.7)
where $A_{eff}$ is the overlap area between the actuation beam and the gate electrode.

## 3.2 NEMS Devices: State of the Art

As described by White in [39], mechanical switches have been candidates for computing going back to the 1930s. However, for the past thirty years, CMOS technology has been dominant. This dominance was the result of the energy efficiency of CMOS (compared to TTL for example as described in Chapter 1). As devices have been scaled to the nano regime, this advantage has been eroded. As explained in Chapter 2, the device supply voltage, $V_{DD}$, and the threshold voltage, $V_{TH}$, should be reduced to minimise power consumption when the transistor is ON (i.e. dynamic power), but this voltage scaling causes an exponential increase in OFF current (static power) due to the finite subthreshold slope ($>60$ mV/dec). With current generation technology, any improvement in dynamic power performance through voltage scaling is immediately negated by poorer static power performance caused by the exponentially increasing leakage (OFF) current. In this scenario, a device that offers near infinite subthreshold slope and essentially zero OFF state leakage is compelling. Several nanoelectromechanical devices have been published in the literature over the past fifteen years or so. Simple cantilever beam (CB) devices are common although more complex devices comprising multiple input and output terminals capable of dynamically configurable logic functions have also been demonstrated. Fig. 3.4 shows three generations of such a device based on a suspended SiGe membrane suspended on four serpentine spring legs [40]. The SiGe is isolated from the tungsten (W) channel by an $\text{Al}_2\text{O}_3$ gate oxide. Dimples are formed in the W channel to
3.2. NEMS DEVICES: STATE OF THE ART

Figure 3.4: Scaling of multi terminal relay device. The device evolved from a 90\(\mu\)m four terminal device to a 15\(\mu\)m six terminal device with multiple source (S) and drain (D) contacts for additional functionality. To switch the device, a control voltage is applied between the gate (G) and body (B) terminals [40].

make contacts on (multiple) source and drain contacts. While this device is micron sized, the device has been scaled from 90\(\mu\)m to 15\(\mu\)m as can be seen in Fig. 3.4, and has been refined to reduce its footprint and increase the number of contacts to allow complex functional logic to be implemented. The CB devices, meanwhile, utilise a variety of materials including single crystal Si, polysilicon, CNTs and various metals, but all typically feature micron dimensions, and feature \(V_{pi}\) in excess of 1V. Metal contact dimples such as Al [41], Ru [42], Cu [43], W [44] and Pt [45] have all been demonstrated. A key advantage of metal beam systems is their compatibility with back end of line (BEOL) CMOS integration. In [41], for example, Muñoz-Gamarra et al
showed how a device based on Al / TiN / W in the back end could deliver a subthreshold slope of 5mV/dec with an $I_{ON}/I_{OFF}$ ratio of $10^3$ (a state of the art MOS device would achieve a ratio of $10^6$).

However, challenges remain. In order to achieve low switching voltages compatible with CMOS technology, beams tend to be long (several $\mu$m). This arises from the difficulty of patterning very small actuation gaps (ideally 10nm or less) and the constraints imposed on gap dimensions by the design rules for the particular technology. In [46], Lee et al proposed a novel design to address this issue based on a novel pipe clip structure to achieve a 4nm thick air gap. More generally, for the larger gaps typically used, the beam stiffness (which is inversely proportional to $L^3$) must be reduced.

While metal contacts provide good electrical behaviour, the metallic bonding, and the associated large van der Waals forces, can be problematic for non volatile operation, given their adhesive nature. Even when stiction can be overcome, reliability remains as a major challenge. The actual contact area is less than the physically patterned contact area, with the actual contact area dominated by contacts via asperities. The current density at these asperities is high enough to cause microwelding and contact ablation. Unlike radio frequency (RF) applications of microelectromechanical systems (MEMS) relays, where contact resistance must be low to minimise insertion losses (i.e. loss of signal power from the resistance of a component inserted in the circuit), a relatively high contact resistance (e.g. 10K$\Omega$) can be tolerated in NEMS applications. This is because the electrical (RC) delay of a nanorelay (~ps) is about a thousand times shorter than the mechanical delay (~ns) arising from the travel time associated with beam displacement. With this in mind, it has been shown that a dielectric material such as TiO$_2$ can be used to coat tungsten to reduce current density and improve reliability, because the
increased electrical time constant (RC delay) remains well below the mechanical time constant. A further advantage of this approach is that oxidation of the metal contact can be avoided (for Ru, however, Lee et al points out that RuO$_2$ is conducting and so less problematic [47]).

Akarvardar et al [48] benchmarked several metals against Si as the beam material. A model was presented that showed that by using a low density 225nm Si beam, 1V operation would be feasible with single crystal Si, but this has not been experimentally verified. Single crystal beams (both cantilever and double clamped) are quite common in resonator applications [49], [50], [51], but beam lengths are in the micron range. The low mass and high resonance frequencies of such suspended Si nanobeams lend these devices to applications in high resolution mass sensing. While the beam lengths are greater than 1$\mu$m, their integration with CMOS signal processing circuitry reflects the promise for wider application of suspended Si NEMS devices for other applications such as digital logic and power gating. In addition to submicron Si beams, a number of avenues have been explored to address the relatively high pull-in voltage required for devices in the literature. The use of a fourth “body” terminal is shown in Fig. 3.5 for a crab like relay device [38]. This idea is similar to a floating bias terminal proposed by Yang et al for MEMS [52]. By applying bias to the body terminal, the switching voltage, $V_{gb}$, can be reduced below 1V. Jeon et al [53], has shown that by judicious application of bias to the body terminal, various logic operations (AND, OR, MAJORITY) can be implemented. As noted earlier, the approach used by Lee et al in [46] seeks to exploit the sensitivity of the pull-in voltage to the actuation gap arising from the quadratic dependence of electric field on the separation between the active element and the gate, as seen earlier in Eqn. 3.2. In this case, a novel design is used to achieve a gap of just 4nm,
Figure 3.5: Schematic of four terminal relay. Device operation is based on the electrostatic force arising from the voltage between the gate and the fourth “body” terminal [38], leading to an operating voltage of 400mV. Fig. 3.6 shows the novel pipe clip structure achieved using high density plasma (HDP) deposition [54] over a self aligned trench to realise the 4nm gap.

Figure 3.6: Schematic of two terminal NEM switch. The expanded image shows how the 4nm gap is realised through the use of a pipe clip structure [54].
As the mechanical requirements of the beam (sufficiently stiff to overcome contact adhesion force, with a low adhesion dimple contact surface) and the electrical current carrying requirement can sometimes be contradictory, a number of designs seek to decouple these requirements. For example, Parsa et al [55] [56] [57], use a compliant polySi beam to define the mechanical performance with a Pt coated beam providing electrical contact. To prevent secondary pull-in, where the active element (beam) causes a short between the source and gate, see Fig. 3.7, the cantilever is modified in the actuation area to provide a non compliant stiffness. In [58], Lee et al use a dual sidewall coating on the beam consisting of HfO$_2$, which provides electrical isolation between the drain and the beam and Pt which provides electrical contact between the source and drain terminals, as shown in Fig. 3.8. This adds

![Figure 3.7: Cantilever beam (CB) device with the beam modified to be non compliant in the actuation area thus preventing secondary pull-in [55].](image-url)
Figure 3.8: CB device with HfO\textsubscript{2} / Pt coating on beam. The HfO\textsubscript{2} isolates the drain from the beam, and Pt provides electrical contact between the drain and source providing the ability to mimic both PMOS and NMOS behaviour. The arrangement also allows independent biasing of the beam [58].

significant flexibility for biasing the devices, which can, for example, act as either PMOS or NMOS in digital circuits. Further, with appropriate beam bias, the pull-in voltage was reduced from 15V to 2.5V. In [59], Qian et al reduced the pull-in voltage to approximately 1V, by increasing the electrostatic force by using a large area (µm scale) capacitive paddle (Fig. 3.9) suspended by two Si beams (the double beam arrangement mitigates torsional effects by balancing the spring restoring force on the paddle). The excellent switching characteristics of this device proved unsustainable due to contact oxidation and high current density in the Si nanowires.
Figure 3.9: This device achieves reduced voltage operation by increasing the electrostatic force with a large area (~2x4μm) capacitive paddle, as shown in the inset [59].
CHAPTER 3. THEORY

3.3 Circuits incorporating NEMS: State of the Art

Due to the relatively long switching time of a NEMS device, associated with mechanical displacement of the active element, it is best to design nanorelay based circuitry to avoid sequential device switching, thereby exploiting the energy efficiency of these devices while mitigating their limited speed. In this way, complex logic circuitry can be configured where all relays switch simultaneously through circuit design tailored to NEMS device characteristics, thus limiting the delay to one mechanical switching period, while simultaneously exploiting the aforementioned leakage and subthreshold characteristics to drive energy efficiency.

In [60], Fedder et al review several strategies that have been pursued for integrating MEMS and CMOS. The first demonstration of functional NMOS - NEMS integration was reported by Chong et al in [61]. In this case, a CB device was fabricated over a 1 μm gate length NMOS transistor, operating at 5V. The actuation gap for the CB device was 100nm, and it was noted that scaling the actuation gap towards 10nm would allow reduction in the operating voltage to 1V. The active element of the device consists of a ~3.5μm Pt (with Ti adhesion layer) ~80nm thick cantilever beam. The NMOS transistor was used to drive the gate of the nanorelay. With the NMOS transistor turned on, the drain voltage was passed to the gate of the nanorelay, and as the drain voltage was continuously swept, the nanorelay was seen to switch ON and OFF repeatedly. Although the nanorelay failed to switch after eight cycles, it was an interesting proof of concept. This group then expanded the work to demonstrate functional CMOS - NEMS circuitry with a CMOS inverter driving a NEMS device [45]. While there were still
reliability problems to be overcome (high ON resistance, $R_{ON}$, and gradual switching transition compared to the abrupt switching of the isolated NEM device), the work demonstrated the potential for integration.

By switching from planar to out of plane (vertical geometry), the transistor interconnect wiring can be used to integrate NEMS with CMOS. State of the art CMOS integrated circuits published by Intel [13] feature 12 layers of interconnect wiring. Xu et al [62] and Muñoz-Gamarra et al [41] showed that the aggressive scaling of interconnect layers can be exploited to incorporate vertical cantilever beam devices in a monolithic integration strategy. An example is shown in Fig. 3.10. It is projected that 1V operation with a switching time of 20ns is feasible, and that by using the vertical out of plane orientation, a footprint of 0.1 $\mu m^2$ is possible. Circuit simulations utilising such devices have been performed for various functions. For example, Xu et al [62] show how energy efficiency can be improved by integrating this NEM device in CMOS buffer circuits, or in power gating, where a non-volatile NEM device based circuit can be used to power off inactive logic blocks in an energy efficient manner. Other proposed circuits which could benefit from cointegration of NEMS and CMOS include field programmable gate arrays (FPGA) [63], static random access memory (SRAM) [64] and header / footer switches [65].

Taking SRAM as an example, Bota et al [66] analyse the performance of NEM relays designed using 65nm node technology in a conventional SRAM cell. The nanorelay is a CB device utilising one of the metal layers with a $17\mu m$ long, 100nm wide, 220nm thick beam with an as fabricated source gate gap of 100nm and stiffness of $8.4 \times 10^{-2}$ N/m. Using this device, they develop a Verilog-A model and compute a pull-in voltage of 0.97V, a release voltage of 0.68V, a gate source capacitance of 0.8 aF with the device in the ON
Figure 3.10: Vertically oriented cantilever beam (CB) device implemented in the back end of line (BEOL) of an integrated circuit. The input node connects to the actuation beam located between two actuation electrodes. Depending on the input signal, the beam is electrostatically attracted to one of the actuation electrodes thus connecting one of the data terminals (\(D_0\) or \(D_1\)) to the output [62].

position, and 0.26 aF with the device in the OFF position. The switching time, \(\tau_{\text{mech}}\), is 361ns, and the ON resistance of the device is 1KΩ. This nanorelay is used in various configurations in the 6T SRAM cell in circuit simulations, and the impact on circuit performance is determined. They find that the leakage power can be reduced by up to 78%. Interestingly, the long switching time of the relay does not impact on the read or write time of the circuit (although a latency delay equal to \(\tau_{\text{mech}}\) is incurred). In addition, the subthreshold slope and hysteretic behaviour of the NEM relays were found
3.3. CIRCUITS INCORPORATING NEMS: STATE OF THE ART

to improve the static noise margin of the circuit.

More complex MEM relay based circuits have also been demonstrated. For example, Fariborzi et al demonstrate a compressor circuit built using 46 scaled MEM relays in [67]. The demonstrator chip shown in Fig. 3.11 contains relay based logic gates, oscillators and memory. This shows great promise if the scaling challenge can be overcome [68].

In addition to logic and memory applications, NEMS are also used as sensors [69], for example in mass detection, where they offer unprecedented resolution. The integration of these NEMS sensors with CMOS reflects the state of the art as it relates to the integration of NEMS devices in CMOS

![Die photo with the relay based logic, oscillator and memory circuitry shown in the inset](image)

**Figure 3.11:** Die photo with the relay based logic, oscillator and memory circuitry shown in the inset [68].
integrated circuits. For example, Esfahani et al demonstrated a resonator based on a 5μm long, 100nm wide, 200nm thick suspended Si beam in a clamped - clamped configuration interfaced to a measurement circuit [70]. A similar monocrystalline Si suspended resonator comprising a 3.2μm long 300nm wide 160nm thick Si nanowire interfaced to circuitry was presented by Sansa et al [71]. While most resonators utilise suspended beams which are a few microns long, Tsuchiya et al [25] recently published the highest resonance frequency (330 MHz) for a lithographically defined beam in a CMOS compatible process. The beam length was 800nm, and it was found that the undercut associated with the isotropic release etch significantly impacts on the resonance frequency (in addition to a thermal oxide passivation of the Si beam). These issues are critical for a complete understanding of the mechanical properties of these beams when used as nanorelays, as will be detailed in Section 4.5.
3.4 NEMS / Nanorelay Technology: Challenges

3.4.1 Miniaturisation

There are two types of nanorelay employed in CMOS circuitry in the literature. The crab like multi-terminal relay developed by Kam et al at Berkeley [40], and simpler cantilever based devices such as that published by Liu et al [72]. While both device types have been scaled, the device active elements remain of micron scale, and operate at a few volts. As noted earlier, Arkarvardar et al [48], proposed that by utilising a low density material such as single crystal Si, 1V operation is possible with a 225nm long beam, if small actuation distances can be manufactured. An actuation gap of 10nm is frequently cited in the literature. With Intel’s 14nm CMOS technology node now in production [73], and 10nm node technology published by Auth et al [13], such patterning tolerances are not unrealistic. However, suspended beams in the literature are rarely submicron. Such nanoscale beams need to be fabricated and their mechanical performance understood, with a need for validated models that predict mechanical properties such as stiffness. This stiffness is a key determinant of operational voltage and must be minimised for low voltage operation, while at the same time retaining sufficient spring restoring force to overcome dimple - source adhesion. In view of these competing requirements, it is important that the stiffness of nanoscale beams is both well understood and controllable through design.
3.4.2 Adhesion

In the CB nanorelay shown in Fig. 3.1, there are multiple forces acting on the active element. The applied voltage to the gate electrode gives rise to an electrostatic force which acts to “pull-in” the active element towards the source electrode. The dynamics of this motion is described by the equation

\[ F_{elec} = m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx \]  

(3.8)

where \( m \) is the mass, \( x \) is the displacement, \( b \) is the damping coefficient caused by displacement of air molecules and anchor losses, and \( k \) is the beam stiffness [74]. The electrostatic force arises from the capacitive charging of the beam with respect to the substrate. As noted earlier, (cf. Eqn. 3.2), this capacitance is given by

\[ C = \frac{\epsilon_0 WL}{d} \]  

(3.9)

and the force \((F = qE)\) acting one of the capacitor plates (the movable beam) is then given by

\[ F = \frac{\epsilon_0 (WL) V^2}{2d^2} \]  

(3.10)

The spring restoring force of the beam, which balances this electrostatic force, is linear

\[ F = -kx \]  

(3.11)

assuming linear Hookean behaviour. As the beam extends, this spring restoring force increases linearly, whereas the electrostatic force increases superlinearly (due to the inverse quadratic dependence on distance from Eqn. 3.10 above). By setting the dynamic terms in Eqn. 3.8 above to zero, it can be shown that when the displacement of the beam equals \( d/3 \), the electrostatic force unconditionally exceeds the restoring force, and the beam abruptly snaps down to the closed position, as discussed in Section 3.1 above. The
voltage at which this condition is met is given by

\[ V_{pi} = \sqrt{\left( \frac{8}{27} \right) \left( \frac{k g_0^3}{\epsilon_0 (W L)} \right)} \]  

(3.12)

with the terms described earlier in Section 3.1. As the gate voltage is now reduced back to zero, an additional force associated with the interface must be considered. At the nanoscale, it is known that adhesive forces, which scale with surface to volume ratio, become important as described by Ramakrishna et al [75] and Gauthier et al [76]. Because of this adhesion force, \( F_{adh} \), the beam remains in the closed position even after the gate voltage falls below the pull-in voltage. This gives rise to the hysteretic device characteristic shown in Fig. 3.12. However, if the spring restoring force of the beam does not exceed \( F_{adh} \), the beam will remain permanently in the closed position. Thus \( F_{adh} \) places a minimum bound on the CB stiffness, and thus, the operating voltage. For this reason the adhesion force is critical. The adhesion force arises from van der Waals interactions, electron exchange in-

\[ \text{Figure 3.12: NEM operating characteristic: note the abrupt switching and hysteretic behaviour.} \]
interactions, unsaturated bonds and capillary effects. The van der Waals forces arise from permanent and instantaneous polarization of atoms, and comprise dipole-dipole force, dipole-induced dipole force and dispersion forces (charge fluctuation). These will be described in detail in Section 6.5. As described by Israelachvili in [77], the adhesive force can be expressed as

\[ F_{adh,v} = \frac{H A_c}{6\pi (D + D_0)^3} \]  

(3.13)

where \( H \) is the Hamaker constant of the contact materials and \( A_c \) is the real area of contact (i.e. actual contact dominated by asperities rather than the physically patterned electrode area), \( D \) is the separation and \( D_0=0.165\text{nm} \) is an effective cutoff distance in this continuum approximation from [78]. To minimise contact resistance, metallic contacts might be used. However, this presents an adhesion problem. These contacts usually feature strong adhesion, arising from a strong electron exchange interaction which dominates the vdW force. An estimation of this force is given by

\[ F_{adh,m} = -2\gamma A_c \frac{d}{dD} \left( \left[ 1 - \frac{D - D_0}{\lambda_M} \right] e^{-\frac{(D-D_0)}{\lambda_M}} \right) \]  

(3.14)

where \( \gamma \) and \( \lambda_M \) are the surface energy and electron mean free path respectively [77]. To address this adhesion problem, metal contacts have been coated with a thin layer of TiO\(_2\) [40], for example, or implemented using rhuthenium, Ru, [42] which oxidises to form conducting RuO\(_2\). It should be noted that the mechanical delay, \( \tau_{mech} \), is about three orders of magnitude greater than the electrical RC time constant, \( \tau_{elec} \), and so there is some margin for increasing the resistance (and thus \( \tau_{elec} \) through \( RC \)) before impacting on device speed, giving some leeway for alternatives to pure metal - metal contacts. A more detailed example of this design aspect on nanorelays is presented in the context of targeting the 22nm technology node in Chapter 5.
Capillary forces also add to adhesion, and can be difficult to avoid if operating in ambient conditions. For Si, approaches such as hydrogen termination or passivating Si with fluorine can reduce the effects of unsaturated bonds and capillary effects. This topic will be dealt with in more detail in Chapter 5 of this thesis.

### 3.4.3 Contact reliability

The importance of minimising adhesion for low operating voltage was addressed above. Nevertheless, the contact must still be conductive with a contact resistance no greater than about 100KΩ (depending on application) according to Kam et al [40]. The scaling of relay devices towards the nano regime exacerbates this problem, as the real contact area is typically even less than the apparent contact area due to roughness, where the current is carried by a number of asperities. This means the current density can be much higher than is apparent from the nominal contact area. Furthermore, the device operates by treating the beam as one plate of a capacitor. The beam charges up (thus generating the electrostatic force), and the stored charge dissipates upon contact. This initial discharge current density can be sufficient to microweld or even ablate the contact. Similar to the problem with adhesive metal contacts, this can be mitigated by coating the metal contact with a dielectric. Kam et al [40] note that of the common dielectric materials, TiO$_2$ reduces the conductivity least, and has the added advantage of preventing oxidation of metallic contacts such as W. However, whatever material system is used, the repetitive impact of switching can cause degradation, ultimately leading to premature failure.

The contact material must meet the criteria of reduced adhesion, acceptable electrical resistance, immunity to ablation from electrical discharge and
resistance to wear and tear simultaneously. Streller et al [79] propose platinum silicide, whose properties can be tuned to achieve 150% of the hardness of as deposited Pt with a contact resistance within 10% of as deposited Pt. Mayet et al [80] demonstrated the use of amorphous WN$_X$ which is electrically conductive and hard (with a hardness of 3GPa and a Young’s modulus of 300GPa) as both a structural element (the cantilever beam) and the contact material. The authors deposited this material using reactive sputtering, and patterned CB devices using electron beam lithography (EBL) and plasma etch. Device operation was demonstrated, but the authors note the requirement for further optimisation to reduce the contact resistance. The performance and reliability of both Pt and RuO$_2$ was investigated by de Boer et al [81], who found that that RuO$_2$ performed better than Pt which failed due to wear after $10^5$ switching cycles, while the RuO$_2$ contacts survived $3 \times 10^8$ switching cycles, albeit with unstable resistance. Silicon Carbide (SiC) is yet another material offering excellent wear resistance as described by He et al [82], as will be described in Section 6.7. Further work on contact materials is required before nanorelay technology can be widely adopted.
Chapter 4

Si nanocantilevers: mechanics, fabrication & testing

4.1 Introduction

This chapter opens with a brief review of cantilever beam mechanics, and introduces an analytical expression for the stiffness of an ideal cantilever beam. The design and fabrication of an array of Si nanocantilever beams is then described, with detailed discussion of the CMOS compatible processing steps (resist coating, lithography, plasma etch, wet etch and strip) involved in their manufacture. This is followed by a description of the use of atomic force microscope (AFM) based force spectroscopy to measure stiffness. Finally experimental results are presented, and fabrication related non idealities are analysed and modelled. The applicability of the analytical model to the nanoscale is then clarified.
4.2 Review of Beam Mechanics

The operation of a cantilever beam (CB) nanorelay was reviewed in Chapter 3. The stiffness, $k$, of a cantilever beam is given by the Euler Bernoulli equation [83]

$$k = \frac{F}{x} = \frac{3EI}{L^3} \tag{4.1}$$

where $F$ is the applied force, $x$ is the beam displacement, $E$ is Young’s modulus, a measure of the stiffness of a material given by the ratio of stress (the applied force per unit area in N/m$^2$ or Pa) to strain (the deformation due to the stress, which is dimensionless) and $I$ is the moment of inertia. By analogy to the way that mass defines the force required for a certain acceleration, the moment of inertia describes the torque required to achieve a certain angular acceleration. The moment of inertia depends not just on the mass of the object, but on the spatial distribution of the mass of a given object. This can be understood intuitively by considering how an ice skater’s rotational velocity increases as the ice skater draws their outstretched arms inward, thus changing the distribution of their mass by reducing the distance from some of their mass to the axis of rotation. For a cantilever beam, the moment of inertia is given by

$$I = \frac{wt^3}{12} \tag{4.2}$$

where $w$ and $t$ represent width and thickness respectively. Substituting Eqn. 4.2 in to Eqn. 4.1 above, the stiffness of a cantilever is given by

$$k = \frac{3Ewt^3}{12L^3} = \frac{Ewt^3}{4L^3} \tag{4.3}$$

This analytical expression is tested against fabricated nanoscale cantilever beams in Section 4.6.
4.3 Design & Fabrication of Nanocantilever Array

It was necessary to validate the analytical models of stiffness presented above based on realistic Si nanocantilevers. The schematic shown in Fig. 4.1(a) is the design concept for a single test nanocantilever, and is based on a suspended beam created on a silicon on insulator (SOI) wafer. An array of nanoscale cantilevers was fabricated on such a wafer. The array is shown in Fig. 4.1(c), with a close up of one nanobeam shown in Fig. 4.1(b). The suspended beams are fabricated in the device layer and suspended by removal of the underlying buried oxide (BOX). The device layer design thickness is 50nm, and the beam widths range from 50nm to 150nm, with lengths from 150 - 1500nm. A schematic representation of the fabrication process is shown in Fig. 4.2, and begins with spincoating electron beam resist, followed by electron beam exposure and development and dry plasma etching of the device layer Si, and the underlying BOX. Then, a wet isotropic etch release process which removes the SiO$_2$ from underneath the nanobeam, creates the suspended active element of the device. The final step is the removal of the electron beam resist in a barrel asher. The process is described in detail below.
Figure 4.1: (a) Schematic of nanocantilever design in SOI wafer (b) SEM micrograph of 500nm long, 75nm wide, 50nm thick fabricated nanocantilever and (c) array of fabricated nanocantilevers.
4.4 Processing

4.4.1 Application of (electron beam) resist

Three e-beam patterning processes were evaluated, one using SML e-beam resist (EM Resists Ltd, SML Resist Technology), one using nLOF 2000 e-beam resist (Micro Chemicals) and one based on ZEP-520 e-beam resist (Zeon Corporation, Tokyo, Japan), which gave the best results. In this step, the ZEP-520 resist is deposited using a spincoat process. A small volume of the resist is dropcast on to the slowly (500 rpm) rotating wafer chuck. This “dynamic dispense” is used to ensure good coverage on the substrate, and helps to eliminate voids in the film coverage. After 5 seconds, the rotational speed is increased to 5000 rpm, and the centrifugal force causes the resist to spread and form a thin film over the substrate. A successful process will yield a uniform film (apart from some inevitable thickness variation at the corners / edges, known as edge bead), with no voids, comets or obvious nonuniformity. The substrate is then baked at 90°C for 180 seconds.

4.4.2 Exposure and development

The patterning of the test array shown in Fig. 4.1(c), is carried out in the scanning electron microscope (SEM) equipped with an add-on Raith exposure system. While electron beam lithography (EBL) is a mature technology originating in the 1960s [84], this approach, described by Jede et al in [85], is a cost effective way to write submicron patterns without recourse to a dedicated e-beam writer. In a conventional SEM, the focused electron beam is rastered across the substrate, and gradually stepped down in the orthogonal direction, analogous to a cathode ray tube (CRT) television imaging system. In the EBL configuration, the rastering is driven by the Raith system, so
CHAPTER 4. SILICON NANOCANTILEVERS

Figure 4.2: Fabrication process flow used for nanocantilever test array: application of e-beam photoresist, e-beam exposure and development, Si plasma etch, SiO$_2$ plasma etch, wet release etch and photoresist strip.

that the beam can be turned on and off (or blanked) to expose the desired pattern in the resist. The SEM used for this work is the Zeiss Supra. The system utilises the Zeiss Gemini column. This column was developed by Zeiss in 1994 [86], and brought a number of innovations to the market. The column is shown in Fig. 4.3, where it can be seen that the beam is formed using a Schottky field emission gun (FEG). This Schottky FEG comprises a zirconium oxide coated tungsten wire sharpened to a point which is heated to facilitate thermionic emission of electrons. This emission is enhanced through the use of the barrier lowering oxide (the zirconium oxide acts to reduce the tip workfunction). These sources are more suited to EBL where the dose is
4.4. PROCESSING

Figure 4.3: Electron optical column of Gemini column by Zeiss. The beam is maintained at 15KV right up to the final electrostatic lens, thus minimising sensitivity to stray fields, and minimising chromatic aberration [86].

current dependent, as cold emission tips suffer from unstable output as the workfunction is continually modified by the adsorption of contaminants on the tip.

The beam is then accelerated up to high voltage (for the exposure in this work, an accelerating voltage of 15KV was used). In the Gemini column, the beam is maintained at this high voltage all the way down through the column thus minimising sensitivity to stray external fields and reducing chromatic aberration. Another development introduced in this column is the position-
ing of the electron detectors inside the column on the optical axis (note the secondary electron (SE) and energy selective backscatter (EsB) detectors). The beam then continues down the objective lens where a combination of magnetic and electrostatic lens focus the beam on to the substrate. Note that it is only at the last electrostatic lens that the beam is decelerated to low voltage for high sensitivity imaging.

The column contains other components not shown in Fig. 4.3 including apertures, an electromagnetic lens, stigmators and alignment systems. The purpose of the aperture is to set the beam current, which is a key determinant in the exposure dose, as explained below. Further, by eliminating higher angle electrons, it enhances the beam convergence angle thus minimising the effects of chromatic aberration. The lens serves to focus the beam, analogous to an optical system. As electrostatic lens do not perform as well in terms of aberrations, focussing is typically implemented using electromagnetic lens. In an electromagnetic lens, the force on the electrons arising from the magnetic flux moves the electrons towards the optic axis. The stigmator is used to minimise astigmatism in the beam. Astigmatism is the condition by which the circular cross sectional shape of the beam becomes oblong as the lens setting is varied. The stigmator is typically a multi (4-8) pole electromagnetic lens, which can shape the beam back to a circular cross section.

In the EBL system used in this work, the SEM is modified to allow blanking of the beam. This beam blanker serves to divert the beam away from the centre of the column thus preventing exposure of the e-beam resist when the beam is blanked. The beam blanker comprises a pair of electrostatic plates across which a voltage is applied to divert the beam. Beam blanking is arranged to be conjugate, in other words there is no movement of the beam at the substrate upon blanking, so that there is no inadvertent exposure of
the resist as the beam moves off centre. To achieve this, the electrostatic plates are located at a beam cross over point in the column.

In addition to the blanking system, the Raith system also controls the stage movement (potentially including a laser interferometer), the digital to analogue convertors (DAC) and deflection amplifiers. The system is based on 16 bit DACs operating at 2.5 MHz. The raster scan signal (both in X and Y) is generated by the DAC output stepping from 0V to full scale deflection (FSD), where each step defines a pixel. Taking the example of a 100µm square area on the substrate or “writefield”, as used in this work, the DAC steps through $2^{16}$ (65536) pixels. For simplicity, taking 65536 pixels $\approx 50,000$, each pixel or “step size” as the beam is stepped through the writefield is

$$\text{Stepsize} = \frac{100\mu m}{50000} = 2\text{nm}$$

Note that the dose delivered per pixel of area $A$, is given by

$$D[C/cm^2] = \frac{i_{beam}[C/s].t[s]}{A[cm^2]}$$

where $D$ is the dose delivered per pixel of size $A$ during the pixel dwell time $t$, based on the beam current $i_{beam}$. For a given current, the minimum dwell time, $t$, is set by the speed of the DAC (1/2.5MHz or 400ns), thus putting a lower bound on the dose. To prevent overexposure (a significant issue with newer generation resists), the pixel size must be increased. In this work, a dose of $27\mu C/cm^2$ was used to expose the array when using zep-520 resist, necessitating a step size of

$$A[cm^2] = \frac{i_{beam}[C/s].t[s]}{D[C/cm^2]} = \frac{(26x10^{-12}.665x10^{-9})}{27\mu C/cm^2} = 64nm^2$$

based on the exposure parameters in Table 4.1 below. Using these parameters, the smallest suspended beam that could be fabricated was less than 30nm wide and is shown in Fig. 4.4.
Table 4.1: Sample EBL parameters for zep-520 e-beam resist.

<table>
<thead>
<tr>
<th>Process parameter</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>27 μC cm$^{-2}$</td>
</tr>
<tr>
<td>Step size</td>
<td>8 nm</td>
</tr>
<tr>
<td>Accelerating voltage</td>
<td>20 kV</td>
</tr>
<tr>
<td>Beam current</td>
<td>26 pA</td>
</tr>
<tr>
<td>Dwell time</td>
<td>665 ns</td>
</tr>
<tr>
<td>Aperture</td>
<td>10 μm</td>
</tr>
</tbody>
</table>

4.4.3 Dry plasma etch

Plasma etching of the structures was carried out in an Oxford Instruments System 100 ICP180 etch tool. This is a load locked tool with a 150 mm

![Image of a narrow beam with dimensions](image.png)

Figure 4.4: The narrowest beam successfully fabricated was less than 30nm in width, but with an increase in edge roughness compared to larger width beams.
lower electrode on which the wafer (carrying the substrate to be etched) is mechanically clamped. The wafer is thermally coupled to the cooled electrode by maintaining a set pressure (10 Torr in this work) of He between the electrode and wafer. This electrode is connected to a 600W 13.56 MHz RF generator via an impedance matching network. This reactive ion etch (RIE) configuration allows creation of a plasma, as electrons are accelerated under the influence of the electric field, collide and create ions. This ionisation proceeds until the gas breaks down and a plasma is formed. In this RIE configuration, an ion density of \(10^{10}\text{cm}^{-3}\) is achieved.

Consider the applied RF voltage on the lower electrode shown in Fig. 4.5 below. The applied voltage is initially symmetrical around 0V. However, after a few cycles, an offset, known as the “DC bias” is generated. To understand this, consider when the applied RF voltage on the lower electrode is positive. Electrons, which are light and therefore fast, are rapidly removed from the

![Graph showing applied RF voltage with ion and electron currents](image)

**Figure 4.5:** Origin of the DC bias: when the ion current, \(i_{\text{ion}}\), and the electron current, \(i_e\), is integrated over a cycle, the net current is zero due to the voltage offset known as the “DC bias”. This DC bias gives rise to an electric field which leads to a directional (vertical) flux of ionised etch species on to the wafer.
Figure 4.6: Si etch profile from ICP plasma etch. Note the vertical etch sidewall. This contrasts with the faceting of Si when wet etching using hydrofluoric (HF) acid, as will be described in Section 4.6, see Fig. 4.14.

plasma to the electrode, giving a relatively large spike in electron current. Now when the applied voltage goes negative, the relatively heavy slow ions are attracted from the plasma. The resulting ion current is relatively small, but lasts longer than the electron current due to the asymmetry in the voltage on the driven electrode. When the current is integrated over the cycle, it is found that the ion current exactly balances the electron current - if this was not the case, the plasma would charge and eventually be extinguished. This “DC bias” provides a vertical directional bias to the trajectory of the etching ion species, which assists in achieving vertical sidewalls, as can be seen from the Si etch profile in Fig. 4.6.

This DC bias is, however, proportional to the RF power, so that selectivity
tends to reduce if the RF power is increased to increase the process rate via the plasma density.

To decouple the process rate (via plasma density) from the DC bias, a second plasma source, such as a microwave magnetron, can be added to the chamber. In the case of the System 100 ICP180 used in this work, a cooled coil is wrapped externally around an Al$_2$O$_3$ tube. This coil is connected to a 3KW RF generator. This power is inductively coupled to the plasma, with no DC bias generated, allowing higher plasma density ($10^{12}$cm$^{-3}$) to be achieved, while retaining independent control over the DC bias via the lower (RIE) electrode.

The ICP etch process such as described above [87] was used to anisotropically etch both the device layer Si and the underlying BOX. The Si etch process utilises SF$_6$ and CHF$_3$ gases. The SF$_6$ is dissociated in the plasma providing fluorine species which react with the Si to produce volatile byproducts including SiF$_3$ and SiF$_4$. The CHF$_3$ in this process is used as a polymerising gas. As can be seen in Fig. 4.7, subject to appropriate bias, Coburn et al [88] demonstrated that when the F:C ratio drops below 4 in fluorocarbon gases (eg CHF$_3$, C$_4$F$_8$ etc), there is a tendency to deposit polymer rather than to etch. In this case, the polymer is deposited on the resist mask, thus maintaining selectivity as the Si and SiO$_2$ are etched. When etching the SiO$_2$, it is necessary to break the strong Si - O bond. The oxide etch process used in this case utilises CHF$_3$ and Ar. The Ar acts, under the influence of the ~130V DC bias, to break these bonds allowing the F to react with the Si to produce volatile SiF$_4$ (and SiF$_3$ etc). The O in the oxide film mitigates against polymerisation by reacting with the carbon in the etching species to produce volatile CO and CO$_2$. The aforementioned parameters of both processes are captured in Table 4.2.
Figure 4.7: Dependence on F:C ratio and ion bombardment energy of polymerising tendency of SiO$_2$ etch process. The process tends towards polymerisation rather than etch as the ratio decreases (adapted from [88]).

4.4.4 Wet release etch

To release the Si cantilever(s) from the underlying BOX, and achieve suspension, the BOX layer was isotropically etched using a buffered oxide etch (BOE) based release process. BOE consists of hydrofluoric acid with ammonium fluoride (NH$_4$F) as a buffering agent, and is widely used for etching silicon dioxide in the semiconductor industry [89]. The isotropic nature of this process can be seen in the undercut profile in Fig. 4.14, and will be described in more detail in Section 4.6 below. For our process, an etch rate of 100 nm/min was recorded and, based on this, an “overetch”, an etch duration of 1.8 mins was used. After this wet release etch, suspended structures can be pulled in to contact with adjacent structures due to capillary forces during
4.4. PROCESSING

Table 4.2: Si and SiO$_2$ plasma etch process parameters.

<table>
<thead>
<tr>
<th>Process parameters</th>
<th>Si etch</th>
<th>SiO$_2$ etch</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process pressure</td>
<td>11</td>
<td>6</td>
<td>mTorr</td>
</tr>
<tr>
<td>SF$_6$ flowrate</td>
<td>15</td>
<td>0</td>
<td>sccm</td>
</tr>
<tr>
<td>CHF$_3$ flowrate</td>
<td>90</td>
<td>10</td>
<td>sccm</td>
</tr>
<tr>
<td>Ar flowrate</td>
<td>0</td>
<td>40</td>
<td>sccm</td>
</tr>
<tr>
<td>He backside pressure</td>
<td>10</td>
<td>10</td>
<td>Torr</td>
</tr>
<tr>
<td>ICP power</td>
<td>1200</td>
<td>1500</td>
<td>W</td>
</tr>
<tr>
<td>RIE power</td>
<td>30</td>
<td>100</td>
<td>W</td>
</tr>
<tr>
<td>DC bias</td>
<td>-88</td>
<td>-136</td>
<td>V</td>
</tr>
<tr>
<td>Etchrate</td>
<td>181</td>
<td>128</td>
<td>nm min$^{-1}$</td>
</tr>
</tbody>
</table>

the drying process, leading to stiction related failures as noted by Guckel et al [90] and Mastrangello et al [91]. The resulting stiction, which was encountered in this work as shown in Fig. 4.8, is a well known phenomenon in MEMS processing. Several approaches to this issue can be found in the literature. For example, Abe et al [92] use elevated temperature to reduce surface tension. Another solution involves the use of temporary structural support. Mastrangelo et al [93] used polymer columns and Kobayashi et al [94] used photoresist to provide such support during the drying process, with the supports subsequently removed using oxygen plasma.

Given the severely reduced dimensions of NEMS compared to MEMS, temporary support structures are probably unsuitable. The supercritical drying process, first reported by Mulhern et al [95], which is based on bringing the working fluid (typically CO$_2$) from liquid to gas without crossing the liquid - gas boundary in the phase diagram, would appear to be a more likely solution. To eliminate this step, a much simpler approach was taken
The substrates were rinsed in an 18 MΩcm deionised (DI) water weir, and then rinsed in a solution of DI water and isopropyl alcohol (IPA), with the IPA concentration gradually increased towards 100%. The substrates were then dried in oxygen free nitrogen (OFN), and inspected in the SEM.

4.4.5 Resist strip

The final step of the process involved the use of a barrel asher to remove the polymer / resist. In the barrel asher, there is negligible capacitive coupling of the RF power, and thus negligible DC bias. An O$_2$ plasma was used to ash or burn off the resist. This process was also implemented in the ICP etch.
4.5 Experimental Measurement of Stiffness

As described previously, the stiffness of the beams was measured by using force spectroscopy performed using an atomic force microscope (AFM). AFM was invented by Binnig in 1985 [96], and is an offshoot of scanning tunnelling microscopy (STM), but is compatible with insulating surfaces unlike the latter. The AFM utilises a cantilever with a sharp (nm scale) tip located on the underside at the unclamped end of the beam. This tip is brought in to contact with a surface, and deflection of the cantilever is monitored by a laser reflecting off the end of the cantilever, on to a four quadrant detector, as shown in Fig. 4.9. There are various modes of operation including topographic imaging where the tip traverses a surface as the substrate / stage vertical position piezo is driven to maintain constant deflection (or force), thus providing a topographic map of the surface, and conductivity mapping, where the tip is under electrical bias, and current is measured as the tip traverses the surface.

In the present case, force spectroscopy is used to monitor the deflection of a fabricated nanocantilever. The technique is shown schematically in Fig. 4.10. Note that this image is not to scale for illustrative purposes, and that the AFM cantilever is typically far larger (100s μm long) than the fabricated cantilever. The substrate containing the fabricated cantilever is brought in to contact with the AFM cantilever (tip). The substrate is then raised further using the AFM piezo drive. This causes both cantilevers to flex. If the stiffness of the AFM cantilever is known, then the stiffness of the
Figure 4.9: AFM schematic: the laser beam reflects off the cantilever and is directed on to the four quadrant detector via a mirror [97]. In this figure, a triangular or double beam cantilever is used. Note that in the experiment, there is the fabricated cantilever being measured, and, separately, the AFM cantilever, see Fig. 4.10.

Fabricated beam at the point where the tip comes in contact can be extracted using the equation

\[
\frac{1}{k_{\text{comb}}} = \frac{1}{k_{\text{ref}}} + \frac{1}{k_{\text{unknown}}}
\]  

(4.7)

where \( k_{\text{comb}} \) is the measured combined stiffness of the AFM probe chip cantilever and the fabricated cantilever, \( k_{\text{ref}} \) is the stiffness of the reference or AFM probe chip cantilever and \( k_{\text{unknown}} \) is the stiffness of the fabricated cantilever. This technique assumes the stiffness of the AFM cantilever is known.

To measure the stiffness of the AFM cantilever, the cantilever reference
4.5. EXPERIMENTAL MEASUREMENT OF STIFFNESS

Figure 4.10: Reference cantilever method: (a) The AFM cantilever of known stiffness is brought in to contact with the fabricated cantilever. (b) With applied force, the combined stiffness can be measured from the deflection, and the fabricated cantilever stiffness extracted.

method described by Torii et al [98] was used. This begins with measurement of the optical lever sensitivity, or the optical detector signal for a given vertical displacement of the cantilever. This involves landing the AFM tip on a hard surface, and then moving the stage vertically using the z piezo drive a known distance while monitoring the detector signal. Such a force-displacement (F-d) characteristic is shown in Fig. 4.11 below. At location 1, the AFM cantilever is located above the surface. As the tip approaches the surface, short range forces begin to pull the tip into contact with the surface. When this short range force exceeds the cantilever spring restoring force, the beam abruptly comes in to contact. This can be seen from the deflection of the AFM cantilever in the schematic (Fig. 4.11(a)) and the downward trajectory of the force curve (Fig. 4.11(b)) both at location 2. The Z piezo continues driving the stage up, and the measured force increases as shown at location 3 (known as the loading region). When the target force is reached, the Z piezo then changes the stage direction, unloading this force, shown at location 4. As the force is reduced, the AFM cantilever returns to the neutral position, while remaining in contact. As the stage continues to move down,
Figure 4.11: A force displacement measurement is shown schematically in (a). Jump to contact is shown at location 2. Locations 3 and 4 show loading and unloading, and surface adhesion can be seen at location 5. A typical measured characteristic is shown in (b).

the adhesive force between the tip and the surface causes the cantilever to deflect until the adhesive force is overcome by the spring restoring force of the AFM cantilever. This is shown at location 5, and the pull-off force can
be read from the characteristic as shown.

The slope of the characteristic in the loading region gives the inverse optical lever sensitivity. The next step is to record a thermal power spectrum to determine the beam resonant frequency. The inverse optical lever sensitivity is then combined with this resonant frequency in Hutter and Bechhoefer’s equipartition theorem [99] to give the beam stiffness.

Once the AFM beam stiffness has been determined, the fabricated beam under test is then mapped by performing an array of force - displacement (F-d) curves along the beam, and the stiffness of the beam at the location along the beam where the AFM probe chip tip comes in to contact, can be extracted from Eqn. 4.7. This reference cantilever method [98] was used over other methods ([100], [101], [102]) for a number of reasons such as inadequate reflectivity from the fabricated nanocantilevers, and the ability of the reference cantilever method to map the stiffness along the length of the fabricated nanocantilever with nanoscale precision.

Indeed it is especially important that the precise location is known in view of the cubic dependence of beam stiffness on length, as was seen earlier in Eqn. 4.3.

4.6 Experimental Results and Analysis

Fig. 4.12 shows a pixelated map where each pixel represents a location where a force - displacement characteristic was recorded (this map is generated from the pull-out force measured at each pixel location, and is for visualisation purposes only). It is helpful to plot this, as the beam outline is clear (cf. 4.12(b)), and the locations along the beam where each individual F-d characteristic was measured is also clear. As noted above, given the cubic
CHAPTER 4. SILICON NANOCANTILEVERS

Figure 4.12: (a) Outline of cantilever, with locations where F-d curves were measured denoted by red markers. (b) Schematic of cantilever showing F-d location markers for the beam in (a).

dependence of stiffness on length, it is important to establish the precise location of each measurement. Based on the schematic showing a fabricated beam in Fig. 4.12(b), the position, noted by markers, along the beam where each F-d measurement has been recorded, is evident. In this way, the stiffness of the beam can be calculated from the F-d characteristic taken at each of the labelled pixel locations in Fig. 4.12(a). To generate the map in Fig. 4.12(a), the AFM steps the probe chip over a given area in 20nm steps in both X and Y directions. The stiffness of one such beam is plotted against location or beam length in Fig. 4.13. This beam has a designed length of 550nm, width of 85nm and thickness of 50nm. It can be seen that the measured stiffness is significantly lower than that predicted by an analytical model, assuming a well defined single crystal beam of these dimensions. There are a number of factors behind this discrepancy. The techniques of nanofabrication introduce a number of non idealities in terms of the beam geometry and composition.
Figure 4.13: Spring constant as a function of beam length for various analytical and FEA models shown on the right hand side. The blue curve models a rectangular beam without undercut, the orange/red curve models a rectangular beam with undercut and the green curve models a trapezoidal beam with SiO₂ coating. The solid curves are from analytical models, the dashed curves are from FEA models and the black dots are measured data.

As noted above, to release the beam (i.e. to achieve suspension), an isotropic hydrofluoric acid etch process was used. The length of the beam is inadvertently increased due to the undercut of the BOX beneath the beam at the clamped end. The isotropic nature of the etch, and this undercut is visible in Fig. 4.14. As noted by Guillon et al [24] and Tsuchiya et al [25], this undercut is known to impact directly on the resonance frequencies of silicon
nanocantilevers fabricated in this way. It is therefore to be expected, that the stiffness is similarly impacted. Indeed, when the beam length is adjusted for the measured undercut, the stiffness, shown by the orange curve, is closer to the experimentally determined value. Notwithstanding this improvement, however, there was still a significant difference between the measured value, and that predicted by the analytical model. To investigate this discrepancy, a cross section was taken through the beam, and the beam geometry and composition were investigated using a transmission electron microscope (TEM). In a TEM an electron beam is transmitted through the sample, and the interaction of the beam with the sample is imaged. It allows the user to image the sample with atomic resolution, and, in this work, it provided important insight into the nature of the fabricated beams. The technique requires the use of electron thin samples (50 - 100nm), as outlined by Giannuzzi [103]. Such a thin lamella was prepared using a focused ion beam (FIB) microscope.

**Figure 4.14:** Cross sectional images of beam revealing nonidealities: (a) FIB cross sectional image through two beams. The isotropic wet etch profile is clear. This undercut also occurs at the clamped end of the beam effectively increasing the beam length, (b) TEM cross sectional image showing trapezoidal beam shape and oxidation of beam caused by resist strip process, and (c) faceting of Si caused by HF based wet etch release process, confirmed by angular measurement.
The FIB (Zeiss Auriga) utilises a liquid metal ion source (LMIS) where Ga metal wets a tungsten filament and is drawn to the tip of the filament, where a Taylor cone [104] is formed by a very high electric field. Ga species are thus ionised, and thermionically emitted from the filament in a focused ion beam with a diameter of 2nm. This beam is then used to sputter away Si until a suitably thin lamella is formed.

A cross section containing two adjacent suspended beams is shown in Fig 4.14(a). There are a number of key observations to be made from this image. Firstly, the aforementioned isotropic etch profile of the HF based wet etch release process is clear from the curved SiO$_2$ sidewall which is labelled in the figure (compare this with the vertical profile of the dry Si etch shown in Fig. 4.6). As already noted, the isotropic undercut at the clamped end of the beam increases the actual beam length beyond the design value. Secondly, close inspection of the FIB cross section would suggest that the Si etch profile is not vertical as shown in the dry etch profile in Fig. 4.6, but angled. This is caused by faceting as is clearly shown in Fig. 4.14(b). Faceting of Si is a well known phenomenon [105], and is used industrially in the creation of via structures in Si substrates. However, the nanoscale faceting of Si arising from HF etching is of less relevance in an industrial context, but clearly of major significance in the present application. Recall that the stiffness of the beam is given by

$$k = \frac{F}{x} = \frac{3EI}{L^3} \quad (4.8)$$

where $E$ is Young’s modulus, $L$ is the beam length and $I$ is the moment of inertia. For a rectangular beam, the moment of inertia is given by Eqn 4.9

$$I = \frac{wt^3}{12} \quad (4.9)$$

where $w$ is the beam width, and $t$, the thickness. For the trapezoidal beam shape that arises from the HF based release process, the moment of inertia
is given by Eqn. 4.10 [106].

\[ I = \frac{d^3 b^2 + 4bc + c^2}{36 b + c} \]  

(4.10)

where \( b \) is the width at the base of the beam, \( c \) is the width at the top of the beam and \( d \) is the height of the beam, see Fig. 4.15(c).

Yet another discrepancy from an ideal beam caused by the nanoscale processing techniques is also visible in Fig 4.14(b). It was noted above, that the e-beam resist was stripped or removed using a relatively high power (2KW) oxygen plasma. Oxygen plasma ashing is a standard technique for resist removal, but in this case, a relatively thick (\(~8\text{nm}\)) oxidation of the Si beam results. Oxidation of Si in ashing processes is a known issue, for example, see analysis by Han et al [107], but the high power process used here is not typical of industrial processes. In Fig. 4.15, the impact of both the trapezoidal beam shape caused by faceting of the Si during the wet release etch, and the compositional inhomogeneity caused by a relatively thick oxide surrounding the beam is quantified. In Fig. 4.15(a), the moment of inertia is calculated for the ideal beam based on the design. This is first modified in Fig. 4.15(b) to reflect a rectangular beam with the dimensions taken from the TEM. In Fig. 4.15(c), Eqn. 4.10 above for the trapezoidal beam is used to calculate the moment of inertia for the trapezoidal beam with dimensions taken from the TEM image. Finally, this value is adjusted to account for the oxidation of the beam, again revealed by the TEM imaging. Overall, the moment of inertia is found to drop from $6 \times 10^{-31}\text{m}^4$ for the ideal designed beam to $9.38 \times 10^{-32}\text{m}^4$ for the beam revealed by TEM inspection.

In the analytical model used to calculate the beam stiffness, these discrepancies are built in by using a correction factor $\alpha_{BCF}$ of 0.15 for the moment of inertia, and by increasing the beam length, $l_a$ based on TEM analysis,
4.6. EXPERIMENTAL RESULTS AND ANALYSIS

Figure 4.15: Moment of inertia for (a) an ideal rectangular beam, (b) a beam with measured dimensions from SEM / TEM, (c) a trapezoidal beam with measured dimensions from SEM / TEM and (d) for beam of measured dimensions and Si / SiO$_2$ composition.

giving a stiffness, $k'$ of

$$k' = \frac{3EI\alpha_{BCF}}{(l + l_u)^3}$$  \hspace{1cm} (4.11)$$

In Fig. 4.13, this corrected analytical model is shown, and is found to be in agreement with the experimental data. To further strengthen the model, it was felt that the complexity of a realistic nanoscale beam would lend itself to a finite element analysis (FEA). FEA was felt to be particularly suitable for the modelling of realistic nanoscale beams due to their complex geometry, specifically the nature of the undercut device layer Si in the vicinity of the clamped end of the beam, and the compositional variation. In FEA, the geometric domain is represented by a mesh of individual elements which make up the entire physical system. Each of these elements is modelled,
and the results are systematically combined to model the behaviour of the overall system. In the present case, an FEA model was constructed to model a rectangular beam clamped at one end by a “shelf” of Si (the undercut device layer Si as revealed by TEM) as shown in Fig. 4.16(a). This approach confirmed the stiffness predicted by the simple analytical model adjusted for the undercut as indicated by the dashed red curve in Fig. 4.13. A second FEA model was then built to account for the trapezoidal shape and composite material. In Fig. 4.16(b), the meshes for the Si beam (or “core”), the SiO$_2$ “sleeve” and the combination with the detail of the clamping is shown. This mesh comprises second-ordered reduced integration tetrahedral elements [108]. With these enhancements to the original FEA model, good agreement with the experimental data was achieved, as can be seen from the dashed green curve in Fig. 4.13.
Figure 4.16: FEA model (a) rectangular model with undercut and clamping details, (b) model modified to account for 8nm SiO₂, (c) Si core and (d) trapezoidal Si / SiO₂ beam [108].
In order to confirm the effect of the oxide on the beam stiffness, a further experiment was carried out, where the stiffness of the beam after standard processing (i.e. with the plasma induced 8nm SiO$_2$ film intact), was measured. This oxide was then stripped with a brief dip in buffered oxide etch. The stiffness was then remeasured. The stiffness of the beam was modelled first to include the 8nm oxide film, and then to include an expected 1-2nm thick native oxide on the Si (the stiffness measurements were performed some time after the oxide strip, and were performed in ambient conditions). The experimental and modelled results are plotted in Fig 4.17.

With the removal of the oxide, the thickness of the beam is reduced,

![Figure 4.17: Stiffness of beam against length or location of applied force before (in blue) and after (in red) removal of the process induced SiO$_2$. With the removal of the SiO$_2$ using buffered oxide etch (BOE), the stiffness of the beam is reduced as predicted by the model.](image-url)
causing a softening of the beam, and this was confirmed by measurement, again with excellent agreement with the model results [108] as can be seen in fig 4.17.
4.7 Conclusions

In this chapter, the fabrication of an array of nano scale Si cantilevers using CMOS type processing was described. While EBL is not CMOS compatible as it is an inherently sequential process, similar lithographic patterning is used in CMOS processing. Although the process was optimised, for example achieving beam widths down to 30nm, with well defined vertical profiles, it was found that the techniques of CMOS type fabrication introduce nonidealities in to the beams.

The stiffness of the beams, which is critical in nanorelay device operation, was measured by performing force spectroscopy in an AFM. In this technique, the AFM probe tip lands on the fabricated cantilever, and loads a pre-defined force. As both cantilevers flex, the stiffness of the fabricated cantilever can be extracted assuming the AFM cantilever stiffness is known (this was measured independently).

Analytical models of beam stiffness were reviewed and tested against the fabricated cantilever array and found to predict a higher stiffness than that recorded using AFM based force spectroscopy. This was found to be attributable to the aforementioned fabrication related non-idealities in the beams. Specifically, the wet release etch undercut the device Si layer lengthening the beams, and this HF based etch also modified the cross sectional shape of the beams by faceting the Si, and the final resist strip process involving an oxygen plasma created a relatively thick oxide film on the beams. The models (both analytical and FEA) were modified to accommodate these non-idealities, and were subsequently validated.

These models will prove useful in predicting the mechanical behaviour of fabricated cantilever beams as nanorelay device technology is scaled to the nano regime.
Chapter 5

Nanorelays at the 22nm technology node

5.1 Introduction

In Chapter 4, the stiffness or spring restoring force of realistic nanoscale single crystal Si cantilever beams was explored using both experimental methods and analytical and finite element modelling. This stiffness or spring restoring force determines the maximum adhesive force that the design can support. While it is desirable to minimise this adhesive force to drive both CMOS operating voltage compatibility and energy efficiency, an additional requirement for current flow must also be met. These criteria can be somewhat contradictory. For example, metallic materials typically offer low resistance contacts, but suffer from relatively high adhesion owing to the metallic bonds and large Hamaker constants associated with metal–metal interactions, whereas dielectric materials might be expected to more easily meet the requirement for low adhesion, but are typically poor conductors. In addition, the contact materials must be physically robust to withstand repeated impact over the
To explore these criteria in more detail, a device specification based on Intel's 22nm technology node [29] (which had just been ramped to production at that time) was developed, and is described in this chapter. These criteria assume a single crystal silicon cantilever beam compatible geometrically with a CMOS transistor at this technology node, and an operational voltage of 1V. A realistic circuit topology is selected, and reasonable drive currents for low power logic and graphics circuits are chosen to determine target electrical specification. This “device specification” was selected in order to identify realistic beam dimensions, and thus stiffness / pull-out force for a pull-in voltage of 1V. In addition, a basic electrical specification based on a typical circuit (4 input AND block) was designed to guide a realistic current which a nanorelay would have to pass. This specification therefore provides a basic framework to assess nanorelay operation at the 22nm node. A more detailed specification would arise for a specific application. For example, NEMS have been proposed for power gating, where the control device (whether MOS transistor or nanorelay) is relatively large, reflecting the high current requirement for this application. In this work, an electrical RC time constant, $\tau_{\text{elec}}$, was somewhat arbitrarily specified to be half the mechanical time constant, $\tau_{\text{mech}}$, based on the modelled beam resonant frequency, as shown in Fig. 5.5, and the required resistance ($\tau_{\text{elec}} = RC$) and current was derived from this. In reality, high frequency operation is challenging as switching speed is governed by the mechanical delay arising from the beam travel time, and circuit design techniques which allow all nanorelays to switch simultaneously are proposed, for example by Chen et al [74] to avoid delays arising from sequential switching.

In Chapter 4, the AFM was used to measure the stiffness of a range of
5.1. **INTRODUCTION**

fabricated cantilevers. The AFM can also be used to test the dimple - source interface in a nanorelay in terms of both adhesion and electrical performance. In this chapter, the use of an AFM to test a range of candidate materials for the contacting interface in a nanorelay is described. To test the adhesion performance, the tip at the end of an AFM probe chip cantilever was engineered to replicate a nanorelay contact dimple, and this was brought in to contact with a substrate representing the source terminal of the nanorelay. By recording F-d measurements as described in Section 4.5, the adhesion of this representative dimple - source combination could be measured. Furthermore, by applying electrical bias in the manner described by Lanza [109], a variation of AFM known as conductive AFM, c-AFM, was used to test the electrical conductivity of this dimple - source interface. The implementation of a low noise c-AFM system is described, and experimental results are presented and discussed.

A range of materials was then selected and tested both experimentally and theoretically to investigate the capacity of these material systems to meet the criteria for incorporation of nanorelays at this technology node. Finally, it should be noted that while the 22nm technology node provided a realistic dimensional guide for this work, the results of this work have implications for the integration of nanorelays in CMOS technology beyond the 22nm node. The fundamental challenges of minimising stiction while maintaining electrical conductivity at the contact is key to the integration of these devices, whatever the technology node.
5.2 The 22nm Technology Node

At the 22nm node, Intel continued on the trajectory outlined by Gordon Moore in 1968 [1]. While maintaining this aggressive scaling, performance was delivered by further developing many innovations such as strained Si [110] and high k metal gate stacks [111], followed by a move from planar devices to the trigate FinFET geometry [29]. Fig. 5.1 shows this FinFET geometry where the channel forms a fin which is surrounded on three sides by the gate (this device also features the aforementioned strained Si and high k / metal gate innovations). It is interesting to note that even with this enhanced trigate design, the subthreshold slope (72mV/dec for PMOS and 69mV/dec for NMOS) cannot compete with that of nanorelays for the reasons described in Chapter 2, see Fig. 5.2. It was at this node that the proposed intercept by nanorelay technology was targeted in this work.

Figure 5.1: Three dimensional schematic and cross section micrograph of 22nm node transistor showing how the non-planar “trigate” design improves coupling of the gate to the channel over planar designs. Adapted from [29].
To determine design rules, the characteristic size of the transistor was used to guide the dimensions of the cantilever beam. Based on these dimensions, and a requirement for operation at a supply voltage of 1V, the dimensions of the beam were optimised using numerical modelling to determine the maximum pull-off force. The electrical performance targets were then determined by the current required in a realistic circuit topology operating in the 10-100MHz speed regime.

![Figure 5.2: Device characteristics for 22nm technology node PMOS and NMOS trigate transistors. Note that with strained Si channels, high k / metal gate stacks and the trigate FinFET geometry, the subthreshold slope values of 69 and 72 mV/dec approach the theoretical limit of 60mV/dec as discussed in Chapter 3. Adapted from [29].](image-url)
5.3 Device Specifications

5.3.1 Mechanical specifications

According to Auth et al [29], critical dimensions at the 22nm node are as shown in Table 5.1 below. Using beam dimensions reasonable for this technology node, and with the stipulation that the minimum gap across the contact in the OFF state was no less than 2nm (imposed by quantum mechanical tunnelling as described by Razavy in [112]), the pull-out force was calculated using a numerical model, based on the device shown in Fig. 5.3. The beam has a width $W$, length $L$, gap $G$ and thickness $T$. In the numerical model, the beam is divided up, as shown in Fig. 5.3, into finite lengths with length $dL$. Recall that the device operates by electrostatic attraction between the beam and the gate electrode. In the model, the capacitance of each segment is calculated using Eqn. 5.1 from Yuan et al [113], and the associated force is calculated using Eqn. 5.2, where $\epsilon_{ox}$ is the dielectric constant of the medium between the gate and beam electrodes.

$$dC = \epsilon_{ox} \left( 1.15 \frac{W}{G} + 2.80 \left( \frac{dL}{G} \right)^{0.222} \right)$$

(5.1)

$$dF = \frac{1}{2} dC \left( \frac{V_{app}}{G} \right)^2$$

(5.2)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transistor fin</td>
<td>60 nm</td>
</tr>
<tr>
<td>Transistor gate</td>
<td>90 nm</td>
</tr>
<tr>
<td>Interconnect pitch</td>
<td>80 nm</td>
</tr>
<tr>
<td>SRAM cell size</td>
<td>0.092 $\mu$m$^2$</td>
</tr>
</tbody>
</table>
As the beam accelerates towards the lower electrode, the electrode separation or gap, $G$, decreases and this capacitance (and force) increases. The model calculates the superposition of the force from each of these segments, and iterates until the cantilever reaches the pull-in condition. The minimum voltage required to attain pull-in is the pull-in voltage, which is determined by the beam length, thickness, width and gap as described by Eqn. 3.4. On the other hand, given that the beam spring constant is given by Eqn. 4.3, the pull-out force can be written as

$$F_{\text{pull-out}} = \frac{EWT^3}{4L^3}G$$  \hspace{1cm} (5.3)

Assuming a Young’s modulus of, $E$, of 170 GPa, Fig. 5.4 maps out the pull-out force as a function of beam dimensions. This is the force available to pull the beam out of contact, and sets the upper bound on surface adhesion forces that must be overcome to support relay operation. In Fig. 5.4(a), the gap and beam thickness are held constant, while the width and length are
Figure 5.4: Nanorelay design space: Pull-out force as a function of operating voltage (target = 1V); (a) varying beam length and width with contact gap and beam thickness held constant and (b) varying contact gap and beam thickness with fixed beam width and length. Note maximum pull-out force at 1V is 0.4nN.
varied. The beam width changes the fringing capacitance, but the beam length has a stronger impact on the stiffness, and therefore actuation and pull-out force, of the beam. In Fig. 5.4(b), the beam length and width are held constant while varying the thickness and gap. It can be seen that the maximum pull-out force of 0.4nN (determined by the stiffness of the beam at these dimensions) at 1V operating voltage is achieved with a beam length of 200nm, width of 8nm, thickness of 18nm and the (quantum tunnelling limited) gap of 2nm. Thus, to support relay operation, the adhesion force of the (conductive) contact must not exceed 0.4nN.

Further modelling was carried out to clarify the attainable operational frequency of the device. The maximum operating frequency of the relay will be determined by the beam resonant frequency. According to Liu [114], the resonant frequency, $f_{\text{res}}$, is given by Eqn. 5.4.

$$f_{\text{res}} = \left( \frac{1.875^2}{2\pi L^2} \right) \left( \frac{E(WT^3)}{12 \rho WT} \right)^{\frac{1}{2}}$$  \hspace{1cm} (5.4)

where $\rho$ is the density of the relay material. In Fig. 5.5, the resonance frequency of the device is plotted against the beam dimensions in a similar manner to the modelling of the pull-out force above. It can be seen that over the range of 100MHz to 1GHz, the resonance frequency is sensitive to length and thickness, and insensitive to gap and width, as predicted by classical mechanics (reducing the length and increasing the thickness makes the beam stiffer, increasing the resonance frequency). From Fig. 5.5, the resonant frequency of a cantilever beam within this dimensional range is from $\sim$100MHz up to $\sim$1GHz, suggesting a mechanical time constant, $\tau_{\text{mech}}$, of $\sim$1ns.

In summary, based on the above considerations, an optimised device compatible with 22nm technology, will have a length of 200nm, a width of 8nm,
Figure 5.5: Nanorelay design space showing resonance frequency as a function of actuation voltage for (a) a contact gap of 2nm and a beam thickness of 8nm, with the beam length and width as variables and (b) a beam length of 200nm and width of 8nm, with the contact gap and beam thickness as variables.
a thickness of 18\,nm and a gap of 2\,nm. It is interesting to compare these dimensions with the fabricated beam described in Section 4.6, where the width of the trapezoidal Si beam was 10\,nm at the top increasing to 30\,nm at the bottom, and the height (or thickness) of the beam was approximately 20\,nm. However, the moment of inertia was modified by the geometrical and compositional nonidealities explained earlier. The proposed nanorelay will support a maximum adhesion force of 0.4\,nN, and will be compatible with the 10-100\,MHz frequency range required for low power / graphics applications.
5.3.2 Electrical specifications

To establish electrical specifications, a nanorelay based implementation of a four input AND gate was considered, as discussed by Chen [74] and Fariborzi et al [67], [115]. Consider the CMOS implementation of this four input AND gate in Fig. 5.6(a). If the CMOS transistors in this circuit were directly replaced with nanorelays, then circuit switching would entail four mechanical delays as the signal propagates from one gate to the next through all four sequential stages. To avoid this problem, the circuit is simplified when implemented using nanorelays as shown in Fig. 5.6(b) so that all relays switch simultaneously resulting in just one mechanical delay, as described by Chen et al [74].

Note that the nanorelays in this circuit are four terminal devices. Fig. 5.7 shows plan view and cross section schematics for 2, 3 and 4 terminal nanorelays. The simplest device is the two terminal switch where the voltage between the source and drain causes the switch to close. While this is of limited circuit use, this design was used in this work for comprehensive analysis of the beam mechanics at the nanoscale. This device can be made more prac-

![Figure 5.6](image-url)

**Figure 5.6:** (a) CMOS implementation of a four input AND gate. The signal propagates sequentially from gate to gate through four stages. (b) Circuit implementation using nanorelays. In this case all devices switch simultaneously, helping to improve circuit speed [74].
5.3. DEVICE SPECIFICATIONS

Figure 5.7: (a) Plan view, (b) side view and (c) circuit symbol for 2 terminal relay where the voltage between the source and drain actuates the device, (d) plan view, (e) side view and (f) circuit symbol for three terminal relay, where a separate gate terminal now controls device actuation, (g) plan view, (h) side view and (i) circuit symbol of four terminal relay, where the device is actuated by the voltage between the gate and body terminals. By independently biasing the body terminal, the device actuation voltage on the gate can be adjusted. Adapted from [21].

tical by adding a third, gate, terminal as shown in Fig. 5.7 (d) - (f). Here the drain terminal is reduced in size, and its actuation function (where it
acts as one plate of the actuating parallel plate capacitor) is performed by a separate gate terminal. In this case, switching is controlled by the input terminal, which is isolated from the output (drain) terminal, analogous to a MOSFET device. The nanorelays shown in Fig. 5.7 (g) - (i) are four terminal devices. In the latter case, switching is controlled by the voltage between the gate and a body terminal, and the source and drain terminals are adjacent and on the same plane. They are connected by a conducting channel, which is mechanically coupled to, but electrically isolated from, the cantilever beam, which acts as the gate. There are several advantages to the four terminal configuration. When three terminal nanorelays are wired in series, the source terminal of one device is connected to the output of the preceding device, so the source voltage is not fixed, leading to instability in the control signal \( V_{GS} \) of the device. This is overcome in the four terminal device, where the control signal \( V_{gb} \) is fixed and independent of the source terminal. The ability to independently bias the body allows dynamic control of the logic function, and additional source drain contact pairs can be added to implement complex logic functions in a single device.

If we specify the electrical time constant of such a nanorelay to be no more than half the mechanical time constant, and assume a mechanical time constant of 1 ns (for a resonance frequency of 1GHz as discussed above), then the RC constant must be no greater than 0.5ns. This RC constant is determined by the relay resistance (dominated by the contact) and the device capacitance plus the capacitance of the nanorelay interconnect wiring. Taking the interconnect capacitance at this technology node to be \( \sim 1.8 \) pF/cm from the ITRS [26], and assuming 1\( \mu \)m of interconnect wiring per relay, the interconnect capacitance per relay is \( \sim 0.18 \) fF. The interconnect capacitance is far greater than the nanorelay device capacitance when modelled as a par-
Table 5.2: Device specification for nanorelay at the 22nm technology node.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage</td>
<td>1 Volt</td>
</tr>
<tr>
<td>Current</td>
<td>100 nA</td>
</tr>
<tr>
<td>Frequency</td>
<td>100 MHz</td>
</tr>
<tr>
<td>$F_{\text{pull-out}}$</td>
<td>0.4 nN</td>
</tr>
</tbody>
</table>

Parallel plate capacitor with the beam and the gate as the electrodes with an air dielectric. Based on the dimensions described above, the device capacitance is about 0.01 fF. Thus to achieve the RC delay of 0.5 ns, the device resistance must be targeted. In the 4 input AND gate in Fig. 5.6(b), the conduction path can include up to four relays, thus

$$
\tau_{\text{mech}} = (4X) RC \implies R = 164.5K\Omega \quad (5.5)
$$

This suggests a relay contact capable of passing a current of 6µA assuming 1V operating bias.

For low speed logic and graphics applications, the frequency can be reduced to 10 - 100MHz, allowing a relaxation of the current requirement to 100nA - 1µA. These requirements are summarised in Table 5.2.
5.4 Experimental Approach

In Section 5.3, nanorelay mechanical and electrical specifications were developed for incorporation of these devices at the 22nm technology node. It was seen that a CMOS compatible nanorelay would have a pull-out force of 0.4nN available to overcome surface adhesion force at the contact. Furthermore, it was noted that the contact must be able to carry a current of 100nA - 1μA. To test candidate material systems against these criteria, AFM was used.

For the mechanical specification (adhesion), it has already been seen in Section 4.5, that AFM can be used to measure surface adhesion force, as shown in Fig. 4.11. In this application, the tip on the underside of the AFM cantilever is used to represent the dimple in a nanorelay to measure adhesion at the dimple-source interface. In the case of Si, the tip could be either Si with a native oxide or Si passivated by hydrogen following a brief etch in HF acid followed by immediate transfer to the AFM vacuum chamber. Diamond tips, shown in Fig. 5.8, were also used. Two types of diamond based probe chips were used. Fig. 5.8(a) shows an ultrananocrystalline diamond (UNCD) probe chip, where both the cantilever and the tip are UNCD. The cantilever is 225 μm long, and is designed to have a spring constant of <1N/m. The probe chip in Fig. 5.8(b) is a nanocrystalline diamond (NCD) probe chip. Unlike the UNCD probe chip, the NCD probe chip is designed to have a stiffness on the order of 100N/m (which is suitable for tapping mode operation). This chip has a composite nature which is shown in Fig. 5.8(b). The chip is fabricated from a tapping mode Si probe chip. This is then coated with a 1μm NCD film. A plasma etch process is then used to shape the tip and reduce the thickness of this film. As shown in Fig. 5.8(b), this process leaves a 400nm thick NCD film coating both the cantilever and the tip. The tip is shown in the SEM micrograph in Fig. 5.8(c). A range of substrates were used to test various
5.4. EXPERIMENTAL APPROACH

Figure 5.8:  (a) UltraNanoCrystalline Diamond (UNCD) AFM probe chip with a 1N/m UNCD cantilever, (b) a NanoCrystalline Diamond (NCD) probe chip, which is fabricated by coating a dynamic mode Si probe chip with 1μm NCD, and shaping of the tip by plasma etching and (c) SEM micrograph of NCD tip with 25nm radius.

source materials including Si, graphene, highly ordered pyrolytic graphene (HOPG) and UNCD. For the electrical specification (current), the technique of conductive AFM (c-AFM) was used. In conductive AFM, the sample is electrically biased, and the circuit is completed by providing a connection to

Figure 5.9:  Low noise DC measurement circuit. Current flows from the biased substrate to signal ground via the AFM tip / cantilever of the probe chip. The signal, guard and ground are taken to the low noise amplifier located immediately adjacent to the AFM head, and the amplified signal is converted to a voltage encoded current signal in the Keithley meter for synchronous display by the AFM mainframe.
CHAPTER 5. NANORELAYS AT THE 22NM NODE

Figure 5.10: Photograph of AFM head with sample *in situ*. Note the proximity of the low noise amplifier.

ground via the AFM probe chip. The circuit schematic for the setup used in this work is shown in Fig. 5.9, and consists of a low noise preamplifier with subfemptoamp input bias current located in close proximity to the sample as shown in Fig. 5.10.

The biased sample is mounted on a custom designed mount which comprises an iron magnetic clasp isolated from an iron guard ring which is in turn isolated from the iron platen on which the sample is mounted, as shown in Fig.5.11. The signal current is fed to the preamp via a coaxial cable, the sheath of which is connected to the guard ring. The AFM probe chip holder was modified to isolate the probe chip from chassis ground, and the probe chip provides the signal ground. The output of the preamp (20 mA current loop) is converted to a voltage encoded current signal by the Keithley meter. This signal is then digitised and displayed synchronously with the force signal by the AFM mainframe digitiser.

To convert the measured current to current density, the contact area must be estimated. In this work, a Hertzian contact was assumed. The Hertzian
model was developed by Heinrich Hertz in 1881 [116], and gives the contact area as

\[ a_{Hertz} = \left( \frac{R_{tip} F}{E_{TOT}} \right)^{\frac{1}{3}} \]  

(5.6)

where \( a_{Hertz} \) is the contact radius, \( R_{tip} \) is the radius of the AFM tip, measured by scanning electron microscope (SEM), and \( E_{TOT} \) is the tip - substrate combined modulus given by

\[ \frac{1}{E_{TOT}} = \frac{3}{4} \left( \frac{1 - \nu_{tip}^2}{E_{tip}} + \frac{1 - \nu_{subs}^2}{E_{subs}} \right) \]  

(5.7)

where \( \nu_{tip} \) is Poisson’s ratio of the tip, \( \nu_{subs} \) is Poisson’s ratio of the substrate, \( E_{tip} \) is the modulus of the tip and \( E_{subs} \) is the modulus of the substrate. For example, in the case of an NCD tip (or dimple) on a Si substrate (source)

\textbf{Figure 5.11}: Substrate mount for c-AFM. The substrate is mounted on a conductive iron plate to which bias is supplied. This plate is isolated from the guard ring which is in turn isolated from the base plate. The base plate acts as a magnetic clasp in this sandwich structure. The figure also shows the AFM probe chip holder, which can be seen \textit{in situ} in Fig. 5.10.
described in detail in Section 5.5, it was found that the NCD tip with radius of 25nm shown in Fig. 5.8(c), loading a force of 100 nN results in an estimated contact area of approximately 17nm$^2$.

In this way, the AFM was used to test both the adhesion and conductivity of a range of materials for use as the dimple and source contact in a nanorelay.

5.5 Results

5.5.1 The (U)NCD - Si dimple - source material system

Using the AFM based techniques described in Section 5.4 to measure adhesion and electrical conductivity, a range of materials was investigated for use as the dimple - source interface in a nanorelay. The material combinations considered included Si (including hydrogen terminated Si), NCD, UNCD (including fluorinated UNCD), graphene and highly ordered pyrolytic graphene (HOPG). In this section the result of the investigation using both NCD and UNCD probe tips on a Si substrate is described in detail. In the experiment, both a UNCD probe chip (shown in Fig. 5.8(a)), and an NCD coated probe chip (shown in Fig. 5.8(b)) were used to replicate a nanorelay dimple, and a Si substrate was used to replicate the source. Using the UNCD - Si combination first, the force - displacement and current voltage data shown in Fig. 5.12 was recorded. To accurately record the adhesive and electrical performance of a UNCD dimple and Si source contact using this technique, it is important to controllably land the AFM tip (or “nanorelay dimple”) on the substrate (or “source terminal”), and then load up a force with minimal lateral movement or skating along the surface, and then to unload and “pull-
Figure 5.12: UNCD - Si force displacement curve. The AFM cantilever is soft, and the deflection exceeds the dynamic range of the AFM optical detector, at the inflection point ($z \approx 0.6 \mu$m) in the load (black) and unload (red) curves. The current is shown by the green and orange traces.

off” the same well defined contact area to record the adhesion, while under electrical bias. In Fig. 5.12, it can be seen that the deflection of the beam exceeds the dynamic range of the optical detector (at a substrate $z$ piezo extension of $\sim 0.6 \mu$m). The optical detector runs out of dynamic range when the cantilever beam flexes excessively and directs the laser reflection off the four quadrant detector (at this point, it can be seen that the load part of the curve abruptly levels off). In fact, it was found that the tip was skating along the surface, and was actually damaging the Si surface by digging in to the Si as shown in Fig. 5.13.

In Fig. 5.12, the current flow is shown by the green curve (extend) and the
orange curve (retract). It can be seen that no current flows until significant force is reached. This is attributed to native oxide on the Si, which must be penetrated before electrical contact is established. Also of note in Fig. 5.12, is the increase in current when the tip changes direction from extend to retract. The maximum current achieved is approximately 20pA under a bias of 1V, but great care must be exercised in interpreting this data due to the unstable and irreproducible contact area at the interface due to the aforementioned skating / machining of the surface.

To address this issue, the AFM probe chip shown schematically in Fig. 5.8(a),
with a long low stiffness UNCD cantilever designed for contact mode imaging, was replaced with a probe chip shown schematically in Fig. 5.8(b), with a short high stiffness cantilever designed for tapping mode operation. As described in Section 5.4, this AFM probe chip has a 1μm thick NCD film deposited along the cantilever out of which a dimple type tip was fashioned, using plasma etch processing, at the unclamped end shown in the SEM micrograph in Fig. 5.8(c). The data for the experiment using this probe chip is shown in Fig. 5.14. Unlike the plot in Fig. 5.12, the deflection of the beam (shown by the black / extend and red / retract curves) does not exceed the dynamic range of the optical detector. In fact the setpoint load force is achieved with a z piezo movement after landing of about 30nm, reflecting minimal horizontal movement of the tip or dimple along the Si surface. In

![Figure 5.14: NCD - Si force displacement curve. Skating has been essentially eliminated (note ~30nm z piezo movement), allowing a more stable and reproducible contact for electrical measurement.](image-url)
Fig. 5.14, the vertical movement of the Z piezo has been reduced from 2μm to about 30nm, and the beam remains stationary on the surface, allowing accurate measurement of the electrical performance of the contact.

The combination of Si and diamond was found to perform well electrically. For example, over a wide range of materials, a Si - UNCD (dimple - source) combination gave the highest current density of approximately 10 nA/nm² (at 1V bias), but was found to be more adhesive than hydrogen terminated Si used as both dimple and source (1.67 nN/nm² compared to 0.6 nN/nm²), as described in Section 5.5.3 below.

Note that the flexible contact mode AFM probe chip used for collecting the data in Fig. 5.12, which offers sufficient sensitivity to record the adhesion, was too soft for the conductivity measurement due to the skating shown in Fig. 5.13, and that the tapping mode AFM probe chip, which was suitable for electrical measurement through a more robust and stable contact, lacked the sensitivity for the adhesion measurement. It is evident that the selection of the appropriate AFM probe chip is imperative for these tests. As the objective of this work was to investigate a range of materials as noted above, it was decided to perform molecular dynamics simulations on various candidate material combinations to predict likely trends in adhesion, and hence to help identify the best dimple - source materials.

### 5.5.2 Molecular dynamics simulations

A battery of molecular dynamics simulations was carried out by industry collaborators [26] to determine adhesion pressures for the various systems. A Molecular Dynamics (MD) simulation is a computer model which calculates the force on each molecule arising from its interaction with all other molecules in the simulation, and determines the response of each molecule by applying
Newton’s equations of motion. These simulations use the ReaxFF force field, a reactive force field that parameterises pairwise van der Waals interactions \cite{117}, \cite{118}. The atomistic models use force fields for Si, O, C, H, F, W and Au. Our collaborators used the ReaxC implementation of ReaxFF in the Large-scale Atomic / Molecular Massively Parallel Simulator MD code \cite{119}, \cite{120}. The results of our simulations are shown together with experimental data in Fig. 5.15, and will be discussed in Section 5.5.3 below.

5.5.3 Performance mapping of candidate material combinations

The approach to measuring contact mechanics and electrical performance of the (U)NCD - Si system described in Section 5.5.1, was used to investigate the simulated candidate material combinations including Si, UNCD, graphene and highly ordered pyrolytic graphene (HOPG) by one of our project collaborators. The adhesion data from this set of experiments is summarised in Fig. 5.15. The AFM tip (or nanorelay dimple) was Si, including one experiment where the tip was hydrogen passivated by dipping in HF acid for a few seconds. The substrate (or nanorelay source) was Si (both with native oxide, and hydrogen terminated), UNCD (both native and after exposure to fluorine based plasma), graphene and HOPG. For each material set, three adhesion values are quoted. The data for the results of a molecular dynamics simulation is shown by the red bar graph. The data plotted in the blue bar graph for each material combination is that recorded at standard temperature and pressure, and the data in the green bar graph is recorded under vacuum at an elevated temperature of 120°C. The purpose of the vacuum and elevated temperature conditions was to minimise the influence of water at the interface.
### Figure 5.15

Adhesion pressure for Si on listed surfaces from molecular dynamics (MD) simulation and force spectroscopy at STP and under vacuum at 120°C. Note that F-UNCD is UNCD exposed to fluorinated plasma prior to the test.

<table>
<thead>
<tr>
<th>AFM tip (dimple)</th>
<th>Substrate (source)</th>
<th>Hamaker Constant</th>
<th>Condition</th>
<th>Adhesion Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>Si</td>
<td>2.37E-19</td>
<td>MD Estimate</td>
<td>2.8±0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.28E-19</td>
<td>STP</td>
<td>2.69±0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.45E-19</td>
<td>Vac, 120°C</td>
<td>1.71±0.19</td>
</tr>
<tr>
<td>Si</td>
<td>H : Si</td>
<td>5.08E-20</td>
<td>MD Estimate</td>
<td>0.6±0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.10E-20</td>
<td>STP</td>
<td>0.96±0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.69E-20</td>
<td>Vac, 120°C</td>
<td>0.44±0.09</td>
</tr>
<tr>
<td>H : Si</td>
<td>H : Si</td>
<td>4.66E-20</td>
<td>MD Estimate</td>
<td>0.55±0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.04E-20</td>
<td>STP</td>
<td>0.6±0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.29E-20</td>
<td>Vac, 120°C</td>
<td>0.27±0.04</td>
</tr>
<tr>
<td>Si</td>
<td>UNCD</td>
<td>1.41E-19</td>
<td>STP</td>
<td>1.67±0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.72E-19</td>
<td>Vac, 120°C</td>
<td>2.03±0.46</td>
</tr>
<tr>
<td>Si</td>
<td>F-UNCD</td>
<td>1.29E-19</td>
<td>MD Estimate</td>
<td>1.52±0.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.85E-19</td>
<td>STP</td>
<td>2.18±0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Vac, 120°C</td>
<td></td>
</tr>
<tr>
<td>Si</td>
<td>Graphene</td>
<td>7.28E-19</td>
<td>MD Estimate</td>
<td>3.4±0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.87E-19</td>
<td>STP</td>
<td>6.44±0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.46E-19</td>
<td>Vac, 120°C</td>
<td></td>
</tr>
<tr>
<td>Si</td>
<td>HOPG</td>
<td>1.19E-19</td>
<td>MD Estimate</td>
<td>1.4±0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.80E-19</td>
<td>STP</td>
<td>2.13±0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.01E-19</td>
<td>Vac, 120°C</td>
<td>1.19±0.18</td>
</tr>
</tbody>
</table>

While direct comparison between MD simulation and experimentally measured adhesion pressure is difficult (due to surface roughness, residues, contaminants, meniscus effect etc), some trends are clear. The lowest adhesion predicted by simulation, and verified experimentally at 0.6 nN/nm² at STP and 0.27 nN/nm² at 120°C under vacuum, is a HF treated Si tip on a HF treated Si surface. In this case, the HF treatment leaves a hydrophobic H terminated Si surface, as shown by Bollani [121]. Unfortunately, this system
Figure 5.16: Relay design space showing that electrical and mechanical specifications cannot be realised simultaneously. With the least adhesive (H:Si) dimple-source system, the maximum contact area for successful relay pull-out is 0.67nm$^2$, and with the most conductive (Si-UNCD) dimple-source system, a minimum contact area of 10nm$^2$ is required. Using a best case scenario, it is possible to realise the specified current requirements with a contact area no less than 10nm$^2$, but a contact area greater than 0.67nm$^2$ will not pull out, even with the least adhesive dimple-source material system.

showed very unstable and irreproducible electrical performance. The best conductivity performance was achieved with the Si-UNCD system where a current density of 5-10nA/nm$^2$ was achieved. Considering the current requirement of 100nA for our circuit application from Table 5.2, a contact area of at least 10nm$^2$ would be required. However, the adhesion (even with the least adhesive H : Si system) for such a contact area was measured to be 6.7nN, well in excess of the 0.4nN spring restoring force available. Although
CHAPTER 5. NANORELAYS AT THE 22NM NODE

the adhesion and electrical specification can be realised separately, our study did not reveal any material combination capable of simultaneously meeting both specifications. Fig. 5.16 shows the pull-out force and the current conduction, both plotted against the contact area. As noted above, nanorelay operation set the maximum adhesion pressure at 0.4nN. Experimentally, the lowest adhesion achieved was 0.6nN/nm², necessitating a sub 1nm² contact area. The region where this condition is met is shaded in green. It was shown in Section 5.3.2 that a current of 100nA is required for the nanorelay based four input AND gate circuit under consideration. To achieve this current with the highest conductivity dimple - source material combination (Si - UNCD), a contact area of at least 10nm² is required. The region where the nanorelay can meet this specification is shown in blue. There is no overlap between these two regimes: with a contact area greater than 1nm², the device will not pull-out, and with a contact area below 10nm², the device will not carry sufficient current.
5.6 Conclusions

Having investigated the stiffness of nanoscale Si cantilevers in Chapter 4, the mechanical and electrical performance of the dimple - source contact was addressed in this chapter. The use of an AFM to measure both the adhesion and conductivity was explained, specifications for both adhesion and conductivity were determined from an analysis of 22nm node technology in terms of device dimensions, and the voltage and frequency requirements. Based on the dimensions of a transistor, a numerical model was used to assess the impact of varying the beam width, length, and thickness, and the contact gap with a view to maximising the pull-out force, thus setting the upper bound on the adhesion at the dimple - source contact. The model was also used to determine the attainable operational frequency for such a device. The electrical specification was derived by setting the nanorelay electrical (RC) time constant at half the mechanical time constant. As the nanorelay capacitance is swamped by the interconnect wiring capacitance, the resistance of the nanorelay is the key factor determining performance. To determine the resistance, a circuit containing four relays operating at 100MHz at 1V was assumed. It was found that to realise this performance, the nanorelay resistance could be no more than 164.5KΩ, and the dimple - source contact would need to pass 100nA.

A range of materials for use at the dimple - source contact was tested against these specifications. The testing of the (U)NCD - Si combination was described in detail, and this approach was extended to look at various material combinations as shown in Fig. 5.15. The lowest adhesion over all materials tested was predicted by molecular dynamics simulations, and confirmed experimentally, to be hydrogen terminated Si (for both dimple and source), with a value of 0.6 nN/nm$^2$ at STP and 0.27 nN/nm$^2$ at 120°C.
under vacuum (from Fig. 5.15). However, this material system failed electrically with unstable and irreproducible conductivity. The best performance electrically was achieved with the UNCD - Si combination, but this system suffered high adhesion (1.5 - 2.2 nN/nm$^2$). None of the material combinations simulated and tested met both the adhesion and conductivity specification simultaneously. Potential solutions to this challenge will be discussed in Chapter 6.
Chapter 6

Possible Future Directions

6.1 Introduction

Nanorelay technology faces several challenges which prevent widespread adoption in CMOS architecture. This work has addressed the scaling of these devices towards dimensions informed by the CMOS 22nm technology node. For integration into such aggressively scaled CMOS technology, it is clear that the mechanical and electrical performance of the contact must be improved. There are also reliability concerns to be addressed. Even if a contact solution is found to meet operational requirements, it must be robust and able to withstand repetitive impact. Other reliability issues arise from oxidation and electrical discharge and ablation or localised welding. Potential solutions to these issues may involve approaches involving liquid immersion or engineered polymer coatings at the contact interface. Future work might also focus on nonlinear mechanics for potential solutions. While this work has demonstrated that the materials sets investigated cannot support usable relay design space, future work should look to expand the palette of materials under consideration.
6.2 The Adhesion and Electrical Challenge

In order to operate at a CMOS compatible voltage, the beam stiffness must be minimised in a CB device. However, the beam must retain sufficient spring restoring force to overcome the contact adhesion force when the electrostatic force (gate voltage) is removed. The balance between low voltage operation and sufficient restoring force must be achieved while simultaneously meeting electrical current requirements. Given that the speed performance of these devices is dominated by the mechanical switching time, the electrical time constant, which is driven by the contact resistance, can be increased somewhat without compromising device speed. This gives scope to consider changing from low resistance metal contacts to less adhesive solutions with higher resistance materials as described by Kam et al [40]. The use of a TiO$_2$ coating on W [40] is one example that exploits this feature as described in Chapter 3. In a similar vein, Dujardin et al [122] functionalised the bottom electrode surface of a nanotube based device with a self-assembled monolayer to successfully reduce stiction. There are, however, limits to increasing contact resistance due to power dissipation, delays and static noise margin reduction [63], [64].

Other approaches reviewed in Chapter 3 seek to decouple the mechanical and electrical performance of the active element (cantilever beam), allowing for structural / materials optimisation for each aspect of device operation. An example is Parsa’s use of a section of non compliant beam whose area can be designed to adjust the pull-in voltage independently of a compliant beam section which determines the stiffness as described previously in Section 3.2, see Fig 3.7 [55], and Lee’s utilisation of a dual sidewall lateral coating to provide electrical isolation of the actuating beam with conductive Pt electrodes to connect source and drain [58], also described in Section 3.2, see Fig. 3.8.
6.3 Device Fabrication

Scaling from MEMS to NEMS poses several challenges from a manufacturing perspective. With the industry currently running 14nm technology in high volume, patterning lateral dimensions on this scale is feasible using advanced optical lithography, thus avoiding the requirement for EBL which is an inherently sequential process. However, in NEMS devices, the pull-in voltage is determined by the separation between the active element and the gate electrode, as discussed previously. While this will be no less than 2nm to avoid tunnelling current, it is expected to be less than 10nm to achieve a low pull-in voltage. In this work, SiO$_2$ was used as the sacrificial layer. This is compatible with high volume manufacturing (HVM) as it can be deposited with a high level of control over thickness and conformality. However, the wet etch approach used in this work will fail at these reduced dimensions due to surface tension effects during the rinse and dry steps. In this work, beam collapse was seen when using DI rinsing, and IPA had to be used, even with a thick 150nm SiO$_2$ layer. An obvious solution would be critical point drying [123], and this has been demonstrated on 15nm and 20nm air gap devices by Jang et al [124] and Jang et al [125] respectively, but this adds complexity to the manufacturing process. Other options include non wet processing such as HF vapour phase etching, and XeF$_2$ (SAMCO Inc, Kyoto, Japan) for alternative (Si) sacrificial layers. Pollet et al [126] demonstrated that by optimising the process parameters, an improvement in release efficiency (eliminating stiction) of 15% was achievable. In [127], Philippe et al present a materials study for etch stop and sacrificial layers for vapour HF release processing in the context of NEMS / CMOS integration. Their study indicates that Boron Nitride (BN) is an excellent etch stop leaving a very clean surface with no discernable etch rate even after 80 minutes
CHAPTER 6. POSSIBLE FUTURE DIRECTIONS

of HF vapour exposure. Several oxides were studied as candidate materials for sacrificial layers, and the authors found that tetraethyl orthosilicate oxide (TEOS), tetraethyl orthosilicate low rate oxide (TEOS LR) and high density plasma “silane oxide” (HDP SiH₄) were promising for this process step.

6.4 Reliability

Even when a contact solution that meets mechanical and electrical performance specifications is identified, a further requirement relates to durability. In operation, the contact interface will be subject to repetitive impact. In [34], Peschot et al note that to guarantee device operation at 100MHz for ten years assuming 1% duty cycle, the device must withstand 3x10¹⁴ switching cycles. While refractory metals such as W have been shown to withstand wear and tear, W is subject to oxidation which leads to unstable contact resistance. As noted in Section 3.2, a thin TiO₂ coating can be used to prevent this oxidation while remaining within the resistance budget. The reliability of the contact is further tested by electrical discharge. Recall that the device utilises electrostatic force to close, and this force exists between the effective plates of a capacitor made up by the CB, the air gap and the gate electrode. When the switch closes, the charge on this capacitor is discharged and this transient current can exceed the steady state current. At the nanoscale, the contact is dominated by asperities that carry this transient current. If the current density is sufficiently high, localised heating and welding, or even ablation can occur. Some examples are shown in Fig. 6.1. Similar to the oxidation issue above, increasing the resistance of the contact can be used to mitigate this reliability risk. Another approach to resolve this issue was proposed by Lee et al [129], who used a liquid medium as a packaging mate-
Figure 6.1: Failure mechanisms for nanorelay technology: (a) Stiction, an example of a nanotube impacted by stiction is shown in (e). (b) Electrical discharge upon contact as described in the text, an example of which is shown in (f) where the active element suffers ablation with each contact. (c) Electrode wear and damage induced or exacerbated by electrical discharge. (d) Fracture or fatigue of active element, shown for example in (g) and failure mode maps for a gold (h) or diamond like carbon (i) electrode [128].

rial. By using an insulating refined transformer oil, the authors were able to suppress arcing at the contact. The authors also point out other advantages of this approach, such as the reduction of van der Waals stiction force. This latter point will be dealt with in further detail in Section 6.5.

In this work, the lowest adhesion amongst a range of materials was found to be H terminated Si (which is hydrophobic) under vacuum conditions at elevated temperature. All this points to the need to avoid water, which causes adhesion by capillary force. Wei et al [130] show that capillary forces are dominant in AFM based adhesion force measurements in ambient air, and develop a model of a liquid bridge arising from (1) water vapour condensation
6.5 Repulsive Behaviour in Liquids

In this work several candidate material systems were tested for both adhesive and conductive properties. Ideally, the adhesion at the interface would be
completely eliminated. In this case, a very flexible beam would enable very low voltage (and energy efficient) operation, as the beam spring restoring force would not need to overcome any adhesive force. The absence of adhesive force would, in principle, completely eliminate hysteresis in the device operating characteristic of a non pull-in mode device, as described in Section 3.1. This is because when a non pull-in mode device is in the ON position, the electrostatic force balances the beam spring restoring force. Thus when the gate voltage (and associated electrostatic force) is reduced, the spring restoring force immediately overcomes the electrostatic force, and, with no adhesive force acting at the dimple-source contact area, the beam is immediately released to the OFF position.

The adhesion force at the contact area between the dimple and the opposing contact is known as the van der Waals force, described in [133] by Dzyaloshinskii et al, and [134] by Zheng et al. This intermolecular force is always present, even in the case of neutral atoms and molecules, and has three components, the Debye or induction interaction, the Keesom or orientation interaction and the dispersion force. The Debye interaction, described by Leite et al in [135] is a dipole-induced dipole interaction. It describes the interaction when the dipole of a polar molecule induces a dipole in a neighbouring non polar molecule. It is given by

$$\omega(r) = -\frac{u^2\alpha_0}{(4\pi\epsilon_0\epsilon)^2r^6} \quad (6.1)$$

where $u$ is the dipole, $\alpha_0$ is the polarisability, $\epsilon_0\epsilon$ is the dielectric constant and $r$ is the distance. The orientation interaction arises from the interaction between permanent (molecular) dipoles [136]. Although the dipoles can have any angle with respect to the electric field, the distribution is described by a Boltzmann distribution which acts to minimise energy rendering the average interaction between two permanent dipoles non zero. This interaction, also
known as the Keesom interaction is given by

\[ \omega(r) = -\frac{u_1^2 u_2^2}{3(4\pi\varepsilon_0)^2 kT r^6} \quad \text{for} \quad kT > \frac{u_1 u_2}{4\pi\varepsilon_0 e^3} \]  

\text{(6.2)}

The third interaction is the dispersion force, whose quantum mechanical origin was described by London [137]. To gain an insight into this interaction, consider a non-polar atom. Although there is a dipole associated with the positive nucleus and negative electron cloud at any instant, the time average of this dipole is zero. However, the instantaneous dipole can induce a dipole in a neighbouring neutral atom, thus creating an attractive force, and this time averaged force is not zero. It is for this reason that van der Waals forces are always present, even in the case of neutral nonpolar atoms / molecules, as noted above. The dispersion force is given by

\[ \omega(r) = -\frac{3}{2} \frac{\alpha_{01} \alpha_{02}}{(4\pi\varepsilon_0)^2 r^6} \frac{h\nu_1 \nu_2}{(\nu_1 + \nu_2)} \]  

\text{(6.3)}

where \( h \) is Planck’s constant and \( \nu \) is the frequency. Noting that all three feature an inverse sixth power dependence on distance, they can be combined to give the van der Waals interaction, \( \omega_{V_{DW}} \).

\[ \omega_{V_{DW}}(r) = -\left[ (u_1^2 \alpha_{02} + u_2^2 \alpha_{01}) + \frac{u_1^2 u_2^2}{3 kT} + \frac{3 \alpha_{01} \alpha_{02} h\nu_1 \nu_2}{2(\nu_1 + \nu_2)} \right] / (4\pi\varepsilon_0)^2 r^6 \]  

\text{(6.4)}

where the first term is the Debye or induced dipole interaction, the second term is the orientation interaction and the third term is the dispersion interaction. Noting that a simple pair potential between two atoms can be expressed as a power law, \( \left( \frac{C}{r^6} \right) \), these terms can be expressed simply as power laws, and Eqn. 6.4 can be expressed as

\[ \omega_{V_{DW}}(r) = -\frac{C_{V_{DW}}}{r^6} = -\frac{C_{\text{ind}} + C_{\text{orient}} + C_{\text{disp}}}{r^6} \]  

\text{(6.5)}

where \( C_{V_{DW}}, C_{\text{ind}}, C_{\text{orient}} \text{ and } C_{\text{disp}} \) are the van der Waals, induced dipole, orientation and dispersion pair potential coefficients respectively.
6.5. REPULSIVE BEHAVIOUR IN LIQUIDS

The interaction between various geometries and surfaces can be expressed using the conventional Hamaker constant after Hamaker[138],

\[ A = \pi^2 C \rho_1 \rho_2 \] (6.6)

where \( C \) is the interaction coefficient and \( \rho_1 \) and \( \rho_2 \) are the atomic number densities of the interacting bodies. The Hamaker constant can also be expressed in terms of bulk properties such as dielectric constant and refractive index as described by Dzyaloshinskii [133] and Lifshitz [139]. Consider a charge, \( Q \), interacting with a molecule (molecule 2) in a medium (medium 3). The interaction is given by

\[ \omega(r) = -\frac{C}{r^6} = -\frac{Q^2 \alpha_2}{2(4\pi\epsilon_0\epsilon_3)^{\frac{1}{2}}r^4} \] (6.7)

If molecule 2 is replaced by a surface (of medium 2), then the interaction (now capitalised to denote a surface) is given by

\[ W(D) = -\frac{\pi Q^2 \rho_2 \alpha_2}{2(4\pi\epsilon_0\epsilon_3)^{\frac{1}{2}}D} \] (6.8)

However, it is well known, that a charge in a medium with dielectric constant \( \epsilon_3 \) near a surface of a medium with dielectric constant \( \epsilon_2 \) experiences a so called image force from an image charge of strength \( -Q \left( \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \right) \), as described by Israelachvili in [77], where \( \epsilon \) represents the respective dielectric constants. This force is given by

\[ F(D) = \frac{-Q^2}{(4\pi\epsilon_0\epsilon_3)(2D)^2} \left( \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \right) \] (6.9)

with an associated interaction energy given by

\[ W(D) = \frac{-Q^2}{4(4\pi\epsilon_0\epsilon_3)D} \left( \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \right) \] (6.10)

Equating Eqns. 6.8 and 6.11 reveals that

\[ \rho_2 \alpha_2 = 2\epsilon_0\epsilon_3(\epsilon_2 - \epsilon_3)/(\epsilon_2 + \epsilon_3) \] (6.11)
CHAPTER 6. POSSIBLE FUTURE DIRECTIONS

Using this, the Hamaker constant, \( A = \pi^2 C \rho_1 \rho_2 \), can now be expressed in terms of bulk properties (dielectric constant, refractive index) where

\[
A \approx \frac{3}{4} \left( \frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3} \right) \left( \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \right) + \frac{3h}{4\pi} \int_a^b \left( \frac{\epsilon_1(i\nu) - \epsilon_3(i\nu)}{\epsilon_1(i\nu) + \epsilon_3(i\nu)} \right) \left( \frac{\epsilon_2(i\nu) - \epsilon_3(i\nu)}{\epsilon_2(i\nu) + \epsilon_3(i\nu)} \right) \, dx
\]

(6.12)

where the first term describes the Keesom and Debye contributions to the van der Waals force, see Eqn. 6.5. The dispersion contribution is captured by the second term where \( \epsilon_1(i\nu), \epsilon_2(i\nu) \) and \( \epsilon_3(i\nu) \) are the AC dielectric constants at imaginary frequency \( i(\nu) \).

This result is important as it shows that the Hanaker constant becomes negative if \( \left( \frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3} \right) \left( \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3} \right) \) is less than zero. Modelling the contact area of a nanorelay as a sphere on a flat surface (see Parsegian [140] for other geometries), the interaction energy, given by \( W = \left( -\frac{AR}{6D} \right) \), where \( R \) is the radius of the sphere and \( D \) is the separation distance, will be positive and the resulting force \( F = \left( \frac{dW}{dD} \right) = \left( -\frac{AR}{6D^2} \right) \) will be negative (i.e. repulsive) if the dielectric constant of the medium is intermediate between that of the two contact materials \( (\epsilon_1 > \epsilon_2 > \epsilon_3) \).

To test this hypothesis, force spectroscopy was carried out using a SiO\(_2\) tip on an AFM probe chip in contact with a Si substrate, all immersed in diiodomethane (CH\(_2\)I\(_2\)). Note that the dielectric constant of CH\(_2\)I\(_2\) (5.32) is less than that of Si (11.7) and greater than that of SiO\(_2\) (3.9). Fig. 6.3 shows a force displacement characteristic recorded in air. The “jump to contact”, loading, unloading and adhesion regions of this plot were already described in Chapter 4, see Fig. 4.11. It can be seen in Fig. 6.3 that this system is strongly adhesive (the red retract curve reaches a maximum of \( \sim 70\)nN at the “pull-off” region of the curve), when the experiment is carried out in air. When this measurement was repeated, but with the system immersed in CH\(_2\)I\(_2\), the adhesion was not merely reduced to zero, but the system became repulsive.
as can be seen from the blue approach curve in the “jump to contact” region, and the red retract curve in the “pull-off” region. This may well point to future solutions, where devices are immersed in dielectric liquids. A more practical implementation of this idea might involve the use of a conductive polymer film at the interface, rather than liquid immersion.
CHAPTER 6. POSSIBLE FUTURE DIRECTIONS

Figure 6.4: (a) CB device in OFF state. The beam length is designed for low pull-in voltage (b) CB device in ON state. The effective beam length is reduced by the mechanical stop, thereby stiffening the beam and increasing the beam spring restoring force.

6.6 Novel Device Design

Novel device designs such as that by Lee et al [46] and Parsa et al [55] were discussed in Section 3.2. Another novel device design might arise from the use of a nonlinear mechanism to address the stiction problem, as shown schematically in Fig 6.4, which shows a mechanical stop located beneath the beam. While the stiffness of the beam is designed for pull-in at a low voltage, the mechanical stop acts to effectively increase the stiffness of the beam by reducing the length of the free beam, i.e. from the mechanical stop. Once the beam comes in to contact with the mechanical stop, it now has increased spring restoring force to overcome stiction, with no penalty in terms of higher pull-in voltage. Given the tight tolerances involved in these devices, with actuation gaps of a few nm, the manufacturing tolerances for such a mechanical stop would be very challenging, but future work in this area might prove fruitful.
6.7 Alternative Materials

Alternative materials to those used in CMOS technology may well meet the mechanical, electrical and reliability criteria described in this thesis. Of the many material sets under consideration, some are particularly interesting. Metals generally offer low resistance, but tend to be adhesive due to their metallic bonding and high dielectric constant, whereas oxides tend to be less adhesive, but suffer high electrical resistance. Diamond like materials such as the nanocrystalline and ultrananocrystalline diamond (NCD & UNCD) used in this work and other “Diamond Like Carbon” (DLC) materials are interesting as they can offer both low adhesion and low resistance as shown by Sumant et al [141] and Loh et al [142]. In [143], Sumant et al suggest that these materials can be integrated with CMOS technology. While the Si-UNCD material set tested in this work failed to meet the criteria for inclusion at the 22nm technology node, further work on this family of materials may lead to improvements. The growth or deposition of this material would have to be achieved within the thermal budget, however, and this could be very challenging if the nanorelay device was to be integrated in the back end of line (BEOL) stage of processing, as described by Sumant et al [143].

Silicon Carbide (SiC) is another candidate that has received attention [144], [145]. SiC technology is more mature than the aforementioned UNCD, for example. Of particular note is the extended lifetime and reliability over extended switching cycles. For example, extended cycle lifetime tests have been run over several days, and have demonstrated reliable performance with stable switching voltages. Simple logic functions have been demonstrated by He et al [146] using SiC NEMS, showing the potential of scaling up these devices.

Since Novoselov et al [147] isolated graphene by mechanical exfoliation...
from graphite over a decade ago, there has been enormous interest in 2 dimensional (2D) materials. This is due to their remarkable electrical, chemical and mechanical properties as described by Geim et al in [148], and these properties have led Lee et al [149], to propose the use of 2D materials for NEMS devices. 2D materials such as graphene and the transition metal dichalcogenides (TMDs) feature very low surface energy due to the lack of dangling bonds on the out of plane surface. In Section 5.5, it was noted that, in this work, the lowest adhesion pressure was found on hydrophobic hydrogen terminated Si. In [149], Lee et al found that graphene had the lowest capillary force, and attributed this to the hydrophobicity of graphene. Graphene, like many TMDs, combines this attribute with chemical and mechanical robustness, and, crucially, excellent electrical performance, with a carrier mobility of $\sim 15000 \text{cm}^2\text{V}^{-1}\text{s}^{-1}$. However, graphene does not have a bandgap, limiting its potential as a channel in a FET device (although a bandgap can be opened by introducing dopants). As TMDs are scaled down to atomic layer thickness, the quantum confinement creates a bandgap, opening up potential for these materials as transistor channels in transistors, where their high carrier mobility might lead to increased drive current as described in a review by Giannazzo [150], and it is this that has motivated the dramatic increase in research activity in this area. In addition to the their electrical performance, the mechanical flexibility and strength of TMDs is desirable as a contact material for NEMS where repetitive actuation can cause wear and tear as noted in Section 6.4 above. The highest quality TMDs have been achieved by mechanical exfoliation, but this produces small flakes, and is not scalable to semiconductor industry requirements. Chemical Vapour Deposition (CVD) will be required for the large scale synthesis of these materials to make them commercially relevant.
6.7. ALTERNATIVE MATERIALS

As noted earlier, the integration of NEMS devices in BEOL CMOS, must be executed within a thermal budget to avoid impacting the underlying CMOS transistors and interconnect wiring (in the context of integrating MEMS and CMOS, Takeuchi et al [151] investigated the impact of processing temperature for BEOL integration of NEMS, and proposed $425^\circ{\text{C}} - 475^\circ{\text{C}}$ as an acceptable range). This poses a challenge for the synthesis of these materials by CVD, where growth temperatures tend to exceed this range. Alternative low temperature approaches are under investigation. Chanyoung et al [152], for example, synthesised PtSe$_2$ thin films at $400^\circ{\text{C}}$ using thermally assisted conversion, suggesting that these materials could well be compatible with CMOS processing, and BEOL NEMS integration.
Chapter 7

Summary and Conclusions

For approximately five decades, the semiconductor industry has evolved in line with the prediction about scaling known as Moore’s Law [1]. Over this time, the industry has scaled from the micro scale to the nano scale, with Intel’s 14nm process [73] in volume manufacture since 2016. Remarkably, in all this time, the basic building block of the integrated circuit, the MOS transistor, has been refined to deliver improved performance and aggressive scaling resulting in the information revolution. From state of the art fabrication such as nanometer scale lithography [153] to Intel’s device innovation such as Ghani et al’s strained Si [110] and Mistry et al’s high k metal gate stacks [111], MOS transistors now represent the world’s most advanced technology. The MOS transistor is now facing a major challenge, however, and this may see the industry move from its historic fast paced evolution to a revolutionary phase, where new devices provide solutions to some of the fundamental challenges facing nanoscale MOS technology.

During the past decade, computing has become mobile, with extremely powerful processors on board mobile devices providing unprecedented connectivity. Whereas the focus of Moore’s law scaling in the past was on speed
performance, energy efficiency is now the key criterion for battery powered mobile devices. The power performance of MOS devices consists of dynamic power, when the device is switching, or is in the ON state, and static power, arising from the leakage current when the device is in the OFF state. Due to the finite subthreshold slope, the threshold voltage can no longer be scaled (to increase gate overdrive) due to the exponential increase in subthreshold leakage. This prevents scaling of the supply voltage in proportion to device dimensions leading to increased power density. The technology is now reaching a fundamental limit for energy per operation which poses a serious threat to further scaling as described by Frank [154]. There are several candidates which address this problem by avoiding the fundamental subthreshold leakage associated with thermionic emission. The tunnel field effect transistor, for example, is based on tunnelling through a narrow gap which is modulated by gate voltage. While these devices offer improved energy performance over CMOS as outlined by Ionescu [32], they typically suffer from poor drive current. Control over threshold voltage is also challenging in these devices.

In this work, nanoelectromechanical devices are investigated as candidates for improving the energy performance of CMOS devices. Microelectromechanical devices (reviewed by Tanaka in [155]) are in widespread use in applications such as accelerometers for airbag deployment in cars and printing heads in inkjet printers. By scaling this technology to the nanoscale, there may be potential for integrating NEMS devices to complement or replace CMOS devices to improve power efficiency. The basic idea is to replace the transistor with a mechanical switch, which features abrupt switching with infinite “subthreshold” slope. Two nanoelectromechanical relay candidates are presented below. Researchers at Stanford University in California proposed a 500nm long 20nm wide cantilever beam device [156], capable of
operating at +/- 180mV. A more complex crab like relay was demonstrated by researchers at UCB [21]. This device consists of a movable SiGe membrane suspended on four serpentine springs. The membrane acts as the gate and is isolated from the channel on the underside of the membrane by a thin gate oxide layer. This is a flexible architecture which can offer dynamic control over switching voltage with multiple source and drain contacts allowing configuration for various logic gates and more complex functions [40].

As NEMS devices are scaled to the nano regime, multiple challenges arise. For the cantilever beam nanorelay devices investigated in this work, there is a compromise to be made between the need to minimise the operating voltage by reducing the stiffness of the beam, and overcoming stiction which is exacerbated by scaling from the micro to nano length scale. The nature of such devices fabricated using conventional semiconductor processing techniques must be understood, and models that accurately predict mechanical properties, such as those developed in this work, are required.

While stiction is a central challenge for these devices, potential solutions are severely constrained by the requirement that the active contact interface remain conductive and handle current densities which are relatively high due to the asperity model of nanocontacts (whereby the actual contact area is much smaller than the physically patterned or apparent contact area). As the speed of these devices is dominated by the mechanical delay due to the relatively slow travel time of the active element, there is scope to increase the electrical RC time constant with no overall speed penalty. Therefore, more resistive non metallic materials, which may give rise to increased RC delay, can be considered, given that such non metallic materials might offer reduced adhesion due to the absence of metallic bonds.

In this work, both these mechanical and electrical challenges were ad-
dressed. The fabrication of arrays of nanocantilevers was presented, and the detailed characterisation of these devices was measured using AFM based force spectroscopy. The nanocantilevers were investigated from both a compositional and geometrical perspective using a combination of focused ion beam (FIB) microscopy, and transmission electron microscopy (TEM), and these insights informed a realistic value of moment of inertia which was then used in both analytical and finite element models. These models were validated against the results of force spectroscopy measurements on fabricated nanoscale beams.

AFM was also used to investigate the adhesive and electrical performance of various CMOS compatible material systems. Conductive AFM was used to characterise various material systems by landing an AFM probe tip on a substrate, and measuring the current under bias while loading the tip up to a predefined force, and then retracting the tip. In this scenario, the tip represents the device dimple and the substrate, the device source terminal. In addition to the electrical performance, these force displacement measurements allow direct measurement of the stiction as the tip (dimple) is pulled off the substrate (source terminal).

Target specifications for both adhesion and current were derived from 22nm technology developed by Intel [29]. Reasonable dimensions based on design rules and transistor device dimensions were identified, leading to a target adhesion, based on the stiffness of a beam of these dimensions. Electrical performance specifications were derived from a realistic circuit implementation using nanorelays. It was shown that while there are material combinations that meet each specification separately, even with the least adhesive hydrogen terminated Si tip on a hydrogen terminated Si surface (both highly doped), there is no design space where both criteria can be met simultane-
ously.

Future work in this field should focus on identifying materials or device geometries that reconcile these two divergent requirements, while remaining robust and able to withstand repetitive impact over the device lifecycle. Some preliminary work in the area of liquid immersion was presented in this thesis. In fact, adhesion was completely eliminated, and repulsive behaviour was demonstrated merely by immersing an otherwise adhesive contact in diiodomethane. This might point to potential solutions to the stiction issue if, for example, conductive polymers which could be deposited on the substrate contact could be implemented. Such future areas for research are worthwhile in view of the fundamental energy crisis facing CMOS as Moore’s Law scaling is continued in to the future.
Appendix A

Processing Details

A.1 Application of Photoresist

Spin zep-520 electron beam (photo)resist using the SSE SpinnerPrimus 15. Select the “50AL_PMM” programme. Dispense 60µL as soon as the rotation starts (the chuck accelerates during the first five seconds of the spin programme). This is followed by a 2 minute 180°C postbake on the hotplate. The resulting resist film should appear reasonably uniform and defect free.

A.2 Electron Beam Lithography

A.2.1 Electron beam exposure

The coupon should be mounted on the square EBL stage, which contains a Faraday cup for beam measurement purposes, and introduced in to the Zeiss Supra SEM. Typical imaging / exposure conditions would be an accelerating voltage of 20KV and an aperture of 10µm, giving a beam current of about
26pA. This is measured using the aforementioned Faraday cup. After optimising the beam conditions to achieve round < 40nm diameter contamination spots, the exposure should be set up.

The GDSII pattern can be found at “C:\quantum\user\Peter Gleeson\GDSII” on the Raith Elphy Quantum PC. The Working Area and Writefield should both be set to 100 µm, and the X-Y to U-V coordinate translation performed as usual. For the exposure, an “Area step size” of 8nm (U and V) is recommended, and an “Area dose” of 27µC/cm² (equivalent to 27 µAs/cm² as displayed on the Raith PC) has been found to give good results.

A.2.2 Development

Development is carried out in Methyl Isobutyl Ketone (MIBK) : Isopropyl Alcohol (IPA) 1:3 at room temperature for 2 mins, followed by rinsing in IPA, and spin drying. The “DRYSP01” programme on the SSE SpinnerPrimus 15 spinner should be used. Inspect in the optical microscope.

A.3 Plasma Etch

A.3.1 Silicon plasma etch

Mount the coupon on the 150mm Si carrier wafer using a minimal amount of fomblin oil to mitigate turbulence during pumpdown and processing. Select the programme “SF6 / CHF3 Si etch”. Edit the etch process step duration to 20 seconds, and run the process. Remove the coupon from the loadlock upon completion of the process, and clean the backside using IPA. Inspect
in the optical microscope.

### A.3.2 Silicon dioxide plasma etch

Mount the coupon on the 150mm Si carrier wafer using a minimal amount of fomblin oil to mitigate turbulence during pumpdown and processing. Select the programme “CHF3 / Ar SiO2 etch”. Edit the etch process step duration to 1 minute, 24 seconds, and run the process. Remove coupon from the loadlock upon completion of the process, and clean the backside using IPA. Inspect in the optical microscope.

### A.4 Wet Release Etch

Observe all safety protocols including training, buddy system, time-of-use run sheets and full personal protective equipment (PPE). Immerse the coupon in the 1:7 BOE bath for 1.8 minutes. Transfer to the 18MΩcm DI weir, and leave for 60 secs. Transfer to a beaker of DI water, and replace 20ml DI with IPA. Repeat until solution approaches 100% IPA. Remove the coupon and dry under OFN.

### A.5 Removal of Electron Beam Resist

#### A.5.1 Barrel asher process

Place the coupon on the sample electrode, carefully noting the coupon location. Pump the chamber down to base pressure, and then open the O₂
rotameter and adjust the flow to achieve a chamber pressure of 0.3 mbar. Confirm the power setting at 5.0 (arbitrary units), and set timer to 30 seconds. Turn power on to ignite plasma. When strip is complete, close the rotameter, evacuate and then vent the chamber. Remove the coupon, and inspect under optical microscope to ensure successful strip.

A.5.2 ICP strip process

Note that the barrel asher based process was found to be somewhat unreliable, and a more robust strip process was implemented in the ICP etch tool. In this case, use the programme “Pure O2 clean”, and run this for 2 minutes.
Bibliography


