Wireless Nano and Molecular Scale Neural Interfacing

Thesis by

Akram Sarwat Sadek

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2017
(Defended August 16, 2016)

For my beloved parents

Mona & Sarwat,

the greatest mercy bestowed upon me.

Acknowledgements

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

I would like to thank my parents Mona Al-Kalagi and Dr Sarwat A.A. Sadek FRCSI, FRCS, FRCS (ORL-HNS), above and before everyone else for helping me complete this thesis. Their whole life, they have tirelessly taken care of me, protected me, fed me, brought me up, educated me and provided for me, with no regard to the personal sacrifice it entailed. They went far beyond what parents may be expected to do for their children, and they are and will always be the most precious blessing of my life. However, things did not end there. My time at Caltech was not easy, and I simply would not have made it through without their constant and tireless support. When my stipend and funding were interrupted, they told me not to worry about a single thing and to concentrate on my work. They paid my stipend for nine months, and also sent me money for supplies to work on my experiment and complete the fabrication of my setup. The vacuum chamber I built and which I describe in Chapter 4 of this thesis, was primarily funded by them. They also purchased one of the lock-in amplifiers I needed for my experiment. I just would not have been able to make it through without them. I will never be able to repay them for everything they have done for me in this life.

I would also like to thank my teachers. I was fortunate in life to have had inspiring and exceptional teachers in school and during my undergraduate years. أُستَاذ فَتحي عبد الدَايِم (*Oustaz* Fathi Abdel Dayem) and أُستَاذ أَحمَد (*Oustaz* Ahmed) taught me the Arabic language and the Holy Qur'an with humbleness, sweetness and optimism. Mathematics was programmed into my neocortex by

Mr Christakis Glykys and Professor Colin Wilkin (UCL), who made it easy and fun for me. I not only learned mathematics from them, but the art of teaching as well. I'd also like to thank Professor Edward P. Glenn (University of Arizona), who was my very first mentor in science, and who I remember with great fondness.

I'd like to especially thank my former mentors and research advisors at University College London, Professor Michael R. Forshaw and Professor Konstantin Nikolić. Everything I know about research I learnt from them. How to think like a scientist, how to select promising ideas and how to write a paper and publish. I will never be able to exceed them. In particular, Mike and I would discuss ideas and problems in physics, computation, nanotechnology and neuroscience for hours at a time, almost every day. This was a unique intellectual experience I was never to experience again, and one I dearly miss. We did amazing research together, and more importantly, working with them moulded the set of ideas in me that formed the entire focus of my research career, including this thesis. If my scientific DNA originates from somewhere, it is most certainly them. The late Professor Marshall Stoneham FRS (1940-2011) and Professor Donald H. Davis were also hugely influential mentors, and people I really looked up to. I wouldn't be at Caltech without them. I'd also like to thank Dr Gordon Wright M.D., University of Cambridge, who was my first neuroscience teacher. I have such enormous respect for him, and will never forget the evening he mentored me by the fireplace at Clare College.

At Caltech, I have many people to thank. Professor Christof Koch is one of the people I respect most in life, and I shan't ever forget all he did for me. It was a privilege to spend time in his group, and I'm so grateful for him giving me a place to work there until I could find my feet. I'll also never forget the time he used his endowment to cover my stipend when it was interrupted, and how he offered to do so again if the need arose. Not only is he one of the most remarkable scientists I have ever met, he is a real mensch and it was an honour to work with him.

It was also an honour to work with the late Professor Ahmed Zewail (1946-2016), and I am extremely grateful for the opportunity he afforded me in his lab. Dr Zewail was one of the most intelligent, visionary and prolific scientists I knew, and he will always be the pride of Egypt. It was in his lab that I accidentally stumbled upon the strange interaction between light and carbon nanotubes, together with Professor Omar F. Mohammed, who was a postdoc in the group. It was these observations that eventually led to the ideas presented in Chapter 5 of this thesis.

John van Deusen was one of my best teachers at Caltech, and was enormously helpful these past years. He taught me everything I know about machining, without which my project would not have been possible. He was also a friend, and his machine shop was a refuge for me when I needed time out. Jody Collier is an exceptional teacher and taught me how to TIG weld despite proffering that his teaching methods are unsuitable for PhD's. I'm grateful to Hiroshi Ito for teaching me patch clamping, and to Anne Taylor for teaching me neural cell culture. Andrew Steele was a great help in getting me started carrying out mouse recordings. I'm also grateful to Rahul Srinivasan and Chris Richards for teaching me molecular biophysics and optics, they were good friends and it was a blast working alongside them. I was very fortunate to take Professor Viviana Gradinaru's exceptional course in optogenetics last year, which was incredibly helpful in crystallising the ideas in the final chapter of this thesis, that I had been ruminating on for many years. The online blog of Dr Shahriar Shahramian (Bell Labs), *The Signal Path*, was an incredible resource for me over the years, and was a major part of my education in electronics, microwave engineering, optoelectronics and signal processing. I am also indebted to Tom Hayes and Professor Paul Horowitz for their exceptional instruction in electronics, which I finally learnt 'properly' in their course Physics 123 at Harvard University after a lifetime of trying to teach myself.

Laura Kim is one of the most gentle and caring people I have ever met, and I appreciate all she did for me during my time at Caltech. I'd also like to thank Kevin Lee and Kurt Murdoch in Human Resources for all their kind help, reassurance and effort. During a very difficult time, they really put my mind at ease, enabling me to concentrate on completing my thesis.

I'm grateful to the staff at the Kavli Nanoscience Institute and UCSB Nanofabrication Facility for running the labs smoothly and enabling my work.

Sotiris Masmanidis and I were classmates in the physics program at University College London. By the time I caught up with him at Caltech, he had just finished his PhD thesis and I was just starting. We both had the same goal of bringing to bear the technology of nanoscience to the problems of neuroscience. Thus was started a fruitful collaboration between us in Christof Koch's lab. I have fond memories of those early days, and really appreciate all he taught me. We had a lot of imaginative ideas, and scarce means to bring them to fruition. His PhD work on piezoelectric NEMS was key to the work I present here.

A very pleasant outcome of working on this thesis was my interaction and discussions with Professor Tom Dean (Google) and Professor Ed Boyden (MIT). I was extremely honoured to be invited by them into an extended email discussion with David Heckerman (Microsoft), which was hugely productive, producing many interesting ideas. They treated me as an equal, which I was very humbled by, and it was a window for me into what it is like to collaborate with great minds as such. I was also honoured t o be invited by Tom to lead a class for his CS379C course at Stanford University.

My distinct gratitude and special thanks go to my thesis advisor, Professor Axel Scherer for taking me in as a refugee, and granting me the latitude and independence to pursue my risky ideas in his lab. I'm extremely lucky to have joined his group; as a graduate student, it is almost unheard of for one to be allowed to work on their own projects instead of a funded project in the lab. The work presented in Chapter 4 of this thesis simply would not have been possible had I been in any other group. The equipment and materials for this experiment were very expensive, and whenever he could, Axel made sure I had what I needed to do my work. He would always tell me, ”I will never get in the way of progress”. Thank you Axel...

The company of good friends made this journey a lot easier and much more memorable. Costas Anastassiou was there from the beginning of Treadstone, and our regimen of burger/movie/deep conversation is sorely missed. Christine & Julien Dubois and their family were lovely. When things got tough once, they extended an invitation to their Malibu home to chill for the weekend and work on an old engine in the garage, knowing it would be the perfect prescription for me. It's something I'll just never forget. Zahi Karam is one of my dearest friends and is the most generous and magnanimous person I know. Together with Alaa Kharbouch and Demba Ba, we had some spectacular adventures and good times. A group hug also goes out to the rest of the deviants; Sanjay Sen, Abdulbasier Aziz, Ahmed Ettaf, Mohamed Aly, Rashied Amini, Aamir Ali and my loyal friend Michette.

Finally, my dearest family. My brother (and fellow adventurer) Omar Sadek B.Sc. (UCL) is the only person to teach me nanofabrication. He actually taught me all about film photography and dark room processing, but those skills are almost identical to those used in micro and nanofabrication, and were key to me fabricating high quality nanostructures. My sister Sarah Sadek LL.M. (UCL), LPC, BVC is my confidante and best friend, and our deep, meaningful discussions have always guided me. My grandmother Adila Shawket always encouraged me and imparted wisdom gleaned through a lifetime of grit and tenacity. I'd like to thank all of them for their love, support, encouragement and sage advice.

Abstract

Nanoscale circuits and sensors built from silicon nanowires, carbon nanotubes and other devices will require methods for unobtrusive interconnection with the macroscopic world to fully realise their potential; the size of conventional wires precludes their integration into dense, miniature systems. The same *wiring problem* presents an obstacle in our attempts to understand the brain by means of massively deployed nanodevices, for multiplexed recording and stimulation *in vivo*. We report on a nanoelectromechanical system that ameliorates wiring constraints, enabling highly integrated sensors to be read in parallel through a single output. Its basis is an effect in piezoelectric nanomechanical resonators that allows sensitive, linear and real-time transduction of electrical potentials. We interface multiple signals through a mechanical Fourier transform using tuneable resonators of different frequency and extract the signals from the system optically. With this method we demonstrate the direct transduction of neuronal action potentials from an extracellular microelectrode. We further extend this approach to incorporate nanophotonics for an all-optical system, coupled via a single optical fibre. Here, the mechanical resonators are both driven and probed optically, but modulated locally by the voltage sensors via the piezoelectric effect. Such *piezophotonic nanoelectromechanical systems* may be integrated with nanophotonic resonators, allowing concordant multiplexing in both the radiofrequency and optical bandwidths. In principle, this would allow billions of sensor channels to be multiplexed on an optical fibre. With view to eventually integrating such technology into a neural probe, we develop fabrication methods for crafting wired silicon neural probes via photolithography and electron beam lithography. Finally, to complement recording, we propose novel ideas for wireless, multiplexed neural stimulation through the use of radiofrequency-sensitive molecular scale resonators.

Published Content

Sadek, A. S., Karabalin, R. B., Du, J., Roukes, M. L., Koch, C. and Masmanidis, S. C. “Wiring nanoscale biosensors with piezoelectric nanomechanical resonators”, *Nano Lett.*, **10**(5):1769-1773, 2010.
DOI: 10.1021/nl100245z
http://pubs.acs.org/doi/abs/10.1021/nl100245z

Sadek, A. S.,“Using piezoelectronics to wire thousands of neural nanosensors into a single optical output”, *Nano Werk*, 30 April 2010.
http://www.nanowerk.com/spotlight/spotid=16063.php

Contents

List of Illustrations and Tables

Nomenclature

MEMS	Microelectromechanical systems
NEMS	Nanoelectromechanical systems
NOEMS	Nano-optoelectromechanical systems
VLSI	Very large-scale integration
ULSI	Ultra large-scale integration
RF	Radiofrequency
SNR	Signal-to-noise ratio
RMS	Root mean square
SSB	Single sideband
TE	Transverse electric
TM	Transverse magnetic
FDM	Frequency division multiplexing
WDM	Wavelength division multiplexing
FPGA	Field-programmable gate array
CNT	Carbon nanotube
NIR	Near Infrared
pin	Positively-doped / intrinsic / negatively-doped
VCSEL	Vertical cavity surface emission laser
MMI	Multimode interference
SOI	Silicon-on-insulator
BOX	Buried oxide
EBPG	Electron beam pattern generator
SEM	Scanning electron micrograph
MOCVD	Metal-organic chemical vapour deposition
PECVD	Plasma enhanced chemical vapour deposition

ALD	Atomic layer deposition
DRIE	Deep reactive ion etching
RIE	Reactive ion etching
ICP	Inductively coupled plasma
LOR	Lift-off resist
PMMA	Poly (methyl methacrylate)
MIBK	Methyl isobutyl ketone
BOE	Buffered oxide etch
HF	Hydrofluoric acid
DI	Deionised water
IPA	Isopropyl alcohol
PCB	Printed circuit board
ENEPIG	Electroless nickel, electroless palladium, immersion gold
ESD	Electrostatic discharge
VEGF	Vascular endothelial growth factor
FGF	Fibroblast growth factor
PDGF	Platelet derived growth factor
TIG	Tungsten inert gas welding
Q	Quality factor
ρ	Density
Y	Young's modulus
d_{31}	Transverse piezoelectric coefficient
E_g	Bandgap energy

Chapter 1

Nano-Neuroscience and the Wiring Problem

> The brain, the masterpiece of creation, is almost unknown to us.
>
> *Nicolaus Steno, 1669*

1.1 The Brain: A 3D Nanoscale Information Processor

Throughout history, the human brain has been compared to the most advanced technology of the era. In ancient Greece, China, the Islamic Empire, and Europe, scientists and philosophers speculated on the nature of the human being, and whether our bodies and minds were in fact machines. The physical body and human mind were compared to hydraulic and pneumatic machines, mechanical clocks and various other automata. In *The Passions of the Soul*, published in 1649, René Descartes put forward the theory of dualism, arguing that whilst the body is a machine, the mind is non-material. Thomas Hobbes suggested however that the mind itself is a mechanism, and that thought arose from minuscule mechanical motions in the brain. In 1748, Julien de La Mettrie put forward that the entire human body, including the brain and human mind, was an intricate machine.

After the electrical nature of the nervous system was elucidated by Luigi Galvani and microscopic examination of the brain started to reveal axons and dendrites, comparison began to be made between the brain and communication networks. Hermann von Helmholtz compared the brain to a telegraph network. Cable theory used to mathematically describe telegraph communication was reapplied to describe passive signal propagation in neurons. A major impetus for the work done by scientists and engineers at Bell Labs in the early 20th century was the semblance between the brain and the world-wide telecommunications network they were creating [1].

In the 1940's, the work of John von Neumann, Alan Turing, Warren McCullough and Walter Pitts in the development of the modern theory of computing, drew heavily upon the brain for inspiration. von Neumann wondered how the brain reliably processed information, given that it's neurons were unreliable components in a warm, wet milieux [2, 3]. His proposed solution to the reliability problem in digital computation, which he called *multiplexing*, was compared to arrangements of neurons in the brain, and von Neumann stated that the brain itself is ‘prima facie digital’ [4]. Development of the theory of artificial intelligence in the 1960's imbued everyone in the field of computation with the belief that the creation of an artificial mind, equivalent or superior to the human mind was inevitable, perhaps just twenty years away. Once the hardware and software inexorably caught up, it would become a reality. This belief was predicated on the assumption that the human brain *was* a computer and performed computation.

On entering the age of micro and nanoscale circuit integration, the widespread expectation was that computer processors would begin to emulate the brain once the transistor count exceeded the number of components in brains, ostensibly, the number of neurons. We have now reached the era where Moore's law has finally broken down, and the metaphor of the brain as a computer looks questionable. The mouse brain contains 75 million neurons, whilst the cat brain has ten times this number. The human brain consists of 86 billion neurons and dissipates ~ 15 W of power.

Present day ultra-large scale integration (ULSI) processors certainly approach this device count and power dissipation. The Intel Broadwell Xeon E5-2644v4 has a transistor count of 7.2 billion spread over 22 cores, and has a power dissipation of 145 W. The Altera Stratix 10 field-programmable gate array (FPGA) contains 30 billion transistors, and supports up to 5.5 million logic elements. Clearly, neither of these chips exhibit brain-like epiphenomena. However, the von Neumann architecture of these processors may not be a fair comparison to neural systems. Neuromorphic processors are designed specifically to try and emulate the functioning of neuronal morphology and physiology (as best understood) with CMOS components, through a synthesis of analogue and digital design. Some iterations also implement this through a non-von Neumann architecture. The most advanced of these chips is TrueNorth, developed by IBM, through the SyNAPSE DARPA project [5]. This processor consists of 5.6 billion transistors comprising 4096 neurosynaptic cores, that are able to emulate 1 million spiking neurons and 268 million synapses. The power consumption is just 70 mW due to its clockless operation. Since the chip is based upon tiled cores that communicate with one another in nearest neighbour fashion in 2-dimensions, it is eminently scalable. Although such chips will be extremely useful in implementing and running neural networks inspired by actual neural architectures in real time, it is still unknown if they will ever emulate a real brain.

The question arises, is the metaphor of the brain as a computer just that: a metaphor? Are we merely caught up in the latest trap of comparing the brain to the most advanced technology of our time? In several respects, this question is central to the work we present in this thesis. The tools we endeavour to build to study the brain will be influenced by our preconceptions about its underlying nature.

The brain certainly processes information. Information processing, however, is distinct from computation. Computation is the application of a set of rules to a collection of symbols; in other words, it is an algorithmic procedure. It is purely mathematical, and the universal Turing machine is *the* model of computation [6]. Information on the other hand is physical, and is identified as the negative of the thermodynamic entropy. Information *processing* is also physical, and is the manipulation of information in space-time through physical laws [7]. Computation is mathematical and exists in Plato's plane of abstraction. Information processing on the other hand is necessarily physical, and is intimately connected to the universe itself. A black hole is an information processor. A Xeon chip running Linux is also an information processor, but is merely set up to perform computation. With information processing, the hardware itself is key, as it is the physical system that is important. Furthermore, the processing of information is intimately connected to noise, a physical phenomenon [7]. With computation though, the algorithm is hardware agnostic, and can run on any suitably prepared physical system that is Turing universal. Only the space-time efficiency of the computation will depend on the particular system chosen.

It is currently unclear if the brain performs what we would call computation. If it did, then the goal of studying the brain would be to elucidate the underlying algorithm, the internal states of the computer, the symbolic representation etc. We would study the brain at the maximum coarseness necessary, perhaps at the level of the spiking output of every single neuron in the brain. In principle, it would then be possible to fully understand it and even emulate it on other hardware. In this thesis we present new technology for recording the global voltage map of the brain. With enough engineering, it should make it possible to one day record the spiking output from every single neuron in a behaving animal, in real time. If the brain is not computational in nature however, even this level of detail will be insufficient to fundamentally understand it.

The human brain's 86 billion neurons are interconnected with 5 million kilometres of wiring and close to 10^{15} synaptic connections. The wiring is only 100 nanometres thick, and the components are packed closely together, tens of nanometres apart. Furthermore, each of the billions of neurons are nothing like transistors, logic gates or even McCullough-Pitts neurons. Instead, they are highly complex information processors themselves [8], having 1000 to 10,000 input and output connections each (c.f. a modern day computer processor, which has ~ 1500 input/output pins). Information processing in single neurons and synapses scales down to the level of nanoscale structures, chemical networks, single proteins and the diffusion of single ions. Communication occurs not just directly via the axonal cabling, but also through functionally significant crosstalk that arises as leaked electrical fields from neighbouring somata and axons interact on one another [9]. The restricted extracellular space means that local depletion of ions may also play some role in computation, akin to cellular automata. The diffusion of low concentrations of ions within small volumes in synapses may also be the basis of low-power Brownian reversible computation at this scale. The brain consists of as many glial cells as neurons, these primarily functioning to support the neurons. Recent work has suggested some types of glia may play a role in information processing themselves, being able to transmit and receive signals between different neurons. This may further increase the complexity of the brain, as we currently understand it, by an order of magnitude.

Clearly, the stupendous multi-hierarchical architecture of the brain makes probing it and understanding it a daunting task. The brain is a physical device obeying physical laws, and we wish to understand the physical principles and functional underpinnings of how the brain generates perceptions, thoughts and actions. Doing so will require us to probe, measure and intervene in the brains of awake behaving animals at every length scale of neural architecture, over large volumes of the brain, concurrently and in real-time. We will need to understand the physics harnessed by neurons for information processing, from the molecular level, to the dendritic tree, individual synapses and spiking axons, how these work together to form microcircuits in cortical columns, and how those columns interact globally through the data buses of the white matter axonal interconnects and the relay stations of brain nuclei. Whilst this is a lofty and perhaps far off goal, the work we present in this thesis is aimed at helping to achieve it by addressing one of the most pressing obstacles to doing so: the problem of wiring.

1.2 Current Techniques for Probing the Brain

Electrophysiology and optical probes are the two primary technologies currently used to study the brain in real-time. Before we outline our proposed approach for large-scale neural interfacing, let us describe these contemporary techniques for studying neural systems.

There are three main signal modalities used during neuron-to-neuron communication, which are probed with current experimental techniques in neuroscience. Neurotransmitter, electrical potential and intracellular calcium are the primary targets of measurement or intervention. A presynaptic bouton will release neurotransmitter (a chemical or peptide) into the synaptic cleft, where it will diffuse to the postsynaptic bouton. There, it will bind to ligand gated ion channels or metabotropic receptors, which will ultimately have a quantal excitatory or inhibitory electrical effect. If the neurotransmitter is excitatory, Na^+ channels will become permeable, depolarising the neuronal membrane. If it is inhibitory, either Cl^- or K^+ channels will be activated, clamping or hyperpolarising the membrane potential. Enough depolarisation at the axon hillock of a neuron can initiate a non-linear, self-regenerating wave of depolarization which propagates down the axon. The electric field disturbance that this ‘spike’ generates can be detected strongly across the cell membrane itself, or more weakly around the neuron from the divergent field lines. Depolarisation can also be *subthreshold*, and be insufficient to initiate a spike. Such electrical ripples can propagate over the scale of dendrites and somata, but decay too quickly to go further. When a spike reaches the end of an axon, it impinges finally on a presynaptic bouton, where it leads to the opening of voltage-gated calcium channels. The concentration of Ca^{2+} is a thousand times smaller inside the neuron than out, and so the rush of Ca^{2+} into the cell is big news. The presence of this ion initiates neurotransmitter release and also activates/modulates chemical networks within the cell. All spiking activity in neurons can be said to ultimately end with a Ca^{2+} influx signal.

1.2.1 Electrophysiology

Electrophysiology has been the most important technique in neuroscience since the pioneering experiments of Galvani over 200 years ago. There is good reason for this: it has a very high signal-to-noise ratio, and excellent temporal resolution. Electrical recording can be intracellular, where measurements are taken inside the cell with respect to the extracellular potential. Typical measurements have an SNR of 10,000:1. The membrane fluctuations recorded can range from $\sim 100\mu\text{V}$ for quantal synaptic events to $\sim 100\text{mV}$ for a spike. Recordings can also be extracellular, where the external electrical potential relatively close to a neuron is measured with respect to a ground electrode placed much further away. These potentials are typically almost a thousand times smaller than intracellular ones, and the SNR of a good recording setup would be 100:1. Recording in either case can be done with wires or with micropipettes. Electrically, micropipettes are superior as the tip can be fashioned to be hundreds of nanometers in diameter, yet the signal-to-noise ratio can be high due to the low impedance 'wiring' to make contact with it. This is a high surface area silver/silver chloride wire placed within the pipette interior in contact with electrolyte. If the electrode tip taper is too long though, it can be a bottleneck and increase the series resistance substantially. Intracellular recording can be done with sharp microelectrodes, with submicron tip diameters [10]. The small tip diameter and long taper increase the electrical impedance and thus the noise level, but the small tip opening limits exchange of the pipette electrolyte with the intracellular fluid, which can otherwise disturb the cell. Usually there will be leakage around the electrode where it penetrates the cell membrane, further impacting SNR. Another type of intracellular measurement is the whole cell patch clamp [11]. Here, a relatively large pipette tip ($\sim 2\ \mu$m diameter) of short taper, which is polished with heat, is brought in contact with the neuron membrane. After several minutes, through an ill understood process, an extremely tight seal spontaneously forms between the lipid bilayer and the rim of the pipette tip. This is called a gigaohm seal, and has an impedance in excess of $1\text{G}\Omega$. Through electroporation, direct contact is then made with the cell interior. The combination of high seal resistance and low tip resistance allows for extremely high SNR measurements to be made.

A big advantage of intracellular measurements is recording can be made in two modes: voltage clamp and current clamp. In current clamp, a feedback amplifier maintains a preset current flow across the membrane as the neuron goes about its activity. The resultant voltage swings by the amplifier are recorded as the membrane potential activity. In voltage clamp, the feedback amplifier maintains the membrane potential, positive or negative current being injected by the amplifier into the pipette to do this. In this way, the membrane currents can be directly measured, whilst having manual control over the intracellular potential. In particular, voltage clamp cannot be achieved by optical or other techniques at present. Current can also be injected by micropipettes to stimulate neurons.

All intracellular measurements with micropipettes suffer the problem of poor long-term stability. The pipette solution eventually kills the neuron, and the pipette-membrane seal breaks down causing leakage of the cell contents. Recordings typically cannot last for more than a few hours. Whole cell patch clamp has been achieved *in vivo* in awake, behaving animals [12], and multielectrode patch recordings have been done on single neurons to elucidate the biophysics of the dendritic arbour [13]. The most advanced patch clamp setup to date can simultaneously patch 12 micropipettes in brain slices [9], although efforts are underway to extend this to *in vivo* preparations. Ideally, measurements of neuronal membrane potentials in awake, behaving animals would be done using intracellular measurements over millions of neurons in a brain region simultaneously, and over hundreds of different points on the individual neurons themselves to map electrical activity in the dendritic tree and spine compartments during network activity. Clearly, this is not possible with bulky intracellular microelectrodes. Presently, extracellular electrodes provide the next best solution electrophysiologically. Simply, an insulated microwire with an exposed tip is inserted into the neural tissue [14]. This can detect the stray electrical potentials given off through the membrane activity of surrounding neurons. A tetrode is an extension of this technique, whereby four staggered microwires are bound together and inserted [15]. The slightly varied placement of the electrodes allows triangulation to be used to map out particular neurons, through the varied amplitude of identical spikes detected by the different electrodes. Conceptually, extracellular measurement is analogous to using a radiofrequency probe to try and elucidate the function of several microprocessors working in concert on a motherboard through their stray radiofrequency emissions. As hopeless as it sounds, it has turned out to be an immensely important tool in neuroscience. Not only can extracellular recording detect spiking activity, but local field potentials can be measured as well. These potentials arise from the sum of dendritic inputs and somatic potentials, and thus something can be determined about the neuronal input in addition to the output train of spikes.

Theoretically, a tetrode should be able to detect on the order of 1000 neurons in its vicinity. In practice, far fewer neurons are detected. There are three reasons for this. First, the brain seems to implement sparse coding, where only a small number of neurons in an ensemble fire very strongly to represent information. This means only a fraction of neurons in a neural circuit fire at any given time. Second, the tetrode microwire is bulky and brain tissue is damaged during its insertion. Neurons in its immediate vicinity likely die after implantation. Damaged blood vessels lead to bleeding at the insertion site, initiating a tissue reaction cascade, which further impedes good recordings. Thirdly, in the immediate vicinity of a neuron, the extracellular spike potential falls off as $\sim \exp(-r/r_0)$, where $r \sim 20$ μm. This means that beyond $\sim$100 μm distance, the potentials due to surrounding neurons are too small to be extracted with current spike sorting techniques with the SNR of extracellular recording.

Neural circuits seem to be organised as a small world network [16]. Cortical columns are very densely connected locally, but sparsely connected globally through white matter interconnects. Ideal extracellular recordings thus have two requirements. They must consist of a large number of recording sites that are densely integrated, in order to capture local circuitry. They must also be widely distributed over the brain, to understand how cortical columns interact globally in the small world network. Sparse coding likewise drives these two objectives. Finally, the electrodes must have minimal cross section, in order to minimise damage to brain tissue and capillary vasculature. Micro electromechanical systems (MEMS) has emerged as a technology to take on this challenge, which we describe in Chapter 3. Typically, probe shafts several millimetres long but tens of microns in thickness and width are crafted from bulk silicon or using silicon-on-insulator (SOI) technology. Multiple planar electrodes are fabricated on the shafts, and routed out using lithographically defined wiring. The technology has matured greatly since its introduction in 1970. State-of-the-art MEMS probes boast multiple shafts with hundreds of electrode recording sites, double sided electrodes on probe shafts and nanoscale lithographic wiring [17, 18, 19]. Recently, integrated circuits that perform low-power amplification and multiplexing have been integrated on the probes themselves, improving SNR and reducing the number of wires leading out of the probes [19]. Furthermore, some MEMS probes have also had radiofrequency wireless circuits integrated in addition, either in separate modules attached to the skull to which the probes are wired, or directly added to the probe itself to untether it from the skull completely [20]. Apart from silicon, several other materials are being explored for MEMS neural probes in order to improve biocompatibility and to match more closely the mechanical properties of the brain tissue they are embedded in. These include polymer neural probes fabricated from polyamide or parylene [21], and even all diamond neural probes where the electrode contacts and wiring are fabricated through doping the diamond with boron [22]. The development of MEMS neural probes are driven not just by applications in neuroscience, but also with a view to brain-machine interfaces, where they are used as the implantable sensor for recording from ensembles of neurons representing movement [23]. They may also be used for stimulation to provide tactile feedback to the somatosensory system from external actuators such as prosthetic limbs [24].

1.2.2 Cyclic Voltammetry

Microelectrodes are not just used to detect electrical potentials in neural systems. They can also be used with fast-scan cyclic voltammetry to detect oxidisable neurotransmitters, which include dopamine, noradrenaline, serotonin, adenosine, histamine and nitric oxide [25]. GABA, glycine and glutamate cannot be detected with this technique as they are not electroactive species. Carbon electrodes as small as a few microns have been used to localize neurotransmitter release around synapses, with a temporal resolution of tens of hertz. Efforts are being made to use even smaller electrodes, crafted from carbon nanofibres or carbon nanotubes to improve the spatial resolution of detection [26, 27]. Implantable, but bulky wireless modules have been developed, to wirelessly record neurotransmitter signals from *in vivo* preparations, as well as more compact wireless integrated circuits [28].

1.2.3 Optical Techniques

In the past decade, optical techniques have caught up with electrophysiology to become one of the most important tools in the neuroscientist's arsenal. Optical probes exist for the stimulation and recording of all three neural signaling modalities; neurotransmitter, intracellular calcium and membrane potential. The great advantage of optical techniques over electrophysiology are their spatial resolution, non-invasiveness and the ability to genetically encode optical probes so that they are expressed in specific neuronal subtypes or subcellular compartments. Furthermore, large areas of the brain can potentially be recorded from simultaneously. Two-photon laser scanning microscopy allows diffraction-limited volumes of tissue to be probed with all these techniques.

Caged compounds can be used to deliver neurotransmitter and intracellular calcium in a form that is inert, until a photolabile bond is broken by ultraviolet light. Combined with 2-photon microscopy, various neurotransmitters can be targeted to specific points on the dendritic arbour mimicking synaptic activity, and likewise, intracellular calcium can be focally increased in specific subcellular neuronal compartments. Photoactivated tethered ligands (PTL's) instead bind neurotransmitter to postsynaptic receptors via azobenzene, which is photoisomerisable. In the trans configuration, the neurotransmitter points away from the active site. On photoactivation, azobenzene goes to the cis configuration, and the neurotransmitter is directed to the active site, triggering the receptor.

1.3 The Wiring Problem

One of the most neglected aspects in the nanoelectronics field is the problem of wiring. How do we wire individual nanoelectronic devices within a nanointegrated circuit together? Furthermore, how do we extract and input information from such a circuit; i.e. how do we let it communicate with the outside world?

The first problem has been looked at to a limited degree, but the second has yet to be explored. A nanoscale circuit or sensor array will lose its advantages of nanoscale integration if the wiring makes the final system just as bulky or even larger than conventional systems. For something like a nanoscale neural sensor array, that could detect electrical potentials or electroactive species on tens of thousands of different channels, ideally we would like something integrated into a minuscule package that could be introduced noninvasively into the brain. If we need physical wires to communicate with this sensor array, this severely impedes its potential advantages of size and mobility. Ultimately, we would like a system for wirelessly interfacing with the sensor array, to extract information and provide power.

Microprocessors are particularly affected by the wiring problem due to their two-dimensional integration. *Rent's rule* is an empirical relation that has been found to apply to integrated circuits, that relates the number T of input/output terminals on the circuit to the number t of terminals on the constituent devices on the chip. In the most optimal case, $T = tN^{2/3}$, where N is the number of components. The problem with a 2D circuit architecture is that as the scale of integration increases, the perimeter of the chip where the input/output terminals are located scales as the square root of area, whilst Rent's rule indicates that the number of terminals increases as the two-third power, a larger scale factor. Evidently, as the scale of integration is increased, a regime is eventually attained where the terminals interfacing with the circuit are log jammed, precluding further scaling: wiring gets in the way.

Migrating to 3D architectures however, the surface area available for terminal placement will scale as the two-third power of volume, which precisely matches the scaling factor in Rent's rule. Incidentally, the brain takes full advantage of this effect. Comparing the cortex to a microprocessor, grey matter is the 'circuitry', and its absolute volume corresponds to N. White matter on the other hand are the long-range interconnects of the brain, and correspond to T, the number of input/output terminals. The amount of white matter in cortex has been found to scale approximately as the four-third power of grey matter volume [16], an observation perhaps rooted in Rent's rule. The brain is also far more efficient at wiring than VLSI and ULSI systems; the total proportion of wiring in the brain ranges from 40 – 60% across different species, which is significantly lower than the 90% typical in their silicon counterparts.

The brain also has a different kind of wiring problem to contend with, namely how to route power to it's constituent components without significantly compromising brain volume. Routing power is different to routing signals in that it can interface in parallel, and does not need a one-

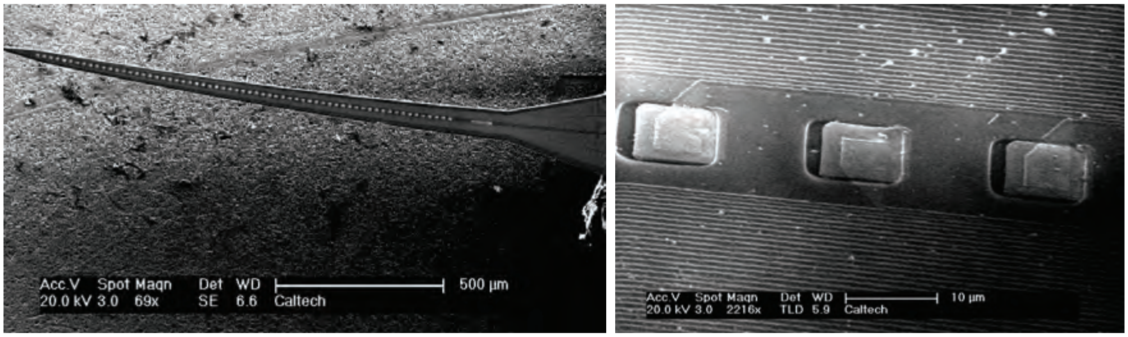


Figure 1.1: **The wiring problem in neuroscience.** Scanning electron micrograph of a silicon neural probe containing 64 recording sites. A magnified view of the probe shows that a significant portion of the shaft width is occupied by wiring. The wiring is the bottleneck to scaling the neural probe without increasing volume, as the electrode is spatially localised, whilst the wiring extends over the full length of the interface and the pathway into the brain. The wires shown here are 200 nm wide and have spacing of 200 nm.

to-one connection. The brain is the most vascularised tissue in the body. A capillary is never more than 10 μm away from a neuron, and thus neurons and their extensive dendritic arbours are densely intertwined with the capillary bed [29]. Despite this, the entire vasculature of the brain only takes up 3% of it's volume [30]. This is achieved in a couple of ways. First, the vasculature is a parallel branching network, with large bus vessels branching down to smaller ones culminating in the capillaries. Those capillaries are in turn highly miniaturised, being only 3 μm wide. This is less than half the width of a red blood cell, which must squeeze into a needle shape to pass through the capillary. Thus an interface only taking up 3% of the brain's volume manages to be intimately connected to every neuron in the brain. If an extracellular recording interface could only achieve a tenth of that resolution, it would be sufficient to answer the challenge of the brain activity mapping project [31]. Such a vision simply cannot be achieved with a one-to-one wiring scheme however, as shown in Fig. 1.1, and must somehow harness the idea suggested by the brain vasculature of a parallel bus interface.

After over a century of technological progress in the fields of telecommunications, the same technology that has allowed billions of people on the planet to seamlessly communicate may enable us to probe how the billions of neurons in the brain do so to process information. Can we harness the principles underlying fibreoptic and cellular communications to find a parallel architectural solution to routing signals from the brain?

1.4 Harnessing Resonators to Probe the Mind

Quantum mechanics and special relativity posit a deep relationship between time and space $(t, \mathbf{x})$ and their corresponding four-vector Fourier transforms in the frequency domain $(\omega, \mathbf{k})$ and energy-momentum space $(E, \mathbf{p})$. Further development of these ideas leads to quantum field theory, which has at its heart the harmonic oscillator. Every particle and wave in the universe, and even space and time itself, are oscillatory excitations of a quantum field defined over all space and time. Oscillation and resonance are thus the fundamental basis of all creation. As the universe is an information processor, it is no surprise then that oscillators are also central to communication and signal processing.

The origin of this thesis, and the solution we present to the wiring problem, is a conversation I overheard between my former advisor at UCL, Professor Mike Forshaw and his colleague Dr David Crawley, whilst we waited in an airport lounge in Würzburg, Germany in 2001. We had just presented our work on the CORTEX European project [7], and they were discussing what could possibly come next. They mentioned what I thought was an intriguing problem, how to ‘scale-up’ or interface whatever circuits and devices we construct at the nanoscale to the macroscopic world, and how to do so without eliminating all the advantages of a nanoscale construct. At the same time, I was reading through Richard Feynman’s *Lectures on Computation* [32], as an introduction to information theory and the physics of computation. The following passage immediately struck me as perhaps being the very solution to the problem my advisors had been thinking about:

> I would like to emphasize that wires represent a real problem in system design. ...the space needed for wiring... is greater than that needed for the functioning components. People occasionally break down and begin to dream, having brilliant ideas such as that of building a machine, by analogy with our broadcasting system, in which each component radiates light of a particular colour, which is broadcast throughout the machine to be picked up and acted on by frequency-sensitive components.

Unfortunately, at the time I had no idea what components could possibly do this. That realisation occurred during my first week at Caltech, when there happened to be a conference on nanoelectromechanical systems (NEMS), which I attended. There, I heard about the work of my former colleague at UCL, Sotiris Masmanidis, on piezoelectric NEMS resonators [33]. This work was significant as it introduced a method for actuating NEMS intrinsically using the piezoelectric effect. Other NEMS devices employed extrinsic methods of actuation, making the entire system bulky.

Importantly, he showed that the devices could be tuned in frequency using a DC bias, making them sensitive detectors of electrical potentials. The nanoscale dimensions of the structure concentrated the electric field across a small width, which imparted excellent voltage sensitivity. Significantly, this tuning effect was linear, unlike other NEMS devices which are nonlinear in response.

With a view to developing these ideas, I began a collaboration with Sotiris in Christof Koch's lab to investigate using this device as a method for multiplexing the electrical signals from potentially tens of thousands of nanoscale neural sensors onto a single optical output [34]. We showed these nanomechanical resonators can be used to transduce analogue AC signals, as long as the frequency is much lower than the resonant actuating frequency. The bandwidth for signal transduction is at least 30 kHz, which is well suited to transducing signals from most analogue sensors. The key property allowing successful signal transduction is the linearity of the response. This is not possible with other types of NEMS devices because of their nonlinear tuning characteristics. We demonstrated the transduction of analogue signals down to 70 μV in amplitude, as described in Chapter 2.

The method presented here for transducing analogue signals is unique, in that unlike the usual method of operating NEMS, where a phase locked loop is used to track the resonant frequency of the structure, the device is resonated at a fixed frequency and the variation in amplitude of the resonance is monitored. This allows real-time transduction of electrical potentials, unlike the changes over minutes to hours that are transduced using conventional methods.

Our technique is very analogous to the operation of the cochlea in the human ear; as a matter of fact, we take advantage of the linearity, sensitivity and multiplexing capability of our devices to build a nanoscale analogue of the cochlea. The details of how the cochlea works aren't precisely known, but current theory posits it acts as a series of independent resonators. The basilar membrane is a triangular piece of tissue that is coiled within the cochlea, and which resonates at different frequencies along its length. The membrane is actively driven to resonance through a bank of outer hair cells, which operate via a piezoelectric effect. The inner hair cells detect the amplitude of the resonance at different points along the membrane, and thus transduce the power at a particular frequency to specific nerve fibres.

The inner hair cells are nanoscale mechanical structures that are highly sensitive to even an atomic width of movement. As sound waves, which are a mix of different frequencies at different powers, impinge upon the cochlea and travel along the basilar membrane, the sound waves are decomposed into component frequencies, the waves at specific frequencies varying the power of the intrinsically resonating structure at those frequencies. This variation in power is transduced by the inner hair cells to nerve fibres. Thus, the cochlea acts as an inverse *Fourier transform*; it takes a single input channel that is multiplexed at different frequencies and decomposes it to multiple, individual output channels.

In our scheme, we implement the inverse of this. We use a bank of piezoelectric nanomechanical beam structures of different length, each resonating via piezoelectric actuation at well separated frequencies. Each NEMS device performs the job of the inner and outer hair cells concurrently. It uses the piezoelectric effect to actuate the devices to resonance, and also harnesses the piezoelectric effect to detect electrical signals applied to the device in a very sensitive fashion.

Each NEMS device is attached to the electrical output from a different nanoscale sensor. The signals from the sensors vary the resonance of each device very slightly. It does this by varying the baseline stress in each beam, much like tuning a guitar string by varying the tension.

What this means is that if we force each NEMS device to resonate at a fixed frequency, 'tuning' the natural frequency of each device using the output signals from a nanosensor will cause fluctuations in the amplitude, and thence power of the mechanical resonance of each NEMS device. If this correlation between power fluctuation and sensor signal is linear, which is the case for our system, then the output from a nanosensor can be encoded using those fluctuations.

Now that each sensor channel has been encoded from the time domain into the frequency domain, we can multiplex the signals in frequency space onto a single output channel. In our scheme, we do this optically by shining a laser on the different resonators concurrently. The mechanical resonance causes power fluctuations in the reflected laser light at different frequencies via interference. This single optical output encodes all the information from the different sensors, and can be decoded to recover the original signal using a reverse process, in our case done using a standard signal processing technique. As each frequency is tied to a specific, known sensor, the system is addressable. This is the broadcast system that Feynman had speculated about in his lectures.

Although NEMS are being extensively studied as nanoscale sensors in their own right, we posit that a system like ours could allow NEMS to in fact interface other types of nanoscale voltage sensors for *in vivo* brain measurement (nanowires, carbon nanotube sensors etc) to the external environment through a single optical output.

Our technique potentially could overcome the wiring limitations of neural probes with a new type of probe design, as shown in Fig. 1.2. A piezoelectric NEMS device could be fabricated beneath every microelectrode to transduce the signal. In the core of the probe, an optical waveguide could be engineered to transmit laser light generated externally and to transmit the reflected signals

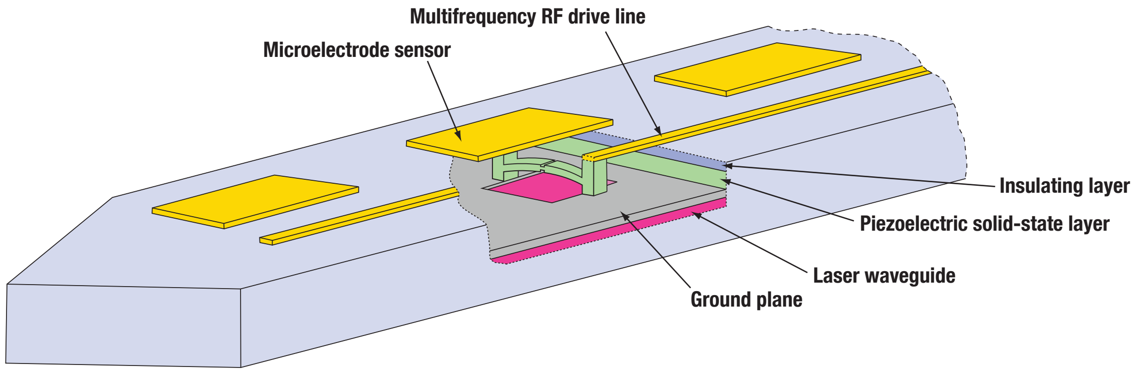


Figure 1.2: **Eliminating wiring on neural probes with piezoelectric nanomechanical Fourier transducers coupled via fibreoptics.** Illustration of proposed implementation for building large-scale neural probes with piezoeolectric nanomechanical resonators interfaced via a single optical fibre

from each NEMS device, all along the same waveguide. This could potentially increase the number of electrodes that can be packed from 100 to the tens of thousands, and could also reduce the cross-section of the probe, minimizing tissue damage. Such a system would be a highly important tool in neuroscience and for neuroprosthetics in medicine.

In our work, we have coupled one of our devices to an extracellular microelectrode and demonstrated that our system is sensitive enough to transduce neuronal action potentials from neurons. This is the first time NEMS have been used for neural recording.

In addition to developing minimally invasive neural interfaces, our system might lead to wireless nanoscale sensor arrays with applications in medical diagnostics and other applications where a mobile, miniature, highly integrated sensor array is required. In our experiments we used an infrared laser to perform the interferometry. Potentially, infra red lasers could be used that are in the range of wavelengths transparent to human tissue (see Appendix D). This would enable their use as non-invasive biological sensors.

The system as we have described it has a few limitations. It requires three separate interconnects to interface with the outside world; an optical interconnect to carry the output signals, a ground wire and a multi-frequency AC drive input wire. Furthermore, it has an ultimate bandwidth of 10,000 sensor channels, under realistic conditions. It also relies on an amplitude modulation scheme, which is more prone to noise. With a clear view to tackling these limitations, I joined Axel Scherer's group during the final phase of this work. In Chapter 4, we explore how we can massively extend the bandwidth and harness frequency and phase modulation through combination of the technology in our initial scheme with nanophotonics. NEMS devices can be actuated through purely optical methods using evanescent wave techniques, meaning actuation and transduction can be done optically. Thus only a single fibreoptic channel would replace the three external interconnects in the original iteration. Furthermore, the integration with nanophotonics will allow the nanomechanical resonators to be coupled with nano-optical resonators, increasing the ultimate multiplexing bandwidth to *billions* of neural sensor channels. If such a fibreoptic system were deployed in a similar manner to the architecture of the brain vasculature, this bandwidth would be sufficient to record the spiking output from every neuron in the mammalian neocortex. Fibreoptic communications will have truly arrived at the brain.

To complete the interfacing loop, in Chapter 5, we explore novel ideas for how microwave communications may be used for wireless *actuation* of neural circuits, using molecular scale resonators delivered to the brain via the vasculature.

Chapter 2

Wiring Nanoscale Sensors with Nanomechanical Resonators

> ...why may we not say that all automata (engines that move themselves by springs and wheels as does a watch) have an artificial life? For what is the heart, but a spring; and the nerves, but so many strings...
>
> *Thomas Hobbes,* Leviathan *(1669)*

2.1 A Nanoscale Cochlea

A multitude of promising nanoscale electronic devices have been developed for the implementation of gain [35], digital logic [36] and analogue sensing [37, 38, 39, 40]. Despite this progress, far less attention has been directed on how to interface systems built from these components with the macroscopic world [41, 42, 43, 44]. The top-down wiring problem is one of several obstacles to the fulfillment of nanotechnology's promise of extremely large scale systems integrated within a minuscule footprint. This is especially apparent for analogue sensors based on carbon nanotubes [37], semiconductor nanowires [38], nanomechanical systems [39, 40] and other devices, where the number of output channels scales in direct proportion to the number of components. Sensor arrays based on these devices have the potential to allow real-time, parallel detection of thousands of different chemical and biological species *in vivo* [38, 45, 46]. Significantly, interfacing with neural systems in a highly integrated and minimally invasive fashion might also become feasible [47]. However, bulky lithographic wires are currently used to connect with individual devices to extract information and provide power. This limits their integration and keeps them tethered to the external environment.

In his lectures on computation, Richard Feynman speculated that the wiring problem might be tackled by an optical interconnection system through free space, acted on by frequency-sensitive components within an architecture [32]. Here we report on a new paradigm for wiring nanoscale sensor arrays to the macroscopic world, employing frequency-sensitive nanoelectromechanical systems (NEMS) coupled to a single optical output. Potentially, thousands of sensor channels could be multiplexed in parallel, eliminating the wiring constraints that arise with scaling.

Our system implements a mechanical Fourier transform for interfacing signals, and is inspired by the mechanism of action of the cochlea in the inner ear. In the human cochlea the basilar and tectorial membranes couple to the three rows of piezoelectric outer hair cells, in a poorly understood fashion, to form a bank of resonant elements that range from 20 Hz to 20 kHz in frequency [48]. The inner hair cells form the electromechanical sensing elements, and can detect deflections as small as several Angstroms [49]. Incoming sound is decomposed in the cochlea by the resonant elements, and variations in the power of the signal at different frequencies are transmitted through individual nerve channels via the respectively coupled inner hair cells [50]. In our scheme we implement the inverse case, coupling electrical signals from individual sensor devices to a bank of piezoelectric nanomechanical resonators, each resonating at a unique frequency. The analogue signals vary the amplitude of the mechanical resonance in a linear fashion, and optical interferometry is used to detect the mechanical motion from all the resonators concurrently. External demodulation of the optical signal can reveal the variations in power at each frequency, and hence can recover the original multiplexed signals in real-time and in an addressable manner.

2.2 Piezoelectric Nanoelectromechanical Systems

The basic component of our technique is a doubly clamped piezoelectric NEMS beam, as shown in Fig. 2.1, which is operated in a novel manner so as to permit real-time signal transduction. Our devices are fabricated from epitaxially grown GaAs, and as depicted in the cross section in Fig. 2.2, comprise of a 200 nm thick *pin* diode, with the *n*- and *p*-doped layers connected to the voltage source and ground respectively. Application of a potential across the *pin* junction results in an electric field that is concentrated across the $\sim$ 50 nm thick charge depletion region. The transverse electric field induces a longitudinal piezoelectric strain within the intrinsic layer,which can be used to resonantly excite the flexural mode of the beam through application of a radiofrequency (RF) drive.

The high impedance of the depletion layer means these devices exhibit very low power dissipation during actuation, on the order of 1-10 nW. As shown in Fig. 2.3, under laser illumination, the impedance of the devices are approximately a few megohms.

2.3 Frequency Tuning

As has been reported previously, clamped beam devices of this class can be tuned in resonant frequency through application of a DC bias potential [33]. Fig. 2.4 illustrates this frequency tuning behaviour, and how the resonance curve of the device shifts under voltage bias. This effect is due to the clamped boundary conditions of the beam, which transduce strain into stress, thus modifying the mechanical properties of the structure and its natural frequency. Importantly, this shift is linear with respect to applied voltage, with a slope of approximately 40 kHz per volt for our devices, as demonstrated in Fig. 2.5. While non-piezoelectric NEMS devices can also transduce a DC bias through electrostatic coupling, their tuning action is nonlinear and thus less suitable for recovery of analogue signals [51, 52, 53].

Here we use this frequency tuning effect to transduce voltage signals in real-time, that are more slowly varying than the radiofrequency drive. Normally, signals from NEMS are extracted by directly monitoring the shift in resonance frequency of the structure via a phase-locked loop, as with mass accretion measurements [40]. In this manner, changes over minutes to hours are tracked. In the present scheme instead of tracking changes in resonance frequency, we drive the NEMS at a fixed frequency and monitor variations in oscillation amplitude as the bias signal varies the fundamental of the resonator. In this manner, analogue voltage signals with a bandwidth of several tens of kilohertz can be transduced.

In addition to bandwidth, an important consideration is signal sensitivity. If signals as small as tens of microvolts can be transduced, this allows for signal transmission from nanoscale sensors that do not have intrinsic gain. Reducing the device thickness has been the primary means of maximising sensitivity [33]. This is the case with piezoelectric NEMS as a thinner beam reduces the amount of force required to deflect it and a small potential can setup a stronger electric field across the intrinsic layer. Here we report on a novel means of improving the sensitivity in adjunct to this, through driving the NEMS devices in the regime of nonlinear bistability. As shown in Fig. 2.6, the devices were driven near the edge of the bistability region, at the jump point of the resonance response

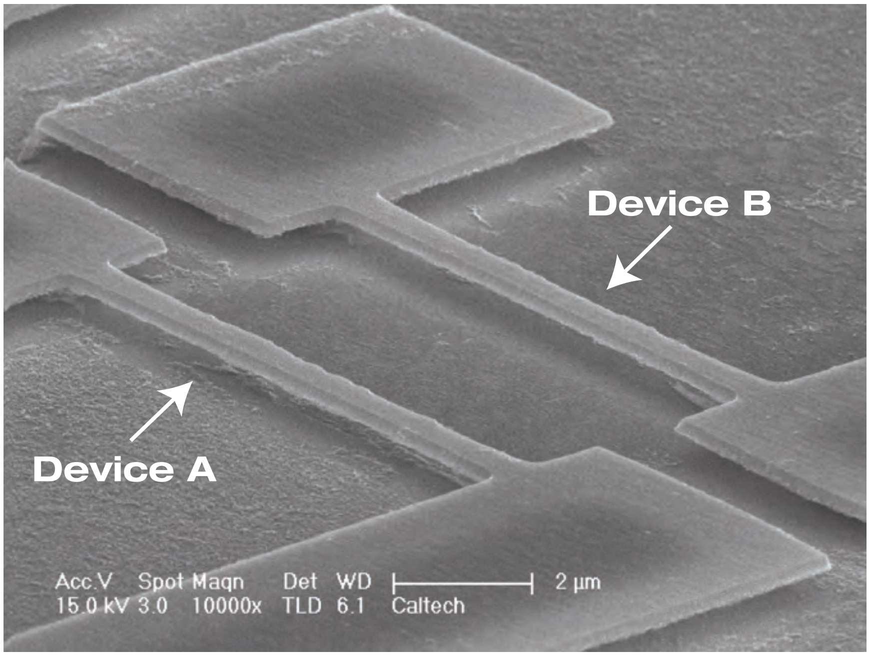


Figure 2.1: **GaAs NEMS devices**. Dual piezoelectric NEMS devices A and B spaced 4μm apart. The dimensions of the beams are (l, w, h) = (9.0, 0.8, 0.2) and (8.5, 0.8, 0.2) μm for A and B, respectively.

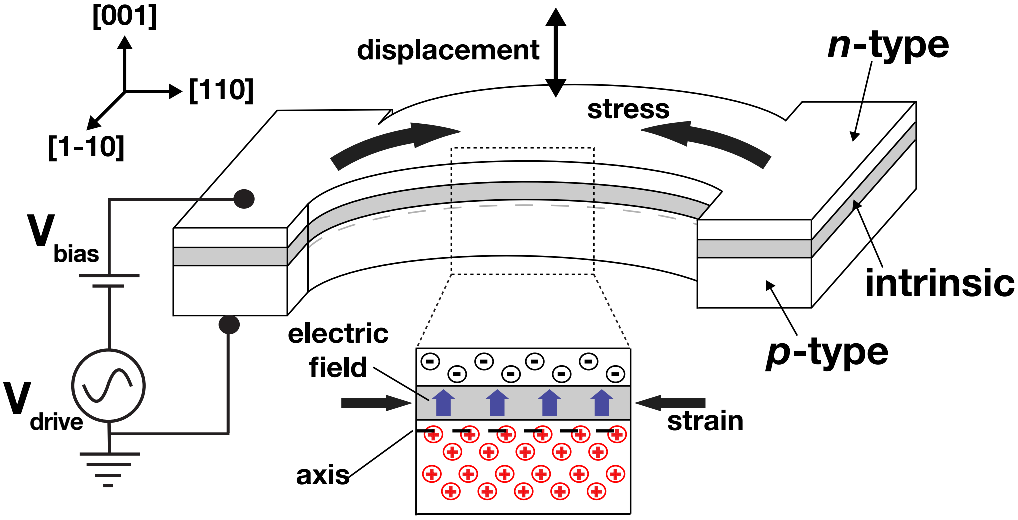


Figure 2.2: **Piezoelectric NEMS**. An electric field across the *pin* junction generates longitudinal piezoelectric strain. An actuating radiofrequency potential, V_{drive}, drives the beam to resonance. A slowly varying potential, V_{bias}, modulates the baseline stress in the beam and tunes the resonance frequency.

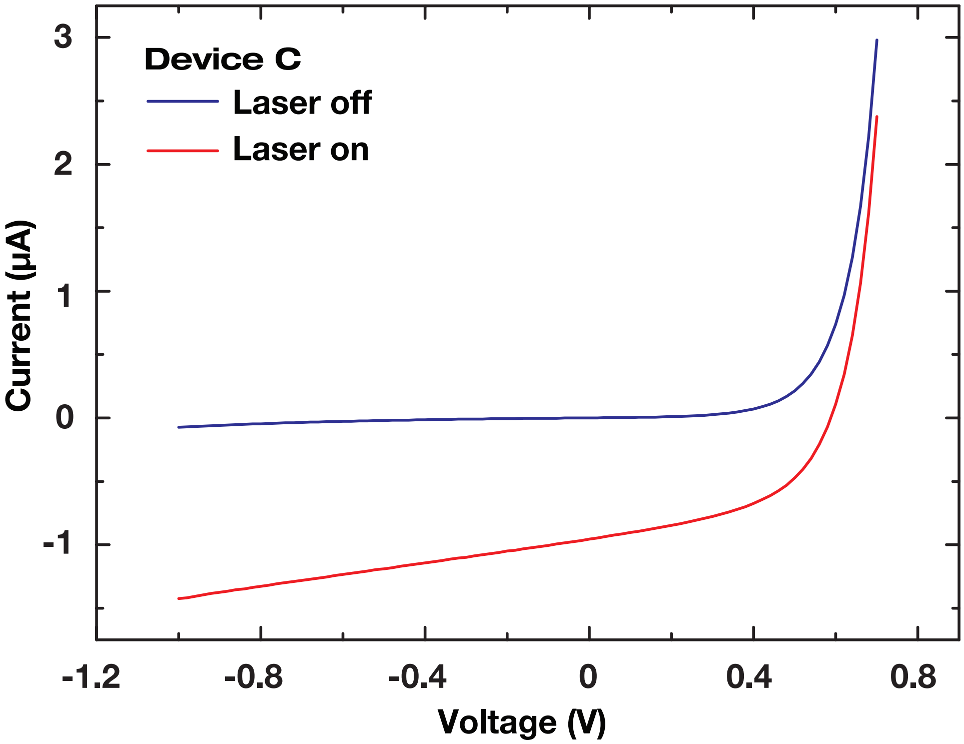


Figure 2.3: **Current-voltage characteristics of devices under laser illumination.** The NEMS devices are constructed as *pin* diodes, and thus share the I-V characteristics and high impedance properties of such diodes. Although the energy of the diode laser used in the interferometry (904 nm) is less than the band gap of GaAs, there was still a small amount of current generated when the devices were irradiated. Thus under operating conditions, the impedance of the devices were reduced, albeit still high. As shown for device C, the impedance dropped from 16 MΩ to 2.1 MΩ under illumination. This translates into a power dissipation of $\sim$ 1 nW when driven in the regime of nonlinear resonance.

curve. [1] The steep response improved sensitivity by an order of magnitude. The demodulated signal amplitude can be expressed by:

$$A(t) = -\frac{d_{31}\eta(Q, |\, V_{\text{drive}}\,|)}{h^2}\sqrt{\frac{3E}{\rho}}V(t)\ , \tag{2.1}$$

where $\eta(Q, |\, V_{\text{drive}}\,|)$ is the slope of the resonance jump point represented in Fig. 2.6, E, ρ and d_{31} are the Young's modulus, density and transverse piezoelectric coefficient of GaAs, h is the device thickness and $V(t)$ is the bias signal. The slope η is steeper for greater quality factor Q and driving potential V_{drive}.

---

[1] Note that the X-component is the imaginary component of the resonance curve. If $|Z|$ is the absolute magnitude of the signal, then $Z = Y + iX$. In the case of an RC resonator, X would be the reactance, Y the resistance, and Z the impedance of the circuit.

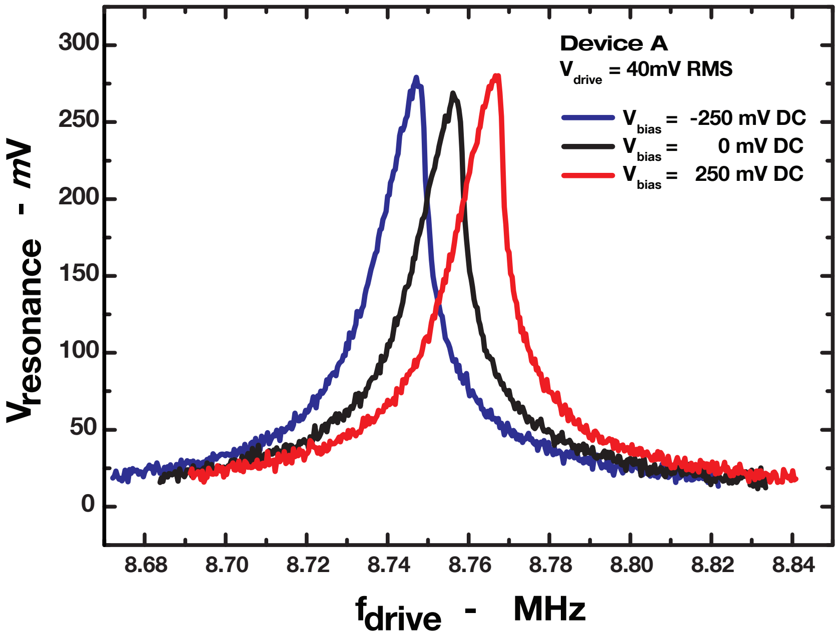


Figure 2.4: **Resonance frequency tuning in NEMS via the piezoelectric effect.** Resonance curves for device A obtained under conditions of positive, negative and zero voltage bias. The resonance peaks shift with voltage bias, illustrating the frequency tuning effect.

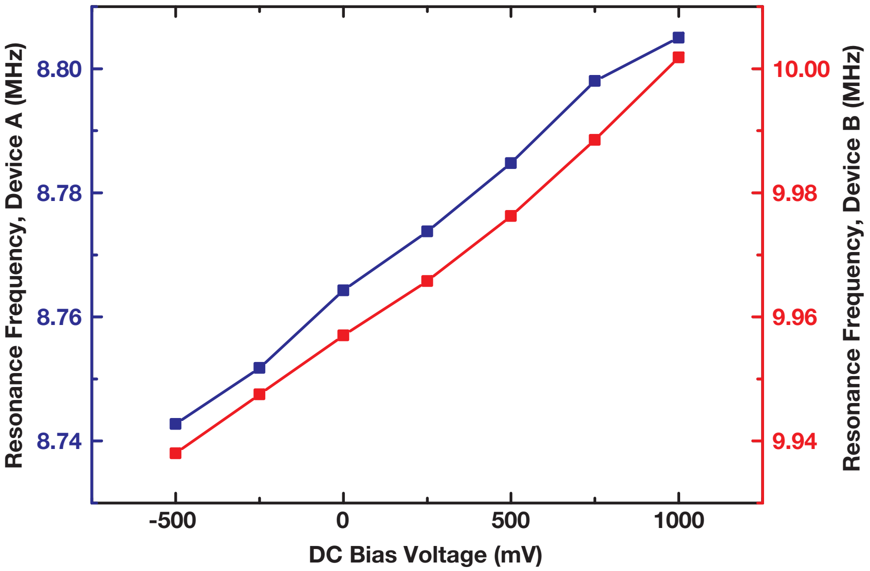


Figure 2.5: **Linearity of frequency tuning.** Resonance frequency versus bias voltage for devices A and B. Resonance frequency corresponded to the peak amplitude in a frequency-swept measurement carried out with a network analyzer. All potentials are measured with respect to the negative (i.e., top) terminal of the *pin* junction comprising the nanomechanical structures. The frequency-voltage curve is linear with a slope of approximately 40 kHz/V for both devices, signifying the tuning effect is independent of device length.

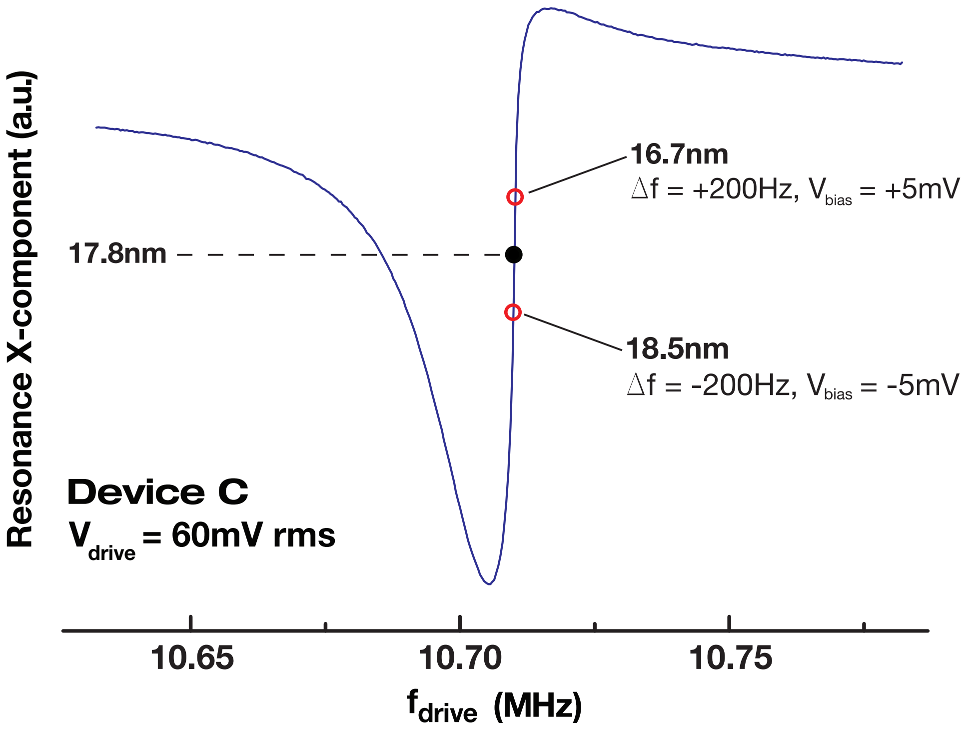


Figure 2.6: **Nonlinear drive.** X-component of resonance plot for device C when driven nonlinearly. The central slope $\eta(Q, \mid V_{\text{drive}} \mid)$ is steep due to high Q $(= 940)$ and V_{drive} $(= 60\ \text{mV}_{\text{rms}})$. A lock-in amplifier locks on to the steepest point (black). When V_{bias} is applied the resonance plot shifts along the frequency axis, and the lock-in point varies in amplitude (red circles). In this manner, signals are encoded as power variations at a specific frequency and can be transduced in real-time. The mechanical oscillation amplitude of the resonator corresponding to bias voltages at +5, 0 and -5 mV are indicated.

2.4 Signal Multiplexing

A clamped beam NEMS resonator can be tailored to a specific fundamental frequency by fabricating it of a particular length. The natural mode frequency for a uniform beam is given by [54]:

$$f_i = \frac{\lambda_i^2}{2\pi L^2}\sqrt{\frac{EI}{m}}\,, \tag{2.2}$$

where I is the moment of inertia of the beam, m is the mass per unit length, L is the beam length, and λ_i is a parameter dependent on the mode and clamping conditions. Since $m = \rho w h$, where w is the beam width, $I = wh^3/12$, and $\lambda_1 = 4.73$ for a clamped-clamped beam, the fundamental frequency of a doubly clamped NEMS resonator is:

$$f_1 \approx 1.028\frac{h}{L^2}\sqrt{\frac{E}{\rho}}\,. \tag{2.3}$$

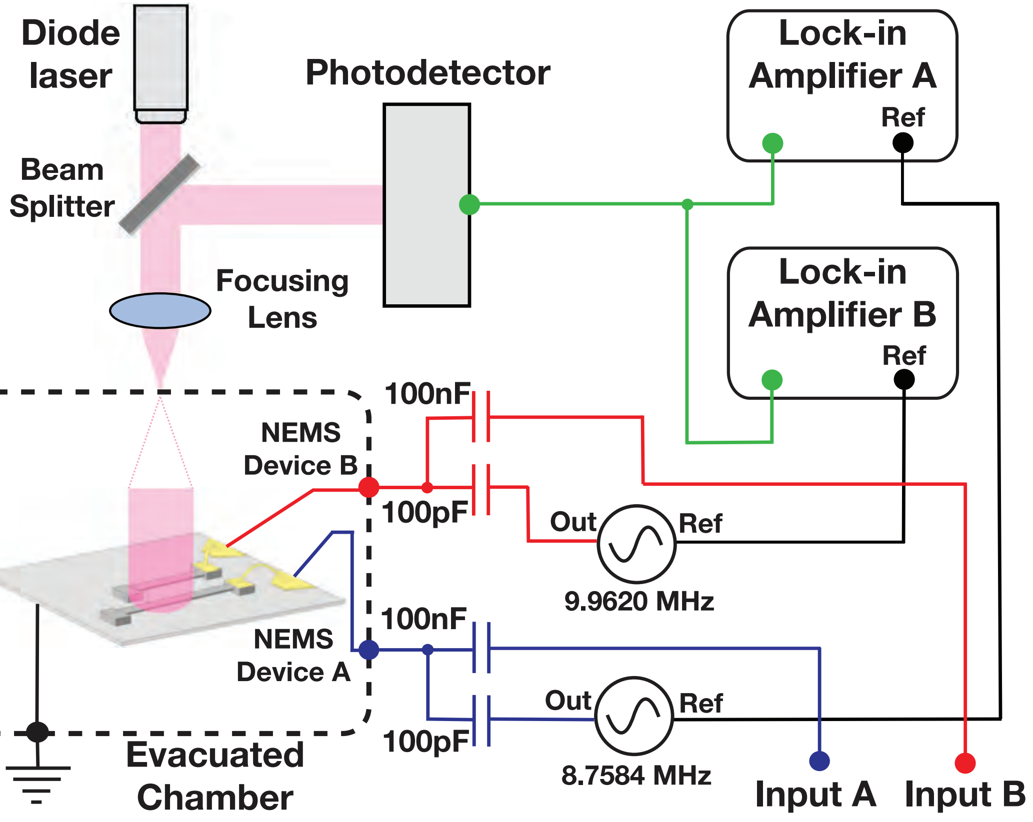


Figure 2.7: **Experimental setup.** Schematic of experimental setup to demonstrate multiplexing through a nanomechanical Fourier transform. All mechanical displacement is read out via a common optical transmission path.

A bank of closely spaced resonators of differing lengths can then be used as a nano mechanical Fourier transform, encoding signals in the time domain into power variations in frequency space using the modulation scheme described. Optical techniques can then monitor the oscillation amplitude of the resonators concurrently through a single optical channel. External demodulation of the optical signal can recover and uniquely identify a signal from a particular input channel based on its RF actuation frequency. NEMS resonators with frequencies ranging from several megahertz [51] to over 1 GHz [55] can be made, suggesting thousands of different channels could be encoded onto a single output. We demonstrate this concept, using laser interferometry [56] to monitor variations in transverse oscillation amplitude from two closely spaced NEMS devices A and B, as shown in Fig. 2.1. They have lengths of 9.0 and 8.5 μm, and resonate at 8.758 and 9.962 MHz, respectively.

Fig. 2.7 illustrates the experimental setup to demonstrate the nanomechanical Fourier transform. The laser was focused to a diameter of $\sim$ 10 μm, allowing concomitant observation of both resonant structures. The devices were operated *in vacuo* to maintain high quality factor and resonance frequency stability. Each device was capacitively coupled to two signals: (i) the driving potential V_{drive}, whose frequency corresponds to the region of highest X-quadrature slope in Fig. 2.6, and (ii) the bias signal used to simulate the output of an analogue nanoscale device. The single channel photodetector signal was demodulated using separate lock-in amplifiers at f_{drive} for each device.

Fig. 2.8 shows successful demodulation of multiplexed bias signals that were applied to devices A and B. We were able to closely reproduce the fine structure of the pulse applied to device A and the sinusoidally varying signal on device B. As shown in Fig. 2.9, we verified that the cross-talk between signals on each channel did not exceed 1.6% at 1 kHz, despite the close spacing of the devices. It was possible to accurately track bias signals for tens of minutes without the use of feedback control of the radiofrequency driving potential parameters. Over longer timescales we observed drift that was compensated by a single-shot manual recalibration of f_{drive}.

Fig. 2.10 shows the boost in NEMS response as a result of driving in the regime of nonlinear bistability. This effect also increases the limiting voltage sensitivity of the device. Shown is the response of device A to a 1 kHz sinusoidal bias signal as the driving amplitude is increased, with the inset showing resonance curves at equivalent values. For device A the maximal response occurs at $V_{\mathrm{drive}} = 51\ \mathrm{mV_{rms}}$. At higher drive levels the tuned frequency of the device appears to become unstable, as suggested by the abrupt decrease in response. We verified that the noise did not appreciably change below the instability (see Fig. 2.16) and thus confirmed that the optimal signal-to-noise ratio (SNR) is achieved by actuating the NEMS devices at levels slightly lower than the critical point.

At low bias levels the NEMS devices exhibit a linear response to V_{bias}, as seen in Fig. 2.11. Linearity is requisite for analogue signal transduction. The earlier onset of nonlinear response in device A at $V_{\mathrm{bias}} \sim 40\ \mathrm{mV_{pp}}$ is due its higher quality factor.

Operation of the devices in the nonlinear bistable region led to attenuated response at greater signal frequencies. Figure 2.12 characterizes the frequency response for devices A and B at different driving amplitudes. For both devices the response curve remains flat at low V_{drive}, but becomes progressively biased toward lower frequencies as V_{drive} approaches the NEMS bistability regime. The observed attenuation effect is more marked for the device with higher quality factor. These observations appear to be a manifestation of the ring-down time of a resonator ($\sim Q/4\pi f_0$), which increases in proportion to Q [57] and is further augmented in nonlinearly excited systems. Our results agree with the theoretical prediction that the greatest frequency shift sensitivity of a NEMS resonator is achieved near the edge of the bistability region, but that this region exhibits the longest

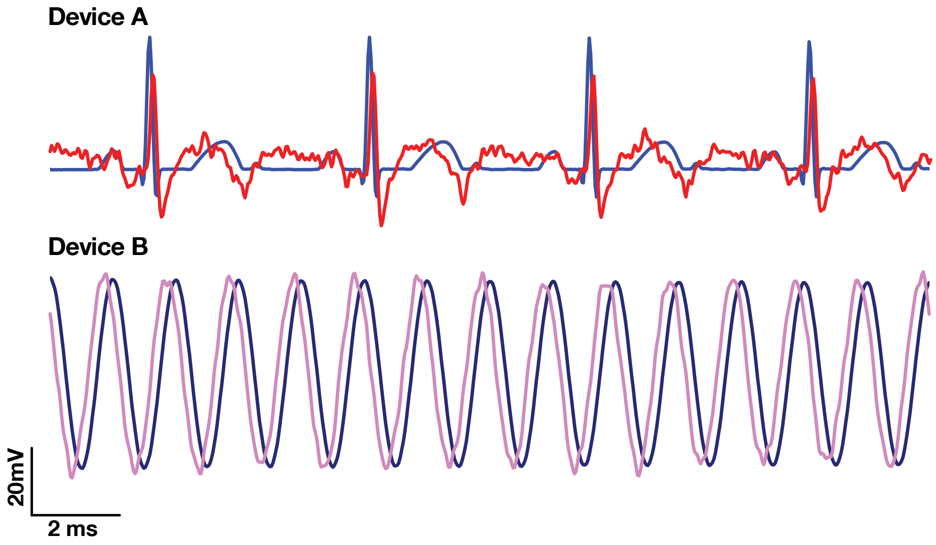


Figure 2.8: **Multiplexing.** Signals (red & purple) recovered from multiplexed optical readout of devices A and B through an inverse Fourier transform. The corresponding bias signal on each device is represented by a dashed blue line. The small phase offset originates rom the low-pass RC filter formed by the signal output impedance and large parasitic capacitance at the device wirebond contact region. Results represent 450 averages.

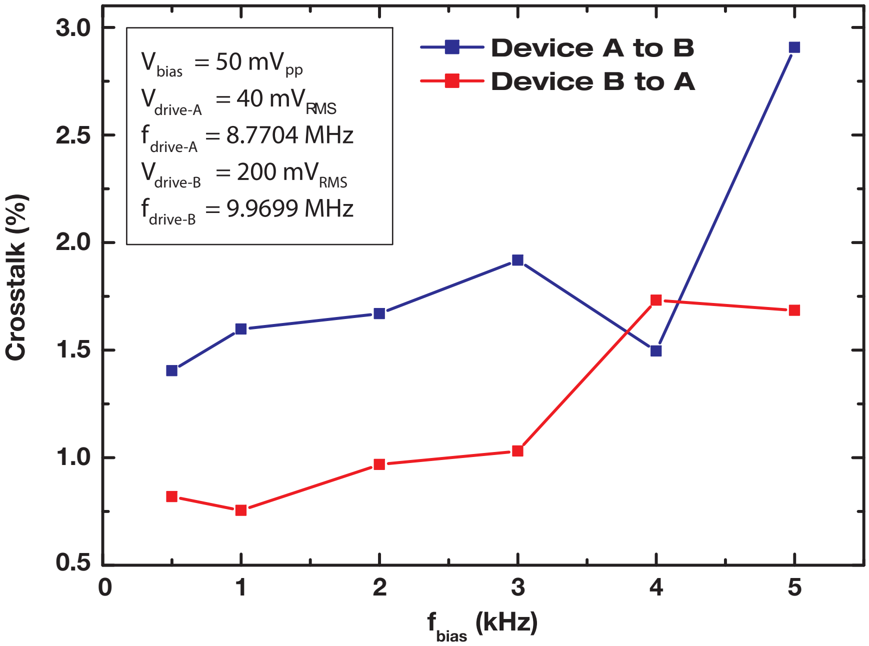


Figure 2.9: **Crosstalk.** Crosstalk versus frequency between devices A and B. A 50 mV_{p-p} sinusoidal bias was applied to device A, and the demodulated photodetector signal was measured from the unbiased device B. Thus any crosstalk would manifest as a spurious signal on device B at the applied bias frequency. The percent crosstalk is calculated from the ratio of the spurious signal on B to the signal that is measured when a 50 mV_{p-p} bias is directly applied to B. A similar measurement was also made with the bias signal applied to device B and crosstalk measured from device A.

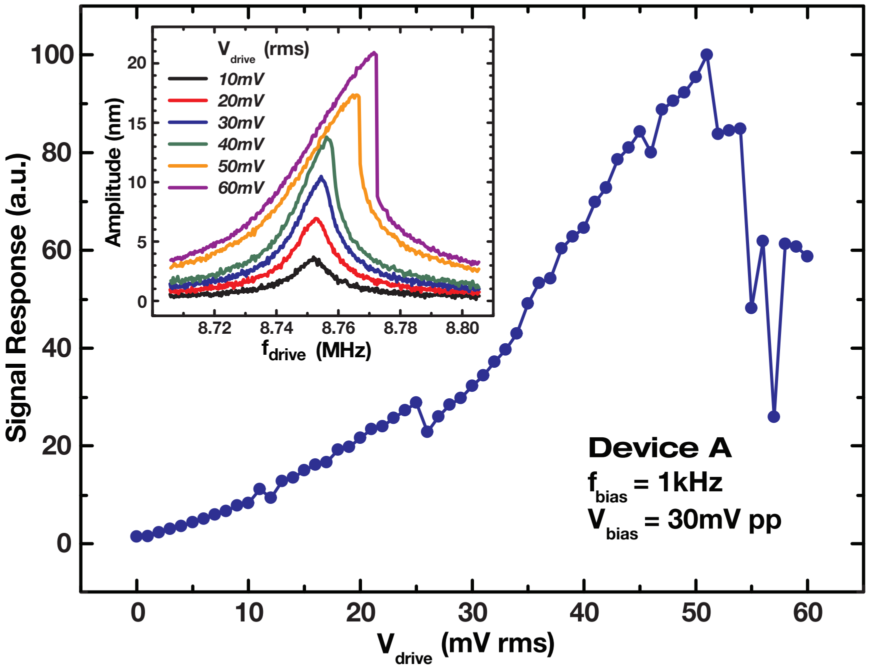


Figure 2.10: **Signal enhancement in nonlinear regime.** Enhancement of device A responsivity with increasing radiofrequency drive amplitude. The bias signal amplitude and frequency were set to 30 $\mathrm{mV_{rms}}$ and 1 kHz respectively.The inset shows resonance plots of the flexural oscillation amplitude at different drive levels. The maximal response is obtained at $\sim 50\,\mathrm{mV_{rms}}$, corresponding to the region of strong nonlinearity.

ring-down effect and hence slowest response time [58]. For electrical sensors requiring both high sensitivity and bandwidth, nanomechanical transduction would therefore necessarily rely on the use of higher operating frequency NEMS resonators. As the system is linear, the original bias signals can be recovered through deconvolution with the frequency response characteristic of each device.

The particular application of the system will determine the device quality factor specification and driving regime. This is due to the tradeoff in ultimate wiring density, signal bandwidth, sensitivity and SNR that arises from these factors. Greater quality factor increases the device sensitivity, SNR and the number of channels that can be multiplexed, but impacts signal bandwidth and dynamic range of the transduced bias signal. Stronger nonlinear drive increases sensitivity and SNR but is disadvantageous to the other parameters. Passive sensors that generate very small potentials would especially benefit from NEMS devices with high quality factor that are driven nonlinearly, and active sensors could operate with lower supply voltage.

Some applications of sensors require measurement of low amplitude, low SNR bias signals. This challenge is particularly salient in the monitoring of electrophysiological activity in the brain

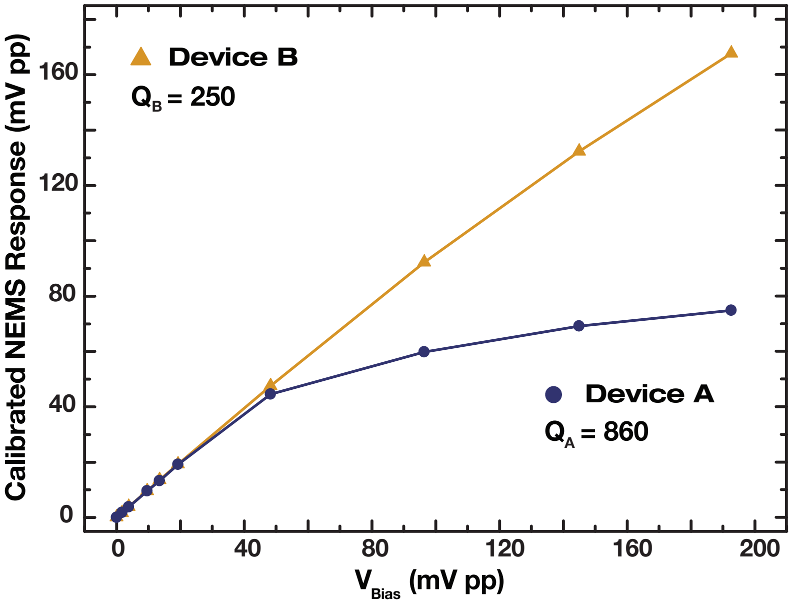


Figure 2.11: **Nonlinearity and resonator quality factor.** Nanomechanically transduced signal amplitude of a 1 kHz sinusoidal signal applied to devices A & B as a function of the applied bias amplitude. The actuation parameters were $V_{\text{drive,A}} = 40\ \text{mV}_{\text{rms}}$ and $V_{\text{drive,B}} = 200\ \text{mV}_{\text{rms}}$. In the regime of low voltage bias, even devices with high quality factor are capable of linear transduction.

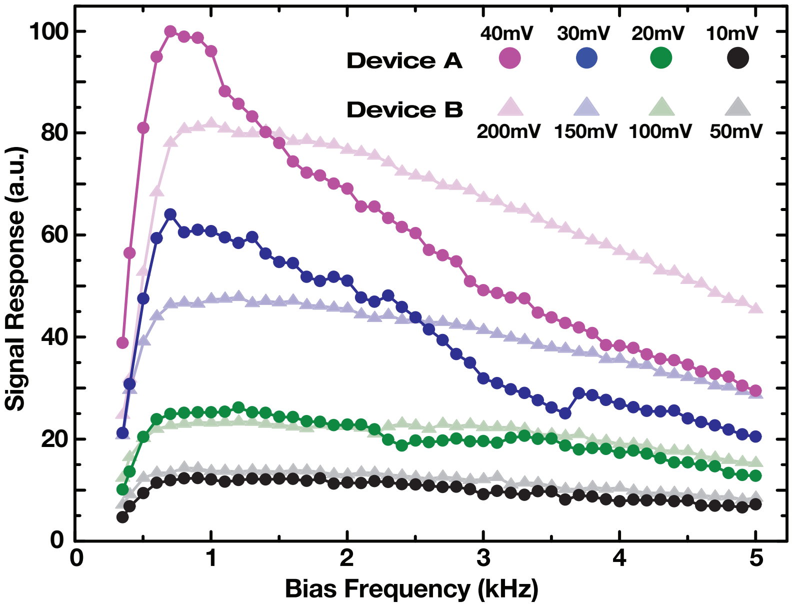


Figure 2.12: **Power spectrum of signal transduction.** Bias frequency response of devices A and B at various driving amplitudes in the linear and nonlinear NEMS actuation regime. The response is comparatively flat in the regime of linear drive, but becomes increasingly biased toward low frequency signals as the device approaches the nonlinear regime. The effect is more marked for device A, which has higher Q and hence lower threshold for nonlinearity. Note that the notch at $f_{\text{bias}} < 500$ Hz is due to our applied high-pass filter and is not an intrinsic NEMS effect. V_{bias} was set to $30\ \text{mV}_{\text{rms}}$ for both devices. For all splots the demodulated amplitude was calculated over 10 periods and averaged 2,500 times.

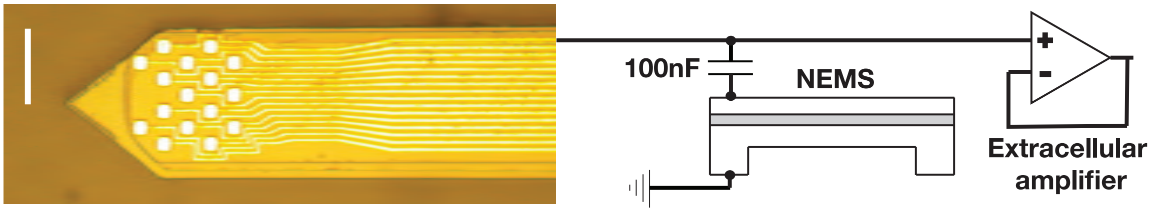


Figure 2.13: **Coupling of neural probe NEMS device.** Neural probe used in our experiments. A single channel connected in parallel to NEMS device C, (l, w, h) = (8.0, 0.8, 0.2) μm, and an extracellular amplifier. The scale bar is 50μm in length.

using multisite extracellular electrode arrays. Such recordings are an invaluable method for studying functional properties of the intact brain from the microcircuit to systems level [15]. Recordings can be carried out with an implantable neural probe that connects a grid of microscale electrodes to amplifying electronics via wires spanning the length of the probe. As the complexity of such neural interfaces increases, the primary size-determining factor of the devices becomes the interconnecting wires, and an arbitrary reduction of wire width is not a practical solution [59].

Application of our technique has the potential to enable fabrication of much higher density and less invasive neural interfaces. We therefore report on the the effectiveness of transducing extracellular action potentials via a NEMS-neural probe interface. A stand-alone resonator (device C) was coupled to a recording site on the probe shown in Fig. 2.13 (see Chapter 3 for fabrication details). The microelectrode array was inserted into a locust metathoracic ganglion preparation induced to sustain rhythmically discharging neurons. The electrode was connected to the resonator via a blocking capacitor, which prevented photoelectric charge generated at the GaA device from coupling to the electrode. A parallel connection went to an extracellular electronic amplifier for cross-validation of the nanomechanical measurement. Extracellular neuronal spikes were identified from the amplifier recording, allowing spike-triggered averages to be obtained on the synchronously collected measurements from the NEMS and electronic amplifier. The mean waveform of an 80 $\mu\mathrm{V}_{\mathrm{pp}}$ spike is depicted in Fig. 2.14 and is simultaneously observed at the demodulated output of the NEMS resonator. We confirmed the mechanical origin of the demodulated signal by verifying that the action potential is no longer measured if the resonator is detuned by $\sim$1% from its fundamental frequency.

2.5 Noise

The SNR is small when transducing signals as diminutive as those from passive extracellular microelectrodes. In order to achieve sufficient SNR for spike detection it was necessary to average the

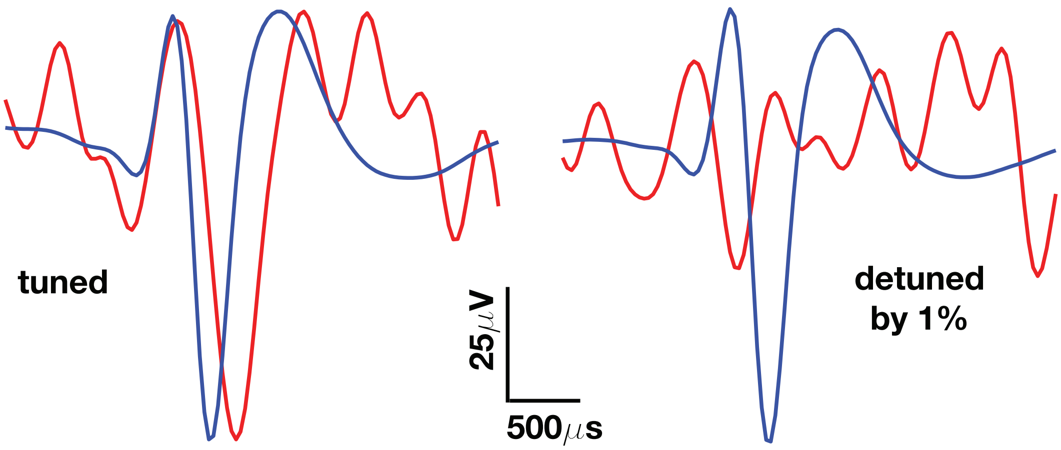


Figure 2.14: **Nanomechanical transduction of extracellular action potentials.** Detection of extracellular action potentials in a locust ganglion preparation using the setup in Fig. 2.13. Demodulated signals from the NEMS device are shown in red, and the corresponding signal recorded by the amplifier is indicated by blue lines. Mechanical readout was confirmed through the extinction of spike detection upon detuning the resonator by 100 kHz. Data is averaged 5000 times.

NEMS signal 5000 times. The best noise performance we could achieve in the frequency band of 350 to 3000 Hz was 400 $\mu\mathrm{V}_{\mathrm{RMS}}$ referred to the microelectrode input, as shown by measurements displayed in Fig.2.15. In contrast, conventional electronic amplifiers designed for electrophysiological recording have a noise floor approaching 1 $\mu\mathrm{V}_{\mathrm{RMS}}$ [60]. Thus for neural recording the SNR would need to improved by at least two orders of magnitude. We verified that in fact approximately 80 % of the measured noise was not intrinsic to the devices but originated in the external demodulation electronics. Fig. 2.17 shows the results of noise measurements on a grounded NEMS device, with the full optical interferometry and demodulation setup taking the measurements. Even with the RF drive cut to the device, most of the noise remained. A power spectrum of the noise, presented in Fig. 2.18, also shows that the major noise contribution arises from $1/f$ noise and 60 Hz noise and its harmonics. Clearly, there is opportunity for substantive improvement in performance.

A number of engineering improvements could be made to the devices themselves to improve SNR. In particular, reduction of parasitic capacitance through miniaturization of device contacts would significantly ameliorate signal attenuation effects. The output impedance of the extracellular electrode coupled with the capacitance of the NEMS device (the wirebonding pad being the major contributor to capacitance) creates a first-order low-pass RC filter. For an electroplated electrode, $R \approx 30$ kΩ, and for the wirebonding pad $C \approx 1$ nF. The cut-off is thus $f_{3\mathrm{dB}} = 1/(2\pi RC) \approx$ 15 kHz. Hence, at 15 kHz the signal attenuation is $\sim$30%, and at 30 kHz it is $\sim$50%. Decreasing the device contact capacitance through more advanced nanofabrication techniques (such as using a

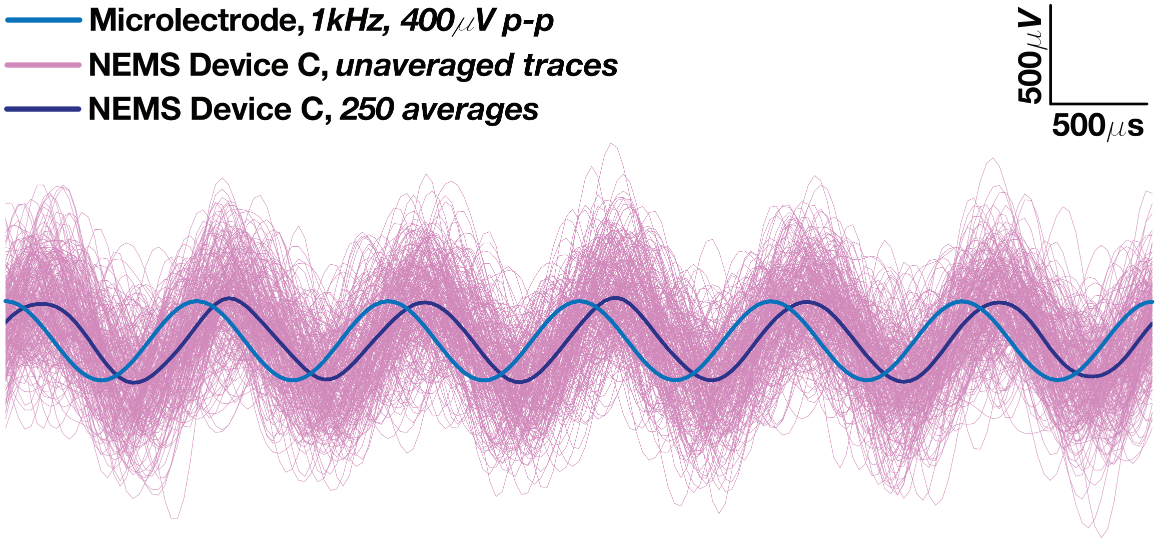


Figure 2.15: **Signal-to-noise ratio.** The neural probe coupled to the NEMS device was used to detect an artificially generated signal in saline solution. Nanomechanical signal appears in dark blue (250 averages). Individual, unaveraged NEMS traces are indicated in magenta. The averaged trace from the extracellular amplifier is in light blue. The SNR of the transduced signal is ~1:1.

bridge structure for electrical contact) can reduce the parasitic capacitance by a factor of 10^3. This would completely eliminate the RC attenuating effect in the passband of interest.

Quality factor enhancement through more meticulous nanofabrication techniques and the use of more strongly coupled and stiff piezoelectric materials such as AlN would improve voltage responsivity [61]. An array of identical NEMS devices serving each channel would also be a straightforward method to boost the optical signal.

The ultimate sensitivity of piezoelectric NEMS resonators to electrical potentials is set by thermomechanical dissipation. The frequency fluctuations of a doubly clamped beam resonator are approximately given by [39]:

$$\delta f \approx \frac{0.036}{h^2(\rho E)^{1/4}} \sqrt{\frac{k_B T L \Delta f}{w}} \,, \tag{2.4}$$

where L, w, h are respectively the length, width and thickness of the beam, Δf is the measurement bandwidth, ρ is the material density and E is Young's modulus. In this approximation, frequency fluctuations are independent of quality factor, because higher Q results in lower critical vibrational amplitude (see Eq. 2.7). In order to convert Eq. 2.4 into effective fluctuations in an applied electrical potential, we must know the coupling between voltage and frequency in doubly clamped piezoelectric resonators [33]:

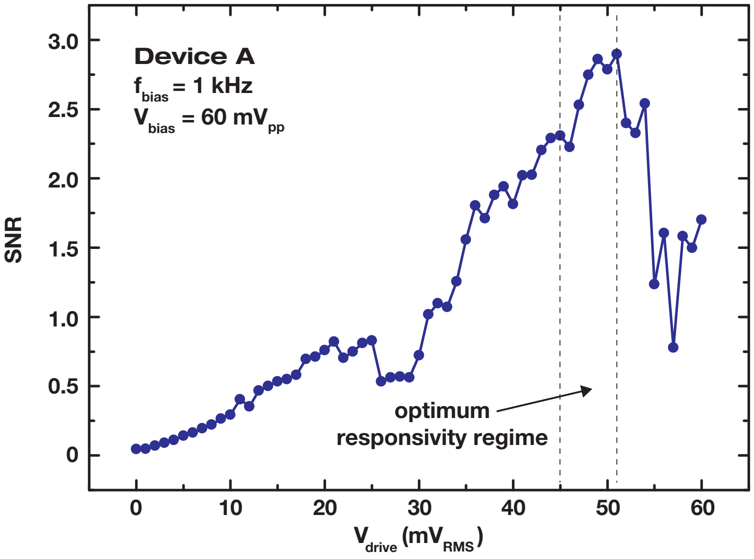


Figure 2.16: **Signal-to-noise ratio enhancement under nonlinear operation.** The SNR for device A is plotted as a function of drive potential. The signal was measured as the root mean square amplitude of the demodulated NEMS response, during application of a 60 $\mathrm{mV_{p-p}}$, 1 kHz sinusoidal bias across the NEMS. RMS noise was measured by turning off the applied bias and recording background activity for $\sim$ 100 seconds. Note that the regime of highest SNR occurs in the region of greatest signal in Fig. 2.10, which corresponds to the region of nonlinear nanomechanical operation.

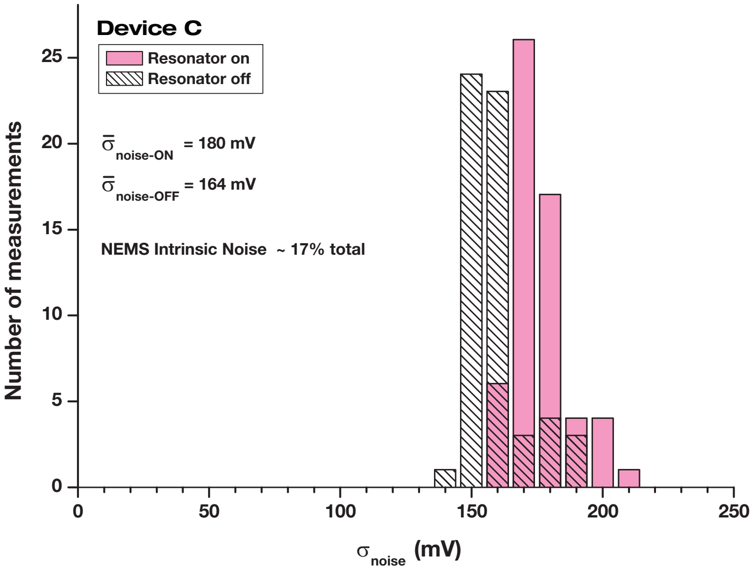


Figure 2.17: **Measurement of intrinsic noise.** The capacitive input bias terminal of device C was grounded and background noise was measured from the optical signal both with the device tuned to resonance and with the radiofrequency drive turned off. 58 recordings of 10 s duration were made under each condition and the noise standard deviation was determined in each case. Over 80% of the measured noise was not intrinsic to the device itself, but originated in the demodulation electronics. The proportion of intrinsic noise was determined using the formula $(\sigma^2_{noise-ON} - \sigma^2_{noise-OFF})/\sigma^2_{noise-ON}$. The drive potential was 40 $\mathrm{mV_{rms}}$ during device actuation. Data was filtered using an 8-pole Butterworth filter in the passband 300-10,000 Hz.

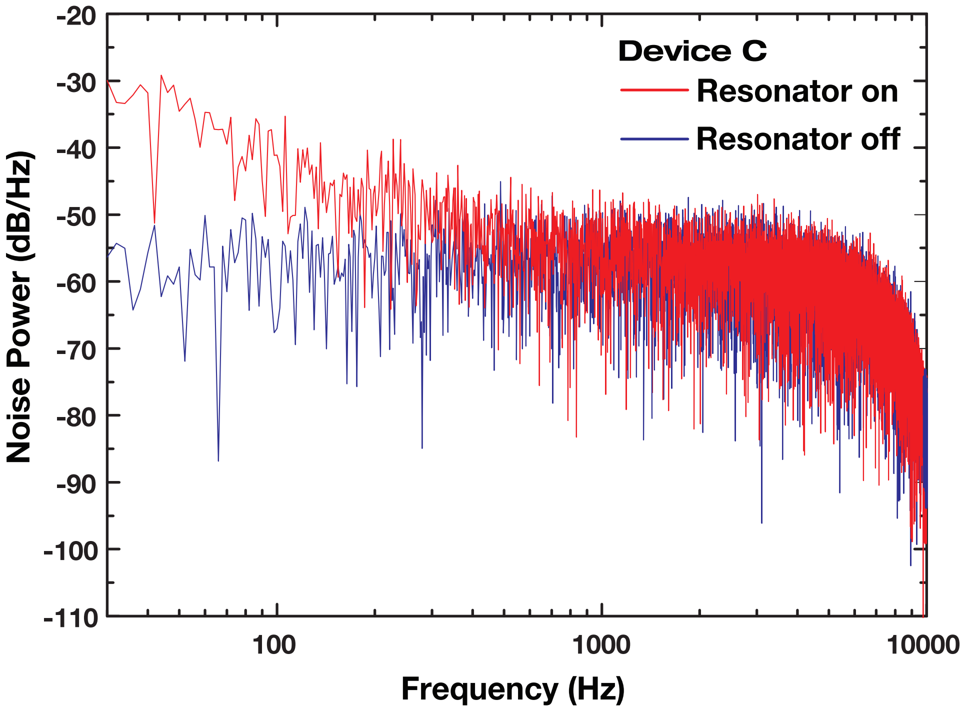


Figure 2.18: **Noise power spectrum.** Spectral analysis of the noise data from Fig. 2.17. Most of the noise during device operation stems from $1/f$ noise originating from the signal generator, and 60 Hz noise and its harmonics. There is thus substantive opportunity for reduction in the intrinsic noise itself. Data for each curve were acquired over a period of 10 minutes.

$$\left|\frac{\delta V}{\delta f}\right| = \frac{d_{31}}{2\pi h^2}\sqrt{\frac{3E}{\rho}}\ , \tag{2.5}$$

where d_{31} is the transverse piezoelectric coefficient. This leads to the expression for the minimum measurable voltage set by thermomechanical fluctuations:

$$\delta V = \frac{\rho^{1/4}}{d_{31}E^{3/4}}\sqrt{\frac{k_B T L \Delta f}{w}}\ . \tag{2.6}$$

We now consider a beam identical to device C with (L, w, h) = (8, 0.8, 0.2) m. The material coefficients of GaAs are $\rho = 5317$ kg/m^3, $E_{[110]} = 101$ GPa, and $d_{31} = 1.4$ pm/V. These values allow us to estimate $\delta V \approx 60\ \mu\mathrm{V}_{\mathrm{rms}}$. The use of AlN instead of GaAs would reduce these fluctuations by a factor of ~ 2. Significant further gains in performance are expected with device miniaturisation to nanoscale dimensions.

Let us return though to the question of intrinsic noise originating from the demodulation electronics, which is $1/f$ in nature, and which comprises the bulk of the noise. $1/f$ noise does not originate from voltage fluctuations (which are due to Johnson noise and have a flat spectrum), but rather arise from thermal noise in switching between two states [62]. This immediately suggests the culprit as the signal generators used to resonantly excite the NEMS devices.

First, consider once again Fig. 2.5 and the measured tuning slope of the NEMS resonators of 40 kHz/V. Thus for a 100 μV bias signal, the fundamental frequency of the NEMS devices shifts by just 4 Hz. For high SNR demodulation of the NEMS signal in neural recording, it's clear the electrical oscillator driving the devices needs to be of extremely high quality. This is a case where the quality and precision of the test equipment used in an experiment is absolutely critical. The actual oscillator used in our experiments was not a signal generator at all (which are designed for high spectral purity and low phase noise), but a function generator, an Agilent 33250A with 80 MHz bandwidth. This oscillator is specified as having a frequency accuracy of 2x10^{-6}, and single sideband (SSB) phase noise of -65 dBc/Hz at 30 kHz offset from 10 MHz. This is really poor performance, and inadequate for our experiments. The frequency accuracy suggests a ballpark figure of 20 Hz of frequency noise in our operating regime. This is about 10 times greater than the typical frequency shift produced by a neuronal spike, and 1000 times greater than the noise floor in a CMOS headstage neural amplifier. This noise alone would explain the poor SNR observed.

In contrast, the Rohde & Schwarz SMC100A signal generator with high stability OXCO reference has a frequency accuracy of 10^{-8} and SSB phase noise of -118 dBc/Hz at 30 kHz offset from 10 MHz. This is a significantly improved specification, and would improve the SNR by over two orders of magnitude. Furthermore, it is possible to connect an external reference to this signal generator with even better frequency accuracy and phase noise specification. For example, a Stanford Research Instruments PRS10 rubidium frequency standard could be used. This reference has a frequency accuracy of 10^{-12} and remarkable SSB phase noise of -155 dBc/Hz at 30 kHz offset from 10 MHz. Use of such a reference on the SMC100A in our experiments would eliminate the signal generators as significant sources of noise, and the intrinsic device noise physics would become dominant. It would also eliminate the frequency drift we observed in our NEMS experiments during long-term measurements over hours to days.

Another important source of noise in our experiments was the method of demodulation used with the lock-in amplifiers. Referring to Fig. 2.7, in our setup we set the reference frequency of the lock-in amplifier by connecting the TTL synchronisation output from the signal generator to the reference input of the lock-in. However doing this introduces phase noise into the demodulated signal, as noise sources perturb the optical signal on its route after reflection off the modulated

NEMS resonator, to the photodetector and its amplifier circuitry, and finally to the lock-in signal input. There is also a phase offset and phase noise between the signal generator output and TTL output, inherent to the circuitry of the signal generator itself. To overcome this, we propose the following. The Stanford Research Instruments SR844 lock-in amplifier used in our experiments has sophisticated amplification and signal conditioning circuitry on the reference input. This allows it to lock-on and generate a reference from the carrier on the signal from the photodetector. Thus instead of connecting the synchronisation output from the signal generator to the reference input, we simply split off the signal input to the lock-in amplifier and connect that to the reference input instead. Although we did not get an opportunity to test this configuration with the NEMS devices, we did try it in a different application, as explained in Appendix D. In that experiment the configuration eliminated the need for 10x averaging and the signal could be transduced in real time.

Finally, another possible path to reducing noise is to use a lock-in amplifier of improved design. The SRS SR844 we used in our experiments was designed 25 years ago. Since the revolution in telecommunications and cellular technology used in microwave communications, the cellular phones ubiquitous among us contain lock-in amplifiers and phase-locked loop circuitry far more advanced than in our instrument. Recently available commercial lock-amplifiers, such as the Zurich Instruments DC-600 MHz UHFLI, have much superior specification. The UHFLI has a reference phase noise of just $1\mu^\circ$, and a frequency accuracy of 10^{-12} due to the built in rubidium standard. In contrast, the SR844 has a reference phase noise of $5000\mu^\circ$ and a frequency accuracy of just 10^{-3}.

2.6 Toward Nanophotonic Integration

In our present scheme, two physical wires are required in addition to the optical interconnect for communication with the external environment; a ground terminal and a path carrying the multi-frequency AC drive generated externally. However recent progress with optically driven NEMS [63], on-chip waveguide fabrication and evanescent wave optical detection schemes [64] raise the intriguing possibility that these techniques could be integrated with piezoelectric NEMS. This might enable realisation of an all-optical nanomechanical Fourier transform that would be a step towards wireless interfacing with nanoscale systems, as both actuation and detection would be optical, whereas local voltage transduction would be piezoelectric. We develop this idea in Chapter 4 of this thesis. Moreover, electronic circuits mimicking the function of the human cochlea could provide a compact method of demodulating signals from our system [65]. Such a vision, where nanoelectronics becomes untethered from the constraints of physical interconnection, might make feasible the development of miniature, mobile nanoscale integrated sensors that could be positioned at any place in the human body, extract vast amounts of information from the brain and that could augment study of the complex biochemical networks within and between cells in real-time.

2.7 Experimental Methods

2.7.1 Nanofabrication

The heterostructure used to fabricate the piezoelectric resonators consisted of (bottom to top, with doping concentration in cm^{-3} in parentheses): a 625 μm highly *p*-doped GaAs substrate, 550 nm *p*-$\mathrm{Al}_{0.8}\mathrm{Ga}_{0.2}\mathrm{As}$ (10^{18}) sacrificial layer, followed by the *pin* structure of 50 nm *p*-GaAs (10^{18}), 50 nm *i*-GaAs ($\sim 5 \times 10^{15}$) and 100 nm *n*-GaAs (10^{19}). All NEMS were aligned along the [110] axis, for which the transverse piezoelectric constant, $d_{31} \approx 1.4\ \mathrm{pm/V}$.

The GaAs wafer was diced into $4\ \mathrm{mm}\ \times\ 5\ \mathrm{mm}$ chips, with the greater length aligned to the [110] axis. This aided in identification during processing. Fig. 2.19 outlines the process flow for fabrication of the chips. The diced chip was cleaned through sonication for 5 minutes in each of acetone, isopropyl alcohol (IPA) and deionised (DI) water. Microchem polymethylmethacrylate (PMMA) 950 A4 electron beam resist was spun on the chip (3000 rpm, 1 min) and baked on a hotplate at 180°C for 90 seconds. Electron beam lithography was then employed to pattern the wirebonding pads and anchor pads of the NEMS devices at a dose of 1300 μC/cm^2 at 100 kV, using a beam current of 10 nA. After the write, the chip was postbaked at 100°C for 90 seconds. Development was then performed in 1:3 methyl isobutyl ketone (MIBK) : IPA with agitation for 30 seconds. It was then placed in a stop bath of IPA for a further 30 seconds, and dried with nitrogen.

A descum was then performed on the chip through a short oxygen plasma etch with a reactive ion etcher (RIE) at 50-50-50 conditions (see Appendix A), for 20 seconds. An electron beam vacuum deposition system was then used to deposit (bottom to top) 10 nm Cr, 70 nm Au, 70 nm Ti. Liftoff was performed in acetone. It was important to use chromium as the adhesion layer instead of titanium in order to prevent delamination of the metallisation during etching in hydrofluoric acid (HF).

The chip was cleaned in oxygen plasma in the RIE at 20-20-80 conditions for 5 minutes, and PMMA resist was once again spun on and prepared as before. The electron beam pattern generator was aligned to the specially defined alignment marks on the chip (20 $\mu\mathrm{m}\ \times\ 20\ \mu\mathrm{m}$ squares) and this

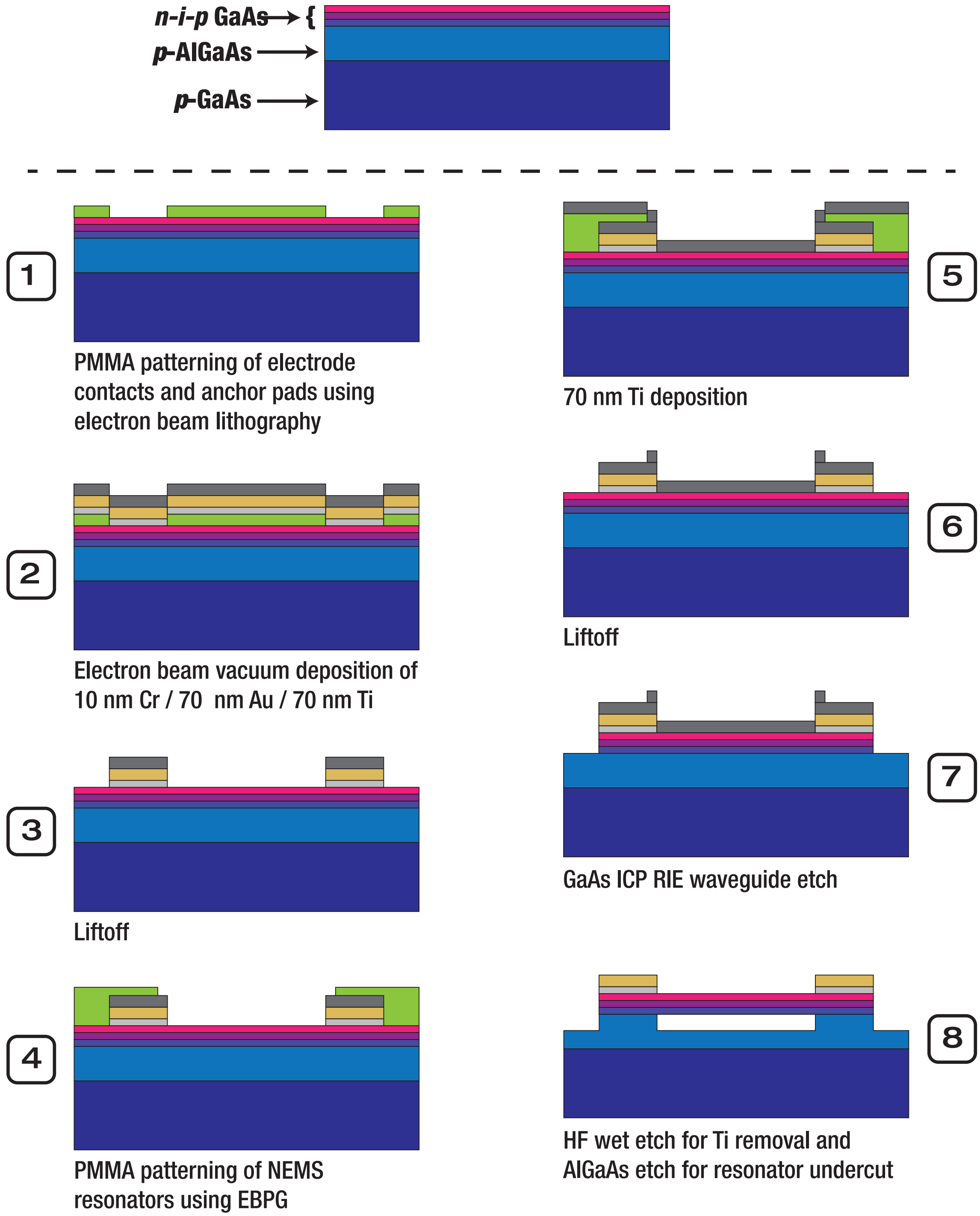


Figure 2.19: **Process flow for fabrication of piezoelectric NEMS devices.** The top electrode of the completed device is the cathode of the *pin* diode, whilst the handle of the chip forms the anode.

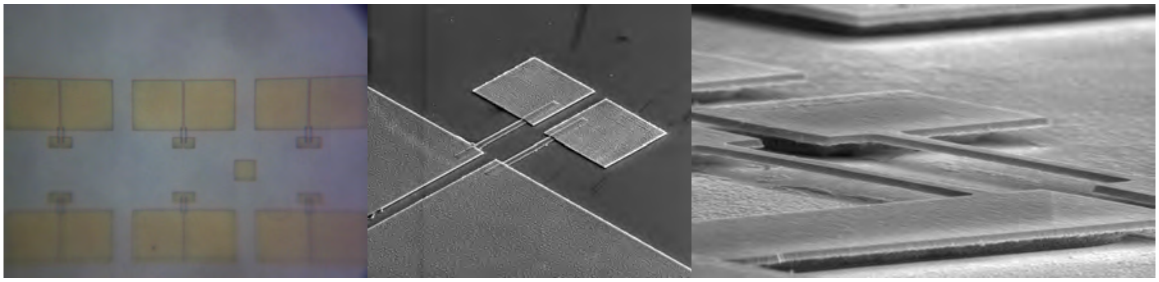


Figure 2.20: **Piezoelectric NEMS fabrication.** First is shown the result of fabrication subsequent to electron beam patterning and metal deposition of the pads and titanium etch mask for the NEMS resonators. The second image shows the result subsequent to GaAs ICP dry etching. The final image shows the result subsequent to undercutting in diluted and buffered hydrofluoric acid.

time the NEMS resonators were patterned over the wirebonding pads and anchor plates. The beam current used was 1 nA. Postbake, development and descum was done as before, and then 70 nm of Ti was deposited. The first picture in Fig. 2.20 illustrates the result subsequent to liftoff.

The chip was then cemented to a 6” wafer using a thin layer of Santovac 5 vacuum pump oil. The entire assembly was mounted in an Oxford Instruments System 100 ICP 380 and a GaAs waveguide etch recipe was run for 1 minute (see Appendix A). The chip was then carefully removed and cleaned in acetone / IPA / DI. The second photograph in Fig. 2.20 illustrates the outcome.

A 7:1 solution was then prepared of DI water to 6:1 buffered oxide etch (BOI). The chip was placed in the solution and gently agitated for 4 minutes to etch away the $\mathrm{Al}_{0.8}\mathrm{Ga}_{0.2}\mathrm{As}$ and undercut the suspended NEMS resonators. The chip was then removed and rinsed in DI water, and dried in nitrogen. The final picture in Fig. 2.20 depicts the end result.

A 68-pin ceramic pin grid array (PGA) package, with a $[0.425\,\mathrm{in}]^2$ cavity (Addison Engineering, Inc., San Jose, CA), was used to mount the chip using a small amount of silver conductive epoxy (MG Chemicals, 8330S). On setting of the epoxy, the chip was then wirebonded to the package using gold leads.

2.7.2 Apparatus & Measurement

Fig. 2.21 depicts the experimental apparatus used to obtain the measurements in this chapter. Compare the components to Fig. 2.7; a Thor Labs 904 nm, 10 mW laser diode (L904P010) was driven using a Newport 501 laser diode driver. The wavelength of the laser exceeded the photonic band gap wavelength of GaAs in order to minimize heating and surface charge excitation effects. Despite this precaution, some DC photoelectric effects were nevertheless observed, which necessitated the

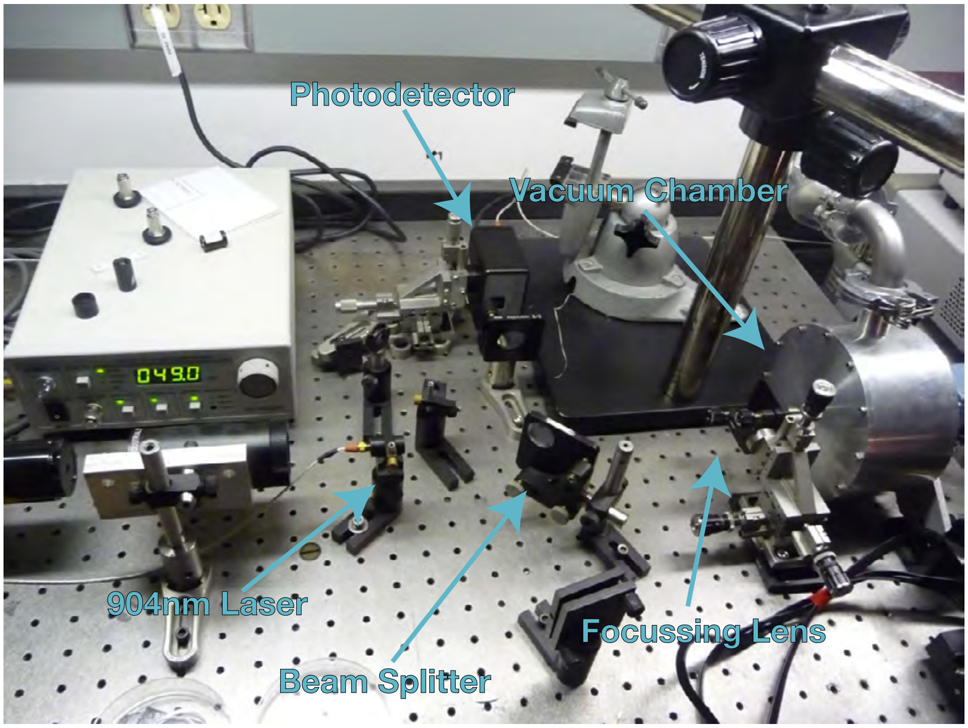


Figure 2.21: **Experimental apparatus.** Shown is the actual equipment setup used in acquiring the measurements made in this chapter. Compare to the schematic in Fig. 2.7.

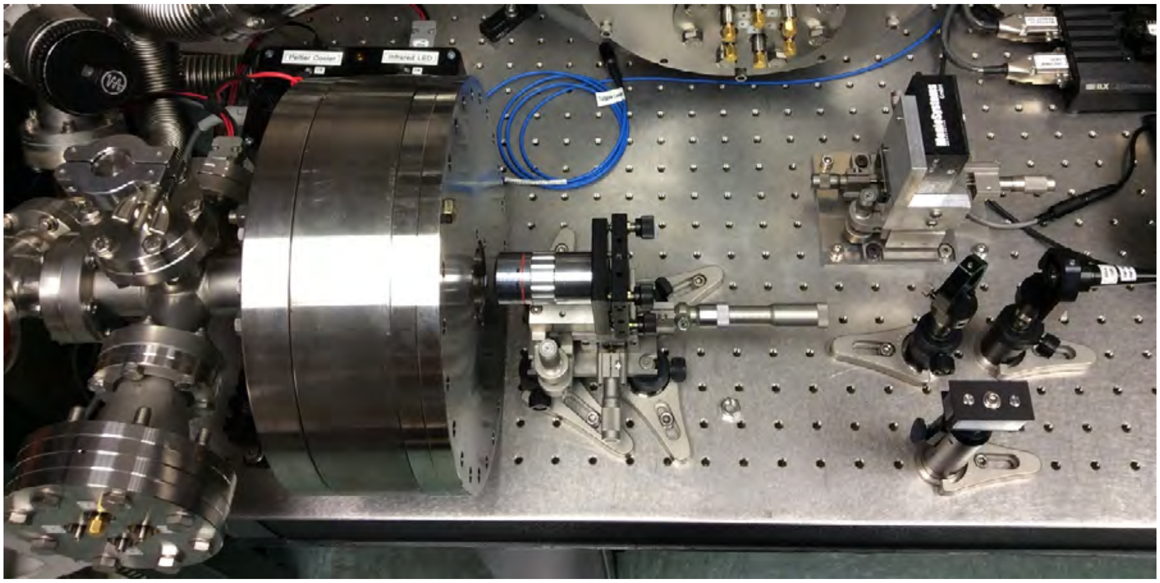


Figure 2.22: **Improved experimental apparatus.** A redesigned and much improved vacuum chamber is coupled with better electrical shielding, an improved laser, better detection and improved modulation/demodulation electronics.

placement of a 100 nF DC blocking capacitor between the NEMS device and voltage bias source. A Thor Labs LT110P-B collimation tube was used to to collimate the laser diode. The focal length of the lens in the tube was $f = 6.24$ mm, and the principal plane of the lens was 8.7 mm from the diode.

A microscope objective of numerical aperture NA=0.4 and focal length $f = 40$ mm was used to focus the laser onto the NEMS device. This was mounted to a three-axis micromanipulator, which allowed precise alignment of the beam. Alignment was made by attaching an ohmmeter to the NEMS device and looking at the change in resistance. With no laser impinging on the device, the resistance measured was ~15 MΩ. As the beam drew closer to the device, the resistance would drop until it became negative and minimal at optimum alignment. This is due to carrier excitation in the NEMS diode.

A Thor Labs BSW26 non-polarising beam splitter (350-100 nm) was used to split the beam and divert the reflected laser light off the NEMS device onto the silicon *pin* photodetector. The photodetector used was a Newport 1801-FS (320-1000 nm, 0-125 MHz).

Lock-in measurements were made with Stanford Research Systems SR844 RF lock-in amplifiers (20 kHz-200 MHz). All data were acquired at a sampling rate of 30 kHz using a National Instruments PCI-6259 data acquisition (DAQ) card. Raw signals from the photodetector and extracellular amplifier were found to contain environmental distortion effects in the form of 60 Hz noise and its multiples extending up to $\sim$ 300 Hz. It was thus necessary to apply an 8-pole high-pass Butterworth filter to our acquired data; f_{3dB} was set to 300 Hz for the data in Figs. 2.8-2.12. This filter was constructed by the author. Moreover a low-pass filter (Stanford Research Systems SR560 preamplifier/filter) was applied to highlight the physiological frequency region (f_{3dB} = 10,000 Hz). Identical filter settings were always used on both the photodetector and bias voltage signals. A meticulous grounding scheme was also implemented to minimise 60 Hz noise. It is believed that closer integration of the neural probe with the nanomechanical resonator will substantially reduce 60 Hz artifacts, and enable measurement of both extracellular action potentials and local field potentials.

The optimal value for f_{carrier} (see Fig. 2.6) was found by applying a 1 kHz sinusoidal bias of known amplitude across the NEMS device terminals (all potentials are with respect to the *n*-doped terminal), and manually sweeping the drive frequency until the largest demodulated signal was observed on an oscilloscope. The demodulated signal amplitude could then be calibrated to the known value of V_{bias}. The nanomechanical displacement in the inset of Fig. 2.10 was estimated from the expression for critical amplitude at the onset of nonlinear bistability [66]:

$$x_{crit} \approx \frac{2.05h}{Q^{1/2}} \, , \tag{2.7}$$

where h is device thickness and Q is resonator quality factor.

In view of the noise sources uncovered after the experiments were completed (as described in Chapter 2.5), and the importance of this multiplexing technique being a real-time transduction system if used for neural recording, it was decided to build an entirely new experiment to overcome the limitations in the original setup and definitively demonstrate real-time operation. This setup is illustrated in Fig. 2.22.

To further suppress carrier excitation during optical interferometry, a Thor Labs 980 nm, 10 mW laser (L980P010) was utilised. A Menlo Systems FPD510-FV *pin* detector, similar in specification to the original photodetector, was used. Very importantly, Rohde & Schwarz SMC100A signal generators, with the optional high stability OCXO installed, were used to drive the NEMS devices. The Stanford Research SR844 lock-in amplifiers were reconfigured to use the improved reference carrier locking scheme described in Chapter 2.5. A pair of Stanford Research Systems SR650 programmable bandpass filters were used to filter the lock-in amplifier outputs before data acquisition. A Mitutoyo 5x microscope objective (NA $= 0.4$, $f = 40$ mm) was used to focus the laser light, and it was calculated to generate a focussed spot size of $d = 7.84$ μm with the LT110P-B laser collimation tube.

Another important modification was the redesigned vacuum chamber used for the experiments, illustrated in Fig. 2.23. The basic components of this chamber were constructed out of 10” conflat CF hardware. A 2” thick double-faced 10” CF flange formed the depth of the vacuum chamber. The rear flange was a zero-length reducer plate that converted to a 2-3/4” CF flange, to which was attached a four-way flange for attachment of electrical feedthroughs and vacuum connection. The front flange was a 1” thick blank that had a 1” hole machined in the middle. To the bore in the front flange, a sapphire viewport was installed flush to the flange and was TIG welded to form a vacuum seal. A sapphire viewport is important to use as it is transparent to the infrared laser light used in our interferometry. A 4-1/2” CF flange was machined to accept four weld-in SMA feedthroughs, which were carefully TIG welded in place. A second flange was fabricated containing a DB9 feedthrough. Both these flanges connected to the rear four-way, and the third port was affixed to a doubled-up valve which controlled the vacuum connection to the pump. Two valves were used to prevent

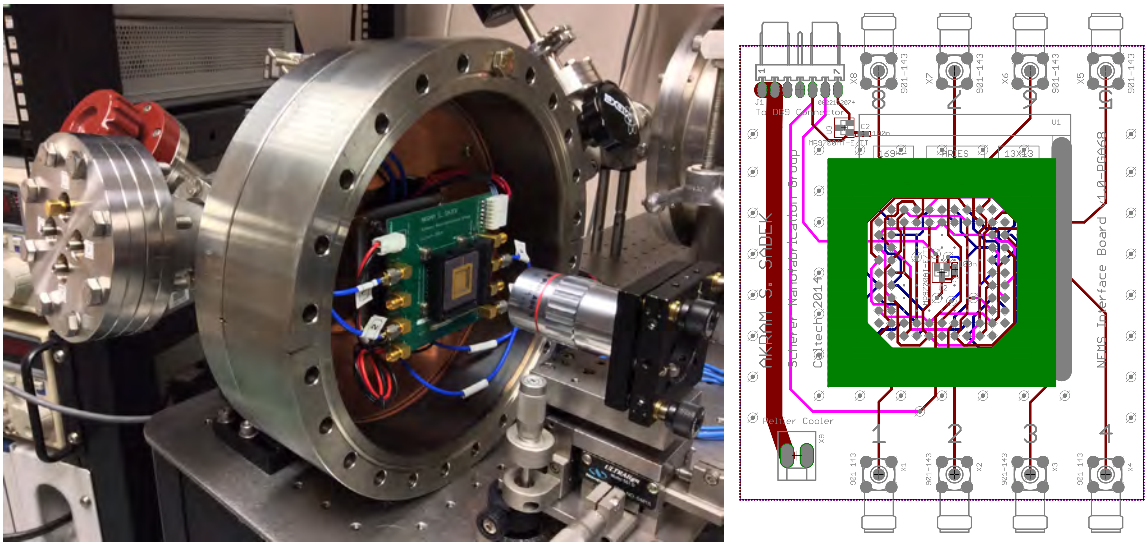


Figure 2.23: **Improved vacuum chamber.** Shown are more details of the redesigned vacuum chamber as well as the custom PCB used to mount the NEMS chip in its 68-pin PGA carrier. SMA connections are used for signal routing and the entire board is mounted to a Peltier cooler for temperature control using a specially machined copper gasket. There are also on-board temperature sensors to monitor the chip temperature during operation. A vacuum of at least 10^{-4} Torr is possible with a Welch 1397 roughing pump.

contamination to the chamber from the pump during startup and shutdown. The vacuum chamber constructed here was capable of much better vacuum than in the original experimental setup. In this particular vacuum system, the only limiting factor was the absence of a turbopump, and instead the Welch 1397 DuoSeal pump was solely relied upon, and could achieve 10^{-4} Torr.

To the rear flange, a special mounting fixture for the NEMS chip was designed and fabricated. This consisted of a custom PCB which was mounted on a high efficiency Peltier cooler. The rear section of the cooler was mounted on a specially machined copper plate, for efficient dissipation of the waste heat. The Peltier cooler was powered and controlled via the DB9 feedthrough, and on-board temperature sensors were also routed through this feedthrough. The purpose of the Peltier cooler was to control the temperature of the NEMS chip and investigate the effect of thermal noise on the total measured noise of the system. The cooler easily allowed the chip temperature to be varied from -10°C to 100°C using a 12 V, 10 A programmable power supply. The PCB mount was very well grounded for electrical shielding as well as efficient cooling, and connections from the 68-pin PGA chip were routed out via SMA connections, which facilitated excellent shielding of the signal. The entire chamber was also grounded to the PCB board, making an excellent Faraday cage. Easy chip loading/unloading was facilitated by the 68-pin PGA ZIF socket on the board.

As of the time of writing this thesis, the new setup has not been used to acquire data, but this is only a week of work away from being achieved.

2.7.3 Electrophysiology

Silicon neural probes with 16-channel microelectrode arrays were fabricated as described on Chapter 3. Nominally, the $10\ \mu\mathrm{m} \times 10\ \mu\mathrm{m}$ microelectrode sensors have an impedance of ~ 1 MΩ at 1 kHz. Use of the microelectrodes unmodified as such in our experiments would have severely attenuated the signals detected due to the RC effects of the large ~ 1 nF parasitic capacitance between the bias and reference terminal, as described in Chapter 2.5. The majority of this capacitance originated at the wirebond contact region for the NEMS device. The recording site impedance was thus reduced via gold electrodeposition to $\sim 30\ \mathrm{k}\Omega$ at 1 kHz, in order to minimise filtering. Electrodeposition produces a three-dimensional highly structured gold surface, and reduces impedance by vastly increasing the effective surface area of the electrode.

An 16-channel extracellular amplifier was custom built for the NEMS experiments, using an Intan Technologies RH1016 amplifier array (see Fig. 2.24). This amplifier allowed specific electrode channels to be switched into the NEMS circuit once a suitable spike was obtained during *in vivo* measurement. Software was written in Labview to control the amplifier via the PCI-6259 DAQ card and trigger the 16-channel multiplexing circuit. The instrument was carefully shielded, and we could routinely achieve a voltage noise of 1.5 $\mu\mathrm{V}_{\mathrm{rms}}$ on grounding the inputs at the probe headstage.

The locust thoracic ganglion was prepared under the protocol described in Ref. [67], and was able to provide stable neuronal recordings for several hours. The first image in Fig. 2.25 depicts the thoracic ganglia just prior to extraction. The isolated ganglia were bathed in 100 μM pilocarpine solution to elicit rhythmic firing of motor neurons. An Ag/AgCl electrode served as the electrical potential reference. The reference electrode was prepared by thoroughly cleaning a silver wire, placing it in 200 mM saline solution, and passing a current of 1 mA through it, with the silver wire connected to the positive electrode. This was done until the wire was a dull grey in colour.

The assembled neural probe was mounted onto a micromanipulator and inserted into the thoracic ganglion until a stable spiking unit was located, as depicted in the second image in Fig. 2.25. After data acquisition, the signal from the extracellular CMOS amplifier was filtered in the pass band of 400 to 4,000 Hz and fed to a spike sorting routine written for Matlab [68]. Spikes were identified based on an amplitude threshold method, and clustered according to the Euclidean distance between their waveforms. After identifying a putative single unit's spike times, they were

Figure 2.24: **Extracellular amplifier.** The amplifier was custom built to facilitate experiments with the NEMS devices, and attention was paid to proper shielding. Based on the Intan RH1016 amplifier, recording channels of interest from the neural probe could be selectively switched into each of the two independent NEMS resonators used in our multiplexing experiments.

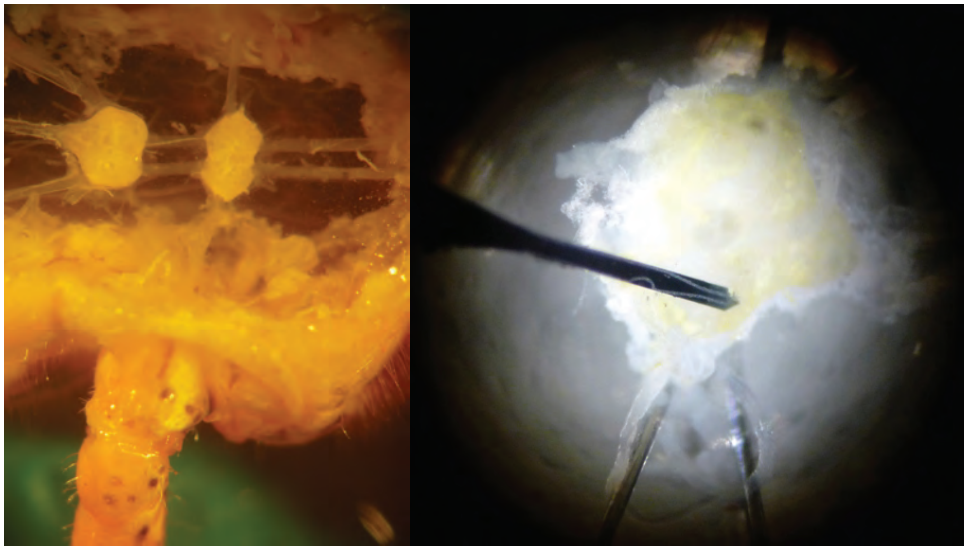


Figure 2.25: **Locust thoracic ganglion preparation and measurement.** Shown are the thoracic ganglia just prior to extrication, and during measurement with the neural probe whilst bathed in PBS solution with 100 μM pilocarpine.

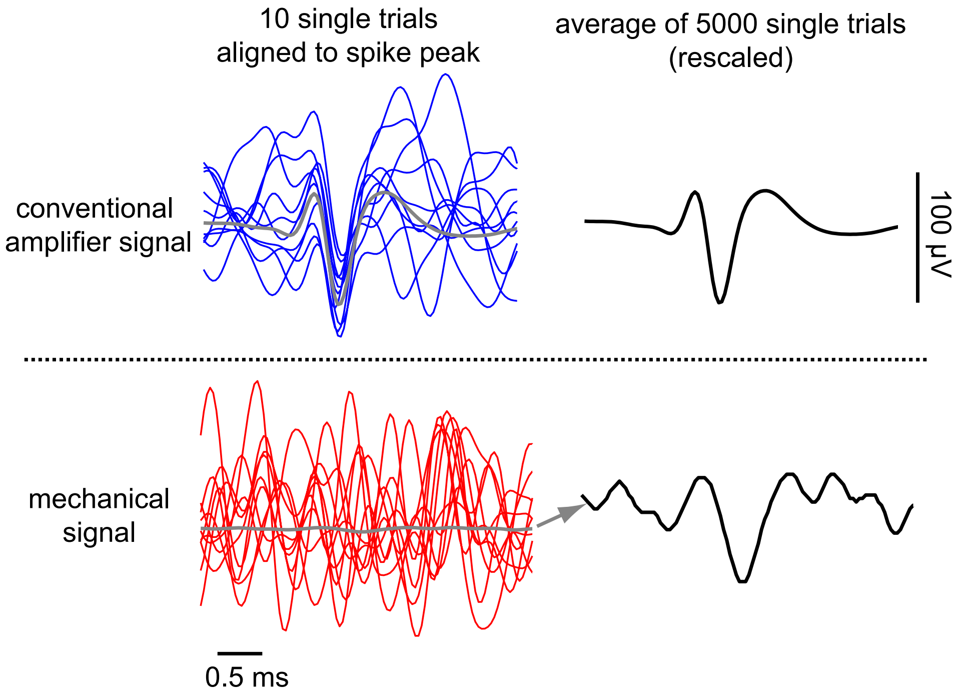


Figure 2.26: **Waveform sorting and averaging to demonstrate nanomechanical transduction of extracellular neuronal spikes.** A spike sorting program was used to extract and align putative spikes using the data from the CMOS extracellular amplifier. The concurrently collected NEMS data was averaged over the same time index to successfully show transduction. This data set is a separate run to that shown in Fig. 2.14.

used to extract the waveforms from the concurrently recorded NEMS signal at the photodetector. Five thousand waveforms were superposed and averaged to yield the results in Figs. 2.14 and 2.26. Before generating these results, the original data were re-filtered to reduce noise artifacts, in the pass band of 300 to 2,500 Hz. The action potential signals were attenuated and eventually abolished upon retracting the neural probe from its nominal recording location, confirming the neuronal origin of the spikes.

Measurements for the plot in Fig. 2.15 were generated by passing a 400 μV_{pp} 1 kHz sinusoidal bias signal into an isotonic buffered saline bath in which the probe was immersed. The bias signal was applied between two Ag/AgCl electrodes, one of which was connected to the NEMS and CMOS amplifier ground reference. Recorded signals were filtered from 350 to 3,000 Hz, and segments spanning a few milliseconds were aligned and averaged.

2.8 Nanomechanical Wiring for Large-Scale Electrophysiology

Our primary goal in developing the technology for wiring sensors via nanomechanical resonators was to enable monitoring of electrophysiological activity across large areas of the brain with high spatial resolution [15]. Ideally, we would like a tool such as this to eventually enable recording of the spiking activity from every neuron in a mammalian brain.

In a typical neural probe interface, an array of microscale electrodes is positioned along an implantable shaft (see Chapter 3). These electrodes measure small extracellular potentials, on the order of tens to hundreds of microvolts. As long as the electrodes are sufficiently small a neural probe can provide information on the firing activity of several single neurons in parallel. The development of more complex neural probes, that can monitor larger areas of the brain and at higher spatial resolution, would provide novel opportunities in the study of the brain, but requires a tradeoff between two important sensor design considerations. On one hand it is desirable to maximize the area covered by the array by employing several hundred to thousands of closely spaced, individually addressable recording sites. On the other hand there is a need to miniaturize the neural probe in order to minimize damage to the surrounding tissue. As with other classes of sensor, as the number of recording channels is scaled up, the primary size-determining factor becomes the interconnecting wiring between the electrodes and macroscopic world. Moreover, a reduction in wiring width and spacing may not be an effective solution to this problem because the high impedance of these wires would exacerbate signal crosstalk and attenuation [18]. Furthermore, fabricating nanoscale wiring below $\sim$100 nm with lithographic techniques makes the wires unreliable, especially on implantation. This is as the flexing and thermal expansion cause cracking of the traces. Thus a technique allowing multiplexing of signals in the immediate vicinity of the recording sites would circumvent the need for using long interconnections and enable development of much higher density and less invasive neural interfaces.

A new type of implantable neural probe containing a planar microelectrode array could be fabricated with the sensors multiplexed onto a single output, as in Fig. 2.27. The electrical potential at each electrode would be altered by local extracellular fields in the brain. However, rather than routing signals from each recording site with a wire spanning the length of the probe, each channel would be locally coupled to a low power piezoelectric NEMS device that resonates at a unique frequency. As illustrated in Fig. 2.28, the nanomechanical components would be fabricated directly beneath the corresponding recording site and be contained in a hermetically sealed vacuum environ-

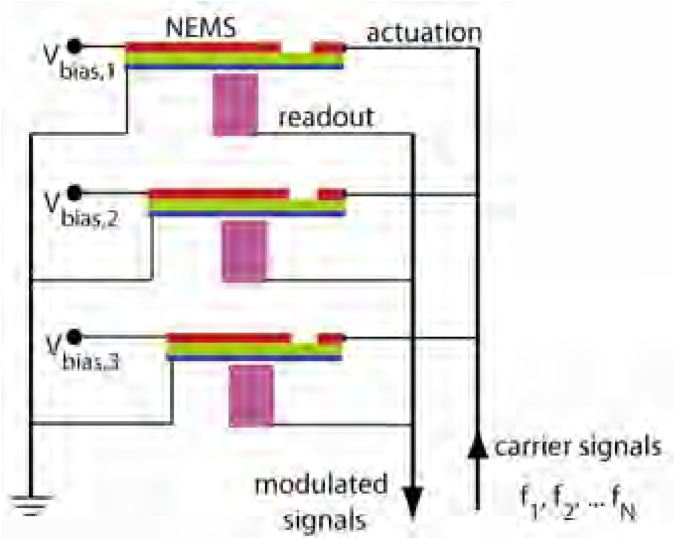


Figure 2.27: **Simplified schematic for multiplexing analogue signals with a nanomechanical Fourier transform.** Most of the wires in a neural probe could be eliminated if electrical potentials were transferred into the frequency domain, which would enable mixing of multiple recording channels on a common output. Here, each NEMS resonator has a different fundamental frequency and is connected to a separately addressable recording site. In addition, all resonators share a common actuation, reference and displacement readout channel. Nanomechanical displacements could be measured via optical or electrical transduction.

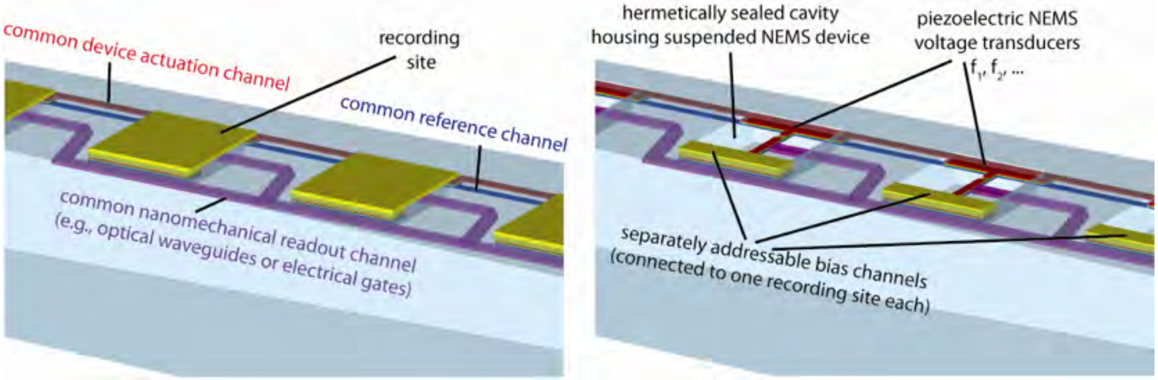


Figure 2.28: **Proposed nanomechanical wiring configuration for neural probes.** In the first figure, the voltage-sensitive, piezoelectrically coupled resonator (not shown) lies directly beneath the rectangular recording site electrode. In the second figure, the recording electrodes have been removed to reveal the cavity that would house the NEMS devices. Nanomechanical oscillation does not exceed several tens of nanometers in amplitude, thus the cavity need not be much thicker than the NEMS device thickness.

ment achieved through a wafer bonding process. The operation of a large set of parallel-connected NEMS devices would require four connections per resonator: the locally connected bias voltage from the extracellular electrode, plus a common multi-frequency drive actuation path, ground terminal, and nanomechanical readout path.

The quality factor of a resonator is given by $Q = f_k/\Delta f$, where f_k is the centre frequency and Δf is the resonance width (full width at half-maximum). Our NEMS resonators could be fabricated with a Q of at least 2000 if done meticulously. Quality factors substantially greater than this are not possible in principle with a piezoelectric NEMS device; as the piezoelectric device contains charge carriers and is effectively a parallel array of capacitors, on mechanical resonance, the oscillation of the charge carriers between the capacitors effectively radiates electromagnetic energy and prevents enough energy being stored in the resonator to achieve a higher Q [69]. This is in fact the dominant mode of damping with this class of device.

If we separate each of our resonators in frequency space by the resonance width Δf, then the neighbouring frequency is given by:

$$f_{k-1} = f_k - \frac{f_k}{Q} \,. \tag{2.8}$$

If we assume an operating NEMS frequency band of 10 MHz to 250 MHz and $Q = 2000$, this computes to a conservative 6500 data channels transmitted off a neural probe in a parallel manner. Even a 1000 site array of 10 $\mu\mathrm{m}^2$ microelectrodes arranged column wise along a single neural probe shaft with 10 μm spacing, could measure activity across the entire depth of a cortical column with 10 μm spatial resolution. The number of channels can be even higher if the bandwidth is expanded to include microwave frequencies [55]. Silicon microelectrode arrays using conventional wiring currently only offer 50-100 recording sites per shaft [17, 19]. We therefore expect to obtain two orders of magnitude increase in recording site density, while allowing arbitrary reduction in neural probe width.

Although neural probe multiplexing on the order of $\sim$10,000 channels through four connections would be a great advance for neuroscience, it is insufficient in itself to answer the challenge of recording the spiking activity from an entire brain concurrently in minimally invasive fashion. Is it possible to go further? Can we somehow harness the enormous advances in modern optical telecommunications and apply them here? Furthermore, the scheme described in Fig. 2.28 is nebulous with regards precisely how an integrated optical waveguide could interface with the NEMS devices to detect their nanomechanical motion. The answer to these questions lies in the synthesis of the ideas in this chapter with nanophotonics. We embark on this new quest in Chapter 4, where we find that in fact multiplexing on the order of a *billion* channels is feasible through a single fibre-optic communications line. This is sufficient to record the spiking activity from the entire primate brain at single-cell resolution.

Chapter 3

Micro- and Nano-Fabricated Neural Probes

> Too many men work on parts of things.
> Doing a job to completion, satisfies me.
>
> *Richard Proenneke,* Alone in the Wilderness *(1968)*

3.1 MEMS Neural Probes

Extracellular recording is one of the most important tools in the neuroscientists arsenal. As electrodes used in this manner can record single-unit activity from multiple neurons concurrently, it has the potential to be extended to map the activity of the brain at a large scale [15]. The technology behind this technique has come a long way since Hubel's pioneering experiments in the 1950's [70]. The first extracellular microelectrode recording arrays produced by lithography were actually made at Intel in the late 1960's, and the first neural probes fabricated with the techniques of microelectromechanical systems (MEMS) were reported in 1970 [71]. A large body of work has been done between then and now, making the probes smaller and more dense [17, 18], fabricating them with biocompatible materials, and integrating microelectronics with the probes to minimise external wiring [19] and make them wireless [72]. In the past few years, MEMS neural probes have become commercially available to the neuroscience community, so they are no longer restricted to the cohort of researchers with the ability to fabricate them.

Our purpose in fabricating MEMS neural probes were many fold. First, neural probes were required to interface the NEMS devices in our experiments with *in vivo* neural recordings. Although the nanomechanical resonators we present in Chapters 2 & 4 are still experimental and not yet in a form that can be directly implanted, taking this next step to create an integrated probe will in fact combine the fabrication techniques used in MEMS neural probes with those we used in our NEMS and nanophotonic resonators. So designing and fabricating the neural probes presented in this Chapter was a great learning experience, and in fact, many of the fabrication techniques worked out here were used in those novel devices as well. Aside from this, the design and fabrication used for the neural probes we introduce here is actually unique, and allows fabrication of substantially higher quality probes than reported in the literature. Another important aspect is that fabrication of MEMS neural probes at the moment is very much an art; the literature on the subject does not detail the precise fabrication methods and protocols to obtain the devices reported. Thus another goal of this work was to outline in detail the methods used to create such probes, and prevent future researchers in the field from having to re-invent the relevant techniques from scratch.

Please note that detailed recipes for different aspects of the fabrication are detailed in Appendix A.

3.2 16-Channel Probes Fabricated by Photolithography

The major advantage to photolithography as compared to a serial process like electron beam lithography is speed; whereas an electron beam pattern generator might require up to several days to pattern an entire 4” wafer, photolithography can do the patterning in seconds. For example, the wafer design presented here can produce up to 172 probes in one fabrication run, with good yield. That is enough probes for years of experiments. I consider the information in this chapter to be an important resource for anyone in the neuroscience field looking to start fabricating their own MEMS probes, who perhaps wish to do so for some custom application that a commercial vendor cannot oblige. It is also a great starting point from which to learn and then go on to produce something entirely new, as all the experience detailed will prevent one from having to needlessly learn the pitfalls of fabrication by trial and error. The fabrication of MEMS probes draws on many different techniques in micro and nanofabrication, and it is also an excellent way to learn the field.

3.2.1 Design

The goal of the design here was to create 16-channel neural probes to be used for acute *in vivo* experiments. The probe had to be easy to mount and remove from the headstage, be easy to handle,

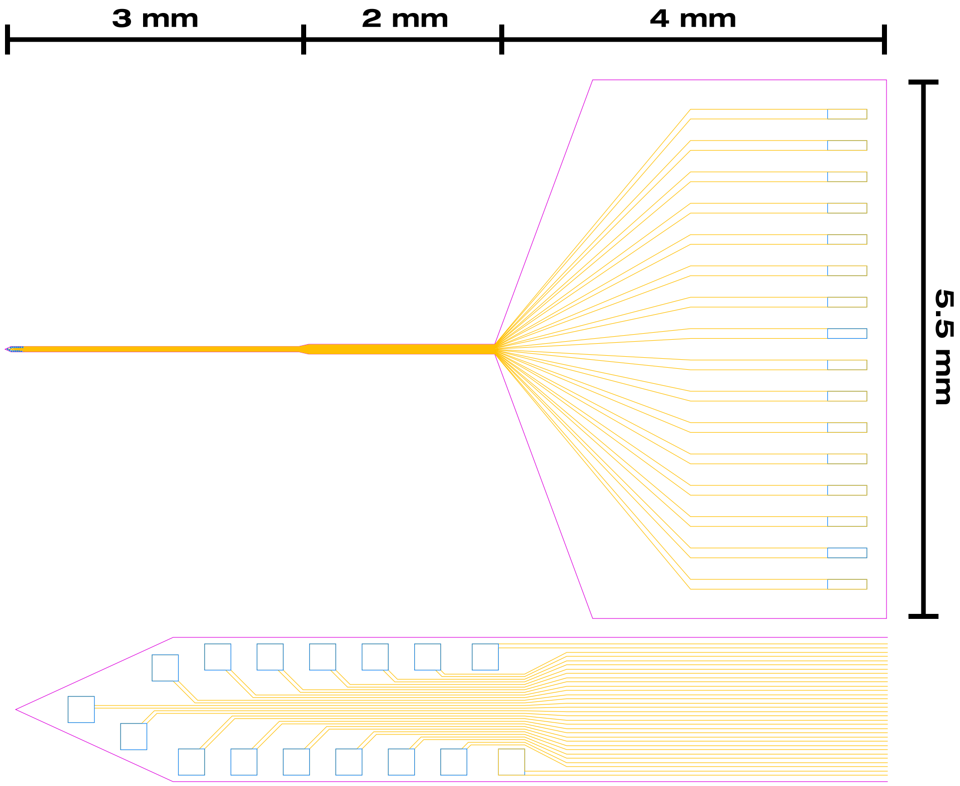


Figure 3.1: **16-Channel neural probe design.** The probe shaft was set at 5 mm length to facilitate penetration through the mouse and rat brain. The wirebond regions are $100\ \mu\mathrm{m} \times 400\ \mu\mathrm{m}$ in area, spaced 200 μm apart, and could be much reduced in size to adapt the design for chronic implantation. The inset shows a blown up view of the probe tip. The probe shaft at the narrowest end is 55 μm in diameter, and the microelectrodes are $10\ \mu\mathrm{m} \times 10\ \mu\mathrm{m}$ and spaced 10 μm apart. Gold represents the metallisation mask, blue the windows through insulation, and purple the probe substrate.

and be robust enough for repeated cleaning and reuse. At the same time, it had to have a minimal cross section to minimise tissue damage during recording, and optimise the quality of single-units detected. The probes were intended to be used with the extracellular amplifiers and headstages presented in Chapters 2 and 4 and detailed in Appendix B. These have 18-DIP zero-insertion force sockets, and so our probes are designed to be coupled with 1.5 mm thick PCB interface boards mounted with 18-DIP pins.

The probes themselves though can be easily adapted for chronic implantation, especially in view of their high quality fabrication, with their triple insulating layers of silicon oxide, silicon nitride and parylene. This is a matter of altering the wirebonding region of the probe design and switching to a flex PCB interface board with miniature Omnetics connectors. Similar probes reported in the literature are only coated with parylene, as this makes the fabrication much more straightforward

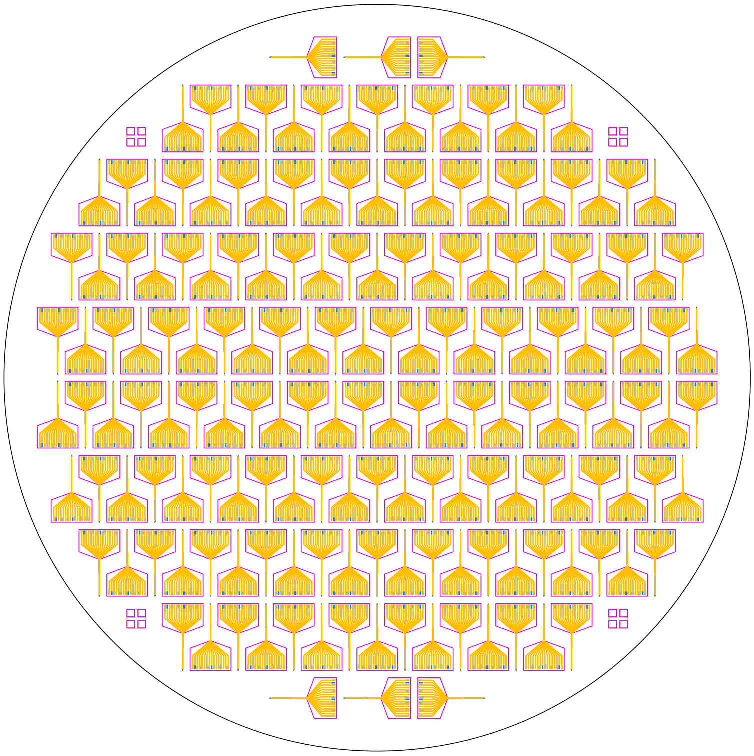


Figure 3.2: **4" Wafer-scale design for 16-channel MEMS probes.** With this design 172 probes are packed into a 4" wafer, providing plenty of probes even in the case of poor fabrication yield. All three photolithographic masks made to this design were implemented in 6" format for ease in alignment. The metallisation mask was brightfield and the masks for the electrode/wirebonding windows and probe shafts were darkfield.

[18, 19]. The drawback with this is long term chronic implantation is impacted; the parylene easily breaks down, and the K^{+} and OH^{-} species present in the extracellular fluid eventually work their way through to the silicon shaft and etch it away. In effect, the parylene is being used for the wrong reason, to physically insulate the probe, instead of to impart improved biocompatibility in neural tissue. The basic fabrication design presented here of: silicon shaft, thermal oxide insulation, metallisation, silicon oxide coating, silicon nitride coating, parylene coating, makes the probes highly resistant to this 'dissolution' effect during long term implantation.

Fig. 3.1 illustrates our basic probe design. The probe length of 5 mm allows full access to the mouse brain with stereotactic *in vivo* recording. The microelectrodes at the tip are $10\ \mu\mathrm{m} \times 10\ \mu\mathrm{m}$, and are spaced 10 μm apart from each other. The finest interconnects are just 1 μm in thickness with a spacing of 1 μm. This is almost at the limit of what photolithography can achieve, the cutoff nominally being 700-800 nm with conventional UV lithography using a mercury lamp. The probe width is only 55 μm, and it is designed to have a thickness of 25 μm. Note that the windows to the microelectrodes and wirebond pads are set slightly smaller in size to their dimensions. This is done so that the overlaying insulation layers overlap the pads slightly and keep them stable, as well as to prevent ionic species in the extracellular fluid from easily accessing the underlying surface of the probe.

Clearly, the wirebonding and PCB mounting area could be reduced in dimension if fabrication of a long-term implantable probe was the goal. It was kept large for practical reasons, primarily to facilitate handling. Since the probe substrates are mounted on large PCB's, there was no real requirement to make this area smaller.

As depicted in Fig. 3.2, probes could be packed on to a 4" wafer with high density. High quality alignment marks were designed to allow easy and precise alignment between the metallisation layer, electrode/wirebond exposure layer and probe substrate definition layer during the various fabrication steps. The three separate masks were ordered in 6" chromium on quartz format, 0.09" thickness from Photo Sciences, Inc (Torrance, CA), using the 1 μm guaranteed resolution service. The 6" size was chosen over the 5" format to facilitate ease of alignment during photolithography. The metallisation mask was fabricated as a brightfield pattern, whilst the electrode/wirebond window mask and probe substrate mask were darkfield.

3.2.2 Microfabrication

The starting point of our fabrication was a 4" silicon-on-insulator (SOI) wafer ordered from Ultrasil Corporation (Hayward, CA). This wafer had a handle thickness of 600 μm, a buried oxide (BOX) thickness of 1 μm and a device layer thickness of 25 μm. We also fabricated some probes with SOI wafers of 50 μm device thickness. Probes made with each thickness had very different mechanical properties. The 25 μm thick devices had flexible shafts, but needed to be introduced into the cortex slowly to prevent bending. The 50 μm thick devices could penetrate more easily, but broke more readily. It is likely possible to fabricate devices down to 10 μm in thickness at which they would become even more flexible, due to the unique mechanical properties of silicon at this length scale. However, it would have been impractical for our application, so we primarily used 25 μm wafers.

Fig. 3.3 depicts a simplified schematic of the process flow for the fabrication of the MEMS neural probes. The fabrication steps are detailed as follows.

The SOI wafer was cleaned in piranha etch (3:1 H_2SO_4:30% H_2O_2) to remove all organic residues. This is a critical first step. After rinsing in DI water and drying, it was then dipped in Transene BOE improved for 10 minutes to remove any native silicon oxide on the surface.

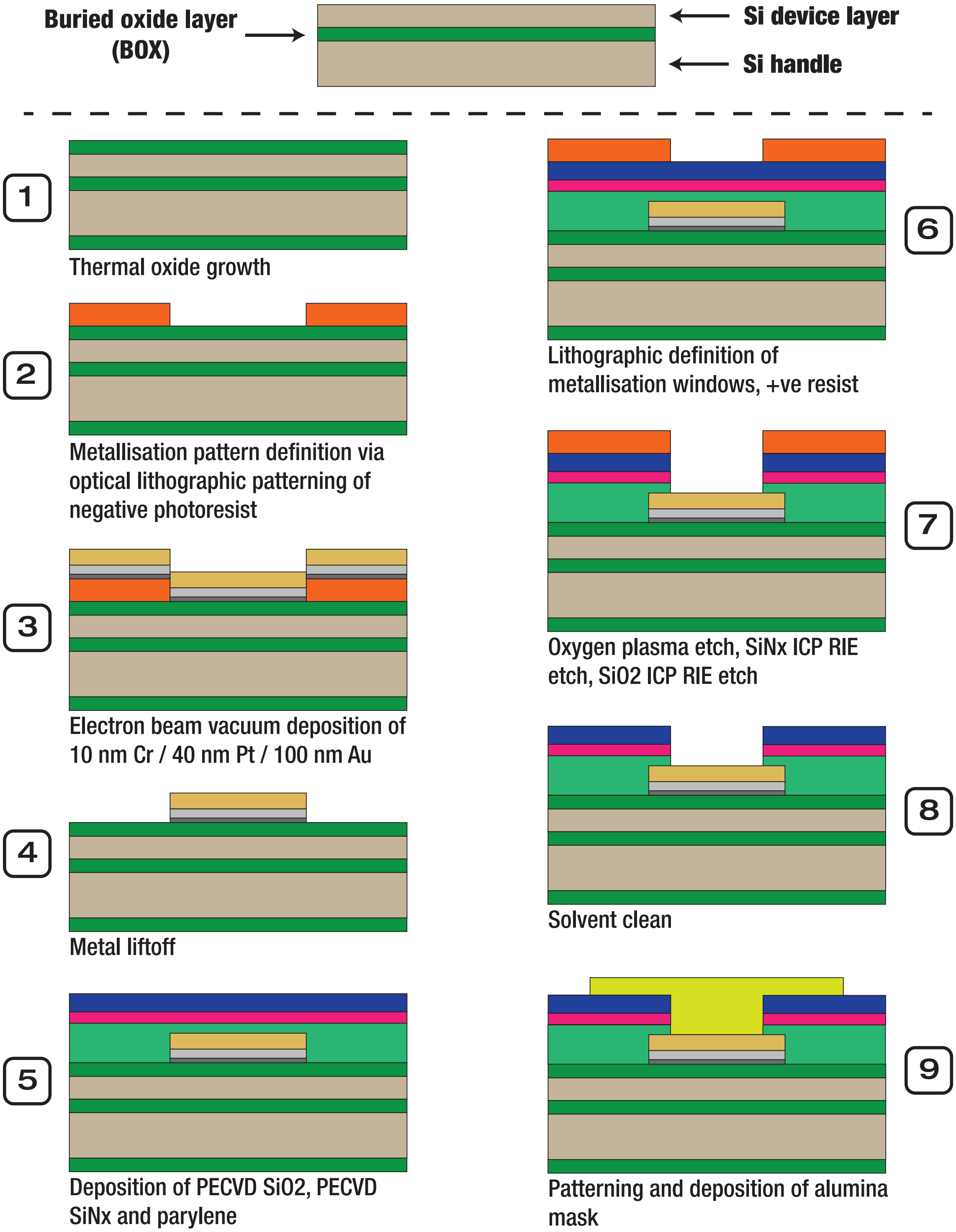
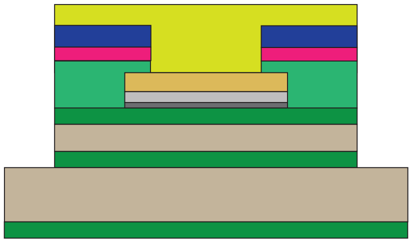


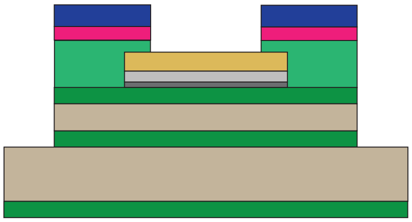
Figure 3.3: **Fabrication process flow for MEMS neural probes.** Diagram continued overleaf.

10



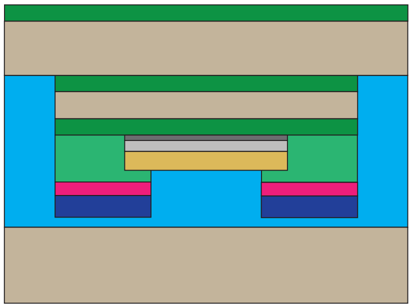
O2 plasma etch, SiNx ICP RIE etch, SiO2 ICP RIE etch, Si Bosch etch, SiO2 ICP RIE etch

11



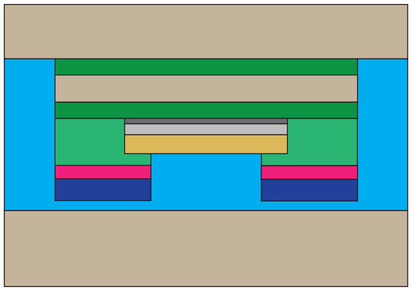
Alumina removal in phosphoric acid

12



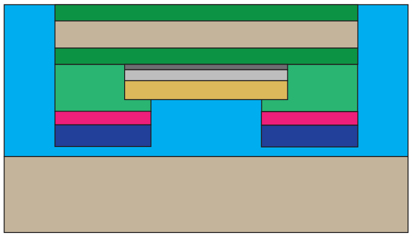
Flip bonding to silicon wafer using Santovac 5 vacuum oil

13



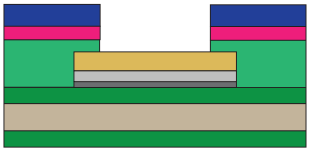
SiO2 ICP RIE backetch

14



Si backside ICP RIE cryoetch

15



Release of MEMS neural probe in isopropyl alcohol

The wafer was then carefully rinsed and dried thoroughly, and placed in an oxidation furnace at 1000°C. A dry oxidation recipe was run on the furnace for 498 minutes, forming a 200 nm thick oxide layer on the surface of the device layer. It is ill advised to grow an oxide layer much thicker than this. The refractive index of silicon oxide is 1.5, compared to that of silicon which is 3.44. If the oxide layer is too thick, this index mismatch causes blurring of the fine 1 μm features of the neural probe tip due to guiding of light through the oxide layer. The electrical conductivity of thermal silicon oxide is just 10^{-16} $\Omega^{-1}\mathrm{cm}^{-1}$ in comparison to silicon which is 10^{10} times greater. Thus only a thin layer is required to form a good insulating surface to deposit the metallisation.

Next, the wafer is sonicated for 5 minutes in Surpass 3000 (Dischem, Inc.) and rinsed gently under running DI water for 30 seconds before drying. This is an adhesion promotor for the hydrophilic oxide layer. Microchem lift-off resist (LOR) 3A is spun on to the wafer and baked. LOR is an excellent method of improving the quality of lift-off; it works by creating a slight undercut just underneath the photoresist layer, preventing bridging of vacuum deposited metal with the sidewalls. Once it has cooled, AZ5214E resist is spun on and soft baked at 95°C for 4 minutes, with the goal of using it as a negative resist in image reversal mode. Instead of AZ5214E, sometimes we use ma-N 2410 negative resist (Micro Resist Technology), and have obtained excellent results.

A Suss MA6 mask aligner is then used to perform the lithography on the wafer. An exposure of 2 seconds is used at a power of 290 W at 365 nm wavelength (i-line). One of the most critical aspects of this step is the type of contact used between the mask and wafer. The MA6 aligner *must* be set to the **vacuum contact** setting before exposure. Simply using hard contact will not get the wafer close enough to the mask to prevent blurring of the fine features. This aspect, together with the oxide thickness, is the key to getting high resolution results. Another very important factor is particulate contamination on the baked photoresist or on the mask itself. Any particles between the mask and wafer will create a gap, compromising the ultimate resolution of the lithography. The mask should often be thoroughly cleaned via overnight immersion in Nanostrip (Cyantek), to remove any particles of photoresist that may have adhered during previous exposures. The wafer itself should be blown off with a dry nitrogen after mounting on the mask aligner chuck, just prior to insertion into the tool. Feature sizes of 700 nm are likely possible if these techniques are kept in mind and attention to detail is maintained in general.

After exposure, the wafer is baked at 115°C for 2 minutes. After cooling down, it is exposed on the MA6 aligner to a flood exposure for 10 seconds. The wafer is then developed in MF-319 with gentle agitation for 160 seconds, before being rinsed in DI water and dried.

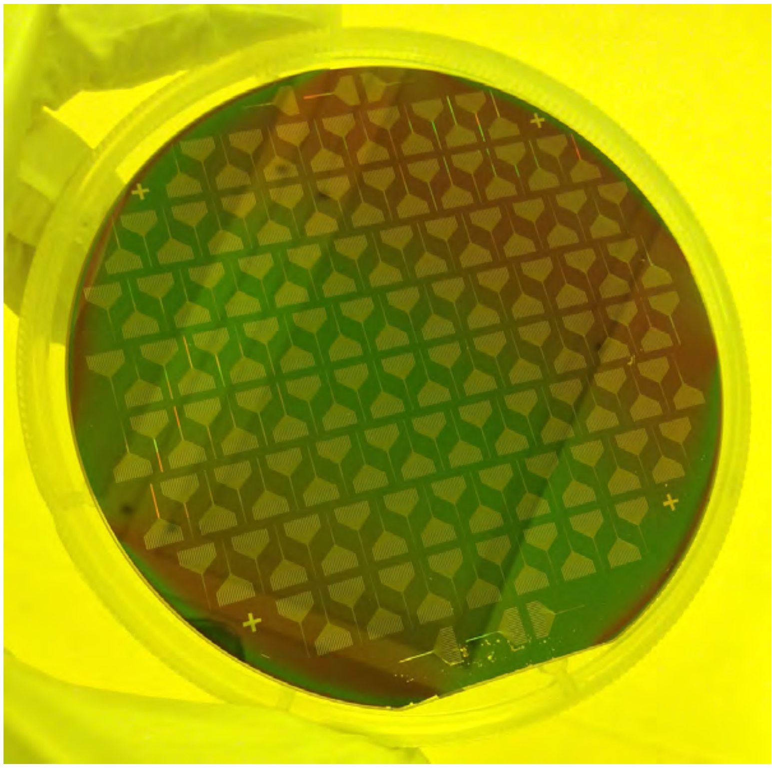


Figure 3.4: **Metallisation lift-off.** Shown is the 16-channel neural probe wafer after successful deposition of Cr-Pt-Au and metal lift-off.

This next step is also quite critical. After the development, a thick layer of undeveloped photoresist or debris may still occur in the developed areas of the wafer. If metallisation were done in this state, there would be poor adhesion between the metal and wafer, causing it to peel off. To prevent this, a ‘descum’ is done on the wafer by performing a 50-50-50 RIE plasma etch for 30 seconds.

The wafer is then immediately placed in an electron beam vacuum deposition system for metallisation. Just prior to placement, it is important to blow off the wafer with a nitrogen gun to remove any dust that may have been attracted to the wafer. The deposition system is used to deposit 10 nm of Cr as an adhesion layer, and then 40 nm of Pt and 100 nm of Au. The platinum layer is used to improve the stability of wirebonding to the bonding pads on the probe.

The metallised wafer is placed in Microchem PG remover (1-methyl-2-pyrrolidon) for liftoff. It is very important to use an NMP based stripper for liftoff instead of acetone as LOR forms an insoluble precipitate in acetone, and must be stripped off in NMP. Fig. 3.4 depicts a wafer after successful lift-off.

After rinsing in acetone, IPA and DI water, the wafer is dried and placed once again in RIE oxygen plasma. This time a 20-20-80 recipe is used for 15 minutes. This step is important to get the wafer thoroughly clean before deposition of silicon oxide, otherwise the film adhesion will be poor.

An Oxford Instruments system 100 plasma-enhanced chemical vapour deposition (PECVD) tool is used for the insulator deposition. Before placing the cleaned wafer in the instrument, a nitrogen gun is again used to blow off any particulates. First, 1 μm of silicon dioxide is deposited on the wafer using silane-argon (SiH_4-Ar) and nitric oxide (N_2O) at 350°C. Then 200 nm of silicon nitride is deposited with silane-argon and ammonia (NH_3). The nitride deposition is what really protects the neural probe against ionic ingress to the shaft, which causes degradation with implantation. It is a high quality film with a tight structure. An even higher quality film of silicon nitride can be deposited if atomic layer deposition (ALD) is used instead of PECVD. The other advantage of ALD over PECVD is that it is conformal, so the nitride film can cover the bare silicon sidewalls of the probe, protecting it against exposure to the extracellular fluid.

The wafer is again cleaned in RIE oxygen plasma with 20-20-80 for 15 minutes, before being blown off with nitrogen and placed in a parylene coater. Approximately 1.5 μm of parylene is deposited. The oxygen plasma treatment is absolutely crucial to creating excellent adhesion between the parylene and wafer, otherwise the parylene will delaminate in time on cortical implantation. Please note that it is just not the cleaning effect of the oxygen plasma that promotes high adhesion, but the activation of the oxide surface, through creation of free oxygen groups. These then readily react with the parylene on deposition, adhering it strongly.

Windows through the insulation layers are now created to access the microelectrodes and wire-bonding pads. A thick positive resist, AZ9245 (Clariant), is spun on the wafer and baked. Note that adhesion promotor is not required because of the hydrophobic parylene coating. The darkfield mask for the access windows is used in hard contact mode to perform the lithography after careful alignment using the alignment marks, and double checking the registration of the windows with the microelectrodes on the probe. On development of the exposure, the wafer is rinsed in DI and dried, then placed in an oven at 80° overnight. This hardens the resist and makes it resilient against all the coming etch steps. It is very important not to set the temperature any higher. Doing so will cause shrinkage of the resist, and will enlarge the feature size of the developed windows.

A series of dry etching steps are then performed to etch through the parylene, silicon nitride and silicon dioxide layers. First, the wafer is etched under 50-50-50 RIE oxygen plasma for 15 minutes to etch through the parylene layer. Next, the devices are mounted on a 6” wafer with Santovac 5 vacuum oil, and the entire assembly is placed in the Oxford Instruments system 100 ICP for nitride

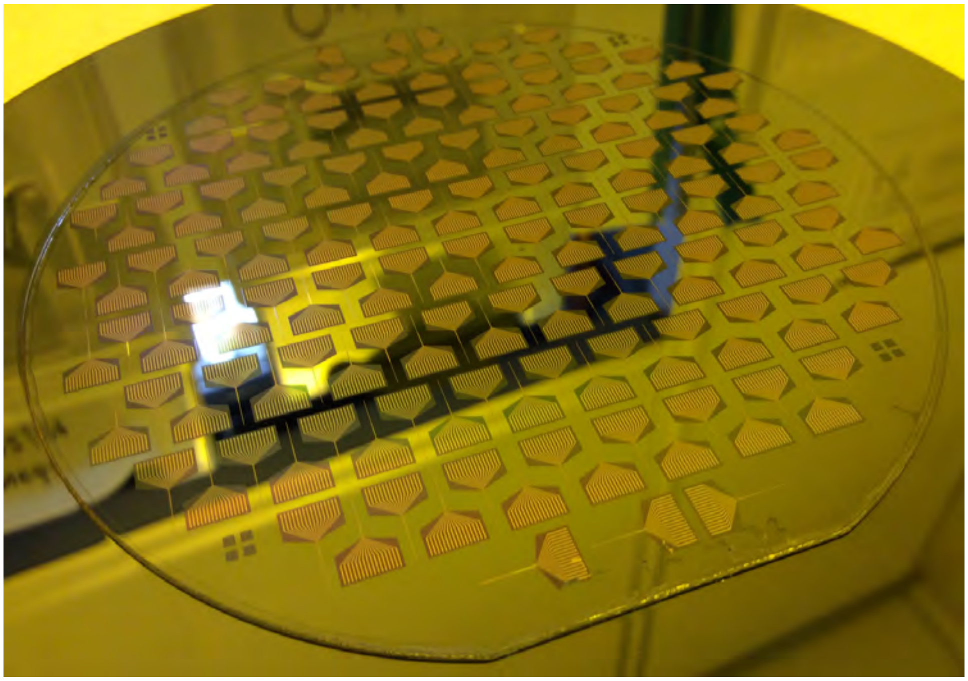


Figure 3.5: **Probe bulk micromachining using alumina mask.** Although it is only 200 nm thick, the alumina mask completely protects the probes beneath them from all the harsh etching steps.

etching. An SF_6 recipe is used. Immediately after, the silicon dioxide layer is finally etched with a C_4F_8 / O_2 recipe. It is important not to over etch on the oxide etch step as the C_4F_8 / O_2 recipe does etch gold, albeit at a low rate.

With the window etch steps complete, the wafer is stripped in PG remover, then cleaned in acetone / IPA / DI water before drying. AZ9245 resist is again spun on to the wafer and baked. Now, the final lithography step is performed to define the probe bulk. The darkfield probe substrate mask is meticulously aligned with the structures on the wafer and hard contact is used to expose the resist. After development in MF-319 and rinse/dry, the wafer is placed in a sputter system and an aluminum target is sputtered under oxygen partial pressure to deposit 200 nm of alumina (Al_2O_3). Alumina is an excellent etch mask for bulk micromachining of silicon via cryoetching and the Bosch process [73].

The wafer is placed in acetone for lift-off overnight. On coming to clean and rinse the wafer, it is extremely important to continuously rinse the wafer with acetone whilst removing it from the lift-off vessel before immediately re-immersing it into IPA. This is to prevent re-adhesion of the alumina particulates from the lift-off process on to the wafer. Once they re-adhere, it is virtually impossible to remove them again without destroying the patterned alumina.

Figure 3.6: **Neural probe substrates in antistatic storage case.** The antistatic case prevents the extremely light probes from spontaneously becoming airborne and adhering to extraneous objects and surfaces, by dissipating static charge.

After rinsing in DI water and drying, the wafer is mounted with Santovac 5 onto a 6” silicon carrier wafer that has previously been thermally oxidised and coated in alumina. An Oxford Instruments system 100 ICP is then used to etch away the parylene with oxygen plasma, and the silicon nitride etch recipe used before to etch the windows is employed to remove that overlaying layer. However to etch the silicon oxide layer, a different recipe is used, employing C_4F_8 and SF_6, instead of O_2. This is because the C_4F_8 / O_2 recipe is very hard on alumina and easily damages it, whilst alumina has very good selectivity against the C_4F_8 and SF_6. In fact, this recipe was developed especially to facilitate the fabrication of edge polished nanotaper couplings for our devices in Chapter 4. The colour change to the wafer makes it very clear when the final silicon oxide etch recipe has completely removed the oxide and reached the underlaying silicon. Now, the system 100 ICP is used to perform deep reactive ion etching (DRIE) on the wafer using the Bosch process. The DRIE is used to completely etch through the 25 μm silicon device layer until it stops at the BOX layer. The C_4F_8 and SF_6 oxide etch recipe is then finally used to etch through the BOX layer until it reaches the silicon handle. Fig. 3.5 shows the wafer with alumina mask *in situ* after all the etch steps are completed.

Prior to further processing, the alumina mask needs to be removed. This is done by placing the

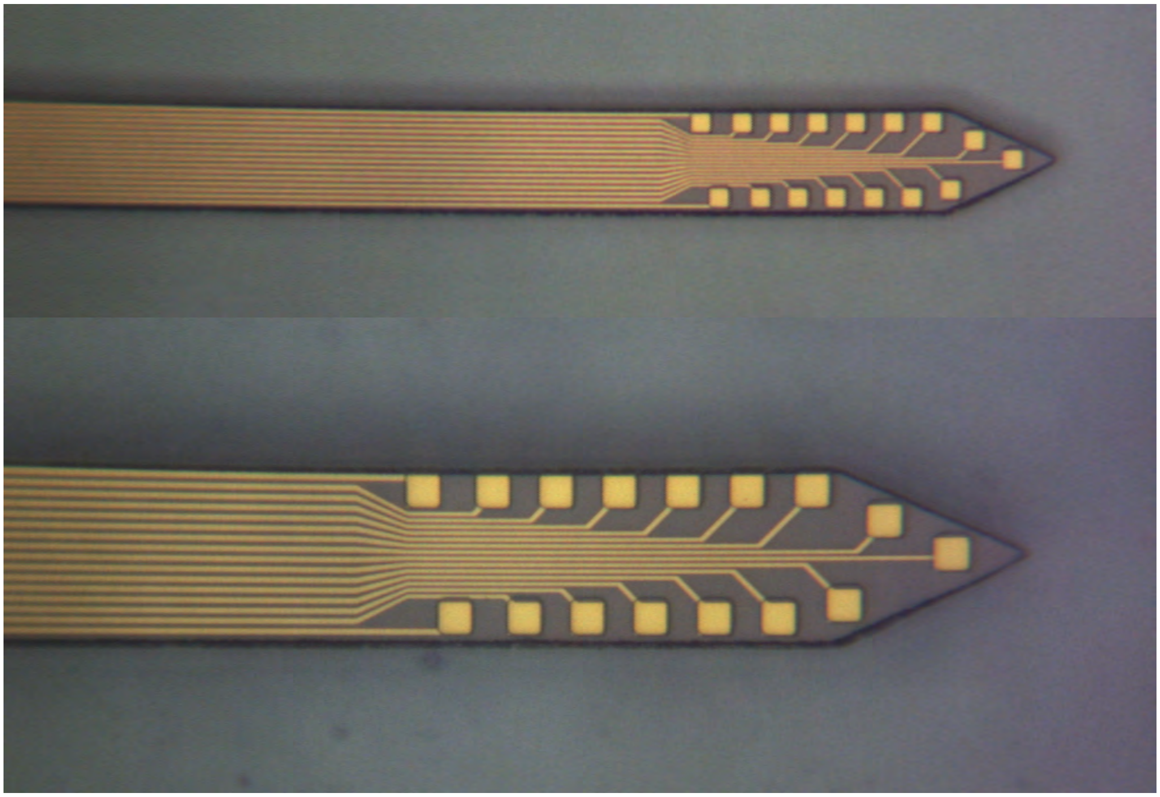


Figure 3.7: **Microelectrode populated shaft tip of completed 16-channel neural probe.** Note the extremely clean fabrication. The spacing between the $\sim$ 1.25 μm traces at the distal end is $\sim$ 500 nm, which is an excellent result for photolithography. Following the fabrication steps as described, results like these are readily obtained.

wafer in phosphoric acid at 60°C for 15 minutes. After rinsing in DI and drying, in preparation for use in the Oxford ICP etcher, Santovac 5 is used to mount the device side of the wafer to a 6" oxidised silicon wafer coated in alumina. It is very important to apply a a good coating of the vacuum oil to the wafer face to fill in the 25 μm deep crevices between the etched probes. Santovac 5 is an excellent material for ensuring thermal contact between wafers in vacuum systems, and is far superior to fomblin oil which is what is commonly used. This is due to its much lower vapour pressure of 10^{-10} Torr and the fact that it easily and readily dissolves in acetone and IPA. Fomblin on the other hand has a vapour pressure of 10^{-6} Torr, and is virtually impossible to completely remove, even using dedicated perfluorinated cleaning solvents. Since the vacuum oil in this application is being applied directly to the probes and their electrode sites to make thermal contact and physically protect them from the etching gases, it's important to use a material that is readily cleaned off and that will not leave a residue on the device surface.

The prepared wafer assembly is placed in the Oxford Instruments system 100 ICP, and the C_4F_8

and SF_6 oxide etch recipe is used to remove the 200 nm thermally grown silicon oxide layer on the handle side of the 4” wafer. Once that etching is complete, it will stop on the bulk silicon of the handle. The table of the ICP etcher is cooled down to -140°C and then a cryoetch recipe is used to etch away the silicon handle. Alumina has an etch selectivity of 5000:1 against silicon with cryoetching, so the alumina coated 6” carrier wafer will not be etched away by the process. Furthermore, silicon dioxide has a high 200:1 selectivity with cryoetching and thus the cryoetch step will stop on the neural probe 1 μm BOX layer without damaging it, as long as the etch is timed correctly not to over etch for too long.

Two more deposition steps are required before the probes are finally ready, and that is to coat the back of the probes with silicon nitride, and subsequently parylene. The 6” carrier wafer with probes *in situ* after the cryoetch process is placed in RIE oxygen plasma under 20-20-80 conditions and is cleaned for 15 minutes. The carrier wafer with probes adhered to it is then placed in the Oxford system 100 PECVD, where 200 nm of silicon nitride is deposited over the BOX layer of the probes. PECVD nitride is not as conformal as ALD nitride, however it will still cover the exposed silicon sidewalls of the probes. After nitride deposition, a further 15 minute O_2 clean is performed in the RIE, and then the wafer assembly is transferred to the parylene coater, where 1.5 μm of parylene is deposited.

This completes all the primary fabrication steps, and now the neural probes are ready and need to be carefully released from the 6” carrier wafer and cleaned. This is done by placing the carrier wafer with probes *in situ* into a bath of acetone for a few hours, and then carefully transferring the wafer to an IPA bath where the assembly is left overnight. The acetone step loosens and lifts the silicon nitride and conformal parylene film from the non-probe areas, which are not completely clean yet due to the Santovac 5 oil. The parylene on the probes are not affected by immersion in acetone for just a few hours, due to the strong adhesion from the oxygen plasma processing. The probes can then ‘lift-off’ from the rest of the silicon nitride, parylene and carrier wafer during the IPA immersion step, and are now free and complete. They are carefully collected from the IPA bath using a small fluid suction tool, rinsed with a bit more IPA, and then carefully dried. The probe shafts are quite flexible, but they will break if they are suddenly smashed against a surface. It is also important to note that they are incredibly light, and will readily stick to objects with static charge. Thus great care should be taken when cleaning and storing the probes. An antistatic vacuum pick-up tool is invaluable for handling them. Fig. 3.6 depicts the probes after collection, cleaning and storage. An antistatic storage case is used to securely store the probes for future use, and prevents

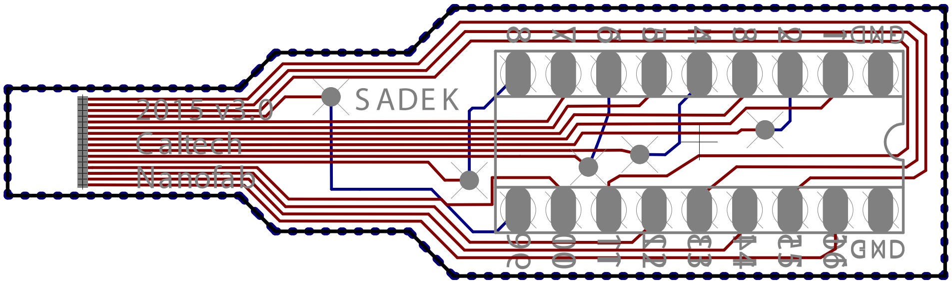


Figure 3.8: **16-Channel probe PCB design.** The PCB is 1.6 mm thick and the contacts are plated with ENEPIG, which enables wirebonding with gold leads. A ground plane throughout the PCB provides shielding for the μV potentials conducted by the traces from the coupled probe.

spontaneous attraction of the probes to other objects and surfaces by dissipating static buildup.

Fig. 3.7 depicts close up optical micrographs of the completed neural probe tip. Comparing the quality of the probe fabrication with that described in the literature with photolithography, it is evidently superior. Although the mask design implemented the distal wire traces with 1 μm width and 1 μm spacing, the lithographic exposure and metal liftoff process left the wiring slightly enlarged to 1.25 μm and the spacing reduced to 500 nm.

3.2.3 Assembly

Fig. 3.8 shows the PCB designed to mount and interface with the 16-channel probe. The board is 1.6mm thick, and the traces from the wirebonding pads to the header are shielded with a ground plane throughout the PCB. An 18-DIL header is used to interface the entire probe to the amplifier headstage, which has an 18-DIP ZIF socket for easy loading and removal (see Appendix B and Chapter 4.4.2). The header is mounted from the bottom of the PCB, and soldered from the top. Sixteen of the pins carry the signals from the neural probe, whilst two of the pins connect the ground from the headstage to the ground plane in the PCB. The PCB is ordered from the manufacturer with electroless nickel, electroless palladium, immersion gold (ENEPIG) plating. It is an expensive process, but essential for reliable gold wirebonding.

First, the header is soldered into place on the PCB with leaded solder. The soldered board is throughly cleaned with flux remover and IPA and dried with a nitrogen gun. A very small droplet of cyanoacrylate adhesive is placed on the probe mounting area of the PCB. A neural probe is collected from its static storage case using a vacuum pick-up tool and is carefully mounted onto the PCB for bonding via the adhesive. Fig. 3.9 depicts the probe after this step.

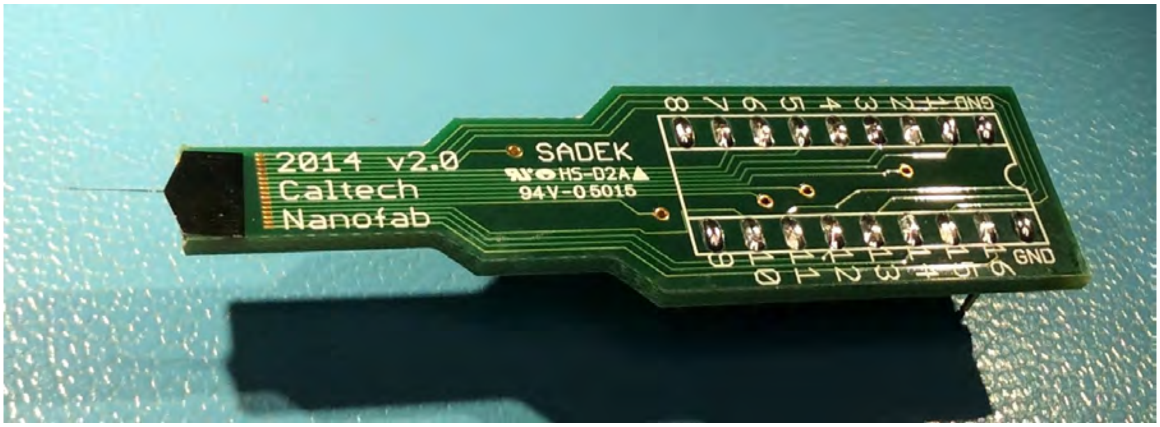


Figure 3.9: **Assembled 16-channel probe.** Shown is the assembled neural probe prior to wirebonding and epoxy encapsulation.

It is now ready for wirebonding. A magnetic chuck was machined, with a dummy 18-DIL ZIF socket with support platform to allow secure mounting of the assembly to the wirebonder stage. Gold leads are used to wirebond the ENEPIG PCB pads to the bonding pads on the probe substrate. Resinlabs EP1238 epoxy is finally used to encapsulate the wirebonds and the probe substrate mounting region on the PCB. It is left for one day to cure, after which the fabrication is complete.

As described in Chapter 2.7.3, it was important for our NEMS experiments to electroplate the microelectrodes on our neural probes to reduce their impedance and ameliorate filtering. To do so, a custom interface was built to mount the assembled probes via ZIF socket on a strand, and permit a BNC connection to be switched in to each of the 16 electrodes on the probe via a rotary switch. The probe tip was immersed in a small vial of Sifco 5355 non-cynanide gold plating solution. A Keithley 2601A sourcemeter was connected to the probe interface, with the positive current connection to the electrode. The ground connection was made to a platinum-iridium wire, which was inserted into the gold plating solution with the probe. The sourcemeter was set up in current pulse mode, and a current of 0.1 μA was programmed. We have seen reported higher currents used during electroplating ($\sim 2 \mu m$), but have found lower currents generate plating of much better quality [74]. A portable impedance tester is used to test the electrodes between plating steps. The electrodes initially have an impedance of $\sim$ 1.5 MΩ. A 15 second pulse is applied to each electrode until the impedance is just under $\sim$ 1 MΩ. Five second pulses are then repeatedly applied until the impedance of each electrode is $\sim$ 100 kΩ. One second pulses are then applied as is needed to reduce the impedance to $\sim$ 30 kΩ.

3.3 64-Channel Probes Fabricated by Electron Beam Lithography

Although it is a serial process and so limited in throughput, electron beam lithography has the advantage in ultimate resolution. Feature sizes of 10 nm are routinely possible, and specialised methods of fabrication can even exceed this limit. We decided to design a 64 channel probe to test the ultimate limits of conventionally wired devices, and also to facilitate our designs for experiments presented in Chapter 4. We plan to demonstrate with piezophotonic NEMS devices the multiplexing of neural potentials from 8 extracellular microelectrode signals during *in vivo* recordings in mouse cortex. Since we wish to demonstrate multiplexing from separate single-units, a one dimensional high density probe that spans the depth of the cortex would be ideal for identifying independent single-units spanning the six layers of cortex. Eight sets of 8-channel clustered sections along the length of the probe would each be dedicated to identifying separate single-units. Once they are identified, the best signal on each section would be ‘switched in’ to a unique NEMS device to demonstrate transduction and multiplexing on a single optical fibre, whilst the rest of the electrodes would concurrently record the signals with a conventional amplifier for cross validation.

3.3.1 Design

The 64-channel neural probe was designed around initial specifications for wiring that was 200 nm thick with 200 nm spacing at the crowded microelectrode area of the probe distally. This is about 100 nm below the best specification found in the literature [19], but having personally made structures much smaller than this using electron beam lithography, this was deemed more than feasible fabrication wise. With a $10\ \mu\mathrm{m} \times 10\ \mu\mathrm{m}$ microelectrode area, the probe end shank dimensions were found to be optimal if they tapered down from 50 μm to 20 μm along a length of 1.9 mm. The wiring was expanded to 350 nm at the 50 μm neck, and then to 1.48 μm, and finally to 2 μm along the rest of the shank. This was done to minimise the impedance penalty of such thin wires. The wirebond pads were defined with a $150\ \mu\mathrm{m} \times 200\ \mu\mathrm{m}$ area and 50 μm spacing, and these specifications drove the dimensions for the rest of the probe, and in turn, the PCB designed to mate with it. The window openings to the electrodes and bond pads are slightly smaller to ensure slight overlapping of the insulation layers over the pads, as with the 16-channel probes.

As electron beam time is very expensive, the metallisation areas were split into electron beam lithography sections and photolithography sections. In the electron beam lithography section, the nanometre scale wiring was defined separately to the microelectrode pads so that they could be writ-

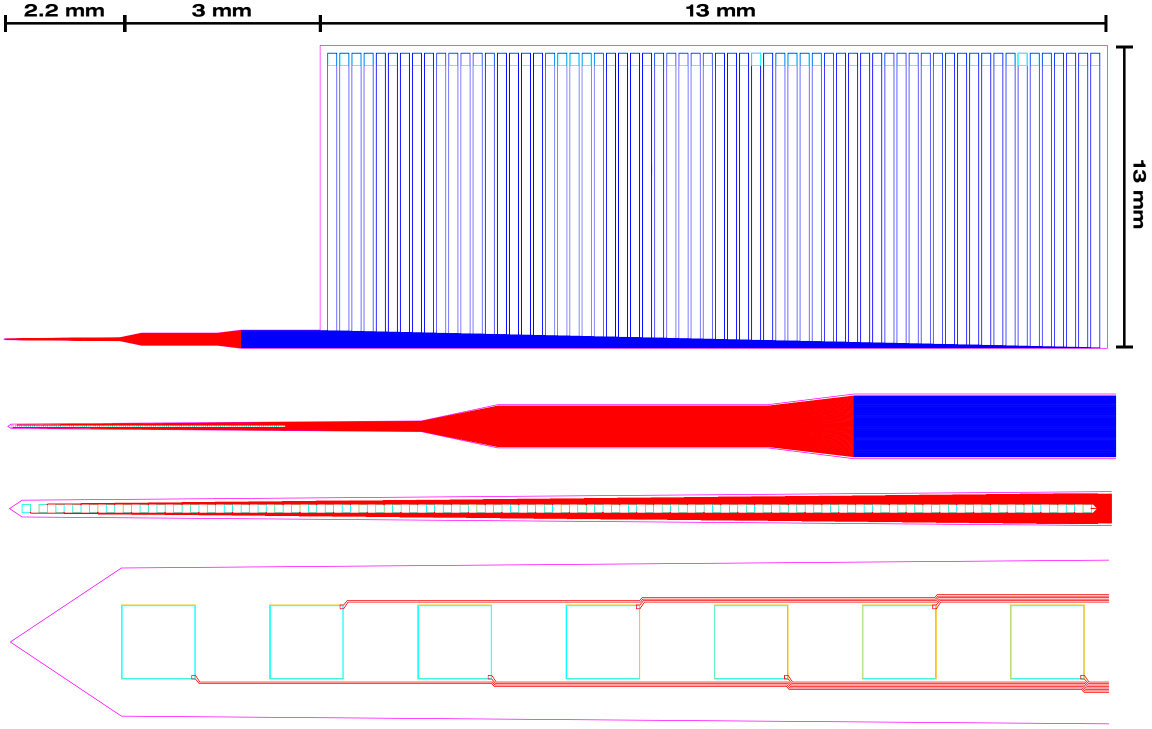


Figure 3.10: **64-Channel neural probe design.** 64 recording channels of 10 μm × 10 μm dimension and 10 μm spacing are arranged linearly over a 1.3 mm length. The wirebond pads are 150 μm × 200 μm and are spaced 50 μm apart. The dimensions of the bulk probe are shown, and the insets show successive magnified views of the probe shaft and tip. Red and gold areas, for the wiring and microelectrodes, are defined by electron beam lithography. Dark blue represents the photolithographic metallisation area. Light blue are the windows through the insulation, and purple defines the probe substrate.

ten with different beam currents to speed up the beam writing process, but also to be able to define different metallisation recipes for the nanowiring and microelectrode sensors. Fig. 3.10 illustrates the the separate lithography sections. The metallisation photolithography mask is brightfield, and the probe substrate and insulation windows were darkfield, as with the 16-channel probes. Note the intersection between the red areas of the electron beam section and dark blue area of the optical lithography section. The traces here were designed to overlap by 20 μm to facilitate accurate registration during the alignment step in the fabrication process.

As the electron beam lithography section is split into different beam writes, they need to be aligned to reference markers on the wafer in the electron beam pattern generator (EBPG) before each write. A separate electron beam lithography layer was created to do just that, and on the wafer scale schematic in Fig. 3.11 they can be seen as the four blue dots clustered around the probe tips. The alignment markers were designed as four 20 μm × 20 μm squares, arranged around the

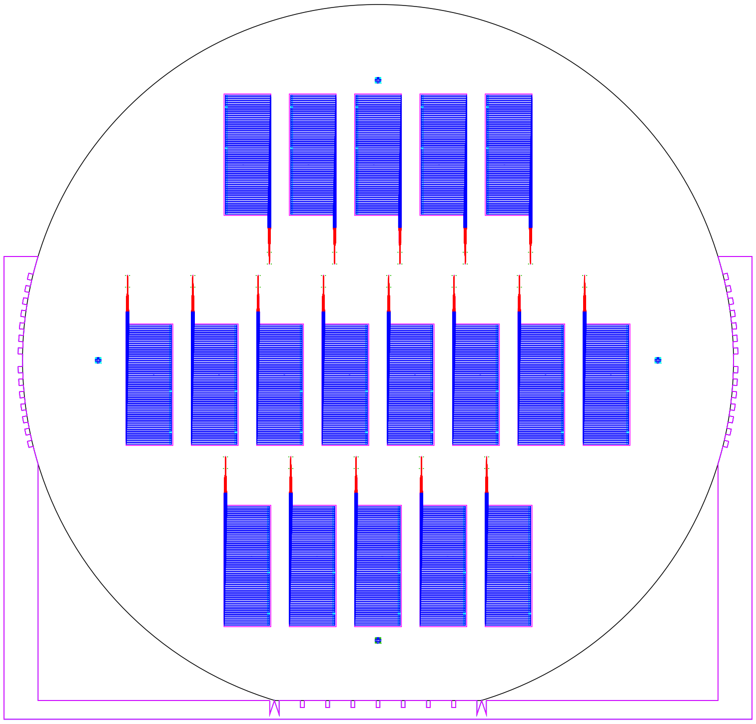


Figure 3.11: **Wafer-scale mask design for 64-channel probes.** 18 probes are packed over 3” wafer. The EBPG write time for each probe is 105 minutes at a beam current of 500 pA for the wires and 1 nA for the microelectrodes.

electrode area of the probe shaft. They can be more easily seen in Fig. 3.12. Better alignment is obtained the closer the markers are to the structure to be written, but they should be spaced at least 200 μm away to prevent the resist patterns being ruined as the EBPG finds the markers with the electron beam.

3.3.2 Nanofabrication

A 3” or 4” SOI wafer (Ultrasil) with a 20 μm thick device layer, 1 μm BOX layer and 600 μm handle layer is cleaned in piranha etch. After rinsing in DI water and drying, the wafer is dipped in Transene BOE improved for 15 minutes and washed thoroughly in DI water. Subsequent to careful drying, the wafer is placed in an oxidation furnace at 1000°C as in Chapter 3.2.2, and 200 nm of oxide is thermally grown.

The first lithographic step in fabricating the probes is to write the alignment marks with the EBPG. The wafer is immersed in Surpass 3000 and sonicated for 5 minutes. It is then rinsed under DI water for 30 seconds and dried. PMMA A4 resist is spun on to the wafer and baked, as detailed in

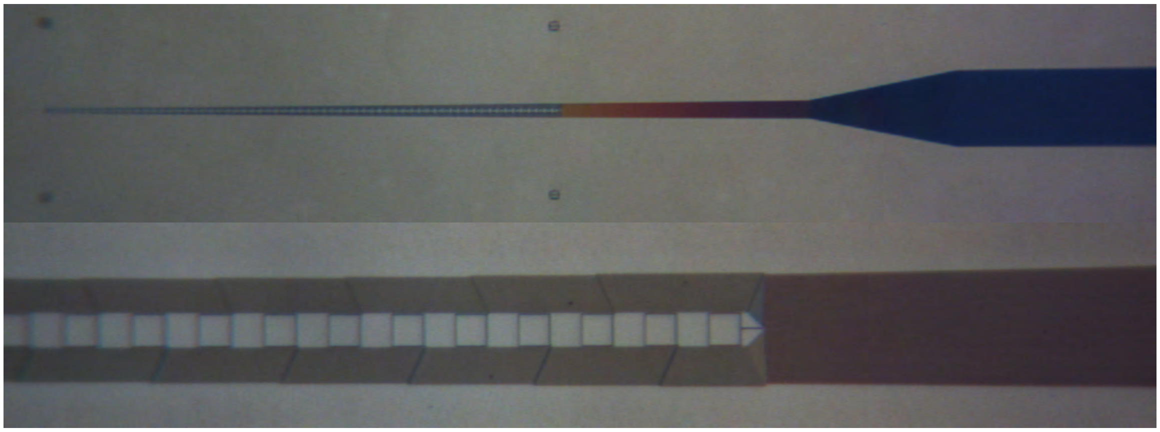


Figure 3.12: **Electron beam lithography of 64-channel probes.** PMMA that has been electron beam patterned with 64-channel probe, prior to metallisation.

Chapter 2.7.1 and Appendix B. The wafer is loaded into the EBPG and the electron beam alignment marks and photolithographic alignment marks are written with a beam current of 1 nA, which takes approximately 115 minutes with a dose of 1300 $\mu\mathrm{C}/\mathrm{cm}^2$ at 100 kV gun voltage.

After development, the wafer is placed in RIE oxygen plasma for a descum under 50-50-50 conditions for 15 seconds. An electron beam vacuum deposition system is then used to deposit a recipe of: 5 nm Ti / 65 nm Au / 70 nm Ti. Lift-off of the metallisation is performed in acetone. This particular metallisation recipe is used as we have found it to produce excellent contrast of the alignment marks in the EBPG, which never fails to find the marks during the alignment step.

The wafer is processed with Surpass 3000, and PMMA A4 is spun on and baked again. It is loaded into the EBPG and the probe nanoscale interconnects are written with a beam current of 500 pA. This step takes 103 minutes per probe. For an entire wafer of 18 probes, the write time is 31 hours. Since each probe has its own electron beam alignment marks, the writes can be done one probe at a time whenever write time on the tool becomes available, until the entire wafer is patterned. EBPG write time is very expensive, so the true cost of fabricating these probes is substantially higher than with the parallel photolithographic process used on the 16-channel probes.

After development of the wafer, it is placed in RIE O_2 plasma at 50-50-50 conditions for a 15 second descum. It is loaded into the electron beam deposition system, and the following metallisation is deposited: 5 nm Ti / 100 nm Pt. Platinum is used as the metal for the interconnects as we have found it to make far more reliable interconnects than gold. It is less prone to cracking with flexure of the probe. The wafer is placed in acetone for lift-off overnight. On removal, it is very important to keep rinsing the wafer with acetone as it is being removed, to ensure that the 200 nm strips of platinum that lift off the wafer are flushed away, and don't re-adhere to the patterns, shorting them out. The wafer is then immersed in IPA, and then rinsed in DI water before drying.

A very important next step is to clean the wafer in RIE O_2 plasma before further processing, which is done at 20-20-80 conditions for 15 minutes. The wafer is processed again with Surpass 3000 adhesion promotor, but this time it cannot be sonicated, otherwise the interconnects would be destroyed. It is instead immersed in the Surpass 3000 and left for a few hours. It is then rinsed for 30 seconds under DI before drying.

PMMA A4 is spun on and baked once more, and then wafer is then loaded in the EBPG. It is again aligned to the alignment marks, and the microelectrode patterns are written at a beam current of 1 nA. This step only takes 2 minutes per probe.

After development and descum, the wafer is metallised with 5 nm Ti / 145 nm Au, and then lift-off is done in acetone. Subsequent to lift-off processing, the wafer is cleaned again in RIE O_2 plasma for 15 minutes (20-20-80), and processed in Surpass 3000 through a few hours of immersion as before. The wafer is then ready for the photolithographic steps.

Photolithographic metallisation of the rest of the probe proceeds in almost exactly the same manner as in Chapter 3.2.2. Of course though, the difference being the mask pattern needs to be critically aligned to the wafer using the Suss MA6 aligner, making use of the specially written alignment marks made during the first EBPG write. However even this level of alignment will be insufficient. A second step of mask alignment needs to be made using the probe traces themselves, moving the alignment microscope to all parts of the wafer to ensure the alignment between traces on the mask and those on the wafer is accurate to within a few hundred nanometers. To verify alignment, contact should be made between the wafer and mask before an exposure is committed, to ensure registration of the interconnects.

After exposure, development, metallisation of 5 nm Ti / 45 nm Pt / 100 nm of Au and lift-off, the wafer is cleaned in RIE O_2 plasma under 20-20-80 conditions for 15 minutes. The next step is one of the most critical steps to the reliability of the probe, namely thermal annealing of the metal traces. As the metal interconnects are so thin and are exposed to stress during flexure of the probe shaft, they need to be annealed to improve their resistance to fracture. It is this step that dramatically improves the yield of viable electrodes on the neural probe. Thermal annealing is done at 250°C for 3 hours in argon using a Jipelec rapid thermal annealer. Fig. 3.13 illustrates the 64-channel probe after completion of all metallisation steps.

The rest of the photolithographic steps, MEMS machining and device processing are identical

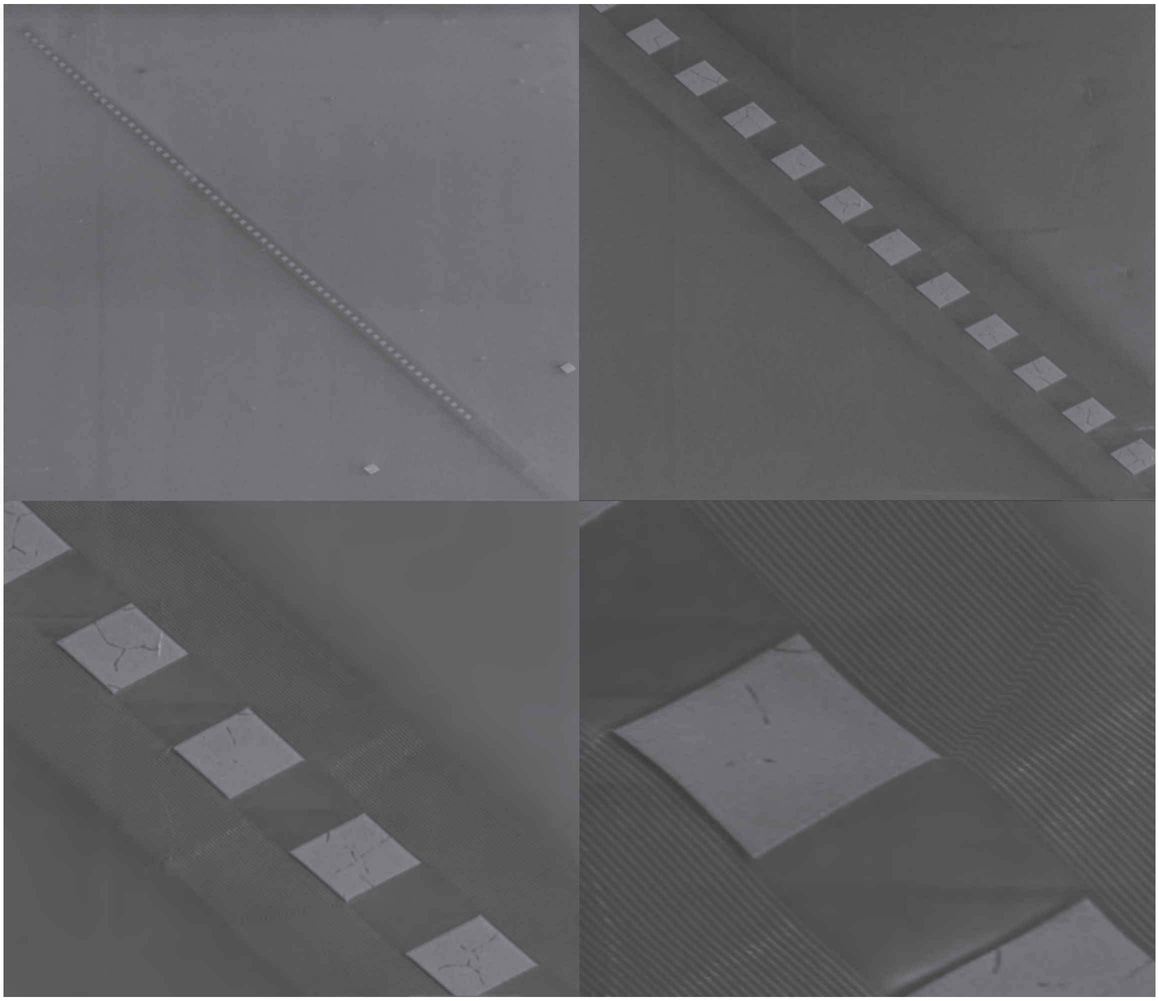


Figure 3.13: **Scanning electron micrograph of 64-channel neural probe.** Image was taken after all metallisation steps and annealing. Note the absence of broken/cracked traces. This particular probe was fabricated slightly different to that described in the text; both the electrodes and nanowires were made of 5 nm Ti / 100 nm Au, which was deposited in one step. The method described in the text yields even better results. Also note the $[20\ \mu\text{m}]^2$ alignment marks for electron beam writing.

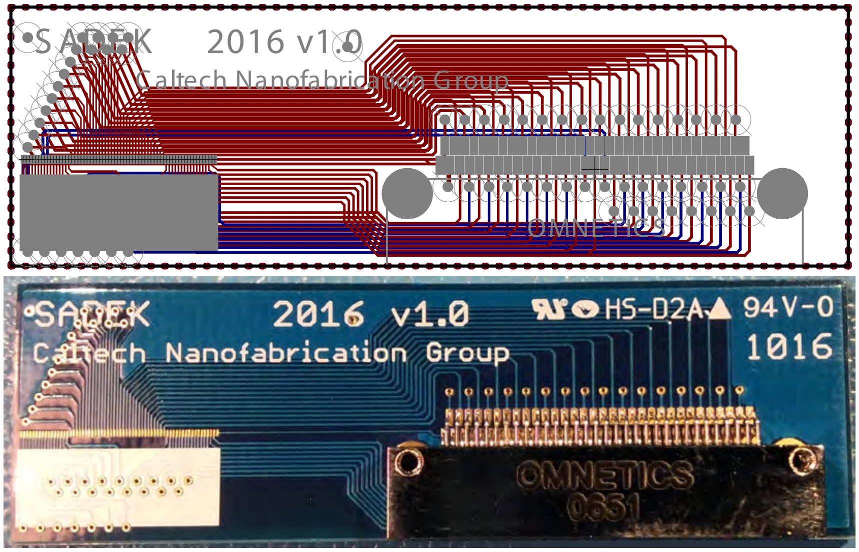


Figure 3.14: **64-Channel neural probe interface PCB.** Depicted is the PCB CAD design, and actual PCB with assembled Omnetics connector.

to the remainder of the process flow described in Chapter 3.2.2 for the 16-channel probes.

3.3.3 Assembly

Fig. 3.14 depicts the design and assembly of the PCB used to interface the 64-channel probe to the headstage amplifier described in Chapter 4.4.2 and Appendix B. This is a 4-layer PCB with a ground plane running through the board. It is ENEPIG plated for wirebonding and easy SMD soldering. The interface between the PCB and headstage is made via an Omnetics 65-pin Bi-Lobe female connector (A29100-065), to mate with the corresponding male connector on the headstage.

The Omnetics connector is first soldered to the PCB. The leads are finely spaced (0.01”) and are double rowed, thus skill is required in soldering the connector. A fine tipped temperature controlled soldering iron is used with Kester 44 / 28AWG solder (24-6337-0007). Solder flux is painted on the board, and the connector is placed in the correct position over the SMD soldering pads. Under a stereo microscope, the end rear leads are soldered first to hold the part in place. A fine pair of tweezers is then used to lift up the front row leads, in order to gain access to the rear row. The rear leads are soldered one at a time, and then the front leads are pushed down to their SMD pads and soldered into place. The flux rosin is cleaned with flux remover, and the entire board is immersed in IPA for several minutes and then dried with a nitrogen gun. #0-80 hex screws, washers and nuts are finally used to securely affix the connector to the PCB using the through holes.

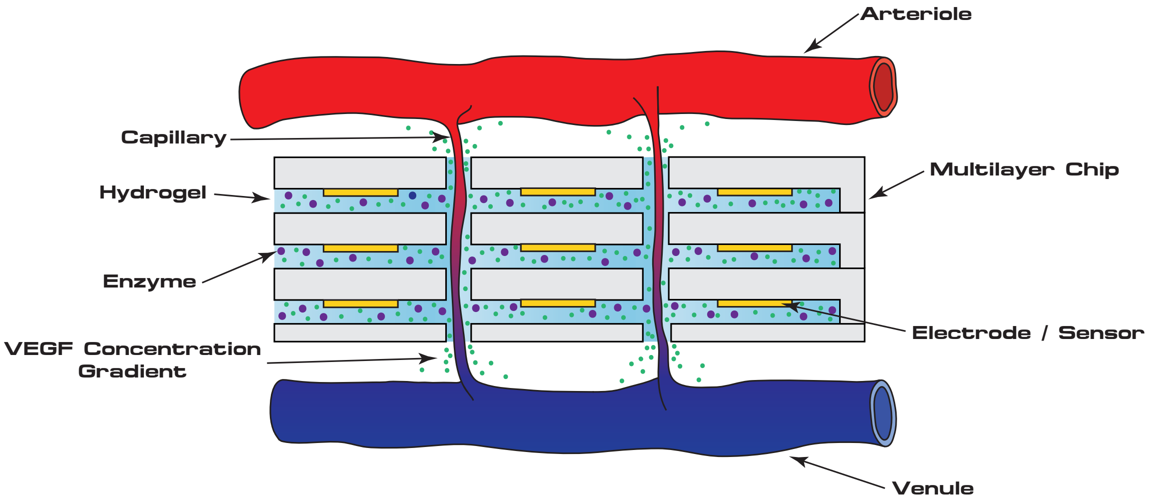


Figure 3.15: **Schematic of proposed microfabricated vascular interface.** Diffusion of angiogenesis growth factors out of the hydrogel matrix embedded in the chip induce capillary growth into the chip. A microenvironment is then formed for rapid diffusional exchange between sensors on the chip and the circulatory system. The same technique could be used to stabilise wireless implantable neural interfaces through anchoring to capillary beds.

As with the 16-channel probes, a dab of cyanoacrylate adhesive is applied to the probe mounting area on the PCB. The 64-channel probe is manipulated into position on the PCB with a vacuum pick-up tool. Each of the 64 wirebonds are then carefully made using gold leads. Resinlabs epoxy is finally used to encapsulate the wirebonds and affix the probe to the PCB permanently.

3.4 Microfabricated Vascular Interface

3.4.1 An Artificial Capillary Bed

There is a need for wireless, implantable sensors that can detect in real-time the concentrations of various proteins, metabolites and ions in blood plasma. Direct implantation of sensors within tissue does not facilitate such detection, as it relies on the diffusion of molecules from the interstitial fluid to the sensor. There is thus a delay of tens of minutes between the variations in concentration of target ions/molecules in the blood and that detected by the sensors. Furthermore, the eventual encapsulation of implanted devices within a thick layer of fibrous tissue due to tissue reaction further impedes the diffusion of such molecules to onboard sensors.

Here we present a proposal for interfacing a sensor chip directly with the vascular system to allow real-time detection of species in blood plasma, through the formation of an artificial capillary bed within the chip. As a result, diffusion of target molecules to the sensors occurs at the same physiological length scale as that from capillaries to the tissue they supply in the body.

Our proposed method for overcoming the problem of sensing from blood plasma is to build a device that interfaces with the vasculature in the same manner of tissue itself. Instead of placing the sensor within the blood proper via the lumen of a blood vessel, which would elicit thrombosis, we induce the body to bring the vasculature to the chip through the formation of an artificial capillary bed via angiogenesis.

The circulatory system consists of the heart, lungs and a system of arteries and arterioles that supply oxygenated blood to the tissues via the capillary bed. A capillary lumen is only $\sim 10\ \mu$m in diameter, and a capillary exists within a few tens of microns of every cell in the body. This ensures oxygen and nutrients in the blood plasma can quickly diffuse to the cells through the capillary walls, which are only one cell diameter thick. Likewise, carbon dioxide and waste products diffuse from the tissue into the blood for removal via the capillary bed. The capillaries then return the blood to the heart via a system of venules and veins. The blood itself never comes into contact with any tissue in the body except for the specialized endothelial cells which line the lumen of every blood vessel, and from which the capillaries are solely comprised. Thus a route to interfacing with the blood stream is to build an artificial capillary bed around a micro/nanostructured biosensor that would allow diffusional exchange of species with it.

Our idea is to build a multilayer chip, embedding microelectrodes and/or other micro/nano sensors (such as nanoelectromechanical systems or optical resonators) within a gelatinous matrix. Micromachined holes 20 - 50 μm in diameter are fabricated vertically throughout the chip, and are also filled with the matrix. The holes are spaced apart similarly to the hole diameter. The gel matrix performs three functions. First, it incorporates angiogenesis growth factors such as vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF)-BB and fibroblast growth factor (FGF)-2 [75] . Once the chip is implanted into tissue, the angiogenesis growth factor concentration gradient that arises from diffusion out of the hydrogel induces arterioles and capillaries to branch off existing vasculature and guide capillaries to grow into the chip. Secondly, the gel matrix acts like a tissue, in the sense of a buffer to absorb species that diffuse from the capillaries that incorporate throughout the chip. Once the capillary bed has been established, diffusional exchange occurs between the capillaries and the sensor array in the chip via the gel matrix. This buffering ameliorates transient noise in the signal. Thirdly, the gel matrix may also incorporate biosensing molecules/proteins that, on binding the primary species of interest, produce secondary species that can be detected by the on-chip electronic sensors.

The area for exchange is determined by the number of sensor matrix layers in the chip, and the number of capillary holes in the chip. The chip can alternatively consist of one layer, incorporating the sensor electronics, power and communication systems. The number of capillary holes would be scaled up to increase diffusional exchange with the biosensor molecules. Real-time detection is ensured as long as every part of the biosensor matrix is within $\sim$10 - 20 μm of a capillary.

3.4.2 Stabilisation and Integration of *in situ* Wireless Neural Interfaces

The technique described in Chapter 3.4.1 may possibly be harnessed for microvascularisation of implantable neural interfaces. Two key problems faced by current long-term implantable microelectrodes are movement of the electrodes in the brain tissue, and the immune response to the implant.

Movement of the implanted probe with respect to the brain tissue occurs as usually the implant is anchored at the skull and cemented in place, whilst the probe fenestrates the brain. Due to the blood circulation, the brain pulsates and thus moves relative to the skull. The brain is also surrounded by cerebrospinal fluid between the arachnoid mater and pia mater, which acts as a shock absorber for the normal stresses encountered with movement in an organism. Thus the brain also normally moves within this region. Despite being untethered, wireless implants that are implanted into the brain without attachment to the skull still move with respect to the surrounding brain as they are inevitably large in size, and constructed of 'hard' materials like silicon that are mechanically mismatched with brain tissue. These issues are being addressed in recent work exploring neural probes with micromechanical electrodes on the probe that can adjust position to account for movement [76]; externally wired electrodes that 'float' in the brain via a flexible interconnect [77]; neural interfaces constructed of flexible polymer substrates to better match the mechanical properties of the brain [78]; and porous and flexible neural interfaces that encourage growth of the brain tissue into the interface [79].

The immunological response is initially triggered by the damage to the brain tissue itself, especially the capillaries and arterioles, as the leakage of blood into the tissue elicits an inflammatory reaction. Longer term, surface chemistry and morphology of the implant are important, and implants eventually suffer the fate of glial encapsulation.

The idea for a vascularised chip, in combination with a highly dense and wireless neural probe as we propose in Chapter 2 and 4, suggests an alternate approach to building a biotolerant long-term neural interface. Biocompatibility could be promoted by making the chip *small*, which would reduce the mechanical discontinuity with the surrounding brain tissue; and if it quickly promotes vascularisation into the chip, this would mechanically anchor it to the brain, and impede glial encapsulation.

If the chip is made small enough, we could envision the following. A small piezophotonic nanomechanical neural interface, a millimetre square, contains ~ 1000 microelectrodes, and 20 μm holes around it's perimeter with access to a hydrogel cavity loaded with angiogenic factors. Brodner & Markowitz demonstrated that intracranial surgery on foetal rats is possible in utero [80]. Such surgery is performed to insert the chips into the brain of a rat foetus. As the foetal brain is still plastic, after the chip has vascularised in the brain tissue, the brain itself models around the implant, accepting it as 'self'. Once the foetus is brought to term and reared into an adult, further surgery can then be done to interface a $\sim$ 10 μm wide single-mode optical fibre into the rat's brain to optically interface with the chip and begin downloading electrophysiological measurements. There is minimal damage to the brain tissue, and the neural probe the fibre interfaces with is surrounded by normal, healthy brain tissue. Such techniques could be a route for performing long-term *in vivo* electrophysiology without damaging the brain one is trying record form.

Recent work has also shown the feasibility of integrating nanophotonics, III-V semiconductors and optoelectronics into a flexible polymer based substrate [81]. Combining this engineering with the ideas presented here could lead to truly large-scale, biocompatible and tissue integrated neural interfaces that could contribute to our understanding of the brain and have application in neuroprosthetics.

3.4.3 Fabrication

We attempted to fabricate angiogenesis promoting chips for *in vivo* testing. Our starting point was SOI with a 100 μm device layer and 1 μm BOX layer. Fig. 3.16 depicts our mask design; chips 3 mm wide were populated over a full wafer. Each chip consists of a central region 1.5 mm wide, with holes 30 μm wide, spaced 90 μm apart. The outer boundary of the chip in purple defines a raised area, which is aligned and bonded to a second 100 μm wafer with holes in it as well. On bonding, a cavity then exists in the area of the holes, which is supported by the bonded areas.

A brightfield mask was ordered for the hole pattern and darkfield mask for the boundary pattern. The SOI wafer was spun with AZ9245 positive resist, and lithography was performed to define the hole pattern on the wafer. This created raised islands of resist where the holes should be placed.

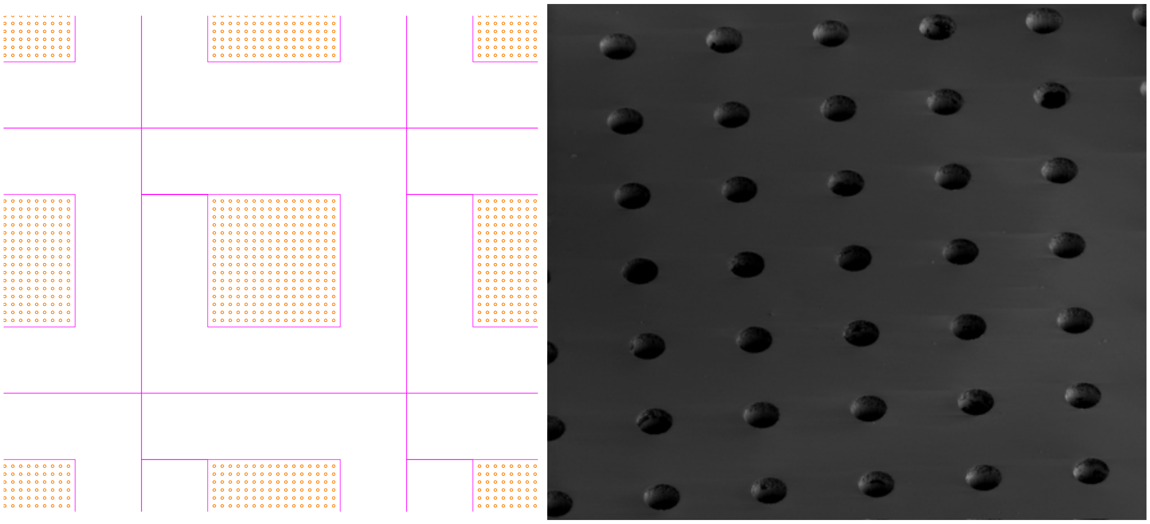


Figure 3.16: **Vascular interface chip fabricated by DRIE.** Depicted is the mask, with hole pattern in orange and boundary pattern in purple. The micrograph depicts successful etching of the through holes on the SOI wafer using cryoetch.

Alumina was then sputtered onto the developed resist. Lift-off was done in acetone under sonication. This created an alumina mask with fenestrations in it corresponding to the hole pattern. The SOI wafer was then placed in an Oxford system 100 ICP where a cryoetch was performed to etch holes through the entire thickness of the 100 μm device layer. The wafer was over-etched slightly so that the cryoetch process stopped on the BOX layer. Another etch step was performed in the ICP to drill through the BOX layer, using the alumina friendly C_4F_8 / SF_6 silicon oxide etch recipe.

The alumina mask was removed in phosphoric acid at 60°C, and AZ9245 resist was spun on and baked. This time, the boundary darkfield pattern was aligned to the hole patterns on the wafer, and photolithography performed. Alumina was deposited onto the wafer and then lifted off. A Bosch process was then run on the wafer in the Oxford ICP to etch away precisely 50 μm of silicon in the holes area of the chips. The alumina mask was then removed as before. The micrograph in Fig. 3.16 depicts the result at this step.

The wafer was briefly dipped in Transene BOE improved to remove native oxide. Subsequently, the wafer was placed in an oxidation furnace at 1000°C where a wet oxidation recipe was run to grow 1 μm of thermal oxide.

A wafer of ultrathin silicon, 100 μm thick, was mounted on a carrier wafer with Santovac 5. AZ9245 resist was likewise processed into the wafer, and in the same manner, the holes pattern was used to drill holes into the wafer with the alumina mask and cryoetch process. After a dip in Transene BOE and rinse/dry, the wafer was likewise oxidised in the furnace to form 1 μm of thermal oxide.

The intention was then to wafer bond the two wafers together using an oxide-oxide bonding process, and then to use a dicing saw to precisely dice out the 3 mm chips. They were then to be loaded with hydrogel mixed with PDGF-BB and FGF-2, which would wick into the cavity before the hydrogel polymerised. Vascularisation was then to be studied after implantation in rats.

At the time these chips were fabricated, the wafer bonding process had not been worked out, and access to a dicing saw was not available. Subsequently, the bonding recipe was perfected for the piezophotonic chips described in Chapter 4, for which a dicing saw was also used. We intend to return to the fabrication of these chips at a later date now that the process flow can be completed.

Chapter 4

Wiring the Brain with Piezophotonic Nano-Optomechanical Resonators

> In things to be seen at once, much variety makes confusion, another vice of beauty. In things that are not seen at once, and have no respect one to another, great variety is commendable, provided this variety transgress not the rules of optics and geometry.
>
> *Sir Christopher Wren, 1632-1723*

4.1 Fibreoptic Communication with Neural Systems

In Chapter 2, we described how it is possible to multiplex the voltage signals from a large number of analogue sensor channels onto a single communications line, by transferring the signals to frequency space. We did so by multiplexing the signals across the *radiofrequency* spectrum, this essentially being frequency division multiplexing (FDM). Significantly, we employed an optical channel as the output line, instead of an electrical one. Until the late 1980's, transcontinental communications were based on FDM on electrical lines (coaxial cables), in the RF band. However, electrical lines are bandwidth limited; the last coaxial transatlantic cable (TAT-7) had a capacity of $\sim$ 50 Mbit/s. The demand for increased communications bandwidth ushered in the switch to using *optical* communications lines, with massively increased bandwidth. The TAT-14 transatlantic fibre optic cable has a capacity of 3.2 Tbit/s, and new technology will soon push this to 16 Tbit/s. Modern fibre optic communications achieve such massive information capacities by combining sophisticated forms of

FDM (in the digital domain) with *wavelength division multiplexing* (WDM). Essentially, the bandwidth of FDM is multiplied by WDM. Our first iteration of FDM in the optical domain presented in Chapter 2 achieved frequency multiplexing at only one optical wavelength. This means the optical information capacity in our system is virtually unused. Our question is: is it possible to improve our neural multiplexing scheme to make full use of the optical bandwidth, and furthermore, is there a way we can do so in an integrated fashion? Can we harness the enormous strides made in optical telecommunications to detect the spiking activity of every neuron in the mammalian cortex, and bring 'fibre to the brain'? [1]

First, let us actually quantify what the required information capacity would be to record the extracellular spiking activity from every neuron in the human brain. If we assume 8.6×10^{10} neurons, a recording bandwidth of 3×10^4 Hz and a bit depth of 16, then the bit rate B is given by: $B = 8.6 \times 10^{10} \times 3 \times 10^4 \times 16 \approx 4 \times 10^{16}\ \mathrm{bit/s} = 40\ \mathrm{petabit/s}$.

Long distance fibre optic communications are driven by a very different set of engineering constraints than would be faced by a fibre optic neural interface. Optical fibres suffer from the effect of dispersion over long distances, and one of the solutions used to counter this is to use laser light over the range where group velocity dispersion is zero (1300-1500 nm). So only a restricted bandwidth of light is used. In addition, frequency modulated data is not sent in analogue form, but digitised (with elaborate modulation schemes). In a neural interface, the constraints would flip. The optical fibre is only traveling a distance of $\sim$ 10 cm, and dispersion is irrelevant. Thus the full optical bandwidth of say 600-2000 nm may be used. Furthermore, frequency modulated data would have to be sent in analogue form, as any kind of digitisation would not be possible at the sensor level. The theoretical capacity of a single-mode fibre is $\sim 1\ \mathrm{petabit/s}$ [82]. This suggests the entire spiking output of the brain could be transduced through just 20-40 single-mode fibers, each being 8 μm in diameter. Those fibers could couple to and branch off larger multimode fibers ($\sim$ 50 μm diameter), and the architecture would be similar to the parallel supply of the brain's vascular system, as described in Chapter 1.3.

The answer to coupling the frequency multiplexing afforded by piezoelectric NEMS to the massive bandwidth of optical communications is the integration with nanophotonics. In 2005, *Povinelli et. al.* theoretically predicted that light traveling through a nanophotonic waveguide could exert piconewtons of force on a neighbouring waveguide through the *gradient optical force* [83], this be-

---

[1] *Fibre to the x* is a generic term for the local delivery of fibre optic communications lines the last mile to their final destination, from the world wide network of cables. e.g. fibre to the home, fibre to the desktop etc. It was my advisor Axel Scherer who coined the term 'fibre to the brain'.

ing large enough to have substantive effect on devices of that mass. *Li et. al.* were the first to use this theoretical prediction to build a practical device. The configuration they fabricated was similar to our own NEMS resonators, with a planar silicon NEMS waveguide suspended over an etched trough of silicon dioxide.

This work is of great significance for our purposes. It means that actuation of the NEMS devices can be separated from their electrical detection functionality, which will vastly improve the noise floor of detection. The measurement of the actuation can also be done in integrated fashion optically, and be independent of actuation and electrical sensing.

However, the device by *Li et. al.* and others reported in the literature are not capable of electrical measurement or actuation. They are crafted from SOI, and as pure silicon devices, are solely optical. Here we propose a new class of photonic NEMS devices we name *piezophotonic nanoelectromechanical systems*. We propose using a piezoelectric pin layer for waveguiding instead of silicon, and making a photonic nanomechanical system that is actuated and probed optically, but that can be tuned electrically via the piezoelectric effect. Such a device would permit a massive bandwidth increase in the potential multiplexing capability of a neural interface. A piezophotonic NEMS device would be able to be coupled with a nanophotonic resonator to form a coupled mechanical-optical resonator, as in cavity optomechanics. Instead of relying on scattering forces for the actuation however, the gradient optical force would actuate the devices. Clearly, it would be possible to make self-resonant devices that mechanically resonate at their natural frequency by injecting high intensity light into such a coupled photonic-NEMS oscillator. Self-resonant oscillators would vastly simplify the operation of these devices and their scalability, and furthermore, would exhibit linewidth narrowing as they are no longer being forced by an external oscillator (thus increasing operating bandwidth). The system carries over very well to our vision of multiplying the FDM by WDM bandwidth for neural multiplexing. The self-resonant coupled system would be electrically tuneable with attached microelectrodes, and that tuneability would be exhibited as a shift in RF frequency. Thus such a system would encode signals via *frequency modulation* instead of amplitude modulation as in Chapter 2, and would need a phase-locked loop to demodulate after optical filtering. [2]

---

[2] In general, we can describe modulation on a signal carrier by,

$$V(t) = A(t)\sin[\omega(t)t + \phi(t)]$$

where A is the carrier amplitude, ω is the carrier frequency and ϕ is the phase offset. The radiofrequency carrier signal ω can be modulated by varying either A, ω or ϕ slowly in time with respect to the carrier frequency. Varying $A(t)$ is termed *amplitude modulation* (AM), and this is how we modulate our carrier in piezoelectric NEMS as described in Chapter 2.

Let us conservatively estimate the potential channel bandwidth of a neural interface based on self-resonant piezophotonic NEMS employing FDM and WDM. We assume the channel bandwidth of 10,000 for the mechanical resonators, as estimated in Chapter 2.8. If WDM is employed in the band 600-2000 nm with 100 pm resolution, this equates to a system capable of carrying $\sim$ 150 million channels over several multimode fibers. Under more optimistic assumptions, which would be at the limits of what may be possible, an ultimate capacity of about a billion channels is predicted.

4.2 The Gradient Optical Force

Radiation pressure has two components of distinct physical origin, which we will calculate from Maxwell's equations. The combination of Maxwell's equations with Lorentz's force law encompasses the entirety of electromagnetic theory, and from it can be derived the following relation, which is an expression of the law of conservation of momentum in classical electrodynamics:

$$\mathbf{f} + \epsilon_{\mathbf{0}}\mu_{\mathbf{0}}\frac{\partial \mathbf{S}}{\partial \mathbf{t}} = \nabla \cdot \sigma$$

Here $\mathbf{S}$ is the Poynting vector, $\mathbf{f}$ is the force density and σ is the Maxwell stress tensor, which is given by:

$$\sigma_{ij} \equiv \epsilon_0 \left(E_i E_j - \frac{1}{2}\delta_{ij}E^2 \right) + \frac{1}{\mu_0}\left(B_i B_j - \frac{1}{2}\delta_{ij}B^2 \right)$$

Let us integrate the force density with respect to volume:

$$\mathbf{F} = \int \mathbf{f}\ \mathbf{dV} = \int_{\mathbf{V}} (\nabla \cdot \sigma)\mathbf{dV} - \epsilon_{\mathbf{0}}\mu_{\mathbf{0}} \int \frac{\partial \mathbf{S}}{\partial \mathbf{t}}\mathbf{dV}$$

$$= \int_S \sigma \cdot \mathbf{d\tilde{n}} - \epsilon_{\mathbf{0}}\mu_{\mathbf{0}} \int \frac{\partial \mathbf{S}}{\partial \mathbf{t}}\mathbf{dV}$$

The force has two components. The second is the scattering force, which is what propels solar sails, pushes the tails of comets, and is what is used in most cavity optomechanics systems to actuate the devices.

The first term is the gradient optical force and what we harness in our devices:

Varying $\omega(t)$ is known as *frequency modulation* (FM), and modulating the phase offset $\phi(t)$ is called *phase modulation* (PM). Piezophotonic NEMS allow us to employ FM and PM modulation schemes, which are more resistant to noise.

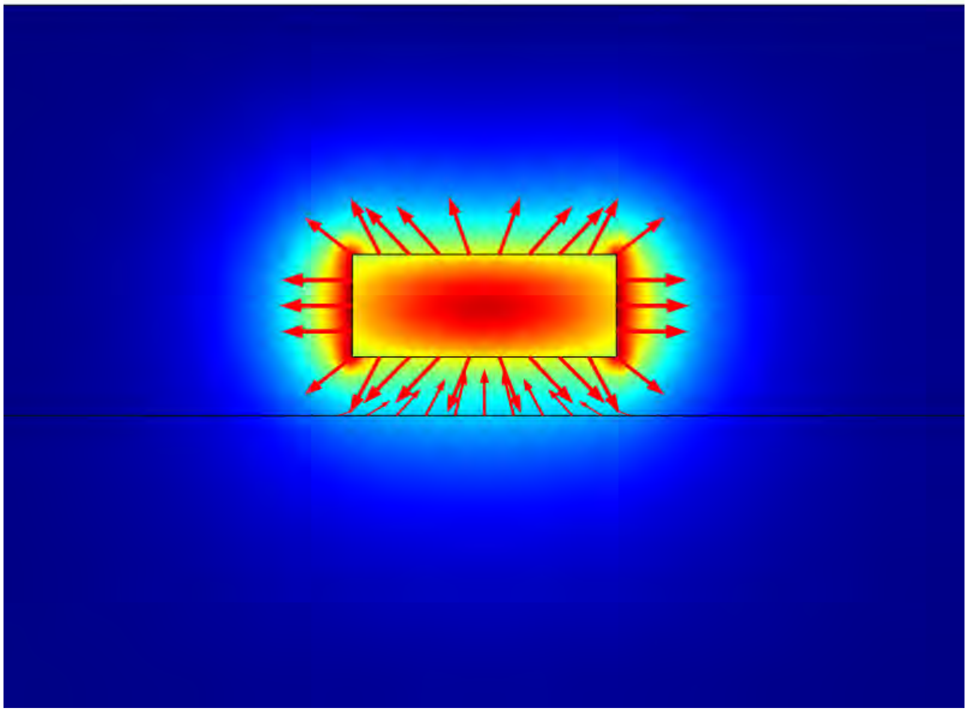


Figure 4.1: **Simulation of the optical gradient force arising from a GaAs waveguide suspended over $\mathrm{SiO_2}$ substrate.** The GaAs waveguide has width 450 nm and thickness 175 nm, and is suspended 100 nm over the silicon dioxide slab, with air as cladding. The first order mode is computed, and it can be seen to bleed into the substrate causing an asymmetric field gradient transverse to the waveguide. The Maxwell stress tensor is also plotted at the surface boundaries (logarithmic weighting), indicating the optical gradient force that arises in this system. Simulation was performed by the author in COMSOL 5.0 using the Electromagnetic Wave, Frequency Domain module.

$$F_{\text{grad}} = \int_S \sigma \cdot \mathbf{d\tilde{n}}$$

It's origin is the polarisation of dielectric atoms in an electromagnetic field to form dipoles. If the field has a gradient, the dipoles will experience a force towards the strongest region of the field. As can be seen from the force equation, full knowledge of the Maxwell stress tensor can then be used to compute the force experienced by the dipoles in the field numerically (although an analytical formulation has also been derived [84]). Fig. 4.1 depicts a COMSOL simulation of a 450 nm wide GaAs waveguide suspended 100 nm over a $\mathrm{SiO_2}$ substrate. The first mode present in the waveguide leaks into the underlying substrate, making a difference in the effective refractive index over the top surface and under the bottom surface of the waveguide. The surface integral of the Maxwell stress tensor is thus positive in the downwards direction.

Povinelli *et. al.* suggested a more straightforward method to compute the optical gradient force. Their argument was a shift in the gap g between the waveguide and substrate would shift the eigenfrequency of the system. Due to conservation of energy, that eigenfrequency shift must be equated into the work done by the waveguide to move down. In this way, the force can be calculated. If we assume $U = N\hbar\omega$ is coupled into the eigenfrequency (i.e. it is the energy in the suspended beam), then:

$$F_{\text{grad}} = -\frac{\partial U}{\partial g} = -\frac{\partial N\hbar\omega}{\partial g}$$

$$= -N\hbar\frac{\partial \omega}{\partial g} = -\frac{1}{\omega}\frac{\partial \omega}{\partial g}U$$

Thus knowledge of the change in effective index ($\omega = \omega_0/n_{\text{eff}}$) will lead to calculation of the force. As with the Maxwell stress tensor, this is usually computed numerically, although an analytical derivation has been carried out [63].

4.3 Device Design

Our goal was to design a piezophotonic nanomechanical resonator, where a piezoelectric material acts as the waveguiding medium, and to which electrical contacts are also made to enable tuning of the mechanical resonance of the NEMS device whilst it is optically guiding light. The first issue was to decide what material to use as the piezoelectric *pin* structure. The table below summarises some important physical properties of GaAs, GaN and AlN that are relevant to us. Note that n is calculated at 1550 nm.

	Y	ρ	n	e_{14}	d_{31}	E_g
GaAs	118 GPa	5.32 gcm^{-3}	3.38	0.16 Cm^{-2}	1.4 pmV^{-1}	1.43 eV
GaN	398 GPa	6.15 gcm^{-3}	2.32	0.65 Cm^{-2}	1.7 pmV^{-1}	3.4 eV
AlN	390 GPa	3.26 gcm^{-3}	2.12	1.55 Cm^{-2}	2.4 pmV^{-1}	6.03 eV

Referring to equation 2.1, one can see the most important parameter for electrical actuation efficiency is the transverse piezoelectric coefficient d_{31}. GaN and AlN are much stiffer materials than GaAs, but the actuation efficiency only depends on the square root of Y/ρ. AlN would seem to be the clear winner, ticking off almost every box with its superior properties. Unfortunately, it has the lowest refractive index, and for our photonic application, this parameter is actually the most important. We thus chose GaAs, especially as its refractive index is almost equal to silicon. This allows us to use silicon dioxide as a cladding material to GaAs.

The basic topology of our device is shown in Fig. 4.2. It consists of a free standing NEMS waveguide in the centre, supported by multimode interference (MMI) converters and silicon dioxide cladding. The MMI converters serve another important function other than to mechanically support the NEMS resonator to maximise Q. They are also there to allow electrical access to the waveguiding material, without affecting the guiding of light. Electrical lines 150 nm wide, below the guiding cutoff, are connected to the MMI converter at the midpoint where there is destructive interference and minimal field intensity. Each line impinges on a wirebonding pad for electrical interfacing with the outside world. One pad connects to the p layer at the top of the stack, whilst the other is etched down into the material to reach the n doped layer for a connection.

The device is actuated via the optical gradient force, and is brought to resonance by oscillating the intensity of the guided light. The actuation is implemented with a pump laser at a frequency in the range 1530-1570 nm. Another laser, also in the range 1530-1570 nm, but at a different frequency to the pump, is called the probe, and acts to measure the mechanical motion of the NEMS device. The physics of how this occurs is actually extremely rich, as a lot of interesting effects occur in this type of material under these conditions. The most dominant effect however is going to be stress-induced birefringence. This is a phenomena whereby the polarisation dependence of the refractive index changes with stress. With our waveguides, we use polarised light in the TE-mode, since this is optimal with waveguides like ours that are thin (out of necessity), but wide. As the beam is stressed however, the refractive index of the beam will change for the TE polarised beam. This will effectively alter the intensity of light transmitted through the waveguide. Thus, at the output, a modulation pattern will be detected on the probe due to the mechanical oscillation of the NEMS beam, even though the probe was unmodulated before being fed to the waveguide.

The thickness of our waveguide material is 175 nm. Simulations in RSoft confirmed that the optimal waveguide width with this thickness is $w_w = 450$ nm, assuming silicon dioxide cladding with refractive index 1.44.

Fig. 4.3 depicts the chip scale configuration of the devices. They are arranged on a 9 mm × 17 mm chip, with the NEMS resonators aligned in the [110] direction, where the piezoelectric coefficient is optimal. Light is guided into and out of the chip with an edge coupler called a nanotaper [85]. This is a compact coupler that converts the mode size and effective refractive index of the chip waveguide into that of an optical fibre brought into close proximity. It works by delocalising the

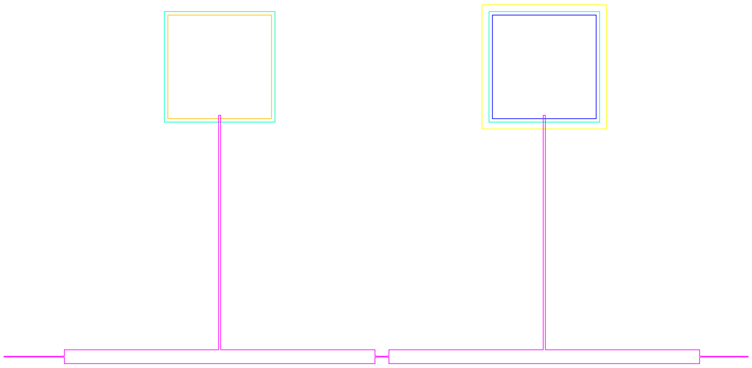


Figure 4.2: **Design of piezophotonic NEMS resonator.** The basic topology consists of a nanomechanical resonator supported by a pair of MMI converters, which improve the Q of the resonance, and provide electrical access to the NEMS device without disturbing its photonic propagation. This construct can be expanded into a number of different topologies, including circuit designs to demonstrate frequency multiplexing, wavelength division multiplexing, and cavity optomechanics configurations where the construct is part of a nanophotonic resonator making a self-resonant coupled optical and mechanical resonator.

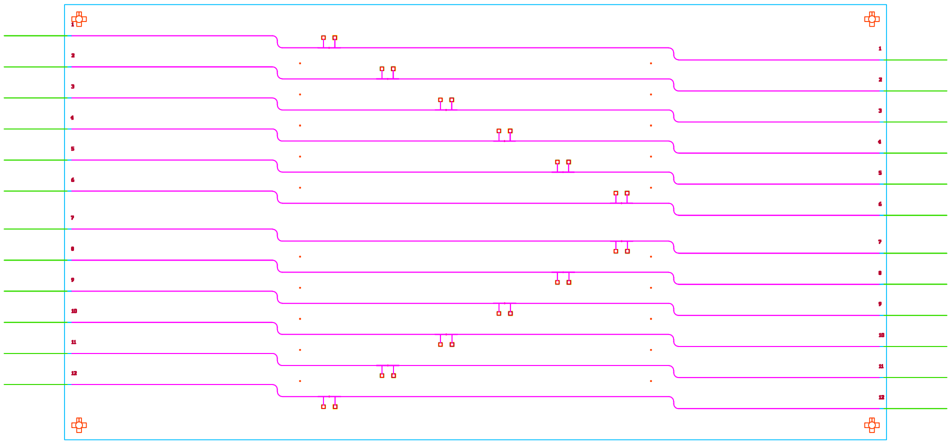


Figure 4.3: **Piezophotonic NEMS chip layout.** Shown in purple are the waveguide sections, whilst the nanotapers are in green. The blue outline is the etched outline of the final chip. Multiple NEMS resonator operating frequencies are populated on the chip to test actuation efficiency across the RF spectrum, and to demonstrate frequency multiplexing.

mode field in the waveguide into the surrounding cladding, where the size of the mode becomes matched with that of the fibre. In addition, since the mode now mostly resides in the cladding, the effective refractive index becomes matched with the optical fibre. These effects maximise coupling and minimise backreflections.

Fig. 4.4 depicts a simulation in RSoft of the nanotaper we designed for our chip. As can be seen, the waveguide mode becomes delocalised into the cladding at the nanotaper, expanding to 2 μm to match the fibre mode size. The design process for the nanotaper is straightforward. The waveguide width w_w is narrowed down after a length l to a width w_t. Over the length l, a parabolic curve connects the two widths. Thus the taper design is a matter of simulating w_t and l for a given w_w to find the optimal parameters for the system. In our case, this was: $w_t = 150$ nm and $l = 100$ μm.

A simulation of the computed dimensions of the multimode interference converters for the 1×1 and 1×2 topology are shown in Fig. 4.5. The 1×2 MMI is used to implement frequency multiplexing in our photonic circuits, by allowing us to split a single optical interconnect from a single nanotaper into two paths for two different RF frequency NEMS resonators, which are then recombined back into a single path with another 2×1 MMI converter.

The basic dimensions of our MMI converters were computed using guided-mode propagation analysis [86]. For the 1×1 case we obtained for the length and width: $L = 224.518$ μm and $W = 10$ μm. However, a basic design like this only operates at a single frequency, here 1550 nm. The design was modified to impart it with 40 nm of bandwidth. This was done by adding simple taper sections between the coupling waveguides and the MMI converter. The 450 nm wide waveguide was expanded into a 900 nm mouth at the MMI converter in linear fashion at both ends, over an axial distance of 3 μm.

The 1×2 MMI converter dimensions were: $L = 112.259$ μm, $W = 10$ μm and $D = 2.5$ μm, where D denotes the distance from the converter axis of symmetry of each of the output ports. The same taper design was implemented here, giving a similar increase in operational bandwidth.

4.4 Experimental Design and Fabrication

The experiment described here combines the techniques in Chapter 2.7 with optoelectronics, photonics and *in vivo* mammalian recording, and is consequently a lot more complex (see Fig. 4.5). Most of the experiment had to be fabricated from scratch, and with great enterprise, some very expensive pieces of equipment needed to be acquired and assembled.

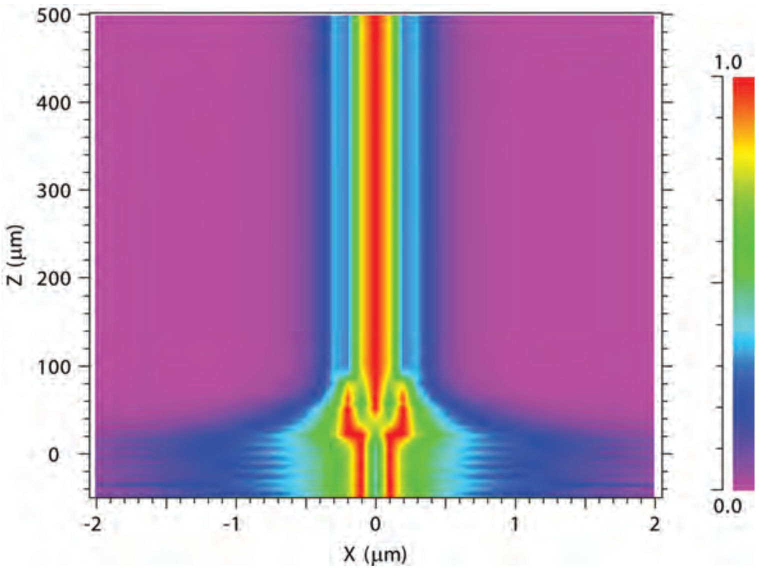


Figure 4.4: **Nanotaper design simulation.** Simulation performed in RSoft Beamprop. The nanotaper was optimised for a coupling mode size of 2 μm, to match our taper lensed fibres. Dimensions of the tip were $w_t = 150$ nm, and the length $l = 100$ μm. The waveguide width $w_w = 450$ nm.

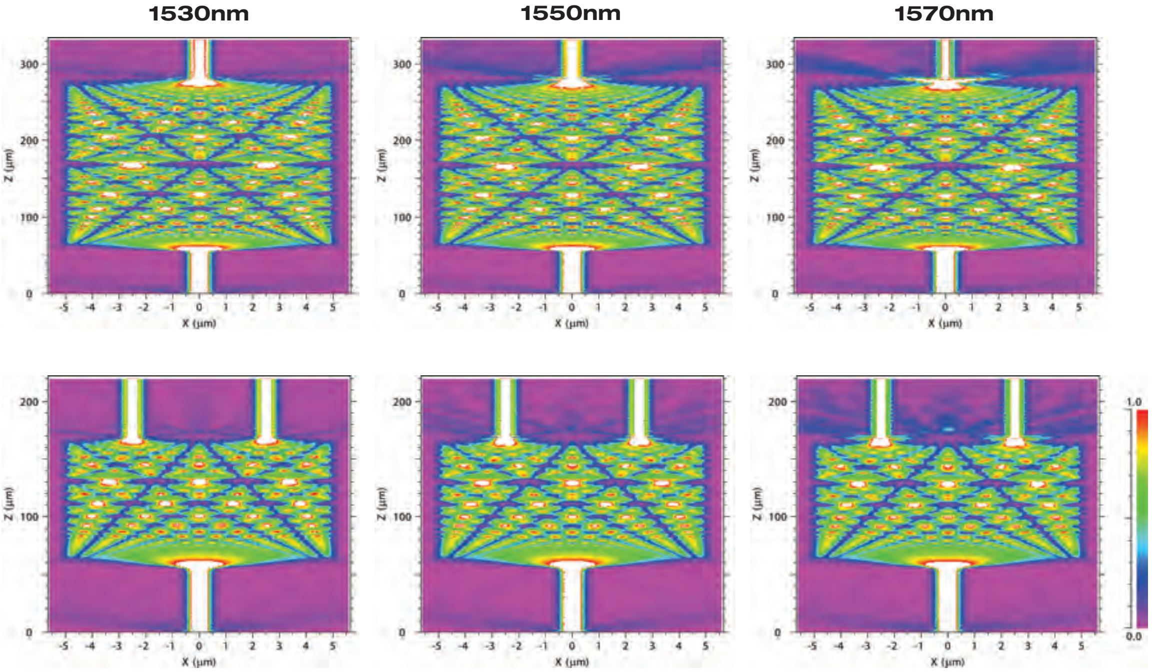


Figure 4.5: **Multimode interference converter design simulation.** The basic computed MMI converter parameters were modified by adding a taper section at the entrance and exit mouths of the 450 nm waveguides to the converter. Adding this taper effectively expanded the bandwidth over which the devices work to ~40 nm. Without the tapers, the devices only operate correctly at a single frequency. Simulations were performed in RSoft Beamprop.

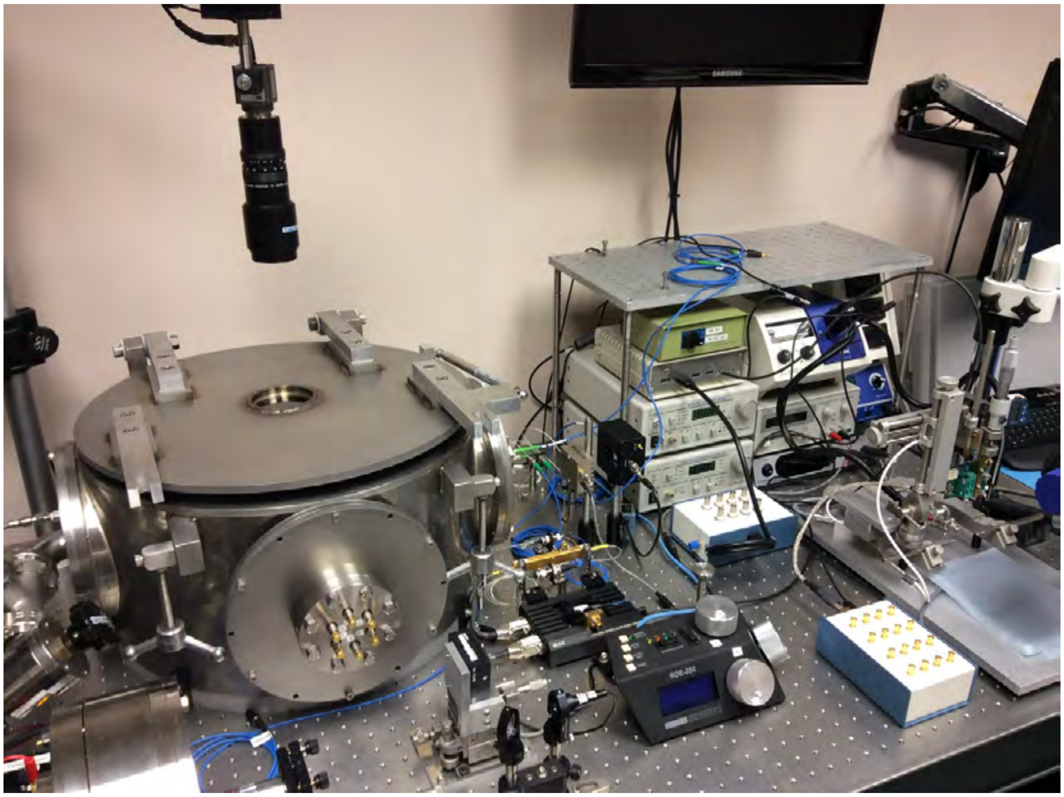


Figure 4.6: **Piezophotonic NEMS experimental setup.**

The basic outline of the setup is an RF electrical and RF optical interface between the NOEMS device described in Chapter 4.3 and the rest of the experiment. Since the NOEMS resonator oscillates mechanically as well as optically, it requires an excellent vacuum to attain a high Q factor, and to minimise noise from molecular adsorption. Thus the entire setup to interface with the chip needs to be housed in a vacuum chamber, and the chamber requires optical, RF and electrical feedthroughs to complete their connection to the experiment.

Pump and probe lasers are routed through the single-mode fibre feedthrough via a splitter / combiner, to interface with the chip's mode converter via a single optical fibre terminated in a taper lensed fibre. The output from the chip is routed out of the vacuum chamber through an identical converter and fibre and interfaced with a narrowband optical filter to isolate the probe signal, which is then connected to a fibre-coupled photodetector. The pump laser is modulated in the RF regime using an electro-optic modulator connected to a signal generator. The probe laser is tuneable. The electrical feedthroughs on the vacuum chamber, connected to the piezoelectric *pin* structure of the NOEMS devices through the ZIF interface, is connected to a breakout box wired directly to the neural probe electrodes, which record *in vivo* by use of a stereotactic stage. The electrodes are also concurrently connected to a low-noise digitising headstage amplifier, which routes out signals via a USB interface box to a computer. The signal from the photodetector is connected to a lock-in amplifier, where the demodulated signal is concurrently recorded using a DAQ card, which is synchronised to the extracellular amplifier interface. In the case of multiplexing, the same photodetector output is connected to multiple lock-in amplifiers for concurrent demodulation and recording by the DAQ.

This setup was designed to demodulate up to six NOEMS multiplexed signals concurrently, and record from up to 64 neural recording channels (upgradeable to 256).

4.4.1 Vacuum Chamber

The heart of this experiment is the vacuum chamber. It was not possible to purchase a vacuum chamber suitable for the experiment in mind, and so one was designed and fabricated by the author. Fig. 4.7 depicts the completed setup. The chamber needed to be of sufficient size to encompass the remote micromanipulators, used for adjusting the position of the taper fibres to the chip, as well as the stage for electrically interfacing with the chip. A sapphire window was designed into the top lid to enable real-time viewing with an infrared camera whilst adjusting the fibres into position. Once the 1550 nm light is successfully guided through the on-chip waveguides by accurately positioning the taper fibers with the manipulators, the scattered infrared emission can be detected by the camera, and thus it is a convenient method of doing an initial rough positioning adjustment. Feedthroughs were required to communicate with the manipulators, couple the single-mode fibres, as well as to route the electrical signals from the chip.

The final dimensions of the vacuum chamber were 19.5” width and 280 mm interior height. All parts were constructed of 304 stainless steel, except for the base, which was constructed of 6061 aluminium. The base was CNC machined to have an optical breadboard integrated into it, for easy affixation of the internal components. Conflat flanges were used for all feedthrough and vacuum connections, but the top lid was sealed with a nitrile o-ring. A gas piston was added to the heavy stainless steel lid to make loading and unloading the chamber safe and easy.

The chamber was constructed over a number of months using a water jet, CNC machine, lathe and milling machine. All stainless steel construction was mated and sealed through TIG welding the components. Many of the parts were also outsourced to a vendor to be planarised and electropolished to make them suitable for use in high vacuum.

Fig. 4.8 depicts the interior of the vacuum chamber, and the primary elements of the interface to

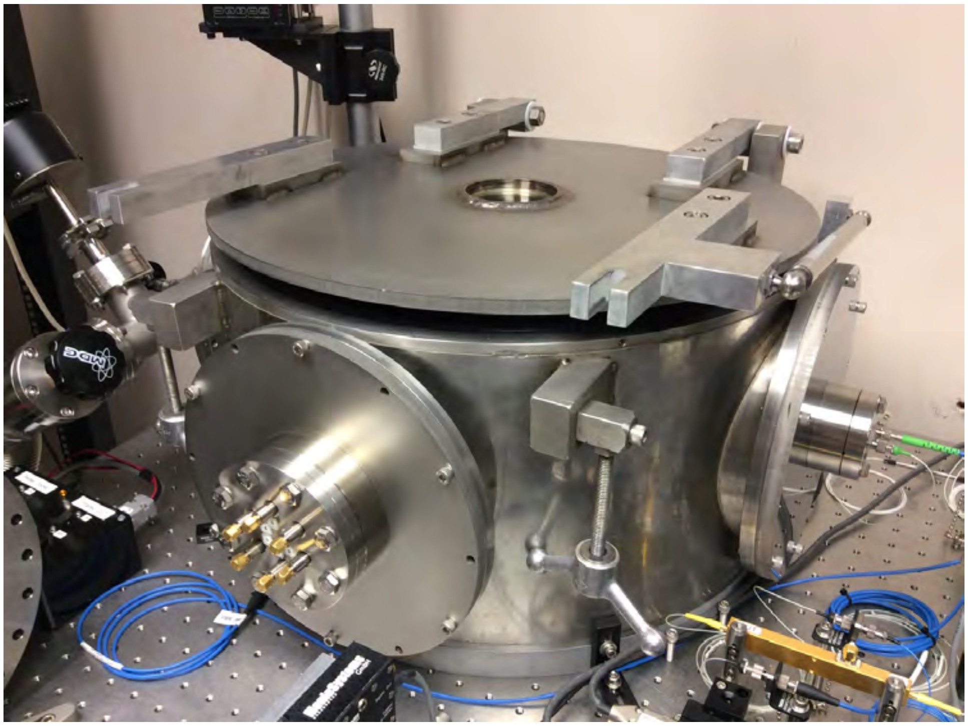


Figure 4.7: **Vacuum chamber constructed for piezophotonic NEMS experiment.** The chamber has SMA, single-mode FC/APC and DE-9 / DB-25 feedthroughs, as well as a sapphire viewport.

the NOEMS chip can be seen. A ZIF socket to accept the piezophotonic NEMS device is mounted on an interface board, that breaks out the electrical connections to SMA coaxial connectors. SMA connections were used for this, due to their compactness and good shielding. Experiments were also intended to be performed on the NEMS devices in the RF band, thus coaxial cables were a necessity. A Peltier cooler is mounted to the board with a specially machined copper gasket, and affixed and bolted using thermal compound. The cooler allows the chip temperature to be varied from approximately $-10°$C - $100°$C. Power is delivered through the DE-9 connector, which routes out signals from temperature sensors on the interface. Sutter MP285 and MP-265 micromanipulators are used to manoeuvre into position taper lensed optical fibres for edge coupling with the photonic chip via mode couplers designed into it. The remote manipulators are necessary as pumping to high vacuum causes a shift in positioning of the fibres, necessitating slight readjustment once vacuum has been achieved. This has to be done with the chamber sealed, thence the sapphire viewport and remote manipulators.

The SMA feedthroughs were machined from a blank CF flange, and CeramTec double-ended grounded SMA weld-in feedthroughs (15263-03-W) were TIG welded in. The single-mode fibre feedthrough was a two-channel FC/APC feedthrough purchased from MPF Products, Inc., designed

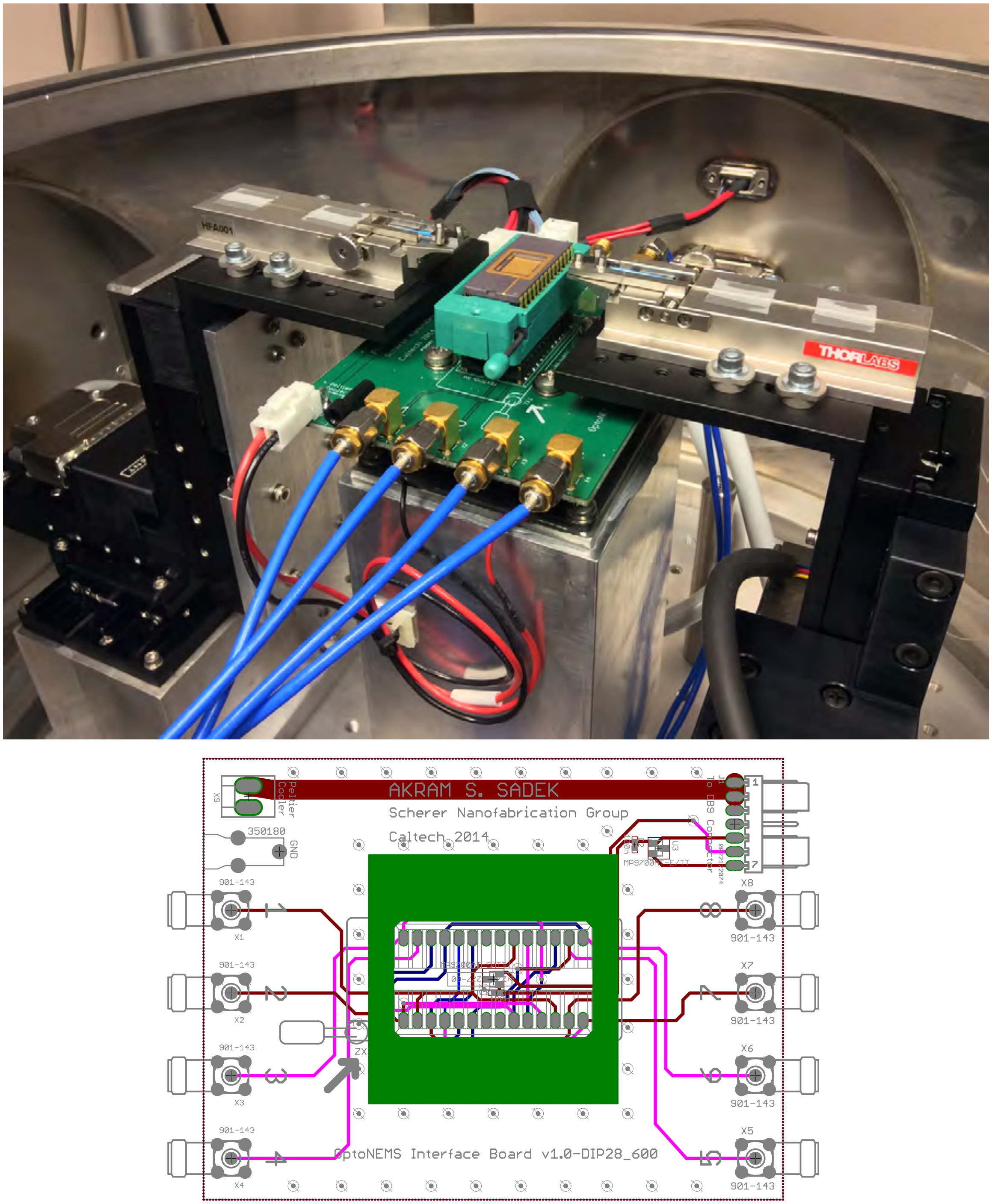


Figure 4.8: **Interior of vacuum chamber depicting chip optoelectronic interface.** Micromanipulators are used to manoeuvre taper lensed fibres into position at the end of the chip, to interface with the on-board mode converters. The ZIF socket interfaces the wirebonded electrical connections on the chip to SMA breakouts. A Peltier cooler interfaced with the mounting PCB allows the chip temperature to be varied from $-10°$C - $100°$C, with monitoring possible through integrated temperature sensors. The design of the interface PCB is also shown.

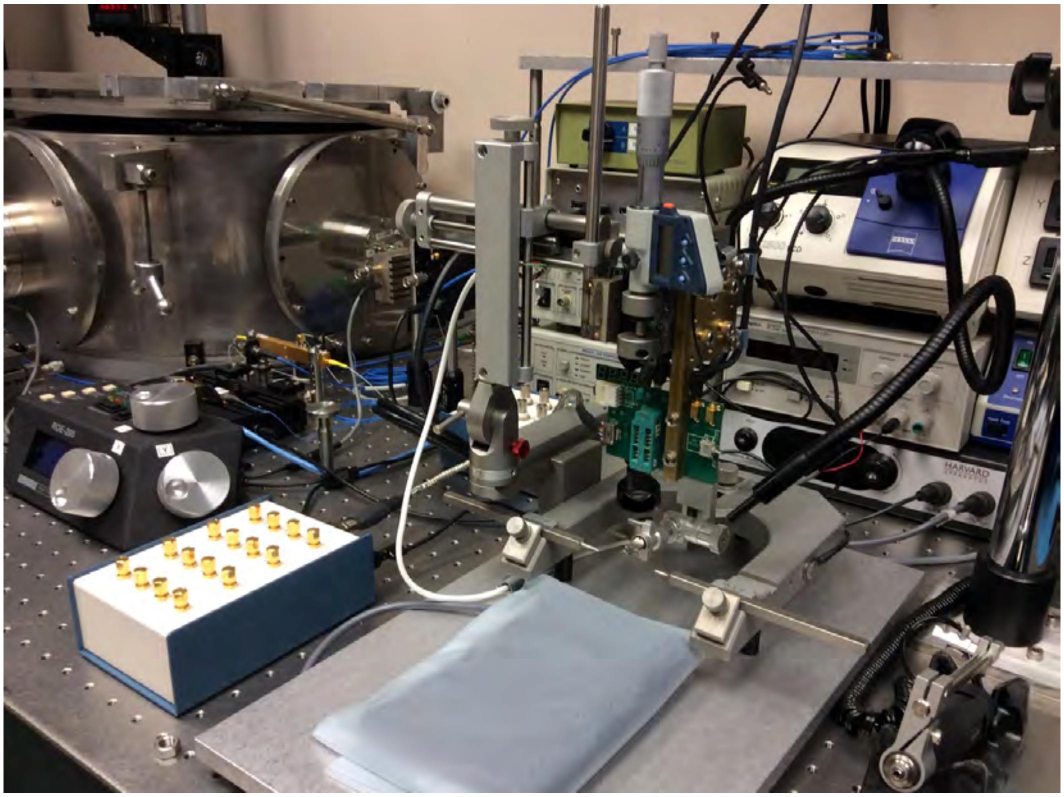


Figure 4.9: **Stereotactic setup for *in vivo* electrophysiology and concordant NOEMS interfacing and recording with a headstage amplifier.** A Kopf 940 stereotactic frame with digital readout is coupled with a micrometer stage bolted to a well-grounded 32-channel multiplexing amplifier headstage. A 16-channel SMA interface box for directly routing out electrode signals is connected to the headstage via a shielded cable. Any of the six SMA feedthroughs from the NOEMS vacuum chamber can be interfaced to any of the 16 SMA channels on the breakout box, through a 100 nF blocking capacitor.

to operate in the 1310-1550 nm bands (part number SM1310).

The vacuum chamber was only connected to a Welch 1397 roughing pump, which allows it to be pumped down to 10^{-4} Torr, but the chamber itself is capable of achieving at least 10^{-6} Torr if used with a turbopump or diffusion pump.

4.4.2 Electrophysiology

Fig. 4.9 shows the electrophysiological setup used to interface the NOEMS chips (and NEMS chips from Chapter 2) to our neural probes from Chapter 3 and a parallel multiplexing headstage amplifier, during small animal stereotactic cortical recordings. Our goal was to do recordings in mouse cortex with our fabricated neural probes, and stream the data with the headstage multiplexer to the computer for visualisation. Once several suitable spike-units were identified, a well shielded breakout box would be used to choose up to six of 16 recording channels for connection to the NOEMS device via capacitively coupled SMA connection using the 6-port SMA feedthrough on the vacuum chamber. Data would then be concurrently recorded by the amplifier and DAQ (connected to the lock-in amplifier output) for cross-validation against each other and confirmation that the NOEMS devices could multiplex and detect in real time the spiking units from multiple electrodes.

Since our setup was necessarily a custom one, we needed to design and construct our own electrophysiological recording system. Fortunately, an enterprising group of students at MIT designed their own USB interface for connecting Intan RHD digital multiplexer amplifiers to a computer for data acquisition. They made their design open source in service to the neuroscience community (Open Ephys). The design plans for the PCB were used to construct our own interface box, and this is shown in Fig. 4.10. The interface is capable of connecting up to four 64-channel multiplexers (for a total of 256 recording channels), but we only plan to use up to 64-channels in our experiments.

Next, a custom headstage was designed and built. This is detailed in Appendix B, and shown in Fig. 4.11. A well shielded PCB was used to interface an 18-DIL ZIF socket and 36-pin Omnetics connector to a 32-channel Intan RHD2132 digitising multiplexer chip. Special care was taken in grounding, with the reference being able to be switched in or out of the ground plane. The PCB ground was connected to a machined brass holder, which interfaced with a Newport micromanipulator stage. The translation stage was manipulated with a digital micrometer for accurate electrode placement within a few microns. Such a manipulation stage was required as the micromanipulators on the Kopf 940 stereotaxic frame were only accurate to about 100 μm resolution. Good grounding was made between all parts of the stereotaxic frame and the headstage PCB. A digital interface cable was constructed to connect the headstage to the Open Ephys USB interface box.

A special shielded breakout box was designed and constructed to tap into the microelectrode channels and allow external connection via grounded SMA cables. A shielded cable was constructed to connect the box to the headstage, using an Omnetics Bi-Lobe 21-pin connector. This is shown in Fig. 4.11 as well.

On testing the headstage amplifier, an RMS noise level of 1.5 μV was recorded with the inputs grounded. This noise level increased to 3 μV RMS when the breakout box was connected to the headstage.

Figure 4.10: **Open Ephys USB interface for using Intan RHD-class digital multiplexer headstages.** This design uses an FPGA to interface the Intan digital multiplexer chips via USB. A BNC connection at the rear was programmed in software to synchronise a pulse signal from the NI PXI-6259 DAQ used to acquire data from the lock-in amplifiers. All construction was done in-house.

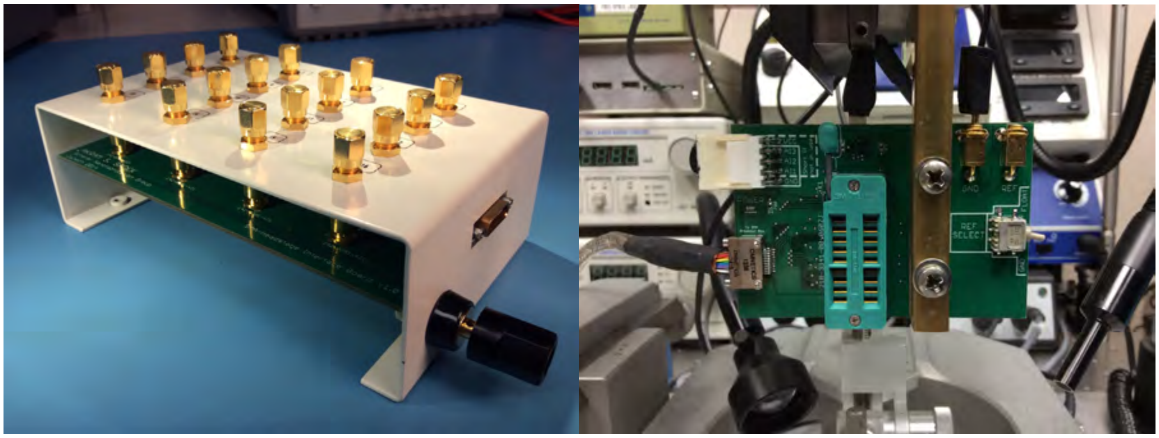


Figure 4.11: **Headstage amplifier and microelectrode breakout box.** The amplifier headstage noise was measured at 1.5 μV RMS with the inputs grounded. This increased to 3 μV RMS when the breakout box was connected to the headstage via the shielded cable.

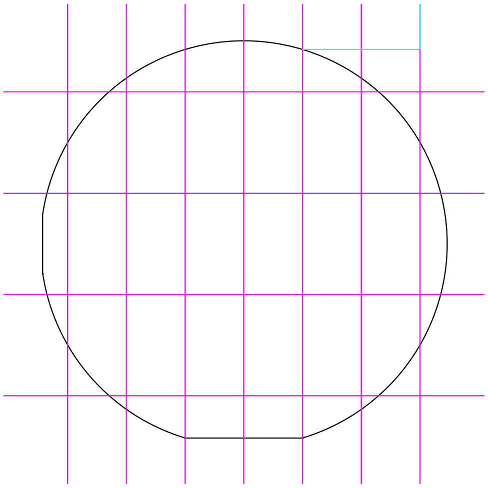


Figure 4.12: **Epitaxial wafer dicing map.** Chips were diced into dimensions 11 mm × 19 mm. Each 3” wafer yielded 12 chips for processing. Note the orientation with respect to the [110] primary flat.

4.5 Device Fabrication

The process flow for these devices was quite complex, and many independent fabrication techniques had to be developed and refined. It is assumed the reader has read through the detailed techniques of fabrication in Chapter 3. Appendix A should be referred to for details of the recipes mentioned.

Our starting point for the fabrication were 3” epitaxial GaAs wafers ordered from Xiamen Advanced Material Co., Ltd (Xiamen, China). Two material stacks were ordered, one for the NOEMS devices described here (175 nm epitaxial thickness) and another for fabricating even thinner NEMS devices (125 nm) as in Chapter 2. The wafers were fabricated using metal-organic chemical vapour deposition (MOCVD), and the primary flat of the wafer was oriented in the [110] direction. The stack layering was as follows:

	Diode			
Layer	NOEMS		NEMS	
n-GaAs	25 nm	10^{18} cm^{-3} (±5%, Si)	25 nm	10^{18} cm^{-3} (±5%, Si)
i-GaAs	100 nm	7×10^{15} cm^{-3} (±5%)	50 nm	7×10^{15} cm^{-3} (±5%)
p-GaAs	50 nm	10^{18} cm^{-3} (±5%, Be)	50 nm	10^{18} cm^{-3} (±5%, Be)
p-Al$_{0.8}$Ga$_{0.2}$As	500 nm	10^{18} cm^{-3} (±5%, Be)	500 nm	10^{18} cm^{-3} (±5%, Be)
p-GaAs	625 μm	10^{19} cm^{-3} (±5%, Be)	625 μm	10^{19} cm^{-3} (±5%, Be)

A 175 nm epitaxial wafer was spin coated with AZ9245 resist and then diced into chips 11 mm ×

19 mm in size, taking care to orient the [110] primary flat in the direction of the long axis of the chips. An ADT 7100 dicing saw was utilised, which produced edges with no chipping, and saw widths of just 50 μm. Fig. 4.12 illustrates the dicing map decided upon for the wafer to maximise yield.

Fig. 4.13 illustrates the process flow for the piezophotonic NEMS chips, which is outlined as follows.

One of the diced chips was cleaned in acetone / IPA / DI water, and then 20-20-80 RIE O_2 cleaned for 15 minutes. 100 nm of SiO_2 was then deposited on the chip by PECVD using silane-argon. This chip was subsequently annealed in the rapid thermal annealer at 250°C for 2 hours. The anneal step is extremely important, as it drives out undesirable gas molecules from the PECVD film before the wafer bonding step.

A 2” silicon wafer (650 μm thick) was thermally oxidised in an oxidation furnace at 1000°C and a wet oxidation recipe was employed to grow 2 μm of silicon dioxide.

Now the process is begun for bonding the substrates together [87, 88], such that the epitaxial surface of the chip is bonded to the surface of the oxide wafer in an upside down fashion. Our goal will be to backetch the non-pin layers of the GaAs chip, so that only the epitaxial layer remains bonded to the silicon oxide surface.

Both the silicon oxide wafer and epitaxial chip are loaded into the RIE concurrently, and a 20-20-80 O_2 plasma recipe is run on both substrates for 5 minutes. This makes their oxide surfaces reactive, and primes them for the next step.

They are then immersed in 0.025% HF solution for 1 minute, and then are immediately spun dry. This step forms a fluorinated oxide network on the surface. A 0.025% HF solution may be obtained by adding 0.5 mL of 6:1 BOE to 140 mL of DI water.

The wafers are finally dipped in 29% (standard) NH_4OH solution and spun dry. The chip and wafer are now ready for mating.

Great care must be taken in mating the surfaces, as once they come in contact, they bond instantly and will only come apart again with great force. If they are pried apart, the silicon oxide on both substrates must be stripped off and the fabrication process restarted. A vacuum pick-up tool is used to hold the epitaxial chip from the bottom side. The chip is then carefully manipulated into to placed to mate the top epitaxial surface to the silicon oxide wafer in a centered, aligned fashion.

A wafer bonder is now used to finalise the bonding process. We use a Suss SB6L wafer bonder, and process the chip at 300°C for 3 hours. The tool pressure used is 600 N of force per chip. For the

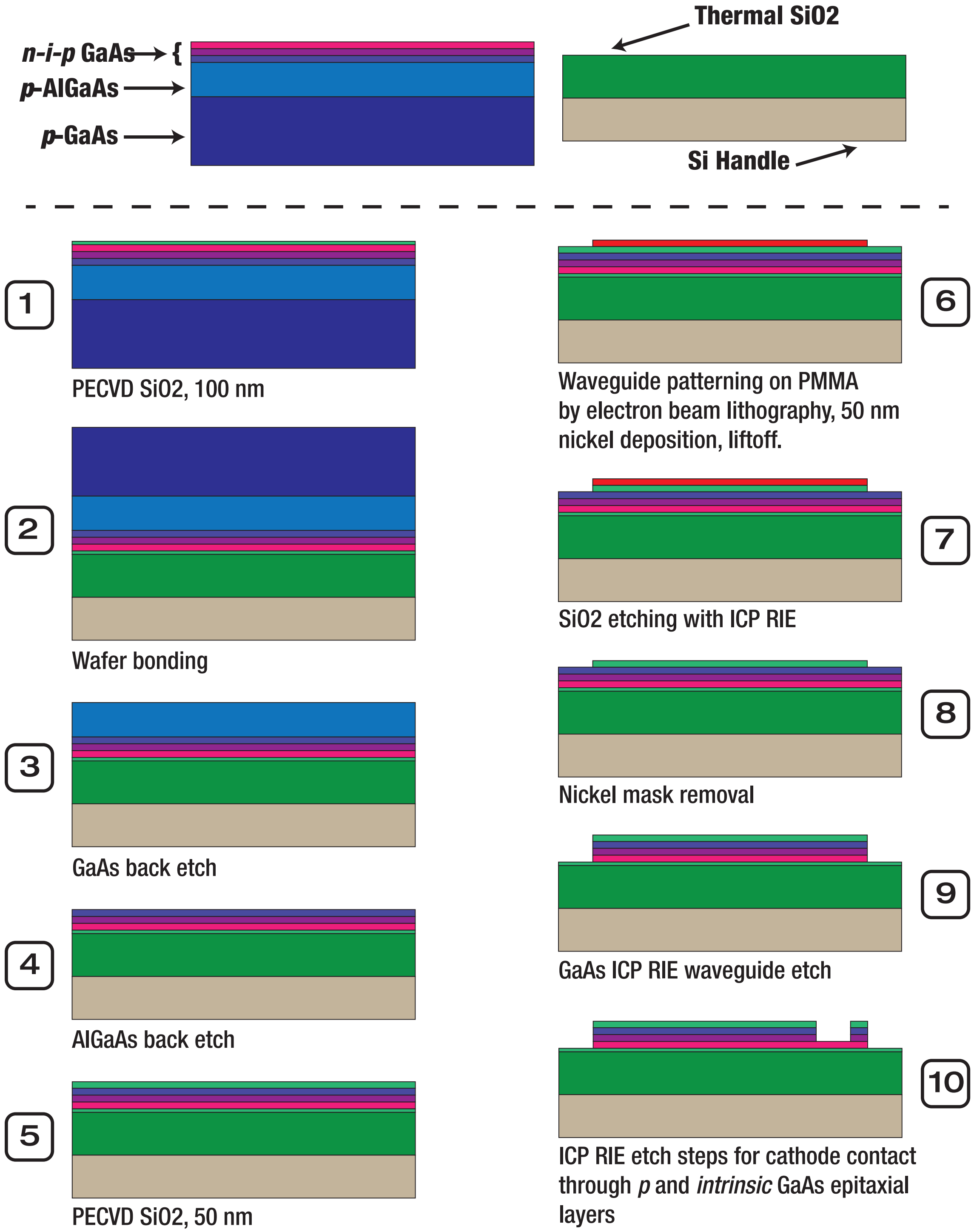
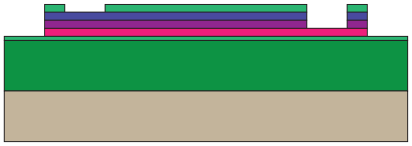


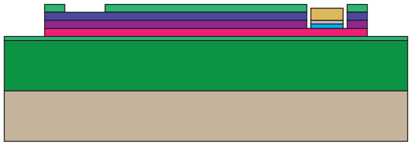
Figure 4.13: **Nanofabrication process flow for piezophotonic NEMS devices.** Note that the fabrication techniques broadly incorporate those in Fig. 2.19 and Fig. 3.3. Thus the fabrication here is compatible with, and designed to be incorporated with the process flow for MEMS neural probes, to enable eventual fabrication of a practical integrated fibreoptic probe system. The steps between 4 and 5 for fabrication of the electron beam and optical lithographic alignment marks have been omitted for clarity. Figure continues on the next page.

11



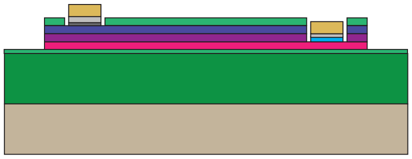
SiO2 etch for anode contact

12



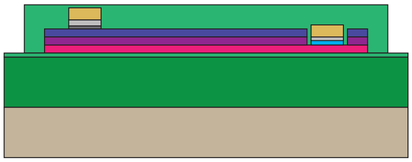
Cathode contact metallisation with 43 nm Ge / 30 nm Pt / 100 nm Au

13



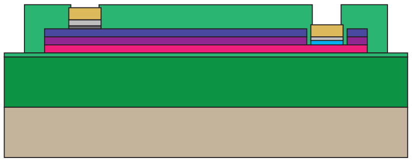
Anode contact metallisation with 25 nm Ti / 50 nm Pt / 100 nm Au

14



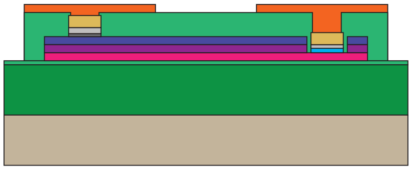
PECVD SiO2 deposition, 2 μm

15



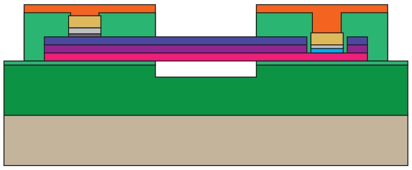
SiO2 window etch to contact pads

16



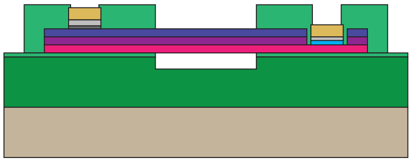
ZEP 520A resist application and EBPG patterning for NEMS resonator access window

17



SiO2 wet etch with HF for access window etching and NEMS resonator undercut

18



Removal and ZEP 520A resist

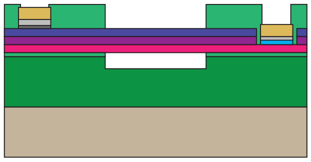
19

Alumina mask application for chip outline waveguide etch

20

Waveguide quality etch through SiO2, GaAs epitaxial stack, SiO2 and Si handle bulk

21



Removal of alumina in phosphoric acid

best performance, it is preferable to process three chips at a time, so that the tool is evenly supported during bonding.

The next step is to etch away the back of the GaAs substrate [89]. Developing a process to do this took some time, as the etch has to be reasonably fast, and selective against AlGaAs without damaging the underlying epitaxial layer. This is not something an Ar/$SiCl_4$ etch can readily achieve. The solution was to migrate to an SF_6/$SiCl_4$ etch, as described in Appendix A. This etch has infinite selectivity against the $Al_{0.8}Ga_{0.2}As$ sacrificial layer used in our devices. Before running the etch, the chip was cleaned in oxygen plasma within the ICP for 15 minutes. It was important to do this otherwise the etch would have cold spots on the surface, where the etch rate was reduced. On completion of the etch, it was also extremely important to clean the chip again in O_2 plasma for another 15 minutes *before* opening the chamber of the ICP. Doing otherwise would cause the leftover etch gases and reactants to react with the air and precipitate on the chip, ruining it.

After completion of the backetch, the next step is to remove the AlGaAs sacrificial layer. The best method to do this would be to use a dry etching technique selective against GaAs and silicon dioxide. Unfortunately, we have not developed such a process yet, although it is possible in principle. Instead, we use HF to remove the AlGaAs. Doing this requires some preparation however in the way of masking off areas undesirable for HF etching with baked photoresist. We do this by spinning AZ9245 onto the wafer/chip, and then using a darkfield mask on it, prepared with dimensions a few hundred microns less than the rectangular dimensions of the epitaxial layer. This effectively makes a top layer mask covering the silicon dioxide areas, just overlapping with the epitaxial layer so that the edge boundary with the underlying oxide is protected. This is important to prevent delamination of the epitaxial GaAs through undercutting by the HF. The AlGaAs layer may then be etched away by dipping the wafer in Transene BOE improved, which is specially formulated to impede this kind of delamination, for a period of 1 minute. The wafer is then cleaned in acetone / IPA / DI water, and then O_2 plasma cleaned with 20-20-80.

Now we begin electron beam lithography. We spin on PMMA 950 A4 and bake the chip. First, we manually centre-align and orient the chip in the EBPG, and run the pattern for the chip alignment marks at 1 nA of beam current. This will write the EBPG and optical lithography alignment marks, as well the device numbering. After development and descum, 5 nm Ti / 65 nm Au / 70 nm Ti is deposited with the electron beam vacuum deposition system, and lifted off in acetone.

The mask we will use for etching the waveguides will be silicon dioxide, and not a metal mask. This is as using a metal mask directly will transfer the crystal structure of the metal to the waveguide sidewalls on etching. Instead, we deposit a layer of silicon dioxide, etch that with a metal mask, and use the oxide as an etch mask for the waveguide etch. This substantively improves the sidewall profile of the waveguides. After solvent cleaning and a 15 minute 20-20-80 O_2 plasma clean, 50 nm of silicon dioxide is deposited in the PECVD using silane-argon.

PMMA 950 A4 is spun on and baked again, and on loading the wafer, the EBPG is aligned to the first set of 20 μm $\times$ 20 μm alignment marks. The waveguide pattern is then run at a beam current of 300 pA. Once the write has finished, alignment to the same set of marks is done again, and the bonding pads are written with a beam current of 5 nA. The wafer is developed, descummed and 50 nm of Ni is deposited with the vacuum deposition system. It is then lifted off in acetone, and then solvent cleaned.

The process flow then proceeds to silicon dioxide mask etching. After O_2 plasma cleaning in the ICP for 5 minutes, the C_4F_8 / SF_6 etch recipe is run for a period of 1 minute. There is now a 50 nm layer of nickel sitting on top of the SiO_2 mask, which needs to be removed. This is done by wet etching it in Transene TFG Nickel Etchant at 90°C for 60 s, before rinsing in DI water and cleaning. This process is safe for the silicon dioxide, silicon, and importantly, for the gallium arsenide epitaxial layer.

After a 15 minute 20-20-80 O_2 plasma clean, the chip is mounted on a 6” carrier wafer using Santovac 5 and the Oxford 100 ICP is used to etch the GaAs waveguides using the Ar / $SiCl_4$ waveguide etch recipe, for a period of 1 minute.

The next step is quite critical. We have two contact pads to interface with the n and p layers of the chip. The p layer is on the surface, and is easily interfaced. However, the n layer is at the bottom of the epitaxial stack. It will be necessary to partially etch the GaAs *pin* structure to reach the n layer and deposit our metallisation there. To do this, we spin on and bake ZEP520A resist, and load the wafer in the EBPG. Alignment is made to the second set of alignment marks, and the pad etch pattern for the n terminal is written with a beam current of 5 nA. It is developed and dried, and then the wafer is mounted in the Oxford ICP. Now, ZEP is a very durable mask, and can be used for etching, unlike PMMA. First, we run the C_4F_8 / SF_6 silicon oxide etch recipe for 1 minute to remove the oxide mask layer overlaying the bonding pad. A second etch is then run, the GaAs waveguide etch, to partially etch the epitaxial structure and reach the n doped layer. This has to be run very carefully not to over etch or under etch. The etch is run for exactly 21 seconds.

The ZEP mask is cleaned off in acetone / IPA / DI water, and the wafer is 20-20-80 cleaned for 15 minutes. ZEP520A is again spun on and baked, and the sample is loaded into the EBPG, which is aligned to the second set of alignment marks. The pad etch pattern for the p contact is run at 5 nA current. After development, the chip is loaded into the Oxford ICP and the C_4F_8 / SF_6 silicon oxide etch recipe is run for 1 minute to remove the insulation.

PMMA 950 A4 is now spun on after cleaning, and this time the n contact metallisation pattern is run at 5 nA. After a descum, 43 nm Ge / 30 nm Pt / 100 nm Au is deposited and lifted off [90, 91].

Likewise, PMMA 950 A4 is then spun on after cleaning, and the p contact metallisation pattern is run at 5 nA. After descum, 25 nm Ti / 50 nm Pt / 100 nm Au is deposited and lifted off [92].

The wafer is solvent cleaned and then 20-20-80 RIE O_2 plasma cleaned for 15 minutes. The electrical contacts are then annealed in the rapid thermal annealer at 450°C for 30 s, to form ohmic contacts [93].

Now, the waveguides are clad in silicon dioxide. The wafer is cleaned in 20-20-80 RIE O_2 plasma, and loaded in the PECVD, where 2 μm of SiO_2 is deposited using silane-argon.

We now have to dry etch the outline of our chip through the entire wafer stack, about the edge of the chip, where the nanotaper waveguides will couple with the lensed fibres. Since this is for optical coupling, all the etch steps have to be of waveguide-like quality. Dry etching must be done through the silicon dioxide layers, the GaAs waveguide layer and the silicon substrate layer, and the etch must be extremely smooth. The mask used to mask off the area of the chip must also be resilient against all the etch processes, and protect the fragile underlying devices. The major stumbling block to achieving this was the silicon dioxide etch. Originally, we used a C_4F_8 / O_2 process which was extremely harsh, producing a poor etch profile, and damaging any compatible mask we tried to use to shield the chip. To overcome this, we developed a C_4F_8 / SF_6 process which is infinitely selective against Al_2O_3, and which essentially produces a waveguide-etch type profile [94].

AZ9245 was spun on the wafer and baked, and a darkfield mask was used to photolithographically define the final outline of the chip (9 mm $\times$ 17 mm), aligning the mask to the optical alignment marks defined in the EBPG during the very first lithography step. After descum, 200 nm of Al_2O_3 was sputtered onto the chip, and lift-off was performed in acetone. Again, on rinsing the wafer, great care was used to irrigate the chip with acetone whilst removing it, and then to immediately immerse it IPA. This was done to prevent redeposition of the alumina lift-off particles back onto the wafer.

Santovac 5 is used to mount the device to a 6” carrier wafer that has been thermally oxidised and sputtered with alumina. The assembly is mounted into the Oxford 100 ICP, and the through stack etch is begun. First, the C_4F_8 / SF_6 silicon dioxide etch is run for 26 minutes. Then the Ar / $SiCl_4$

GaAs waveguide etch is run for 1 minute. The C_4F_8 / SF_6 silicon dioxide etch is then run again for another 26 minutes. On the completion of the etch, the table is cooled to -140°C and then the cryoetch recipe is run to etch through the entire silicon substrate thickness. Finally, the C_4F_8 / SF_6 silicon dioxide etch is run once more to etch through the last thickness of SiO_2 unremoved during the last HF etch step. The chip is now free, and is removed from the carrier by immersing it in acetone and then IPA. After removal and further solvent cleaning, the chip is placed in phosphoric acid at 60°C for 20 minutes to remove the alumina mask. It is rinsed thoroughly in DI water on completion of the etch, and a further 20-20-80 O_2 plasma clean is run for 15 minutes to prepare the chip for the next step.

Finally, ZEP520A is spun on and baked, and the chip is loaded into the EBPG. Alignment is made to the final set of alignment marks, and the window etch pattern is run at 1 nA, to define access windows through the oxide insulation/cladding down to the wirebond pads and to free and undercut the NEMS resonator. AZ9245 resist is then hand applied to the edges of the chip to protect this area, and is subsequently baked on. The chip is then immersed in 6:1 BOE for 10 minutes whilst gently agitating. It is removed, rinsed and dried to check the progress of the etch to ensure overetching does not occur. PECVD silicon dioxide etches much more quickly than thermal oxide. It is re-immersed in BOE for up to another 12 minutes, periodically removing and checking the etch progress under the electron microscope, until the NEMS resonator is suspended $\sim 50 - 100$ nm above the thermal silicon dioxide substrate. To conclude the nanofabrication, the chip is rinsed in acetone / IPA / DI water and dried.

Now that the chip is ready, it can be mounted in its chip carrier and wirebonded. We use an NTK-104032 28-pin sidebrased ceramic package which is ideal for our optoelectronic interfacing configuration. This particular package has all the wirebonding pads on the north and south ends of the package, matching the bonding area of our chip. The chip is mounted in the east-west direction, so that the optical coupling edges protrude across the side of the chip carrier. This allows easy interfacing with the lensed fibres using the micromanipulators in our vacuum test system. To mount the chip, a spacer was constructed out of 3/64" aluminium sheet using a water jet cutter, and machined to dimensions 8 mm $\times$ 9 mm. This spacer fits perfectly within the chip carrier cavity, and is flush with the surface of the ceramic carrier packaging. Silver conductive epoxy (MG Chemicals, 8330S) was used to bond the spacer to the chip carrier cavity, and then the NOEMS chip was bonded over the spacer and ceramic body with the silver epoxy. An antistatic vacuum pick-up tool was used to handle the chip, and strict ESD precautions were observed. On setting of the epoxy, the NEMS devices are finally wirebonded to the chip carrier with gold leads, and the chip is mounted into the vacuum chamber interface board using the ZIF socket.

Chapter 5

Radiofrequency Control of Ion Channels for Wireless and Multiplexed Manipulation of Neural Circuits

> To understand in depth what is going on in a brain, we need tools that can fit inside or between neurons and transmit reports of neural events to receivers outside... There is no law of physics that declares such an observational tool to be impossible.
>
> *Freeman Dyson,* Imagined Worlds *(1997)*

5.1 Radiogenetics

We propose to develop a new class of neuroscience and neurotherapeutic ion channel nanomolecular tools we term *radiogenetics*. In contrast to optogenetics, these tools are injectable into the bloodstream, are completely wireless, can be programmably multiplexed on potentially thousands of separate radio channels, and can be used with native ion channels or ones introduced through viral expression or genetic engineering. This is achieved through separation of the ion channel itself from the molecular machinery for sensing and actuation and from that for molecular and genetic targeting. This allows physically well characterised solid state nanomaterials to be used for high-bandwidth wireless sensing and actuation.

We will develop two different classes of tools based on *single-walled carbon nanotubes conjugated with single-stranded DNA*. The first class will be based on *semiconducting* carbon nanotubes, and will be sensitive to microwave radiation in the 1-10 GHz range. A section of the DNA strand will be a G-T polymer whose length will coil around and select semiconducting nanotubes of a specific length. Nanotubes of defined length will be sensitive to microwaves of a particular frequency, and will mechanically resonate with high quality factor in fluid on irradiation with this frequency. The terminal end of the DNA strand will be a DNA aptamer. This aptamer can be designed to target a native ion channel directly, or a membrane bound adaptor protein that is introduced through a viral vector. The nanotube-DNA constructs are biocompatible, cross the blood-brain barrier and are cleared from the body like a drug. As a consequence, specific native ion channels can be programmably targeted and wirelessly actuated in real-time with a defined microwave frequency, using intravenously injected nanodevices.

The second class of tools will be based on *metallic* carbon nanotubes and will be sensitive to near-infrared (NIR) radiation in the range 700-1100 nm. Using a similar construct as with the first scheme, a proximal DNA strand with a tailored selection sequence will select for nanotubes of specific diameter and chirality. As such, they will select for nanotubes sensitive to a particular infrared frequency. The terminal DNA sequence will be an aptamer as above. On irradiation with infrared light, the metallic nanotubes increase in temperature. As they are bound closely to the native ion channels, the temperature increase is sufficient to actuate the channels. Since biological tissue exhibits an ‘optical window’ in the NIR region, this technique is wireless. It is also capable of multiplexing, as it allows programmable selection of nanotubes sensitive to defined infrared frequencies.

Finally, with both techniques, the same basic technology can be further adapted to deliver DNA or RNA into cells, instead of the use of viruses. Since carbon nanotubes act as microwave nanoantennas and concentrate microwave fields, if a carbon nanotube is in the vicinity of a cell membrane, even very low power microwave radiation will electroporate the membrane via the nanotube. In this manner, DNA or RNA can be delivered into a cell. A restriction enzyme may then snip the gene from the nanotube-nucleotide delivery vector via an encoded restriction site, making it active.

5.2 Background & Significance

A major effort has been initiated in the past few years to bring to bear major advancements in nanophysics and nanotechnology to the problems of neuroscience [95, 31]. The goal of this endeavour is to develop the capability to simultaneously manipulate and record from every single neuron in an awake behaving brain, in a completely wireless and non-invasive fashion. Wireless and minimally invasive large-scale neural recording may become possible through the integration of frequency multiplexed nanomechanical oscillators with nanophotonics [34]. The potential bandwidth of such recording interfaces are on the order of billions of channels on a single optical fibre. However, thus far, the most useful tools for large-scale recording and manipulation have come from the world of biology, harnessing fluorescence and light-sensitive proteins.

Optogenetics has emerged in the past decade as an extremely powerful tool in neuroscience, that allows electrical modulation of single neurons *in vivo* with high temporal precision through the action of visible light. Furthermore, cell specificity can be encoded chemically through the use of a promotor in tandem with the opsin gene. Having control over single neurons in this manner has for the first time enabled complex dissection of neural circuits in awake, behaving animals.

Despite it's experimental power, optogenetics has a number of limitations. Although light itself has enormous bandwidth (the basis of fibre optic communications), optogenetics cannot take advantage of it as channel rhodopsins have very broad spectral sensitivity. This makes it impossible to multiplex more than two optical channels at a time. Since the opsins used in optogenetics are merely repurposed genes found in nature, and those genes specifically evolved to have such broad sensitivity, it's highly unlikely opsin genes will ever be able modified to have better bandwidth. Multiplexing is thus the primary limitation of optogenetics. This is an especially severe limitation as individual cortical neural circuits can have millions of components - dissecting such circuits in an optogenetic-like manner requires far more channel bandwidth.

A second limitation is the use of light for actuation. Biological tissue scatters visible light very readily, more strongly for shorter wavelengths. This necessitates the use of implanted optical fibers for manipulation of circuits since light won't penetrate the skin, let alone the skull or brain tissue itself. Thus the technique is not wireless. This limits the effective number of neural circuits that can be manipulated at the same time, without significant damage to the brain through implanted waveguides.

A third limitation is the size of the opsin genes themselves. Their large size necessitates the use of larger viruses, which have inherent drawbacks, including poor and short-term expression. It would be advantageous to use smaller viruses, with genes that are small enough to have sufficient room for several promotor sequences.

Other drawbacks to optogenetics include the eventual death of transfected neurons due to overexpression of the channel, and for therapeutic applications, the expression of a foreign protein, which may elicit an immune response.

In this proposal we present a novel set of ideas that may represent a new path in the evolution of the concept behind optogenetics. The primary reason for the drawbacks listed above is that it is currently very difficult to perform functional protein modification or *de novo* protein design. This requires quantum simulation, which is still several decades into the future at the very least. Here we propose a new paradigm - *the conjugation of solid state crystalline nanoparticles with programmable biological aptamers*. Since solid-state nanoparticles are crystalline and highly regular, we can use solid state physics and highly simplified quantum mechanical models to precisely predict the electrical, optical and mechanical properties of these nanocrystals. This allows tailored design of communication and actuation in our molecular devices. We can then couple such devices to various biological aptamers for interaction with native proteins, which is well understood on the biology side. It also becomes possible to use smaller viruses to deliver genes for small proteins designed to mate with the aptamers, potentially decoupling the genetic targeting of our devices from the communication/actuation machinery itself, which is no longer encoded genetically or manufactured by the cell.

A small collection of efforts have been initiated to develop ideas parallel to this. This includes initial work by Pralle [96] and follow up work by Anikeeva [97] to develop a technique for wireless control of single neurons through heating of TRPV1 channels via radiofrequency heating of manganese ferrite nanoparticles. A downside to this method is that the magnetic nanoparticles are freely diffused in the extracellular fluid, and are not bound to the membrane or the ion channel. This is why specialised heat-sensitive TRPV1 are required for the scheme to work - the temperature increase is too small to affect native ion channels. At present, frequency multiplexing and programmed targeting of native ion channels is not available with this technique, although the possibility of multiplexing has been theoretically explored [98]. Functionalisation of ferrite nanoparticles with antibodies has also been demonstrated, opening the possibility of use with native ion channels [99]. Finally, work by Pepperberg and Bezanilla has demonstrated targeted optical stimulation of single neurons using membrane-bound gold nanoparticles [100]. This also operates through heating, which induces changes in membrane capacitance, in turn activating native ion channels.

5.3 Carbon Nanotube Constructs

5.3.1 Mechanical Perturbation of Ion Channels

An intravenously delivered semiconductor carbon nanotube mechanical resonator, conjugated with a programmable nucleic acid, for addressable microwave control of native ion channels.

Semiconductor carbon nanotubes (CNT's), of diameter 1-1.5nm and with lengths in the range of 50nm-5μm, mechanically oscillate with high quality factor in fluid [101], when exposed to light [102] or microwave radiation [103, 104]. This mechanical oscillation is resonantly tuneable [105], and nanotubes of different lengths will be sensitive to different microwave frequencies. Semiconductor CNT devices like these are *resonant nanoelectromechanical systems* (NEMS). Conversely, metallic carbon nanotubes heat up when exposed to electromagnetic radiation [106]. The physics behind this effect is complex, and based on photon-electron-phonon interactions as described by quantum field theory [107].

Carbon nanotubes in their isolated form are toxic to mammals [108]. Naked CNT's are highly hydrophobic, and congregate to form rafts in lipids, causing cell lysis. However, carbon nanotubes can be conjugated with DNA or RNA to make them hydrophilic and biocompatible [109]. A simple guanine-thymine (G-T) polymer will coil around a carbon nanotube, and furthermore because of energetics, a G-T repeat of a specific length will coil around and 'select' carbon nanotubes of a specific length [110]. The length of the G-T repeat thus allows programmable selection of CNT's of a particular length, and hence a specific microwave frequency sensitivity.

DNA-CNT conjugations like this have already been shown to cross the blood-brain barrier after intravenous injection [111]. Furthermore, through currently poorly understood pathways, CNT's conjugated to DNA are actually cleared by the body and excreted [112, 113]. Thus, it acts like a drug with a half-life after intravenous injection, and does not remain permanently *in vivo*.

Our idea is to construct a new type of molecular device with a G-T repeat to select a semiconductor carbon nanotube of programmed length. The G-T repeat sequence is then followed by a DNA (or RNA) aptamer sequence tailored to bind to a native neuronal ion channel [114]. Further intermediate sequences could possibly be included to combine this technology with the techniques of DNA computing [115]. We name this construct a *cNEMS-DNA device*, and depict it in Fig. 5.1. Since most ion channels are mechanosensitive to some degree because of their evolutionary origin, after the device has bound to the channel, upon mechanical actuation of the nanotube by the ad-

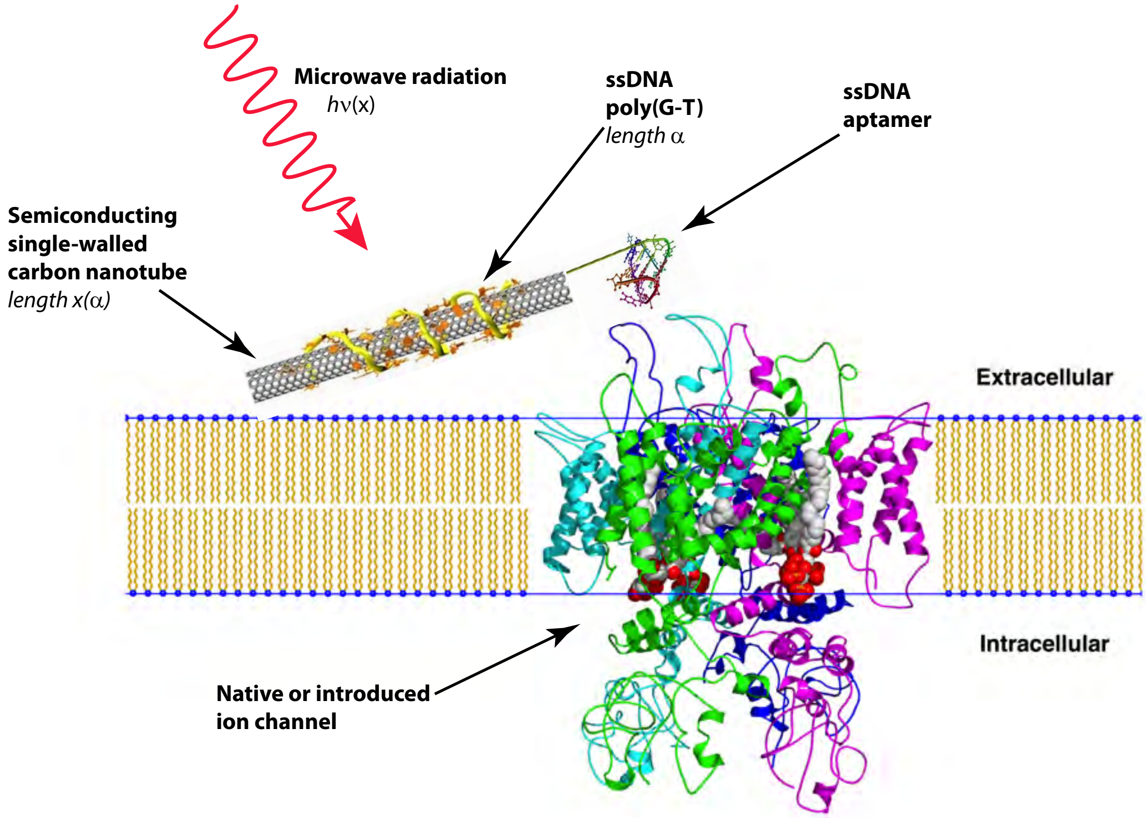


Figure 5.1: **cNEMS-DNA ion channel construct.** On absorbing microwave radiation at the resonant frequency, the carbon nanotube mechanically oscillates. This disturbs the lipid packing and membrane mechanics, perturbing the kinetics of the ion channel, thence switching its conductive state. A poly(G-T) repeat of length α coils around and selects a semiconducting carbon nanotube of length $x(\alpha)$, which is sensitive to microwave radiation at frequency $\nu(x(\alpha))$. Thus the initial programmed strand of DNA links a particular ion channel using the aptamer sequence to microwave sensitivity at a specified frequency, using the poly(G-T) sequence.

dressed microwave frequency, the resonating nanotube should disrupt the lipid membrane around the ion channel. Disruption of the interaction between the transmembrane motifs of the channel and the lipid bilayer may alter the conductive state of the channel. In this fashion, it may be possible to develop all manner of excitatory and inhibitory microwave sensitive ion channels, with a frequency bandwidth potentially in the thousands. Significantly, these devices would be able to be introduced intravenously, obviating the need for invasive neurosurgery.

In addition to its utility for multiplexing, such microwave sensitive ion channels would allow fully wireless and non-invasive single neuron control, addressable over the entire brain. The human body acts as a broadband microwave antenna, and allows deep penetration of radiation in the range 10MHz-10GHz. It is anticipated that power in the μW range would be more than sufficient to actuate the devices, well below the threshold where heating would be a concern.

Finally, since foreign proteins are not introduced, it may have therapeutic application, for example for wireless and non-invasive deep brain stimulation for treatment of Parkinson's disease. The biocompatibility and bioclearance of the cNEMS-DNA devices would be highly advantageous here. Furthermore, the ability to multiplex may allow more complex intervention at several concurrent points in the relevant brain circuits.

A more sophisticated iteration of this concept might enable novel multichannel wireless neural interfaces. For example, if targeting of unique frequency cNEMS-DNA devices could be achieved over primary visual cortex V1 with sufficient resolution, visual projection might be achieved remotely via microwave transmission.

The technique is flexible enough to several modalities for targeting native or introduced ion channels:

- The DNA aptamer could directly target a native ion channel.
- The DNA aptamer could target a specialised ion channel introduced via injected virus or expressed through genetic engineering, as with optogenetics.
- The aptamer could target a small membrane-bound adaptor protein, introduced virally or via genetic engineering. The adaptor protein would directly mate the cNEMS-DNA device to the target ion channel. Since the adaptor protein is so much smaller than an optogenetic channel, it could be packaged with several promotor sequences in a small, well expressed virus such as an AAV.
- The aptamer could target a small membrane-bound protein, introduced virally or via genetic engineering. The protein would serve as an anchor point for the cNEMS-DNA device to disrupt the membrane mechanics of the cell more broadly, affecting neighbouring ion channels.
- The cNEMS-DNA device could be directly tethered to an antibody or nanobody, specific for a native or introduced ion channel or adaptor/anchor protein.

5.3.2 Thermal Perturbation of Ion Channels

To develop an intravenously delivered metallic carbon nanotube nanoscale heater, conjugated with a programmable nucleic acid, for addressable infrared control of native ion channels.

A variation on the technique described in 5.3.1 is to instead use metallic carbon nanotubes, which increase in temperature on irradiation with near infra-red (NIR) light [116, 117]. Since biological tissue exhibits an optical window in the 700-1100nm range, this may allow wireless control of ion channels with infrared radiation. All ion channels are temperature sensitive, the probability distributions for their conductive states shifting with changes in temperature. Whilst there are channels specialised to respond to small temperature changes, such as those in the TRP family, a native ion channel will respond to the temperature increase from the carbon nanotube provided it is directly bound. The terminal DNA targeting motif in this case is as laid out in 5.3.1.

As the NIR frequency selectivity for metallic nanotubes would not be determined by the length of the nanotube, but by the tube diameter and chirality, the potential channel bandwidth might be smaller than with mechanical actuation, although still substantial. A comprehensive study of various DNA motifs that select CNT's of specific chirality and nanotube diameter has already been carried out [118], thus programmable frequency selection may likewise be possible with the *cHeater-DNA* configuration.

With both the configurations described in 5.3.1 & 5.3.2, targeted transfection of single neurons with DNA may still be required in order to express adaptor proteins or specialised ion channels, such as TRPV1, which are both mechanosensitive and heat sensitive. In optogenetics, such targeted protein expression is achieved using viral vectors or by genetically engineering the organism. In recent years, substantive work has explored the utility of carbon nanotubes for transporting DNA and RNA into cells, as an alternative to viruses. Although viral vectors exhibit excellent gene expression, concerns of safety currently preclude their use in therapeutic applications, most especially in the brain. The primary modality for DNA transport used in reported schemes with carbon nanotubes is via endocytosis, which is an active process requiring energy from the cell. Carbon nanotubes may be covalently bound to DNA or RNA using polyethylene glycol (PEG), or through coiling of the nucleic acid around the nanotube, as already described [116, 119]. Once the complex is within the cell, one of the schemes reported by Dai releases the DNA from the nanotube through irradiation with NIR, which heats the nanotube, causing uncoiling of the DNA strand [119]. The drawback with relying on endocytosis is the DNA tends to be hydrolysed within the endocytotic vesicles. To overcome this, other work has demonstrated directed transport of DNA with carbon nanotubes using magnetic fields (*magnetoporation*) [120]. Here, the left over nickel catalyst particles at the tips of PECVD grown carbon nanotubes are used to direct and motivate the carbon nanotubes with magnetic force, to 'spear' the cells and drive the nanotube-DNA complex directly into the cell without relying on vesicular endocytosis. This produces excellent levels of transfection, similar to that achieved with viral vectors.

5.3.3 Electroporation

Microwave fields have also been shown to nanoporate cell membranes, to such degree that artificial ion channels form, inducing spiking activity in neurons [121, 122]. Although at low power, this occurs spontaneously at frequencies from 30-300 GHz. The cNEMS-DNA and cHeater-DNA constructs we have put forward suggest a further construct for gene transfection, drawing on the kinds of ideas just described. Since metallic carbon nanotubes acts as nanoantennas and concentrate microwave fields [123], we propose exploring ideas for a cNanoporate-DNA construct for gene delivery. Here, similar to the cHeater-DNA construct, a selection sequence will coil around and select a metallic carbon nanotube. Downstream to this sequence will be a restriction site, terminated with the actual gene for transfection. On microwave irradiation, the metallic nanotubes will concentrate and re-emit the microwave field, inducing nanoporation of the cell membrane because of its dielectric properties. This will cause the entire construct to become internalised into the cell, shortcircuiting the endocytotic pathway. Once inside the cell, if the correct restriction enzyme is present, the gene will be released from the rest of the cNanoporate-DNA device. In this manner, the carbon nanotube always remains conjugated to DNA, biocompatible and subject to degradation. Furthermore, since the gene will only be released and become active in specific cells with the correctly designed restriction enzyme, the construct could be administered intravenously for generalised distribution. Other techniques could be drawn upon to target the particular neurons of interest with the compatible restriction enzyme.

5.4 Summary

Hence, in the fashion described, the same basic construct could be used for a suite of tools to:

- mechanically perturb an ion channel (**cNEMS-DNA**: *semiconducting CNT +* μ*wave radiation*)
- heat an ion channel (**cHeater-DNA**: *metallic CNT + NIR radiation*)
- deliver an ion channel (**cNanoporate-DNA**: *metallic CNT +* μ*wave radiation*)

in a wireless and frequency programmable manner, for non-invasive, large-scale and multi-channel control of neural circuits.

Bibliography

[1] J. Gertner. *The Idea Factory: Bell Labs and the Great Age of American Innovation*. The Penguin Press, New York, 2012.

[2] J. von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 329–378. Princeton University Press, Princeton, NJ, 1956.

[3] A. S. Sadek, K. Nikolić, and M. Forshaw. Parallel information and computation with restitution for noise-tolerant nanoscale logic networks. *Nanotechnology*, 15(1):192–210, 2004.

[4] J. von Neumann. *The Computer and the Brain*. Yale University Press, New Haven, 1958.

[5] P. A. Merolla *et. al.* A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197):668–673, 2014.

[6] R. Penrose. *The Emperor's New Mind*. Oxford University Press, New York, 1989.

[7] A. S. Sadek, K. Nikolić, and M. Forshaw. Fault tolerance and ultimate physical limits of nanocomputation. In D. Crawley, K. Nikolić, and M. Forshaw, editors, *3D Nanoelectronic Computer Architecture and Implementation*. Institute of Physics Publishing, Bristol and Philadelphia, 2005.

[8] C. Koch. *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, Inc., New York, 1999.

[9] C. A. Anastassiou *et. al.* Ephaptic coupling of cortical neurons. *Nature Neurosci.*, 14:217–223, 2011.

[10] G. Ling and R. W. Gerard. The normal membrane potential of frog sartorius fibers. *J. Cell. Physiol.*, 34:383–396, 1949.

[11] O. P. Hamill *et. al.* Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Arch.*, 391:85–100, 1981.

[12] A. K. Lee *et. al.* Whole-cell recordings in freely moving rats. *Neuron*, 51:399–407, 2006.

[13] J. T. Davie *et. al.* Dendritic patch-clamp recording. *Nat. Protoc.*, 1:1235–1247, 2006.

[14] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurons in the cats striate cortex. *J. Physiol.*, 148:574–591, 1959.

[15] G. Buzski. Large-scale recording of neuronal ensembles. *Nature Neurosci.*, 7:446–451, 2004.

[16] S. B. Laughlin and T. J. Sejnowski. Communication in neuronal networks. *Science*, 301:1870–1874, 2003.

[17] T. J. Blanche *et. al.* Polytrodes: high-density silicon electrode arrays for large-scale multiunit recording. *J. Neurophysiol.*, 93:2987–3000, 2005.

[18] J. Du *et. al.* High-resolution three-dimensional recording of neuronal activity with microfabricated electrode arrays. *J. Neurophysiol.*, 101:1671–1678, 2009.

[19] J. Du *et. al.* Multiplexed, high density electrophysiology with nanofabricated neural probes. *PLoS ONE*, 6:e26204, 2011.

[20] R. R. Harrison *et. al.* Wireless neural recording with single low-power integrated circuit. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 17:322–329, 2009.

[21] K. C. Cheung. Implantable microscale neural interfaces. *Biomed. Microdevices*, 9:923–938, 2007.

[22] M. W. Varney *et. al.* Polycrystalline-diamond mems biosensors including neural microelectrode-arrays. *Biosensors*, 1:118–133, 2011.

[23] E. J. Hwang and R. A. Andersen. Brain control of movement execution onset using local field potentials in posterior parietal cortex. *J. Neurosci.*, 29:14363–14370, 2009.

[24] J. E. O'Doherty *et. al.* Active tactile exploration using a brain-machine-brain interface. *Nature*, 479:228–231, 2011.

[25] K. Murari *et. al.* Integrated potentiostat for neurotransmitter sensing. *IEEE Eng. Med. Biol. Mag.*, 24:23–29, 2005.

[26] T. E. McKnight *et. al.* Resident neuroelectrochemical interfacing using carbon nanofiber arrays. *J. Phys. Chem. B*, 110:15317–15327, 2006.

[27] B. E. K. Swamy and B. J. Venton. Carbon nanotube-modified microelectrodes for simultaneous detection of dopamine and serotonin *in vivo*. *Analyst*, 132:876–884, 2007.

[28] M. Roham *et. al.* A wireless IC for wide-range neurochemical monitoring using amperometry and fast-scan cyclic voltammetry. *IEEE Trans. Biomed. Circuits Syst.*, 2:3–9, 2008.

[29] P. Blinder *et. al.* The cortical angiome: an interconnected vascular network with noncolumner patterns of blood flow. *Nature Neurosci.*, 16:889–897, 2013.

[30] A. D. Wong *et. al.* The blood-brain barrier: an engineering perspective. *Front. Neuroeng.*, 6:7, 2013.

[31] A. P. Alivisatos *et. al.* Nanotools for neuroscience and brain activity mapping. *ACS Nano*, 7(3):1850–1866, 2013.

[32] R. P. Feynman. *Feynman Lectures on Computation*. Addison-Wesley, Reading, MA, 1996. A. J. G. Hey and R. W. Allen, editors.

[33] S. C. Masmanidis *et. al.* Multifunctional nanomechanical systems via tunably coupled piezoelectric actuation. *Science*, 317:780–783, 2005.

[34] A. S. Sadek *et. al.* Wiring nanoscale biosensors with piezoelectric nanomechanical resonators. *Nano Lett.*, 10(5):1769–1773, 2010.

[35] C. Joachim and J. K. Gimzewski. An electromechanical amplifier using a single molecule. *Chem. Phys. Lett.*, 265:353–357, 1997.

[36] A. Bachtold *et. al.* Logic circuits with carbon nanotube transistors. *Science*, 294:1317–1320, 2001.

[37] J. Kong *et. al.* Nanotube molecular wires as chemical sensors. *Science*, 287:622–625, 2000.

[38] Y. Cui, Q. Wei, H. Park, and C. M. Lieber. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science*, 293:1289–1292, 2001.

[39] K. L. Ekinci, X. M. H. Huang, and M. L. Roukes. Ultrasensitive nanoelectromechanical mass detection. *Appl. Phys. Lett.*, 84:4469–4471, 2004.

[40] Y. T. Yang *et. al.* Zeptogram-scale nanomechanical mass sensing. *Nano Lett.*, 6:583–586, 2006.

[41] Z. Zhong *et. al.* Nanowire crossbar arrays as address decoders for integrated nanosystems. *Science*, 302:1377–1379, 2003.

[42] R. Beckman *et. al.* Bridging dimensions: demultiplexing ultrahigh-density nanowire circuits. *Science*, 310:465–468, 2005.

[43] Z. L. Wang and J. Song. Piezoelectric nanogenerators based on zinc oxide nanowire arrays. *Science*, 312:242–246, 2006.

[44] D. R. Walt. Nanomaterials: top-to-bottom functional design. *Nature Mater.*, 1:17–18, 2002.

[45] G. Zheng *et. al.* Multiplexed electrical detection of cancer markers with nanowire sensor arrays. *Nature Biotechnol.*, 23:1294–1301, 2005.

[46] A. K. Naik *et. al.* Towards single-molecule nanomechanical mass spectrometry. *Nature Nanotechnol.*, 4:445–449, 2009.

[47] F. Patolsky *et. al.* Detection, stimulation, and inhibition of neuronal signals with high-density nanowire transistor arrays. *Science*, 313:1100–1104, 2006.

[48] R. D. Rabbitt *et. al.* Evidence of piezoelectric resonance in isolated outer hair cells. *New J. Phys.*, 88:2257–2265, 2005.

[49] W. Denk and W. W. Webb. Thermal noise limited transduction observed in mechanosensory receptors of the inner ear. *Phys. Rev. Lett.*, 63:207–210, 1989.

[50] R. J. Pumphrey and T. Gold. Hearing I: the cochlea as a frequency analyser. *Proc. R. Soc. London Ser. B*, 135:462–491, 1948.

[51] A. N. Cleland and M. L. Roukes. A nanometre-scale mechanical electrometer. *Nature*, 392:160–162, 1998.

[52] P. A. Truitt *et. al.* Efficient and sensitive capacitive readout of nanomechanical resonator arrays. *Nano Lett.*, 7:120–126, 2007.

[53] Q. P. Unterreithmeier, E. M. Weig, and J. P. Kotthaus. Universal transduction scheme for nanomechanical systems based on dielectric forces. *Nature*, 458:1001–1004, 2009.

[54] S. Timoshenko. *Vibration Problems in Engineering*. Wiley, New York, 1974.

[55] X. M. H. Huang *et. al.* Vhf, uhf and microwave frequency nanomechanical resonators. *New J. Phys.*, 7:247, 2005.

[56] D. W. Carr and H. G.Craighead. Fabrication of nanoelectromechanical systems in single crystal silicon using silicon on insulator substrates and electron beam lithography. *J. Vac. Sci. Technol. B*, 15:2760–2763, 1997.

[57] N. Liu *et. al.* Time-domain control of ultrahigh-frequency nanomechanical systems. *Nature Nanotechnol.*, 3:715–719, 2008.

[58] E. Buks and B. Yurke. Mass detection with a nonlinear nanomechanical resonator. *Phys. Rev. E*, 74:046619, 2006.

[59] K. Najafi, J. Ji, and K. D. Wise. Scaling limitations of silicon multichannel recording probes. *IEEE Trans. Biomed. Eng.*, 37:1–11, 1990.

[60] R. R. Harrison and C. Charles. A low-power low-noise cmos amplifier for neural recording applications. *IEEE J. Solid-State Circuits*, 38:958–965, 2003.

[61] N. Sinha *et. al.* Piezoelectric aluminum nitride nanoelectromechanical actuators. *Appl. Phys. Lett.*, 95:053106, 2009.

[62] N. Gershenfeld. *The Physics of Information Technology*. Cambridge University Press, Cambridge, U.K., 2000.

[63] M. Li *et. al.* Harnessing optical forces in integrated photonic circuits. *Nature*, 456:480–484, 2008.

[64] I. De Vlaminck *et. al.* Detection of nanomechanical motion by evanescent light wave coupling. *Appl. Phys. Lett.*, 90:233116, 2007.

[65] S. Mandal, S. M. Zhak, and R. Sarpeshkar. A bio-inspired active radio-frequency silicon cochlea. *IEEE J. Solid-State Circuits*, 44:1814–1828, 2009.

[66] H. A. C. Tilmans, M. Elwenspoek, and J. H. J. Fluitman. Micro resonant force gauges. *Sens. Actuat. A*, 30:35–53, 1992.

[67] S. Ryckebusch and G. Laurent. Interactions between segmental leg central pattern generators during fictive rhythms in the locust. *J. Neurophysiol.*, 72:2771–2785, 1994.

[68] U. Rutishauser, E. M. Schuman, and A. N. Mamelak. Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, *in vivo*. *J. Neurosci. Methods*, 154:204–224, 2006.

[69] R. Tabrizian and M. Rais-Zadeh. The effect of charge redistribution on limiting the $kt^2 \cdot Q$ product of piezoelectrically transduced resonators. In *18th International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers '15)*, pages 981–984, Anchorage, AK, June 2015. IEEE.

[70] D. H. Hubel. Tungsten microelectrode for recording from single units. *Science*, 125:549–550, 1957.

[71] K. D. Wise, J. B. Angell, and J. B. Starr. An integrated-circuit approach to extracellular microelectrodes. *IEEE Trans. Biomed. Eng.*, BME-17:238–247, 1970.

[72] R. Sarpeshkar *et. al.* Low-power circuits for brain-machine interfaces. *IEEE Trans. Biomed. Circuits Syst.*, 2(3):173–183, 2008.

[73] M. D. Henry *et. al.* Alumina etch masks for fabrication of high-aspect-ratio silicon micropillars and nanopillars. *Nanotechnology*, 20:255305, 2009.

[74] J. E. Ferguson, C. Boldt, and A. D. Redish. Creating low-impedance tetrodes by electroplating with additives. *Sens. Actuators A. Phys.*, 156:388–393, 2009.

[75] D. Cai *et. al.* Angiogenic synergism, vascular stability and improvement of hind limb ischemia by a combination of PDGF-BB and FGF-2. *Nature Medicine*, 9(5):604–613, 2005.

[76] N. Jackson *et. al.* Long-term neural recordings using MEMS based movable microelectrodes in the brain. *Front. Neuroeng.*, 3:10, 2010.

[77] S. Musallam *et. al.* A floating metal microelectrode array for chronic implantation. *J. Neurosci. Methods*, 160:122–127, 2007.

[78] J. Viventi *et. al.* Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity *in vivo*. *Nature Neurosci.*, 14:1599–1605, 2011.

[79] B. Tian *et. al.* Macroporous nanowire nanoelectronic scaffolds for synthetic tissues. *Nature Mater.*, 11:986–994, 2012.

[80] R. A. Brodner and R. S. Markowitz. Feasibility of intracranial surgery in the rat fetus: model and surgical principles. *Neurosurgery*, 14:553–556, 1984.

[81] Y. Chen, H. Li, and M. Li. Flexible and tunable silicon photonic circuits on plastic substrates. *Sci. Rep.*, 2(622), 2012. DOI:10.1038/srep00622.

[82] R. J. Essiambre *et. al.* Capacity limits of optical fiber networks. *J. Lightwave Technol.*, 28(4):662–701, 2010.

[83] M. Povinelli *et. al.* Evanescent-wave bonding between optical waveguides. *Opt. Lett.*, 30(22):3042–3044, 2005.

[84] W. H. P. Pernice, M. Li, and H. X. Tang. Theoretical investigation of the transverse optical force between a silicon nanowire waveguide and a substrate. *Opt. Express*, 17(3):1806–1816, 2009.

[85] V. R. Almeida, R. R. Panepucci, and M. Lipson. Nanotaper for compact mode conversion. *Opt. Lett.*, 28(15):1302–1304, 2003.

[86] M. Bachmann, P. A. Besse, and H. Melchior. General self-imaging properties in $n \times n$ multimode interference couplers including phase relations. *Appl. Opt.*, 33:3905–3911, 1994.

[87] Q. Y. Tong, G. Fountain, and P. Enquist. Room temperature SiO_2/SiO_2 covalent bonding. *Appl. Phys. Lett.*, 89:042110, 2006.

[88] D. Liang *et. al.* Low-temperature, strong SiO_2-SiO_2 covalent wafer bonding for III-V compound semiconductors-to-silicon photonic integrated circuits. *J. Electron. Mater.*, 37:1552–1559, 2008.

[89] R. J. Olson *et. al.* Alternate backside thinning of GaAs-based devices. In *GaAs Mantech*. The International Conference on Compound Semiconductor Manufacturing Technology, 1999.

[90] C. Lin and C. P. Lee. Harnessing optical forces in integrated photonic circuits. *J. Appl. Phys.*, 67:260–263, 1990.

[91] A. G. Baca *et. al.* A survey of ohmic contacts to III-V compound semiconductors. *Thin Solid Films*, 308-309:599–606, 1997.

[92] G. Stareev. Formation of extremely low resistance Ti/Pt/Au ohmic contacts to *p*-gaas. *Appl. Phys. Lett.*, 62:2801–2803, 1993.

[93] Y. Wang *et. al.* Research on rapid thermal annealing of ohmic contact to GaAs. In *Optoelectronics and Microelectronics (ICOM), 2012 International Conference on*, pages 61–63. IEEE, August 2012.

[94] K. Kolari. High etch selectivity for plasma etching SiO_2 with AlN and Al_2O_3 masks. *Microelectronic Engineering*, 85:985–987, 2008.

[95] A. H. Marblestone *et. al.* Physical principles for scalable neural recording. *Front. Comput. Neurosci.*, 7(137):1–34, 2013.

[96] H. Huang *et. al.* Remote control of ion channels and neurons through magnetic-field heating of nanoparticles. *Nature Nanotech.*, 5:602–606, 2010.

[97] R. Chen *et. al.* Wireless magnetothermal deep brain stimulation. *Science*, 347(6229):1477–1480, 2015.

[98] M. G. Christiansen *et. al.* Magnetically multiplexed heating of single domain nanoparticles. *Appl. Phys. Lett.*, 104:213103, 2014.

[99] S. A. Stanley *et. al.* Radio-wave heating of iron oxide nanoparticles can regulate plasma glucose in mice. *Science*, 336:604–608, 2012.

[100] J. L. Carvalho de Souza *et. al.* Photosensitivity of neurons enabled by cell-targeted gold nanoparticles. *Neuron*, 86:207–217, 2015.

[101] R. B. Bhiladvala and Z. Jane Wang. Effect of fluids on the q factor and resonance frequency of oscillating micrometer and nanometer scale beams. *Phys. Lett. E*, 69:036307, 2004.

[102] D. J. Flannigan and A. H. Zewail. Optomechanical and crystallization phenomena visualised with 4d electron microscopy: interfacial carbon nanotubes on silicon nitride. *Nano Lett.*, 10(5):1892–1899, 2010.

[103] K. Jensen *et. al.* Nanotube radio. *Nano Lett.*, 7:3508–3511, 2007.

[104] A. Ziaei *et. al.* Microwave applications of carbon nanotubes: nano-antennas and nano-switches. In V. K. Varadan, editor, *Proceedings of SPIE*, volume 6931 of *Nanosensors and Microsensors for Bio-Systems*, page 69310R, 2008.

[105] V. Sazanova *et. al.* A tunable carbon nanotube electromechanical oscillator. *Nature*, 431:284–287, 2004.

[106] T. J. Imholt *et. al.* Nanotubes in microwave fields: light emission, intense heat, outgassing and reconstruction. *Am. Chem. Soc.*, 15:3969–3970, 2003.

[107] M. Verissimo-Alves *et. al.* Polarons in carbon nanotubes. *Phys. Rev. Lett.*, 86:3372–3375, 2001.

[108] G. M. Mutlu *et. al.* Biocompatible nanoscale dispersion of single-walled carbon nanotubes minimizes *in vivo* pulmonary toxicity. *Nano Lett.*, 10:1664–1670, 2010.

[109] M. Zheng *et. al.* DNA-assisted dispersion and separation of carbon nanotubes. *Nature Mater.*, 2:338–342, 2003.

[110] M. Zheng *et. al.* Structure-based carbon nanotube sorting by sequence-dependent DNA assembly. *Science*, 302:1545–1548, 2003.

[111] H. Kafa *et. al.* The interaction of carbon nanotubes with an *in vitro* blood-brain barrier model and mouse brain *in vivo*. *Biomaterials*, 53:437–452, 2015.

[112] R. Singh *et. al.* Tissue biodistribution and blood clearance rates of intravenously administered carbon nanotube radiotracers. *Proc. Natl. Acad. Sci.*, 103(9):3357–3362, 2006.

[113] Z. Liu *et. al.* Circulation and long-term fate of functionalized, biocompatible single-walled carbon nanotubes in mice probed by raman spectroscopy. *Proc. Natl. Acad. Sci.*, 105(5):1410–1415, 2008.

[114] H. Ulrich *et. al.* *In vitro* selection of RNA molecules that displace cocaine from the membrane-bound nicotinic acetylcholine receptor. *Proc. Natl. Acad. Sci.*, 95:14051–14056, 1998.

[115] A. Condon, D. Harel, J. N. Kok, A. Salomaa, and E. Winfree, editors. *Algorithmic Bioprocesses*. Natural Computing Series. Springer, 2009.

[116] N. Wong Shi Kam *et. al.* Carbon nanotubes as multifunctional biological transporters and near-infrared agents for selective cancer cell destruction. *Proc. Natl. Acad. Sci.*, 102(33):11600–11605, 2005.

[117] S. Ghosh *et. al.* Increased heating efficiency and selective thermal ablation of malignant tissue with DNA-encased multiwalled carbon nanotubes. *ACS Nano*, 3(9):2667–2673, 2009.

[118] X. Tu *et. al.* DNA sequence motifs for structure-specific recognition and separation of carbon nanotubes. *Nature*, 460:250–253, 2009.

[119] Z. Liu *et. al.* siRNA delivery into human T cells and primary cells with carbon-nanotube transporters. *Angew. Chem. Int. Ed.*, 46:2023–2027, 2007.

[120] D. Cai *et. al.* Highly efficient molecular delivery into mammalian cells using carbon nanotube spearing. *Nature Methods*, 2(6):449–454, 2005.

[121] P. H. Siegel and V. Pikov. THz in biology and medicine: towards quantifying and understanding the interaction of millimeter- and submillimeter-waves with cells and cell processes. In *SPIE Photonics West, BiOS*, page 75620H, San Francisco, CA, USA, January 2010.

[122] V. Pikov *et. al.* Modulation of neuronal activity and plasma membrane properties with low-power millimeter waves in organotypic cortical slices. *J. Neural Eng.*, 7(4):045003, 2010.

[123] J. A. Rohas-Chapana *et. al.* Enhanced introduction of gold nanoparticles into vital *acidothiobacillus ferrooxidans* by carbon nanotube-based microwave electroporation. *Nano Lett.*, 4(5):985–988, 2004.

Appendix A - Micro and Nanofabrication Recipes

Photolithography

Microchem LOR 3A

Spin	500 rpm for 30 s, 6000 rpm for 45 s
Bake	180°C hotplate for 5 minutes

Clariant AZ5214E, Image Reversal Mode (-ve)

Spin	500 rpm for 30 s, 6000 rpm for 45 s
Soft Bake	95°C hotplate for 4 minutes
Exposure	2 s, 350 nm (i-line), 290 W
Post Bake	115°C hotplate for 2 minutes
Flood Exposure	10 s
Development	160 s, MF-319 with gentle agitation

Micro Resist Technology ma-N 2410 (-ve)

Spin	500 rpm for 30 s, 5000 rpm for 45 s
Soft Bake	150°C oven for 30 minutes
Exposure	60 s, 350 nm (i-line), 290 W; divide exposure into twelve 5 s exposures, spaced 1 minutes apart
Development	90 s, MF-319 with gentle agitation

Clariant AZ9245 (+ve)

Spin	500 rpm for 30 s, 2000 rpm for 45 s
Soft Bake	110°C hotplate for 2 minutes
Exposure	50 s, 350 nm (i-line), 290 W
Development	8 minutes, MF-319 with gentle agitation
Post Bake	80°C oven for 6 hours, if using as a hardmask

Electron Beam Lithography

Microchem PMMA 950 A4 (+ve)

Spin	3000 rpm for 60 s
Soft Bake	180°C hotplate for 90 s
Exposure	1300 μC / cm^2, 100 keV gun voltage
Post Bake	100°C hotplate for 90 s
Development	30 s in 1:3 MIBK / IPA, with gentle agitation
Stop Bath	30 s in IPA

Zeon ZEP520A (+ve)

Spin	500 rpm for 30 s, 3000 rpm for 60 s
Soft Bake	180°C hotplate for 180 s
Exposure	150 μC / cm^2, 100 keV gun voltage
Development	60 s in ZED-N50, with gentle agitation
Stop Bath	20 s in MIBK

Deposition

Alumina Sputtering

Target	Al on DC gun
Gas 1	O_2, 13 sccm
Gas 2	Ar, 100 sccm

Chamber Pressure	10 mTorr
Power	400 W
Deposition Rate	10 nm / min

Dry Etching

Use Santovac 5 vacuum pump oil as thermal backing for mounting all samples on carrier wafer in the ICP etcher. It is far superior to Fomblin, and dissolves easily in IPA and acetone.

Cryoetch, Oxford Instruments system 100 ICP

Table Height	20 mm
Helium Backing	10 sccm
Chamber Pressure	10 mTorr
Temperature	-140°C
Gas 1	O_2, 10.5 sccm
Gas 2	SF_6, 90 sccm
RIE Power	10 W
ICP Power	1000 W

Bosch Etch, Oxford Instruments system 100 ICP

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	20 mTorr
Temperature	15°C
Etch Rate	3.6 μm / min
ETCH STEP	
Gas 1	SF_6, 160 sccm
RIE Power	30 W
ICP Power	1750 W
Time	15 s

DEPOSITION STEP

Gas 1	C_4F_8, 140 sccm
RIE Power	10 W
ICP Power	1750 W
Time	10 s

GaAs Waveguide Etch, Oxford Instruments system 100 ICP

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	4.6 mTorr
Temperature	20°C
Gas 1	$SiCl_4$, 7 sccm
Gas 2	Ar, 10 sccm
RIE Power	60 W
ICP Power	250 W
Etch Rate	440 nm / min

GaAs Backside Etch, Oxford Instruments system 100 ICP

This etch stops on $Al_{0.8}Ga_{0.2}As$ with infinite selectivity. Run the ICP in RIE mode. Do a 15 minute O_2 plasma clean on sample first, and also before removing sample after completion.

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	70 mTorr
Temperature	30°C
Gas 1	$SiCl_4$, 13.33 sccm
Gas 2	SF_6, 6.67 sccm
RIE Power	75 W
ICP Power	0 W
Etch Rate	1.83 μm / min

Silicon Nitride Etch, Oxford Instruments system 100 ICP

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	10 mTorr
Temperature	15°C
Gas 1	SF_6, 45 sccm
RIE Power	50 W
ICP Power	2000 W
Etch Rate	150 nm / min

Silicon Dioxide Etch, C_4F_8 / O_2, Oxford Instruments system 100 ICP

This etch is fast, but will damage alumina masks (and whatever is underneath) and will oxidise chromium.

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	7 mTorr
Temperature	20°C
Gas 1	C_4F_8, 40 sccm
Gas 2	O_2, 3 sccm
RIE Power	200 W
ICP Power	2100 W
Etch Rate	550 nm / min

Silicon Dioxide Etch, C_4F_8 / SF_6, Oxford Instruments system 100 ICP

This etch has infinite selectivity against alumina, and is excellent for waveguides.

Table Height	20 mm
Helium Backing	5 sccm
Chamber Pressure	30 mTorr
Temperature	20°C
Gas 1	C_4F_8, 17 sccm

Gas 2	SF_6, 70 sccm
RIE Power	150 W
ICP Power	800 W
Etch Rate	80 nm / min

RIE Oxygen Plasma, 50-50-50

Use this recipe to ‘descum’ lithography before deposition. It can also be used to etch parylene (etch rate 100 nm / min).

Chamber Pressure	50 mTorr
Gas 1	O_2, 50 sccm
RIE Power	50 W

RIE Oxygen Plasma, 20-20-80

Use this recipe for cleaning.

Chamber Pressure	20 mTorr
Gas 1	O_2, 20 sccm
RIE Power	80 W

Wet Etching

Alumina Etch

Heat phosphoric acid (85%) on hotplate to 60°.
Etch rate is 10 nm / min.

Nickel Etch

This etch is safe for GaAs, Si and SiO_2.
To remove 50 nm of Ni, heat Transene TFG Nickel Etchant on hotplate to 90°. Etch chip for 60 s.
Rinse in DI water.

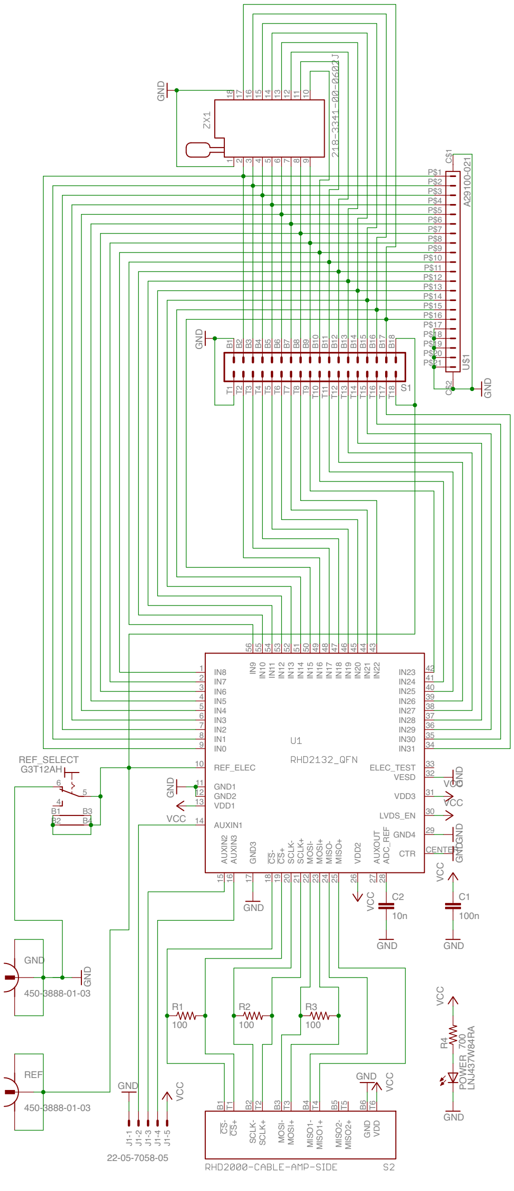
Wafer Bonding

SiO_2-SiO_2 Bonding

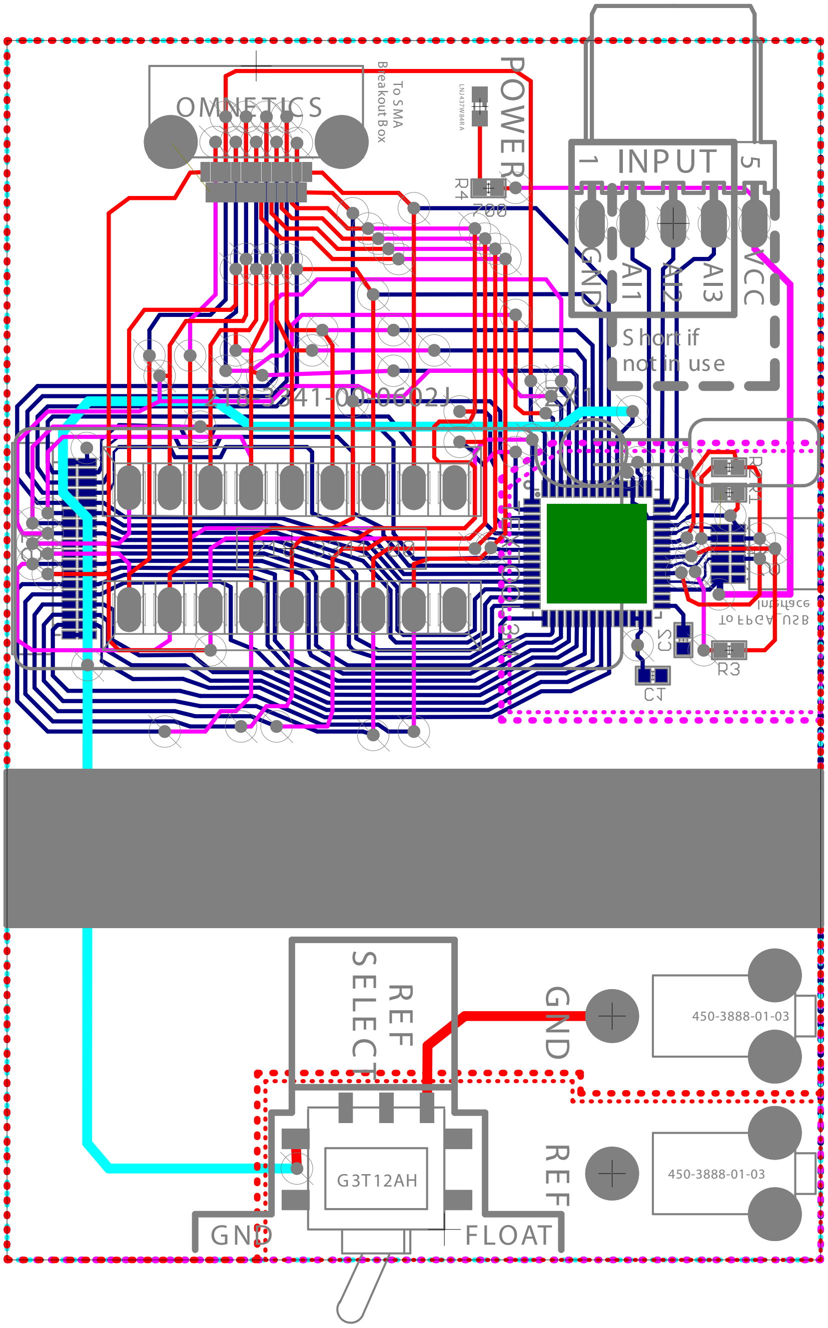
- Perform a 20-20-80 RIE O_2 plasma treatment on both wafers for 2 minutes
- Dip each wafer in 0.025% HF for 1 minute, immediately after O_2 plasma treatment
- Spin dry
- Dip each wafer in NH_4OH (29%) solution
- Spin dry
- Manual mating (will bond on contact, so align carefully). Once bonded, cannot be dismantled and reprocessed unless silicon oxide is stripped away and regrown / redeposited.
- Bond in wafer bonder (Suss SB6L) at 300°C for 3 hours. Tool pressure will depend on your sample / wafer size. Use 500 N of force for every 10 mm × 10 mm of area.

Appendix B - 32-Channel Headstage

Schematic



Printed Circuit Board Design



List of Components

1x	Intan Technologies RHD2132 Digital Electrophysiology Interface Chip
1x	Omnetics Bi-Lobe Connector, Female (21), A29100-021
1x	Omnetics Nano Connector, Female (36), Dual w/Guide Posts, A79025-001
1x	Omnetics Nano Connector, Male/Female Polarised (12), A79623-001
1x	3M Textool 2.54mm DIP Socket, 218-3341-00-0602J
1x	Molex KK Right-Angle Gold 2.54mm, 0022122054
1x	NKK Switches G3T12AH Process Sealed Ultra-Miniature SMT Toggle
2x	Cambion Test Jack 1.2mm, 450-3888-01-03
1x	Panasonic Amber LED 0603 590nm, LNJ437W84RA
3x	100Ω 5% 1/10W, Panasonic Resistor 0402, ERJ-2GEJ101X
1x	698Ω 1% 1/10W, Panasonic Resistor 0402, ERJ-2RKF6980X
1x	10nF 25V, Murata Ceramic Capacitor 0402, GRM155R71E103KA01D
1x	100nF 25V, TDK Ceramic Capacitor 0402, C1005X5R1E104K050BC

Appendix C - List of Key Equipment

Test Equipment

Stanford Research Systems SR844 RF Lock-In Amplifier

Stanford Research Systems SR810 DSP Lock-In Amplifier

Stanford Research Systems SR560 Low-Noise Voltage Preamplifier

Stanford Research Systems SR650 Programmable Bandpass Filter

Rigol DG5072 70 MHz Function Generator, 2-Channel

Rohde & Schwarz SMC100A Signal Generator 9 kHz - 3.2 GHz with OCXO

Agilent 53230A Universal Frequency Counter 350 MHz, 15 GHz with OCXO

Agilent 33250A 80 MHz Function Generator

Agilent 33220A 20 MHz Function Generator

Agilent DSOX4024A Oscilloscope: 200 MHz, 4-Channel

Agilent 8753ES Vector Network Analyzer 30 kHz - 6 GHz with OCXO

Hewlett Packard 3577A Vector Network Analyzer 5 Hz - 200 MHz

Tektronix 2430A 150 MHz, 100 MS/s 2-Channel Oscilloscope

Keithley 2601A Sourcemeter

Photonics

FiberControl MPC1-01-1-2-FC-APC-D Motorized Polarization Controller

Newport λ Commander OSP-9500C Programmable Spectral Source

Newport 501 Laser Diode Driver

Newport 325 Temperature Controller

ILX Lightwave LDM-4984 Laser Diode Mount

Agilent 8163B Lightwave Multimeter with 81689A tuneable Laser, 81533B & 81524A Optical

Head

Thor Labs 904 nm, 10 mW Laser Diode, L904P010

Thor Labs 980 nm, 10 mW Laser Diode, L980P010

TriQuint Optoelectronics 1539.77 nm, 10 mW Electroabsorption Modulated Laser, E2502G47

Philips 10 Gbps 850 nm VCSEL, 1x4 array, ULM850-10-TN-N0104U

Menlo Systems FPD510-FV High-Sensitivity PIN Detector, 400-1000 nm, Free Space, 0-250 MHz

Menlo Systems FPD510 High-Sensitivity PIN Detector, 850-1650 nm, Fibre Coupled, 0-250 MHz

Newport Si Photoreceiver, 1801-FS, 320-1000 nm, 0-125 MHz

Newport InGaAs Photoreceiver, 900-1700 nm, 30 kHz - 1 GHz, FC Connector

Lucent Technologies 2623N 10 GHz Electro-optic Modulator

Oz Optics Fused Polarisation Maintaining Splitter FUSED-12-1550-8/125-50/50-3A3A3A-1-1-PM

WT&T Technology CL5 Tapered-Cone Lensed Polarisation Maintaining 1550 nm Optical Fibre ($2.5\ \mu$m $\times$ $2.5\ \mu$m spot size, mounted in silicon v-groove with slow axis of fibre oriented in the plane of the substrate)

Sutter MPC-200 Dual Micromanipulator Controller

Sutter MP-285/M Micromanipulator Mechanical

Sutter MP-265/M Narrow Format Micromanipulator Mechanical

Electrophysiology

Kopf Instruments Model 940 Digital Stereotaxic Frame

Zeiss Stereo Discovery V8 Zoom Microscope, 0.5x Achromatic Objective

Intan Technologies RHA1016 16-Channel Biopotential Amplifier Array

Intan Technologies RHD2132 Digital Electrophysiology Interface Chip

Intan Technologies RHD2164 Digital Electrophysiology Interface Chip

Data Acquisition

National Instruments PXI-6259 DAQ Card

National Instruments PXI-1033 Chassis

National Instruments BNC-2090A Terminal Box

Appendix D - Wireless Transdermal Optical Communication

Introduction

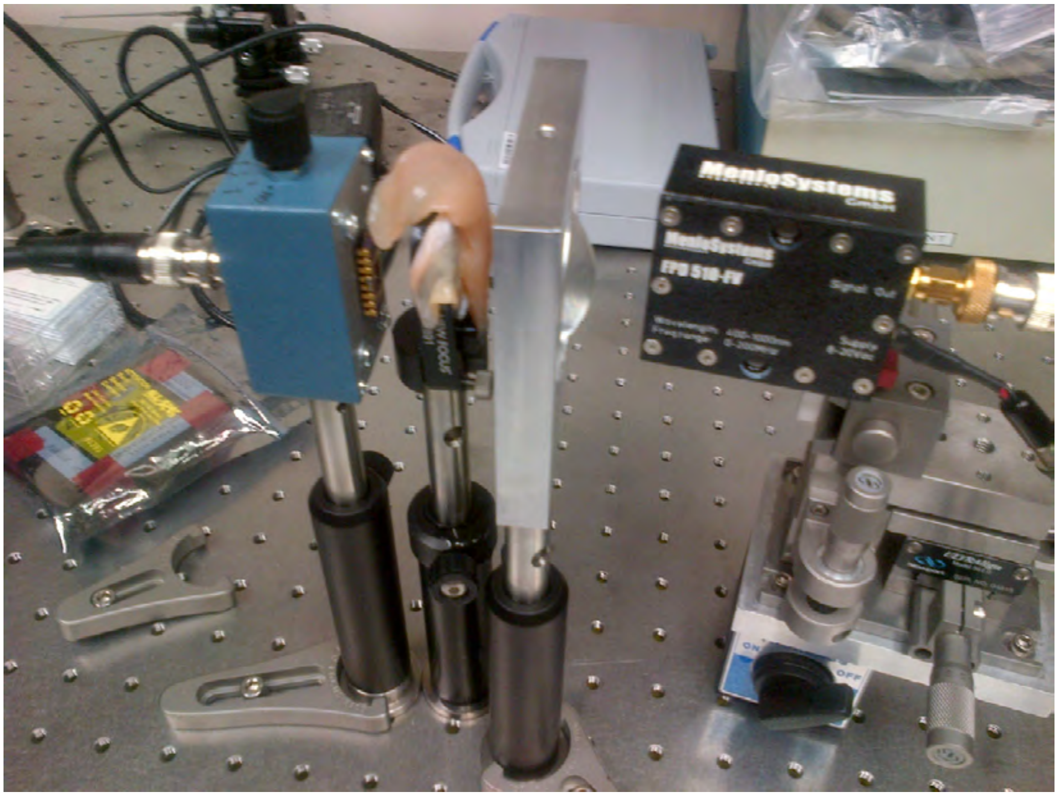
The purpose of the following set of experiments were to test whether 850 nm emission from a 10 Gbps vertical cavity surface emission laser (VCSEL) could be modulated to transmit data of 30 kHz bandwidth in real-time on a radiofrequency carrier wave, through skin and muscle tissue. We propose to use a VCSEL with such performance characteristics as the output for a micron scale neural probe and implantable glucometer, and thus ascertaining whether they are realistic specifications is an important part of the device development.

Experimental Setup

A 10 Gbps 850 nm VCSEL (Philips ULM850-10-TN-N0104U, 1x4 array) was wire bonded to a chip carrier, and wired to a BNC connector though a custom built mounting enclosure. A negative meniscus lens was used to collect the divergent laser light and focus it onto a *pin* InGaAs photodetector, with a 400-1000nm flat response and DC-200 MHz bandwidth (Menlo Systems FPD510-FV). Chicken muscle and skin tissue $\sim$5 mm thick was interspersed between the VCSEL output and collecting lens.

A signal generator (Agilent 33220A) was used to generate a carrier signal, which we varied from 1-10 MHz. This had an offset voltage of 1.6 V and a peak-to-peak amplitude of 200 mV. A second identical signal generator was used to modulate the carrier generator using either amplitude or frequency modulation, with frequencies ranging from 100 Hz-30 kHz.

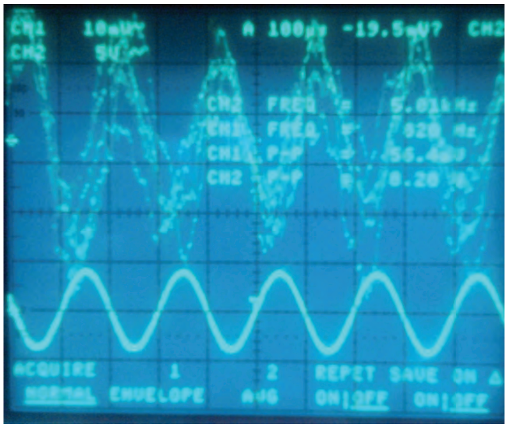
The carrier generator was connected directly to the VCSEL, and the TTL synchronization output from the carrier generator was connected to the frequency input of an RF lock-in amplifier (Stanford Instruments SR844). The output from the photodetector was then connected to the lock-in amplifier input, which locked on to the TTL sync output. This enabled demodulation of amplitude modulated signals. For frequency modulation of the carrier, to demodulate the signal the internal lock-in amplifier signal generator was used instead of triggering off the carrier generator TTL sync output, and manually set as close to the possible to the carrier frequency. The modulation generator and lock-in outputs were concurrently displayed on an oscilloscope for comparison. The lock-in output was filtered through an 8-pole 300 Hz high-pass filter to ameliorate 60 Hz interference.

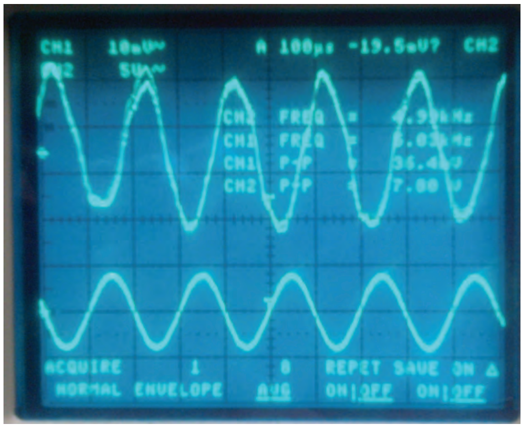


Results

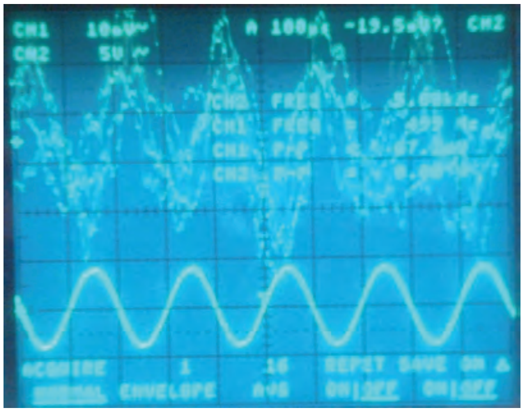
Initially, the collecting lens was placed in front of the VCSEL, and the photodiode was placed at the focus. An attached 3-axis micromanipulator was used to maximize the demodulated signal output to ensure optimum alignment. As an initial test, a hand was placed in the path of the VCSEL beam and this completely eliminated the demodulated signal.

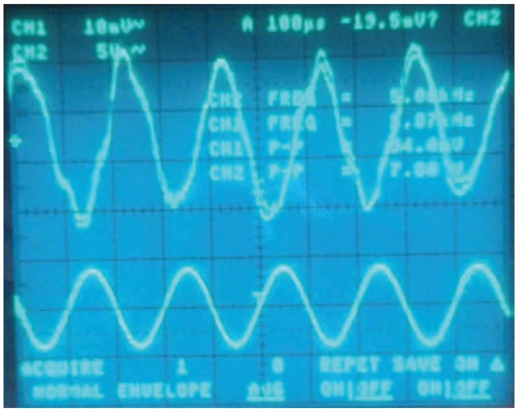
As the VCSEL is intended to transmit through human tissue that is substantially thinner than this, we tried interspersing chicken skin and muscle. First skin was tried. This substantially scattered the laser light, and the photodetector position was no longer optimal for signal detection. It had to be moved much closed to the lens, and was manually adjusted to maximize the demodulated signal. As shown below, a sinusoidal modulation signal of 5 kHz can clearly be detected in real-time (a 3 MHz carrier was used). The trace on the right is with 8x averaging.





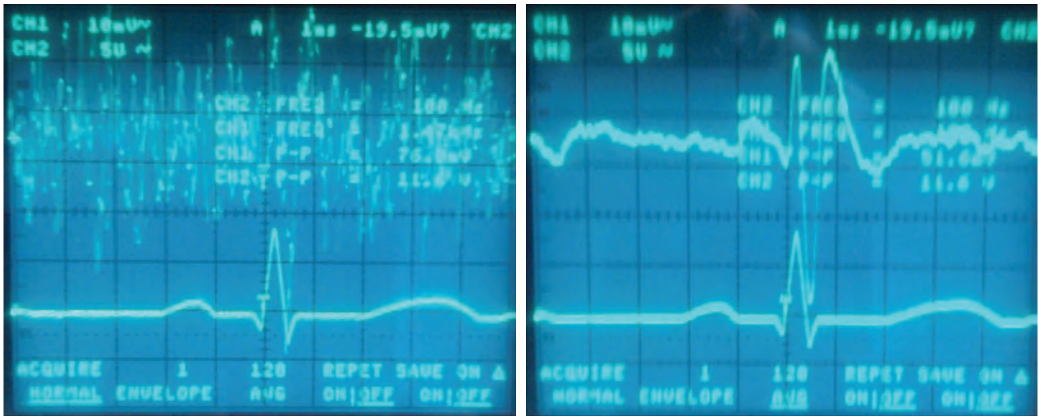
Next, muscle was interspersed together with the skin. The amplitude of the signal dropped, but it was still easily detectable as shown below (trace on the right is with 8x averaging).





Finally, we attempted to transmit a cardiac pulse signal, synthesized by the modulation signal generator. Typical action potentials spike at a frequency of around 100 Hz, and thus the cardiac pulse signal was set to this value. As shown below, it was difficult to detect the signals in real time, but the SNR improved greatly with some averaging. Better band-pass filtering should enable real-time detection. The oscilloscope trace on the right is with 120x averaging.

Note that the demodulated signal (top trace) is a differentiated version of the original signal (due to high-pass filtering somewhere in the instrumentation). This plus uneven frequency response distorts the demodulated signal. Both are linear effects and can be compensated for with integration



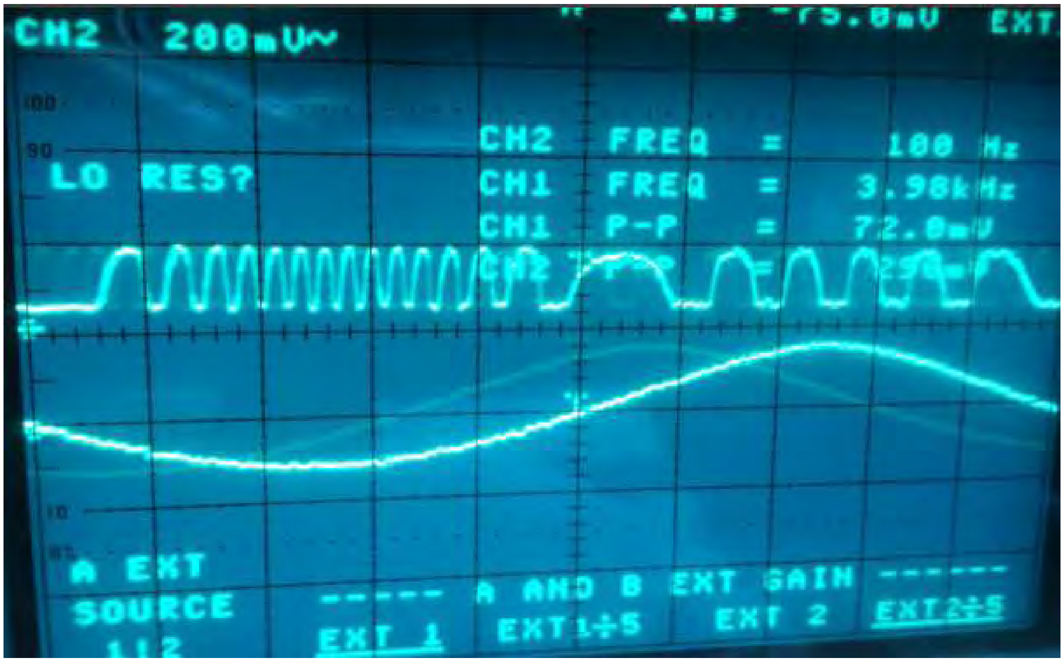
and deconvolution.

Clearly though the SNR of the signal is not as good as we'd like it to be, and some averaging was required. To overcome this, the setup was modified. The output from the photodetector was connected to the signal input of the lock-in amplifier as before. However, the TTL synchronisation output from the signal generator was *not* connected to the frequency input of the RF lock-in amplifier. Instead, the output from the photodetector was concurrently connected to the frequency reference input, and the internal electronics of the amplifier conditioned and locked on to the carrier wave. This kind of demodulation setup reflects more closely how an actual implanted device would be operated, since carrier generation would occur on-chip in the implanted device. Furthermore, through this method phase noise that is introduced in between the laser modulation and signal demodulation is stripped out, since the amplifier is using the carrier on the signal itself to set the reference frequency.

The output below shows the results of using this method. Here though a third signal generator was used to frequency modulate the second one, which in turn amplitude modulated the carrier generator which drives the VCSEL at 3-MHz. The lock-in amplifier was used to demodulate the carrier and recover the 3 kHz frequency modulated signal riding on it. As can be seen, the SNR of the signal is very high and no averaging was required.

Conclusion

Our initial experiments indicate that using a VCSEL for wireless data transmission from a neural probe or glucometer is feasible, with subcutaneous placement in the wrist, or with embedment in the neural cortex provided the detector is placed at the skull with an optical window through it.



Our experiments were successful up to a modulation frequency of 10 kHz, but detection became difficult or impossible above this cut off, probably due to filtering somewhere in the setup and instrumentation. Thus this is not a fundamental limitation.

Amplitude modulation worked best, and we were able to recover signals in real-time with no averaging. We tried carrier frequency modulation, and were successful in demodulating signals, but a lack of a phase locked loop in our setup made consistent detection difficult. We will address this with future experiments.

Further experiments will use a band-pass filter and preamplifier to improve the signal from the lock-in amplifier, and more consistent experiments will be performed to investigate scattering and the effects of blood. A VCSEL will also be wirebonded to a microprobe which will be implanted subcutaneously in a rat, to test whether real-time transmission is feasible through living tissue.