COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author’s moral rights if you:
- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author’s reputation

For further information contact the University’s Copyright Service.

sydney.edu.au/copyright
Readout and Control Beyond a Few Qubits:
Scaling-up Solid State Quantum Systems

John Murray Hornibrook

A dissertation submitted at
The University of Sydney

in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

School of Physics,
The University of Sydney
December 29, 2015
Abstract

This thesis presents new methods and approaches for controlling and reading out qubits towards the operation of a scalable quantum machine. The experiments demonstrate advances over existing brute-force techniques to the problems of readout, applying qubit operations, and managing an interface to classical computation. The results include approaches to suppressing parasitic fields in superconducting resonators, custom software for superconducting electronics, dispersive readout and multiplexed readout of quantum dots, and cryogenic high frequency switches for qubit control. Finally, a scalable distributed architecture for a solid state qubit system is presented incorporating these individual components. The demonstrations presented in this thesis typically use basic elements and individual spin qubits, however the focus is on approaches that can scale to larger numbers of qubits necessary for useful computation. The thesis looks out beyond the single and few qubit experiments of today to the capabilities needed in order to control large-scale quantum systems.

The first series of experiments involves superconducting coplanar waveguide resonators, low-loss circuits used in the field as a quantum bus, enabling coupling and readout of both superconducting and semiconductor qubits. The effect of parasitic electromagnetic fields is explored, and resonator losses are shown to increase due to dissipation in the coupled environment. Low temperature measurement and simulation demonstrates suppression of this effect,
and a loaded $Q$-factor of $2.4 \times 10^5$ is shown. In the limit where intrinsic losses dominate, the dependence of two-level system effects on drive power and temperature can be directly measured. An understanding of this electromagnetic environment, so-called “microwave hygiene”, will become more important as systems with dozens of quantum elements and microwave ports emerge.

Two techniques to improve readout of qubit arrays are detailed. Dispersive readout of a spin qubit is presented using a dot-defining electrode as a sensitive detector. Similar performance to existing charge sensing methods - the quantum point contact - is shown, allowing readout of arrays of qubits with no increase in the footprint and the device level. Readout from multiple channels - using both gate-sensors and quantum point contacts - is demonstrated via frequency domain multiplexing with a low loss, lumped element superconducting chip. Three channel readout of a double quantum dot and ten channel readout using field-effect transistors is shown, and limits due to bandwidth and footprint are given for larger systems.

The scalable readout provided by these two techniques is incorporated into a detailed architecture for scalable manipulation of solid state qubits. A control scheme is presented as an alternative to prohibitive brute force wiring of individual qubits from room temperature. The building blocks of this scheme - reflective switches operating at millikelvin temperatures - are designed and characterised, and compound devices with multiple switches and planar lumped element components are demonstrated. As a proof of concept for the architecture, a double quantum dot is manipulated through a millikelvin switch matrix which is itself controlled by a cryogenic semiconductor field-programmable gate array. These preliminary efforts are, in principle, scalable and form part of ongoing work towards a scalable quantum processor.
## CONTENTS

1. *Introduction* ...................................................... 1

2. *Background* ...................................................... 6
   2.1 Quantum Computation ........................................... 6
      2.1.1 The Qubit .................................................. 6
      2.1.2 Decoherence and Error Correction ....................... 8
      2.1.3 Quantum Algorithms and Scaling ......................... 9
   2.2 Gallium Arsenide and Quantum Dots ......................... 12
      2.2.1 Gallium Arsenide Heterostructure ....................... 12
      2.2.2 1-D and 0-D Semiconductor Systems ..................... 14
      2.2.3 Double Quantum Dots .................................... 15
   2.3 Superconducting Systems ..................................... 19
      2.3.1 Transmission line resonators ........................... 19
      2.3.2 Resonator Decay .......................................... 22
      2.3.3 Josephson Junctions ..................................... 24
      2.3.4 Superconducting Qubits .................................. 25
3. Superconducting Resonators with Parasitic Electromagnetic Environments .................................. 28
   3.1 Abstract ............................................................... 28
   3.2 Introduction ....................................................... 28
   3.3 Experimental Design .......................................... 31
   3.4 Asymmetric Lineshape .......................................... 32
   3.5 Two Level Systems ............................................. 36

4. Simulation Software for Single Flux Quantum Circuits ............ 39
   4.1 Abstract ............................................................... 39
   4.2 Introduction ....................................................... 39
   4.3 Nodal Analysis ................................................. 44
   4.4 Once-Only Matrix Inversion ................................. 48
   4.5 Other Features .................................................. 51
   4.6 Discussion .......................................................... 53

5. Dispersive Readout of a Few-Electron Double Quantum Dot with Fast rf Gate-Sensors .................. 54
   5.1 Abstract ............................................................... 54
   5.2 Introduction ....................................................... 54
   5.3 Experimental Design .......................................... 57
   5.4 Characterisation .................................................. 59
6. Frequency Multiplexing for Readout of Spin Qubits .......................... 65
   6.1 Abstract ................................................................. 65
   6.2 Introduction ............................................................ 65
   6.3 Experimental Design ................................................. 67
   6.4 Quantum Dot Readout ................................................. 69
   6.5 Scaling ................................................................. 73

7. Cryogenic Control Architecture for Large-Scale Quantum Computing 76
   7.1 Abstract ................................................................. 76
   7.2 Introduction ............................................................ 77
   7.3 Control Micro-architecture ........................................... 79
   7.4 Implementation of the Control Architecture ......................... 80
   7.5 Switching Matrix .................................................... 81
      7.5.1 HEMT Switching Elements ........................................ 83
      7.5.2 Capacitive Switching Elements ................................... 84
      7.5.3 2:2 Switch Matrix ............................................... 87
   7.6 Cryogenic Logic ..................................................... 90
   7.7 Semiconductor Qubit Control ....................................... 91
   7.8 Discussion ........................................................... 94

8. Future Direction .............................................................. 97
Appendix

A. Niobium Fabrication ........................................ 104

B. GaAs Switch Fabrication .................................. 109

C. Circuit Components used in SPICE Software .......... 112
Statement of Student Contribution

The work in Chapter 3 is based on the following papers:


The work in Chapter 5 is based on the following paper:


The work in Chapter 6 is based on the following paper:


The work in Chapter 7 is based on the following paper:

If not otherwise acknowledged, the work presented in this thesis is my own.

Signed .........................................

John Murray Hornibrook
Date: December 29, 2015
Acknowledgements

Quite early on in my PhD, I saw a brilliant production of Goldini’s comedy “Servant of Two Masters”, where the central character aims to satisfy his hunger by simultaneously satisfying the whims of two masters, enduring as best he can the associated tribulations.

On an unrelated note, I had two supervisors during my PhD.

First, I am grateful to Emma Mitchell for her significant investment in time, energy and patience. I’m not sure any single story would do justice to Emma’s diligence, although I will try: The only reason I was granted an hour break in our nine hour cleanroom training sessions was so that Emma could go for a run. Emma also fitted our lab with much appreciated beaurocracy-shielding, so that my torment and head-banging were at least the consequence of physics.

The CSIRO team are, without exception, generous with their time in helping students. In particular, I should mention the combination of microwave engineering and motorcycle repair I learned from Chris Lewis (I don’t ride, of course), and the cleanroom hours made less terrible by Jeina Lazar’s deep well of nanofabrication advice and stories of the Kurdish people.

I should thank also Simon Lam for help with niobium, Rex Binks and Peter Sullivan who can make anything, and in Shane Keenan I discovered the remarkable therapeutic benefits derived from Irish post-docs willing to believe
anything about Australian spiders.

I am also very fortunate to have had the supervision and tutelage of David Reilly. To accompany an encyclopedic knowledge, David projects a measured serenity that was, against all odds, unshaken throughout my PhD. Although Giorgio Frossati’s attempt to lift a dilution refrigerator with an underspecced pallet lifter while condensing $^3$He was a commendable effort. I should in particular remark on David’s meditative calm as I patronisingly explained how we perform rf sensing - considering reference [131], he has probably seen it before.

I should also mention (and where else could I?) that David has either a curious lack of circadian rhythm or a well-configured email client, the result of which being an unbroken chain of emails through some nights. Even more suspicious is the “sounds good” reply that arrives some 17 seconds after you have sent your many-page treatise.

Experimental physics is necessarily a team effort, however my fellow lab rats at Sydney University provided support far beyond helping to turn the often-heavy hamster wheel. I’ve probably burned the midnight oil with James Colless more than anyone else - thinking on it now, the people I’ve spoken to between the hours of 2am and 5am over the last few years are strong favourites to be either drunk or James. I should probably also take this opportunity to apologise for the large number of live shrimp all over our post-doc Sylvain Blanvillain’s kitchen floor - in my defence, that wasn’t what I had in mind when I offered to help with dinner.

I am grateful to Xanthe Croot for her blind optimism helping me to hope for the best and to Alice Mahoney for her dry cynicism helping me to laugh at the worst, and after sharing a lab with Ewa Rej I encourage other groups to employ at least one person to make questionable claims about Polish history.

---

and dating strategies. In addition, I derived no small amount of enjoyment watching Sebastian Pauka and Ian Conway Lamb’s impossibly fast progress in transitioning from minions to fully functioning serfs.

I must also thank my Father for encouraging me to study anything I chose, so long as ‘anything’ was not Law, did not rhyme with Law, and did not look like Law if you squint. And of course my Mother for her rose-tinted glasses crediting my absent-mindedness to “thinking of higher things” when I was probably thinking about Transformers or Batman.

Finally, I need to thank my wife Ivy, for amongst many things her incredible patience and unwavering calm. I recall Ivy explaining her understanding of Schrodinger’s Cat, after coming across it in an unclear resource, where she finished with the rather despondent “But it doesn’t matter what you do, because the cat always dies”. In my post-thesis years, I’ll endeavor to show that there’s scope for more optimism than this - both in our marriage and in quantum computation.
1. INTRODUCTION

The inclusion of entanglement and superposition of quantum states as resources in a computing scheme provides new tools for the development of quantum algorithms. These properties can, in-principle, be exploited to solve a variety of problems that are presently intractable for a classical machine.\textsuperscript{1} This section will outline the well-traversed motivation for quantum computation, and flag the challenges posed to the quantum information processing community in realising such a machine.

The quantum algorithm typically presented is Shor’s algorithm for integer factorisation and the discrete logarithm problem. This algorithm solves in polynomial time a problem that takes classical computers exponential time\textsuperscript{2} and has alluring consequences for public-key cryptography schemes. Closely following is Grover’s algorithm, where lookups to a black box - equivalent to inverting an unknown function - run in $O(\sqrt{n})$ time. In addition to the banal application of database searching, this provides a quadratic speedup to problems involving a black box lookup, and is equivalent to halving the bit-length of various cryptographic cyphers.

Considering the large collection of algorithms that have been developed in a short space of time and in the absence of hardware, it seems reasonable to

\textsuperscript{1} Complexity theory enthusiasts, fearful for various versions of the Church-Turing thesis, can find solace in the now-legitimate use of words like “hypercomputation” and “super-Turing”.

\textsuperscript{2} It should be noted that while this is thought to be the case, an exponential time lower bound has not been demonstrated.
anticipate further contributions. For instance, integer factorisation - solved in polynomial time by Shor’s algorithm - is a special case of the Hidden Subgroup Problem. A more general solution would have implications for several intractable graph and lattice problems.

In addition to the development and implementation of powerful algorithms, a machine based on quantum phenomena can be used in the simulation of quantum systems. The complexity of a classical simulation of a system is exponential in the number of particles in the system (as is the degrees of freedom). It is conceivable that such a system can be simulated by a quantum computer by utilising its own (large) number of degrees of freedom, making possible the modelling of, for instance, large biological molecules.

The absence of sufficient hardware to run these powerful algorithms points to the difficulty in constructing such a machine. The state of a quantum bit is fragile, to put it mildly, especially when compared to a static voltage on a capacitor as in classical machines. For successful computation, the interaction between the qubit and the environment must be weak enough so that the qubit decoherence is on a longer timescale than the qubit operations. For existing solid state qubit architectures, this necessitates low temperature, low noise and precise control over both static and changing electromagnetic fields. An additional difficulty in achieving this isolation is that qubits must also be coupled to other qubits in the system and to classical readout electronics so that computation can be performed.

These difficulties preclude qubit fidelities sufficient to perform computations beyond the capabilities of classical computers, at least without some error correction scheme. Not only is a fault-tolerant scheme required, but classical redundancy-based approaches to error correction are unusable in a quantum machine due to the inability to clone a quantum state. An alternative method implementing additional qubits is to store the information of a single logical qubit in the highly-entangled state of multiple physical qubits. This approach forms the basis of quantum error correction, and permits fault-tolerant oper-
Quantum fault-tolerant error correction schemes are certainly promising, although in order to realise them we are faced with engineering a quantum processor with coherent manipulations of millions of qubits. This is a difficulty that encompasses not only the qubit device, but also a wide range of associated electronics to facilitate qubit operations and readout. These tools include a host of cryogenic electronics (filters, attenuators, bias tees, couplers, amplifiers) and room temperature instrumentation (oscilloscope, arbitrary waveform generator, digital-analog-converter, pulse generator).

Abstracting these requirements of the system slightly, a quantum processor needs slow-changing electromagnetic control (for instance, defining semiconductor quantum dots or biasing superconducting qubits), fast-changing electromagnetic fields (various qubit operations) and readout. In addition, implementing quantum algorithms requires a high-level logic including assigning voltages, directing pulses and performing calculations on readout signals. The difficulty in implementing this control, and the point of this introduction, is that naive brute force approaches scale unfavourably with the number of qubits, so that systems with (as few as) tens of qubits become difficult to engineer.

The development of cryogenic control techniques, both on-chip and otherwise, to alleviate these constraints for scalable quantum information experiments, is the theme of this thesis. In the following chapters, techniques and experiments are presented working towards this goal, addressing the various qubit

---

\[3\] Following calculations in [1] and assuming a generous 99.9% qubit fidelity and 100 ns measurement time, factorisation of a 1000-bit number could be completed in 3.3 hours with \(2 \times 10^8\) qubits. This number could be reduced by higher fidelities (99.99% \(\Rightarrow\) \(2 \times 10^6\) qubits) or tolerance for longer computation time (24 hours \(\Rightarrow\) \(3 \times 10^7\) qubits).
1. Introduction

requirements.

In Chapter 2, ideas touched upon only briefly in the present chapter are re-
visited, with a more detailed overview of qubits and qubit decoherence. The
remainder of the chapter is split between background information pertinent
to two solid state architectures. First, heterostructures based on gallium ar-
senide materials, used in this thesis for quantum dots and cryogenic electron-
ics, are introduced, and the subsequent formation of spin qubits is discussed.
Second, superconducting electronics is described as background for Chapters
3 and 4.

Superconducting coplanar resonators, a promising contender for qubit read-
out and coupling distant qubits - especially in superconducting qubit systems
- are discussed in Chapter 3. Specifically, we describe losses associated with
coupling between the resonator and the electromagnetic environment, and
steps that can be taken to increase the loaded resonator quality factor.

Chapter 4 looks at Josephson junction-based superconducting circuits as a
contender for use in quantum control circuitry. An in-house software package
is described, with speed and features well-suited to designing superconduct-
ing electronics.

Chapters 5 and 6 examine techniques for readout of scalable spin qubit sys-
tems. In 5, we read out the state of a quantum double dot using an in-situ
dot-defining gate, obviating the need for quantum point contacts near ev-
ery qubit. In 6, we present and characterise a cryogenic multiplexed readout
scheme so that readout hardware and instrumentation scales favourably with
the number of qubits in the system.

In Chapter 7, the focus is shifted to the problem of qubit control. Since indi-
vidual room temperature control lines for each qubit is not feasible, a scheme
is presented where cryogenic switches direct a small\(^4\) number of bus lines to

\(^4\) Constant in the number of qubits
multiple qubits. We develop components of such a system and demonstrate
its use with cryogenic logic, a quantum double dot, and frequency multiplexed
readout.

Finally, the implications for constructing a scalable quantum computer are
discussed in the context of future work in this field.
2. BACKGROUND

2.1 Quantum Computation

The introductory chapter broached the idea, and the challenges, of using quantum superposition and entanglement for calculations, and now these will be revisited in more detail. The qubit is outlined as a two-level system that will form the building block of a quantum processor. Details are given of other ideas previously touched upon: loss of information to the environment, possible error correction schemes, and the motivations and requirements for scaling to large numbers of qubits.

2.1.1 The Qubit

The qubit is a quantum mechanical two-level system and the fundamental unit of a quantum processor, analogous to the classical bit. A two-level system can be used to implement a qubit provided it allows manipulation of the qubit state, entanglement with other qubits, and readout. The quintessential example is the spin of an electron, and while spin qubits and other architectures will be detailed later, this section will discuss the qubit more abstractly.

The two levels of the qubit can be written as $|0\rangle$ and $|1\rangle$, and the state of the qubit as the superposition

$$
\psi = \alpha |0\rangle + \beta |1\rangle .
$$

This superposition is the first of the significant advantages mentioned for a
quantum machine: the qubit is intrinsically parallel and the qubit space is a complex vector space rather than a more limiting binary space.

The coefficients $\alpha$ and $\beta$ can be interpreted as probabilities so that $|\alpha|^2$ is the probability that measurement returns $|0\rangle$. Imposing the normalisation restriction $\alpha^2 + \beta^2 = 1$ suggests representation on the surface of a sphere, so the wave function can be re-written as

$$\psi = \cos(\theta/2) + e^{i\phi} \cos(\theta/2),$$

for $\theta$ in $[0, \pi]$, $\phi$ in $[0, 2\pi)$, where a global phase has been removed since state measurement yields $|0\rangle$ or $|1\rangle$ only.

The sphere formed is known as the Bloch sphere, and is shown in Figure 2.1. Single qubit manipulations — required for quantum computation — can then be thought of as rotations around various axes on this sphere.

The second distinction between qubits and their classical counterparts is the possibility of forming entangled states between multiple qubits. An entangled state is a multi-qubit state for which there is no Kronecker product of single-qubit states. By way of example, $\phi = \frac{1}{\sqrt{2}} |0\rangle_A |1\rangle_B + \frac{1}{\sqrt{2}} |1\rangle_A |0\rangle_B$ cannot be written as $\phi = (\alpha_A |0\rangle_A + \beta_A |1\rangle_A) \otimes (\alpha_B |0\rangle_B + \beta_B |1\rangle_B)$ for any $\alpha_i, \beta_i$. Entangled states allow the encoding of large amounts of information in small numbers of qubits.

So far, the qubit has been treated as an abstracted and idealised two-level system, and achieving this in a practical
implementation is not straightforward. As alluded to previously, the most significant challenge is managing the coupling between the qubit state and the environment. Preservation of the qubit state requires the qubit be as isolated as possible. At the same time, a degree of coupling to other qubits is required to realise multi-qubit gates. In addition, the qubit needs to be coupled to external control electronics to facilitate single-qubit operations and readout. While different architectures have different decoherence rates, no approach has a comprehensive solution to the difficulty of balancing coupling and isolation.

Some specific solid state qubit implementations will be described in this chapter — semiconductor quantum dots in Section 2.2 and superconducting systems in Section 2.3 — since these are the most relevant to this thesis, although this is far from exhaustive. Qubit realisations not mentioned include ion traps [2, 3], NV centres in diamond [4], nuclear spins [5, 6] and photonic systems [7, 8, 9].

### 2.1.2 Decoherence and Error Correction

The deterioration of the quantum state due to coupling to the environment is known as decoherence, and occurs via a range of processes. Considering the Bloch sphere in Figure 2.1, decoherence can be classed in one of two ways. Relaxation occurs when the $|1\rangle$ state dissipates energy into the environment and relaxes to the $|0\rangle$ state, and the timescale on which this occurs is referred to as $T_1$. Dephasing is an introduction of random rotations on the Bloch sphere, on a timescale referred to as $T_2$.

Decoherence in today’s qubit implementations is severe enough to make scalable quantum computation infeasible [10], and so some error correction scheme is required to allow computation in the presence of imperfect gate operations. A further complication is that classical error correction techniques — largely based on duplication of a digital signal — are not applicable to
quantum systems for two reasons. First, a quantum computation is sensitive to phase errors as well as bit errors. Second, the no cloning theorem disallows perfect copies of an unknown quantum state [11].

Quantum error correction schemes are based upon redundant data encoding, where a single logical qubit is encoded among multiple entangled physical qubits. The simplest is a 3-qubit scheme, where bit flip errors or phase flip errors are corrected [12]. Full quantum error correction, where both bit and phase errors are corrected, first appeared in [13], where the state of a logical qubit is encoded among 9 entangled qubits as an extension of the simpler 3-qubit scheme. The Steane code [14] represents an improvement on this, providing error-correction with only 7 physical qubits for each logical qubit.

The difficulty in developing a viable error correction scheme lies in the practical implementation in addition to the scheme’s fault-tolerant threshold. For instance, a code that involves arbitrary entanglement between more distant neighbours or a three-dimensional lattice is not easily realised. Practical codes have been designed, notably the range of surface codes defined on a two-dimensional lattice requiring only nearest neighbour interaction [1, 15]. The new difficulty is this: the fragility of the qubit states compared to their classical counterparts suggests an error correction scheme will be required, which in turn implies that the number of physical qubits required will be significantly larger than the number of logical qubits needed to implement an algorithm.

2.1.3 Quantum Algorithms and Scaling

The introductory chapter touched upon some quantum algorithms to motivate the scaling of quantum computation. These will now be revisited in more detail with a view to understanding the capabilities and requirements of a useful quantum computer. Here, ‘useful’ means a machine able to solve problems that cannot be solved by the (impressive) classical computational
resources used today. In the language of complexity theory, this is a class of problems $BQP$ — bounded error quantum polynomial time — that can be efficiently\(^1\) solved by a quantum computer that lie outside of $P$, the class of problems efficiently solved by a classical computer.

Problems in this class include integer factorisation and the related discrete logarithm problem. The current best classical approach to integer factorisation is to use a number field sieve [16], with exponential computational time $t \sim \exp\left(2L^{1/3}(\log L)^{2/3}\right)$ where $L = \ln x$ is the size of the input and $x$ the integer being factorised. Shor’s algorithm [17] solves this problem on a quantum computer with computational time $t \sim L^3$. By far the most significant consequence of this algorithm is the implications for the security of RSA cryptography systems\(^2\).

This algorithm is one of several based on the quantum Fourier transform — analogous to the discrete Fourier transform with the point that it can be executed efficiently on a quantum computer — such as the somewhat artificial Deutsch-Jozsa algorithm [18] or algorithms for solving linear systems by constructing an appropriate Hamiltonian [19]. Integer factorisation is a particular instance of the broader Hidden Subgroup Problem, which has implications for a range of graph problems, with other such sub-problems also in $BQP$ such as finding the shortest vector on a lattice [20].

Earlier than the development of these algorithms was the idea to use a quantum computer to simulate quantum systems [21]. A quantum computer with $n$ logical qubits can simulate a $2^n$-dimensional Hilbert space, where the classical machine would require arrays with $2^n$ complex elements. In this way, a quantum computer can efficiently simulate several systems for which there is no efficient classical algorithm [22]. A particular such problem is determining

---

\(^1\) That is, solved in a time $t = f(x)$ where $x$ is the size of the input and $f$ is polynomial.

\(^2\) In RSA, a message $M$ is encrypted as $E = M^s \mod c$, where $s, c$ are public. Decryption is $M = E^t \mod c$ where $t$ is calculated from $s$ and the factors of $c$. If $c = pq$ is chosen for primes $p, q$, determining $t$ depends on finding $p, q$. 

the ground state energy of a molecule, which can be solved in polynomial time on a quantum machine, which also maps to various problems in mathematics [23, 24].

With these concepts in mind, it is possible to estimate the number of qubits required for a useful demonstration of some of these algorithms. In [1], calculations to determine the number of resources needed to implement Shor’s algorithm with a surface code are given. For factorisation of a 2000-bit number with an error rate of $1 \times 10^{-3}$ and 100 ns qubit measurement time, the authors give $\sim 1.0 \times 10^9$ qubits and a computation time of 26.7 hours. Following the calculations in this reference, a smaller number (1000-bit) with a lower error rate ($1 \times 10^{-4}$ — roughly two orders of magnitude better than the error rate threshold) still requires $\sim 2.0 \times 10^6$ qubits and 3.3 hours for operation, with fewer (or more) qubits required at the cost of more (or less) time. Resource estimates for some other algorithms are given in [25] for a variety of architectures and quantum error correction schemes. Of the algorithms mentioned already in this section, the authors give at minimum $10^7$ qubits for estimating the ground state energy of a molecule and the prohibitive minimum of $10^{21}$ qubits for the shortest vector on a lattice.

Looking at demonstrations to date, the ubiquity of Shor’s algorithm provides a useful benchmark. For instance, the number 15 has been factorised with seven nuclear spin qubits [26], with three superconducting phase qubits [27] and with four qubits in a photonic system [7, 8, 9]. The field is faced with a discrepancy between current qubit demonstrations and the eventual number of qubits required for useful computation. Closing this gap requires the development of novel techniques across all aspects of quantum technologies.
2.2 Gallium Arsenide and Quantum Dots

This thesis presents several techniques for scaling up qubit control and read-out, and provides proof-of-concept demonstration with simple quantum systems. For many of the experiments, this involves devices made using gallium arsenide heterostructure, where electrons are confined to a two-dimensional plane at a GaAs / Al$_{0.3}$Ga$_{0.7}$As interface. The readout techniques in Chapters 5 and 6 and the control architecture in Chapter 7 are demonstrated using quantum dots defined in gallium arsenide. In addition, the same gallium arsenide heterostructure is used for a range of switching devices in Chapter 7 designed to manage control pulses among large numbers of qubits.

2.2.1 Gallium Arsenide Heterostructure

In gallium arsenide heterostructures, a quantum well is formed at the interface between GaAs and Al$_x$Ga$_{1-x}$As, where $x$ is usually around 0.3. The layers in the heterostructure are grown using molecular beam epitaxy (MBE), making use of the fact that the lattice parameters of the two materials are nearly identical. A typical layering is shown in Fig. 2.2(a).

The bandgap difference between the two materials forms one wall of the quantum well. The other wall is formed using modulation doping — silicon donors are placed some distance from the interface. The resulting potential well is triangular, and shown in Fig. 2.2(b). When the material is cooled, only one of the sub-bands of the quantum well is below the Fermi energy, confining electrons to a two-dimensional plane referred to as a two-dimensional electron gas (2DEG).

The heterostructure wafer has some additional treatments, as shown in Fig. 2.2(a). Below the MBE-grown gallium arsenide layers is a superlattice: alternating layers of GaAs / AlGaAs to smooth the effect of impurities in the GaAs substrate. Finally, a gallium arsenide capping layer is used above the
Fig. 2.2: Illustration of gallium arsenide heterostructure. (a) Typical profile of a GaAs heterostructure, where electrons are confined at the GaAs / Al-GaAs interface. Several GaAs and AlGaAs layers smooth imperfections in the substrate, and a GaAs capping layer prevents oxidation. (b) The triangular potential well. Only one of the sub-bands lies below the Fermi energy so that other energy levels are unoccupied. (c) Negative voltages on the metallic top gates deplete the 2DEG beneath them and can be used to confine electrons to pockets or to constrict a conducting channel.
final AlGaAs layer to prevent oxidation at the surface.

Electrical contact to the 2DEG is made using a eutectic alloy. AuGe metal is deposited with a Ni adhesion layer and then annealed. The heated metal diffuses into the semiconductor and electrically contacts the 2DEG.

Electrons in the 2DEG are confined in directions parallel to the interface by negative voltages applied to metallic top gates. In this thesis, the metallic gates are gold with a titanium adhesion layer, evaporated after electron-beam lithography patterning. A Schottky barrier is formed at the metal-semiconductor junction so that the top gates can be held at a negative voltage relative to the 2DEG. This depletes regions of the 2DEG directly beneath them and confines the electrons as shown in Fig. 2.2(c). Voltages on the metal top gates are also used during cooling — both the quantum dot devices in Chapters 5, 6 and 7 and the switches in Chapter 7 are cooled with a positive voltage to improve performance. This reduces noise in the case of quantum dot devices and insertion loss in the case of cryogenic switches.

2.2.2 1-D and 0-D Semiconductor Systems

Electron confinement in the two dimensions parallel to the 2DEG is used to explore quantum effects in heterostructure devices. When measuring a current flowing through a section of 2DEG, top gates can be used to restrict conduction to a narrow channel, as can be seen in Fig. 2.2(c). This confinement results in a discretisation of the transverse motion and the conductance through the channel becomes quantised. The channel is referred to as a quantum point contact (QPC) and becomes very sensitive to the surrounding electronic environment\[28, 29, 30\]. It serves as a sensitive electrometer and is used as one of two readout techniques for quantum dots in this thesis (the other technique is presented in Chapter 5).

When top gates are used to confine electrons in the 2DEG to small regions
on the order of the de Broglie wavelength, called quantum dots, the system is considered 0-dimensional and the energy levels of the bound states are discrete. The system can be used to probe quantum effects since the behaviour is analogous to that of a single atom. The number of electrons on the dot is typically small, and can be precisely tuned by varying the top gate voltages.

The spin of an electron in a quantum dot provides an ideal qubit candidate [31]. Spins do not strongly couple to the environment and the electron’s charge allows electronic manipulation and readout. Information can be encoded in the spin state in the presence of a high magnetic field $B$, and a changing magnetic field with frequency $\Omega = g\mu_B B/\hbar$ drives oscillations between the two spin states. Here, $g$ is the gyromagnetic ratio, $\mu_B$ is the Bohr magneton and $\hbar$ is the Planck constant. The use of gate voltages to define the dot allows precise control over the confinement potential and also over the coupling between multiple dots.

A limitation of gallium arsenide quantum dots is decoherence due to the nuclear spins of the host lattice, which electrons will couple to via the hyperfine interaction. These spins result in an effective magnetic field known as the Overhauser field. Spin qubits in these systems are still viable due to the fact that the nuclear spins evolve slowly compared to the electron spins, so the nuclear field can be considered static for electron spin operations. Suppression of decoherence due to this field can then be achieved by applying pulse echo sequences [32, 33] or fast feedback on a shorter timescale than the nuclear bath fluctuations [34, 35].

### 2.2.3 Double Quantum Dots

Computation with the quantum dots in the previous section poses some challenges that can be alleviated by storing the qubit information in the combined spin state of two quantum dots. In addition to avoiding the requirement of a large magnetic field, the basis states can be chosen with $m = 0$ so that
they are unaffected by changes in $B$. Readout of an electron spin can be avoided by exploiting Pauli exclusion to read out the qubit state via a charge measurement.

An electron micrograph of a double quantum dot is shown in Figure 2.3(a), where the gate geometry allows the formation of two adjacent quantum dots with a gate-controlled inter-dot coupling. The area of 2DEG to which electrons are confined is indicated in blue. In (i), by applying a relatively small negative voltage on the centre gates, the wells are not separate and a single quantum dot can be defined. By contrast, a strong negative voltage on the centre gate defines two independent quantum dots (ii). By choosing voltages between these two extremes, two coupled quantum dots can be defined (iii). This is the same device as used in experiments in Chapter 6.

The gate voltages are used to control the electrostatic energy and consequently the number of electrons on each dot. Fig 2.3(b) shows regions corresponding to different numbers of electrons as the gate voltages $V_L$ and $V_R$ are swept. Where a single quantum dot has been defined, increasing either $V_L$ or $V_R$ (that is, so that they become less negative) increases the number of electrons on the dot, indicated by the corresponding number in the figure. Where the dots have been defined and well separated (ii), increasing $V_L, V_R$ increases the number of electrons on the left, right dot respectively. Regions with $m$ electrons on the left dot and $n$ on the right dot are denoted $(m, n)$. If the dots are completely uncoupled, the lines defining the regions are horizontal and vertical.

When the dots are close together (iii), gate voltages on the walls affect the electron occupancy of both dots, so the lines between regions are no longer vertical. Electrons on one dot affect the charging energy of the other, so the lines are no longer perpendicular. In addition, transitions are present (that is, the regions are hexagonal rather than diamonds) due to the possibility of electrons tunnelling between the two dots.
Fig. 2.3: Some key features of a double quantum dot.  
(a) Electron micrograph showing metallic gates in white whose voltages determine the shape of the potential. In (i), a single potential well is defined, forming a single quantum dot. In (ii) and (iii), there are two distinct wells that are either completely independent (ii) or coupled (iii).  
(b) Charge stability diagram showing regions corresponding to different numbers of electrons present in the quantum dots in (a), as the voltages $V_L$ and $V_R$ are swept.  
(c) Charge stability diagrams from the experiment in Chapter 7. The measurement is of conduction through the QPC marked in (a) in the case of a single dot and a coupled double dot.
It should be noted that the quantum dots can be completely emptied with a sufficiently negative gate voltage. Once transitions are no longer seen (in the bottom-left of the images in Figure 2.3(b)), the dot is inferred to be empty, and can consequently be filled with an arbitrary number of electrons by counting transitions.

These diagrams are referred to as charge stability diagrams. Figure 2.3(c) shows charge stability diagrams for the experiment in Chapter 7, where the resistance of a nearby quantum point contact that is sensitive to electrons on each dot is measured. By plotting the derivative with respect to $V_R$, transitions where electrons leave the dots can be seen in the cases of a single dot and a coupled double dot.

To form a spin qubit with a coupled double quantum dot, the qubit information is encoded in the combined spin state of two electrons[36]. Such a pair of spin-1/2 particles can form one of three triplet (spin 1) states and one singlet (spin 0) state. For our qubit bases, the singlet

$$S = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$$

and triplet

$$T_0 = \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}$$

are chosen.

The advantages of such a scheme are numerous. Qubit operations are completely electrical and fast [32]. Both basis states have $m=0$ so are insensitive to magnetic field changes.

Readout of the qubit state, typically difficult for electron spins, can be performed via a spin-to-charge conversion by attempting to move both electrons to the same dot. The singlet state is anti-symmetric and can have both electrons in the ground state. The triplet state is not, so Pauli exclusion forces one electron to remain in each dot. The charge state is easily measured using
a quantum point contact as described previously with details in Chapters 5, 6, and 7.

One possibility, outside the scope of this thesis, is to encode the qubit information in the combined spin state of three electrons. Now both single qubit and two qubit operations can be performed via the exchange interaction. This idea is presented in [37] and realised in [38, 39].

2.3 Superconducting Systems

Superconducting systems show promise for quantum computation due to their inherent low loss and the existence of a non-linear element, the Josephson junction. Low dissipation preserves the quantum state, and a non-linearity in the qubit energy levels allows specific transitions to be addressed. First, superconducting distributed resonators, the focus of Chapter 3, are introduced. Second, three types of qubits based on the Josephson junction are discussed. While the techniques reported in Chapters 6 and 7 are demonstrated using gate-defined quantum dots, it is envisaged that these techniques will also be of use in superconducting systems.

2.3.1 Transmission line resonators

Superconducting distributed resonators based on a coplanar transmission line have applications in quantum computing as a quantum bus for coupling qubits and for reading out qubits. A coplanar waveguide resonator (CPWR), shown in Fig. 2.4(a), is formed when a coplanar transmission line is cut in two places (additional details in, for instance, [40]). An rf tone coupling into the resonator will form a standing wave for the resonant frequency of the fundamental mode

\[ f_0 = \frac{c}{\sqrt{\epsilon_{\text{eff}}} \cdot 2l}, \]
where \( c/\sqrt{\epsilon_{\text{eff}}} \) is the phase velocity and \( 2l \) is the wavelength. The resonator is distributed due to the series inductance per unit length and a parallel capacitance per unit length. The effective dielectric constant \( \epsilon_{\text{eff}} \), and the characteristic impedance \( Z_0 \) of the transmission line can be determined by conformal mapping techniques [41] using the cross-section in Fig. 2.4(b).

The effective dielectric constant is given by

\[
\epsilon_{\text{eff}} = 1 + q(\epsilon_{\text{sub}} - 1),
\]

where \( \epsilon_{\text{sub}} \) is the dielectric constant of the substrate. The filling factor \( q \) is given by

\[
q = \frac{1}{2} \frac{K(k_1)K(k_0')}{K(k_1')K(k_0)},
\]

where the parameters \( k_i \) are given by

\[
\begin{align*}
k_0 &= \frac{w}{w + 2g}, \\
k_1 &= \frac{\sinh \left( \frac{\pi w}{4h} \right)}{\sinh \left( \frac{\pi(w+2g)}{4h} \right)}, \\
k_i' &= \sqrt{1 - k_i^2},
\end{align*}
\]

and the functions \( K \) are complete elliptic integrals of the first kind. The characteristic impedance of the line is given by

\[
Z_0 = \frac{30\pi}{\sqrt{\epsilon_{\text{eff}}} K(k_0)}.
\]

The high mode density (that is, high electric field per photon) and low intrinsic loss of superconducting CPWRs leads to their use in circuit quantum electrodynamics (cQED), coherently coupling a quantum object to photons in the resonator. In the context of this thesis, the quantum object is a qubit, and the system is described by the Jaynes-Cummings Hamiltonian [42, 43]:

\[
H_{\text{JC}} = H_{\text{resonator}} + H_{\text{qubit}} + H_{\text{interaction}} = \hbar \omega_r \left( a^\dagger a + \frac{1}{2} \right) + \frac{1}{2} \hbar \omega_a \sigma_z + \hbar g (\sigma_+ a + a^\dagger \sigma_-)
\]
2. Background

Fig. 2.4: Coplanar waveguide resonator. (a) A coplanar transmission line is cut in two places forming a resonator with length equal to half the wavelength of the fundamental mode resonance. The transmission line has a series inductance per unit length and parallel capacitance per unit length. (b) A cross-section of the coplanar line with centre width $w$ and gap $g$ used in calculations for $\epsilon_{\text{eff}}$, $Z_0$ and $f_0$. (c) The electric field of a coplanar line has a high mode density between centre conductor and ground planes, making it ideal for qubit coupling.
For the resonator term, $a^\dagger$ and $a$ are the photon creation and annihilation operators, and $\omega_r = 2\pi f_0$ is the resonant frequency of the resonator. For the qubit, $\omega_a$ is the frequency corresponding to transitions between basis states. Finally, in the interaction term, the operators $\sigma_\pm$ swap the qubit state.

When the resonator and qubit transition are on resonance, the eigenstates are

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle|0\rangle \pm |\downarrow\rangle|1\rangle),$$

and energy is exchanged between the qubit excitation and the photon number in the resonator.

Coupling a qubit to a superconducting CPWR has been demonstrated in several experiments and is not limited to superconducting qubits, although this is the natural choice. Successful couplings include a superconducting charge qubit [44, 45, 46, 47], a superconducting flux qubit [27], a gallium arsenide double quantum dot [48, 49], an indium arsenide nanowire double quantum dot [50] and an electron spin ensemble [51].

### 2.3.2 Resonator Decay

Successful operation of superconducting CPWRs as a quantum bus requires the strong coupling regime, where the coupling between the cavity and qubit is large compared to both the qubit decoherence and the rate at which photons decay from the resonator. This photon loss rate is given by

$$\kappa = \frac{\omega_r}{Q},$$

where $Q$ is the Quality ($Q$-) factor of the resonator and is the commonly used metric of resonator performance. When the frequency response of the resonator is Lorentzian, the $Q$-factor can be determined from the full width
at half maximum $df$ as$^3$

$$Q = \frac{f_0}{df}.$$ 

The resonator $Q$–factor is determined by several loss mechanisms, each with a corresponding decay rate. Since decay rates from multiple effects add, the $Q$–factors sum as

$$\frac{1}{Q} = \frac{1}{Q_1} + \frac{1}{Q_2} + \ldots$$

Loss mechanisms are typically classed as either intrinsic to the resonator itself or due to coupling to the environment, characterised by $Q_{int}$ and $Q_{ext}$, respectively.

Losses intrinsic to these devices are small due to the low dissipation of the superconducting state and the temperatures required for qubit operation ensure $T \ll T_c$ for the commonly used superconducting materials. The dielectric loss tangent varies with substrate (see, for instance, the list in [55]), and at low powers is typically dominated by photons in the resonator coupling to quantum mechanical two-level systems in the substrate and substrate / resonator interface. This effect has been studied extensively in superconducting CPWRs [52, 55, 56], with underlying physics described in [57].

A significant contributor to external losses is the coupling capacitors on each end of the resonator, which can be varied according to the requirements of the experiment. Where the qubit or resonator can be tuned, a large coupling is typically chosen for fast readout, since at other times the qubit and resonator can be decoupled. For a fixed frequency, one aims to choose a coupling small enough that the cavity decay rate is less than the decoherence rate, but large enough that state readout is possible [58]. An additional contribution that reduces $Q_{ext}$ is the coupling to the local electromagnetic environment, and this is discussed in Chapter 3.

$^3$ In practice, the loading of the external circuit leads to a non-Lorentzian frequency response, and more complex methods are required to extract the $Q$–factor, for instance expressions in [52, 53] or geometrical techniques in [54]. This loading is detailed in Chapter 3.
2. Background

2.3.3 Josephson Junctions

A Josephson junction [59] is a two-terminal device that is a weak contact between two superconducting leads. This is typically a $\sim 1$ nm insulating layer between the two superconductors, although it can also be a normal metal or a physical deformation that prevents superconductivity at that point. In the context of this thesis, Josephson junctions can be used to form a variety of solid state qubits and are the active component in the Single Flux Quantum (SFQ) circuits discussed in Chapter 4.

In superconductors, Cooper pairs obey Bose statistics and can be described by the wavefunction

$$\psi = |\psi| \exp(i\varphi),$$

where $\varphi$ is the global phase.

Consequently, the phase drop across the junction $\phi$ can be defined as the difference in phase of the superconducting wavefunctions on either side. With $V$ the voltage across the junction and $I$ the rate of Cooper pairs moving across the junction, the governing equations are

$$\frac{d\phi}{dt} = \frac{2e}{\hbar} V \quad \text{and} \quad I = I_c \sin \phi,$$

where $e$ is the electronic charge and $I_c$ is the critical current of the junction.

That a junction can be either superconducting or normal suggests its use in the processing of digital information. That is, zero voltage and finite voltage across the junction can be used to represent binary 0 and 1. Building a circuit around this basic switch has been pursued extensively [60, 61, 62], although the $\sim 1$ ns time taken to reset the switch significantly limits the computing speed.

An alternative method of encoding digital information is based on the property that a superconductor will quantise magnetic flux in a closed loop. This
can be easily seen by considering an open superconducting loop and then closing it. Combining the first equation above with Faraday’s Law, \( V = \frac{d\Phi}{dt} \), and integrating with respect to time, yields

\[
\Phi = \frac{\hbar}{2e}\phi \quad \text{and} \quad \phi = 2\pi \frac{\Phi}{\Phi_0},
\]

where \( \Phi_0 = \hbar/2e \) is the magnetic flux quantum. If the loop is now closed, the superconducting phases must be equal to within a multiple of \( 2\pi \), so the flux also satisfies

\[
\Phi = n\Phi_0,
\]

for some integer \( n \). This flux quantisation can also be used for digital processing, representing 0 and 1 with flux states \( n \) and \( n + 1 \). Switching between these states is fast and low loss, and is discussed in more detail in Chapter 4.

### 2.3.4 Superconducting Qubits

Qubits based on the superconducting Josephson junction are a natural choice for coupling to superconducting CPWRs explored in Chapter 3. In addition, the methods in Chapter 7 are applicable to a variety of solid state qubit architectures.

Josephson junctions are ideal elements for qubits for two reasons. First, they operate at low temperatures and without dissipation so that highly coherent systems are possible. Second, they have a non-linear energy level spacing, so that individual transitions can be selectively addressed. Typical energy level spacings (\( \sim 5 - 10 \text{ GHz} \)) correspond to both reasonable superconducting CPWR lengths (\( \sim 10 - 20 \text{ mm} \)) and a large range of commercially available microwave electronics. The various types of superconducting qubits will be listed here rather than thoroughly described, with various reviews available for more detail [63, 64, 65].
Fig. 2.5: Types of superconducting qubits. (a) A charge qubit encodes the qubit state in the number of Cooper pairs on a superconducting island. It is coupled to a superconducting reservoir via a Josephson junction and the charging energy tuned via the voltage $V_g$. (b) A flux qubit is a superconducting loop, flux-biased with a value of $\Phi_0/2$. The qubit states are the two nearly degenerate values of flux, corresponding to persistent currents in the loop with opposite direction. (c) A phase qubit is similar to a flux qubit, although the bias is such that a washboard potential is formed, and the two lowest energy levels of the higher well are used as the qubit states.

The simplest Josephson junction qubit is an island of superconductor connected via a Josephson junction to a large superconducting reservoir, as shown in Fig. 2.5(a). The number of Cooper pairs on the island is quantised and non-linear since the Josephson junction adds a potential $U = -E_J \cos \phi$. As such, it can be used as a charge qubit with the qubit state encoded in the number of Cooper pairs on the island [66].

The relevant energy scales are the Josephson energy $E_J = \Phi_0 I_c / 2 \pi$ and the charging energy $E_c = e^2 / 2C$, with the most promise for superconducting charge qubits operating in the regime $E_C < E_J$. Here, the qubit is less susceptible to changes in background charge and various qubit structures have been developed to take advantage of this, called variously the ‘transmon’ [67], the ‘quantronium’ [68] and the ‘Xmon’ [69].
An alternative type of superconducting qubit, shown in Fig. 2.5(b), uses a superconducting loop broken by a Josephson junction to take advantage of flux quantisation. The flux through the loop is biased to \((n + 1/2)\Phi_0\) so that the qubit states are the flux states \(n\Phi_0\) and \((n + 1)\Phi_0\). This arrangement is referred to as a flux qubit or persistent current qubit, since the two states correspond to persistent currents in the superconducting loop with opposite sign [70].

The flux qubit forms a double well potential, with each well corresponding to a different current direction. If a similar device is biased away from \((n + 1/2)\Phi_0\), the wells are not equal and the device resembles a washboard potential. This can also be operated as a qubit, known as a phase qubit, using the lowest two energy levels of one of the wells as the qubit states \(|0\rangle\) and \(|1\rangle\). This is shown in Fig. 2.5(c) [71].
3. SUPERCONDUCTING RESONATORS WITH
PARASITIC ELECTROMAGNETIC ENVIRONMENTS

3.1 Abstract

Parasitic electromagnetic fields are shown to strongly suppress the quality ($Q$)-factor of superconducting coplanar waveguide resonators via non-local dissipation in the macroscopic environment. Numerical simulation and low temperature measurements demonstrate how this parasitic loss can be reduced, establishing a Lorentzian lineshape in the resonator frequency response and yielding a loaded $Q$-factor of $2.4 \times 10^5$ for niobium devices on sapphire substrates. In addition, we report the dependence of the $Q$ and resonance frequency shift $\Delta f_0$ with input power and temperature in the limit where losses from two-level systems in the dielectric dominate.

3.2 Introduction

The viability of quantum information hardware based on condensed matter is dependent on isolating and protecting quantum systems from environments that lead to dissipation and uncontrolled evolution [72, 73]. On-chip microwave resonators are key components in quantum technology, enabling readout [44], creating strong interaction between distant qubits [45], and providing a means of transporting a quantum state between different architectures, for instance, in the coherent coupling of superconducting qubits.
3. Superconducting Resonators with Parasitic Electromagnetic Environments

Understanding the mechanisms that lead to dissipation in resonators [74, 75, 76] is thus of central importance in scaling-up quantum information processing and in the construction of supporting quantum technology such as single photon detectors [77] and parametric amplifiers [78].

At low temperatures and frequencies below a few 100 GHz, intrinsic dissipation in superconducting devices is dominated by loss from dielectric materials [55] and radiative processes that depend strongly on the electromagnetic environment [79, 80, 81]. Recent work has also investigated loss from trapped Abrikosov vortices [82] and stray infrared light [83]. In addition to these loss mechanisms, practical resonator circuits are always loaded by lossy external measurement and excitation circuitry. For device applications that require in situ high quality (Q) factor resonators, loading from external circuitry should not lead to a further suppression in the Q-factor from its intrinsic limit.

We investigate how a dissipative parasitic environment loads superconducting coplanar waveguide (CPW) resonators, strongly suppressing the Q-factor and leading to a Fano lineshape in the frequency response of the resonator. We demonstrate via electromagnetic (EM) simulation and low temperature measurements, how this dissipative parasitic coupling can be reduced, restoring a Lorentzian line shape and yielding a Q-factor of $2.4 \times 10^5$ for niobium devices on sapphire substrates. Having suppressed extrinsic losses from this parasitic coupling, we show the dependence of the loaded Q and resonance frequency with input power and temperature. In the low power regime, measurements are consistent with intrinsic dissipation from two-level systems (TLS) associated with defects in the dielectric [55, 84, 56]. These intrinsic losses establish a bound for the extend to which parasitic dissipation must be suppressed.
3. Superconducting Resonators with Parasitic Electromagnetic Environments

(a) Photograph of a niobium CPW resonator on sapphire, inset shows 0.5 fF coupling capacitors. (b) Equivalent (highly simplified) circuit model for the resonator and parasitic path \((C_P)\) with dissipative element \((R_P)\). (c,d) show transmission measurements with -90 dBm power after attenuation at a temperature of 10 mK. In (c), a parallel dissipative path reduces the \(Q\) and results in an asymmetric lineshape. In (d), the parallel path is suppressed resulting in a higher \(Q\) and Lorentzian lineshape.
3. Superconducting Resonators with Parasitic Electromagnetic Environments

3.3 Experimental Design

CPW half-wavelength ($\lambda/2$) resonators are patterned using optical lithography and argon ion-beam milling of 150 nm thick niobium films on r-cut sapphire substrates as shown in Fig. 3.1(a). Substrates are first cleaned using the ion-beam before sputter deposition of Nb at $3 \times 10^{-7}$ mb. Niobium films yield critical current densities of $J_c \sim 14$ MAcm$^{-2}$ and critical temperatures of $T_c \sim 8.3$ K. The geometry of the CPW resonator comprises a 10 $\mu$m wide central track separated from ground on both sides by gaps, 4.6 $\mu$m wide, defining a characteristic impedance $Z_0 = 50$ $\Omega$. The resonator is under-coupled to highly filtered and attenuated input and output ports via two gaps in the central track that define 0.5 fF capacitors [Fig. 3.1(a), inset]. Numerous aluminium wire bonds are used to ground the Nb device to a PCB (Rogers dielectric RO6010) that is soldered into a light-tight copper enclosure and mounted at the mixing chamber stage of a dilution refrigerator (with base temperature 10 mK) after 60 dB of attenuation from room temperature. $Q$-factor measurements are made using a vector network analyser to determine S-parameters after cryogenic [85] and room temperature amplification.

A highly simplified lumped element equivalent circuit of the resonator is shown in Fig. 3.1(b), neglecting time and spatially dependent fields. In addition to the capacitance and inductance per unit length that define the resonator, the equivalent circuit accounts for a parallel circuit path arising from parasitic coupling between the ports and resonator. The addition of this parasitic path $C_P$, which is always present to some degree, leads to a Fano lineshape in the frequency response of the resonator, shown in Fig. 3.1(c) for a $\lambda/2 = 20.8$ mm device at base temperature. In our equivalent circuit, $C_k$ represents the combined parallel capacitance of the coupling ports and parasitic coupling from the central conductor. We extract a $Q$-factor of 13500 from this measurement by fitting [red line in Fig. 3.1(c)] to the functional form of a Fano resonance, $\sigma(\epsilon) = (\epsilon + q)^2/(\epsilon^2 + 1)$, where $\sigma$ is the frequency...
dependent amplitude, $\epsilon$ is the frequency detuning and $q$ is a parameter that characterises the strength of the Fano lineshape.

### 3.4 Asymmetric Lineshape

A Fano lineshape has been observed previously for CPW resonators [86, 52, 54] and can be understood as arising from the coherent interaction between the single frequency mode of the resonator and a continuum of accessible modes from a parallel circuit path that has a flat frequency response [87]. Importantly, we note that the presence of this parasitic parallel path, and its Fano lineshape signature, do not constitute a loss mechanism unless the parallel path is dissipative. We find however, that for practical circuit implementations, this parasitic path involves resistive losses from normal metals in the PCB and dissipation in dielectric materials well beyond the neighbourhood of the resonator. Below we show how this parasitic dissipation leads to a strong reduction in the loaded $Q$-factor.

We first compare the frequency response of two very similar resonators, each bonded to different PCBs and sample enclosures. The results from these two devices are indicative of several other devices we have measured, including resonators fabricated on different substrates\(^1\). Comparing the lineshape, Fig. 3.1(c) and (d) show that the frequency response is very sensitive to the details of the PCB and sample enclosure interconnects. For the Fano lineshape shown in Fig. 3.1(c), measurements were made in the sample mount shown in Fig. 3.2(a), while the Lorentzian lineshape data in Fig. 3.1(d) was taken using a sample mount designed to suppress parasitic coupling, shown in Fig. 3.2(b). In addition to the variation in lineshape between the two data sets, we note the 20-fold difference in $Q$-factor that results from different parasitic

---

\(^1\) In resonators on magnesium oxide substrates, suppression of the parasitic environment recovered a symmetric lineshape with $Q$-factor of $5.5 \times 10^4$, lower than sapphire due to the more lossy dielectric.
3. Superconducting Resonators with Parasitic Electromagnetic Environments

Fig. 3.2: (a), (b) Two different sample setups with different parasitic coupling used in measurements shown in Fig. 3.1(c) and 3.1(d), respectively. (c) Electric field $E$ in the PCB dielectric for the slices indicated by the axes in (a), (b) and (d) current density $J$ in the lower PCB ground plane along the same slice. (e) Comparison of simulated (dashed lines) and measured (solid lines) transmission response for setups in (a) and (b). Data taken at room temperature in the absence of a superconducting chip.
dissipation in the two sample mounts. Fits to the data [shown in red in Fig. 3.1(c) and (d)] yield an estimate of the strength of the parasitic coupling characterised by the Fano parameter $q$. For low parasitic coupling ($q \sim 1000$) the lineshape becomes Lorentzian, characteristic of a driven oscillator.

To better understand this parasitic dissipation, we simulate the environment of the resonator, PCB, and sample enclosure using a finite element 3D EM field solver [88]. Taking a horizontal slice across the device, we compare the simulated electric field and current densities present for the two different setups shown in Fig. 3.2(a) [red] and (b) [black]. The magnitude of the E-field slice [Fig. 3.2(c)] is taken in the dielectric, 250 $\mu$m below the copper surface of the PCB and the current density is calculated at the lower PCB ground plane [Fig. 3.2(d)]. As is evident in the simulation, the parasitic field density has been significantly lowered for the sample mount arrangement shown in Fig. 3.2(b). To confirm that these EM simulations accurately capture the response of our system, we also compare the simulated response of the enclosure and PCB, (with no resonator chip) to transmission data measured for both configurations as shown in Fig. 3.2(e). We note that for the sample mount shown in Fig. 3.2(b) we have suppressed the parallel coupling between ports by $\sim 60$ dB [see Fig. 3.2(e)].

The large electromagnetic cross-section of superconducting resonators makes them susceptible to parasitic effects on macroscopic scales. To visualise the extent of non-local EM fields, Fig. 3.3(b-d) shows further results of our EM simulations for the low parasitic setup shown schematically in Fig. 3.2(b) and as a photograph in Fig. 3.3(a). Even for this optimised setup, we find significant electric field and current density “hot-spots” far from the resonator. These are largely associated with regions in which the geometry of the planar transmission line varies, despite a constant impedance, such as at the bondwire interface between the PCB and superconducting chip [see Fig. 3.3(d)]. We believe it is these regions of appreciable current and E-field density that constitute the parasitic environment of the resonator, producing
Fig. 3.3: (a) The chip with CPW is wire bonded to the PCB that incorporates a coplanar line and via fencing between sample and SMA connectors. Simulations of the parasitic EM signal using Ansoft’s HFSS at -90 dBm input powers for a frequency near $f_0$. (b) E-field magnitude in the PCB dielectric, (c) current density in the PCB ground plane, and (d) the current density in the chip, top PCB ground planes, and wire bonds.
dissipation in the PCB normal metal, dielectric, and bondwire interconnects.

In moving from the sample mount design shown in Fig. 3.2(a) to the low dissipation mount shown in Fig. 3.2(b) we have reduced parasitic coupling by adding numerous tightly spaced vias to ground that strongly confine the E-field and current [89]. In addition, the design does not taper the transmission line [90], employing microwave launchers and CPW track widths of the smallest possible size. A further reduction in parasitic coupling was evident for well-matched transmission line geometry at the PCB-chip interface. Numerous wirebonds are used to reduce on-chip ground current density [83].

3.5 Two Level Systems

Finally, having alleviated parasitic coupling as the dominant source of dissipation, we report the dependence of the loaded $Q$ on input power and temperature. It is now well established that defects or TLSs in the resonator dielectric lead to loss by absorbing microwave photons at low power and temperature [55, 84, 56]. For the low dissipation setup, we observe an improvement in $Q$ with increasing input power [Fig. 3.4(a)] consistent with TLSs being driven into long-lived excited states that cannot absorb further microwave energy. At still higher powers a strong reduction in $Q$ and a distortion in the lineshape is evident [see Fig. 3.4(b)], consistent with a non-linear surface impedance from large-angle grain boundaries [91, 92]. Further evidence that the loss is now dominated by TLSs is seen in the non-monotonic temperature dependence of the loaded $Q$ [Fig. 3.4(c)] and fractional frequency shift $\Delta f/f_0$ [Fig. 3.4(d)]. The component of the loss due to TLSs in this sample can likely be decreased by using ultra-high purity sapphire to remove paramagnetic impurities [93, 74]. At temperatures below 100 mK the loaded $Q$ and $\Delta f/f_0$ exhibit a slight inflection that is presumably derived from the thermal population difference between the excited and ground states of the TLSs [94].
Fig. 3.4: (a) Power dependence of the loaded $Q$ for the two setups with different parasitic dissipation where black points are from the setup in Fig. 3.2(a) and red points from setup in Fig. 3.2(b). Data taken at 10 mK. (b) Nonlinear response of the resonator at high input power. (c) shows the temperature dependence of the loaded $Q$ and (d) the resonance frequency fractional shift $\Delta f/f_0$ attributed to the presence of TLSs in dielectric environment of the resonator.
The presence of macroscale parasitic channels, as investigated here, lead to an unwanted coupling between the resonator and its dissipative environment. For complex device architectures that require many ports and planar microwave feed-lines, parasitic modes that inadvertently couple energy far from the resonator present a key technical challenge for low loss quantum circuits. We anticipate that the mitigation of parasitic dissipation using the methods reported here will be of interest for the design of scaled-up quantum hardware.

We acknowledge C. J. Lewis and J. I. Colless for technical assistance and thank J. Martinis, A. Houck, J. Aumentado and T. Duty for useful discussions. This research was supported by the IARPA/MQCO program through the U. S. Army Research Office and the Australian Research Council Centre of Excellence Scheme (EQuS CE110001013).
4. SIMULATION SOFTWARE FOR SINGLE FLUX QUANTUM CIRCUITS

4.1 Abstract

Superconducting electronics based on the quantisation of magnetic flux are well-suited for control hardware in quantum systems due to their fast, low-dissipation cryogenic operation. Software for the simulation and development of Single Flux Quantum (SFQ) circuits is presented based on lumped element nodal analysis. While software in this family has been in use for decades, custom implementations offer advantages over more general programs due to the specific requirements of superconducting electronics. The software described here has several features well-suited to SFQ circuits, including a re-formulation of the circuit equations so that a matrix inversion at each time step is avoided. The improvement in asymptotic complexity and parallel computation makes feasible the optimisation of SFQ circuits with thousands of components.

4.2 Introduction

The scaling of qubit systems to a size sufficient to implement quantum algorithms with error correction introduces a range of problems for the associated classical electronics, such as heat dissipation, bandwidth, noise and wiring. Novel hardware for readout and control will need to be developed in a scal-
able manner, and on top of this, algorithm implementation requires high-level logic to steer pulses with low latency - less than the qubit lifetime.

Superconducting electronics with Josephson junctions as the active component is a technology with attributes well-suited to quantum control problems. Their operation is based on short voltage pulses associated with a change in flux through a superconducting loop, and they have low power dissipation, low latency logic and lossless superconducting interconnects [95, 62, 96]. The low power dissipation and low temperature operation allows for their use in the same cryogenic environment as the qubit system so that associated latency is small. These systems can provide cryogenic operation in the 100s of GHz [97], can interface with electric signals used in qubit control [98], and can be fabricated together with superconducting qubits [99].

Quantum computation aside, these circuits are being developed at a time where conventional CMOS-based semiconductor computing is becoming limited by heat dissipation, with processor improvements in recent years being driven more by architecture than by an increased clock rate. In addition, rising energy demands of large scale computation\(^1\) are motivating research into alternative computing schemes.

The circuits described in this chapter exploit the quantisation of magnetic flux, where flux in a closed superconducting loop is a multiple of the flux quantum \(\Phi_0 = \frac{2e}{h}\). Single flux quanta (SFQ) provide a natural encoding of digital information - flux states \(\Phi = n\Phi_0\) and \(\Phi = (n + 1)\Phi_0\) can be used to represent logical 0 and 1. The Josephson equation is

\[
\frac{d\phi}{dt} = \frac{2e}{h} V,
\]

where \(\phi\) is the difference in phase between superconducting states on either side of the junction and \(V\) the voltage across the junction. From this, we can see that a transition from one flux state to another produces a transient

---

\(^{1}\) While this is a record holder rather than an indicative number, the Tianhe-2 supercomputer in Guangzhou draws 18 MW with another 6 MW in cooling.
voltage pulse with integrated amplitude
\[ \int_0^t V(t)dt = \Phi_0. \]

This pulse can be passed along matched, lossless transmission lines to other components of the circuit, where further changes in flux can be triggered. The SFQ exists as either voltage pulses or persistent currents in a superconducting loop with magnitude given by \( \Phi_0 = 2 \text{ mV ps} = 2 \text{ mA pH} \).

Fig. 4.1(a,b) show examples of SFQ circuits, in this case simple circuits in the Rapid Single Flux Quantum (RSFQ) family [95]. Fig. 4.1(a) generates an SFQ pulse at the output on a rising edge at the input. The circuit is based around the superconducting loop indicated in red. When the input current rises such that the current through the junction labelled in red is above its critical value, the junction switches and produces an SFQ pulse that propagates to the circuit output. The second part of the circuit is a Josephson Transmission Line (JTL) segment, which propagates SFQ pulses and serves as a noise discriminator. The profile of a typical SFQ pulse produced in this way is shown in Fig. 4.1(d).

A stand-alone JTL is shown in Fig. 4.1(b), where the junctions are current-biased below their critical value. A voltage pulse arriving at the input switches the first junction and produces an SFQ pulse, which in turn switches the second junction producing an SFQ pulse at the output. Fig. 4.1(c) shows a simple compound circuit created with the DC-to-SFQ converter in (a) and two JTLs in (b), and simulation results are shown in (e,f,g). On a rising edge from the voltage input (Fig. 4.1(e)), a voltage pulse propagates through the circuit (Fig. 4.1(f)). Fig. 4.1(g) shows the phase across the last Josephson junction in the circuit, demonstrating the \( 2\pi \) phase change associated with each SFQ pulse. The simulations were generated using in-house software described in this chapter.

This RSFQ family of circuits has had considerable interest and success in both the range and clock speed of devices demonstrated. These include, but
Fig. 4.1: (a, b) Two basic RSFQ circuits. The DC-to-SFQ converter (a) produces an SFQ pulse at the output on a rising edge at the input. The Josephson Transmission Line (b) propagates SFQ pulses. (c) A simple compound circuit used in this chapter as an example and for testing software. (d) Voltage profile of a typical SFQ pulse. (e, f, g) Simulation results using in-house software of the circuit in (c). An SFQ pulse is produced at the output (f) coinciding with a rising edge at the input (e), and corresponding to a $2\pi$ flip in the last junction in the circuit (g). (h) Typical parameter sweeping showing regions where rising edges produce SFQ pulses as intended. (i, j, k) Confirming correctness of software by generating equivalent plots using WRspice [100].
are far from limited to, logic gates [95, 101], a 15-bit analog-to-digital con-
verter [98], a 750 GHz flip-flop [97], 60 GHz multiplexing and demultiplexing
[102], and a 20 GHz, 8-bit processor [103].

One particular drawback of the RSFQ design is the static power dissipation
of most implementations. To bias large numbers of junctions with DC cur-
crents, a resistor network is typically employed which dissipates more heat
than the circuit operation. Some potential ways to overcome this include
using inductor and junction arrays to distribute bias currents [104], or an
alternative approach Reciprocal Quantum Logic [105], where junctions are
inductively coupled to an AC rail, and bits are encoded as pairs of SFQ pulses
with opposite polarity.

In all SFQ approaches, the circuit behaviour is simulated in the time do-
main using nodal analysis on lumped elements prior to device fabrication.
In addition to confirming the intended behaviour of an SFQ circuit, further
simulations can be used to optimise or evaluate circuit designs. As an exam-
ple, a measure of the robustness of a design can be determined by simulating
critical margins (by how much a parameter needs to move before the circuit
fails) or yield (Monte Carlo simulation of fabrication variations to determine
what proportion of devices will perform as intended). Various high-level
methods have been proposed to efficiently navigate the parameter space and
increase margins and yield in SFQ circuits such as inscribed hyperspheres
[106], genetic algorithms [107] and particle swarm optimisation [108].

A simple example for the circuit described in Fig. 4.1(c) involves sweeping the
parameters of the JTL, which is shown in Fig. 4.1(h). Here we are sweeping
the area of the Josephson junction\(^2\) and the bias current, with circuits with
the appropriate behaviour (6\(\pi\) phase change after 3 rising edges) indicated
in blue.

\(^2\) The junction area affects the capacitance, resistance and critical current of the junc-
tion.
This chapter describes a software package implementing fast nodal analysis with Josephson junctions, with features suited to SFQ circuit design. There are several programs in existence used for such circuit simulation from the SPICE family of solvers\textsuperscript{3}, however there are shortcomings. Programs not implementing an intrinsic Josephson junction may have functionality for user-defined circuit elements, however this is often prohibitively slow. Many, typically older, solvers do not provide parallel processing - a significant hindrance considering both the circuit evaluation methods described above and the ubiquitous support for multiple threads on modern processors, not to mention the increasing availability of cluster computing. Discontinued solvers are available online, although with scant support and documentation. Finally, a designer’s high level tools for circuit evaluation and optimisation are limited by what is implemented in that particular program, which is often unnecessarily general in order to handle a wide range of circuit problems.

We present a software package that seeks to address some of these concerns and provide custom functionality for SFQ circuit design. We will give a brief introduction to nodal analysis, provide circuit verification using existing commercial software, and then describe advantages of this package.

\section*{4.3 Nodal Analysis}

Nodal analysis is a method of calculating the potential difference between various nodes in an electrical circuit using Kirchhoff’s current law. It is the basis of the family of SPICE solvers, and is used in the simulation of SFQ-based superconducting circuits.

Before considering SFQ circuits in particular, we will illustrate nodal analysis through the trivial circuit in Fig 4.2, where the aim is to find the voltage

\textsuperscript{3}SPICE is an electronic simulator originally developed at UC Berkeley and now with dozens of implementations. Publicly available solvers suitable for superconducting circuits include JSIM, WRspice, Xyce, PSCAN, with a (now slightly outdated) review [109]
at each node in the circuit in a manner scalable to an arbitrary network of
components. We write Kirchhoff’s Current Law at each node in the circuit:

\[
\begin{align*}
\text{Node 1:} & \quad -I_s + \frac{V_1 - V_2}{R_1} = 0 \\
\text{Node 2:} & \quad \frac{V_2 - V_1}{R_1} + \frac{V_2}{R_2} = 0
\end{align*}
\]

This is in the general form

\[ G.V = I, \]

with the conductance matrix \( G \), and current \( I \), given by

\[
G = \begin{pmatrix}
\frac{1}{R_1} & -\frac{1}{R_1} \\
-\frac{1}{R_1} & \frac{1}{R_1 + 1/R_2}
\end{pmatrix}
\quad \text{and} \quad I = \begin{pmatrix} I_s \\ 0 \end{pmatrix}.
\]

We can then write down the solution via a matrix inversion as

\[ V = G^{-1}I. \]

Using this strategy, we can solve an arbitrary circuit provided we can write the current contribution to node \( n \) from a particular element in the form \( I_n = f(V_i) \) for \( f \) a linear function. We will demonstrate this procedure for capacitors, voltages, inductors and Josephson junctions, at which point we have the elements required for simulation of SFQ circuits. Where solutions vary with time due to time-varying sources or components such as capacitors and inductors, the process described is repeated after a time interval \( \delta t \).

**Fig. 4.2:** Trivial circuit for illustration of nodal analysis. The aim is to provide a solution in a manner scalable to an arbitrary collection of nodes and components.
Capacitors

The governing equation of a capacitor is

\[ I = C \frac{dV}{dt}, \]

for \( C \) a constant property of the component. We can approximate the slope \( dV/dt \) using the last time step\(^4\):

\[ I = \frac{C}{\delta t} V(t) - \frac{C}{\delta t} V(t - \delta t). \]

This is now linear in \( V_t \), and represents one equation in the system \( G.V = I \) (and is equivalent to modelling the capacitor as a resistor with \( R = \delta t/C \) and a parallel current source with value \( -CV(t - \delta t)/\delta t \)).

Inductors

The governing equation of an inductor is

\[ V = L \frac{dI}{dt} \]

Isolating the current contribution using \( x(t) = x(t - \delta t) + \dot{x} \delta t \) gives

\[ I(t) = I(t - \delta t) + \frac{V \delta t}{L}. \]

Since recording currents in the software is not as straightforward as voltages, it can be easier to use

\[ I(t - \delta t) = \int_0^{t-\delta t} \frac{V}{C} dt, \]

where this accumulated current is stored for each inductor and incremented by \( V \delta t/C \) after every \( \delta t \).

Voltage Sources

\(^4\)While this is accurate for sufficiently small \( \delta t \), a more precise form is used in the software, and detailed in Appendix C along with similar expressions for inductors and Josephson junctions. The increase in computation associated with these corrections is far outweighed by the speed-up afforded by a longer \( \delta t \).
A voltage source can be implemented in various ways, the simplest being to consider a current source in parallel with a very small resistor. If the resistance is negligible compared to the circuit load, most of the current flows through the resistor to the other node of the current source, resulting in a potential between the nodes equal to a constant $V = I_s R$.

**Josephson Junction**

In this implementation, we use the resistor and capacitively shunted junction (RCSJ) model for a Josephson junction. That is, a perfect junction element in parallel with a resistor and capacitor. Since resistors and capacitors have been addressed above\(^5\), we now consider only the Josephson junction. The governing equations of this element are

$$I_s = I_c \sin \phi \quad \text{and} \quad \frac{d\phi}{dt} = \frac{2e}{\hbar} V$$

for $I_s$ the superconducting current flowing through a junction, $V$ the potential difference across the junction and $\phi$ the phase difference between the superconducting wave functions on either side of the tunnel barrier. $I_c$ is the (constant) junction critical current.

Considering these equations, it becomes apparent that we need to consider $\phi$ as a variable, so that one of the $V_i$ in the expression $G.V = I$ is the phase across the junction. The current contribution $I = f(V_i)$ of the junction element is

$$I = I_c \sin \phi.$$  

Since we have increased the dimensionality of the system, equivalent to adding an additional node to the circuit, we need another equation to prevent the system becoming under-determined. Kirchoff’s Current Law does

\(^5\) There are two points to note with the resistor. First, junctions are often shunted with a parallel thin film resistor to decrease the junction resistance. Second, the quasiparticle current varies depending on the voltage across the junction with respect to the gap voltage $V_g = 2\Delta/e$, and this is described in Appendix C
not apply to our additional node, so we use the second expression in 4.1:

\[
\phi(t) = \phi(t - \delta t) + \delta t \times \frac{d\phi}{dt} = \phi(t - \delta t) + \frac{2e\delta t}{\hbar} V
\]

The correctness of our implementation of these components is verified by comparison to exact analytic expressions and existing commercial software. Fig. 4.1(i, j, k) show simulation results from the same circuit using the program WRSpice from Whiteley Research [100]. Comparison to Fig. 4.1(f, g, h) suggests the solver is correct for this range of circuit elements.

### 4.4 Once-Only Matrix Inversion

The computationally expensive step in nodal analysis is the matrix inversion from \( G.V = I \) to \( V = G^{-1}I \). A standard way to perform this computation is LU decomposition, where the matrix \( G \) is written as the product \( G = LU \) where \( L \) is a lower triangular matrix and \( U \) an upper triangular matrix. Since the matrices are triangular, solving the systems \( L.y = I \) and \( U.V = y \) is straightforward. A significant speed-up for SPICE solvers (and other areas where linear systems dominate) is that the matrix \( G \) is sparse, especially for large circuits. This allows for various methods to speed up the decomposition such as Gilbert-Peierls’ algorithm [110] and KLU [111].

In our software, we use a different approach based on the fact that for each circuit, SFQ design involves a large number of solves of the system \( G.V = I \). A given simulation (that is, a given circuit with a particular set of parameters) involves a large number of time steps \( \sim 10^3 - 10^5 \). Consequently, one-time simplifications of the system \( G.V = I \) become advantageous regardless of their computational cost.

We formulate the equations \( G.V = I \) so that the matrix \( G \) is constant throughout a simulation. In this case, a single matrix inversion is performed
once only, and during subsequent time steps, we evaluate only $V = G^{-1}I$, asymptotically faster than matrix inversion. The process involves the following steps, which we will state and then describe:

- Form equations $G.V = I$ so that currents, voltages and junction phases are in the variables $V$.
- Perform variable substitution to reduce the dimensionality
- Invert the matrix $G$ once only, and at each time step compute $V = G^{-1}I$.

We first prepare the governing equations differently to standard nodal analysis. In the set of equations $G.V = I$, we include the currents through the components as variables in the vector $V^6$. The increased dimensionality can be later reduced using variable substitution, but this allows greater flexibility over which variables are removed.

The increase in dimensionality has two adverse affects on the subsequent computation. First, the initial matrix inversion is more time consuming, however this is not significant since it occurs once only. More detrimental is the increased time to compute $V = G^{-1}I$ at each time step. To overcome this, the system is reduced in size using standard variable substitution. That is, the software selects a row $i$,

$$\sum_{j=1}^{n} G_{ij} V = I_i,$$

and a term $j$, so that we can write

$$V_j = \frac{1}{G_{ij}} \left( I_i - \sum G_i V \right).$$

---

6 Previously the $V$ were node voltages and also Josephson junction phases, and now it includes component currents as well. We will keep the notation $G$, $V$ and $I$, since the subsequent techniques are similar to standard nodal analysis.
This expression is substituted into the other equations of $G.V = I$, resulting in an updating of the components $G, I$, and a reduction in the dimensionality of the system. The row $i$ and term $j$ are chosen at each iteration via a heuristic that reduces the complexity of the coefficients in the solution and maintains a constant $G$.

This process is computationally expensive, however it is undertaken once only for each circuit to be simulated. In addition, many computationally intensive SFQ simulations involve sweeping parameters on the same circuit with a view to optimising the parameters involved. In this instance, the form of the system $G.V = I$ is identical, even if $G$ varies. In these simulations, the variable substitution is performed once only, and the calculation of $G^{-1}$ is performed once for each parameter set\(^7\).

The computation at each time step is then two-fold. First, the multiplication $V = G^{-1}I$ is carried out. Second, the equations used for substitution are invoked to determine the value of the variables that were eliminated at the start of the simulation. The multiplication $V = G^{-1}I$ is, at worst, $O(n^2)$.

Performance of this procedure is shown in Fig. 4.3. In this case, we generate circuits of arbitrary size using the simple components in Fig. 4.1, using a single DC-to-SFQ converter followed by a train of JTLs. Fig. 4.3(a) shows the time taken\(^8\) per time step as a function of the number of nodes in the circuit, fitted to a curve of the form $\alpha n^2$. The agreement shows the performance is as expected and dominated by a multiplication rather than an inversion. The matrix $G$, despite the substitutions, is still mainly zeroes and further speedup is available using algorithms for sparse matrix operations, although this has not yet been implemented.

\(^7\) The varying resistance of a Josephson junction and variable time steps result in the occasional need to perform additional matrix inversions. Details are given in Appendix C

\(^8\) The hardware for results shown here is two Intel Xeon X5550, 2.67 GHz, 8 MB cache, 4 cores / 8 threads per processor. 24 threads were tested using two Intel Xeon X5690, 3.47 GHz.
4. Simulation Software for Single Flux Quantum Circuits

4.5 Other Features

The program will distribute work among multiple processor cores. Speedup is available in this way for the multiplication $V = G^{-1}I$, however multi-threading is more suited to the parameter sweeping that is done in SFQ circuit design during optimisation or testing circuit yield. There is no upper bound on the number of threads that can be used, however it has not been tested past 24 threads on 12 physical cores\(^8\).

In most SFQ optimisation, a simulation with a particular parameter set is deemed to succeed or fail only, since the change in flux through a superconducting loop is intrinsically discrete. However during design, a simulation can provide more detailed information. For instance, how far along a circuit an SFQ pulse propagates, or the width of a voltage pulse produced at a particular point.

Fig. 4.3(b) shows a typical output from the program, in this case a two-dimensional parameter sweep similar to Fig. 4.1(h). Parameter values for which a rising edge fails to produce an SFQ pulse are indicated in red. Where a pulse is produced, the voltage amplitude is indicated in blue-green, with greater amplitudes - and sharper pulses since $\int V\,dt = \Phi_0$ - away from the critical parameters.

Multi-dimensional parameter sweeps such as this one are not iterated over sequentially. Instead, parameter values distributed across the swept space are chosen, with subsequent choices filling in the gaps left by earlier choices so that a rough picture of the parameter space is established quickly. Fig. 4.3(b) was produced on 16 concurrent threads after three iterations.
Fig. 4.3: Results from SFQ simulation software. (a) The computation time scales as $\alpha n^2$ for $n$ the circuit size. This is a result of the once-only matrix inversion with the multiplication $V = G^{-1}I$ at each time step. (b) A typical output of the software, taken after 3 iterations on 16 threads, similar to the parameter sweeps in 4.1(h,k). Where the circuit is successful, sharper SFQ pulses are represented as blue, and broader pulses in green. Circuit failures are indicated in red.
4. Simulation Software for Single Flux Quantum Circuits

4.6 Discussion

Circuits based on SFQ logic are necessarily numerically simulated before fabrication and testing. Fast simulation that scales well with circuit size enables the design of complex circuits as well as a range of high level optimisation techniques that would otherwise be computationally prohibitive. Since custom software is not constrained by unnecessary generality, it’s possible to achieve advantages over their commercial cousins.

The circuits in this family are well suited to the engineering challenges associated with scaling systems of solid state qubits. The hardware required to interface with a quantum circuit includes fast electric pulses, slow voltage bias and fast readout. This hardware needs to be faster than qubit decoherence, low noise and scalable to large numbers of qubits. Components located within a dilution fridge have the advantage of low latency and reduced thermal links to room temperature, but also require very low heat dissipation.

SFQ circuits operate at low temperatures and dissipate very little heat. Their fast operation and proximity to a qubit circuit implies that classical computation could be carried out on a timescale shorter than the qubit lifetime as part of an adaptive control scheme or other quantum algorithm. Computation aside, superconducting circuits can be used as fast, cryogenic data converters, providing an interface between a quantum system and classical semiconductor electronics. These architecture ideas will be explored in more detail in Chapter 7.
5. DISPERSIVE READOUT OF A FEW-ELECTRON DOUBLE QUANTUM DOT WITH FAST RF GATE-SENSORS

5.1 Abstract

We report the dispersive charge-state readout of a double quantum dot in the few-electron regime using the \textit{in situ} gate electrodes as sensitive detectors. We benchmark this gate-sensing technique against the well established quantum point contact (QPC) charge detector and find comparable performance with a bandwidth of \(\sim 10\) MHz and an equivalent charge sensitivity of \(\sim 6.3 \times 10^{-3} \text{e/}\sqrt{\text{Hz}}\). Dispersive gate-sensing alleviates the burden of separate charge detectors for quantum dot systems and promises to enable readout of qubits in scaled-up arrays.

5.2 Introduction

Non-invasive charge detection has emerged as an important tool for uncovering new physics in nanoscale devices at the single-electron level and allows readout of spin qubits in a variety of material systems [112, 113, 114, 29, 30, 32, 115, 116, 117]. For quantum dots defined electrostatically by the selective depletion of a two dimensional electron gas (2DEG), the conductance of a proximal quantum point contact (QPC) [29, 30, 32, 115, 117] or single electron transistor (SET) [114, 116] can be used to detect the charge config-
uration in a regime where direct transport is not possible. This method can, in principle, reach quantum mechanical limits for sensitivity [118] and has enabled the detection of single electron spin-states [29, 115, 119] with a 98% readout fidelity in a single-shot [120].

An alternate approach to charge-state detection, long used in the context of single electron spectroscopy [121], is based on the dispersive signal from shifts in the quantum capacitance [122, 123] when electrons undergo tunnelling. Similar dispersive interactions are now the basis for readout in a variety of quantum systems including atoms in an optical resonator [124], superconducting qubits [125, 126, 127] and nanomechanical devices [128].

In this Letter we report dispersive readout of quantum dot devices using the standard, in situ gate electrodes coupled to lumped-element resonators as high-bandwidth, sensitive charge-transition sensors. We demonstrate the sensitivity of this gate-sensor in the few-electron regime, where these devices are commonly operated as charge or spin qubits [129] and benchmark its performance against the well established QPC charge sensor. We find that because the quantum capacitance is sufficiently large in these devices, gate-sensors have similar sensitivity to QPC sensors. In addition, we show that gate-sensors can operate at elevated temperatures in comparison to QPCs.

Previous investigations, in the context of circuit quantum electrodynamics (c-QED), have engineered a dispersive interaction between many-electron dots and superconducting coplanar waveguide resonators [48, 49, 130, 50, 131]. Recently, the charge and spin configuration of double quantum dots has also been detected by dispersive changes in a radio frequency resonator coupled directly to the source or drain contacts of the device [132, 133, 50, 134]. The present work advances these previous studies by demonstrating that the gates, already in place to define the quantum dot system, can also act as fast and sensitive readout detectors in the single-electron regime. This is a surprising result, given the small capacitive coupling between the gate and dot, but lifts a barrier to qubit readout in large scaled-up quantum dot arrays.
Fig. 5.1: (a) Micrograph of a similar device to the one measured and circuit schematic. One of the *in situ* dot-defining gates (blue) is coupled via a bondwire to an off-chip Nb/Al₂O₃ superconducting lumped-element resonator to enable dispersive readout. (b) Amplitude $S_{11}$ (blue) and phase response (red) of the resonator. (c) Illustration of the charging energy spectrum for a quantum dot. The resonant rf gate voltage $V_{rf}$ induces tunnelling at the charge degeneracy point (green oscillation) leading to a dispersive shift that is suppressed for configurations of stable charge (orange oscillation).
by alleviating the need for many ohmic contacts, large on-chip distributed resonators, or proximal charge detectors.

5.3 Experimental Design

Our gate-sensor, shown in Fig. 5.1(a), comprises an off-chip superconducting Nb on Al$_2$O$_3$ spiral inductor ($L \sim 210$ nH) in resonance with the distributed parasitic capacitance ($C_p \sim 0.23$ pF) that includes a TiAu gate electrode used to define the quantum dots (resonance frequency $f_0 = 1/2\pi \sqrt{LC_p} = 704$ MHz, $Q$-factor $\sim 70$). As the sensitivity of the resonator is improved by minimizing this parasitic capacitance, we deep etch the sapphire substrate between windings of the Nb inductor (lowering the dielectric constant) and make use of short bondwires between the inductor and GaAs chip. The dots are 110 nm below the surface of a GaAs/Al$_{0.3}$Ga$_{0.7}$As heterostructure (electron density $2.4 \times 10^{15}$ m$^{-2}$, mobility 44 m$^2$/V s at 20 Kelvin) that is mounted on a high-frequency circuit board at the mixing chamber of a dilution refrigerator with base temperature $T \sim 20$ mK. The electron temperature $T_e$, determined by Coulomb blockade (CB) thermometry, is below 100 mK. The amplitude and phase response of the resonator is measured, following cryogenic amplification, using a vector network analyzer, as shown in Fig. 5.1(b).

Dispersive gate-sensors (DGS) detect charge-transitions (rather than absolute charge) by sensing small changes in the polarizability or quantum admittance when an electron tunnels in response to the alternating rf gate voltage. Tunnelling modifies the resonator capacitance beyond the geometric contribution (at the position of green symbol in Fig. 5.1(c)) compared to the regime where tunnelling is suppressed (orange symbol in Fig. 5.1(c)). Changes in the quantum capacitance alter the resonator frequency, which in

---

1 Using EM simulation software (Ansoft HFSS), we determine the inductor contributes 0.14 pF to the total parasitic capacitance $C_p \approx 0.23$ pF
Fig. 5.2: (a) Dispersive signal from the gate-sensor showing transitions in electron number for a large single quantum dot. Green and orange symbols correspond to positions of symbols in Fig. 1(c). (b) Derivative of the QPC conductance signal with gate voltage $V_g L$ in a region of gate-space similar to (a). The slight shift in gate voltage and period of the oscillations in comparison to (a) is due to the presence of the QPC gate bias. (c) Phase response of the gate-sensor showing peaks corresponding to single electron transitions. (d) Vertical slice through the conductance signal in (b), at $V_g R = -723$ mV. (e) SNR of the gate-sensor as a function of the modulation frequency of a signal applied to a nearby gate. (f) SNR for the gate-sensor as a function of carrier frequency. (g) Width and height of the DGS response signal with power applied to the resonator (before $\sim$ 44 dB of attenuation). (h) Coulomb charging diamonds for the quantum dot, measured using the gate-sensor in a regime where direct transport is not possible. Colour scale is the derivative of the dispersive signal. Labels indicate number of electrons.
5. Dispersive Readout of a Few-Electron Double Quantum Dot with Fast rf Gate-Sensors

This response of the resonator $\Delta \phi$ is detected by fast sampling of the in-phase and quadrature components of the reflected rf to produce a baseband signal, $V_{DGS}$, proportional to the dispersive shift [135].

5.4 Characterisation

Our device integrates a QPC charge sensor together with the DGS and allows simultaneous readout of the quantum dot system using both detectors. A comparison of the relative sensitivity of the QPC and DGS is shown in Fig. 5.2(a-d) where the response of each detector is measured as a function of the gate voltages $V_gL$ and $V_gR$ used to define a large, single quantum dot in the Coulomb blockade regime. The dispersive signal $V_{DGS}$ from the gate resonator is shown in Fig. 5.2(a,c), with Fig. 5.2(b,d) showing the derivative of the conductance $G$ of the QPC with respect to $V_gL$.

The sensitivity of both sensors is quantified by applying a small modulation voltage to a nearby gate, inducing periodic variation in conductance of the QPC or dispersive response of the DGS. We calibrate the detector signal $dG$ or $dV_{DGS}$ due to this modulation by comparing its amplitude to the signal response from a single electron transition. A measurement of the signal-to-noise ratio (SNR) in a given bandwidth yields the detector sensitivity. For the QPC we measure a typical charge sensitivity at 36 Hz of $\sim 3 \times 10^{-3}$ e/$\sqrt{\text{Hz}}$, corresponding to an integration time $\tau_{int}$ of 9 $\mu$s required to resolve a change of a single electron charge on the dot. The DGS method yields a $\tau_{int} = 39$ $\mu$s to resolve a single electron transition (equivalent to $6.3 \sim \times 10^{-3}$ e$/\sqrt{\text{Hz}}$). The sensitivity of the DGS compares favourably to ref. [132], where a $\tau_{int}$ of 4 ms is required to resolve a single electron charge using a resonator connected to a lead via an ohmic contact. In comparison to the

\[2\] A modulation in the dispersive response by a nearby gate can be understood as the variation in capacitance with depletion of electrons surrounding the DGS.
Fig. 5.3: (a) Dispersive response $V_{DGS}$ from the gate-sensor for a few-electron double quantum dot. Labels indicate number of electrons in the left and right dot. (b) Equivalent charge sensing signal from the QPC detector confirming the few-electron regime. The shift in location of the transitions compared to the data in (a) is due to the required QPC gate bias. (c) Temperature dependence of the sensing signal for both the DGS (left axis, solid line data) and QPC (right axis, dashed line data). The transitions are taken at a fixed $V_g R$ and offset vertically for clarity. Both detectors resolve clear sensing peaks at $T \sim 20$ mK, with the QPC losing all sensitivity at elevated temperature $T \sim 1100$ mK. (d) DGS signal in a zoomed-up region showing a double dot charge transition. (e) The calibrated phase response from the DGS for a slice through (d) with $V_g R$ held at -608 mV. (f) Dispersive response of the gate-sensor where tunnelling to the reservoirs is suppressed in the few-electron limit. High tunnel rate, intra-dot transitions remain visible.
rf-QPC ($\tau_{int} \sim 0.5 \mu s$ [135]) and rf-SET ($\tau_{int} \sim 100$ ns [136]) however, there is considerable room for improving the sensitivity of the DGS, for instance, by further decreasing the parasitic capacitance.

To determine the bandwidth of the dispersive gate-sensor the SNR of its response is measured with increasing frequency of the small modulation voltage applied to a nearby gate (Fig. 5.2(e)). This method gives a detection bandwidth of $\sim 10$ MHz, limited by the $Q$-factor of the resonator, and consistent with the dependence of SNR with carrier frequency, as in Fig. 5.2(f). We further characterize the DGS by measuring how the height and width of the electron transition signal (see Fig. 5.2(c)) depends on applied resonator power, as shown in Fig. 5.2(g). Optimal SNR is achieved for a power at the resonator of $\sim -80$ dBm. At these powers, the transition width is $\sim 1$ mV, putting an upper bound on back-action onto the qubit.

Finally, we extract the relative geometric capacitive coupling between the sensor-gate and the quantum dot. The charging energy of the dot $E_c = e^2/2C_\Sigma$, can be measured by using the DGS to sense Coulomb diamonds as a function of source-drain voltage across the dot $E_c = eV_{sd}$, as shown in Fig. 5.2(h) (where $e$ is the electron charge and $C_\Sigma$ is the total dot capacitance). By measuring the period of CB oscillations we estimate that the gate-sensor geometric capacitance $C_g \sim 10$ aF contributes $\sim 5$ percent of $C_\Sigma$.

For a single quantum dot biased at the point where electron $n$ and $n+1$ are degenerate, the quantum capacitance is given by $C_Q = (e^2/4k_BT_e)(C_g/C_\Sigma)^2$ [133, 137], when the dot tunnel-rate is much larger than the resonator frequency ($k_B$ is the Boltzmann constant). This quantum capacitance shifts the resonance frequency by an amount $\Delta f \simeq C_Q f_0/2C_p$, ($C_p$ is the resonator parasitic capacitance). This frequency shift results in an observed phase response $\Delta \phi \simeq \alpha C_Q Q/C_p$, ($Q$ is the $Q$-factor of the resonator). The constant of proportionality $\alpha$ is of order unity at resonance and is related to the transmission coefficient of the resonator. For $T_e \sim 100$ mK and $C_g/C_\Sigma \simeq 0.05$ this formula gives $C_Q \simeq 9$ aF which is broadly consistent with our observed
phase shifts of $\Delta \phi \times 180/\pi \simeq 0.2$ degrees.

Having quantified the sensitivity of the gate-sensor, we now configure a double dot and show that this gate readout method can operate in the few-electron regime, where these devices are commonly operated as qubits. The double dot charge-stability diagram is detected using the dispersive gate-sensor as shown in Fig. 5.3(a), where regions of stable electron number are labeled $(n,m)$, corresponding to the number of electrons in the left and right dots. We confirm that the double dot is indeed in the few electron regime by also detecting the charge configuration using the proximal QPC charge sensor, as shown in Fig. 5.3(b).

Charge sensing using QPCs or SETs requires that the sensor be kept at a value of conductance where sensitivity is maximized. This is typically achieved by applying additional compensating voltages to gates when acquiring a charge-stability diagram. It is worth noting that gate-sensors do not require such offset charge compensation or gate voltage control. Of further practical use, we find that DGSs are robust detectors at elevated temperatures, in contrast to QPC charge sensors which suffer from a thermally broadened conductance profile and suppressed sensitivity with increasing temperature. The single-electron response of both QPC and DGS can be compared in Fig. 5.3(c) for $T \sim 20$ mK and $T \sim 1100$ mK.

The gate-sensor can be made to detect both intra- and inter-double dot tunnelling transitions, as shown in Fig. 5.3(d) which depicts a close-up region of the charging diagram. A line-profile of the transitions (Fig. 5.3(e)) indicates that the DGS is most sensitive to electron transitions from the right reservoir, due to its position, but is capable of distinguishing all transitions. Near an intra-dot transition, the quantum capacitance for the double dot can be shown to be $C_{dd}^Q = (e^2/2t)(C_g/C_\Sigma)^2$ where $t$ is the tunnel coupling energy of the double dot [132]. As for the single dot above, the phase shift (in radians) is $\Delta \phi \simeq \alpha C_{dd}^Q Q/C_p$. The measured phase shift $\Delta \phi \simeq 0.1$ degrees for the intra-dot transition is near half the shift for transitions to the leads,
consistent with a tunnelling coupling $t/h \approx 8$ GHz.

Increasing the tunnel barriers between the double dot and the reservoirs suppresses the gate sensing signal when the tunnel rate drops substantially below the detector resonance frequency ($f_0 \sim 704$ MHz). This regime is reached in Fig. 5.3(f), where transitions to the reservoirs are suppressed, but intra-dot transitions remain visible as these occur at a tunnel frequency above $f_0$. The observation of the intra-dot transition in the few-electron regime is important since it is this signal that forms the basis of spin qubit readout in these devices [132, 50, 129]. Of further note, in contrast to QPC or SET detectors that exhibit a broadband back-action spectrum [138], gate-sensors act-back on the qubit at a single, adjustable frequency.

The demonstration that *in situ* surface gates also serve as readout detectors with comparable sensitivity to QPCs is perhaps unexpected, given that the geometric gate-to-dot capacitance is only $\sim 5$ percent of the total capacitance. Readout using gate-sensors, however, makes use of the quantum capacitance which as we have shown, can be of the same order as the geometric contribution ($C_g \approx C_Q$). Gate-based readout then, has potential to address the significant challenge of integrating many QPC or SET detectors into large arrays of quantum dots, for instance, in the scale-up of spin qubit devices. The use of wavelength division multiplexing techniques [139, 140] would further allow each gate in an array to be independently and simultaneously read out at a unique frequency. Such an approach will also likely apply to systems without source-drain reservoirs altogether, such as donor qubits [141], or in the readout of Majorana devices [142].

We thank X. Croot and S. D. Bartlett for technical assistance and discussions. J.M.H. acknowledges a CSIRO student scholarship and use of CSIRO facilities for Nb inductor fabrication. We acknowledge funding from the U.S. Intelligence Advanced Research Projects Activity (IARPA), through the U.S. Army Research Office and the Australian Research Council Centre of Excellence Scheme (Grant No. EQuS CE110001013).
6. FREQUENCY MULTIPLEXING FOR READOUT OF SPIN QUBITS

6.1 Abstract

We demonstrate a low loss, chip-level frequency multiplexing scheme for readout of scaled-up spin qubit devices. By integrating separate bias tees and resonator circuits on-chip for each readout channel, we realise dispersive gate-sensing in combination with charge detection based on two rf quantum point contacts (rf-QPCs). We apply this approach to perform multiplexed readout of a double quantum dot in the few-electron regime, and further demonstrate operation of a 10-channel multiplexing device. Limitations for scaling spin qubit readout to large numbers of multiplexed channels is discussed.

6.2 Introduction

Scaling-up quantum systems to the extent needed for fault-tolerant operation introduces new challenges not apparent in the operation of single or few-qubit devices. Spin qubits based on gate-defined quantum dots [129] are, in principle, scalable, firstly because of their small (sub-micron) footprint, and secondly, since spins are largely immune to electrical disturbance, they exhibit low crosstalk when densely integrated [143]. At the few-qubit level, readout of spin-states is via quantum point contact (QPC) or single electron transistor (SET) charge sensors, proximal to each quantum dot [30, 114, 29,
These readout sensors pose a significant challenge to scale-up however, in that they require separate surface gates and large contact leads, crowding the device and tightly constraining the on-chip architecture.

The recently developed technique of dispersive gate-sensing (DGS) overcomes this scaling limitation by making use of the gates, already in place to define the quantum dots, as additional charge sensors [144]. The gates act as readout detectors by sensing small changes in the quantum capacitance associated with the tunnelling of single electrons. In turn, shifts in capacitance are measured by the response of a radio-frequency (rf) LC resonator that includes the gate. In principle, all of the quantum dot gates used for electron confinement can also be used as dispersive sensors, simultaneously collecting more of the readout signal that is spread over the total device capacitance and thus increasing the signal to noise ratio. Enabling all-gate readout, as well as multichannel rf-QPC or rf-SET charge sensing, requires the development of multiplexing schemes that scale to large numbers of readout sensors and qubits.

Here we report an on-chip approach to frequency multiplexing for the simultaneous readout of scaled-up spin qubit devices. We demonstrate 3-channel readout of a few-electron double quantum dot, combining two rf-QPCs and a dispersive gate-sensor as well as the operation of a 10-channel planar multiplexing (MUX) circuit. Similar approaches to frequency multiplexing have been demonstrated for distributed resonators in the context of kinetic inductance detectors [77], superconducting qubits [145, 146] and rf-SETs [139, 140, 147]. The present work advances previous demonstrations by lithographically integrating the feed-lines, bias tees, and resonators, which are fabricated on a sapphire chip using low-loss superconducting niobium. By putting these components on-chip, the size of the entire MUX circuit is reduced far below the wavelength of the rf signals, suppressing impedance mismatch from the unintentional formation of stub-networks [40] that otherwise occur in macroscale multi-channel feedlines. Finally, we briefly discuss
the ultimate limitations to scaling frequency multiplexing for spin qubit readout.

6.3 Experimental Design

Our readout scheme (Fig. 6.1(a)) comprises a multiplexing chip fabricated from a single layer of superconducting niobium film (150 nm, \( J_c = 15 \text{ MAcm}^{-2} \), \( T_c = 8.4 \text{ K} \)) on a sapphire substrate (r-cut, 3 mm × 5 mm × 0.5 mm) using optical photolithography and argon ion beam milling. The niobium remains superconducting at the moderate magnetic fields needed to operate spin qubits. Each inductor \( L_i \) in resonance with the parasitic capacitance \( C_p \) defines a unique frequency channel \( f_i = 1/(2\pi\sqrt{L_iC_p}) \) for addressing each readout detector. This multiplexing chip is mounted proximal to the spin qubit chip, consisting of a GaAs/Al\(_{0.3}\)Ga\(_{0.7}\)As heterostructure with two dimensional electron gas (2DEG) 110 nm below the surface (carrier density \( 2.4 \times 10^{15} \text{ m}^{-2} \), mobility 44 m\(^2\)/V s). Ti/Au surface gates define the quantum dots and readout sensors. Bondwires connect the inductors \( L_i \) on the multiplex chip to rf-QPCs via an ohmic contact \(^{135}\) or directly to the gates for the DGS readout \(^{144}\). The labels (i) - (iii) in Fig. 6.1(b) are used to identify frequency channels for the separate readout detectors. Each resonant circuit contains an integrated bias tee for independent dc voltage biasing. Both the multiplexing chip and qubit chip are housed together in a custom printed circuit board platform \(^{89}\) mounted at the mixing chamber stage of a dilution refrigerator with base temperature 20 mK.

The on-chip bias tees are constructed using inter-digitated capacitors (Fig. 6.1(d)) with critical dimension 3 µm and have size-dependent values between 3 pF and 5 pF, with lower frequency channels requiring a larger capacitance for similar insertion loss. To further increase the coupling capacitance we spin-coat the interdigitated sections with photoresist (AZ6612, \( \epsilon \approx 4 \)) to yield

\(^1\) Ansoft HFSS
Fig. 6.1: (a) Three channel frequency multiplexing scheme for spin qubit readout. The individual $LC$ resonator circuits comprise a matching inductor $L_i$, parasitic capacitance $C_p$ and a bias tee for independent biasing of each gate sensor. (b) Micrograph of the GaAs double dot device. Individual channels of the multiplexing chip are connected via bondwires to either a gate sensor (labelled (ii)) or an ohmic contact on one side of a QPC (labelled (i), (iii)). (c) Optical micrograph of the multiplexing chip which is patterned using niobium on a sapphire substrate, comprising interdigitated capacitors (d) and spiral inductors (e). (f), (g) Microwave transmission through bias tee components - measurement via a vector network analyser (VNA) and 3D numerical simulation$^1$.  

---

1. Note: The reference to the vector network analyser (VNA) and 3D numerical simulation seems to be a typographical error. The correct notation should be VNA and HFSS.
a larger dielectric constant than free space. The inductors (red, Fig. 6.1(e)), used in both the resonant circuit and bias tees, are spiral shaped with critical dimension \(3 \, \mu m\). The measured inductances (170, 250 and 400 nH) are in agreement with analytical calculations based on their geometry [148]. The self-resonance frequency of all the inductors is increased by over-etching the sapphire dielectric between adjacent turns, decreasing the effective dielectric constant and reducing the capacitance. Measurements of the transmitted power for the individual planar components are shown in Fig. 6.1(f,g) (blue, red trace) and yield agreement with numerics based on a 3D electromagnetic field simulation (black trace) \(^2\).

### 6.4 Quantum Dot Readout

The multiplexing scheme is implemented using a 3-channel chip to read out the state of a double quantum dot. The frequency response of the chip strongly depends on the state of the readout detectors, as shown in Fig. 6.2(a). In the absence of gate bias (black trace), the QPCs are far from pinch-off and the corresponding resonances are not apparent since the impedance of the \(LCR\) network is well away from the characteristic impedance of the feedline \((Z_0 \sim 50 \, \Omega)\). The resonances are formed (red trace) with the application of negative gate bias, depleting the electron gas and increasing the resistance of the QPC to bring the combined \(LCR\) network towards a matched load. Larger gate bias subsequently pinches-off the rf-QPC, further modulating the amount of reflected rf power at the resonance frequency. The response of the gate-sensor with bias is significantly different to that of the rf-QPC. For the gate-sensor, depleting the 2DEG beneath the gate also increases its resonance frequency, as shown in Fig. 6.2(c). This frequency dependence arises from the change in parasitic capacitance as the electron gas is depleted. With the gate voltages typically needed for defining quantum dots, the parasitic capacitance \(C_p\) is of the order of 0.3 pF. Electromagnetic field simulation suggests
contributions to $C_p$ are roughly equal between 2DEG, bondwires and adjacent turns of the planar inductors. Given the large separation in resonance frequencies, crosstalk is negligible in this 3-channel implementation.

We now demonstrate charge sensing measurements of a double quantum dot in the few-electron regime using this MUX configuration. The three independent readout channels (i, ii, iii) are separately addressable by selecting the rf carrier to match the respective resonance frequency. We note that direct digital synthesis can be used to create a single waveform that contains all of the separate carrier frequency components for each channel. The rf signal reflected from the MUX chip is first amplified at cryogenic temperatures before demodulation by mixing the generated and reflected rf tones. Low-pass filtering removes the sum component and, after further baseband amplification, yields a voltage $V_{rf}$ proportional to the response of the resonance circuit [135]. Alternatively, high bandwidth analog to digital conversion can dispense with the need for separate mixers for each channel by directly acquiring the reflected waveform and performing demodulation in software.

Readout via the QPCs (i and iii) exhibits a typical charge stability diagram in the few-electron regime as a function of gate bias $V_L$ and $V_R$ as shown in Fig. 6.3(a,c). The label $(m,n)$ denotes the number of electrons in the left and right quantum dot respectively with the colour axis proportional to the derivative of the readout signal with gate bias. In comparison to the rf-QPCs, the dispersive gate readout channel is insensitive to charge transitions that occur with tunnel rates below the resonator frequency [144]. Note that biasing the gates to tune the QPCs also shifts the voltages $V_L$ and $V_R$, such that their values are dependent on which sensor is being read out.
Fig. 6.2: (a) Frequency response of MUX circuit separating left rf-QPCs (i), dispersive gate sensor (ii), and right rf-QPC (iii) into separate frequency channels. With negative voltage applied to the gates, the frequency response (shown in red) exhibits resonances as the impedance of the readout sensors approach the characteristic impedance of the feedline. (b) and (d) show the frequency response of the left and right rf-QPCs as the gate voltage modulates the conductance. (c) shows the frequency response of the dispersive gate sensor with gate bias. Note the significant shift in resonance frequency as the gate capacitance is reduced by depleting the electron gas beneath.
Fig. 6.3: Multiplexed readout of double quantum dot in the few-electron regime. The derivative of $V_{rf}$ with respect to $V_L$, in arbitrary units, is shown as a function of the voltages on the left and right gates, $V_L$ and $V_R$. Charge stability diagrams (a), (b), (c) correspond respectively to readout using the separate channels (i), (ii) and (iii) as indicated in Fig. 2. Electron occupancy in the left and right dots is indicated by the labels (m,n). Note that when biasing the left and right QPC gates (needed for (a) and (c)) a different gate bias $V_L$ and $V_R$ is required for the same electron number.
Having demonstrated our approach to frequency multiplexing, we investigate the scalability of this scheme by operating a 10-channel chip shown in Fig. 6.4(a). The 10-channels are defined using inductors $L_i$ with values between 60 nH and 250 nH that form a resonant circuit with parasitic capacitance $C_p$ as described above. Each channel again integrates a bias tee, needed for independent biasing of the gate sensors. Operation of the 10-channel chip is tested at 4.2 K using a series of high electron mobility transistors (HEMTs) fabricated from a GaAs/Al$_{0.3}$Ga$_{0.7}$As heterostructure and connected to the MUX chip via bondwires. These HEMTs, shown in Fig. 6.4(c), act as independent variable resistors and mimic the response of 10 different QPCs for the purpose of testing the MUX scheme. With each HEMT connected to its corresponding resonator, the frequency response of the chip is shown in Fig. 6.4(b), firstly with all HEMTs in the high resistance state (black trace). Selectivity of each frequency channel is demonstrated by alternatively biasing even-numbered (blue trace) and then odd-numbered (red trace) HEMTs. The exact resonance frequency is set by the contribution to the parasitic capacitance from the HEMT, which depends on the extent to which it is depleted. In this demonstration we have not carefully adjusted the resistance of the HEMTs to optimize the $Q$-factor of each resonator.

Frequency multiplexing allows simultaneous readout but requires separate resonator and bias circuits for each readout channel. Although the size of our demonstration devices are large, the use of alternate fabrication methods will likely alleviate any road-block to scaling based on footprint. For instance, the use of multilayer processing for the capacitors $C_{bias}$ can shrink their footprint to $\sim 15 \mu m \times 15 \mu m$ for similar capacitance. The space occupied by the bias tee inductors $L_{bias}$ can be suppressed by using resistors instead of inductors to achieve high impedance. Reducing the critical dimension of the resonator inductors to $\sim 100$ nm results in a $55 \mu m \times 55 \mu m$ footprint for the largest (400 nH) inductor used here. Taken together, and assuming these
Fig. 6.4: (a) Optical micrograph of a 10:1 MUX chip with integrated bias tees. (b) Frequency response of the MUX chip with inductors $L_i$ each connected to GaAs HEMT with a gate-controllable conductance to mimic the response of 10 QPCs. Data shows the response with all HEMTs in the resistive state (black), odd HEMTs resistive (blue) and even HEMTs resistive (red). (c) Shows an optical micrograph of a section of the HEMT device with dashed lines indicative of bondwires.
superconducting circuits are fabricated on the same GaAs chip as the qubits, these dimensions suggest that thousands of readout channels are feasible in a moderately sized $1 \text{ cm} \times 1 \text{ cm}$ area.

A more serious challenge is frequency crowding arising from the limited bandwidth available using planar lumped element inductors. For a maximum resonance frequency of $\sim 5 \text{ GHz}$ and given the need to separate channels by several linewidths to suppress crosstalk, the total number of independent gate sensors that can be read out simultaneously is $\sim 100$. Beyond this number several approaches are possible. These include a brute force method, duplicating the reflectometry circuit, including cryogenic amplifiers for every bank of 100 channels. Alternatively, the available bandwidth can be extended by making use of distributed resonators [149], but these typically have larger footprints. Finally, if the constraint of simultaneous readout is relaxed, time domain multiplexing via cryogenic switching elements would allow readout of banks of frequency multiplexed channels to be interleaved in time. Whether qubit readout via such a time sequenced scheme is possible is likely dependent on the details of the particular quantum algorithm being implemented.

This research was supported by the Office of the Director of National Intelligence, Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office grant W911NF-12-1-0354 and the Australian Research Council Centre of Excellence Scheme (Grant No. EQuS CE110001013). Device fabrication is made possible through fabrication facilities at CSIRO (Lindfield) and the Australian National Fabrication Facility (ANFF). J. M. H. acknowledges a CSIRO student scholarship.
7. CRYOGENIC CONTROL ARCHITECTURE FOR
LARGE-SCALE QUANTUM COMPUTING

7.1 Abstract

Solid-state qubits have recently advanced to the level that enables them, in principle, to be scaled-up into fault-tolerant quantum computers. As these physical qubits continue to advance, meeting the challenge of realising a quantum machine will also require the development of new supporting devices and control architectures with complexity far beyond the systems used in today’s few-qubit experiments. Here, we report a micro-architecture for controlling and reading out qubits during the execution of a quantum algorithm such as an error correcting code. We demonstrate the basic principles of this architecture using a cryogenic switch matrix, implemented via high electron mobility transistors (HEMTs) and a new kind of semiconductor device based on gate-switchable capacitance. The switch matrix is used to route microwave waveforms to qubits under the control of a field-programmable gate array (FPGA), also operating at cryogenic temperatures. Taken together, these results suggest a viable approach for controlling large-scale quantum systems using semiconductor technology.
7.2 Introduction

Realising the classical control and readout system of a quantum computer is a formidable scientific and engineering challenge in its own right, likely requiring the invention of a suite of new devices with tailored physical properties. Already underway for this purpose is the development of near quantum-limited amplifiers [150, 151, 152, 153], small footprint circulators [154], ultralow loss resonators [75, 155], cryogenic filters [156], and interconnect solutions [81, 89, 157]. The hardware for classical data conversion and processing however, has yet to be tightly integrated with the quantum technology. Such a classical control interface must be fast, relative to the timescales of qubit decoherence, low noise, so not to disturb the fragile operation of qubits, and scalable with respect to physical resources[158, 159, 160]. A particular challenge is ensuring that the footprint for routing signal lines or the operating power does not grow rapidly as the number of qubits increases[161, 162]. As solid-state quantum processors will likely operate below 1 kelvin[143, 163, 80, 31, 164], there are advantages to also locating components of the control system in a cryogenic environment, adding further constraints.

Similar challenges have long been addressed in the satellite and space exploration community[165], where the need for high-frequency electronic systems operating reliably in extreme environments has driven the development of new circuits and devices [166]. Quantum computing systems, on the other hand, have to date largely relied on brute-force approaches, controlling a few qubits directly via room temperature electronics that is hardwired to the quantum device at cryogenic temperatures.

Here we present a control architecture for operating a cryogenic quantum processor autonomously and demonstrate its basic building blocks using a semiconductor qubit. This architecture addresses many aspects related to scalability of the control interface by embedding multiplexing sub-systems at cryogenic temperatures and separating the high-bandwidth analog control
`Prime-line / Address-line’ (PL/AL) architecture that separates prime analog waveforms, used to manipulate qubits, from the addressing data used to select qubits. (a) An example sub-section of a quantum algorithm shown using quantum circuit notation. The highlighted clock cycles include single-qubit rotations (yellow), a two-qubit gate (green) and readout operation (red). Note that multiple operations are intended in a given clock cycle such that the required analog waveform for control or readout can be connected in parallel to any qubit. (b) Prime-lines corresponding to a universal gate set are routed to qubits via a switch matrix controlled by the address-lines. Coloured paths correspond to the highlighted clock cycles in (a). Vertical dashed lines indicate the clocking of the analog prime-waveforms which occurs at a rate that is 10-100 times slower than the clocking of the address bus. The clock rate of the address bus will depend on its width and qubit coherence times.
waveforms from the digital addressing needed to select qubits for manipulation. Our demonstration makes use of a semiconductor switch matrix constructed using high electron mobility transistors and a new type of microwave switch element based on the gate-tuneable capacitive response of a heterostructure device. Under the control of a commercial, field-programmable gate array (FPGA) made to operate at 4 kelvin, the switch matrix is used to route microwave signals to selected quantum dot qubits at 20 mK. Bringing these sub-systems together in the context of our control architecture suggests a path for scale-up of control hardware needed to manipulate the large numbers of qubits in a useful quantum machine.

7.3 Control Micro-architecture

Our control micro-architecture executes a quantum algorithm by decomposing it into a sequence of universal quantum gates, allowing for arbitrary logic operations to be realized using a small set of repeated single- and two-qubit unitaries applied in sequence. At the level of physical qubits in the solid-state, whether they are spins [120], transmons [80, 143], or quasi-particles [142], these elemental gate operations amount to applying calibrated electrical waveforms to a particular set of qubits or pairs of qubits each clock cycle as determined by a quantum algorithm.

A key aspect of our control architecture is the separation of these analog 'prime waveforms', which are typically pulses at microwave frequencies, from the digital qubit addressing information that determines which waveform is directed to which qubit, at a particular point in the code. In comparison to brute-force approaches, this scheme lifts the need of having a separate waveform generator and transmission line for each qubit, taking advantage of a small universal gate set that uses the same analog waveforms over-and-over throughout the algorithm. As realistic qubits will inevitably include variations in their physical parameters, the control architecture must also
incorporate means of calibrating and adjusting the response of the qubit to the control waveforms, as described below.

Our ‘prime-line / address-line’ (PL/AL) architecture is shown schematically in Fig. 7.1, where we have drawn part of a circuit for implementing a quantum error correcting surface code [167, 168]. Precisely timed analog prime waveforms, generated at cryogenic or room temperature, propagate cyclicly on a high-bandwidth prime-line bus that is terminated with a matched impedance at a location in the system where heat can be dissipated. The quantum algorithm is then executed exclusively via the digital address-line bus, selecting qubits and qubit pairs to receive the appropriate prime waveform at the correct clock cycle in the circuit. Readout proceeds in a similar way, with the digital address bus selecting a particular qubit (or readout device) for interfacing with multiplexing devices [146, 169] and analog readout circuitry such as a chain of amplifiers and data converters.

### 7.4 Implementation of the Control Architecture

Realising our PL/AL architecture requires integrating multi-component control and readout hardware with the quantum system of qubits fabricated on a chip. Owing to the large number of qubits that are likely to be needed for quantum computation and the timescales involved in their control, there are advantages to locating sub-systems of the control architecture at cryogenic temperatures, either on-chip with the physical qubits, or in close proximity and connected via integrated multi-chip modules [170] and compact transmission lines. Aspects of the control system will however, generate significant heat or fail to function at the millikelvin temperatures needed for qubit operation. The competing constraints of interconnect density, heat generation, signal latency, footprint, and noise performance suggest a control architecture that is distributed across a cryostat, taking advantage of the significantly different thermal budgets available at each temperature stage. This distributed
architecture is illustrated in Fig. 7.2, where control sub-systems are positioned at different temperature stages of a cryogen-free dilution refrigerator. Below we describe and provide a basic demonstration of these sub-systems.

7.5 Switching Matrix

The key sub-system underpinning the control micro-architecture is a switch matrix, or routing technology that steers the prime waveforms to particular qubits based on a digital address. This technology is ideally located in close proximity to the qubits to avoid latency and synchronization challenges that arise when signals propagate over length-scales comparable to the electromagnetic wavelength (typically centimetres for quantum control waveforms). Physically integrating the switch matrix and qubit system has the further advantage of significantly reducing the wiring and interconnect density by making use of lithography (or multi-chip module packaging) to provide connection fan-out. In this way we envisage a switch matrix that receives multiplexed data on a small number of transmission lines and decodes this address data to operate large numbers of parallel switches (see Fig. 2). Multiplexing of this kind will likely be essential for operation in cryogenic environments where large numbers of parallel transmission lines add a sizeable heat load when carrying signals between stages that are at different temperatures. The use of superconducting materials is key as these can dramatically reduce the cross-section and thereby thermal load of transmission lines without degrading electrical performance [170].

A switch matrix with elements that act as variable impedances can also be configured to enable the amplitude and phase of the prime waveforms to be individually adjusted before arriving at each qubit. By incorporating a calibration routine or feedback scheme, this approach can be used to account for the variation in physical parameters that will inevitably occur with systems comprising large numbers of qubits.
Fig. 7.2: Schematic of a control micro-architecture that distributes subsystems across the various temperature stages of a dilution refrigerator, depending on the available cooling power (image is of a Leiden Cryogenics CF450). A millikelvin switch matrix, on the same chip as the qubit device or close to it, steers a small number of control pulses to qubits using addressing information from cryogenic logic at 4 K. The matrix will incorporate a level of digital decoding to enable switch addresses to be transmitted on a relatively small number of serial lines. The cryogenic logic also interfaces with multiplexed readout and digital-to-analog converters. The 4 K stage typically has a cooling power $\sim 1$ W, with the 20 mK stage having less than 10 $\mu$W.
Various technologies appear suitable for constructing such a switch matrix, including semiconducting devices [171, 172, 173], mechanical systems [174, 175], and superconducting logic [105, 99]. For qubit technologies built from semiconductors [144, 164], field-effect based devices are ideally suited owing to their sub-nanosecond switching-speed, gigahertz transmission bandwidth, low dissipation, small footprint, cryogenic compatibility, and opportunity for integration with qubits. Below we demonstrate the operation of such devices using GaAs high electron mobility transistor (HEMT) circuits, configured as a switch matrix with variable amplitude and phase response. We note that complex circuits constructed from HEMTs demonstrate that these devices are well suited to extensive fan-out [176].

7.5.1 HEMT Switching Elements

A prototype HEMT-style microwave switch based on a GaAs/AlGaAs heterostructure is shown in Fig. 7.3(a,b). Fabrication of these switching elements follows a similar procedure to quantum dot qubit devices (allowing easy integration). The mesa is wet etched using sulphuric acid, before Au/Ge/Ni ohmic contacts are thermally evaporated and annealed at 470 degrees for 100 seconds. The final metal layer is thermally evaporated TiAu (10 nm / 100 nm). In the on-state, the switch is configured to have a characteristic impedance of $\sim 50 \, \Omega$, owing to its coplanar waveguide (CPW) geometry. Prime waveforms are fed to and from the HEMT two-dimensional electron gas (2DEG) via eutectic ohmic contacts and TiAu planar transmission lines. In the off-state a negative voltage applied to the TiAu top gate pinches-off the electron gas channel, reflecting the prime waveform signal due to the large impedance of the HEMT relative to the characteristic impedance of the $\sim 50 \, \Omega$ feedline. The transmission response of the switch is shown in Fig. 7.3(c), with an on/off ratio (OOR) above 40 dB in the frequency range 0 - 2.5 GHz, suitable for control of spin qubits [177]. For these prototype devices a large insertion loss of 10-20 dB is observed, owing mostly to the
resistance of the ohmic contacts, which is currently a few hundred ohms in our process. Precise control of the contact resistance and capacitance using ion-implantation can overcome this limitation and also dramatically shrink the footprint of these devices [178, 179].

The time-domain response of the switch is demonstrated by amplitude modulating an applied 120 MHz constant wave tone, as shown in Fig. 7.3(d). To determine the maximum switching time of the HEMT we modulate a 5 GHz carrier tone with a sinusoidal waveform applied to the gate and measure the depth of modulation as a function of gate frequency, as indicated in Fig. 7.3(e). For these prototype devices the switching time is of order 1 ns.

7.5.2 Capacitive Switching Elements

Microwave switching devices based on the depletion of an electron gas also enable a new capacitive mode of operation. In this configuration the CPW feedline transitions to a microstrip geometry by contacting the electron gas to the planar ground planes using ohmic contacts, as illustrated in Fig. 7.4(a,b). The two conductors in the microstrip transmission line are thus constructed using the top gate and electron gas as ground. This device can act as a reflective switch by depleting the effective ground plane using a negative bias on the gate. Depletion reduces the capacitance between the conductors of the microstrip and modulates the device impedance. Transmission through the switch is shown in Fig. 7.4(c) in the on (blue) and off (red) state, with an OOR greater than 25 dB for 0 - 8 GHz. To the best of our knowledge, a switching device based on a depleted ground plane has not been reported previously.

The switch is capacitively coupled to the input and output ports, with a planar spiral inductor at one port forming a bias tee to provide the dc gate voltage needed to deplete the electron gas. In place of a planar interdigitated capacitor, we make use of the GaAs heterostructure to provide a low footprint
Fig. 7.3: **Characterisation of a HEMT switch as a building block for the PL/AL architecture.** (a) Microscope photograph of the device fabricated on GaAs/Al$_{0.3}$Ga$_{0.7}$As heterostructure. (b) Schematic cross-section showing the coplanar line diverted through the 2DEG. A negative voltage (-300 mV) on the top gate increases the impedance of the switch, reflecting the input signal. (c) Transmission as a function of frequency for the on (blue) and off (red) state. (d) Example of time-domain response. When the gate voltage (green) is zero, the 120 MHz sine wave provided at the switch input is propagated to the output (blue), and not otherwise. (e) Modulating a carrier signal through the 2DEG with a sinusoidal gate voltage creates sidebands. The amplitude of the sidebands as a function of frequency indicates a 1 - 2 ns switching time.
7. Cryogenic Control Architecture for Large-Scale Quantum Computing

Fig. 7.4: A switch design that produces an impedance mismatch by depleting the transmission line ground plane. Shown is an image (a) and cross-sections (b) of the device. The input is a coplanar line (i) which transitions to a microstrip using the 2DEG as the ground plane (ii,iii). The ground plane is tapered so that it lies beneath the signal track (iv,v). When a negative voltage is applied to the signal track, the ground plane is depleted (v) and the impedance mismatch reflects the input signal. (c) Transmission through the switch in the on (blue) and off (red) states. (d) Frequency response of capacitors formed using surface gates and 2DEG as a parallel plate (inset).
parallel plate capacitor, formed between the CPW central track and the electron gas, as shown in the inset to Fig. 7.4(d). The frequency response of this capacitor is shown in Fig. 7.4(d).

The capacitance-based switch has improved performance at higher frequency than the HEMT-based switch, although it has a larger footprint due to both the length of line needed for adiabatic tapering from 50 to 200 Ω and for the coplanar-to-microstrip transition. Working with a characteristic impedance of \( \sim 200 \Omega \) minimises the area of electron gas and reduces ohmic loss. The improved frequency performance stems from the absence of a gate structure, which in the HEMT switch capacitively couples the source and drain contacts, even in the off state. The required footprint is reduced significantly in an all-microstrip circuit that is designed to operate at a characteristic impedance close to 200 Ω. In their current form the performance of both kinds of switches is better suited to controlling spin qubits, where the frequency of signals are of order 1 GHz. For superconducting qubits, we envisage extending the operation of these switches to frequencies in the 4 - 12 GHz range by shrinking their footprint to suppress parasitic capacitances and inductances that lead to resonances in the present design.

### 7.5.3 2:2 Switch Matrix

We demonstrate cryogenic operation of a prototype routing matrix based on HEMT switches with two input and two output ports. A magnified image of the device is shown in Fig. 7.5(a) with associated schematic in (b). Each input port is split and connected to each output port via a switch so that the transmission parameters \( S_{ij} \) of the device are controlled by the respective gate voltages \( V_{ij} \). The output ports include bias tees, which are needed for use with qubits based on semiconductor quantum dots. Bias tees are constructed using planar spiral inductors and 2DEG-based capacitors as illustrated in the inset of Fig. 7.4(d).
Fig. 7.5: Small scale 2-input, 2-output switch matrix based on HEMT switches, with on-chip bias tees for quantum dot operation. Device image is shown in (a) with associated circuit diagram in (b). (c) Transmission measurement with path A (blue) in the on-state and path B (red) in the off-state. (d) Voltage output with a 1 GHz input tone where path A is in the on-state and path B is (i) off, and (ii) half-on. (e) An example of IQ modulation, implemented by feeding the input ports of the 2:2 matrix with signals that have a 90° phase offset. Arbitrary amplitude and phase is produced at the output (data shown in figure) by selecting the appropriate $V_{i,j}$ (see main text). (f) Example voltage output for one of the constant amplitude quarter circles in (e).
Operation of this switch matrix is demonstrated by comparing the transmission of signals as a function of frequency for path A (blue) and path B (red), as indicated in Fig. 7.5(b). The response through both paths when path A is on ($V_{3,1}=0$) and path B is off ($V_{4,1}=-500$ mV) is shown in Fig. 7.5(c). The corresponding time-domain response for a 1 GHz tone is shown in Fig. 7.5(d)(i). We observe a negligible ($< 0.05$ dB) change in the response of one path when the other is path is switched from the on state to the off state.

An advantage of semiconductor-based switching elements is their ability to be configured as variable impedances, producing arbitrary amplitude output, as shown in Fig. 7.5(d)(ii).

We also demonstrate basic IQ modulation using our switch matrix by applying rf tones at both inputs with a 90° phase offset between them. The 90° shift can be produced by a length of transmission line (with narrowband response) or as a separate quadrature prime waveform. The output waveform at angular frequency $\omega$ is $A \sin \omega t + B \cos \omega t = R \sin(\omega t + \phi)$, where the magnitude $R$ and phase $\phi$ are determined by the amplitudes $A$ and $B$, controlled by the gate voltages $V_{i,j}$. After the calibration function $R, \phi = F(V_{i,j})$ is generated once, we can select the appropriate $V_{i,j}$ to produce a tone with arbitrary phase and amplitude in the first quadrant of the complex plane, as shown in Fig. 7.5(e). The corresponding voltage output along a quarter circle of constant amplitude is shown in Fig. 7.5(f). By controlling the amplitude and phase shift using the integrated switch matrix, the connection between each qubit and the prime line bus can be specifically adjusted to compensate for the inevitable variation in parameters between physical qubits\(^1\).

\(^1\) Calibration of the switch response and qubit can be performed at the same time, measuring the qubit evolution as a function of switch gate voltage.
7. Cryogenic Control Architecture for Large-Scale Quantum Computing

7.6 Cryogenic Logic

For controlling and programming the switch matrix via the address bus, we envisage a layer of fast, classical logic that serves as an interface between the physical qubits and the compiled quantum algorithm (which will likely comprise mostly an error correcting code). This layer of classical logic is also needed for executing various automatic sequences associated with fast feedback for qubit stabilisation, readout signal conditioning, or open-loop error suppression[180, 34]. For controlling a large-scale quantum computer there are many advantages to locating this classical logic and associated data converters close to the qubits, inside a dilution refrigerator. In comparison to room temperature based control systems, cryogenic operation results in an enhanced clock speed, improved noise performance, reduced signal latency and timing errors, and larger bandwidth. Some of these aspects stem from the ability to make use of compact superconducting transmission lines and interconnects at cryogenic temperatures. We note that locating control electronics inside the vacuum space of the refrigerator allows it to be positioned physically close to qubit device, even if qubits and control systems are at moderately different temperatures.

The choice of technology for constructing this layer of classical control is largely dictated by the qubit coherence times, control signal bandwidth, and the number of simultaneous qubits under control. With a convergence of solid-state qubit coherence times now approaching 1 millisecond [181, 33, 182], present day CMOS-based FPGAs or application specific integrated circuits (ASICs) operating at 4 kelvin are a viable control platform. Higher performance control systems that are likely to be realized in the longer term include technologies based on InP devices [183], SiGe BiCMOS [184, 185], and superconducting flux logic [105, 99].

For the basic demonstration of the PL/AL scheme considered here the classical logic is implemented using a commercial FPGA (Xilinx Spartan-3A)
that we have made operational at the 4 K stage of a dilution refrigerator. To achieve cryogenic operation the FPGA chip was mounted on a custom, cryogenic printed circuit board that includes components which vary little in their parameters at cryogenic temperatures [89, 157]. Power and clock signals to the FPGA are adjusted for cryogenic operation using room temperature sources and a semi-rigid coax line is configured for sending serial commands, with reprogramming of the low temperature array occurring via a dedicated ribbon cable. With the FPGA mounted at the 4 K stage we measure an idle power dissipation of $\sim 30$ mW, with negligible increase during dynamic logic operations for the simple code executed here. We estimate a dynamic power dissipation of $\sim 100$ mW for computational operations that use most of the gates in the Spartan-3 array (further details of cryogenic operation of FPGAs are given elsewhere [186]). The FPGA is programmed to interpret serial communication and output a 3.3 V signal on selected pins to activate prime waveform routing in the switch matrix. These outputs are combined with a negative voltage provided from room temperature via a cold resistive adder so that the switch matrix gates receive -50 mV for the on-state and -380 mV for the off-state voltage.

7.7 Semiconductor Qubit Control

We combine the building-blocks of our micro-architecture described above, to demonstrate that a semiconductor qubit can feasibly be controlled autonomously without introducing additional noise or heating to the quantum system. The qubit is a GaAs double quantum dot configured as a charge or spin qubit in the few-electron regime (the heterostructure has a carrier density $2.4 \times 10^{-15}$ m$^{-2}$ and mobility 44 m$^2$/Vs at 20 K). These qubits are commonly controlled using dc-pulse waveforms on the gates to rapidly manipulate the energy levels of the quantum dots [32]. A typical setup connects a waveform generator to each gate using a separate high bandwidth coaxial
cable and bias tee.

For this demonstration we connect a single coaxial cable from a waveform generator at room temperature to the input of the 2:2 switch matrix, with the two matrix output ports connected to the two plunger gates \( LP \) and \( RP \) of the double dot, as shown schematically in Fig. 7.6(a). The waveform generator produces a prime waveform consisting of a 100 kHz square wave (shown in Fig. 7.6(b)) which is then steered by the 4 kelvin FPGA by opening and closing switches in the matrix depending on commands sent from room temperature.

The charge state of the double dot is sensed using an rf quantum point contact [135, 169], which provides a readout signal \( V_{rf} \) as a function of the gate voltages \( V_L \) and \( V_R \) indicated in (c). With both switches of the matrix set to the off state, a standard charge stability diagram is detected indicating that the off state provides sufficiently high isolation between input and output ports, as shown in Fig. 7.6(d)\(^2\). In contrast to using the qubit decoherence time to detect additional noise sources from the control circuits, we note that the width and jitter of a quantum dot charge transition provides a broadband probe of electrical noise, including fluctuations that occur on timescales much longer than the qubit coherence.

Sending a command to the cold FPGA allows the prime waveform to be directed to the left, right, or both plunger gates, producing two copies of the charge stability diagram. These copies appear because, on the timescale of the readout, a square wave with 50% duty-cycle configures the double dot in two distinct charge states that are offset from one another by the voltage \( \Delta V_R \) or \( \Delta V_L \), as shown in Fig. 7.6(e-g). We note that the shift measured in Fig. 7.6(g) is the vector sum of the shifts in (e) and (f), account for the cross capacitance between each gate and each dot [187]. In comparison to data taken on the bare quantum dot, we are unable to detect any additional noise.

\(^2\) A very small amount of jitter in the charge transitions can be seen due to coupling of the rf-QPC carrier to the gates.
Fig. 7.6: **Simple implementation of the micro-architecture introduced in Fig. 7.2.** (a) Experimental setup for measuring a double quantum dot, using a cryogenic FPGA to steer pulses via a millikelvin switch matrix. Charge-state readout is performed using an rf-QPC. (b) Switch matrix output showing a 100 kHz square wave directed to plunger gates of the quantum dot. (c) Micrograph of the quantum dot device. The shaded gates, labelled LP and RP, are connected to the switch matrix output. (d-g) Charge sensing of the double quantum dot in the few-electron regime, with electron occupancy indicated by the labels \((m, n)\). The colour axis is the derivative of the sensing signal \(V_{rf}\) with respect to \(V_R\). When the FPGA-controlled switch matrix blocks waveforms (d), a standard double dot stability diagram is detected. When the square wave is directed to either LP (e), RP (f) or both (g), copies of the stability diagram appear (see text). These measurements demonstrate that the double dot potential can be controlled autonomously by the switch matrix and cold FPGA.
or an increase in the electron temperature (which is of order 100 mK) when configuring the charge-state using the cryogenic FPGA and switch matrix.

7.8 Discussion

Our simple demonstration of a multi-component control architecture provides a path for scaling up the classical support system needed for operating a large-scale quantum computer. Aspects of this demonstration will also likely find immediate use in improving the performance of few-qubit experiments using electron spins in quantum dots. For example, in using the switch matrix to produce multiple out-of-phase copies of control waveforms, crosstalk can be suppressed by cancelling the voltage that is capacitively coupled to neighbouring surface gates [188]. Using the switch matrix as a high frequency cryogenic multiplexer will also enable the automated testing and characterisation of many devices in a given cool-down experiment. In the longer term, our micro-architecture can be extended to allow additional functionality of the switch matrix, providing qubit control frequency correction by using the HEMTs as mixers, or as cryogenic adder circuits that reduce the complexity or resolution needed for biasing surface gates that define quantum dots.

We comment here also on the possibility of implementing our PL/AL control architecture using a switch matrix based on single flux quanta superconducting logic [99]. Such logic already appears well suited to control flux-based qubits at high speed and with low dissipation. To what extend these devices are compatible with magnetic fields and the need to generate and steer large voltage waveforms required in the operation of semiconductor qubits remains an open question.

At the layer of classical logic, our demonstration shows that commercial FPGA devices can be configured to work at cryogenic temperatures and are compatible with controlling qubits in close proximity. Beyond the control ar-
chitecture presented here, the use of cold, low-latency classical logic will likely improve the performance of feedback systems generally needed for adaptive measurement, quantum state distillation, and error correction protocols. A further consideration is the heat generated by the switch matrix, which must operate at the mixing chamber stage of a dilution refrigerator. Given that these switches are reflective, rather than dissipative at microwave frequencies, heat generation will be dominated by the charging of the gate capacitance with each switch, as is the case for today’s room temperature CMOS technology. For 1000 HEMT switches of the kind shown here operating at a clock frequency of 1 GHz, we estimate a total power dissipation 100 $\mu$W. Straightforward improvements in switch design, such as a reduction in sub-threshold voltage swing at low temperature, can likely reduce dissipation by a factor of 100. Even so, improvements in cryogenic refrigeration technology, both at the chip-level [189] and cryostat, similar to what has been achieved in rare-event physics [190], will likely be needed to enable large-scale quantum information processing.

We have proposed a micro-architecture for the control of a large-scale quantum processor at cryogenic temperatures. The separation of analog control prime waveforms from the digital addressing needed to select qubits offers a means of scaling this approach to the numbers of qubits needed for a computation. To demonstrate the feasibility of our scheme we have shown that a semiconductor qubit can be controlled using a cryogenic FPGA system and custom switch matrix for steering analog waveforms at low temperature. We anticipate that integrated, autonomous control systems of this kind will be increasingly important in the development and demonstration of fault tolerant quantum machines.

We thank B. Smith, D. Tuckerman, D. Wecker, K. Svore, C. M. Marcus, L. DiCarlo, L. P. Kouwenhoven, and M. Freedman for useful conversations. Devices were fabricated at the NSW Node of ANFF. This research was supported by the Office of the Director of National Intelligence, Intelligence Ad-
vanced Research Projects Activity (IARPA), through the Army Research Office grant W911NF-12-1-0354, the Australian Research Council Centre of Excellence Scheme (Grant No. EQuS CE110001013), and Microsoft Research.
8. FUTURE DIRECTION

Quantum computation can access a range of powerful algorithms for problems that cannot otherwise be efficiently solved.

Before these advances are to be realised, the number of qubits coherently coupled in a system must increase substantially from current state-of-the-art. It is likely that brute force approaches, in which each qubit requires a direct duplication of classical hardware resources, will become infeasible around a few hundred qubits. This thesis has explored various techniques with this goal in mind: the development of quantum hardware facilitating development of a scalable quantum machine. This chapter will review the key experiments from the thesis and discuss natural next steps for these ideas, in the context of this scaling up.

In Chapter 3, the effect of an inevitable parasitic electromagnetic environment on superconducting coplanar waveguide resonators is investigated. The presence of parasitic channels loads the resonator $Q$-factor in a way that cannot be circumvented by, for instance, reducing the coupling capacitance.

These effects will become more important as devices become more complex, where high coherence times are required in systems with many quantum elements and microwave ports. Precise management of the electromagnetic environment combined with several recent techniques (for instance, [192, 74, 75]) will be needed for large-scale use of superconducting resonators as quantum busses. An example of a such a design is the Kagome lattice, shown in Fig. 8.1(a), reproduced from [191].
Fig. 8.1: Ideas for scalable qubit layouts. (a) A Kagome lattice, reproduced with permission from [191] - a planar array of superconducting coplanar waveguide resonators. (b) Dispersive readout and frequency multiplexing allows readout of arrays of quantum dots. Shown here is six dots, fabricated by Sylvain Blanvillain.

In Chapter 5, a readout technique for gate-defined quantum dots is presented using a surface gate as a readout detector. From a response to the quantum capacitance, these gate-sensors have a demonstrated sensitivity comparable to measurements using a quantum point contact, and in-principle a sensitivity beyond that [193].

The footprint of a gate-sensor on the quantum dot chip is no more than the dot-defining gate itself, and readout can be performed using any of the available gates. This paves the way for scalable arrays of quantum dots without requiring QPC detectors near each dot. A small scale example of such an array is shown in Fig. 8.1(b), with six quantum dots indicated by shaded regions.

The next consideration is managing readout from multiple sensors, and this is addressed in Chapter 6. A scheme for low-loss chip-level frequency multiplexing is demonstrated that incorporates both traditional QPCs and the
8. Future Direction

Fig. 8.2: The next iteration of frequency multiplexing chips, as described in Chapter 6. The six-channel devices have lower loss across the bias tee capacitors and occupy a smaller footprint. Fabrication carried out at 8.2.

dispersive gate-sensors introduced in Chapter 5, and three channel readout of a double quantum dot was demonstrated.

Once the scheme is verified, Fig. 8.2 shows a wafer of foundry-produced [194] frequency multiplexing chips for six-channel readout and occupying a smaller footprint (\(\sim 1 \text{ mm} \times 2 \text{ mm per channel}\)). State of the art electron-beam lithography techniques combined with multi-layer devices could reduce this significantly, to the point where bandwidth becomes the limit to further scaling. If non-simultaneous readout is permitted, one can use the cryogenic RF switches in Chapter 7 to multiplex readout in the time domain. Logic-controlled switches can then steer readout between banks of frequency-multiplexed qubits. Alternatively, switches close to the device can be used to select a set of sensors for simultaneous readout at a given point in the algorithm.

Chapter 7 brings the ideas of the previous chapters together in a demonstration of the prime-line / address-line architecture outlined in Figures 7.1
8. Future Direction

Fig. 8.3: A PCB with an FPGA, digital-to-analog converter and analog-to-digital converter for cryogenic qubit control, designed by Ian Conway Lamb.

and 7.2. A quantum control scheme suitable for large numbers of qubits is presented, which incorporates cryogenic pulse steering for qubit operations as well as the readout techniques presented earlier in the thesis. A simple demonstration using a double quantum dot is shown, where an FPGA at 4 K directs a square wave to either dot via a 2-input, 2-output switching matrix.

This steering arrangement is based on two ideas: cryogenic logic and a millikelvin switching array. The logic demonstrated in Chapter 7 was simplistic, and more complicated arrangements are envisaged. Fig. 8.3 shows a more advanced board for cryogenic control. The board includes a digital-to-analog converter and analog-to-digital converter in addition to the FPGA. With this arrangement, qubit pulses and RF reflectometry can be performed within the dilution refrigerator, reducing the number of lines from room temperature and permitting adaptive qubit control without the delay imposed by communication with a room temperature instrument.

This logic need not be exclusively CMOS or even semiconductor-based. Su-
perconducting circuits based on SFQ pulses are fast and with low heat dissipation - well suited to qubit environments - and can be developed for computation and data conversion as well as scaling tools such as multiplexing. Chapter 4 presented software useful for the development of large SFQ circuits. There is scope for improvements in speed, employing sparse matrix algorithms and SSE instructions for modern processors, and in functionality, implementing various algorithms to efficiently traverse the large parameter space of these circuits.

Chapter 7 presented two reflective switching ideas based on the GaAs heterostructure used for quantum dot devices, and incorporated them into a 2-input, 2-output device. Fig. 8.4 shows how this might scale beyond a single qubit, with a 4-input, 5-output device with 20 switches. This device is fabricated using standard optical photolithography and further scaling is possible within existing technology. In the case of gallium arsenide, it is not
infeasible to have the quantum dots, the switching matrix and the superconducting readout scheme implemented on a single compact chip.

Taking a slightly more distant view, one can imagine the hardware for controlling a quantum processor as a series of modular components distributed across all stages of a dilution refrigerator. Examples of such basic modules can be found in this thesis: a frequency multiplexing chip, 4 K semiconductor electronics, or a switch matrix. The experimentalist can then select components from this toolbox depending upon the qubit architecture and intended experiment. A number of (frankly, exciting) technological advancements still need to occur before useful algorithms with an error correction scheme become feasible. These areas include, but are far from limited to, high density superconducting wiring, low power classical processors and data converters, and multi-layer nano-scale fabrication, in addition to advances in the qubit devices themselves. Nothing in this list is prohibitive, and the rate of development in recent years suggests cautious optimism is not an unreasonable position to take.
APPENDIX
A. NIOBIUM FABRICATION

This appendix details fabrication techniques for the superconducting coplanar waveguide resonators in Chapter 3, spiral inductors in Chapter 5 and frequency multiplexing chips in Chapter 6. Both devices are made using niobium thin film, usually with a sapphire substrate. The Nb is sputtered across the entire substrate, followed by optical photolithography and ion beam milling.

A.1 Nb sputtering

Niobium film tends to lift-off substrates, and some treatment of the substrate is required to get a film that will not lift off during sonication or wire bonding. The substrates used are typically polished r-cut sapphire, and they are prepared before sputtering:

- Sonication with detergent. This produced a cleaner wafer under magnification than a clean with traditional solvents only, although it may be this was an unclean batch of substrates

- Sonication with dionised water (DIH₂O) then drying with nitrogen (N₂) gas, to completely remove the detergent

- Sonication with acetone

- Sonication with isopropanol (IPA), N₂ drying
The Nb films are further cleaned using an ion beam mill at a lower voltage (300 V, 5 minutes) to provide a surface that is more even than the polishing, verified by profilometer measurement. The Nb film is then sputtered without breaking vacuum. The system is typically evacuated overnight, reaching a base pressure of $\sim 7 \times 10^{-7}$ mbar. A cooling stage inside the vacuum system can be filled with liquid nitrogen, producing a lower pressure ($\sim 4 \times 10^{-7}$ mb) via cryotrapping, although this is unnecessary.

The Nb films are magnetron sputtered in an argon environment (300 W, 0.7 Pa) at a rate of $\sim 0.9$ nm s$^{-1}$. Films produced in this way are smooth and have no problems with lift-off either in fabrication or in use.

### A.2 Patterning

Films are patterned using optical photolithography. Following sputtering, chips are spin-coated in photoresist (resist S1813, 4000 rpm, 60 seconds produces a 1.5 µm layer) and baked for 60 seconds at 105 degrees. The chips are then dipped in developer for 5 seconds before rinsing in DIH$_2$O and drying. This slight hardening of the S1813 produces a sharper edge profile after exposure and developing.

The chip is exposed using photomasks from [195] using a vacuum pressure (produces sharper edges, required for the moderately small features over a large area in the multiplexing chips) for a time depending on the intensity of the globe, which varies slightly over time. A typical exposure is 13 seconds with an intensity 6.5 mW cm$^{-2}$.

The chip is baked again (60 seconds, 135 degrees) and then developed for 20 seconds, rinsed with DIH$_2$O and dried. Spinning and photolithography would often have to be re-done for the devices in Chapter 6 due to the large area over which a small imperfection could make the device unusable. If the resist looks precise, the ion beam milling result will also.
A.3 Ion Beam Milling

Excess Nb is removed using argon ion beam milling (500 V, base pressure $1 \times 10^{-6}$ mb, etch rate $\sim 12$ nm min$^{-1}$). To ensure the Nb is completely etched, the substrate ends up being etched to some extent. This is not a problem for the devices in this thesis. In Chapter 3, over-etching reduces the number of substrate defects\footnote{Defects in the metal oxide and at the substrate / metal interface are, of course, unaffected, and these are thought to be the dominant source of TLS inducing defects, e.g. [56]} (resulting in quantum mechanical two-level systems and contributing to resonator loss), although it will change the effective dielectric constant $\epsilon_{eff}$ and consequently the characteristic impedance $Z_0$ of the line. In Chapter 6, over-etching in spiral inductors reduces the parasitic capacitance between adjacent turns. It also reduces the capacitance of finger capacitors, although this is ameliorated to some extent by spin-coating the completed capacitor.

The photoresist is removed using sonication in acetone for at least 5 minutes followed by sonication in IPA. This aggressive resist removal is typically required since the ion beam heats the photoresist during the etching process, even with the chip well thermalised to a water cooled stage. For long etching processes, a beam shutter can be periodically closed then opened to allow the chip to cool.

A.4 Other Comments

Niobium Oxide

Niobium forms a thin oxide layer, primarily Nb$_2$O$_5$ of 2-3 nm. This layer does not affect wire bonding, so there is no need for a capping metal layer. For some earlier devices, a gold capping layer was used at wire bond locations.
Gold is magnetron sputtered immediately following niobium (without breaking vacuum, so an oxide layer will not form) at 60 W at a rate of 0.5 nm s\(^{-1}\). The gold is patterned in the same way as above, using reactive ion etching to remove excess gold. This process is not as precise as ion beam milling, so not used for critical features, but it is significantly faster.

\textit{Rabbit Ears}

A potential feature that can appear with ion beam milling is shown in Fig. A.1(b). This is a coplanar track as part of a superconducting coplanar waveguide resonator as studied in Chapter 3. The cross-section of the feature is shown in Fig. A.1(a). During ion beam milling, niobium is re-sputtered along the edge of the photoresist walls, and it remains once the photoresist is removed. The cross section of the metal is shown in (iv), verified by a profilometer measurement, and gives the feature its informal name ‘rabbit ears’.

The feature can be suppressed by adjusting the exposure time, developer time and pre-exposure developing to achieve a vertical edge profile, combined with a moderate to high power sonication in acetone after ion beam milling. Fig. A.1(c) shows a device treated in this way without the feature.

\textit{Other Substrates}

Superconducting coplanar waveguide resonators were also fabricated on magnesium oxide (MgO), c-cut sapphire and silicon. In the case of MgO and c-cut sapphire, the fabrication is identical other than the change in dielectric constant\(^2\) and, consequently, characteristic impedance. Fabrication on silicon is similar although the vacuum pressure at exposure must be reduced to avoid cracking the more brittle substrate.

\(^2\) r-cut sapphire: 9.4; c-cut sapphire: 11.5; magnesium oxide: 9.8
Fig. A.1: Formation of “rabbit ears” in ion beam milling. (a) During ion beam milling ((i) and (ii)), metal is re-sputtered against the walls of the photoresist. This can persist after the resist is removed ((iii) and (iv)). (b) The feature as it appears under a microscope image, for a section of a coplanar track. (c) A section without the feature.
B. GAAS SWITCH FABRICATION

The fabrication process for the GaAs switches in Chapter 7 undergoes a different treatment to the superconducting devices. The material used is the same material as used for quantum dot devices, allowing for ease of integration onto the same chip. In this thesis, the devices are grown by A. C. Gossard at University of California, Santa Barbara, and M. Manfra at Purdue University.

The process is to etch the 2DEG from where it is not required, evaporate and anneal the eutectic ohmic contacts, then evaporate the top metal layer.

B.1 Wafer Preparation

The wafers are diced using a diamond tipped scribing tool. The chips are then cleaned before use as follows:

- Sit in warm (80 degrees) N-Methyl-2-pyrrolidone (NMP) for 15 minutes
- Sonication in NMP
- Sonication in IPA, dry
- Bake at 200 degrees for 5 minutes to remove any residual water
B. GaAs Switch Fabrication

B.2 2DEG Etching

The chip is spin-coated with photoresist (AZ6612, 10000 rpm for 20 seconds, 4000 rpm for 20 seconds) then baked at 95 degrees for 60 seconds. The chip is exposed at 10 mW cm\(^{-2}\) for 1.6 seconds then developed for 40 seconds, rinsed in DIH\(\text{H}_2\)O and dried. Similarly to silicon substrates mentioned in Appendix A, low pressure contact is required to prevent the wafer from cracking.

The gallium arsenide is etched past the 2DEG using a sulfuric acid solution (0.4 \% H\(_2\)SO\(_4\), 3 \% H\(_2\)O\(_2\)), which etches at a rate of roughly 1 nm s\(^{-1}\), then cleaned with DIH\(\text{H}_2\)O and dried. With quantum dot devices, a profilometer measurement after half the 2DEG has been etched can help prevent excessive over-etching, but for switch devices this is not a concern.

B.3 Metal Layers

Metal deposition for both the eutectic alloy and the final metal layer is done via evaporation and then lift-off using a bi-layer resist.

The first photoresist spun on is LOR 5A (10000 rpm for 1 second, 4000 rpm for 60 seconds), which is baked at 170 degrees for 5 minutes, followed by AZ6612 (10000 rpm for 20 seconds, 4000 rpm for 20 seconds), baked at 95 degrees for 60 seconds. Similar to 2DEG etching, the resist is exposed at 10 mWcm\(^{-2}\) for 1.6 seconds and developed for 40 seconds.

The metal layers are evaporated at base pressure \(\sim 3 \times 10^{-6}\) mbar. For the top metal layer (signal tracks and gate lines), 100 nm thick gold is used with a 10 nm titanium adhesion layer. For the ohmic contacts, the following stack is used:

- 5 nm nickel, as an adhesion layer
B. GaAs Switch Fabrication

- 35 nm germanium, followed by
- 72 nm gold, forming the eutectic alloy
- 18 nm nickel
- 50 nm gold as a capping layer

Lift-off is performed in warm (80 degrees) NMP for at least an hour, with low power sonication if required. The ohmic contacts are formed by annealing the chips (before the final metal) at 470 degrees for 110 seconds.

A low vacuum pressure and clean gallium arsenide surface improves adhesion of the metal layers to the substrate. Contamination during evaporation can adversely affect both adhesion and ohmic contact diffusion through the gallium arsenide.
C. CIRCUIT COMPONENTS USED IN SPICE SOFTWARE

C.1 Precise Component Terms

In Chapter 4, we introduced SPICE software with features useful for simulating superconducting circuits. Simple expressions for capacitors, inductors and Josephson junctions were given, and the more complex forms will be detailed here. Nodal Analysis uses Kirchhoff’s Current Law to generate a series of linear equations which can be solved via a matrix inversion. The aim, then, is to find linear functions $f$ such that a component’s contribution to a node can be written as $I = f(V_i)$.

C.1.1 Capacitors

The characteristic equation of a capacitor is

$$I = C \frac{dV}{dt},$$

and a simple approach is to use

$$\frac{dV}{dt} = \frac{V(t) - V(t - \delta t)}{dt}.$$  

As can be seen from Fig. C.1, this expression is not identical to $dV/dt$ at $t$ or $t - \delta t$, although it will be between these two values assuming the concavity does not change over such a short time span. A more precise expression is

$$\frac{V(t) - V(t - \delta t)}{\delta t} = \frac{dV}{dt}(t) + \frac{dV}{dt}(t - \delta t) \frac{\delta t}{2}.$$
Fig. C.1: Illustration for a more precise expression for the gradient $dV/dt$.

Isolating $dV/dt$ and using the characteristic equation above gives us the current contribution for a capacitor:

$$I = \frac{2C}{\delta t} V - \frac{2}{\delta t} V(t - \delta t) + \frac{dV}{dt}(t - \delta t)$$

Storing node potentials and their derivatives for the next $t$ is required, although this is computationally straightforward.

### C.1.2 Inductors

To generate an expression for inductors, we use the same technique as above, noticing that

$$\frac{I(t) - I(t - \delta t)}{\delta t} = \frac{dI}{dt}(t) + \frac{dI}{dt}(t - \delta t).$$

For an inductor, we can use $dI/dt = V/L$ to find the current contribution

$$I = \frac{dt}{2L} V(t) + I(t - \delta t) + \frac{dt}{2L} V(t - \delta t),$$

and we store $I(t)$ for the previous time step by incrementing by $dt/2L(V_t - V_{t-\delta t})$ after each iteration.
The Josephson junction current contribution is found using the resistor and capacitively shunted model (RCSJ), where we use a perfect junction element in parallel with a resistor and a capacitor. The perfect junction element produces an extra node in the circuit corresponding to the phase $\phi$, and two equations:

$$I_s = I_c \sin \phi \quad \frac{d\phi}{dt} = \frac{2e}{\hbar} V$$

To make the phase equation more precise than the expression given in Chapter 4, we use the same approach as above an the second of these two equations:

$$\frac{\phi(t) - \phi(t - \delta t)}{\delta t} = \frac{\frac{d\phi}{dt}(t) + \frac{d\phi}{dt}(t - \delta t)}{2}$$

$$\phi(t) = \phi(t - \delta t) + \frac{e\delta t}{\hbar} (V(t) + V(t + \delta t))$$

The current contribution of the junction element is

$$I = I_c \sin \phi.$$

This is not linear in the $V_i$ (of which $\phi$ is a part), so we estimate what $\phi$ will be for this time interval:

$$\phi(t) = \phi(t - \delta t) + \delta t \frac{d\phi}{dt}$$

$$= \phi(t - \delta t) + \delta t \frac{2e}{\hbar} (V(t) - V(t - \delta t)),$$

where $V(t)$ is first estimated as $V(t) = V(t - \delta t) + \delta t dV/dt(t - \delta t)$.

The current from the resistor and capacitor is as we have described previously:

$$I_C = \frac{2C}{\delta t} V - \frac{2}{\delta t} V(t - \delta t) + \frac{dV}{dt}(t - \delta t)$$

$$I_R = \frac{V}{R}.$$

There is one caveat here concerning the resistance $R$. The quasiparticle current through the junction changes depending on whether the voltage across
the junction is more or less than the gap voltage $V_{gap} = 2\Delta/e$. For implementation, the software uses the same prediction for $V(t)$ described above before determining the appropriate $R$.

C.2 Remarks on Constant $G$ in $G.V = I$

Reformulating the equations $G.V = I$ so that $G$ is constant is not without its complications, and we will mention two here. First, the matrix $G$ is a function of the time step $\delta t$, apparently forcing a constant time step for a simulation. To keep the advantages of a variable $\delta t$ depending on what is occurring in the circuit, matrix inverses for several values can be computed for a given circuit, and the appropriate $G^{-1}$ used depending upon the $\delta t$ chosen.

One term that is difficult to remove from $G$ is the quasiparticle current through a Josephson junction, and consequently the resistance of the junction in the RCSJ model. This resistance takes on different values depending on whether the voltage across the junction is above or below the gap voltage $V_{gap} = 2\Delta/e$.

In most simulations, the junction voltage is less than the subgap voltage for all junctions and most time steps. Consequently, it is possible to recompute $G^{-1}$ in these situations and maintain the speed advantages of a constant $G$. An improvement on this follows from the observation that throughout a simulation, the same junctions tend to be above the gap voltage for many time steps. In this way, a limited number of matrix inverses can be stored, indexed by which junctions have a different quasiparticle current.

The resistance of a junction in our software is determined by whether or not the voltage is above the gap voltage. This is correct for $T = 0$, although for $T \geq T_c/2$, thermally activated quasiparticle conduction is possible when the voltage is lower, and this poses problems for our once-only inversion. We propose that in this case a piece-meal function be used for quasiparticle
conduction so that the junction resistance takes one of a small number of values for $V \sim V_{gap}$. A list of recently used $G^{-1}$, indexed by the junction voltages and with size determined by available hardware and circuit size, could be maintained with $G^{-1}$ being re-computed when a new arrangement is seen, although this has not yet been implemented.
BIBLIOGRAPHY


Bibliography


