Overlapping Aluminum-Gate Quantum Dots for Valley-Orbit Based Qubits in Si/SiGe

By

J.P. Dodson

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Physics)

at the
UNIVERSITY OF WISCONSIN-MADISON
2021

Date of final oral examination: 04/14/2021

The dissertation is approved by the following members of the Final Oral Committee:
Mark A. Eriksson, Professor, Physics
Susan N. Coppersmith, Professor, Physics
Robert F. McDermott, Professor, Physics
Jason Kawasaki, Professor, Materials

Copyright © 2021
by J.P. Dodson

To my family. I am so thankful for their unwavering love and support that has helped get me where I am today.

Abstract

The research effort toward quantum information processing has risen dramatically in the last two decades, and for good reason: quantum information promises a host of new technologies and problem-solving capabilities, ranging from more accurate simulation of quantum systems for research purposes to applications in cryptography, artificial intelligence and medicine. All of this hinges on the ability to create a robust quantum processor—a nearly unrivaled modern technological challenge. By implementing quantum algorithms, quantum computers may outperform classical computers, enabling particular problems to be solved in relevant time scales that otherwise would be intractable.

There are many different approaches to building a quantum processor. The largest and most robust quantum processors have been built using superconducting qubits and trapped ions. A rivaling technology which is slightly behind the curve, but has a big upside, is the topic of this thesis: semiconductor quantum dot qubits. Even within the narrow topic of semiconductor quantum dots there are many different implementations, including silicon-metal-oxide-semiconductor (Si-MOS), Si/SiGe, GaAs and donor systems. Furthermore, within each of these unique material approaches, there exist several different types of qubit architectures that can be implemented, all with fundamentally different degrees of freedom for forming the logical basis for a qubit

($|0\rangle$ & $|1\rangle$). In this thesis, we focus on valley-orbit-based qubits in Si/SiGe.

Variability in valley-orbit state splittings is a problem for large-scale implementations of silicon-based quantum processors that must be addressed. Depending on the particular qubit architecture, the role of valley-orbit states varies; however, a common theme for nearly all silicon-based qubits is that valley-orbit splittings must be precisely engineered or have large in situ tunability. In particular, valley-orbit-based qubit energies are defined by the valley-orbit state splitting in Si/SiGe quantum dots, making it especially important for these splittings to be well engineered for this approach in Si/SiGe quantum dots.

In this thesis, we demonstrate an improved device gate architecture for in situ tunability of valley-orbit splittings. First, we identify yield limiting mechanisms during the fabrication of devices and present an improved fabrication process flow. Then, we show the quantitative relationship between low-lying valley-orbit states and identify a robust in situ tunability technique. Finally, by exploiting the dependence of valley-orbit state splittings on electrostatic confinement and electron number, we show progress toward single-shot readout in the (4,1)-(3,2) electron regime, allowing for in situ tunability of the qubit frequency and enhancement of the readout window.

Contents

1 Semiconductor quantum dots **1**
1.1 Introduction 1
1.2 Semiconductor quantum dots overview 5
1.3 Outline of thesis 11

2 Quantum dots in Si/SiGe **15**
2.1 A brief theoretical introduction 15
2.2 Electrostatic characterization of devices 19
2.3 Valley-orbit states 28

3 Device Fabrication **41**
3.1 Device architectures overview 43
3.2 Fabrication methods 49
3.3 Yield engineering of overlapping aluminum-gate architecture 73
3.4 Conclusion 87

4 QDscript **91**
4.1 Introduction 91
4.2 Package structure 92
4.3 Automated tuning 98
4.4 Future work 99

5 One- and two-electron valley-orbit states **101**
5.1 How valley-orbit states probe quantum well interfaces 101
5.2 Supplemental material 113
5.3 Extended discussion 123

6 Outlook: Si/SiGe quantum dots **131**

A Fabrication recipes **133**
A.1 Photolithography 133

A.2 Nanolithography 137

Bibliography **141**

Chapter 1

Semiconductor quantum dots

1.1 Introduction

The use of computation as a means of studying processes and performing tasks has ushered in a new age of technological advancements. After the invention of the transistor in 1947 at Bell Laboratories by Shockley, Bardeen, and Brattain, digital computing progressed quickly in the following decades. Modern classical computers now take many forms, including powerful supercomputers that can perform computations at up to 442 PFLOPS [1].

While classical computers have changed nearly every facet of our world's technology, quantum computers have the potential to solve particular problems that are intractable for classical computers—even the most powerful modern supercomputers. In fact, the first demonstration of a quantum processor outperforming a supercomputer was made very recently [2]. With hundreds of already known quantum algorithms—collectively known as the "Quantum Algorithm Zoo" [3]—it is difficult to say exactly how useful

quantum computers will ultimately be. However, as many of these algorithms already have proposed uses in simulation of quantum systems, search and optimization, cryptography, and solving large systems of linear equations, there is a high level of interest in advancing the technology quickly to utilize this new type of computing.

The idea behind quantum computing first came from Richard Feynman's seminal work in 1982 [4], which proposed using quantum bits (qubits) to compute. Qubits operate in a fundamentally different way than classical bits, which comprise modern classical processors. A classical bit can be in one of two states: 0 or 1. For a CMOS transistor, these states physically correspond to macroscopic parameters such as $0 \rightarrow 0\,\mathrm{V}$ and $1 \rightarrow 5\,\mathrm{V}$. Feynman's proposal was to instead encode information in discrete energetic states. These states would form a two-level quantum mechanical system and would be the quantum analog of the classical bit.

The fundamental difference between quantum bits and classical bits is due to the principle of superposition. While each contains a set of basis states 0 and 1, a qubit can be in a linear superposition of both states whereas a classical bit must be in one or the other. The qubit state vector can then be described by the state $|\Psi\rangle$

$$|\Psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{1.1}$$

subject to the normalization condition

$$\alpha^2 + \beta^2 = 1 \tag{1.2}$$

where α and β describe the amplitude of the state $|\Psi\rangle$ in each of the logical basis

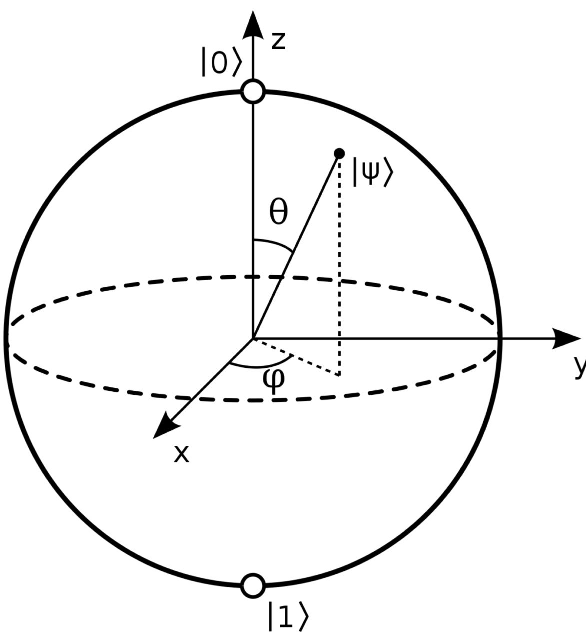

Figure 1.1: Bloch sphere representation of a qubit. Basis states $|0\rangle$ and $|1\rangle$ are at the $\pm z$ poles of the Bloch sphere, with angles θ and ϕ denoting the amplitude and phase of the qubit state $|\Psi\rangle$.

states [5]. It is not possible for a classical bit to simultaneously be in both basis states at the same time. One may naturally ask "what happens when the CMOS transistor voltage is at 2.5 V, exactly between the high and low values?" In this case, a particular threshold would be put in place. For the sake of argument, if that threshold was at 2.5 V, any voltage measured at 2.5 V and above would be categorized as a logical 1, and anything below would be a logical 0. There is no in-between; thus the principle of superposition is unique to quantum computing.

While the notation above is mathematically useful, it is often helpful to represent a qubit state as a vector in space. By setting $\alpha = \cos(\theta/2)$ and $\beta = e^{i\phi}\sin(\theta/2)$, Eqn 1.1 becomes

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle \tag{1.3}$$

This state can then be shown as a vector on a sphere, known as the Bloch sphere. Fig. 1.1 shows the Bloch sphere representation of a qubit. Basis states $|0\rangle$ and $|1\rangle$ are at the $\pm z$ poles of the Bloch sphere, with angles θ and ϕ denoting the amplitude and phase of the qubit state $|\Psi\rangle$.

Another key difference between classical bits and quantum bits is that the information contained in the phase and amplitude of the qubit state collapses to a single state when measured. For example, if the state given in Eqn. 1.1 was measured, $|0\rangle$ would be measured with probability $|\alpha|^2$ and $|1\rangle$ would be measured with probability $|\beta|^2$; however all information about the amplitudes α and β would be lost.

The power of quantum computation becomes apparent when considering multiple qubits. For two classical bits, there are four basis states: 00, 01, 10, and 11. This is also true for two qubits, whose four basis states are $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$. The difference between classical computing and quantum computing can be understood from the concept of quantum entanglement. While a classical two-bit processor must be in one of the four basis states at a given moment, a two-bit quantum processor can be in all four states *simultaneously.* This is represented below in Eqn 1.4

$$\Psi = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \tag{1.4}$$

A classical two-bit processor could be described as containing two pieces of information: one describing the state of the first bit and the second describing the state of the second

bit. Additionally, each of these coefficients can only take on values of 0 or 1. For a two-bit quantum processor, we see that there are four pieces of information describing the state: α_{00}, α_{01}, α_{10} and α_{11}. Additionally, each of these coefficients is a complex number and is not restricted to integers 0 and 1 like classical bits. As the number of qubits composing a quantum process increases, the amount of extra information available in the computational basis compared to a classical processor grows rapidly. For a classical processor, there are N coefficients describing the state of a processor with N bits, but for a quantum processor there are 2^N coefficients for a quantum processor with N qubits. It remains a challenge of quantum computing to utilize this extra information available via quantum algorithms.

1.2 Semiconductor quantum dots overview

There are many different physical realizations of quantum processors. Some of the leading candidates include trapped ion qubits, which have been shown to have the highest qubit fidelities of any platform [6], superconducting qubits, where large 2D arrays of qubits have been demonstrated with high single and two-qubit gate fidelities [2], semiconductor quantum dots, and many others not mentioned here. Semiconductor quantum dots will be the focus of this thesis, where, in particular, the quantum dots are formed in silicon.

Semiconductor quantum dots began making significant experimental progress in the early 2000s, shortly after the Loss and DiVincenzo proposal in 1998 to use the spin states of electrons in semiconductor quantum dots as a logical basis for qubits [7]. Semiconductor quantum dots are a very flexible platform for qubit design since the logical basis of a qubit can be encoded in several distinct degrees of freedom. Fig. 1.2

summarizes the main approaches for encoding a qubit in quantum dots. For each of the five different types of qubits listed, the logical basis is shown schematically with yellow circles representing electrons and arrows indicating the direction of the electron spin. The charge qubit is a single electron combined with two quantum dots where the logical basis is defined by which dot the electron is currently occupying. This means that the charge qubit can be operated via all electrical control because gate voltages are used to control the electron occupation of each dot. All electrical control of qubits is an advantage since integrating magnetic fields into a semiconductor quantum dot platform and addressing multiple qubits is difficult. Although the charge qubit can be quickly manipulated purely by gate voltages, it decays and dephases rapidly, making it a sub-optimal qubit candidate. The fast decay time, also known as the T_1 time, is due to the electron-phonon coupling. The dephasing time, or T_2 time of the charge qubit, is short due to charge noise.

Charge noise is generally the main source of noise that decoheres qubits in semiconductor quantum dots. The origin of charge noise has been investigated thoroughly and is understood to be charge traps that arise within dielectrics and material interfaces. Suppressing charge noise is difficult because it involves incredibly precise engineering of material stacks, so oftentimes the best way to suppress the effects of charge noise is to design qubits that are insensitive to it. This approach has worked extremely successfully for superconducting qubits, from which the transmon became the cornerstone of qubit architectures [8]. Just as superconducting qubits designed the transmon, there are encoded qubits in quantum dots that are highly insensitive to charge noise. Fig. 1.2 shows several examples of these. The single-spin qubit is the simplest, where a single electron is in one quantum dot and the logical basis is encoded as the spin state of the electron. The advantage of single-spin qubits lies in their simplicity. Since it is

just a single electron in one dot, the device fabrication overhead per qubit is much smaller than other architectures. Additionally, encoding the qubit basis in the spin degree of freedom makes the qubit highly insensitive to charge noise and less susceptible to relaxation. As mentioned earlier though, addressing single-spin qubits using all electrical control is not possible, making global magnetic fields and/or integrated micromagnets a necessity. This poses a significant challenge in scaling up to larger arrays of qubits. Furthermore, two-qubit gates are generally not as insensitive to charge noise as one-qubit gates.

Other types of qubits bypass the magnetic field integration by using clever architectures combining spatial, spin, and charge degrees of freedom. The hybrid qubit, shown schematically in Fig. 1.2, does not need a global magnetic field or micromagnets, since the qubit splitting is determined by the intrinsic singlet-triplet splitting of one of the quantum dots. The hybrid qubit—a particular implementation of a valley-orbit qubit—requires three electrons in two dots where a pair of electrons is in either a singlet or triplet state. An additional benefit of the hybrid qubit is it combines spin and charge degrees of freedom such that fast all electrical control is possible while maintaining long relaxation and coherence times. The main challenge for hybrid qubits is precisely engineering the valley splitting, which directly corresponds to the singlet-triplet splitting/qubit energy. The valley splitting is difficult to control for reasons discussed in Chapter 2.

Finally, the exchange-only qubit uses three electrons across three dots to encode the logical basis. The exchange-only qubit is a very promising candidate for many reasons. One and two-qubit gates can be operated via the exchange interaction, which can be turned on and off using all electrical control; thus magnetic fields and micromagnets are not necessary. The exchange-only qubit contains symmetric operating points that

make it highly insensitive to charge noise. Additionally, baseband pulse-sequences with modest bandwidth may be used instead of high frequency resonant gating, potentially reducing the stringent demands high frequency wiring places on experimental setups. While there are many reasons to make exchange-only qubits, they suffer from their complex pulse schemes and large fabrication overhead per qubit.

Since there are so many different qubit architectures and material systems that can be used for semiconductor quantum information, navigating the literature and progress throughout the last two decades is challenging. In 2000, the Loss and DiVincenzo criteria were proposed as a means to help guide and gauge progress in quantum computing. The five principles that are necessary for quantum computation are:

1. A physical system that is scalable
2. Robust initialization of a qubit into a fiducial state
3. Long decoherence and relaxation times
4. A universal set of quantum gates (i.e., two-axis control of bloch sphere for a single qubit, and two-qubit gates)
5. High fidelity qubit readout capability

Some of the first major milestones in semiconductor dots were accomplished in GaAs/AlGaAs heterostructure. Many important works demonstrated key elements that help satisfy the first DiVincenzo criteria [9–14]. These included observation of photon assisted tunneling in quantum dots, allowed and forbidden states in one- and two-electron quantum dots, integrated charge readout of quantum dots using a quantum dot as a charge sensor, excited-state spectroscopy, single-electron manipulation in a double quantum dot, and measurement of the singlet-triplet relaxation times. Another big achievement was single-shot readout of an electron spin in 2004 [15], making spin-qubits a promising candidate due to the long relaxation times.

The first demonstrations of qubits in GaAs were singlet-triplet and single-spin qubits [16, 17], in 2005 and 2006, respectively. The single-spin qubit experiment performed coherent rotations using a resonant magnetic field, commonly referred to as electron spin resonance (ESR), but in 2007 an all electrically controlled single-spin qubit was shown [18]. This was an important development as it allowed for different types of architectures to be implemented for single-spin qubits and eventually led to an experiment where a single-spin qubit was manipulated using electric dipole spin resonance (EDSR) [19]. Major advancements in single-shot readout capabilities using a radiofrequency quantum point contact (RF-QPC) were demonstrated in 2007 [20] and using dispersive gate readout in 2013 [21]. Over this time period, it was shown that each of the five DiVincenzo criteria were possible in quantum dots, so it became a matter of improving fidelities to fault tolerant levels [22].

Quantum dots in GaAs/AlGaAs heterostructure encountered a fundamental limitation though. The presence of nuclear spins creates magnetic noise which couples directly to spin qubits, limiting coherence times. This makes it difficult to achieve the fidelities necessary for a fault-tolerant quantum processor. Silicon remedies this because the nucleus of the ^{28}Si isotope is spinless. ^{28}Si is naturally abundant, comprising 92.3% of naturally occurring silicon. There also exist isotopic purification techniques where virtually all other isotopes can be removed from the ^{28}Si substrate. This platform is highly attractive for spin-based qubits since magnetic noise due to nuclear spin in isotopically purified ^{28}Si is negligible.

Additionally, quantum dots can be formed in Si/SiGe heterostructure instead of just silicon (the latter often referred to as silicon metal-oxide semiconductor, or Si-MOS). Qubits in both platforms have progressed in parallel, each taking significant steps toward developing robust quantum processors in the last decade. 2012 marks the

first time coherent oscillations of a silicon-based quantum dot qubit was successfully demonstrated [23]. This experiment was conducted using a double dot structure with two electrons, forming a singlet-triplet qubit in Si/SiGe. The next realization was the quantum dot hybrid qubit (QDHQ) in 2014 [24, 25], showing that encoded qubits using both charge and spin degrees of freedom enabled fast gating with long coherence times. The coherence time of this qubit was extended to be over 100 ns in 2017 [26], further improving the prospect of the QDHQ. Many other types of qubits were demonstrated around the same time, including single-spin, singlet-triplet, and charge qubits [27–30].

Silicon as a platform for quantum computing really began gaining momentum in 2014 when a single-spin qubit was shown to have gate fidelities exceeding 99% in Si-MOS [31], as for the first time, silicon demonstrated gate fidelities exceeding the fault-tolerant threshold. In 2018, a similar architecture was used in Si/SiGe to obtain gate fidelities above 99.9% in isotopically purified silicon [32]. Two-qubit experiments followed in both Si-MOS and Si/SiGe [33–36], although in each of these cases, the two-qubit gate fidelity was below fault-tolerant thresholds, reaching a peak gate fidelity of 98%. Thus, one of the outstanding tasks for silicon quantum dots is to design a two-qubit architecture that can exceed fault-tolerant thresholds and scale to larger arrays.

An emerging approach for achieving this is the exchange-only qubit. Recently (2019), a single exchange-only qubit was shown to have gate fidelities exceeding 99% [37], but in addition to this, two-qubit operations use the exact same physical interaction (exchange interaction) as single-qubits. Although two-qubit gate operations take longer than one-qubit gate operations, this is promising for achieving fault-tolerant two-qubit gate fidelities. Additionally, in this experiment, state preparation and measurement (SPAM) errors were demonstrated to be <0.8%. Gate fidelities are not the

only factor determining the effectiveness of a quantum processor. As the DiVincenzo criteria states, achieving high fidelity initialization and readout is just as important. Several different high bandwidth readout schemes have been implemented in silicon to achieve high readout fidelities, including cryogenic amplification using a heterojunction-bipolar transistor (HBN) or high electron mobility transistor (HEMT), dispersive gate readout, and the RF-QPC [38–43].

Many challenges have been overcome since the first silicon quantum dot qubit demonstration in 2012. The main focus of the field is on further development of qubit architectures that are suitable for larger arrays of qubits with fault-tolerant performance. While there has been some headway, such as four qubit experiments [44] and 9-dot arrays [45], there are still many fundamental challenges for the future of silicon quantum dots, including robust fabrication processes, repeatable high fidelity fabrication, and control of qubits and fundamental research into physical processes affecting the performance of silicon qubits. Steady progress is still being made, making silicon quantum dots a strong candidate for quantum processors in the future.

1.3 Outline of thesis

In Chapter 2, a brief theoretical overview of Si/SiGe heterostructure and Si/SiGe quantum dots is given to motivate the main results shown in Chapters 3–5. Additionally, quantum dot devices fabricated using an overlapping aluminum gate architecture are shown and example measurements characterizing devices that are typically performed are used to help explain the experimental principles behind tuning dot devices into the qubit regime. This includes gate electrode performance characterization, sensor dot formation, screening gate optimization, and experimental techniques used to detect

the electron occupation of a triple-quantum dot array.

Next, Chapter 3 provides an overview of device architectures that can be used for quantum dots in silicon with an emphasis on overlapping aluminum gate architectures. The process for fabricating these devices is discussed in detail at a photolithography and electron-beam lithography level. A yield engineering study is performed on overlapping aluminum gate devices where several yield limiting steps are investigated. The native aluminum oxide that electrically isolates aluminum gate layers is characterized. A proposal for mitigating gate-to-gate leakage is put forth using a UV-ozone or plasma ash process to enhance the thickness of the inter-gate oxide. The morphology of overlapping gate structures is also analyzed using transmission electron microscopy (TEM). Critical dimensions in the geometry of gate designs are identified, such as the width and thickness of barrier gates. Finally, the interconnect structure from bond pad to active region is discussed, where in particular, the thermal budget of devices is considered. An optimized interconnect fabrication process is discussed to optimize yield of devices.

Chapter 4 shows progress development of a quantum dot measurement environment called QDscript. QDscript is a package written in Python that hooks into a data acquisition platform named Labber. QDscript allows for flexible script-based measurements in Python. This helps when tuning quantum dot devices since the large amount of gate electrodes necessary for quantum dot devices makes for complex measurement environments. Tuning algorithms and integrated analysis tools in QDscript have improved measurement efficiency and also aided in machine learning auto-tuning of a double dot device [46].

In Chapter 5, the relationship between one- and two-electron valley-orbit states are studied. Pulsed-gate spectroscopy and magnetospectroscopy are used to probe these

states as a function of electrostatic confinement. This led to the measurement of the interplay between the valley, singlet-triplet, and orbital splitting in a Si/SiGe quantum dot. Theoretical simulations using a tight binding approach with full configuration interaction calculations are found to be in agreement with the experimental data and help explain the physical origin of the relationship between these valley-orbit states. Electron-electron interactions and the atomic details of the quantum well interface fully explain the phenomenon observed in the experimental data.

Finally, in Chapter 6, the main results of this thesis are summarized. The projects discussed in this thesis have led to important developments for overlapping aluminum gate devices and valley-orbit qubits, and help form a better understanding of the future direction of device geometries/gate stacks and valley-orbit qubit implementations. Proposals to improve device yield and valley-orbit qubit performance are made as a conclusion to this thesis.

Qubit Type	Logical basis	Main advantage	Main challenge
Charge	$\|0\rangle$ $\|1\rangle$	All electrical control of 1 and 2 qubit gates	Rapid dephasing due to charge noise
Single-Spin	$\|0\rangle$ $\|1\rangle$	Simple spin encoding that is insensitive to charge noise	Scaling and integration of micromagnets Charge noise sensitive two-qubit operation
Singlet-Triplet	$\|0\rangle$ $\|S_0\rangle$ $\|1\rangle$ $\|T_0\rangle$	State-dependent dipole moment	Charge noise during two-qubit operations
Hybrid	$\|0\rangle$ $\|S_0\rangle$ $\|1\rangle$ $\|T_0\rangle$ – $\|T_-\rangle$	Charge noise insensitive qubit operation	Precise valley splitting engineering
Exchange-Only	$\|0\rangle$ $\|S_0\rangle$ $\|1\rangle$ $\|T_0\rangle$	Universal qubit control via exchange interaction	Complex pulse sequence for multi-qubit gates

Figure 1.2: Encoded qubit architectures for semiconductor quantum dots. For each encoded qubit, the logical basis is shown schematically with yellow circles representing electrons and arrows indicating the direction of the electron spin. The charge qubit is a single electron combined with two quantum dots where the logical basis is defined by which dot it is currently occupying. The single-spin qubit is a single electron in one quantum dot where the logical basis is given as the direction of the spin. The singlet-triplet qubit is formed by two-electrons in a singlet or triplet state across two dots. The hybrid qubit is more involved, requiring three electrons in two dots where a pair of electrons is in either a singlet or triplet state. Finally, the exchange-only qubit utilizes three electrons across three dots, where the two electrons in the left-most two dots form either a singlet or triplet state.

Chapter 2

Quantum dots in Si/SiGe

2.1 A brief theoretical introduction

Quantum dots are physical systems with tight confinement in all three spatial dimensions. For quantum dots in Si/SiGe heterostructure, electrons are confined to be in a two-dimensional plane—known as a two-dimensional electron gas (2DEG)—by using two heterojunctions which form a quantum well. The conduction band energy inside the quantum well of Si/SiGe heterostructure is lower than outside, confining electrons in the quantum well along the growth direction.

This is illustrated in Fig. 2.1, where the material stack for a typical Si/SiGe quantum dot device is shown. First, a graded SiGe buffer layer is grown on silicon where the ratio of germanium to silicon concentration increases to approximately $Si_{0.7}Ge_{0.3}$. The silicon well is grown after the SiGe buffer layer, which is where electrons are confined in a two-dimensional electron gas (2DEG). The 2DEG is located just below the SiGe spacer interface. The SiGe spacer is subsequently grown above the silicon quantum well. Finally, the heterostructure is typically capped with 1–2 nm of silicon to make

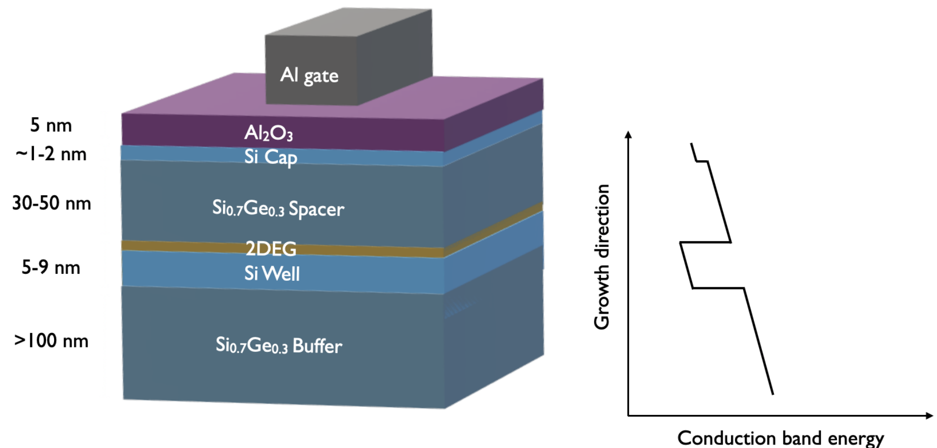

Figure 2.1: Typical material stack for Si/SiGe quantum dot device. First, a graded SiGe buffer layer is grown on silicon where the ratio of germanium to silicon concentration increases to approximately $Si_{0.7}Ge_{0.3}$. The silicon well is grown after the SiGe buffer layer, which is where electrons are confined in the two-dimensional electron gas (2DEG). The 2DEG is located just below the SiGe spacer interface. The SiGe spacer is subsequently grown above the silicon quantum well. Finally, the heterostructure is typically capped with 1–2 nm of silicon to make fabrication easier. The conduction band energy of the heterostructure is shown adjacent to the 3D schematic. Silicon's conduction band is lower in energy than germanium, causing the conduction band to dip where silicon is present in the growth. The slope of the conduction band is due to an assumed non-zero gate voltage on the metallic gate shown above the heterostructure which is used to define quantum dots. The combination of the quantum well and electric fields from gate electrodes causes a sharp dip in the conduction band energy at the interface between the Si and SiGe spacer, forming the 2DEG.

fabrication easier.

The conduction band energy of the heterostructure is shown adjacent to the 3D schematic in Fig. 2.1. Silicon's conduction band is lower in energy than germanium, causing the conduction band to dip where silicon is present in the growth. The slope of the conduction band is due to an assumed non-zero gate voltage on the metallic gate shown above the heterostructure which is used to define quantum dots. The combination of the quantum well and electric fields from gate electrodes causes a sharp minimum in the conduction band energy at the interface between the Si and SiGe

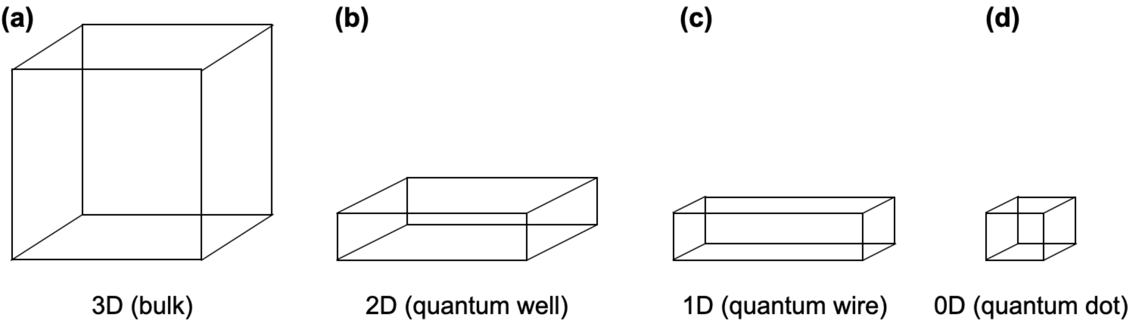

Figure 2.2: Confined systems in semiconductors. (a) In bulk semiconductor, charge carriers may freely move in all three dimensions. (b) In quantum wells, charge carriers are confined to a plane such that there is one-dimension of tight confinement. This is a similar type of system to the one Si/SiGe heterostructure forms with electrons. (c) A quantum wire contains two-dimensional confinement where charge carriers may freely move in a single direction. (d) 0D systems, such as quantum dots, form when tight confinement in 3-dimensions is present.

spacer, forming the 2DEG. Thus, tight confinement in the growth direction is achieved (typically referred to as the z-dimension).

These material systems are further confined by patterning particular geometries of metallic gate electrodes. Gates are then used to selectively accumulate and deplete particular regions in the 2DEG. This creates electrostatic confinement in the x- and y-dimensions of the 2DEG. Combined with the z-dimension confinement from the heterostructure, a "0-dimensional" system is formed. The term quantum dot is a convenient description of these systems since a physical system resembling a point in three-dimensional space is formed.

While not a true 0D system (the electron wavelength is confined to be on a ∼nms scale in all three dimensions), we can begin to understand the physical characteristics of this system by considering this case. Fig. 2.2 shows four distinct systems that can arise in semiconductors. In the bulk case (Fig. 2.2(a)), charge carriers move along the

conduction band in all three dimensions of a semiconductor. The wavefunction of a particle for such a bulk system is given by

$$\Psi(x, y, z) = A \sin(k_x x) \sin(k_y y) \sin(k_z z) \tag{2.1}$$

$$k_{x,y,z} = \frac{n\pi}{L}, \qquad n = 1, 2, 3, ... \tag{2.2}$$

if the 3D box has side length L. The dispersion relation of a free particle in a semiconductor is given by

$$E(k) = \frac{\hbar^2 k^2}{2m^*} \tag{2.3}$$

where $\hbar$ is Planck's constant and m^* is the effective mass. From the dispersion relation, we can calculate what is known as the density of states (DOS). The DOS is useful when considering the density of available energy levels for a system within a given volume. Although the wavenumber k takes on discrete values, the number of available energetic states can be calculated by considering a sphere of radius k in k-space. Then, the number of states is given by

$$N = 2 \cdot \frac{1}{8} \left(\frac{L}{\pi}\right)^3 \frac{4\pi k^3}{3} \tag{2.4}$$

where the factor of 2 is due to the spin degeneracy of an electron and the factor of 1/8 arises because n can only take on positive values, constraining the volume of the sphere to positive $k_{x,y,z}$ integers. The DOS per unit volume is then calculated as

$$\text{3D:} \qquad g(E) = \frac{dN}{dE} = \frac{\sqrt{2}}{\pi^2 \hbar^3} m^{*3/2} \sqrt{E} \tag{2.5}$$

In lower dimensional systems, the dependence of the density of states on energy changes. As discussed earlier, Si/SiGe heterostructure confines electrons in a 2DEG, effectively creating a 2-dimensional system. Other systems, such as semiconductor nanowires [47], confine electrons in a single spatial dimension. The DOS for each of these systems can be calculated in a similar way as the bulk DOS above, yielding

$$\begin{aligned} \text{2D:} \qquad g(E) &= \frac{m^*}{\pi\hbar^2} \\ \text{1D:} \qquad g(E) &= \frac{1}{\pi\hbar}\sqrt{\frac{m^*}{2E}} \end{aligned}$$

The final extension to quantum dots is confinement in all spatial dimensions. At this point, the electron can no longer move freely in any direction and a discrete energy level spectrum arises. The density of states is given as

$$\text{0D:} \qquad g(E) = 2\sum_i \delta(E - E_i) \tag{2.6}$$

This discrete energy spectrum allows for a physical system that can be used for quantum information. In the following section, experimental characterization of devices is shown, where this discrete level structure is probed and controlled into the few-electron regime.

2.2 Electrostatic characterization of devices

In this section, measurements of a triple-quantum dot device are presented as a way of describing the important characteristics and tuning techniques that are necessary in

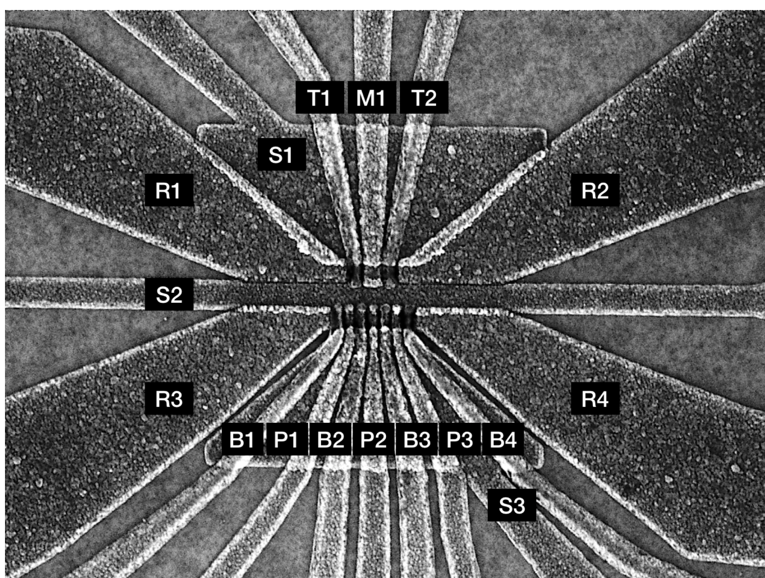

Figure 2.3: Scanning electron microscope (SEM) image of triple-quantum dot device. Three dots are formed beneath plunger gates P1–P3. The charge occupation of these three dots is detected by an integrated charge sensing dot beneath gate M1. Tunnel barrier gates B1–B4 control the tunnel coupling between dots and reservoirs. Screening gates help screen electric fields from gate layers 2 and 3, and help confine quantum dots along the y-axis.

forming robust quantum dots for qubit operation. This includes charge sensor tune-up, qubit-operational quantum dot tune-up, optimization of screening gate voltages, charge noise characterization, and electrostatic characterization of the device as a whole.

Fig. 2.3 shows a scanning electron microscope (SEM) image of a triple-quantum dot device. A more complete description of the device can be found in Chapter 3. Three dots are formed beneath plunger gates P1–P3. The charge occupation of these three dots is detected by an integrated charge sensing dot beneath gate M1. Tunnel barrier gates B1–B4 control the tunnel coupling between dots and reservoirs. Screening gates help screen electric fields from gate layers 2 and 3, and help confine quantum dots along the y-axis. As discussed in the previous section, discrete states form in quantum dots due to strong confinement in 3-dimensions. While the heterostructure confines

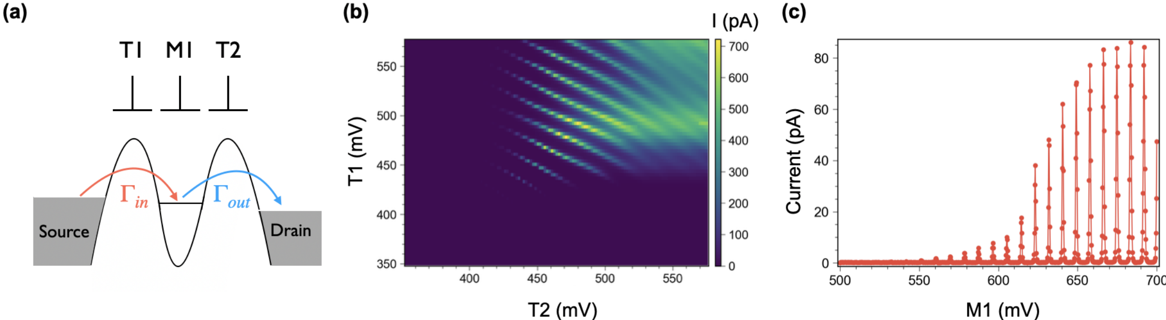

Figure 2.4: Transport through a single quantum dot. (a) A discrete spectrum of energy levels arises in a tightly confined quantum dot. Gate electrode M1 controls the chemical potential of the dot. When an energetic state lies between the fermi energy of the source and drain leads, single-electron transport may proceed. The tunnel rates into and out of the dot are controlled by gates T1 and T2, setting tunnel rates Γ_{in} and Γ_{out}. (b) If the tunnel rates are high enough, current through the dot may be observed. This type of plot is named a "wall-wall" plot because at low T1 and T2 voltages, two "walls" appear where current is blocked. Above these values, fast tunneling occurs, resulting in measurable current. (c) Coulomb blockade where the chemical potential of the dot is modified by M1. As the gate voltage increases, the chemical potential of the dot is lowered, causing several discrete states to pass through the source-drain bias energy window.

electrons in the z-dimension, the geometry of gate electrodes shown in Fig. 2.3 has the ability to confine four separate quantum dots. This is possible because of the confining effect electrostatic fields from the gate electrodes have on conduction electrons within the 2DEG.

Fig. 2.4 shows an example of transport through a single quantum dot. In Fig. 2.4(a), a discrete spectrum of energy levels arises in a tightly confined quantum dot. Gate electrode M1 controls the chemical potential of the dot, which sets the energy of the available states with respect to the source and drain leads. If the energy of a state is between the source and drain chemical potentials, an electron may tunnel onto the dot from the source, and subsequently off to the drain. While there may be many electrons "bound" to the quantum dot at lower energy levels, only a single

electron may participate in transport at a time, i.e., two electrons would not be able to simultaneously tunnel onto the dot. This is forbidden due to Coulomb interactions. The tunnel rates into and out of the dot are controlled by gates T1 and T2, setting tunnel rates Γ_{in} and Γ_{out}. In Fig. 2.4(b), current through the quantum dot is observed when barrier gate voltages T1 and T2 are high. This type of plot is named a "wall-wall" plot because at low T1 and T2 voltages, two "walls" appear where current is blocked. Above these values, fast tunneling occurs, resulting in measurable current. In Fig. 2.4(c), Coulomb blockade is observed, where the chemical potential of the dot is modified by M1. As the gate voltage increases, the chemical potential of the dot is lowered, causing several discrete states to pass through the source-drain bias energy window.

For devices with screening gates, an additional layer of tuning complexity is added. Fig 2.5 shows how to optimize screening gate voltages. In Fig. 2.5(a) current is measured between source-drain contacts beneath reservoir gates R1 and R2 as the voltages on screening gates S1 and S2 are swept, resulting in a "triangle plot". At low S2 voltages and high S1 voltages, a vertical wall of current forms, where current passes directly beneath the S1 gate. For low S1 voltages and high S2 voltage, the same phenomenon occurs but beneath the S2 gate. Between these two walls lies a triangular region of current. A slope of 45° marks the region where the current path is equally affected by both S1 and S2. This region is where current passes through the channel beneath finger gates T1, M1 and T2. Fig. 2.5(b) shows a triangle plot for S2 and S3 on the bottom side of the device. For each plot, gate voltages for the screening gate that lie within the triangular region always produce quantum dots when finger gates are tuned to the correct voltages.

After optimizing screening gate voltages, the sensor dot can be used to detect the

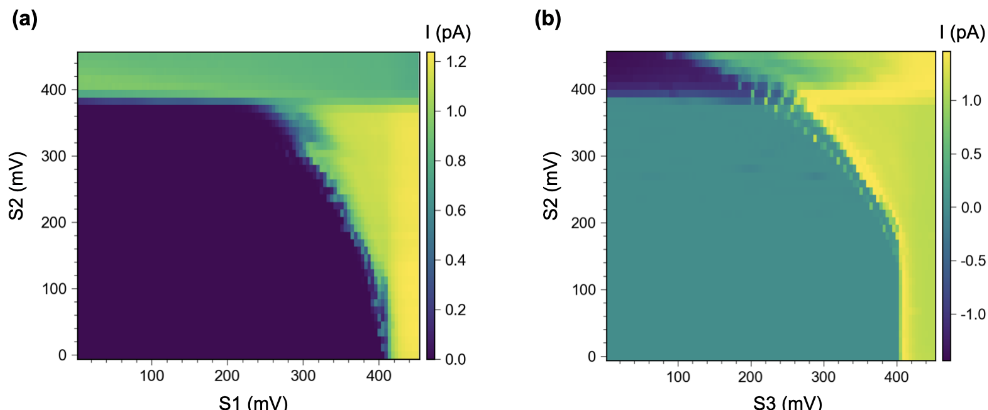

Figure 2.5: "Triangle plot" for screening gate optimization. (a) Current is measured between source-drain contacts beneath reservoir gates R1 and R2 as the voltages on screening gates S1 and S2 are swept. At low S2 voltages and high S1 voltages, a vertical wall of current forms, where current passes directly beneath the S1 gate. For low S1 voltages and high S2 voltage, the same phenomenon occurs but beneath the S2 gate. Between these two walls lies a triangular region of current. A slope of 45° marks the region where the current path is equally affected by both S1 and S2. This region is where current passes through the channel beneath finger gates T1, M1 and T2. (b) Triangle plot for S2 and S3 on the bottom side of the device. For each plot, gate voltages for the screening gate that lie within the triangular region always produce quantum dots when finger gates are tuned to the correct voltages.

electron occupation of quantum dots P1–P3. This is achieved using the M1 sensor dot. The gate voltage of M1 is set such that the current through the sensor dot is highly sensitive to changes in the local electrostatic potential. Since the current through the M1 sensor dot is directly capacitively coupled to the M1 gate voltage, cross-capacitances from nearby quantum dots and gate electrodes can induce measurable changes in the current through the sensor dot.

This is shown in Fig. 2.6. In Fig. 2.6(a), current through sensor dot M1 is measured as the voltage on gate P2 is increased. Sharp changes in current mark electron loading events into the P2 dot. Due to the cross-capacitance between the P2 dot and M1 sensor dot, the change in local electrostatic potential caused by an electron loading onto the

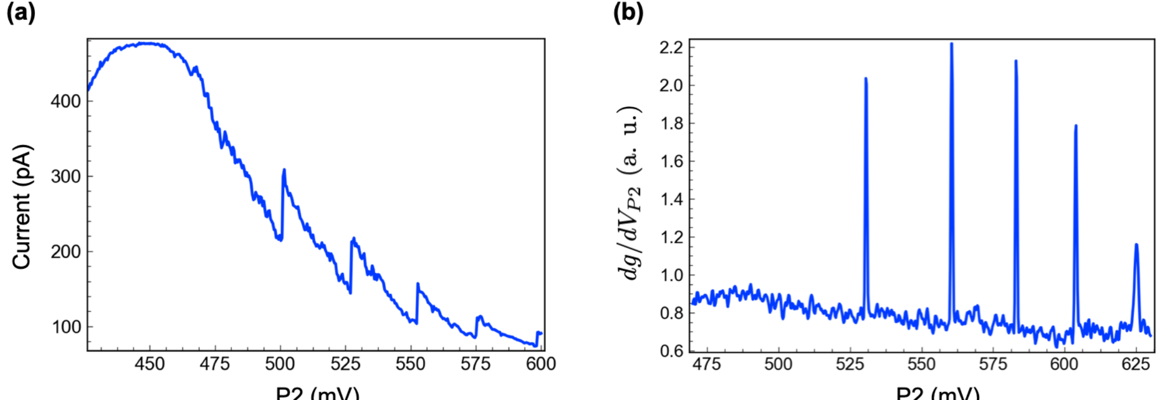

Figure 2.6: Charge sensing using the integrated sensor dot. (a) Current through sensor dot M1 is measured as the voltage on gate P2 is increased. Sharp changes in current mark electron loading events into the P2 dot. Due to the cross-capacitance between the P2 dot and M1 sensor dot, the changes in the local electrostatic potential caused by an electron loading onto the dot significantly shifts the current through the sensor dot. (b) A similar measurement is taken where instead a small sinusoidal signal is applied to the P2 gate in addition to the DC voltage. A lock-in amplifier is used to detect the signal at that frequency, effectively differentiating the DC current measurement shown in (a).

dot significantly shifts the current through the sensor dot. In this way, the number of electrons on a dot can be counted. The change in background of the sensor dot current is due to the cross-capacitance of the P2 gate electrode, whereas the sharp changes are due to the change in capacitance of the P2 quantum dot from electron loading events. In Fig. 2.6(b) a similar measurement is taken, where instead of measuring the current through the sensor dot, a small sinusoidal signal is applied to the P2 gate in addition to the DC voltage. A lock-in amplifier is used to detect the signal at that frequency, effectively differentiating the DC current measurement shown in Fig. 2.6(a). This often results in a better signal-to-noise (SNR) ratio since a lock-in measurement filters noise outside of the designated measurement frequency.

The next stage of tuning involves counting the number of electrons in a single or

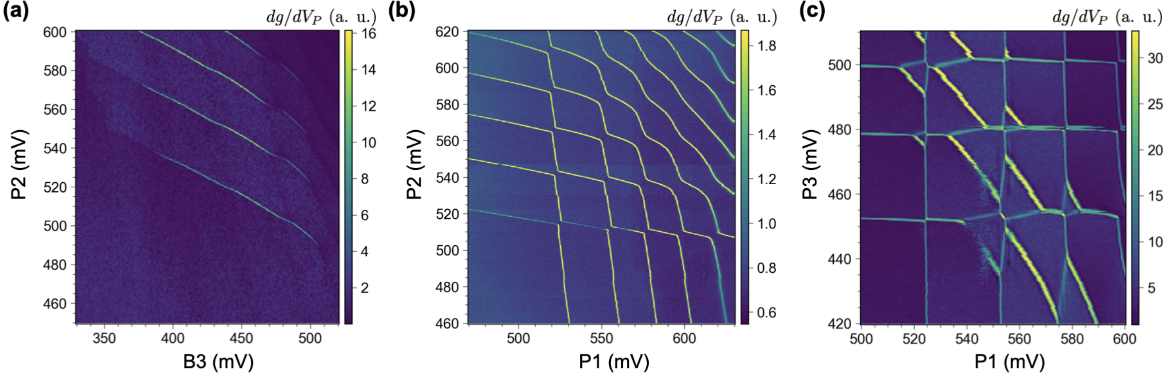

Figure 2.7: Lock-in detection of stability diagrams. (a) A dot beneath P2 is measured, each sloped line corresponding to a change in electron occupation of the P2 dot. B3 controls the tunnel rate into the dot. At low(high) B3 voltages, the tunnel rate is slow(fast). (b) Stability diagram of double dot. Since transitions stop occurring at regular intervals in the bottom left region of the plot, it can be assumed that each dot is fully emptied of electrons. The number of electrons can then be counted by adding an electron to each dot for each line that is crossed along the x- and y-axes. (c) A triple quantum dot is formed. Several transitions are shown for the P1 and P3 dots and the 0→1 transition for the P2 dot manifests as the diagonal line that jumps back and forth.

multi-dot system by using stability diagrams. Fig. 2.7 shows stability diagrams for one, two and three dot systems acquired using lock-in detection. In Fig. 2.7(a), the first three electron loading events into P2 are measured by changing the gate voltages on P2 and B3. P2 controls the chemical potential of the quantum dot, modifying the total number of electrons within the dot, and B3 controls the tunnel rate into and out of the dot. On the far left at low B3 voltages, the tunnel rate is slower than the lock-in frequency; thus the signal fades. At higher B3 gate voltages, the signal fades due to tunnel rate broadening. Once tunnel rate broadening far exceeds broadening due to the thermal energy of electrons, the signal is obscured. Fig. 2.7(b) shows a stability diagram of a double dot. Since transitions stop occurring at regular intervals in the bottom left region of the plot, it can be assumed that each dot is fully emptied

of electrons. The number of electrons can then be counted by adding an electron to each dot for each line that is crossed along the x- and y-axes. In Fig. 2.7(c), a triple quantum dot is formed. Several transitions are shown for the P1 and P3 dots and the $0 \rightarrow 1$ transition for the P2 dot manifests as the diagonal line that jumps back and forth. The diagonal transition line occurs because the P1 and P3 gate electrodes have a nearly equal cross-capacitance to the P2 quantum dot. The jumps in the transition line are due to cross-capacitance of the P1 and P3 quantum dots to the P2 dot. The P2 transition line jumps whenever an electron is loaded into P1 or P3 due to the cross-capacitance of the dots.

The previous figures in this section demonstrate well-controlled electrostatics in a triple-quantum dot device. Beyond electrostatic control, there are many other characterization techniques that are useful in determining the quality of a device for qubit operation. As discussed in Chapter 1, charge noise is the main source of noise contributing to decoherence of qubits in silicon; thus, it is important to have an efficient and accurate technique for measuring the charge noise in a device.

Fig. 2.8 shows a charge noise measurement technique that can be used in transport (discussed in further detail in Ref. [48]). In Fig. 2.8(a) measurements at two locations on a Coulomb peak are used to measure current noise spectra, denoted by the orange and green stars. Current fluctuations are proportional to gate voltage fluctuations induced by charge noise. This is captured at the point marked by the orange star since the current is highly sensitive to the gate voltage. At the green star, the derivative of the current with respect to the gate voltage is nearly zero, so the current noise is no longer representative of the charge noise in the system. Instead, other sources of noise such as the wiring and electronics of the experimental setup are captured. The total current noise of the system, excluding noise from wiring and electronics, can be

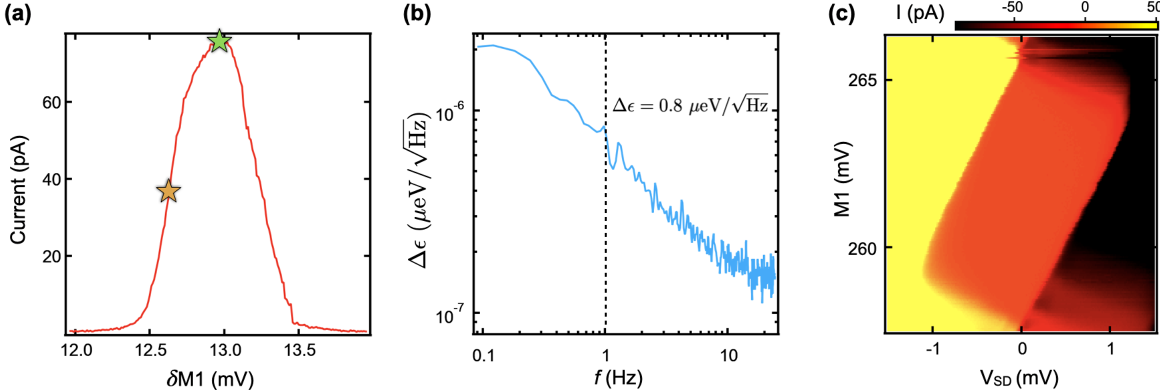

Figure 2.8: Charge noise measurement using the sensor dot beneath M1. (a) Measurements at two locations on a Coulomb peak are used to measure current noise spectra, denoted by the orange and green stars. Current fluctuations are proportional to gate voltage fluctuations induced by charge noise. This is captured at the point marked by the orange star since the current is highly sensitive to the gate voltage. At the green star, the derivative of the current with respect to the gate voltage is nearly zero, so the current noise is no longer representative of the charge noise in the system. Instead, other sources of noise such as the wiring and electronics of the experimental setup are captured. (b) The noise spectra at the green point is subtracted from the noise spectra at the orange point and plotted in units of $\mu\text{eV}/\sqrt{\text{Hz}}$. The charge noise value is often reported at 1 Hz, which here yields a value of 0.8 $\mu\text{eV}/\sqrt{\text{Hz}}$. (c) The current noise spectra is converted into detuning noise spectra by analyzing a Coulomb diamond in the few electron regime. Measurement of α, the gate voltage to energy lever arm, converts voltage noise into detuning noise. Current noise is converted to voltage noise using the slope at the orange star in (a).

summarized as

$$\delta I(t) = \frac{\partial I}{\partial \epsilon}\delta\epsilon(t) + \frac{\partial I}{\partial \Gamma_S}\delta\Gamma_S(t) + \frac{\partial I}{\partial \Gamma_D}\delta\Gamma_D(t) \tag{2.7}$$

where Γ_S and Γ_D are the tunnel rates to the source and drain contacts, respectively, and ϵ represents the charge noise. By subtracting the noise spectra at the green star from the orange star, noise contributions to the total noise from tunneling noise are also subtracted away, leaving only the charge noise. In Fig. 2.8(b), the noise spectra at

the green point is subtracted from the noise spectra at the orange point and plotted in units of μeV/$\sqrt{\text{Hz}}$. The charge noise value is often reported at 1 Hz, which here, yields a value of 0.8 μeV/$\sqrt{\text{Hz}}$. The conversion between current noise and detuning noise is given by Eqn. 2.8

$$\Delta I_\epsilon \alpha = |dI_{SD}/dV_L| \Delta\epsilon \tag{2.8}$$

where ΔI_ϵ is the measured current noise, $|dI_{SD}/dV_L|$ is the slope at the orange star in Fig. 2.8(a), and $\Delta\epsilon$ is the detuning noise. Fig. 2.8(c) shows measurement of α by analyzing a Coulomb diamond in the few electron regime. Using this charge noise technique, device charge noise can be quickly measured at a transport level while tuning devices into the few electron regime.

2.3 Valley-orbit states

The energy level structure in silicon quantum dots is more complicated than GaAs because of the presence of multiple valley states. In Fig. 2.9, the Brillouin zone of a silicon crystal is depicted in reciprocal lattice space. Unlike GaAs, the conduction band energy is minimized at $k_0 = 0.85X$ rather than at Γ ($k_0 = 0$). This leads to six degenerate valleys in bulk silicon's conduction band.

For quantum dots in Si/SiGe heterostructure, the silicon well experiences tensile strain from the SiGe buffer and spacer. This occurs due to the lattice mismatch of Si and SiGe, where SiGe's lattice constant is slightly larger than Si. This tensile strain works to lift the energy of the $\pm x$ and $\pm y$ valleys with respect to the $\pm z$ valleys, as shown in Fig. 2.9 [49].

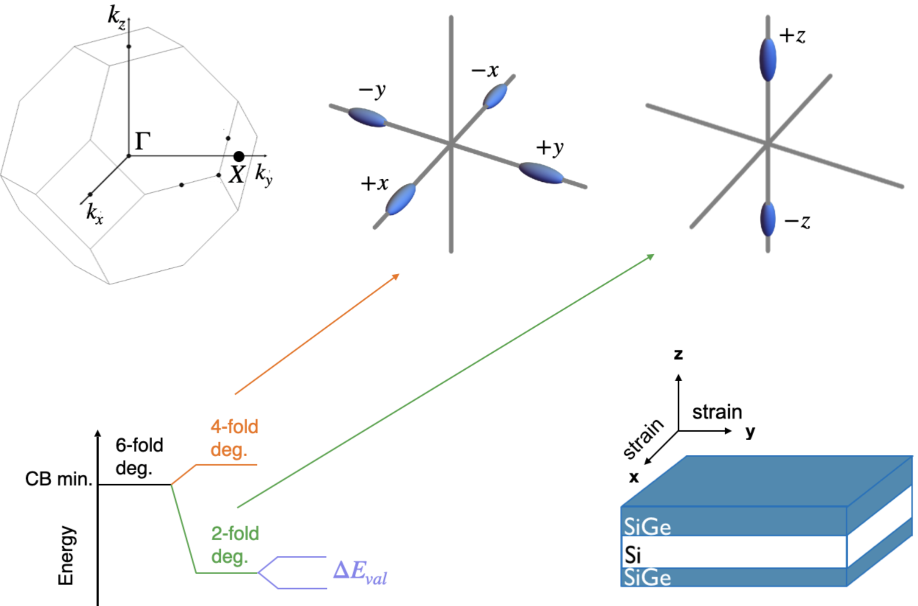

Figure 2.9: Valley states in Si/SiGe heterostructure. In bulk silicon, there is a 6-fold degeneracy of the conduction band in the $\pm x, \pm y$ and $\pm z$ directions. The silicon well experiences tensile strain due to the lattice mismatch between the Si well and the SiGe buffer/spacer. This tensile strain lifts the conduction band energy in the $\pm x$ and $\pm y$ directions with respect to the $\pm z$ directions. The remaining 2-fold degeneracy is lifted due to the phase difference between the $\pm z$ electron wavefunctions. This is illustrated in Fig. 2.10.

The remaining two-fold degeneracy is further split due to a number of physical phenomena, including electronic z-confinement from the quantum well, electric fields from gate electrodes, and the atomic details of the Si/SiGe interface near the 2DEG. In Fig. 2.10, the conduction band energy is plotted as a function of the heterostructure material stack, shown in red as $V(z)$. The slope of $V(z)$ is due to assumed metallic electrodes that appear as part of the device geometry at the top of the material stack. The $\pm z$ valley wavefunctions are plotted in orange and blue. Since the extent of the wavefunction is finite due to the confining potential of the Si/SiGe heterostructure, a phase difference arises between the two valley states. There is a non-zero component

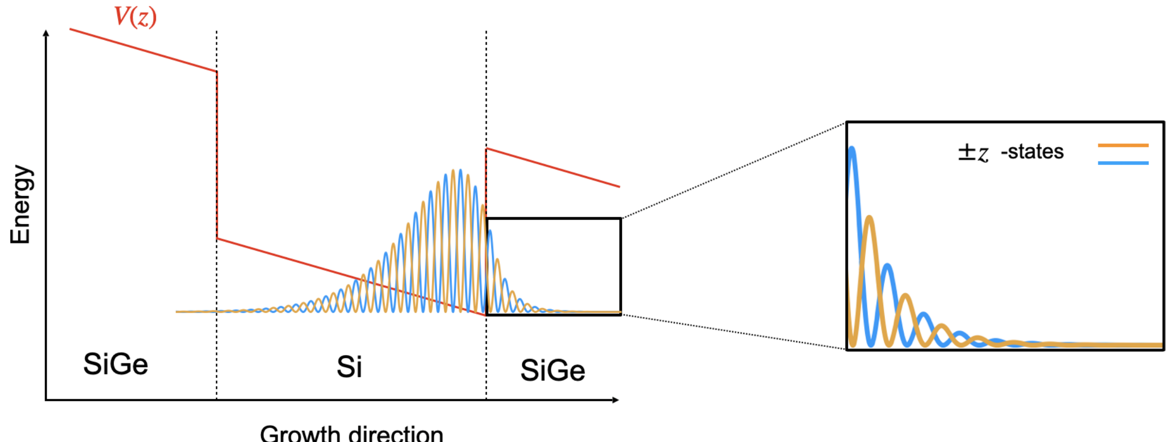

Figure 2.10: The origin of the valley splitting in Si/SiGe heterostructure. The conduction band energy is plotted as a function of the heterostructure material stack, shown in red as $V(z)$. The slope of $V(z)$ is due to assumed metallic electrodes that appear as part of the device geometry at the top of the material stack. The $\pm z$ valley wavefunctions are plotted in orange and blue. Since the extent of the wavefunction is finite due to the confining potential of the Si/SiGe heterostructure, a phase difference arises between the two valley states [50]. There is a non-zero component of each wavefunction in the higher energy SiGe spacer, depicted as the bordered inset. The phase difference gives rise to a different total amplitude of the $\pm z$ wavefunctions in the SiGe region, creating an energy difference between the two states. This energy difference splits the remaining 2-fold degeneracy and is known as the valley splitting.

of each wavefunction in the higher energy SiGe spacer, depicted as the bordered inset in Fig. 2.10. The phase difference gives rise to a different total amplitude of the $\pm z$ wavefunctions in the SiGe region, creating an energy difference between the two states. This energy difference splits the remaining 2-fold degeneracy and is known as the valley splitting.

In the absence of interface roughness, the valley splitting is predicted to be between 0.1–1 meV [51]; however, interface roughness in the form of atomic steps plays an important role in determining the valley splitting and is unavoidable in chemical vapor deposition (CVD) and molecular beam epitaxy (MBE) growths of Si/SiGe heterostruc-

ture. The valley splitting for Si/SiGe quantum dots has been measured to be between 20–270 µeV across many experiments [24, 52–60]. The large range of valley splittings occurs since heterostructure interfaces are not atomically pristine.

The orbital splitting is another important part of the level structure in silicon quantum dots. Quantum dots experience a confining potential that is similar to many simple quantum confinement problems such as the infinite square well, finite square well, or harmonic oscillator. The discrete states that arise from solving the Schrodinger equation for a given confinement potential leads to the orbital splitting for quantum dots.

The infinite square well is one of the simplest cases, leading to energy eigenvalues and wavefunctions

$$E_n = \frac{\pi^2 \hbar^2}{2m^* L^2} n^2 \tag{2.9}$$

$$\Psi_n = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L} x\right) \tag{2.10}$$

Plugging in a diameter of 38 nm for a dot [45], the orbital splitting is calculated as

$$E_{\text{orb}} = E_1 - E_0 = 4.11 \text{ meV} \tag{2.11}$$

which is very close to the average value obtained in Ref. [45] of 3.0 ± 0.5 meV. Although the infinite square well is a crude model of the confinement potential of an electron in a quantum dot, it reproduces the single-particle orbital splitting quite closely.

The harmonic oscillator is more representative of the confinement potential for quantum dots. The solutions to the energy eigenvalues and wavefunctions for the

harmonic oscillator are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{2.12}$$

$$\Psi_n = \frac{1}{\sqrt{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right), \qquad n = 1, 2, 3, ... \tag{2.13}$$

where $H_n(z)$ are Hermite polynomials. The orbital splitting is then set as

$$E_{\mathrm{orb}} = E_1 - E_0 = \hbar\omega \tag{2.14}$$

so in this case, the strength of the confinement potential directly corresponds to the orbital splitting. A representative length scale is given by $L = 2\sqrt{\hbar/m^*\omega}$ [61], which for a diameter of 38 nm leads to an orbital splitting of 1.11 meV; however, we note that this length scale is not necessarily the same length scale as in the infinite square well. Nonetheless, orbital splittings are reproduced closely here as well.

The deviation from the true confinement potential is apparent in Fig. 2.11(a). A parabolic potential is compared with an electrostatic simulation (courtesy of H. Ekmel Ercan) of the confinement potential in a quantum dot. Generally, at lower electron occupations and lower energy orbital states, a parabolic potential is very close to the full simulation potential, giving good insight into the behavior of electron wavefunctions and eigenvalues. In Fig. 2.11(b), Schrodinger's equation is solved for a parabolic potential. The wavefunctions and energy eigenvalues of the first four orbital states are plotted and the parabolic potential is shown as the shaded gray region. In more realistic simulations of the confinement potential, there is a slight anharmonicity in the orbital splittings. For higher orbital states, the splitting becomes smaller and smaller

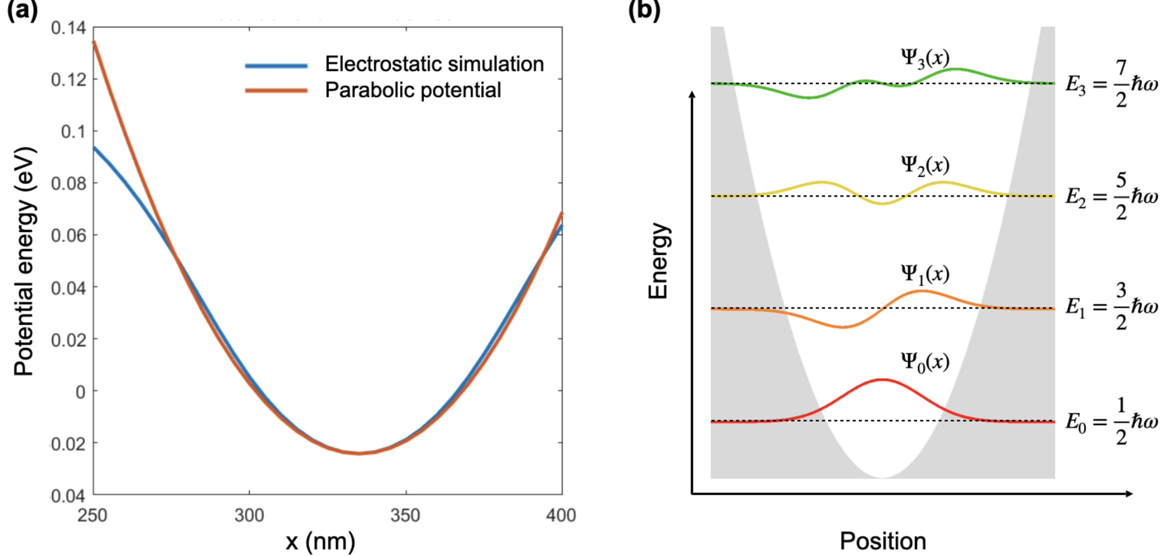

Figure 2.11: Simulating orbital splittings using a parabolic potential. (a) A parabolic potential is compared with an electrostatic simulation (courtesy of H. Ekmel Ercan) of the confinement potential in a quantum dot. Generally, at lower electron occupations and lower energy orbital states, a parabolic potential is very close to the full simulation potential, giving good insight into the behavior of electron wavefunctions and eigenvalues. (b) Schrodinger's equation is solved for a parabolic potential. The wavefunctions and energy eigenvalues of the first four orbital states are plotted and the parabolic potential is shown as the shaded gray region.

due to the weaker confining potential as compared to a parabolic potential.

The full level structure for a silicon quantum dot incorporates both the valley splitting and orbital splitting. For each orbital state, there exists a ground and excited valley state an electron can occupy. This would double the amount of single particle energy levels shown in Fig. 2.11(b). Additionally, valley-orbit coupling perturbs the bare valley and orbital states. Valley-orbit coupling occurs in silicon quantum dots due to the interplay between the finite size of the electron wavefunction and quantum well interface roughness. The position of the Si/SiGe interface affects the valley splitting because as the position changes, the phase of $\pm z$ wavefunctions changes. Intuition for

this behavior can be gained from Fig. 2.10. If the position of the interface moves, the total amplitude in the SiGe region of the heterostructure of each valley state changes. Thus, if the quantum well interface is not atomically pristine, the valley phase changes as a function of position in the x-y plane. Since the wavefunction of the electron is confined in the x-y plane, the orbital component of the wavefunction is directly coupled to the valley splitting. Valley-orbit coupling has a number of implications, which are discussed in detail in Ref. [62].

The level structure is further complicated in the multi-electron regime. When two electrons are on the same dot, four spin states can arise: a singlet and three triplets. Combined with the spatial wavefunction, the total wavefunction for the $|S_0\rangle$ and $|T_0\rangle$ spin states are

$$|\Psi_{S_0}\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) |gg\rangle \tag{2.15}$$

$$|\Psi_{T_0}\rangle = \frac{1}{2} \left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) \left(|ge\rangle - |eg\rangle\right) \tag{2.16}$$

where $|g\rangle$ denotes the ground spatial state and $|e\rangle$ denotes the first excited spatial state. These total wavefunctions must be antisymmetric since they represent fermionic systems. This is shown in Eqn. 2.16, where the spatial state for the symmetric spin-triplet state is antisymmetric. Spatial-like states are given by both orbital and valley states, for example, $|g\rangle = |E_{\mathrm{orb}_0}, E_{\mathrm{val}_0}\rangle$. The excited state could include either the ground orbital state and the first excited valley state, or the ground valley state and the first excited orbital. In almost all cases for silicon quantum dots, however, the first excited valley state is much lower in energy than the first excited orbital state.

For the triplet state, there is partial occupation of an excited orbital or valley state.

This results in the triplet state being higher in energy than the singlet state. This singlet-triplet splitting is very similar to the valley splitting, but with one fundamental difference: The singlet-triplet splitting is the two-particle splitting between the ground and first excited state, and the valley splitting is the single-particle splitting between the ground and first excited state.

In the limit that the confinement potential is infinitely strong, the singlet-triplet splitting equals the valley splitting [63]. As electron-electron interactions become on the order of the strength of the confining potential, the singlet-triplet splitting can be suppressed from the non-interacting limit. This is demonstrated for Si/SiGe quantum dots in Ref. [64]. The suppression of the singlet-triplet splitting has important implications for designing and operating valley-orbit qubits in Si/SiGe.

The quantum dot hybrid qubit (QDHQ) is a type of valley-orbit qubit which gets its name because it combines both charge and spin degrees of freedom. Having a charge degree of freedom allows for strong electromagnetic interactions between electrostatic gates and qubit states since a dipole moment forms between the logical $|0\rangle$ and $|1\rangle$ states. The spin degree of freedom enables qubits to be encoded with spin in addition to the charge degree of freedom, drastically increasing coherence times over pure charge qubits. While unwanted charge noise couples to the charge degree of freedom, the decoherence rate due to this coupling is suppressed by detuning "sweet spots" where the energy level structure is flat.

The logical states of the hybrid qubit are [65]

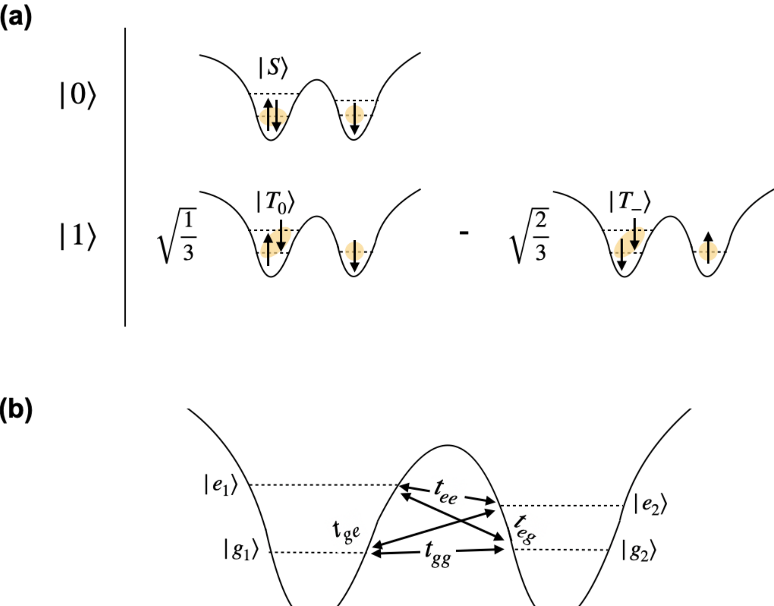

Figure 2.12: (a) The logical states of the QDHQ are shown spatially. For $|0\rangle$, two electrons form a singlet state in the left dot, and a third spin-down electron is in the ground valley state of the right dot. For $|1\rangle$, the system is in a superposition of triplet states. The first state is given as two electrons in a $|T_0\rangle$ configuration in the left dot and one down-spin electron in the right dot. The other state contains two electrons in a $|T_-\rangle$ configuration in the left dot and one up-spin electron in the right dot. (b) There are four distinct tunnel coupling elements between the ground and excited valley-orbit states. This tunnel coupling are not necessarily the same and dictate key qubit parameters and qubit performance.

$$|0\rangle = |S\rangle\,|{\downarrow}\rangle \tag{2.17}$$

$$|1\rangle = \sqrt{\frac{1}{3}}\,|T_0\rangle\,|{\downarrow}\rangle - \sqrt{\frac{2}{3}}\,|T_-\rangle\,|{\uparrow}\rangle \tag{2.18}$$

where Fig. 2.12(a) captures the spatial position of the electrons within the double-dot system. For $|0\rangle$, a singlet state is formed with two electrons in one dot and an additional spin down electron in the second dot. For $|1\rangle$, a superposition of two states

is formed where the superposition is between $|T_0\rangle$ and $|T_-\rangle$ in the left dot. Then, there is an additional spin down electron in the right dot, pairing with the $|T_0\rangle$ state, and a spin up electron paired with the $|T_-\rangle$ state. Thus, both logical states have the same S^2 and S_z. This is important because spin flips need not be induced when gating this qubit, allowing for very fast Rabi flopping.

Fig. 2.12(b) shows that there is a distinct tunnel coupling element for each of the four possible transitions within the double quantum dot. Each tunnel coupling element is labeled with subscripts t_{ij} where i corresponds to the left quantum dot and j corresponds to the right quantum dot. i and j can take on possible values g and e which denote the ground or first excited valley-orbit states. The tunnel couplings are not necessarily the same and dictate key qubit parameters and qubit performance. The Hamiltonian in the charge basis for this system is derived in Refs. [66, 67] and given here as

$$H|\Psi\rangle = \begin{pmatrix} \epsilon/2 + E_{ST_L} & 0 & t_{ee} & t_{ge} \\ 0 & \epsilon/2 & -t_{ge} & t_{gg} \\ t_{ee} & -t_{ge} & -\epsilon/2 + E_{ST_R} & 0 \\ t_{ge} & t_{gg} & 0 & -\epsilon/2 \end{pmatrix} \begin{pmatrix} |(2,1)_e\rangle \\ |(2,1)_g\rangle \\ |(1,2)_e\rangle \\ |(1,2)_g\rangle \end{pmatrix} \tag{2.19}$$

where ϵ is the double dot detuning, E_{ST_L} is the singlet-triplet splitting of the left dot and E_{ST_R} is the singlet-triplet splitting of the right dot. Solving for the eigenvalues yields the solutions shown in Fig. 2.13 as a function of detuning. The curvature in the eigenstates as a function of detuning is present due to the non-zero tunnel couplings hybridizing the charge states. Thus, the flatness of each energy level with respect to

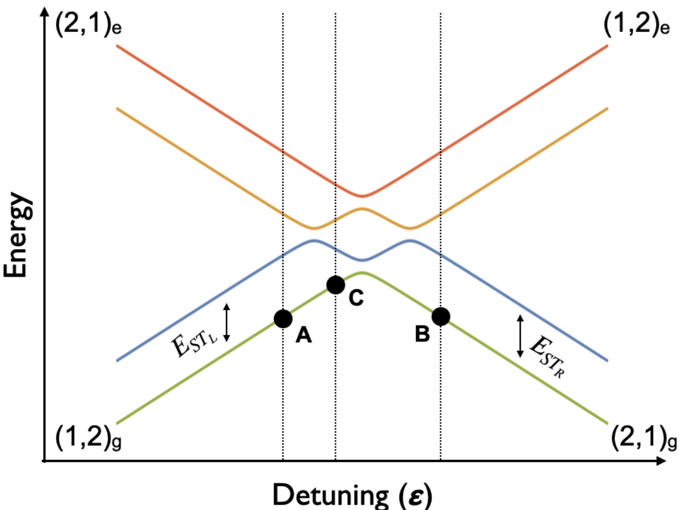

Figure 2.13: QDHQ dispersion. The ground and first excited valley-orbit states are shown near the double dot avoided crossing. Several avoided crossings form due to distinct tunnel coupling elements between the ground and excited states of each quantum dot. In this case, the energetic separation of the ground and first excited state is given by the singlet-triplet splitting of each dot, E_{ST_L} and E_{ST_R} since there are two electrons in the charge configurations shown. Qubit initialization (A), manipulation (B), and readout (C) points in detuning are shown. By adiabatically ramping from A to B, the qubit is loaded into the logical $|0\rangle$ state, where a resonant microwave drive can be applied to gate the qubit. After the gate has been applied, the qubit is adiabatically ramped to C, where the charge configuration varies depending on the resulting logical state after manipulation.

one another is largely determined by the single-triplet splittings, the tunnel couplings, and the detuning.

Several avoided crossings form in Fig. 2.13 due to distinct tunnel coupling elements between the ground and excited states of each quantum dot. In this case, the energetic separation of the ground and first excited state is given by the singlet-triplet splitting of each dot, E_{ST_L} and E_{ST_R} since there are two electrons in the charge configurations shown. Fig. 2.13 shows the qubit initialization (A), manipulation (B), and readout (C) points in detuning. By adiabatically ramping from A to B, the qubit is loaded into the

logical $|0\rangle$ state, where a resonant microwave drive can be applied to gate the qubit. After the gate has been applied, the qubit is adiabatically ramped to C, where the charge configuration varies depending on the resulting logical state after manipulation.

Chapter 3

Device Fabrication

Robust fabrication of quantum dot devices in Si/SiGe heterostructure is a demanding task and remains one of the biggest outstanding challenges for the field of silicon-based quantum processors. There are several factors that play into the challenging fabrication for Si/SiGe devices. First, the critical dimensions of devices are on a nanoscale due to the size of the electron wave function within the quantum dot. The wave function is even smaller in silicon than in other material platforms, such as GaAs/AlGaAs quantum dots, which is due to the large effective mass of electrons in the conduction band of silicon. Since the effective mass of the transverse component of the electron wave function is $0.19m_e$ in silicon, and the effective mass for GaAs devices is $0.067m_e$, silicon devices must be patterned smaller in order to achieve equivalent energy level spacings. Due to the small critical dimensions of devices in Si/SiGe, there is little room for error when lithographically patterning devices. Precise and repeatable patterning of devices at a 50 nm length scale becomes a necessity.

Additionally, quantum dots fabricated in Si/SiGe add a degree of complexity over devices fabricated in Si. The additional constraints come from the heterostructure

itself. The main difference is that Si/SiGe heterostructure places an upper limit on the thermal budget of devices because the topmost Si/SiGe heterojunction must be sharp. When Si/SiGe heterostructure experiences a temperature exceeding 800 °C, interdiffusion of the Si and SiGe [68] can occur. This may lower the valley splitting in silicon quantum dots, making it difficult to coherently manipulate, readout, or initialize qubits. The consequence of this lower thermal budget is that thermal growth of high quality SiO_2 becomes difficult, and other dielectric growth or deposition processes must be used.

Finally, choosing a device architecture and geometry is a complex and multifaceted task. On the one hand, covering the substrate with as much gate metal as possible is advantageous due to a higher degree of control of the chemical potential of the active region of the device. On the other hand, increasing the number of gate electrodes per quantum dot increases the complexity of tuning devices and decreases the yield of devices due to poorly formed gate electrodes. A careful analysis of varying architectures and geometries is discussed further in the Section 3.1.

This chapter is a detailed discussion of fabrication for quantum dots devices in Si/SiGe. Section 3.1 is an overview of varying device architectures and geometries is given. The advantages and disadvantages of using an overlapping aluminum gate architecture instead of other architectures is discussed thoroughly. Section 3.2 is a step-by-step guide for fabricating overlapping aluminum gate devices, with careful consideration for particular process steps and alternatives that exist. Section 3.3 details a yield engineering investigation into overlapping aluminum gate devices and shows measurements on a device using an improved fabrication process. Finally, in Section 3.4, the results of this chapter are summarized and a proposal for future generations of devices is given.

3.1 Device architectures overview

There are several different general considerations to make when designing a gate electrode architecture and geometry for quantum dots in Si/SiGe. First, an architecture which forms robust quantum dots with easily tunable chemical potentials and tunnel couplings is vital. Spurious (unintentional) dots that may form while the intentional dots are tuned must be limited as they can negatively impact device performance. The ability to suppress undesired cross-talk or tunneling events between pairs of quantum dots is also required. There are several other crucial elements to consider when designing a device that will affect its performance, including charge sensitivity, orbital confinement, gate stack for minimal charge noise, and gate architecture to maximize device yield. What are the general approaches for simultaneous optimization of these parameters? Below two main architectures are discussed, including several variations: “open-style” designs and “tight” designs.

Open and tight designs

For GaAs/AlGaAs heterostructure, open-style designs were prevalent during the early 2000s, when a number of seminal demonstrations for semiconductor quantum dots as a platform for quantum computation took place [11, 13, 15–17, 69]. Open-style devices are particularly suitable for GaAs since they are depletion mode devices — devices are “on” when gate voltages are set to zero such that current flows between source-drain contacts. Undoped silicon-based quantum dots are accumulation mode devices — a positive gate voltage is needed for non-zero current to flow between source-drain contacts. In GaAs, where the effective mass of conduction band electrons in is small and depletion mode devices are used, it makes open-style devices a logical choice

due to its simplicity. In silicon, however, the choice becomes much more ambiguous.

Modulation-doping, a technique pioneered in GaAs/AlGaAs superlattices [70] which was later studied in Si/SiGe heterostructures [71], can be used for Si/SiGe devices, making them similar to GaAs devices in that the dopant layer produces depletion mode devices. This led many first generation device architectures in Si/SiGe to take on an open-style similar to GaAs devices, in which a number of fundamental milestones were accomplished, including measurements of T_1, valley splitting, Pauli-spin blockade, single-shot readout and coherent control of charge and spin-based qubits [25, 52, 54, 72–77]. However, doped Si/SiGe devices were phased out in favor of undoped devices because it was found that intentional dopants increased noise and instability of devices [78]. This switch in part helped push Si/SiGe devices into a more viable platform for quantum computation, in which the first singlet-triplet and single-spin qubits were demonstrated [23, 27].

Overlapping-gate designs

The final pivotal change in device architecture and geometry to date is more subtle. The first undoped Si/SiGe devices used a global top gate to accumulate the two-dimensional electron gas (2DEG) and depletion gates to form quantum dots, making it an open-style device. A different architecture (with many variations) became a staple in Si/SiGe, largely known as the overlapping gate architecture, or “tight” design. The fundamental modification of these devices is that each quantum dot has its own top gate which is responsible for controlling the chemical potential of that dot. In some designs, depletion gates are adjacent to these top gates, which control the tunnel coupling into and out of neighboring dots and reservoirs. This varies from the global top gate devices in that each gate electrode has a specific and unique function in tuning

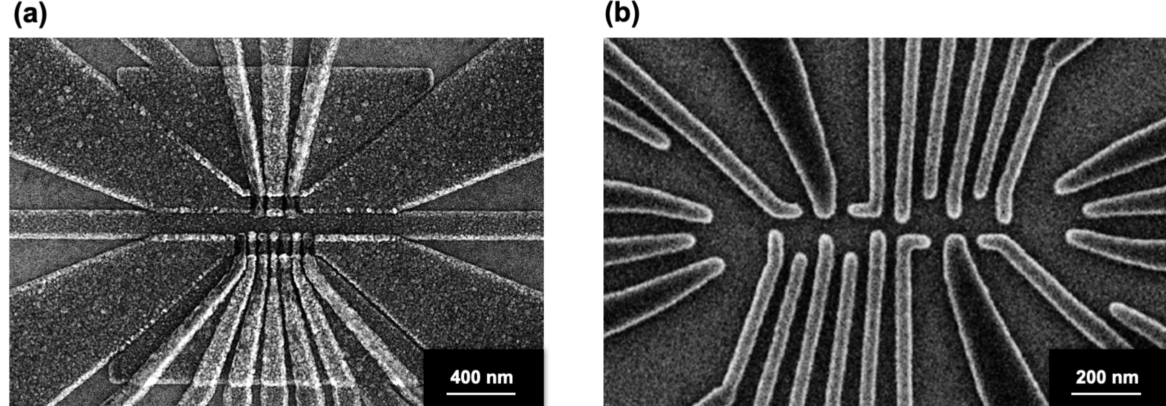

Figure 3.1: Top down scanning electron microscope (SEM) image of overlapping gate and open-style devices for quantum dots in Si/SiGe. (a) Example of a triple-dot defined using an overlapping gate architecture. The gates are patterned such that quantum dots are formed directly beneath particular gates, which control the chemical potential of the dots. For each dot, there exists a neighboring barrier gate that controls the tunnel coupling into and out of adjacent dots and reservoirs. (b) Example of a quadruple quantum dot defined using an open-style architecture. Multiple gates are used to "corral" electrons into quantum dots. Above the depletion gates, shown in Fig. 3.1(b), there are top gates used for accumulation (not shown). The pockets in gate metal formed by the depletion gates near the center of the device are where the quantum dots are formed.

and controlling the chemical potential and tunnel coupling of devices. Although more advanced open-style devices use multiple top gates, the difference between these designs and tight designs are that tight designs minimize gaps between gate electrodes, often using multi-layer fabrication processes that involve overlapping gates.

The main difference in the two gate architectures is observed in the top down scanning electron microscope (SEM) images shown in Fig. 3.1. Fig. 3.1(a) is an SEM image of an overlapping aluminum gate triple-dot device. The finger gates that define the charge sensor (top side) and triple-dot array (bottom side) overlap each other such that the substrate is completely covered by aluminum gate metal. This is a more robust approach to tuning devices since there is a higher degree of control of

the chemical potential of the 2DEG than in other designs. Additionally, devices with gaps in gate metals are susceptible to charge impurities that can create spurious dots. In overlapping gate designs, this problem is mitigated. Another beneficial feature of the overlapping gate architecture is the use of screening gates. Screening gates are patterned on the first layer, and suppress formation of spurious dots by accumulation gates in subsequent layers. Finally, the overlapping gate architecture can be easily tiled to form larger arrays of quantum dots [45]. This is another advantage over open-style devices, whose dense array of depletion gates makes it difficult to use integrated charge sensing techniques beyond four or eight dots. Fig. 3.1(b) is an SEM image of the depletion gate layer of a quadruple quantum dot using an open-style (or "stadium-style") architecture. The integrated charge sensors are found on the left and right sides of the device. Current through the charge sensors is accumulated using top gates (not shown). Additionally, separate top gates control the chemical potential of the quantum dot. The quantum dots are located within the gaps in depletion gate metal near the center of the device. While open-style devices are highly flexible because gates can be multi-purposed, the ability to scale and tile this architecture becomes cumbersome beyond four dots. For this reason, overlapping gate architectures have become a natural choice for new generation devices.

Overlapping gate architectures encompass a wide range of material stacks and gate geometries, each with their own respective benefits and challenges. The first overlapping gate architecture used in silicon was for Si-MOS dots, which used an overlapping aluminum gate architecture without deposited dielectric between gates [79]. In terms of the gate stack, there are two main approaches. These approaches are illustrated in Fig. 3.2(b,c) alongside an open-style device in Fig. 3.2(a). In Fig. 3.2(a), the open-style device shows gaps between gates in the first layer. While the function of these gates

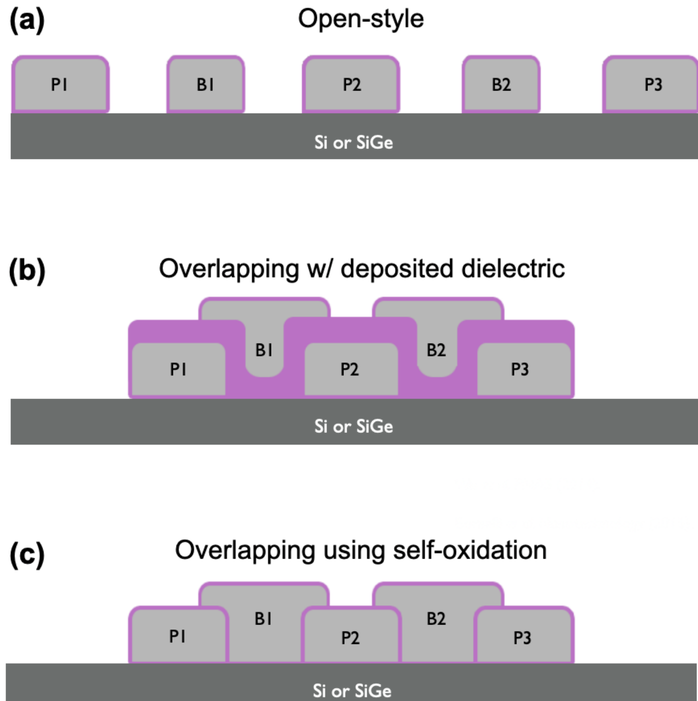

Figure 3.2: Schematic cross-sections of different device architectures. (a) Open-style device cross-section. Gates are not overlapping and gate metal does not completely cover the substrate in the active region. (b) Schematic cross-section of an overlapping gate architecture with a deposited dielectric. By depositing dielectric between gate layers, devices may be more robust to electrostatic discharge (ESD) events and the gate metal does not have to self-oxidize, allowing for more flexible fabrication processes. (c) Schematic cross-section of an overlapping gate-architecture without deposited inter-gate oxide. By using aluminum gate metal, the exterior of aluminum gates naturally grows oxide, electrically isolating gate layers from each other.

varies between different open-style designs, the defining characteristic of them is the existence of gaps in the gate metal where either the chemical potential of the dot is controlled or the tunnel coupling between dots is tuned.

In overlapping gate approaches, there are no gaps in the gate metal between lithographically defined quantum dots. A key difference between two overlapping-gate architectures is shown in Fig. 3.2(b) and Fig. 3.2(c). In Fig. 3.2(b), dielectric is deposited

(purple) between gate layers, creating an electrically isolating barrier. In Fig. 3.2(c), a thin native oxide grows on the gate metal, and no deposited dielectric is necessary. There are many considerations needed when choosing between one of these two general architectures. Depositing dielectric between gate layers enhances electrical isolation, which helps protect devices from electrostatic discharge (ESD). Additionally, there is greater flexibility in choosing the type of gate metal (Pd, Au and Al are common choices) and potential for higher quality post-metallization anneals. For overlapping devices *without* deposited dielectrics, the metallic gate stack may have a smoother topology, less physical space between gates, and mitigate potentially noisy interfaces created by deposition processes. However, these devices rely on the native oxide that grows on gate metal when exposed to air, thus the metal must self-oxidize. Aluminum is the only viable option, and is discussed in more detail in Section 3.3.

The choice of gate metal affects the fabrication process of quantum dot devices in different but important ways, which must be considered. As mentioned above, choosing between Au, Al, or Pd is only an option in device architectures with deposited inter-gate oxide. Typically, Au and Pd are easier to work with since they do not oxidize in air. This allows for simpler interconnect designs since these metals can be directly contacted electrically with subsequent gate layers whereas this is not possible with Al. Additionally, Au and Pd have smaller grains sizes than Al, leading to smoother gate metal profiles [80]. Finally, Au and Pd are not etched by common etchants such as hydrofluoric acid (HF) [81], making process design simpler and more robust. While there are many clear advantages in choosing gate metals besides Al, there may be more subtle reasons for choosing Al over Pd or Au. Superconducting material for gate metal may be beneficial for a number of reasons, including a reduction in noise coupling between gates and quantum dots, and reduction of power dissipation by electrically

driven gates. For the scope of this thesis, Al is chosen only for its self-oxidation property, allowing for the omission of deposited inter-gate dielectric.

3.2 Fabrication methods

In this section, a detailed step-by-step fabrication guide for overlapping aluminum-gate devices without deposited inter-gate dielectric is presented and discussed. A process flow diagram is shown in Fig. 3.3. Photolithography steps are carried out at a wafer-level and electron-beam (e-beam) lithography is performed at a chip level. There are several advantages to processing at a wafer level: many fabrication tools are designed and optimized for wafers, higher throughput of devices, easier substrate handling and the ability to quickly pattern hundreds of devices with a robust photomask. All of these factors work to increase the yield and performance of devices. Additionally, a mesa design that patterns a large number of interconnects and corresponding bond pads allows for flexible quantum dot device designs at the e-beam level. Thus, a single design for a device at the microscale level accommodates many different designs at the nanoscale level, allowing for quick iteration of device designs and geometries in the active region. Carefully designing a fabrication process that is both robust and efficient is a crucial part of semiconductor quantum dot research due to the length and difficulty of device fabrication. In addition to presenting a step-by-step fabrication guide in this section, bring-up of this fabrication process is discussed and motivated in the context of chip-level lithography and the challenges that were overcome in the fabrication of Si/SiGe devices at the University of Wisconsin-Madison (UW-Madison).

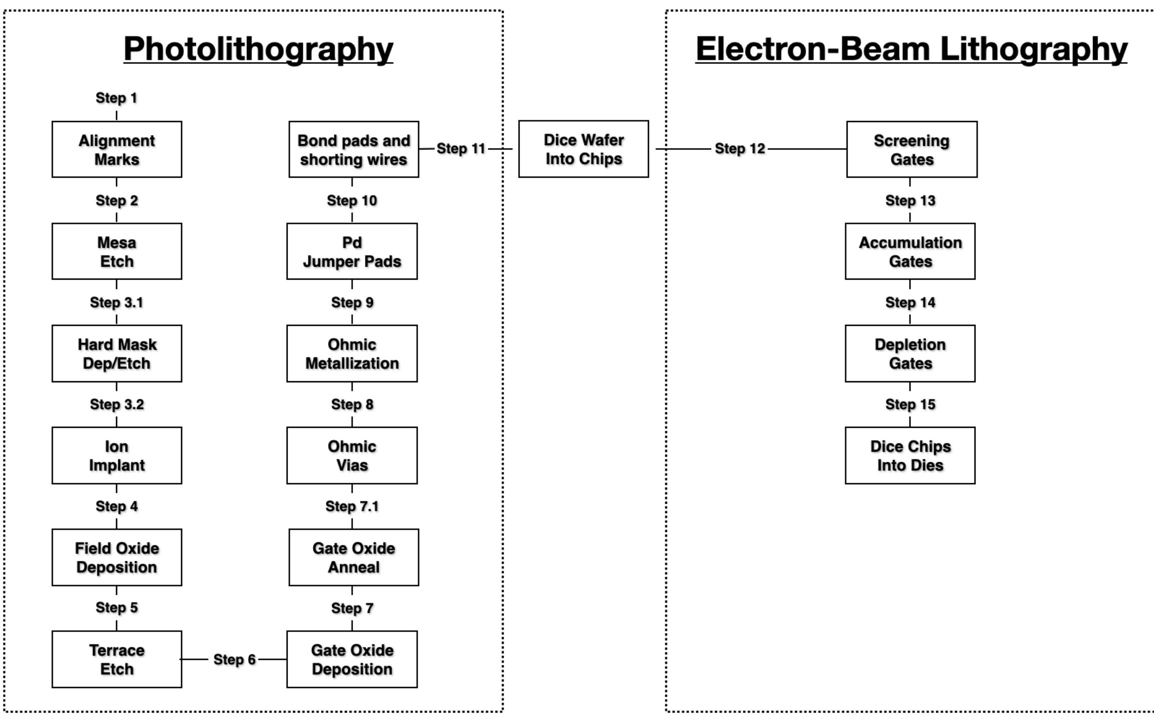

Figure 3.3: Flow chart for fabrication process used in fabricating overlapping aluminum-gate devices without inter-gate dielectric. First, a number of photolithography steps are carried out at a wafer level to create an interconnect structure between on-chip bond pads and the active region of the device. Later, the wafer is diced into chips, and electron-beam lithography is used to pattern and metallize the active region of the device in three separate steps.

3.2.1 Microscale fabrication

Nikon stepper

The UW-Madison clean room, or Nanoscale Fabrication Center (NFC), is a nanofabrication facility with over 70 instruments designed to aid in microscale and nanoscale fabrication. It contains a number of tools that can be used for photolithography, including an MJB3 aligner, MA6 aligner and a Nikon i-line stepper (NSR-2005i8A). While contact lithography (MA6 aligner) is a viable option for microscale fabrication of quantum dots, projection lithography using the Nikon stepper is preferred due to its higher

resolution and alignment precision. It is also noted that there are other instruments such as a Heidelberg direct write system that work well for the photolithography steps; however, there are no direct write laser systems in the NFC.

The Nikon stepper is a fully automatic photolithography patterning system that can accommodate wafers of 2”, 3” or 4” diameter. The stepper uses i-line irradiation, resulting in a resolution as small as 0.5 µm. Additionally, the stepper has an internal alignment protocol using on-chip alignment marks that are patterned by the user. This allows for overlay accuracy as precise as 150 nm. The resolution and overlay accuracy exceeds the constraints for the steps in the fabrication of quantum dot structures, which is 1 µm resolution and 0.5 µm overlay accuracy in the process described here. The i-line irradiation passes through a series of mirrors, lenses and a reticle (or mask) to transfer patterns onto the substrate. The stepper also has “blinding” capabilities, allowing for exposure of designated regions of the reticle, giving high flexibility in pattern development. The Nikon stepper is currently the sole photolithography patterning tool used for fabrication of quantum dot devices in the Eriksson Group.

Chip photolithography

Work using the stepper began at a chip-level. This is both unconventional and less than ideal, but it was necessary due to the substrate growth process at UW-Madison. Si/SiGe heterostructure was grown using chemical vapor deposition (CVD) on rectangular substrates due to the constraints of the growth chamber. Since it is difficult to purchase or grow high quality Si/SiGe heterostructure wafers, chip-lithography was the only option up until 2016 at UW-Madison. Despite being constrained to chips, there is still motivation for using photolithography processes over e-beam processes. For large area writes (such as bond pads), e-beam lithography is slow, making large batches of

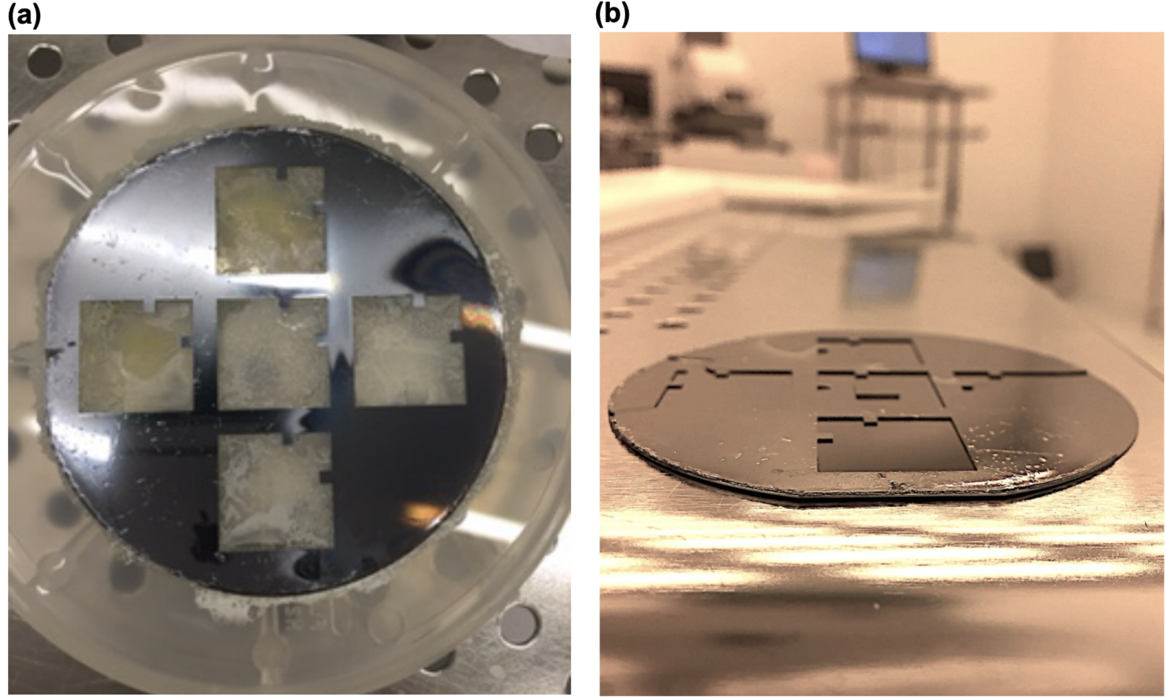

Figure 3.4: Fabrication of carrier wafer for chip lithography in the stepper. (a) Top wafer of double-stack carrier wafer after it has been etched using the Bosch process on the STS Multiplex direct reactive ion etcher (DRIE). Five cavities are etched in this carrier wafer for simultaneous patterning of up to five chips. (b) Carrier wafer after the top etched wafer is glued to the bottom wafer using photoresist.

devices time consuming to pattern. Furthermore, photoresists are more tolerant to certain fabrication processes, such as HF etches, where bleed-through or delamination of e-beam resists such as PMMA are more problematic.

A chip-level photolithography process using the stepper was brought up to improve the yield and throughput of devices beginning in 2015. Since the stepper only accepts substrates in the form of wafers between 2-4”, this involved using a carrier wafer with an etched out area containing mountable slots for chips. The carrier wafers are fabricated using two wafers. The bottom wafer is a standard silicon wafer with dimensions that are acceptable for the stepper. The top wafer is a silicon wafer which has a cavity

milled completely through it using an STS Multiplex direct reactive ion etcher (DRIE). By using the Bosch process, a highly anisotropic, fast etch through the silicon wafer is achieved, creating a high resolution cavity that can be flexibly designed for carrying chip substrates. The top carrier wafer and chip substrate must be of similar thickness, otherwise the stepper has focusing issues. The top and bottom wafers are glued together using photoresist and the double-stack wafer containing an etched cavity is the final product for enabling chip-lithography. A completed mountable wafer is shown in Fig. 3.4. Fig. 3.4(a) shows the top wafer post-etch in the STS DRIE, demonstrating a carrier wafer with five cavities for patterning of up to five chips simultaneously in the stepper. Fig. 3.4(b) shows the completed wafer, after the top etched wafer has been glued to the bottom wafer via photoresist.

Wafer photolithography process

After HRL began supplying the Eriksson Group with Si/SiGe wafers, the chip-lithography process was converted into a nearly identical wafer-level process. This became the main fabrication process for the group and for overlapping aluminum-gate devices. Chip photolithography is still useful for novel Si/SiGe heterostructures grown in non-standard dimensions. An overview of the process is shown in Fig. 3.3, and below it is described step-by-step.

1. First, alignment marks are patterned and etched using the stepper. The stepper needs a particular spacing of marks such that it can align to them for multi-layer processes. Usually alignment marks and the first patterning step for a fabrication process are done at the same time, but it was found that a systematic offset between the first layer and subsequently aligned layers occurs on this step-

per. More modern steppers would not have this issue, allowing for the first and second steps to be written simultaneously. However, for this process, they are patterned in separate steps due to this systematic offset. Alignment marks are etched into the substrate using a Uniaxis 790 RIE with an SF_6 process gas. All RIE processes are run with a carbon platen to prevent etch grass from contaminating the surface [82]. The lithography process for this step is described in Appendix A.1.1. Once the etch is finished, the photoresist is stripped using the process in Appendix A.1.5.

2. The second step is the mesa etch. The mesa etch results in an isolating trench surrounding the active region of the device. This is required for the interconnect structure used in this process since bond pads are only evaporated to be 200 nm thick. When wire bonding with a wedge bonder, it is recommended that bond pads be 750–1000 nm thick such that the bonding wire does not penetrate the substrate. Since the bond pads are thinner than this, 2DEG may be accumulated beneath bond pads and their corresponding gate leads due to the fact that there is contact between the metallic gate and the Si well. To cut off this accumulation, the mesa etch step etches completely through the 2DEG. The process gas for this etch is CHF_3 rather than SF_6 due to the lower selectivity between Si and SiGe; that is, the etch rate for both materials are more similar than for SF_6, allowing for a less stringent and heterostructure dependent etch time. After the etch is completed, the photoresist is stripped using the process in Appendix A.1.5.

3. The third lithography step is to define ion implantation windows that form ohmic contacts. This allows for electrons to be supplied to the device, creating source-drain contacts. For many charge sensing schemes there is a source-drain current

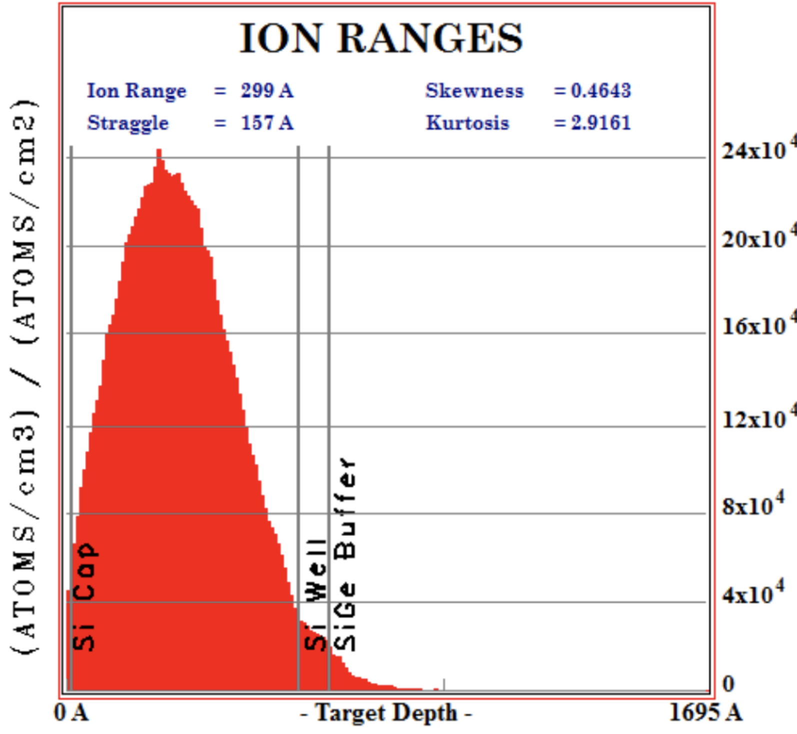

Figure 3.5: Monte-Carlo simulation using Stopping and Range of Ions in Matter (SRIM). The heterostructure spacer is 60 nm (high end of typical range), thus the implant energy is chosen to be 30 keV. From the simulation, 3×10^4 atoms/cm^2 is reached at the 2DEG. The dose is chosen to be 5×10^{15} ions/cm^2, reaching a density of 1.5×10^{20} ions/cm^3. This is well above the silicon metal-insulator transition of roughly 3.74×10^{18} ions/cm^3 necessary to create a robust ohmic contact.

that runs through these ohmics. Ion implantation is a process by which ions are accelerated towards a substrate and are physically embedded within that material. In the case of Si/SiGe heterostructure ion implantation for ohmic contacts, phosphorus ions are used ($^{31}P^+$). The accelerating energy is typically between 20–30 keV for Si/SiGe heterostructure spacers of 20–60 nm, the delivered dose is 5×10^{15} ions/cm^2, and the implantation angle with respect to the substrate is 7°. This is to ensure proper scattering of the ions once they enter the substrate, otherwise the scattering length may be too long and lead to unreliable

results. The accelerating energy is determined using a Monte-Carlo simulator for the Stopping and Range of Ions in Matter (SRIM) [83]. SRIM calculates the density of ions that reach the Si well given a delivered dose, accelerating energy, material stack and ion type. The ion density at the Si well should be above the metal-insulator transition of Si (roughly 3.74×10^{18} ions/cm^3) [84]. Fig. 3.5 shows a SRIM simulation for heterostructure with a 60 nm spacer. From the simulation, 3×10^4 atoms/cm^2 is reached at the 2DEG when using an implant energy of 30 keV. The dose is chosen to be 5×10^{15} ions/cm^2, reaching a total density of 1.5×10^{20} ions/cm^3. This is well above the silicon metal-insulator transition of roughly 3.74×10^{18} ions/cm^3 necessary to create a robust ohmic contact.

The masking material and cleaning recipe used post-ion implantation are non-trivial decisions. For all devices fabricated in this thesis, a patterned photomask was used to protect the substrate from damage during ion implantation using the recipe detailed in Appendix A.1.3, and a cleaning recipe described in Appendix A.1.4 was used to remove cross-linked resist. However, if possible, using a hard mask may be favored over a photomask. This is because cross-linked resist after ion implantation is notoriously difficult to remove, even with aggressive cleaning procedures. The residual resist lowers yield of devices, often creating electrical discontinuities in on-chip interconnects. One way to use a hard mask process is by depositing thick SiO_2 and then etching ion implantation windows into the SiO_2. During the cleaning process after ion implantation, a long HF dip can be used to remove the SiO_2 from the substrate with minimal damage to the heterostructure surface since there is a silicon cap. Another way to utilize a hard mask is by depositing a thin screening dielectric (10–20 nm) before baking on the

photomask. This in effect achieves the same result as the thick SiO_2 hard mask: photoresist does not become crosslinked to the silicon surface and is removed by using HF during the post-ion implantation cleaning step. It is vitally important to remove all photoresist from the substrate before proceeding because following ion implantation, a 15 second rapid thermal anneal (RTA) using forming gas is performed to heal and repair the implanted region and create a uniform density of ions within the ohmic region. If photoresist remains on the substrate during this process it will become cross-linked and be nearly impossible to remove.

4. Deposition of the field oxide comes after ion implantation. The field oxide used is 15 nm of Al_2O_3 grown using atomic layer deposition (ALD) at 200°C. The thickness is chosen such that gate-to-2DEG current injection (or gate-to-2DEG leakage) above the ohmic region is suppressed. While samples presented in this thesis did not include a post-ALD forming gas anneal, it is recommended to do so if possible. This is discussed in further detail in Section 3.3.5. A cross-sectional schematic of the device is shown in Fig. 3.6 where the field oxide is shown as part of the device material stack.

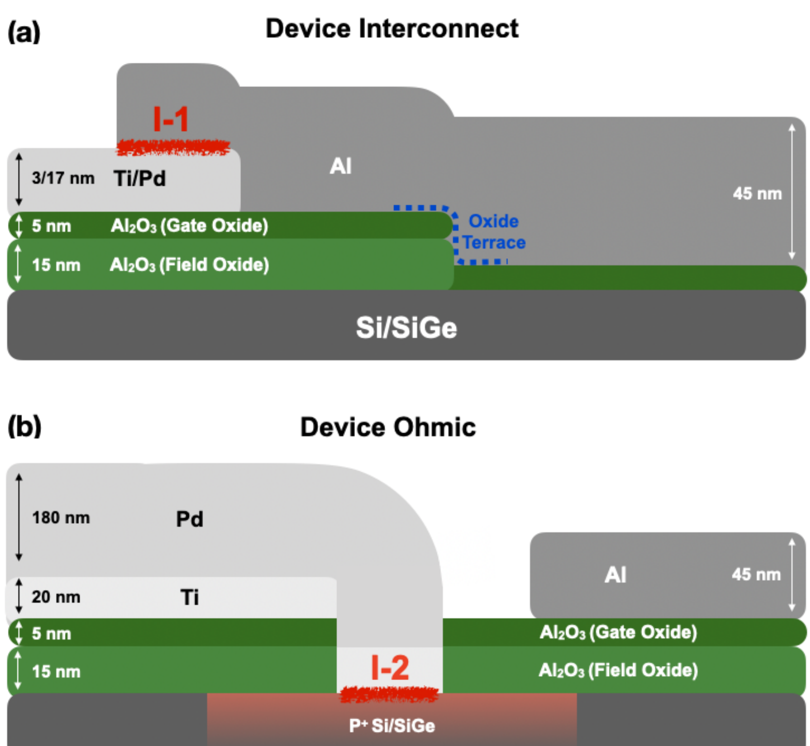

Figure 3.6: (Reproduced from Ref. [85] with permission from IOP Nanotechnology) Cross-sectional schematic of device interconnect jumper pad region and ohmic region. (a) The device interconnect is shown where the Al gate metal meets the Pd jumper pads. The field oxide is terraced via an etch such that it is removed in the active region and only the gate oxide remains. (b) The implanted region is shown to be contacted by Ti/Pd gate metal, forming the source-drain contacts. The aluminum gate that is used in accumulation is separated from the 2DEG by both the field and gate oxide for robust electrical isolation.

5. The terrace etch allows for a thick dielectric above the ohmic region but a thinner gate oxide beneath the gates in the active region, illustrated in Fig. 3.6(a). The thicker oxide over the ohmic region is used because the device is more susceptible to gate-to-2DEG leakage. An etch window is patterned using the process in Appendix A.1.3, then the wafer is dipped in 20:1 buffered oxide etchant (BOE) to etch away the Al_2O_3. Finally, the photoresist is stripped using the process in Appendix A.1.5.

6. The gate oxide deposition is much thinner than the field oxide deposition. This is important for a few reasons. First, there is evidence that charge noise increases with bulk dielectric near the active region [86]. Furthermore, as the distance between gates and the 2DEG is increased, it becomes more difficult to precisely control the chemical potential of the 2DEG. Thus, the gate oxide should be as thin as possible. While some devices fabricated on Si/SiGe contain only the native SiO_2 as a gate oxide, this was found to be a low-yield process in the Eriksson Lab. 5 nm of ALD-grown Al_2O_3 is a compromise between maximizing yield and minimizing the bulk dielectric, which is what was chosen for this fabrication process. The gate oxide is grown using the same ALD process as the field oxide and is shown schematically in Fig. 3.6.

7. After deposition of both the field and gate oxides, all blanket dielectrics for this fabrication process have been completed and etch windows can be opened for ohmic contact. While some fabrication processes may combine the ohmic via etch with ohmic metallization, it is preferred to separate these out into two separate lithography steps. This is because metallization (and the subsequent lift-off process) should be done with negative resist that undercuts during development. The undercut helps break continuous metallic films during evaporation and improves the quality of lift-off. On the other hand, resists that undercut are more susceptible to delamination, and etchants such as 20:1 BOE may etch unintended areas beyond the patterning area. For this reason, a positive resist is used for the ohmic via etch using the process in Appendix A.1.3. It is noted that HMDS is used to promote adhesion between the resist and substrate surface in this step to limit the possibility of delamination. The field and gate oxide above the ohmic

region are etched away using 20:1 BOE as illustrated in Fig. 3.6(b). After removal, there is a window or via above the ohmic which can then be electrically contacted. Finally, the resist is stripped using the process in Appendix A.1.5.

8. This is the first metallization step, and the thermal budget of the device significantly decreases afterwards. Thus, it is important for all high temperature processes (including anneals) to be performed before this step. By metallizing the ohmics, electrical contact to the 2DEG can be made via a wire bond to the bond pad contacting the ohmic. Negative resist AZ-5214 is used for all patterning steps involving lift-off for reasons discussed above, using the process in Appendix A.1.2. Before metal deposition and after patterning the ohmic via windows, the sample is dipped in 20:1 BOE for 20 s to remove the SiO_2 that grows natively on the substrate surface above the ohmic contact region. This step is required, otherwise the native oxide will electrically isolate the deposited metal contacts from the implanted regions. After the BOE dip, the sample is quickly transported into an electron-beam system (within 15 minutes) and the main chamber is pumped out below atmosphere so that the silicon on the ohmic surface no longer oxidizes. The electron-beam system used for deposition of metal is a Lesker PVD 75 electron-beam evaporator. The pressure in the chamber during deposition is kept at $<1 \times 10^{-6}$ Torr. For the ohmic metallization step, Ti/Au or Ti/Pd is typically used. Here, Ti/Pd is used because it slightly increases the thermal budget of the device. The Ti is used as an adhesion layer and is evaporated at $0.3\,\text{Å}\,\text{s}^{-1}$ to be 5 nm thick. 40 nm of Pd is evaporated afterwards at $1.0\,\text{Å}\,\text{s}^{-1}$. Following deposition of the ohmic metal, lift-off is achieved using the process in Appendix A.1.6.

9. The Pd jumper pads are evaporated next. This step is necessary when fabricating devices using Al gates since it is difficult to electrically contact Al wires. Al grows a native oxide so it must be ion milled or etched *in situ* during an evaporation step. This is further complicated for these devices because the Al gate thicknesses range from 30–65 nm. Since the selectivity of ion milling Al vs. AlO_x is high (i.e., Al mills much faster than AlO_x), it is difficult to ion mill the Al gates without partially milling through them. The next best solution is to have thin jumper pads that the Al gates can electrically contact. It is also noted that jumper pads are not necessary if the fabrication process is reordered such that the active region is patterned first and the bond pads are patterned last. Although this is a standard way of fabricating quantum dot devices, fabricating all interconnects and bond pads before the active region is a better approach for overlapping aluminum-gate devices because an on-chip electrostatic discharge (ESD) protection circuit can be implemented before patterning the active region—the most sensitive component of the device. A cross-sectional schematic of the Pd jumper pads is illustrated in Fig. 3.6. Al gates comprising the active region are evaporated later on in the process and electrically contact the jumper pads. Since the jumper pads are evaporated to be only 20 nm thick, the Al gate metal does not experience step coverage issues when contacting them. The Ti/Pd is evaporated to be 3/17 nm thick using the same recipe as discussed above.

10. The final photolithography step in this process is the deposition of bond pads and shorting wires comprising the ESD protection circuit. The bond pads are evaporated as a 20/180 nm stack of Ti/Pd using the same recipe as discussed above. Fig. 3.7 shows a completed device with an ESD protection circuit in-

tegrated on-chip. Fig. 3.7(a) is an optical image of a full device with 40 total bond pads. By patterning many devices at the wafer-level with 40 bond pads, a flexible interconnect design is fabricated, allowing for selection of wires used for devices when patterning the active region. Using the overlapping gate architecture, quadruple quantum dots can easily be accommodated. The ESD protection circuit is observed as the outer ring around the bond pads, where all bond pads are electrically tied together on-chip. Fig. 3.7(b) shows the device after it has been fabricated and packaged onto a printed circuit board (PCB), where the shorting wires have been physically disconnected using a diamond scribe attached to a wire bonder. This process allows for the device to be electrically grounded at all stages of fabrication and packaging.

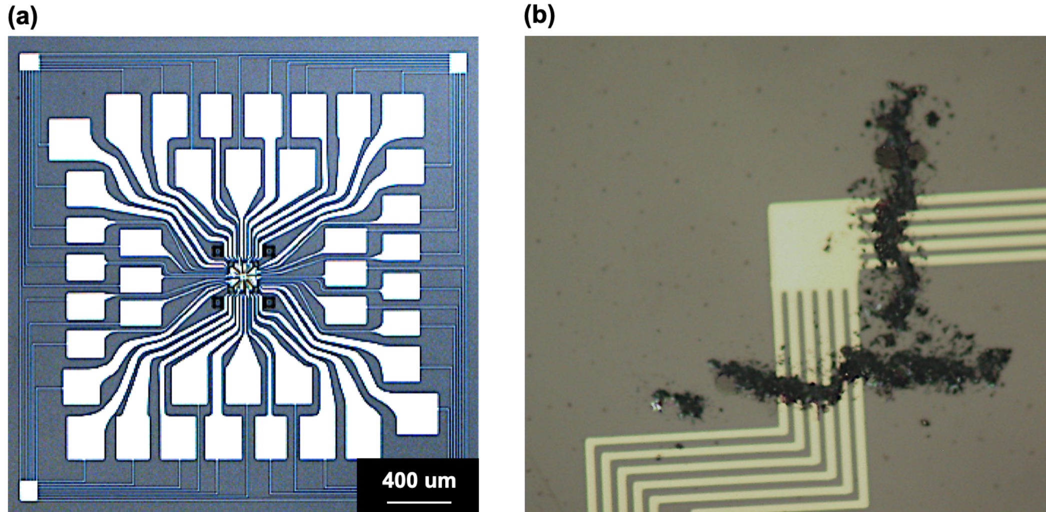

Figure 3.7: Completed device with electrostatic discharge (ESD) protection circuit. (a) Optical image of a full device with 40 total pond pads. By patterning many devices at the wafer-level with 40 bond pads, a flexible interconnect design is fabricated, allowing for different device geometries and sizes. Using the overlapping gate architecture, quadruple quantum dots can easily be accommodated. The ESD protection circuit is observed as the outer ring around the bond pads where all bond pads are electrically tied together on-chip. (b) After fabrication and packaging of the device onto a printed circuit board (PCB), the shorting wires are physically disconnected using a diamond scribe attached to a wire bonder. This process allows for the device to be electrically grounded at all stages of fabrication and packaging.

11. The final step at the wafer-level is dicing the wafer into chips to prep for e-beam lithography. An image of a completed wafer using the above process is shown in Fig. 3.8. For this step, S-1813 resist is spun onto the wafer at 4000 RPMs for 30 s to protect from contamination caused by the dicing saw. The wafer is then diced into 11 mm × 11 mm chips such that up to 16 devices can be patterned per chip at the e-beam level.

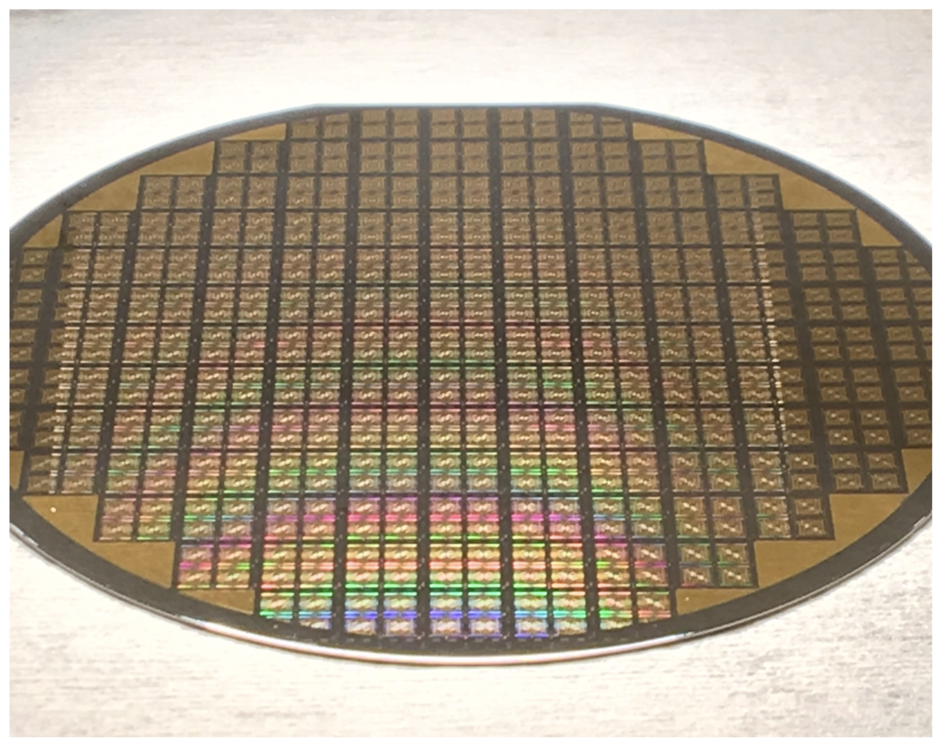

Figure 3.8: Wafer after all photolithography steps have been completed. The wafer is later diced into 11 × 11 mm dies such that up to 16 devices can be patterned per chip using e-beam lithography.

3.2.2 Nanoscale fabrication

Electron-beam lithography

After all photolithography steps have been completed and the wafer has been diced into chips, the chips can be patterned using electron-beam (e-beam) lithography. The e-beam system at UW Madison is an Elionix ELS G-100 system and is used for all e-beam lithography steps in the fabrication of the quantum dot devices in this thesis. While the Elionix system has wafer writing capabilities, I prefer to write on chips such that device designs can be quickly iterated. Since the full wafer photolithography process is long and Si/SiGe wafers are expensive/time consuming to grow, if possible, it is best to fabricate an interconnect structure that can accommodate many different

device designs which can be iterated through at the e-beam level.

The Elionix ELS-G100 system is a state-of-the-art lithography system that is capable of patterning features as small as 6 nm. When the alignment protocol is properly used, consistent overlay accuracy of better than 10 nm is possible, and often the alignment is within 5 nm of desired overlay. The Elionix system uses a 100 kV electron beam, limiting proximity dose effects. It allows for flexible on-chip alignment marks, making it simple to pattern alignment marks at the photolithography level (done during the bond pad step in the process described above). Additionally, the Elionix system comes with a range of wafer mounts (2”–6” in size) and chip mounts, allowing for reliable patterning of small and odd shaped substrates.

While the Elionix may pattern features as small as 6 nm, this is only possible when using ultra-thin resists (<30 nm). For the process described here, a three layer overlapping aluminum gate device is fabricated with gate layer thicknesses of 30, 45, and 60 nm for the three layers, respectively. Evaporating gates as thick as 60 nm is not possible when using thin e-beam resists, because metal on the substrate will form a continuous film with metal on the resist surface, causing problems during lift-off. Instead, PMMA 495 A4 resist is used, which is spun onto the substrate at 5500 RPMs, resulting in a thickness of approximately 180 nm. Fig. 3.9 shows patterning tests, pushing feature linewidths and gaps to the limit using this resist recipe. Fig. 3.9(a) demonstrates Al gates patterned as narrow as 55 nm without failure. In Fig. 3.9(b), Al gates are patterned as close as 25 nm without failure, but when the spacing is closer than 25 nm, resist wall collapse is observed. A visual signature known as “flagging” of gate metal is observed in the far right structure, which occurs when the gap between features is too small, resulting in improper lift-off of aluminum.

Effects of cold developer and cold-stage evaporation

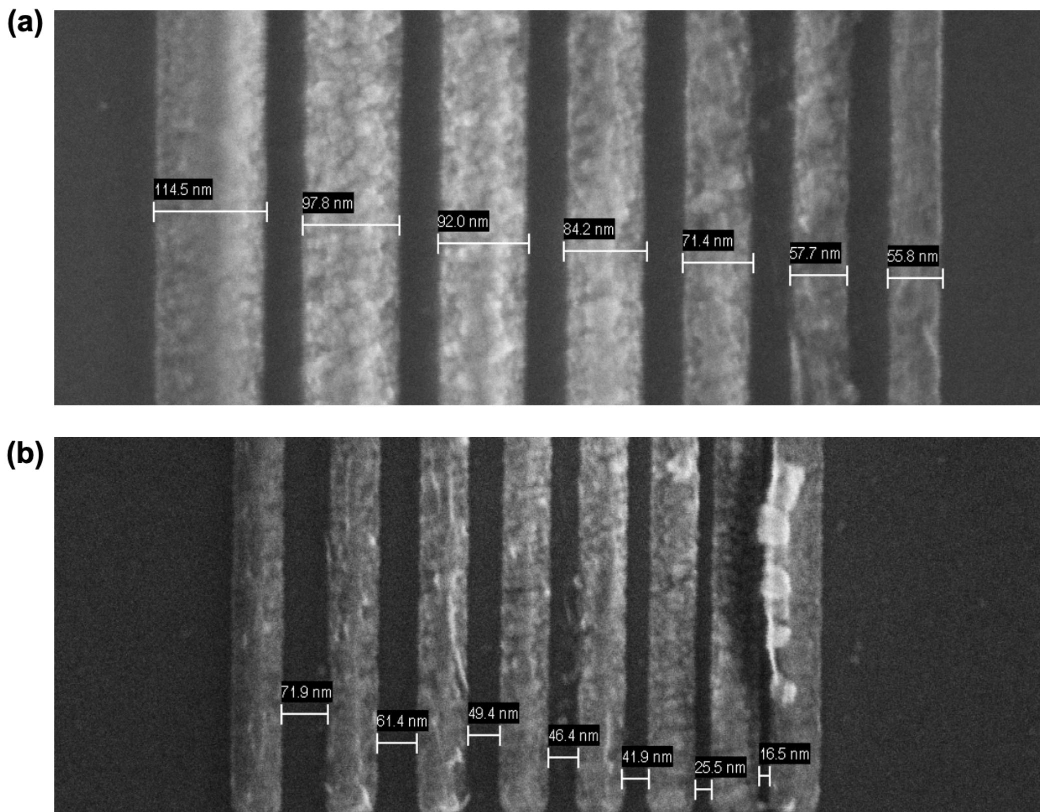

Figure 3.9: Pushing the limits of the Elionix using PMMA 495 A4 resist. (a) Linewidth test, showing Al structures ranging from 115 nm down to 55 nm using resist intended to be 180 nm thick. By using thick resists, Al gates as thick as 65 nm with 55 nm linewidths can be lifted-off with high yield. (b) Resist wall collapse test. Below 25 nm, resist wall collapse is observed. A visual signature known as "flagging" of gate metal is observed in the far right structure, where aluminum is not properly lifted-off due to the small gap between features.

There are many different techniques that can be used to enhance pattern transfer and improve metal deposition during nanofabrication of overlapping aluminum-gate quantum dot devices. During the e-beam resist development process, exposed resist chemically interacts with the developer. The exposed resist becomes soluble due to the e-beam exposure breaking the resist polymer chains into smaller polymers, which dissolve preferentially in developer solution over the larger ones.

Chilling the developer used, or using a "cold" developer process can be useful for increasing the quality of pattern transfer. By decreasing the temperature of the developer, the development rate is reduced and the e-beam dose needed to dissolve exposed resist increases. By increasing the dose, the resolution of features may be increased because partially exposed regions of the resist that would normally develop at room temperature remain undeveloped in the cold developer process.

One of the main purposes of using cold developer is to limit the proximity effect that is prevalent in e-beam lithography. When primary electrons enter the resist from the e-beam, a forward scattering process occurs. This process immediately limits the feature resolution and is highly dependent on the resist thickness, type of e-beam resist used, and e-beam energy. After the primary electrons have completely penetrated the resist, backscattering of primary electrons and emission of secondary electrons from the substrate occurs. Both of these processes work to expose unintended regions of the resist, leading to the proximity effect. By using cold developer, proximity dosed regions of the resist are less likely to be dissolved in the developer than when using room temperature developer.

While cold developer was not used in the samples presented in this thesis, the proximity dose was still problematic and was dealt with using a different technique. The proximity effect is most problematic when large exposed features neighbor small features. This creates a large background dose nearby the larger feature, which can alter the dose needed for the smaller feature. This can be dealt with in a number of different ways, including software that calculates and compensates for proximity dose, reducing larger feature size (if possible), or using cold developer. All three techniques were explored as a solution for overlapping aluminum-gate devices. Luckily, the larger feature size (reservoir gates, discussed in more detail below) could be reduced without

impacting device performance. However, cold developer may be extremely useful for other device geometries.

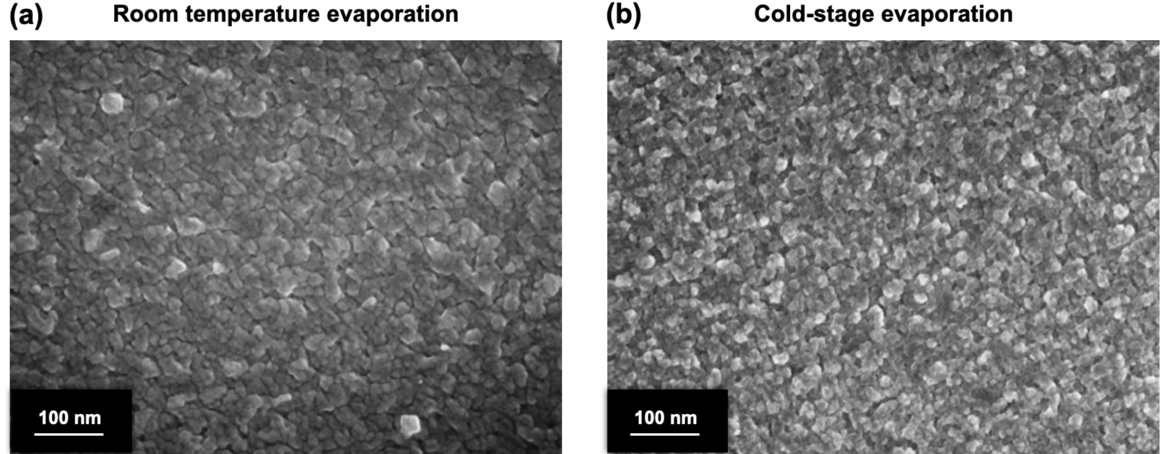

Figure 3.10: Comparison of room temperature vs. cold-stage evaporated aluminum. (a) SEM image of e-beam evaporated aluminum while the substrate stage is held near room temperature. A number of heat shields block the stage from the evaporator hearth, keeping the stage only slightly above room temperature. The average aluminum grain size is estimated to be approximately 40 nm. (b) SEM image of e-beam evaporated aluminum while the substrate stage is cooled by liquid nitrogen. The cold stage limits the size of aluminum grains. The average aluminum grain size is estimated to be approximately 20 nm when the stage is cooled to 77 K.

Cold-stage evaporation is another technique that can be used to improve the uniformity of metal depositions and overall quality of the gate stack for quantum dot devices. Cold-evaporation was not used for the devices presented in this thesis, but it was explored and found to improve certain aspects of devices. Ultimately, this technique was not used because of the difficulty of integrating the process into the e-beam evaporator used for metal depositions.

In particular, for aluminum depositions, cold-stage evaporation reduces the grain size, leading to more uniform gate electrodes. Fig. 3.10 compares the grain size of e-beam evaporated aluminum when using a room temperature stage and a stage cooled

to 77 K using liquid nitrogen. Fig. 3.10(a) shows an SEM image of e-beam evaporated aluminum while the substrate stage is held near room temperature. A number of heat shields block the stage from the evaporator hearth, keeping the stage only slightly above room temperature. The average aluminum grain size is estimated to be approximately 40 nm. Fig. 3.10(b) shows an SEM image of e-beam evaporated aluminum while the substrate stage is cooled by liquid nitrogen. The cold stage limits the size of aluminum grains. The average aluminum grain size is estimated to be approximately 20 nm when the stage is cooled to 77 K. Since gate electrode widths are often patterned to be 50 nm or less for quantum dot devices, a grain size of 40 nm can be problematic and cause issues with device yield.

Nanolithography process

By patterning all interconnect steps using photolithography at the wafer-scale, the number of e-beam lithography steps can be limited to only 3 steps for overlapping aluminum-gate devices. This increases the ability to iterate on device designs and geometries in the active region. Below, the fabrication process at the e-beam level is described in detail.

1. After dicing the wafer in 11 mm $\times$ 11 mm chips, each chip houses 16 devices that can be patterned using the Elionix e-beam lithography system. The chips are mounted to a stage specialized for substrates in the form of chips or pieces, fastened on by metallic clips. Alignment marks that are compatible with the e-beam system are patterned onto the chips during a photolithography step (either the bond pad step or ohmic metallization step). Using Au alignment marks produces the highest contrast between alignment marks and the silicon substrate,

but Pd is also a good choice for alignment mark material. Al alignment marks are discouraged since the contrast with Si is very low due to the similarity in atomic number.

The first aluminum gate layer in the active region is the screening gate layer. The screening gates are used to screen stray electric fields from subsequent layers and prevent formation of spurious dots. They also are used to lithographically define the position of quantum dot arrays or sensor dots, and can effectively be used as depletion gates. Generally, the e-beam dose for each layer is roughly the same (around $1400\,\mu\mathrm{C\,cm^{-2}}$); however, each layer is always dose tested before final devices are patterned. This is because drift of key parameters can occur on a weeks or even days timescale, such as the age of resist, developer, beam current, etc. The active region is always written using a $250\,\mu\mathrm{m}$ write field to produce the best feature resolution.

The screening gate layer is patterned using the process in Appendix A.2.1. An SEM image of the screening gate layer is shown in Fig. 3.11(a). The screening gate layer consists of two large triangular gates and a central dividing gate, shown as S1–S3. The screening gates act to deplete the 2DEG beneath them, defining quantum dot channels for the triple-dot array on the bottom side and the sensor dot on the top side. The screening gate layer is metallized using e-beam evaporated Al at $0.3\,\text{Å}\,\mathrm{s^{-1}}$. The Al is evaporated to be 30 nm thick such that it is thicker than the Ti/Pd interconnects (20 nm). This is to ensure yield problems due to step coverage failure are reduced. Next, lift-off is performed using the process in Appendix A.2.2, completing the pattern transfer. Finally, the aluminum gates are further oxidized beyond the native oxidation that occurs

in air using a plasma oxidation technique, detailed in Appendix A.2.3. This is discussed further in Section 3.3.

2. The second layer, which is functionally an accumulation gate layer, contains plunger gates (P1–P3), a sensor dot gate (M1), and reservoir gates (R1–R4). The plunger gates define quantum dots beneath them and control their electron occupation. The reservoir gates accumulate reservoirs of electrons that can tunnel into dots defined beneath the plunger gates. This layer is also patterned using the process in Appendix A.2.1, shown in Fig. 3.11(b). It is noted that the accumulation gate layer patterns the largest area of the three active region layers, thus cold developer or other techniques discussed above that limit the proximity effect can be beneficial. The second layer is evaporated to be 45 nm thick such that step coverage issues that may arise between the screening gate layer and accumulation gate layer are mitigated. Finally, the same lift-off (Appendix A.2.2) and plasma oxidation (Appendix A.2.3) processes are used to complete the second layer.

3. The third and final nanolithography step patterns the depletion gate layer, or barrier gates (B1–B4, T1–T2). Barrier gates work to deplete the 2DEG beneath them and control the tunnel coupling between dots and reservoirs. The aluminum gate metal is deposited to be 60 nm thick such that it can overlap both the first and second layers without electrical discontinuities forming. An SEM image of the barrier gate layer is shown in Fig. 3.11(c). Identical processes are used in this layer as the previous two, except the plasma oxidation is excluded. The final device with all three layers is shown in Fig. 3.11(d). Nearly the entire substrate is covered in gate metal, allowing for precise control of the 2DEG chemical potential

in all parts of the active region.

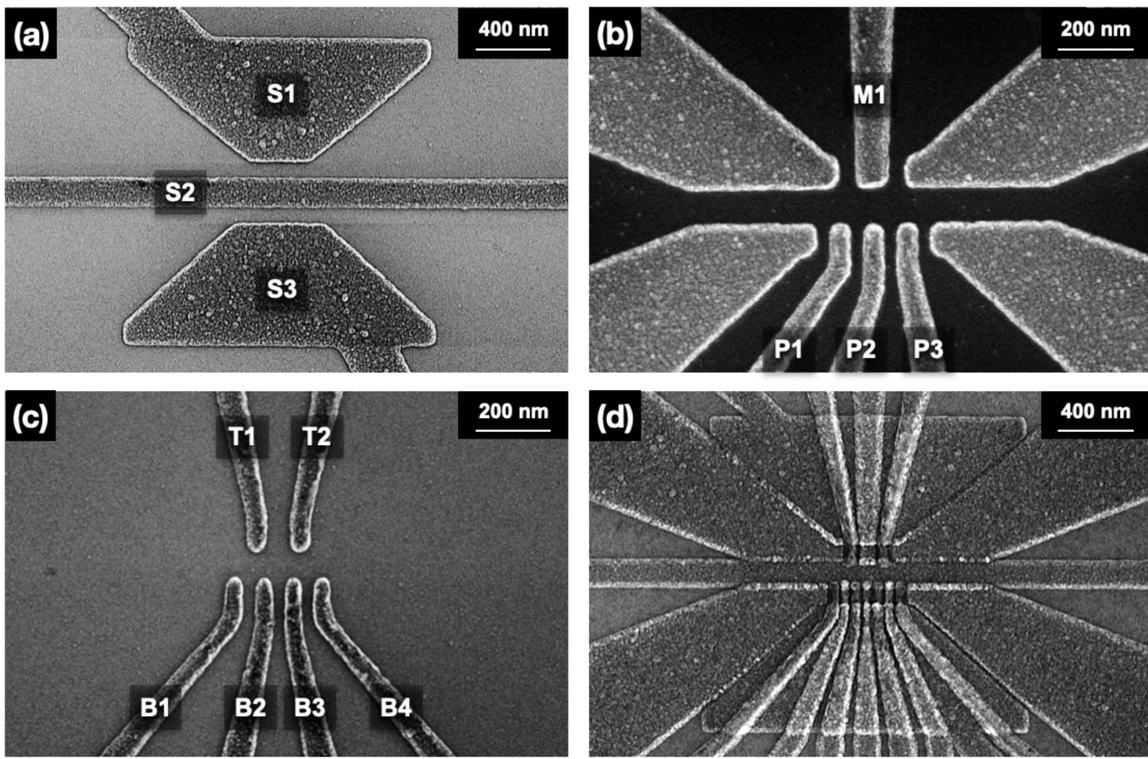

Figure 3.11: Individual gate layers for overlapping aluminum-gate devices. (a) The screening gate layer (S1–S3) consists of two large triangular gates and a central dividing gate. The screening gates act to deplete the 2DEG beneath them, defining quantum dot channels for the triple-dot array on the bottom side and the sensor dot on the top side. (b) The accumulation gate layer is the second layer, which contains plunger gates (P1–P3), a sensor dot gate (M1), and reservoir gates (R1–R4). The plunger and sensor dot gates define quantum dots beneath them and control their electron occupation. The reservoir gates accumulate reservoirs of electrons that can tunnel into dots defined beneath the plunger gates. (c) The third and final layer is the barrier gate layer (B1–B4 and T1–T2), or depletion layer. Barrier gates work to deplete the 2DEG beneath them and control the tunnel coupling between dots and reservoirs. (d) A full device SEM shows all three layers in the active region. Nearly the entire substrate is covered in gate metal, allowing for high control of the 2DEG chemical potential in all parts of the active region.

3.3 Yield engineering of overlapping aluminum-gate architecture

The following section is an Accepted Manuscript from IOP Nanotechnology and is reproduced here with permission. The manuscript published by IOP Nanotechnology is Ref. [85].

3.3.1 Introduction

In this section, an improved fabrication process for overlapping aluminum gate quantum dot devices on Si/SiGe heterostructures is presented, which incorporates low-temperature inter-gate oxidation, thermal annealing of gate oxide, on-chip electrostatic discharge (ESD) protection, and an optimized interconnect process for thermal budget considerations. This process reduces gate-to-gate leakage, damage from ESD, dewetting of aluminum, and formation of undesired alloys in device interconnects. Additionally, cross-sectional scanning transmission electron microscopy (STEM) images elucidate gate electrode morphology in the active region as device geometry is varied. The overlapping aluminum gate layers are shown to conform homogeneously to the topology beneath them—independent of gate geometry—and critical dimensions in the gate geometry are identified where pattern transfer becomes non-ideal, causing device failure.

Developing a suitable physical system for quantum computation has received much attention in the past two decades. Since Loss and Divincenzo's proposal [7], significant progress has been made using spins in solid-state systems as qubits. Coherent control of semiconductor quantum dots using spin degrees of freedom was first demonstrated

in $GaAs/Al_{0.3}Ga_{0.7}As$ heterostructures [16, 17]. This particular heterostructure found initial success due to the small electron effective mass in GaAs and depletion mode operation of devices. This allows for large, single-layer gate geometries to tune devices into the few-electron regime [87]. Although fabrication and characterization of one-qubit devices in GaAs has become routine [13, 16, 17, 19, 88], short coherence times of spin-qubits due to the presence of nuclear spins [89] make it difficult to achieve fidelities necessary for fault-tolerant operation [22].

Silicon-based approaches have significantly improved qubit performance in solid-state spin systems in part due to the spinless nucleus of ^{28}Si. Experiments using isotopically purified ^{28}Si have shown average single-qubit control fidelity in excess of 99.9% [32]. However, in Si, the lithographic demands are more stringent because of the larger electron effective mass. Different gate designs have been explored, including open geometries, which use global top gate(s) and several depletion gates to form each quantum dot [29, 90] and tight geometries which use linear, overlapping gates with dedicated accumulation and depletion electrodes for each quantum dot [45, 79]. While both general designs have generated two-qubit devices in Si/SiGe [34–36] and Si-MOS [33, 91], the overlapping gate architecture has clear advantages in scaling to larger systems. A 9-dot array has already been demonstrated [45] and similar architectures can be scaled into much larger arrays. Additionally, quantum dots form in predictable locations with tunnel couplings that can be well-controlled using a single gate [92, 93]. While there are benefits in choosing a linear, overlapping gate architecture, developing a high yield fabrication process is challenging.

In this section, many yield limiting steps in the fabrication of linear, overlapping aluminum gate quantum dot devices are investigated, showing failure analysis of critical interfaces and providing improved process steps for difficulties experienced in typical

fabrication procedures. Results are presented on three main topics: low-temperature oxidation of inter-gate aluminum oxide (AlO_x), cross-sectional scanning transmission electron microscopy (STEM) analysis of overlapping Al gate geometries, and characterization of interconnects between the device bond pads and active region—the local region surrounding quantum dots. Four low-temperature oxidation techniques are compared for enhancement of the native AlO_x that electrically isolates subsequent gate electrode layers from each other. STEM analysis investigates test structures with varying dot-to-dot pitch, characterizing how different gate geometries affect gate electrode morphology and the filling of barrier gates in gaps between plunger gates. For the interconnects, the process flow is optimized to allow for thermal annealing of the Al_2O_3 grown by atomic layer deposition (ALD), incorporation of the low-temperature oxidation techniques presented, and integration of on-chip electrostatic discharge (ESD) protection.

3.3.2 Device electrostatic characterization

Fig. 3.12(a) shows a cross-sectional schematic of the quantum dot channel in Fig. 3.11(d), where screening gates S2 and S3 border the channel. The width of all plunger/barrier gates is increased to 100 nm before crossing the back edge of the screening gates in order to minimize step coverage failure, which is also shown in Fig. 3.11(d). In this particular geometry, plunger gates are nominally 70 nm wide with a 120 nm pitch, and barrier gates are 60 nm wide, filling a 50 nm gap. The device is cooled in a dilution refrigerator with a base temperature of <50 mK and electron temperature $T_e = 100 \pm 10$ mK. Fig. 3.12(b, c) show charge stability diagrams in the few-electron regime between each pair of adjacent quantum dots P1/P2 and P2/P3, respectively, as measured by an in-

tegrated charge sensing dot beneath M1. M1 is biased into a configuration such that it is on the highest sloped region of a Coulomb blockade peak, where changes in charge occupancy of quantum dots P1-P3 result in detectable shifts in current through the sensor dot [94]. The differential conductance dg_{M1}/dV_P is measured using standard lock-in techniques as detailed in Ref. [11], where the voltage on both plunger gates are modulated simultaneously. High resolution plots were taken over a 24 hour period to demonstrate stability in the few-electron regime.

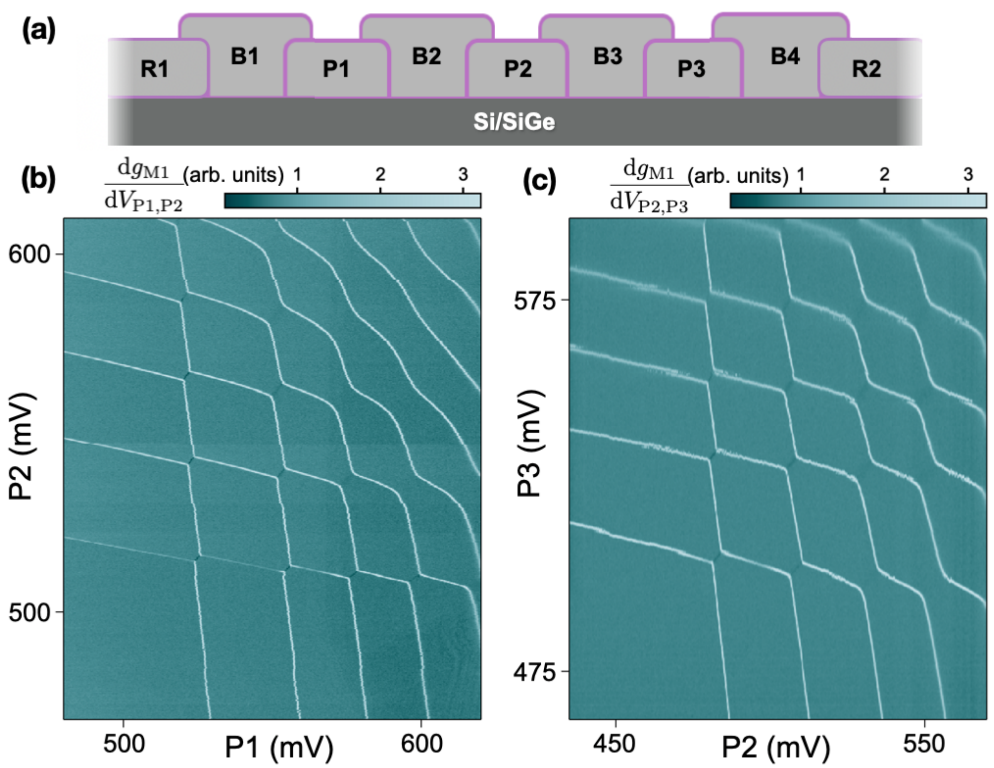

Figure 3.12: (a) Cross-sectional schematic of the gate stack in the quantum dot channel between gates S2 and S3, where the light purple border around the Al electrodes is native AlO_x. The screening gates are not shown because they border the quantum dot channel. (b, c) Stability diagrams of plunger gate pairs P1/P2 and P2/P3 where the differential conductance dg_{M1}/dV_P beneath charge sensing dot M1 is measured by modulating the voltage on both plunger gates simultaneously, demonstrating electrostatic control of the few-electron regime in each quantum dot pair.

3.3.3 Low-temperature inter-gate oxidation

Gate layers are electrically isolated via the AlO_x that grows natively on the Al gate electrodes. This allows for the omission of blanket dielectric films such as ALD-grown Al_2O_3 typically present between gate layers, which may cause increased charge noise in Si/SiGe quantum dot devices [86]. Removal of grown dielectrics for electrical isolation comes at a cost though—one must rely on the AlO_x that grows natively on evaporated Al, which is reported to grow between 1.6–3.0 nm by a variety of techniques [80, 95–99]. This makes high yield fabrication of devices difficult due to gate-to-gate leakage and damage from ESD.

In order to determine if this native oxide is sufficient to prevent leakage, four different oxidation techniques of Al gates are compared: native oxidation (NO), thermal annealing at 250 °C (TO), plasma ashing (PA), and UV-ozone treatment (UV). All samples are natively oxidized at standard temperature and pressure before a 15 minute treatment by each oxidation method. For TO, the anneal is at 250°C and 45% humidity. For PA, the plasma asher used is a YES R3 Downstream Plasma Cleaner at a pressure of 5 Torr with 80 sccm O_2 using a power of 250 W. For UV, a Samco UV-1 model is used with an oxygen flow rate of $0.5\,\mathrm{L\,min^{-1}}$, giving an ozone concentration of $6\,\mathrm{g\,m^{-3}}$ with its substrate platen heated to 250 °C.

To characterize how each oxidation treatment affects the electrical isolation properties of the Al gates, two-layer test structures are fabricated and designed to replicate the overlap between plunger/barrier layers and the screening gate layer of the device shown in Fig. 3.11. The test structure device design includes the portion of the plunger/barrier gate that climbs onto the screening gate lying beneath it to ensure this rugged interface was included as part of the breakdown test. The first layer is oxidized

using one of the four methods, and subsequently the breakdown voltage is measured using standard current-voltage measurements. Devices are fabricated on [100] Si wafers, electrically isolated from the Si by 100 nm of ALD-grown Al_2O_3. The two layers are patterned using EBL. 30 and 50 nm of Al are deposited for the first and second layer, respectively. The structures have an overlap area of 1 μm × 0.1 μm, similar to the gate overlap in Fig. 3.11(d). Each electrode layer forming the test structures is electrically connected to bond pads, as described in detail in Section 3.3.5, which is used to apply a differential voltage to the device and measure leakage current using a Keithley Model 2700 Multimeter, current limited to 5 nA. Devices are tested cryogenically at 2 K by increasing the differential voltage between electrode and counter-electrode pairs until breakdown is observed, defined here to be when 100 pA of gate-to-gate leakage is measured (a measured current density of 0.1 A cm^{-1}). Devices are inspected after measurement in an SEM to ensure they have not been destroyed by ESD.

Fig. 3.13(a) shows the results of the breakdown test. For each of the four oxidation methods, 10, 41, 11, and 35 samples are measured for NA, TO, UV, and PA, respectively, to determine an average breakdown voltage, shown as a solid data point. Adjacent to solid data points are the breakdown voltages of all devices measured for each method, shown as hollow, semi-transparent points. The average breakdown and standard deviations are summarized in Table 3.1. The spread and standard deviations measured for TO and PA are a more accurate representation of their true values due to the larger number of devices measured.

The minimum breakdown voltages observed in PA and UV, shown in column four of Table 3.1, are significantly above the estimated maximum differential voltage needed for overlapping Al gate devices in Si/SiGe (∼1–2 V), whereas it is only modestly higher for NO and TO. As device size is increased to include more and more overlapping

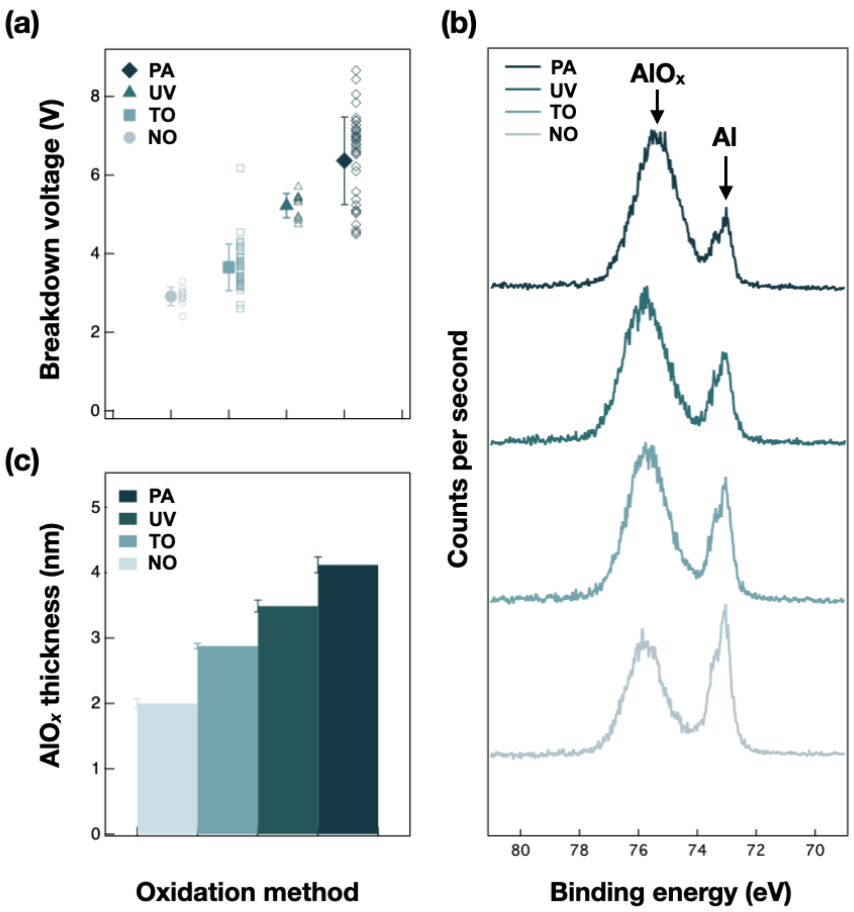

Figure 3.13: Low-temperature oxidation characterization of AlO_x. (a) Breakdown voltage data for native (NO), thermally annealed (TO), UV-ozone (UV), and plasma ashed (PA) inter-gate AlO_x. Solid data points show the average breakdown voltage and standard deviation measured for each method. Adjacent to solid data points are the breakdown voltages of all devices measured for each method, shown as hollow, semi-transparent points. (b) X-ray photoelectron spectroscopy (XPS) spectra taken for each method; traces offset for clarity. The AlO_x thickness is extracted from the relative intensities of the oxidic/metallic Al 2p peaks [100]. (c) Extracted thickness of AlO_x from XPS spectra.

aluminum gates, it becomes increasingly important that the native AlO_x is further oxidized to be more robust to electrical breakdown, thus PA and UV represent two low-temperature oxidation methods that mitigate failure due to gate-to-gate leakage

Table 3.1: Electrical characterization summary of low-temperature Al oxidation methods. Oxide thickness was determined from x-ray photoelectron spectroscopy (XPS). The breakdown voltage (BDV) is defined here to be when >100 pA of current is measured between electrode and counter-electrode pairs.

Method	Thickness (nm)	BDV (V)	Min. BDV (V)
PA	4.12 ± 0.12	6.36 ± 1.12	4.50
UV	3.49 ± 0.09	5.23 ± 0.31	4.75
TO	2.88 ± 0.04	3.65 ± 0.59	2.60
NO	2.00 ± 0.07	2.91 ± 0.24	2.80

at an acceptable level for larger devices. The distinct increase in breakdown voltage of the PA and UV techniques suggests increased AlO_x thickness and/or higher quality AlO_x. To isolate these variables, bulk Al films are analyzed using x-ray photoelectron spectroscopy (XPS) to determine oxide thickness. The XPS spectra are shown in Fig. 3.13(b). The AlO_x thickness can be extracted from the relative intensities of the oxidic/metallic Al 2p peaks, as detailed in Ref. [100]. There is a slight red shift observed in the AlO_x XPS peak of the PA samples which is likely due to PA being the thickest of the four films. Since the Al-O bonding energy is ∼0.6 eV lower than the Al-OH bonding energy [101], the surface layer of AlO_x is more hydroxyl rich than bulk AlO_x, and the red shift occurs simply because PA is thicker than the other films. The spectra shown in Fig. 3.13(b) are a subset of a larger dataset taken to determine the average thickness and fluctuations in thickness for each method. Six XPS spectra are taken at different positions on the bulk film for each oxidation method to obtain statistical fluctuations in the AlO_x thickness. It is noted that the AlO_x thickness of amorphous and crystalline oxidized Al are similar [99], so our bulk measurements are closely representative of the much smaller device gate electrodes, which are composed

of Al grains on the order of the gate electrode width. Fig. 3.13(c) shows the results from all measured samples, and the values are summarized in Table 3.1. The thickness for PA, UV and TO all exceed NO. PA displays the fastest oxidation rate.

Calculating a breakdown field of the AlO_x layer is not as straightforward as taking the ratio of columns 3 and 2 of Table 3.1, because of the complex oxidation behavior of the Al-AlO_x-Al stack. Nonetheless, it is useful to note that such a simple division would yield results of ∼12–15 MV cm^{-1}. For context, the breakdown values reported for *bulk* AlO_x in literature are ∼3–8 MV cm^{-1} [102–104]. There are three reasons for the significant difference between the previously reported bulk values and the results of the ratio of columns 3 and 2 of Table 3.1. First, column 2 reports the measured thickness of films that did not have Al evaporated on top of the oxide. The bottom part of Al electrodes has been shown to oxidize when in contact with oxygen-rich materials such as SiO_2 [105] and ALD-Al_2O_3 [106]. Thus, the AlO_x thickness between the two Al electrodes may increase upon deposition of the top Al electrode, presumably with a less oxygen-rich composition. Second, the AlO_x in these samples is in the ultrathin regime, and such films display an enhanced breakdown field [107]. Third, the substrate temperature is 2 K rather than room temperature for these measurements, and such a decrease in temperature can increase breakdown fields [104, 108]. All three of these mechanisms are likely to contribute to the large ratio of columns 3 to 2 in Table 3.1.

The oxidation rate of Al for both UV and PA follows a $d \sim \sqrt{t}$ dependence [95, 96], where d is thickness and t is time, which can be used to further increase the AlO_x thickness if desired. PA is implemented in the fabrication of the triple-dot device shown in Fig. 3.11. Low-frequency charge noise was measured in two devices using PA on separate chips, using the technique detailed in Ref. [48], obtaining values of 2.31 and 0.89 μeV Hz$^{-1/2}$ at 1 Hz where the noise power spectral density follows a $1/f$

dependence with exponents 1.03 and 1.04, respectively. This is consistent with other recent results for Si/SiGe quantum dots using ALD-grown Al_2O_3 as a gate dielectric [48, 86].

3.3.4 Transmission electron microscopy for failure analysis

To further investigate potential failure modes in overlapping Al gate devices, test structures are fabricated using PA between gate layers and take cross-sectional STEM images (Fig. 3.14). Two-layer test structures are fabricated to investigate filling of the gaps between plunger gates (first layer) by barrier gates (second layer) for varying gate widths.

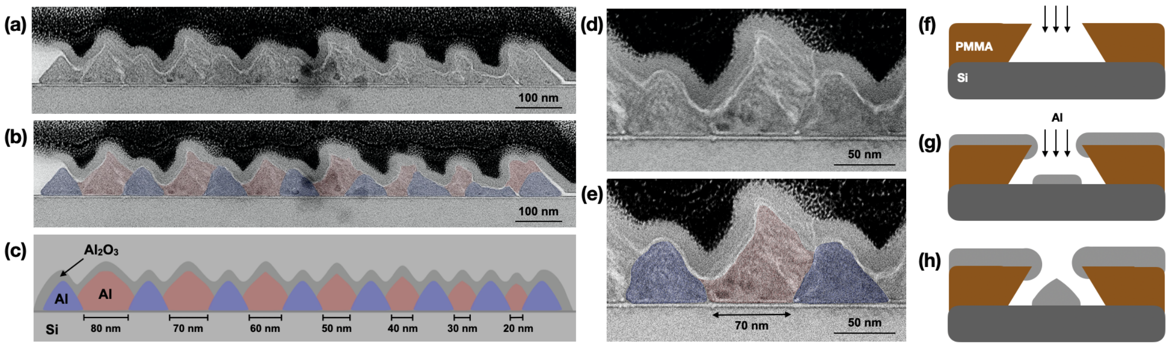

Figure 3.14: Failure analysis for varying gate widths. (a, b) Scanning transmission electron microscopy (STEM) images of an overlapping Al gate structure, where (b) has been false-colored. Al plunger gates (blue) are deposited nominally 50 nm thick and 70 nm wide on a Si substrate with gate-to-gate pitches varying from 90–150 nm. Al barrier gates (red) are deposited subsequently, nominally 65 nm thick with widths varying from 40–100 nm. The test structure is capped with 20 nm of ALD-grown Al_2O_3 (gray) and a protective platinum layer (black). (c) A schematic of the expected cross-section of the overlapping Al gate structure. The sidewalls are assumed to be $\sim 60°$ as is observed on average in the STEM image. (d, e) Zoom in of Fig. 3.14(a), where (e) has been false-colored. The 70 nm gap between plunger gates is filled homogeneously by the barrier gate. (f-h) Schematic showing before (f), during (g) and after (h) the Al e-beam evaporation. The sloped sidewalls observed for Al gates is due to accumulation of Al on PMMA sidewalls during evaporation [80].

The test structures are fabricated using the same procedure described for the device shown in Fig. 3.11(a-d), except the first layer is omitted. The evaporator hearth is water-cooled and copper radiation shielding is used to keep the sample stage near room temperature to reduce high-temperature induced morphological effects such as large grain size and sloped sidewalls of the aluminum [109]. The gate geometry is shown in Fig. 3.14(a, b), where the plunger gate layer consists of eight 70 nm wide, 50 nm thick gate electrodes with gate-to-gate pitches varying from 90–150 nm in steps of 10 nm (increasing right to left). This leaves nominal gaps 20–80 nm wide for the barrier gates to fill. The barrier gates are evaporated 20 nm wider than each gap, ranging from 40–100 nm (increasing right to left) and the thickness of the barrier gate layer is nominally 65 nm. A schematic of the expected cross-section is shown in Fig. 3.14(c), where the sloped sidewalls and the effect on gate morphology has been taken into account. This schematic can be compared to the region of interest of the triple-dot device, which is shown earlier in Fig. 3.12(a). In Fig. 3.14(b), the plunger gate layer (blue) and barrier gate layer (red) have been false-colored using the dark-field STEM image shown in Fig. 3.15 as a guide.

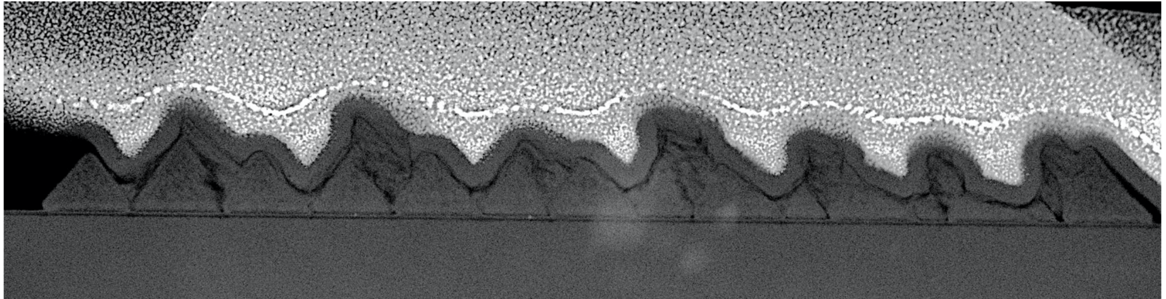

Figure 3.15: STEM image of overlapping aluminum gate test structure using the high-angle annular dark-field (HAADF) detector. This high contrast image was used to aid in identifying interfaces between plungers and barriers and false-coloring the plunger and barrier gate layers shown in the Fig. 3.14.

Several striking features are revealed from the STEM images. The sidewall slope of the plunger and barrier gate electrodes is found to be between 45–60°. This is consistent with AFM profiling shown in Ref. [109] and can be explained by the process illustrated in Fig. 3.14(f-h). During evaporation, Al accumulation narrows the opening of the PMMA mask. In extreme cases, the opening closes completely, leaving gate electrodes thinner than intended. This symptom is visible on the far right of Fig. 3.14(a, b), and can lead to yield problems associated with step coverage failure.

On the far right of Fig. 3.14(a, b), the plunger gates themselves are deformed and reduced in thickness. This is attributed to resist wall collapse [110] since the thickness is skewed to one side. Top-down SEM images (not shown) are consistent with such resist collapse. Resist wall collapse is an issue in these devices because the barrier gates are relatively thick (65 nm) to guarantee high-yield connection over the many underlying gates in this device. As a consequence, the PMMA resist is also chosen to be thick (180 nm), for high-yield lift-off.

The barrier gates fill and contact the underlying surface across all regimes, independent of the gap width. The only case in this test structure where a barrier gate is not in contact with the substrate is on the far right side, where the intended gap of 20 nm completely closed off. In the narrower regime, the barrier gate layer fails for a different reason than the plunger gate layer. The barriers do not hold their intended thickness and shape, due to the process illustrated in Fig. 3.14(f-h). Given the 45–60° sloped sidewalls, the thickness of the gate cannot reliably exceed the width using this recipe. A way to design around this is to pattern barrier gates as wide as possible when devices contain small gaps between plunger gates.

3.3.5 Device interconnects

Another challenge in fabricating quantum dots in Si/SiGe heterostructures is the design considerations needed for developing on-chip interconnects between bond pads and the active region. High temperature processes are desirable at different stages of device fabrication, but oftentimes they cause failure of interconnects. Below, critical interfaces affecting the thermal budget at various stages of fabrication, how to design around the thermal budget, and integration of on-chip ESD protection is discussed.

For the interconnect design, a mesa is etched between the bond pads and active region, shown in Fig. 3.16(b), preventing gate-to-ohmic leakage in the event that wire bonds punch through the substrate into the 2DEG. The yield of devices can be severely limited by damage due to ESD after the fabrication of the active region. To ameliorate this, on-chip ESD protection is implemented. Shorting wires are patterned before fabrication of the active region in the same lithographic step as bond pads, preventing build-up of charge between gates. The shorting wires can be seen Fig. 3.16(a) bordering the bond pads. An equipotential for all gates is maintained through device packaging by wire bonding, grounding through a printed circuit board, and physically scribing away the shorting wires on-chip afterwards. Alternatively, the leads can be electrically disconnected after fabrication using an etch step; however, this does not protect the device during packaging.

Consideration of the thermal budget is important when determining a process flow for quantum dot devices. Significant interdiffusion of the Si/SiGe interface at the quantum well can occur above 800°C [68], lowering the valley splitting [111]. This makes growth of high quality SiO_2 [112, 113] on Si/SiGe heterostructures difficult, so instead ALD is used to grow amorphous Al_2O_3 or HfO_2 for gate-to-2DEG isolation.

After the field and gate oxide is grown, thermal annealing can be used to reduce interface trapped charge [114, 115], which has been shown to reduce threshold voltages and increase transconductance in Si-MOS quantum dots [79].

The presence of interfaces (Fig. 3.16(c), I-1 and Fig. 3.16(d), I-2) further limits the thermal budget of devices, and the presence of thin Al films imposes its own set of thermal constraints, including dewetting and void formation, occurring at 400 °C [80] and between 300–500 °C [116], respectively. To maximize the thermal budget of devices, we choose palladium instead of gold for interconnect metallization. Au is often used, but formation of $AuAl_2$ (purple plague), a non-conductive alloy, is observed in thin Au-Al films at temperatures as low as 217°C on a minutes time scale [117]. Additionally, at the ohmic metallization site, the Au/Si interface experiences an interstitial diffusion process that affects thin film electrode morphology at temperatures as low as 200 °C [118]. A similar process is observed when annealing a Au/Ti/Si structure analogous to I-2 at 250 °C for 30 minutes. When using Pd, the thermal budget is found to increase. Formation of Pd-Al alloys in thin films near 300 °C [119] can cause electrical discontinuities, which we verified with a 30 minute anneal at 300 °C on a 30/17/3 nm Al/Pd/Ti gate stack. Additionally, we observe diffusion processes at 400 °C occurring at the Pd/Ti/Si interface, affecting ohmic gate morphology. The consequence of these critical temperatures is that they cannot be exceeded by processes such as post-metallization anneals or inter-gate oxidations with the expectation of high yield. However, the UV and PA methods presented in this paper do not exceed any of these critical temperatures when using Pd as an interconnect/ohmic gate metal (Fig. 3.16(b), inset), and thus are good choices for oxidation of inter-gate oxide.

3.4 Conclusion

In this chapter, an improved fabrication process was demonstrated and failure analysis was performed on many critical interfaces in overlapping aluminum-gate quantum dot devices. The main takeaway from the AlO_x characterization is that PA and UV mitigate failure due to gate-to-gate leakage and ESD, and their process temperatures are compatible with the thermal budget of overlapping Al gate quantum dot devices. Optimization of individual oxidation techniques, not considered here, may improve oxide quality and allow further tuning of the AlO_x thickness. For UV oxidation of Al, relative humidity, time of oxidation, temperature, and partial pressure of oxygen all affect the growth rate [96, 97, 120]. For PA, time of oxidation, excitation frequency, power, and partial pressure of oxygen affect the growth rate [95, 121].

Overlapping aluminum-gate two-layer structures were fabricated and analyzed using cross-sectional STEM imaging, which helped identify failure modes due to particular gate geometries. The sidewalls of the plungers/barriers were found to be between 45–60°, creating potential issues with step coverage near the active region. This can often be designed around by maximizing the barrier gate widths when devices have small gaps between plunger gates. Failure modes associated with critical dimensions in device geometries were also identified. Gates with <40 nm gaps between plungers showed abnormalities due to resist wall collapse and thinner than intended deposition thickness.

Additionally, we showed an interconnect fabrication process that implements on-chip ESD protection and identified critical temperatures that can result in electrical discontinuities at different stages of device fabrication. By designing a fabrication process compatible with these temperatures, field/gate oxides may still be annealed

at 450 °C in forming gas before metal deposition, and inter-gate oxide in Al devices can be further oxidized using UV or PA. The result of a process with these changes implemented was shown in Figure 3.12(b, c), where the quantum dots were notably stable and all gate electrodes worked as intended.

Over the years of fabricating these devices and performing a yield engineering study, many new insights were gained into improving future generations of devices. Some of these insights were not applied to the current generation of devices that are shown in this thesis. Based off of the well formed sensor dot (M1) on the top side of the device, it is possible to form dots using larger lithographic dimensions than the dots P1–P3. For the current generation of devices, the width of the channel formed between S1/S2, and S3/S2 is 90 nm on both sides. The width of M1 is 100 nm while the width of P1–P3 is 70 nm. Additionally, the width of the barrier gates adjacent to the sensor dot (T1 and T2) are 80 nm which fill a gap of 70 nm while barrier gates B1–B4 are 60 nm wide, filling a gap of 50 nm. Since one of the most common failures of these devices was barrier gate failure, an important question to answer is what is the ideal gap width and barrier gate width? For future generations of devices, it is suggested that the barrier gate gap for all quantum dots be increased to 60 nm, and the barrier gate width be increased to 85 nm. Additionally, the plunger gate widths for P1–P3 should be increased to 80 nm to allow for the wider barrier gates.

Although larger dots are not desired for ideal performance, robust fabrication is such an outstanding issue for quantum dots in Si/SiGe that the more important immediate task for the field is increasing the yield of devices so larger arrays of qubits can be studied. The devices shown in this thesis prove that larger quantum dots can be formed, and should be used in the immediate future. Once a robust fabrication has been established, techniques into reducing the size of dots should then be further

investigated.

Finally, it is noted that switching the ordering of lithographically defined layers may be advantageous. Instead of patterning the plunger gates and reservoir gates on layer two, it may be beneficial to change the ordering such that barrier gates are patterned in the second layer and plungers/reservoirs are patterned in the third layer. The reason for this is that the lithographic dimensions of the barrier gates are more demanding than all other layers, but by patterning barrier gates on the third layer, the metal deposition must be very thick ($\sim$60 nm). By moving the barriers to layer two, they could be patterned much thinner ($\sim$45 nm), creating a more robust device fabrication process.

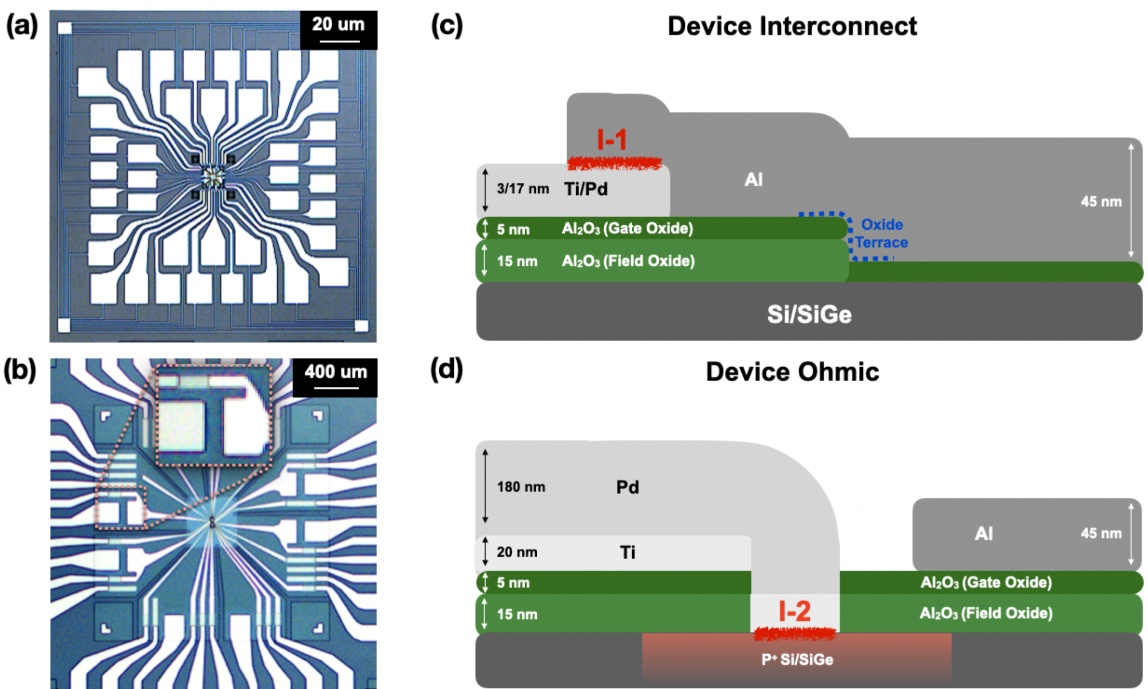

Figure 3.16: Interconnect and ESD protection architecture. (a) Optical image of full device. The electrostatic discharge (ESD) protection wiring borders the device, shorting all gate leads together during fabrication of the device. (b) Optical image of the device mesa. Interconnects between the active region and bond pads are joined by 3/17 nm thick Ti/Pd pads. (Inset) Top-down zoom-in of a Pd-Al interconnect and ohmic, corresponding to the cross-sectional schematic shown in (d). (c) Schematic cross-section (not to scale) of a device interconnect where the Al gate lead meets the Pd interconnect. An oxide terrace is used to reduce bulk oxide beneath the active region. At I-1, an interface between Al and Pd is formed, resulting in problematic alloys at ∼300 °C. (d) Schematic cross-section (not to scale) of a device interconnect at the ohmic site. An interface (I-2) between Ti/Pd and the P^+ doped Si/SiGe heterostructure forms. Diffusion processes are observed to begin at ∼400 °C, which can affect the interconnect morphology and electrical connectivity.

Chapter 4

QDscript

4.1 Introduction

Measurement of quantum dot devices requires the use of various types of electronics including the use of dozens of voltage sources. The size and geometry of devices dictate the number of voltage sources needed, but generally anywhere from 10–50 voltage sources are necessary for current architectures where the number of quantum dots used for both qubits and qubit detection are between 2–12. Due to this large overhead in controlling dozens of voltage sources, it is highly beneficial to have control software that is flexible and easy to use.

The Eriksson Lab began using Labber as a data acquisition platform in 2017. Labber comes with a suite of measurement and analysis tools, including instrument drivers, instrument communication via PyVISA, plotting tools, analysis tools, and a graphical user interface (GUI) for organizing groups of instruments into a single file. Additionally—and most importantly to this chapter—Labber includes a Python API that extends its functionality into Python. This allows for highly flexible measurement

setups that can be programmed via Python scripts.

QDscript is a python package written in the Eriksson Lab that creates a quantum dot measurement environment within Python that hooks into Labber. QDscript offers a host of measurement, analysis, and automation tools catered for quantum dot tune-up and measurement. Additionally, it uses Labber's measurement and plotting protocol with very low latency, enabling all of the features of Labber while also maintaining a convenient scripting environment for measuring quantum dot devices.

4.2 Package structure

4.2.1 Initializing a script

Labber has three main components: the instrument server, the measurement editor, and the log browser. The instrument server allows for communication between instruments and Labber via GPIB, serial, USB, TCP/IP, or any other interface that is supported by the VISA standard. Instruments can be named, configured and imported into the measurement editor. The measurement editor defines a measurement file that contains a subset of the active instruments from the instrument server. Within each measurement file, "step channels" and "log channels" can be defined. Step channels are instrument values that are swept or stepped whereas log channels are instrument values that are measured. The ranges of step channels are configured for each given measurement from the measurement file. When the measurement begins, Labber uses PyQt5 to show the measurements in a live plotting view. When the measurement is completed, a log file is created and displayed in the log browser, which contains the measured data and swept quantities. The log browser is a library of measurement files

that has embedded analysis and plotting tools.

Labber contains a module named `ScriptTools` which utilizes Labber's Python API such that Python scripts can call internal Labber functions. One of Labber's most convenient features is the live plotter. A template script for running a measurement via Python that uses the live plotter is shown below.

```
ScriptTools.setExePath("\path\to\ScriptTools")
meas_object = ScriptTools.MeasurementObject("file_path_name", "
    log_file_path_name")
meas_object.setMasterChannel("log_channel_name")
meas_object.setOutputFile("log_file_path_name_out")
meas_object.performMeasurement(return_data=False)
```

In this example, `"file_path_name"` is the absolute path of the measurement configuration file as a string variable, `"log_file_path_name"` is the path of the output log file as a string variable, `"log_channel_name"` is the name of a log channel defined in the Labber measurement configuration file, and `"log_file_path_name_out"` is the name of the log file.

This script is simple and useful, but only applies when the step channel range has already been defined in the measurement file, and a single log channel is to be measured. However, step channel values can also be set by the measurement object in Python

```
meas_object.updateValue("step_channel", 0.0, "SINGLE")
meas_object.updateValue("step_channel", 0.0, "START")
meas_object.updateValue("step_channel", 1.0, "STOP")
meas_object.updateValue("step_channel", 101, "N_PTS")
```

where the step channel in the measurement file named "step_channel" sweeps from 0.0 to 1.0 in 101 steps with a point spacing of 0.01 in this example. In this way, step

channels and log channels can be defined and measured solely from a Python script.

4.2.2 QDscript measurement environment

For more complex situations, catering a scripting environment to quantum dot devices significantly improves measurement efficiency. QDscript creates that environment by creating simple plotting functions, instrument control functions, and embedded analysis tools. Additionally, QDscripts improves measurement efficiency in Labber by performing a number of automated tasks before and after each plot to ensure the measurement file is up to date. This removes the painstaking task of manually updating all control voltages in the measurement file by hand every time a measurement is taken.

When setting up a script in QDscript, there are a few imports and functions that make up the header which must be run every time. Luckily, these parts of the script can be left unchanged while the rest of the script changes depending on step channel and log channel parameters. The header is shown below.

```
1 ScriptTools.setExePath("\path\to\ScriptTools")
2 scenario = Scenario(file_name="\path\to\measurement\file")
3 ProjSettings.__init__("user", "project", "msName", "fileExt", "msPath
    ","msFile", scenario)
```

`ScriptTools` has already been discussed. The second line initializes the scenario object where `Scenario` is a class that contains all information pertaining to the Labber measurement file. `ProjSettings` is a static class that is used for initializing project data into the QDscript environment. It is also used for persistent data that lasts between script sessions so quantum dot parameters can be measured and used in later scripts.

The backend of QDscript contains four core functions that are used often. They are `set()`, `setnow()`, `get()`, and `plot1d()`. `set()` is a utility function that sets a measurement

file instrument value. It is noted that this is not the same thing as directly setting the instrument value. QDscript is set up in this way because setting file values takes a very short amount of time compared to setting instrument values, and oftentimes many instruments (in particular voltage values) are being set for each script. This allows QDscript scripts to take only a fraction of a second from the time it begins to when the first Labber plot runs. Then, before the measurement begins, all instruments are set to the corresponding file values simultaneously using Labber. In other words, `set()` enables batch writing of instrument values from scripts. `setnow()` sets the instrument value directly, but as discussed above, may take a significant amount of time to run depending on the communication protocol of the instrument. This allows for flexible setups where instrument values *can* be directly set, but when initializing many different instruments before a plot, `set()` can be used to reduce the overhead. `get()` is used to get any instrument value by name, and is useful when setting up more complex scripts. For example, many tuning algorithms use feedback from previously measured data sets when deciding what parameters to use for the next tuning sequence.

Finally, `plot1d()` is the core plotting function of QDscript. While there also exists `plot2d()`, `plot3d()`, and `plotNd()`, they are only wrapper functions for `plot1d()`, which is the master plotting function in QDscript. `plot1d()` takes in tuples defining the dimensionality of the scan, allowing this single master function to be run every time a plot is taken. This is useful since other utility functions can be built into `plot1d()` which are then guaranteed to run every single time a measurement is taken. For example, the measurement file instrument values that are used at the beginning of a scan can be updated and set to the final instrument values after the plot has finished. This enables a measurement file that is automatically kept up to date with current instrument values, eliminating the need to edit values by hand each time a new measurement is taken.

An example script for initializing, setting two instrument voltages named "P1" and "P2" to 1.0 V, setting a third voltage named "P3" to 0.5 V, and then plotting "P3" from 0.5–1.0 V is shown below.

```
ScriptTools.setExePath("\path\to\ScriptTools")
scenario = Scenario(file_name="\path\to\measurement\file")
ProjSettings.__init__("user", "project", "msName", "fileExt", "msPath
    ","msFile", scenario)

set("P1;P2", 1.0)
set("P3", 0.5)
plot1d("P3", 0.5, 1.0, 101, None, "log_channel")
```

As shown in line 5, multiple different step channels can be specified by name in a single line. This also applies to plotting where any number of step channels can be concatenated using a semicolon delimited string. These plotting and set functions are the backbone of QDscript. Additional functionality for analyzing data sets and performing tuneup algorithms is discussed in the next sections.

4.2.3 Data analysis in QDscript

In Labber, data files are saved in a .hdf5 format, named "log files." While .hdf5 files are flexible and use a dictionary-based format, it is difficult to reliably import data from these log files since dictionary keys change from file to file. These dictionary keys include the data set dimensionality, number of step channels, number of log channels and whether or not data was taken using a buffer. Buffers are a temporary storage unit in instrument memory that allow for blocks of data to be stored on the instrument and read in bulk to reduce data acquisition time. In a Labber log file, this results in "trace data", which is saved in a different format than data acquired point-by-point.

QDscript contains a class named `LabberDataset` that imports log files and automatically organizes the data set depending on dictionary keys in a log file. The class automatically detects the dimensionality of the data set, the number of step and log channels, and whether or not data is in a trace format.

Another class named `QDdataObject` is used in tandem with `LabberDataset` to perform user-friendly data analysis. An instance of `QDdataObject` can be created by using an instance of `LabberDataset` as an initialization parameter to produce a data object in Python. This data object contains familiar data attributes such as "xdata", "ydata" and "zdata" for a given data set. All data arrays are converted into `numpy` arrays.

Below is an example of plotting a log file from Labber using QDscript and `matplotlib`
.

```
1 from matplotlib import pyplot as plt
2 dataset = analysis.LabberDataset(file_name)
3 data_object = analysis.QDdataObject(dataset)
4 plt.plot(data_object.xdata, data_object.ydata)
5 plt.show()
```

An instance of `LabberDataset` is created, importing the Labber log file into a Pythonic object. Using the `QDscript.analysis` module, a `QDdataObject` is initialized as `data_object`, from which analysis can be easily performed. Labber log file data is reproduced in Line 4 using `matplotlib`. The `data_object` contains attributes `xdata` and `ydata` that are plotted. By default, `QDdataObject` automatically assigns the xdata and ydata attributes to the first step channel and log channel, respectively. However, if there are multiple step channels or log channels, these data arrays can be selected using the built-in methods `get_stepchannel_data("step_channel_name")` and `get_logchannel_data("log_channel_name"` `)`. In this way, more complex log files can be dealt with using `QDdataObject`. Addition-

ally, since all `QDdataObject` arrays are `numpy` arrays, fitting procedures can naturally be applied using curve fitting libraries such as `scipy`.

4.3 Automated tuning

Since QDscript builds an environment allowing for full instrument control and data analysis from Python, complex scripts involving automation tasks can be utilized. There are several useful automation functions embedded in QDscript that help ease the quantum dot tune-up process. A cross-capacitance matrix can be measured and persistently stored using function `get_comp_vals("P1", "M1", ...)`, where in this example a capacitance value between P1 and M1 is measured. This cross-capacitance matrix can then be used to dynamically compensate gate voltages while measuring stability diagrams using the `comp` flag from `plot1d()` as shown below

```
1 plot1d("P1", 0.5, 1.0, 101, None, "log_channel", comp="M1")
```

Now, the voltage on gate M1 will dynamically compensate while P1 is swept from 0.5 to 1.0. This is highly useful when trying to optimize the sensitivity of the sensor dot M1 while other gates are being swept. Without compensation, the optimal gate voltage value for M1 quickly shifts due to the cross-capacitance from P1.

Other automation functions that are often used include `retune_charge_sensor("M1", ...)`, `hysteresis_check("step_channel", ...)`, `get_electron_temperature(...)`, and a host of others which all help tune devices quickly while measuring and extracting important information and parameters from the device.

Using these core measurement and analysis functions, QDscript helps enable auto-tuning of double dot devices *in situ* using machine learning [46]. Machine learning was successfully used to tune a quantum dot device into the double dot regime. QDscript

helped close the feedback loop between the machine learning program and the measurement tasks. To reduce the amount of total data needed to complete the double dot tuning algorithm, rays of data in the stability diagram phase space are taken and pieced together by the machine learning program. QDscript enables the ray scans by allowing for compensation between different gates that can be manually swept. QDscript passes the data into the machine learning program, the machine learning program identifies gate voltages that need to be adjusted to form a better double quantum dot, and then QDscript repeats the scans with new parameters. In this way, automated tuning of double dots was accomplished.

4.4 Future work

QDscript has become a powerful Python scripting environment in tandem with Labber for measuring and controlling quantum dot devices. Since QDscript is written in Python, it is easily accessible to the scientific community, where Python is one of the leading programming languages. In the future, “cold-start” auto tuning—where a device is tuned into the few electron regime without any manual tuning—could be accomplished by developing a few more simple algorithms. Identifying turn-on thresholds and optimal screening gate voltages by detecting key values in wall-wall plots are all that remains in the cold-start tune-up automated procedure.

Chapter 5

One- and two-electron valley-orbit states

Section 5.1 and 5.2 of this chapter are reprinted from arXiv:2103.14702.

5.1 How valley-orbit states probe quantum well interfaces

5.1.1 Introduction

The energies of valley-orbit states in silicon quantum dots are determined by an as yet poorly understood interplay between interface roughness, orbital confinement, and electron interactions. Here, we report measurements of one- and two-electron valley-orbit state energies as the dot potential is modified by changing gate voltages, and we calculate these same energies using full configuration interaction calculations. The results enable an understanding of the interplay between the physical contributions

and enable a new probe of the quantum well interface.

The ability to make uniform and tunable qubits is crucial for large-scale applications. Modern computers use one control electrode per field effect transistor with excellent uniformity, and proposed architectures for quantum chips also rely on a small number of control lines per qubit, in order to minimize the density of control wires [122, 123]. Progress has recently been made enhancing the homogeneity of the electrical environment by using quantum dot designs that eliminate modulation doping and instead make use of metal surface electrodes to both accumulate and deplete electrons [45]. The resulting structures enable good control over electron occupation, gate voltages, and tunnel couplings between quantum dots with a small number of gate electrodes per quantum dot [35, 37, 85, 93, 124–127].

Uniformity remains a challenge with regard to conduction band valley energies in silicon [128, 129], and important physical questions need to be addressed. Atomistic disorder is known to play a particularly important (and typically uncontrolled) role in determining the energies of electrons at the bottom of the valleys [111, 130–134], resulting in a wide range of observed valley splittings in Si/SiGe quantum dots (20–270 µeV, [24, 52–60]). Critically, the interplay between the factors that determine the valley splitting in quantum dots — the atomic details of the interface (which vary with lateral position), the degree of lateral confinement, and electron-electron interactions within a quantum dot — are not yet fully understood.

This chapter reports quantitative characterization of the relationship between low-lying one- and two-electron valley-orbit states and the quantum dot confinement strength, shape, and position. The pulsed-gate spectroscopy and magnetospectroscopy measurements reveal valley splittings in the range 36–87 µeV, two-electron singlet-triplet splittings between 22–59 µeV, and orbital splittings that can be tuned from 1.69–2.26 meV.

Simulations combining full configuration interaction (FCI) [135] with empirical tight-binding (TB) theory [136] are shown to be in good agreement with the experimental results, and together these methods enable an understanding of the interplay between effects arising from quantum well interface roughness, orbital confinement strength, and electron-electron (e-e) interactions. This combination of experiment and theory not only explains the origin of the energy spectrum but also provides a new method for probing the quantum well interface.

5.1.2 Experimental setup

Spectroscopy of one- and two-electron valley-orbit states is performed in a device fabricated using a three-layer overlapping aluminum gate architecture [45], as shown in Fig. 5.1(a). A detailed fabrication process can be found in Ref. [85]. The integrated sensor dot under gate M measures the electron occupation of the central quantum dot under gate P2. The triple-dot on the bottom side is tuned into a regime where B1 and P1 form a large tunnel barrier on the left side of P2, suppressing the tunnel rate into reservoir R3. Gates P3 and B4 extend the reservoir R4 into the quantum dot channel, as shown by the electron density heat map in Fig. 5.1(b), allowing for the tunnel barrier beneath gate B3 to tune the tunnel coupling between the P2 dot and right reservoir under R4. The electron temperature is measured to be $T_e = 100\,\mathrm{mK}$.

One- and two-electron valley-orbit splittings are measured by pulsed-gate spectroscopy using the experimental setup shown in Fig. 5.1(c). A square voltage pulse with amplitude V_{pulse} and frequency f_{pulse} is applied to gate P2, pulsing the ground state of the quantum dot between two levels: E_{Lg} and $E_{\mathrm{Ug}} = E_{\mathrm{Lg}} + \alpha V_{\mathrm{pulse}}$, where $E_{\mathrm{Lg}}(E_{\mathrm{Ug}})$ denotes the ground state of the dot in the loading(unloading) position, and

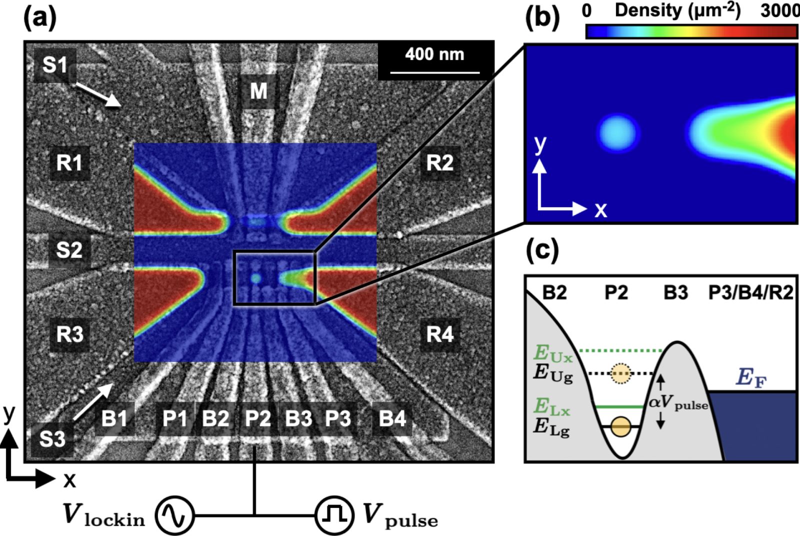

Figure 5.1: Device layout and experimental setup. (a) A scanning electron microscope (SEM) image of a device lithographically identical to the one measured shows the gate electrode layout in the active region. A COMSOL Multiphysics Thomas-Fermi simulation of the electron density is overlaid on the SEM image. The sensor dot under gate M measures the average charge occupation $\langle n \rangle$ of the P2 quantum dot via lock-in amplifier detection at frequency f_{lockin}. (b) The COMSOL simulation shows the charge density of the P2 dot, where the tunnel barrier to the left(right) reservoir is opaque(transparent). (c) Valley-orbit state splittings are measured using pulsed-gate spectroscopy. A square pulse with amplitude V_{pulse} at frequency f_{pulse} is applied to gate P2, rapidly pulsing the chemical potential of the P2 dot between E_{Lg} and E_{Ug}. The change in chemical potential induces detectable shifts in the tunnel rate into and out of the P2 quantum dot, allowing for measurement of excited state energies.

α is the lever arm for gate P2. A measurable change in the average electron occupation of the dot $\langle n \rangle$ occurs when an excited state provides an additional channel for the electron to enter the dot, yielding a measurement of the energy of this excited state [12, 137, 138].

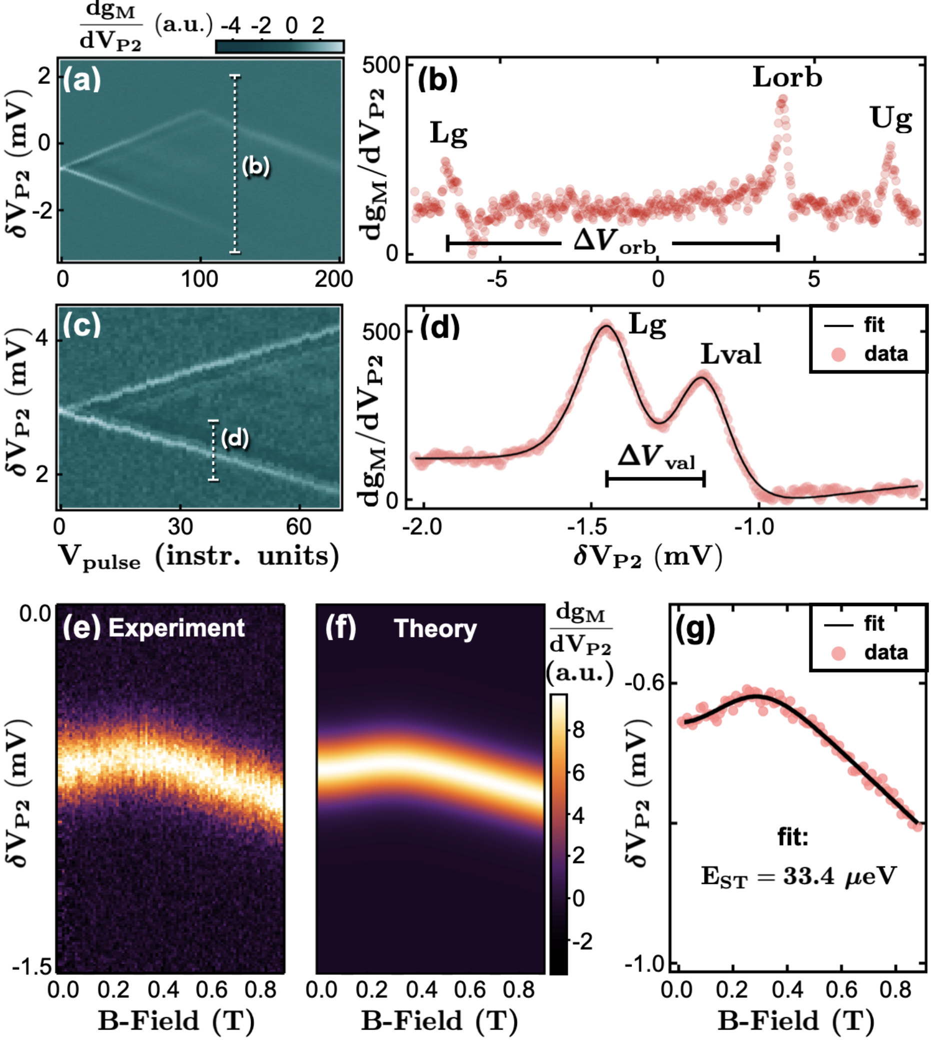

Figure 5.2: Experimental methods (a, b) Measurement of orbital splitting using pulsed-gate spectroscopy. The pulse amplitude is increased until the lowest-lying orbital state is visible, shown in (a). A linecut (dashed line) is taken when the pulse amplitude exceeds the orbital splitting, allowing detection of the ground (Lg) and excited orbital (Lorb) states, shown as the peaks in (b). The orbital splitting is given by $\alpha\Delta V_{\mathrm{orb}}$. (c, d) Using the same method, the valley (E_{val}) and singlet-triplet splittings (E_{ST}) are measured. The red data points clearly show two distinct peaks in (d). The data is fit using Eqn. 5.1, shown as the black line, giving $E_{\mathrm{val}} = \alpha\Delta V_{\mathrm{val}} = 53.2\,\mu\mathrm{eV}$. (e–g) Experimental magnetospectroscopy data (e) is reproduced in (f) by treating the dot-reservoir system as a grand canonical ensemble. The experimental peak locations in voltage space are extracted from (e) and plotted in (g) as red circles, which are then fit to Eqn 5.2, yielding $E_{\mathrm{ST}} = 33.4\,\mu\mathrm{eV}$.

5.1.3 Measurement of valley-orbit states

Fig. 5.2(a, b) show pulsed-gate spectroscopy measurements of the one-electron E_{orb}. The excited orbital state is separated well enough from the ground state such that each peak position is found by fitting to $\cosh^{-2}(\alpha V_{P2}/2k_BT_e)$ [69], where $E_{\text{orb}} = \alpha\Delta V_{\text{orb}}$. As shown in Fig. 5.2(c, d), a similar procedure is used to measure the valley splitting. In this case, there is overlap of the ground and lowest excited state signals that arises from thermal broadening. To extract the peak locations, we make use of an expression for $\langle n\rangle$,

$$\langle n\rangle = \sum_{i=g,x} \Gamma_i \frac{e^{(E_i - E_F)/E_{0i}}}{e^{(E_i - E_F)/k_BT_e} + 1}, \tag{5.1}$$

where $E_i = \alpha V_i$ is the position of each peak in energy, and E_{0i} and Γ_i are fitting parameters for each peak. Eqn. 5.1 is derived in Section 5.2.1. As shown by the solid line in Fig. 2(d), the experimental data are fit by the derivative of Eqn. 5.1 with respect to the gate voltage ($d\langle n\rangle/dV_{\text{P2}}$), enabling extraction of $E_{\text{val}} = \alpha\Delta V_{\text{val}}$.

The two-electron singlet-triplet splitting E_{ST} is measured using both pulsed-gate spectroscopy (using a similar procedure as above) and magnetospectroscopy [54, 55, 139, 140], as shown in Fig. 5.2(e–g). A full theoretical model for magnetospectroscopy is developed in Section 5.2.2 which allows for the experimental data shown in Fig. 5.2(e) to be closely reproduced by the model in Fig. 5.2(f). This model enables fitting the peak position of the data in Fig. 5.2(e), using

$$V_{P2}(B) = \frac{1}{\alpha\beta} \ln \left(\frac{e^{\frac{1}{2}\kappa B + \beta_e E_{\mathrm{ST}}} \left(e^{\kappa B} + 1 \right)}{e^{\kappa B} + e^{2\kappa B} + e^{\kappa B + \beta_e E_{\mathrm{ST}}} + 1} \right), \tag{5.2}$$

where V_{P2} is the gate voltage, $\kappa = g\mu_B\beta_e$ where $\beta_e = 1/k_B T_e$, g is the electron g-factor, μ_B is the Bohr magneton, and B is the magnetic field. The peak positions from Fig. 5.2(e) are extracted and plotted as red circles in Fig. 5.2(g), and the fit (solid line) to Eqn. 5.2 yields for this example $E_{\mathrm{ST}} = 33.4\,\mu\mathrm{eV}$.

5.1.4 Comparison between measured values and COMSOL simulations

Using the spectroscopic techniques from Section 5.1.3, E_{val}, E_{ST} and E_{orb} are measured as a function of electrostatic confinement in the x-y plane. The confinement is varied by changing the S3 gate voltage (V_{S3}) between 260–420 mV while compensating with neighboring barrier/plunger gates to maintain a constant electron occupation and tunnel rate into the dot.

Fig. 5.3 shows the effects of changing V_{S3} on quantum dot orbital energy, shape, and position. The device schematic pictured in the top left inset shows the approximate location of the P2 dot as the shaded blue region. Changes to the voltage V_{S3} applied to the screening gate, shown as the shaded gray region in the inset, modify the P2 dot confinement, orbital shape, and position. The minimum orbital splitting, plotted as solid circles in Fig. 5.3, is found to be non-monotonic with V_{S3}, because of the strong effect S3 has on the electrostatic confinement of the dot along the y-axis. At high V_{S3}, the y-confinement becomes weak, and the minimum orbital splitting falls off rapidly due

to elongation of the dot along the y-axis. Toward the center, the dot becomes isotropic, increasing the minimum orbital splitting up to 2.26 meV at $V_{\mathrm{S3}} = 370\,\mathrm{mV}$. At low V_{S3}, suppression of the minimum orbital splitting occurs due to the compensating barrier and plunger gate voltages needed to stay in the one-electron regime, elongating the dot along the x-axis. Since low V_{S3} creates a tight confinement potential along the y-axis, the x-orbital is the most weakly confined orbital below $V_{\mathrm{S3}} = 370\,\mathrm{mV}$.

This behavior is well explained by semiclassical (Thomas-Fermi) electrostatic simulations using COMSOL Multiphysics, shown in the bottom inset of Fig. 5.3. The position and orbital shape are simulated at the four points indicated by the shaded purple circles from the experimental data. The outline of each oval represents the shape that encloses 50% of the electron wave function, and the small circles show the center of the electron density for each simulation. The Thomas-Fermi (TF) simulations qualitatively match the experimental findings, where the x-orbital is most weakly confined at low V_{S3} and the y-orbital is most weakly confined at high V_{S3}. In addition to the changes in shape, significant change in the position of the dot is observed in the COMSOL simulations. Over the range shown, the quantum dot position slides down and to the left as V_{S3} increases for a total change in position of 27.5 nm. We note that COMSOL simulations were not used to determine the dot shape and position above 390 mV, because this regime becomes close to the accumulation threshold for the S3 gate, where uncertainties become large.

5.1.5 Measurement and simulation of valley-orbit splittings

Atomic steps in the quantum well interface play an important role in determining E_{val} and E_{ST}. Fig. 5.4(a) plots E_{val} (filled blue squares) and E_{ST} (filled red circles) as the

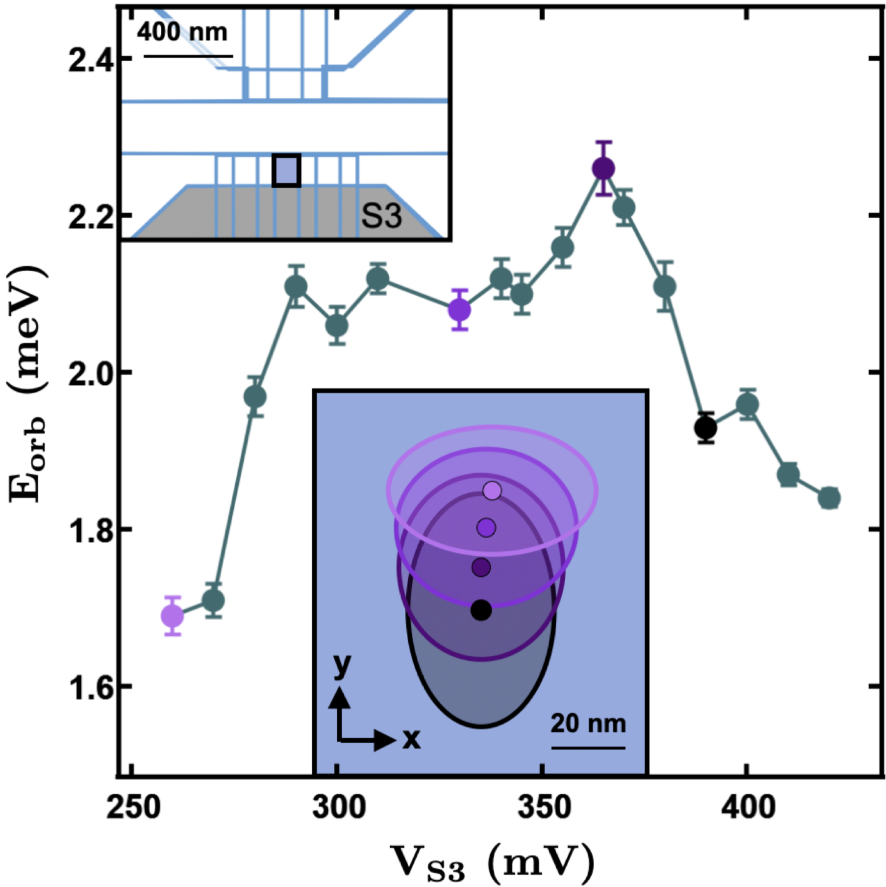

Figure 5.3: Quantum dot orbital shape and position. The top left inset shows a device schematic where the P2 quantum dot is located within the filled blue region and the S3 gate is shown as the filled gray region. The minimum orbital splitting is plotted as a function of V_{S3}. At low V_{S3}, the dot has strong y-confinement and weak x-confinement, leading to a drop in the minimum orbital splitting. As V_{S3} is increased, the dot becomes isotropic, increasing the minimum orbital splitting. At high V_{S3}, the dot has weak y-confinement and strong x-confinement, reducing the minimum orbital splitting again. Thomas-Fermi electrostatic simulations are used to calculate the quantum dot shape and position within the blue region shown in the device schematic. Four points are simulated, shown as the shaded purple circles in the experimental data. The change in shape of the electron density is found to be in agreement with the experimental data. Additionally, the position of the quantum dot for each simulation is illustrated as the small solid circle in the inset, showing that the quantum dot slides down and to the left 27.5 nm.

orbital shape and dot position are varied with V_{S3}, where the shaded region represents the measurement uncertainty. The measured E_{val} and E_{ST} change substantially across the electrostatic configurations, and their large *in situ* tunability arises from motion of the P2 quantum dot with respect to atomic steps at the quantum well interface. As the dot approaches an atomic step, more of the wave function overlaps the step, thereby increasing the valley-orbit coupling which suppresses E_{val} and E_{ST} [62]. To determine the position of atomic steps relative to the quantum dot, we make use of a combination of the experimental measurements, the COMSOL simulations just described, and FCI calculations, the latter of which incorporate valley-orbit coupling that arises from interface roughness. The FCI calculations use a single fitting parameter in order to match the measured values of E_{val} and E_{ST} namely, the position of the dot relative to atomic steps at the quantum well interface.

Figure 5.4(a) shows the results of FCI calculations for E_{val} and E_{ST} as open blue squares and open red circles, respectively. Close agreement with the experimental data is found when the atomic steps are separated by 35 nm. The positions used in the FCI calculations to produce the best fit to the data at each point are plotted in Fig. 5.4(b) as solid colored points. Each point has a colored gradient overlay that represents the spatial extent of the singlet (ground) state with respect to the atomic steps, which are shown as black dashed lines. The gradient steps from darkest to lightest, corresponding to wave function probability thresholds of 75%, 50%, and 25% of its maximum value.

A linear fit to the dot positions calculated from the FCI simulations reveal a total change in position of the dot of 24.7 nm, which is remarkably close to the 27.5 nm change in position extracted from the COMSOL simulations, given that these two methods are based on different physical inputs. This correspondence can also be seen by comparing the FCI results to the open purple circles in Fig. 5.4(b), which shows the central dot

position extracted from the four COMSOL simulations. The close agreement provides a measure of validation for an approach that allows an *in situ* method of probing the quantum well interface through a combination of spectroscopic measurements of the quantum dot and theoretical simulations. We note that in future work, more detailed information could be determined by moving the dot along two axes.

The inset of Fig. 5.4(a) reports the ratio $E_{\mathrm{ST}}/E_{\mathrm{val}}$ as a function of E_{orb}, showing that this ratio is significantly below unity for the entire range of parameters measured. Electron-electron interactions suppress E_{ST} below the non-interacting energy E_{val} [63], revealing that there are strong e-e interactions across all orbital splittings studied here (1.69–2.26 meV). To look for any correlations between dot position relative to the steps and the ratio $E_{\mathrm{ST}}/E_{\mathrm{val}}$, the dot positions in Fig. 5.4(b) are color coded by their corresponding orbital energies, with the mapping shown in the inset plot. The random distribution of colorings in Fig. 5.4(b) demonstrates that orbital energy spacing is more important than dot position in determining the suppression of E_{ST} below its non-interacting limit of E_{val}.

5.1.6 Conclusion

In this chapter, large *in situ* tunability of valley, singlet-triplet, and orbital splittings was measured, allowing for determination of the quantitative relationship between these three important energy scales. FCI simulations of the measured valley-orbit states were found to be in good agreement with the data, pointing to a combination of two primary physical parameters driving the relationship between these valley-orbit states: electrostatic confinement strength and quantum dot position relative to steps in the quantum well interface. This improved understanding enables a new *in situ*

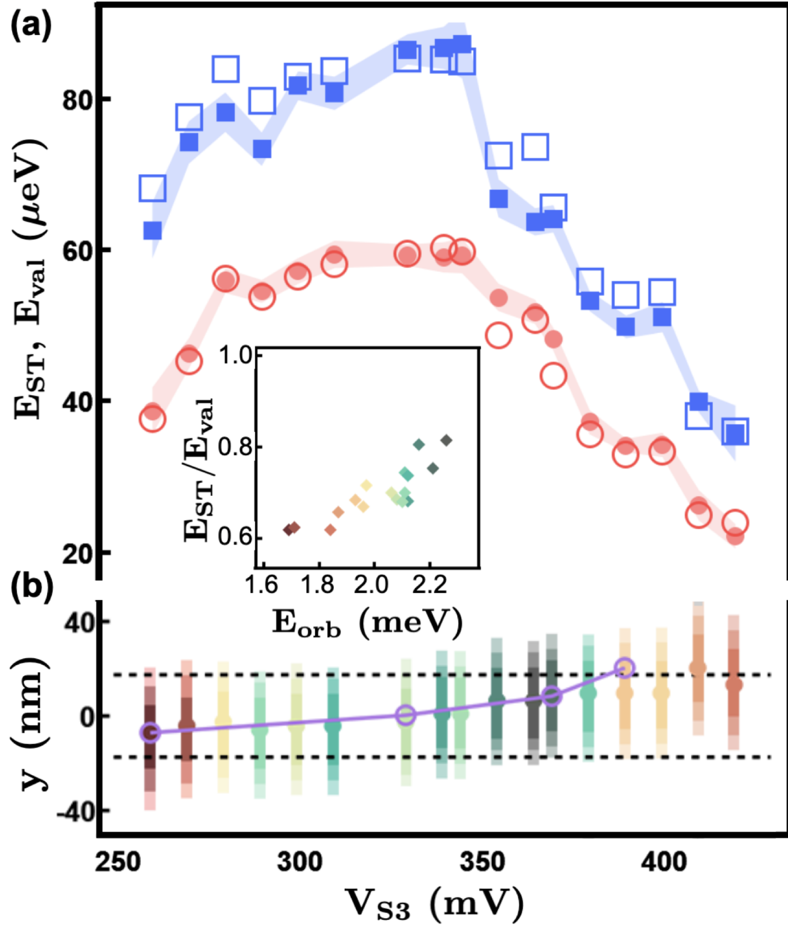

Figure 5.4: (a) The measured valley splitting (E_{val}) and singlet-triplet splitting (E_{ST}) are plotted as a function of V_{S3} as filled blue squares and red circles, respectively, where the measurement uncertainty is indicated by the width of the shaded regions. FCI simulations for E_{val} and E_{ST}, shown as open blue squares and red circles, respectively, quantitatively reproduce the experimental data once the effects of disorder in the form of atomic steps at the quantum well interface are included. (b) The non-monotonic behavior observed in both E_{ST} and E_{val} as a function V_{S3} is well explained by the quantum dot position changing with respect to distinct atomic steps at the quantum well interface. The wave function and position of the P2 dot singlet state is plotted with respect to atomic steps (dashed lines) at the interface, shown as the colored gradients and points. A best fit to the data produces a smooth change in position of the dot across steps spaced approximately 35 nm apart. In addition, the change in position from the Thomas-Fermi simulation, plotted as the purple line and open circles, agrees well with the FCI simulations. The inset plots the ratio E_{ST}/E_{val} as a function of the orbital splitting (E_{orb}). The suppression of the ratio E_{ST}/E_{val} below unity is a consequence of strong electron-electron interactions.

method of probing the quantum well interface through a combination of spectroscopic data and theoretical simulations.

Additionally, the relationship between the singlet-triplet and valley splittings is important for the future of valley-orbit based qubits. Since the qubit energy for valley-orbit qubits is determined largely by the valley splitting or singlet-triplet splitting, this method of *in situ* tunability can help for improved reliability and performance of valley-orbit qubits. Although large *in situ* tunability was possible by moving the quantum dot, tuning the singlet-triplet splitting by varying the quantum dot confinement is a more robust method of tuning qubit energy since it doesn't disturb neighboring dots nearly as much.

5.2 Supplemental material

5.2.1 Pulsed-gate spectroscopy

Consider a dot tunnel coupled to a reservoir, illustrated in Fig. 5.5(a). The dot is pulsed in a two-step duty cycle between voltages V_L and V_U, as shown in Fig. 5.5(b). The corresponding chemical potentials of the ground states of the dot, E_{Lg} and E_{Ug}, are related to V_L and V_U via a lever arm: $E_{\nu g} = \alpha V_\nu + E_0$, where $\nu = U, L$ and E_0 is the ground state (g) of the dot in the absence of a pulse. If the dot also has an excited state (x), then $\Delta E \equiv E_{Ux} - E_{Ug} = E_{Lx} - E_{Lg}$, and $E_{\nu x} = E_{\nu g} + \Delta E$. Tunneling can occur if there is a state in the reservoir that electrons can tunnel out of or into, as determined by the Fermi function

$$f_{\nu i} = \left[e^{(E_{\nu i} - E_F)/k_B T} + 1\right]^{-1}, \tag{5.3}$$

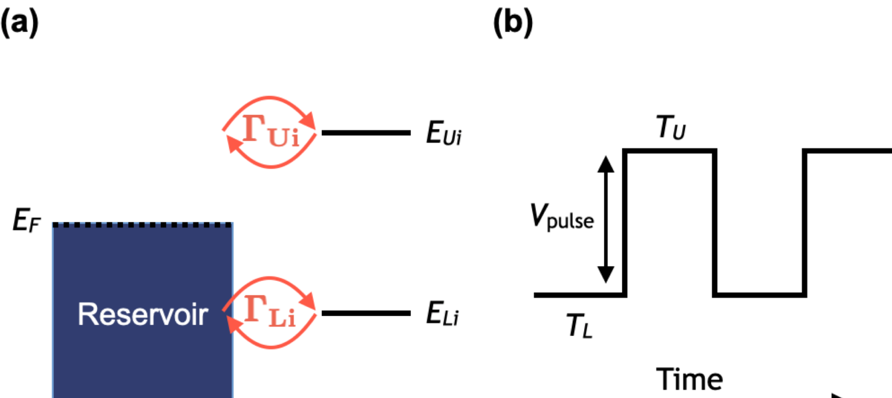

Figure 5.5: Parameter definitions for pulse spectroscopy. (a) E_F is the Fermi energy of the reservoir. E_{Ui} is the chemical potential of the dot eigenstate state i at the top edge of the pulse. Here, i can refer to the ground state (g), the first excited state (x_1), etc. E_{bi} is the chemical potential of the dot eigenstate i at the bottom edge of the pulse. Γ_{Ui} and Γ_{Li} are the corresponding tunnel couplings. (b) The duty cycle for the pulse applied to gate P2. $T_L = T_U$ for valley-orbit measurements. The pulse frequency $f_{\text{pulse}} = 1/(T_L + T_U)$ is on the order of the tunnel rate into/out of the dot during the bottom part of the pulse cycle. The tunnel rate out of the dot during the top part of the pulse is much faster than the pulse frequency such that the dot is completely empty at the beginning of the bottom part of the pulse.

where $i = g, x$.

The tunnel rates, $\Gamma_{\nu i}$, depend on the chemical potential $E_{\nu i}$, and can differ for the different dot eigenstates ($\nu = g, x$). As discussed in Ref. [53], $\Gamma_{\nu i}$ is a complicated exponential function. However, since there is not enough structure in the data to fully characterize the features of $\Gamma_{\nu i}$, we can simply linearize this function about the Fermi level, obtaining

$$\Gamma_{\nu g} \simeq \Gamma_{0,\nu g} e^{(E_{\nu g} - E_F)/E_{0,\nu g}}, \tag{5.4}$$

$$\Gamma_{\nu x} \simeq \Gamma_{0,\nu x} e^{(E_{\nu g} + \Delta E - E_F)/E_{0,\nu x}}. \tag{5.5}$$

While there also exists a decay rate from the excited state to the ground state, Γ_{xg}, here we do not consider its effect since recent measurements [67] indicate the decay rate is much slower than our pulse frequency $f_{\text{pulse}} \simeq 1\,\text{MHz}$.

For a single ground state in the quantum dot, the rate equation for its level filling n_g is derived using detailed balance

$$\dot{n}_{\nu g} = \Gamma_{\nu g}(f_{\nu g} - n_g) \tag{5.6}$$

Before solving Eqn. 5.6, we note that the steady state filling of the ground state, n_g, can be computed by setting $\dot{n}_g = 0$. The steady state filling is then simply the Fermi function, $n_g = f_g$. However, this is only valid in the limit that $\Gamma_{\nu g} \gg f_{\text{pulse}}$. The measurements discussed in Section 5.1.4 tune the tunnel rates such that $\Gamma \simeq f_{\text{pulse}}$, giving high sensitivity for pulsed gate spectroscopy measurements [12]. Thus, we must find the general solution to Eqn. 5.6 for $t > t_0$

$$n_{\nu g}(t) = f_{\nu g} - f_{\nu g}\left(1 - \frac{n_{\nu g}(t_0)}{f_{\nu g}}\right) e^{-\Gamma_{\nu g}(t - t_0)} \tag{5.7}$$

where $n_{\nu g}(t_0)$ is the occupation of the dot at $t = t_0$. Since αV_{pulse} is much larger than the singlet-triplet splitting (E_{ST}) and valley splitting (E_{val}), $\Gamma_{Ug} \gg f_{\text{pulse}}$ in the fit

regime [141]. Thus, when fitting our data, it is assumed that the dot is completely empty at the beginning of the bottom part of the duty cycle. Then $n_{Lg}(t_0) = 0$ and Eqn. 5.6 simplifies to

$$n_{Lg}(t) = f_{Lg}\left(1 - e^{-\Gamma_{Lg}(t-t_0)}\right) \tag{5.8}$$

The average electron occupation of the ground state of the dot during the entire pulse cycle is found by integrating from $t = t_0$ to $t = t_0 + T_L$ and dividing by $1/(T_U + T_L)$, where it has been assumed the average electron occupation of the dot during the top part of the duty cycle is 0 and $T_U = T_L$. Integrating Eqn. 5.8, we obtain

$$\langle n_g \rangle = \frac{1}{2} f_{Lg} + \frac{1}{2T_L}\frac{f_{Lg}}{\Gamma_{Lg}}\left(e^{-\Gamma_{Lg}T_L} - 1\right) \tag{5.9}$$

where the first term on the right hand side of Eqn. 5.9 represents the steady state occupation of the dot during the bottom part of the pulse, and the second term represents the perturbation due to energy dependent tunneling [53]. Finally, assuming that $\langle n_L \rangle = \langle n_{Lg} \rangle + \langle n_{Lx} \rangle$, the average electron occupation for the ground and excited states in the dot during the bottom part of the pulse is

$$\langle n \rangle = \frac{1}{2}\sum_{i=g,x}\left[f_{Li} + \frac{1}{T_L}\frac{f_{Li}}{\Gamma_{Li}}\left(e^{-\Gamma_{Li}T_L} - 1\right)\right] \tag{5.10}$$

This form can be simplified further into a convenient fitting form by expanding the exponential in Eqn. 5.10. This is valid for our measurements since $\Gamma \sim f_{\text{pulse}}$ and

$T_L = 1/(2f_{\text{pulse}})$, making $\Gamma_{Li}T_L < 1$. Expanding to second order, we obtain the final fitting form for determining valley-orbit state splittings using pulsed-gate spectroscopy

$$\langle n \rangle = \sum_{i=g,x} \Gamma_i \frac{e^{(E_i - E_F)/E_{0i}}}{e^{(E_i - E_F)/k_B T_e} + 1} \tag{5.11}$$

where $E_i = \alpha V_i$ is the position of each peak in energy, and E_{0i} and Γ_i are fitting parameters for each peak.

5.2.2 Magnetospectroscopy

Magnetospectroscopy was used in addition to pulsed-gate spectroscopy to supplement measurements of E_{ST}, providing better precision at low valley-orbit splittings and quantitative verification of pulsed-gate valley-orbit splitting measurements. In this section, a model is developed that can be used to extract E_{ST} from magnetospectroscopy data sets, where the plunger gate voltage is swept on one axis and the magnetic field is swept on the other. This 2-dimensional data set is then converted into a 1-dimensional data set by peak-fitting along the voltage axis, where the differential conductance is fit using the derivative of Eqn. 2 in Ref. [69] with respect to gate voltage. The peak-fitted data set is then curve fit using Eqn. 5.21, which we now derive.

Consider a quantum dot that is tuned electrostatically such that the one- and two-electron charge states are nearly degenerate. The system is found to be in one of six spin states that is dependent on the number of electrons in the quantum dot. For one electron, it is found to be in the $|\uparrow\rangle$ or $|\downarrow\rangle$ spin-state. For two-electrons, it is found to be in one of the singlet ($|S_0\rangle$) or triplet ($|T_-\rangle, |T_0\rangle, |T_+\rangle$) spin-states. Including charge and spin degrees of freedom and excluding charging energy, the energy of each state is

given by [142, 143]

$$E_{\downarrow} = -\frac{1}{2} g\mu_B B - \alpha V_{\mathrm{P2}} \tag{5.12}$$

$$E_{\uparrow} = \frac{1}{2} g\mu_B B - \alpha V_{\mathrm{P2}} \tag{5.13}$$

$$E_{S_0} = -2\alpha V_{\mathrm{P2}} \tag{5.14}$$

$$E_{T_-} = -g\mu_B B - 2\alpha V_{\mathrm{P2}} + E_{\mathrm{ST}} \tag{5.15}$$

$$E_{T_0} = -2\alpha V_{\mathrm{P2}} + E_{\mathrm{ST}} \tag{5.16}$$

$$E_{T_+} = g\mu_B B - 2\alpha V_{\mathrm{P2}} + E_{\mathrm{ST}} \tag{5.17}$$

where $g = 1.99$ is the electron g-factor in silicon, μ_B is the Bohr magneton, α is the gate lever arm, and V_{P2} is the gate voltage.

Using standard lock-in detection techniques, detailed in Ref. [11], the derivative of the average electron occupation of the dot with respect to gate voltage is measured, shown in Fig. 5.2(e). A derivation for the average electron occupation of a system in thermal and diffusive contact with a large reservoir is detailed in Ref. [144]. The total electron occupation $\langle n \rangle$ is determined by treating the system as a grand canonical ensemble

$$\langle n \rangle = \frac{1}{\mathcal{Z}} \sum_i N_i e^{(N\mu - E_i)/k_B T_e} \tag{5.18}$$

$$\mathcal{Z} = \sum_i e^{(N_i\mu - E_i)/k_B T_e} \tag{5.19}$$

where $\mathcal{Z}$ is the grand partition function, μ is the chemical potential of the reservoir, T_e is the electron temperature and N_i is the number of electrons on the dot for each given state (two electrons for the $|S_0\rangle$ and $|T\rangle$ states, and one electron for $|\uparrow\rangle$ and $|\downarrow\rangle$

states).

A 4-state model (inclusion of states $|\downarrow\rangle, |\uparrow\rangle, |S_0\rangle, |T_-\rangle$) may be sufficient in the limit of low T_e and high E_{ST}; however, we find that in the experimentally relevant range ($T_e \simeq 100\,\mathrm{mK}$, $E_{\mathrm{ST}} = 25$–$60\,\mu\mathrm{eV}$), the 6-state model is requisite since significant occupation of higher lying states at $B = 0$ occurs. We do not attempt to quantitatively match the raw voltage V_{P2}, instead plotting δV_{P2}. This simplifies the final result (Eqn. 5.18), allowing $\mu = 0$ as this quantity only provides an offset in V_{P2}. Additionally, this justifies excluding the charging energy from the states at the beginning of the derivation. The average electron occupation is then

$$\langle n\rangle = 1 + \frac{e^{\alpha\beta_e V_{\mathrm{P2}}}\left(1 + e^{\kappa B}\left(1 + e^{E_{\mathrm{ST}}\beta_e} + e^{\kappa B}\right)\right)}{e^{\alpha\beta_e V_{\mathrm{P2}}} + e^{(\alpha V_{\mathrm{P2}} + E_{\mathrm{ST}})\beta_e + \kappa B} + e^{E_{\mathrm{ST}}\beta_e + \frac{1}{2}\kappa B} + e^{\alpha\beta_e V_{\mathrm{P2}} + \kappa B} + e^{E_{\mathrm{ST}}\beta_e + \frac{3}{2}\kappa B} + e^{\alpha\beta_e V_{\mathrm{P2}} + 2\kappa B}} \tag{5.20}$$

where we have introduced $\beta = 1/k_B T_e$ and $\kappa = g\mu_B\beta_e$. The derivative of Eqn. 5.20 with respect to gate voltage is plotted in Fig. 5.2(f).

When fitting for E_{ST}, the peak value position along the gate voltage axis (V_{P2}) is extracted as described above (shown in Fig. 5.2(g)). The curve fit is obtained from 5.20 by solving the equation $d^2\langle n\rangle/dV_{\mathrm{P2}}^2 = 0$. The resulting curve fit as a function of magnetic field is

$$V_{P2}(B) = \frac{1}{\alpha\beta_e}\ln\left(\frac{e^{\frac{1}{2}\kappa B + \beta_e E_{\mathrm{ST}}}\left(e^{\kappa B} + 1\right)}{e^{\kappa B} + e^{2\kappa B} + e^{\kappa B + E_{\mathrm{ST}}\beta_e} + 1}\right), \tag{5.21}$$

5.2.3 COMSOL simulation procedure

The COMSOL Multiphysics simulations are full device simulations where the dimensions for each gate electrode are modeled out to 500 nm away from the center of the device. Using a realistic quasiclassical (Thomas-Fermi) electrostatic simulation [145, 146], the electron density is calculated as a function of gate voltages. The simulation is used to gain intuition and provide estimates for quantum dot shape and position as the electrostatic configuration is varied across the experimental range.

The presence of offset charge is incorporated into the simulation by first calculating the electron density with the measured device voltages with no offset charge. Offset charge is then added using a constant gate voltage offset to all gates until the electron density beneath the reservoirs plateaus. This is the theoretical analog to the experimental device tuning procedure used, where reservoir voltages are increased until a plateau in transport is observed. Then, while fixing S3 to 370 mV plus the original offset, a new offset to all other gates is applied until the integrated charge in the quantum dot beneath P2 is roughly one electron. This procedure preserves the differential voltage between each gate as closely as possible.

The final step in simulating the experimental range of electrostatic tunings is adjusting V_{S3} and the gates neighboring the P2 quantum dot to closely approximate the gate voltage changes made in the experiment. The gate lever arms (α) are determined for each gate in the COMSOL simulation, and are found to be in close agreement (but not identical) to the experiment. Table 5.1 shows the experimentally measured α for each gate neighboring the P2 quantum dot. To keep the P2 quantum dot electron occupation close to 1.0 throughout the different electrostatic regimes, COMSOL offset voltages were determined using the ratio of experimental to COMSOL gate lever arm

Table 5.1: Measurements of α for gates S3, S2, P2, B2 and B3 at varying electrostatic configurations (denoted by the S3 gate voltage value). In addition to the changing S3 gate voltage, a compensating voltage to neighboring plunger and barrier gates was applied to maintain the electron occupation of the P2 dot. The change in α values clearly shows the dot moving away from S2 toward S3 as the S3 gate voltage changes from 260–420 mV. This exact behavior is observed in the COMSOL simulations.

S3 (mV)	α_{S3}	α_{S2}	α_{P2}	α_{B2}	α_{B3}
260	0.231	0.263	0.209	0.063	0.098
330	0.226	0.213	0.207	0.064	0.089
370	0.230	0.179	0.198	0.061	0.079
420	0.277	0.134	0.189	0.064	0.075

values as $\Delta V_{\text{sim}} = \alpha_{\text{exp}}/\alpha_{\text{sim}} \Delta V_{\text{exp}}$. By applying offsets in this way, the differential gate voltage used in the experiment is preserved without significantly changing the electron occupation of the simulated dot.

The COMSOL simulated orbital splittings and change in position of the quantum dot qualitatively match the experimental results, giving insight into the physical behavior of the device across the electrostatic configurations. Both the simulation and measurements indicate $\hbar\omega_x < \hbar\omega_y$ at low S3 gate voltages, and $\hbar\omega_x > \hbar\omega_y$ at large S3 gate voltages. Between the S3 voltages of 330–370 mV, the dot is found to be roughly isotropic in the simulation. As noted in the Section 5.1.4, voltages of V_{S3} >390 mV did not produce physical results in the simulation due to electron accumulation directly beneath the S3 screening gate. This same phenomenon occurred at 430 mV in the experiment. The discrepancy is most likely due to a small difference in α_{S3} between the simulation and experiment since the α of each gate varies across the electrostatic tunings, as shown in Table 5.1.

5.2.4 TB and FCI methods

Full configuration interaction (FCI) simulations aided by tight-binding (TB) calculations of the electronic wave functions are used to accurately model E_{val} and E_{ST}, as shown in Fig. 5.4. The TB model calculations result in wave functions with fast oscillations in the z-direction which breaks the valley degeneracy. Thus, effects on the valley splitting due to atomistic disorder, and the disorder-induced valley-orbit coupling are captured in the TB model. FCI methods allow for accurate calculation of two-electron energies through diagonalization of the interacting two-electron Hamiltonian in the basis of 3160 Slater determinants generated by using the TB solutions. Therefore, the combination of these two procedures capture the interplay among the valleys, interface disorder and e-e interactions.

The FCI calculations use a number of experimentally driven input parameters. This includes the measured E_{orb} for each point, a vertical electric field of $0.7\,\mathrm{MV\,m^{-1}}$, a 9 nm thick quantum well and atomic steps at the quantum well interface separated by 35 nm. The confinement potential of the quantum dot is based on E_{orb}, which is used in the form of a harmonic potential for determining the electron wave function in the x-y plane of the quantum dot. Additionally, this influences the strength of e-e interactions. The procedure for simulating valley and singlet-triplet energies is detailed more thoroughly in Ref. [147].

For the calculation of quantum dot position with respect to the position of atomic steps at the quantum well interface, the following procedure is used: first, the singlet-triplet and valley splittings are simulated as the quantum dot position changes. The dot position is varied in steps of 1.76 nm such that a matrix of 20 values are simulated between two atomic steps separated by 35 nm. Next, the simulated values are compared

to experimental values, and the point which simultaneously has the smallest total error between simulation and experiment for both singlet-triplet and valley splittings is chosen. Finally, the dot is only allowed to move in one direction, otherwise the results would be unphysical. This condition is enforced by noting that the simulations produce a reflectional symmetry halfway between the steps. Two low-error simulations are found on both sides of this symmetry point. Thus, once the quantum dot has moved past the symmetry point, it cannot move backwards across it.

5.3 Extended discussion

5.3.1 Comparison between pulsed-gate spectroscopy and magnetospectroscopy

Since Fig. 5.3 reports values of E_{ST} and E_{val} measured using both magnetospectroscopy and pulsed-gate spectroscopy, it is important to confirm quantitative agreement between the two measurement techniques. Pulsed-gate spectroscopy is the more powerful of the two techniques for determining valley-orbit state splittings since single and multi-electron valley-orbit states can be detected. However, in some cases, excited-state signals are too weak due to tunnel rate constraints on the system, making magnetospectroscopy the superior method. This is especially true in regimes where valley-orbit energies approach the electron temperature T_e. For our measurements, it became difficult to distinguish between the ground and excited state signals below $\sim$35 μeV due to thermal broadening.

Fig. 5.6 shows the comparison of E_{ST} measured by magnetospectroscopy (blue squares) and pulsed-gate spectroscopy (orange circles). Magnetospectroscopy data

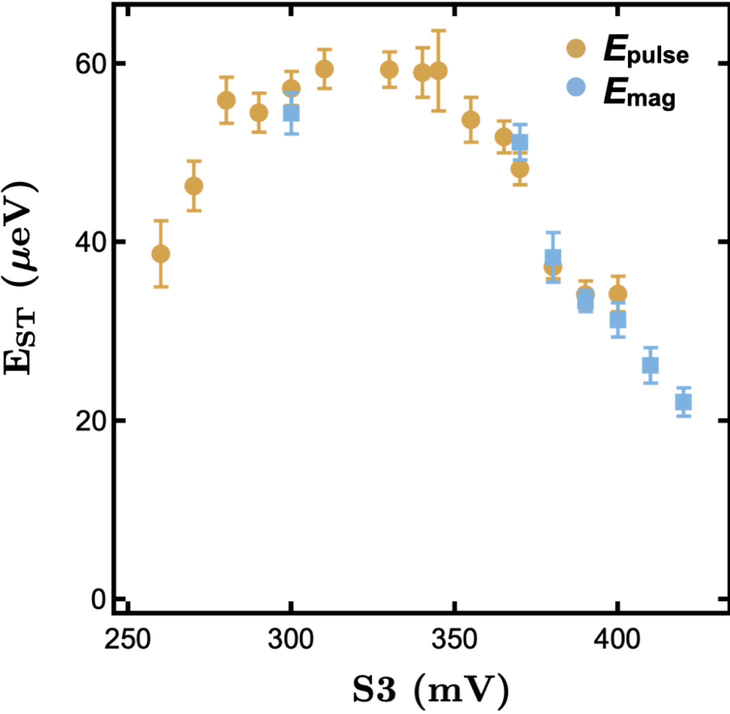

Figure 5.6: Comparison of singlet-triplet splitting E_{ST} measured at different electrostatic configurations. The screening gate voltage V_{S3} changes the electrostatic confinement potential of the dot. E_{ST} is measured as a function of the V_{S3} using both pulsed-gate spectroscopy (red) and magnetospectroscopy (blue). Good agreement is found between the two methods, where magnetospectroscopy is used for smaller values of E_{ST} since pulsed-gate spectroscopy has difficulty resolving two distinct peaks below 35 μeV at an electron temperature of 100 mK.

was not taken at all electrostatic configurations due to the length of the measurement. Since magnetospectroscopy was the primary method used for low E_{ST}, we emphasized measurements in that regime. Agreement within the error bars (given as the standard deviation) is found between the two methods.

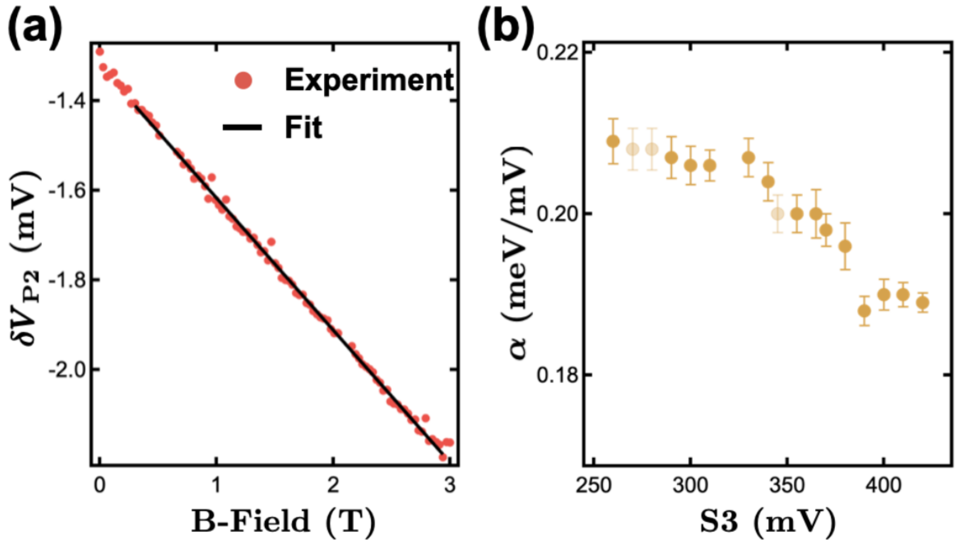

Figure 5.7: Measurement of the gate voltage to energy lever arm α. (a) An example of determining α using magnetospectroscopy. The P2 gate voltage is swept such that the quantum dot beneath it undergoes a charge transition from 0 to 1 electrons (plotted as red circles). α is calculated by determining the slope of (a) and using Eqn. 5.22. (b) The change in α as a function of V_{S3} is plotted. Opaque orange data points show measured values of α where the error bars are given as the standard deviation. Semi-transparent points are extrapolated based on neighboring values such that valley-orbit energies can be calculated at these points. Since the quantum dot moves monotonically in one direction with V_{S3}, α is also expected to be monotonic as a function of V_{S3}.

5.3.2 Measurement of alpha

Measurement of alpha (α), the gate electrode lever arm, is necessary in converting measured valley-orbit voltages via pulsed-gate spectroscopy into energies. There exist a variety of techniques to measure α but we find magnetospectroscopy to be the best method since α measurements can be taken in the same electrostatic tuning as the valley-orbit state measurements, giving the most accurate estimate of α in the experimental regime. By sweeping to large magnetic fields (3–4 T), low-error measurements of α are obtained at each of the electrostatic configurations where valley-orbit state splittings were measured.

An example measurement is shown in Fig. 5.7(a), where the peak position of the differential average electron occupation with respect to gate voltage ($d\langle n\rangle/dV_{\mathrm{P2}}$) has been extracted at each point (shown as red circles) from the 2D magnetospectroscopy data set, as described in Section 5.3.2. At positive magnetic fields, the change in energy across an electron loading line is well described by the Zeeman effect [142]. When converted to a gate voltage, α is determined as

$$\alpha = \frac{1}{2}\frac{g\mu_B}{\hbar}\left(\frac{dV_g}{dB}\right)^{-1} \tag{5.22}$$

where the pre-factor of 1/2 is due to the change in spin of the quantum dot across the loading line. α is measured at many of the electrostatic configurations used for valley-orbit splitting measurements, shown in Fig. 5.7(b). A few points were not measured (shown as semi-transparent orange circles). This was because a higher density of valley-orbit state measurements was needed than initially planned for the experiment, and α measurements are extremely time consuming. Where α is not measured explicitly, it is extrapolated to be the average of the measured points that neighbor it for valley-orbit splitting calculations.

5.3.3 Anisotropic electrostatic confinement

The orbital anisotropy is defined to be the ratio of orbital splitting in the x- and y-dimensions. Here, we define the ratio to be ω_x/ω_y, such that $\omega_x/\omega_y = 1$ is identified as the maximally isotropic case. Large orbital anisotropy can change the magnitude of valley-orbit state splittings [147]. In the single-particle case, E_{val} in an anisotropic dot is perturbed from the isotropic case due to the change in distribution of the electronic wave function. This results in the electron sampling a different area of the Si/SiGe

quantum well interface which will change E_{val} due to valley-orbit coupling. In addition, anisotropic multi-electron quantum dot valley-orbit states are perturbed from the isotropic case due to electron-electron (e-e) interactions. The weights of excited valley-orbit states occupied by a multi-electron dot are impacted by the anisotropy of the dot [147] which directly affects the valley-orbit splitting. For a two-electron quantum dot, anisotropy suppresses E_{ST} [147]. If orbital anisotropy can be *in situ* tuned, it may lead to *in situ* tuning of qubit parameters such as E_{ST} and E_{val}. We note that, while non-zero anisotropy is measured on the extreme ends of electrostatic tunings, the quantum dot is nearly isotropic across the vast majority of the different electrostatic tunings.

The gate geometry of the device presented in Fig. 5.1 is well suited for probing *in situ* orbital anisotropy tuning. Gate electrode S3 is used to pinch the P2 quantum dot orbital confinement in the y-dimension ($\hbar\omega_y$), resulting in high tunability. This procedure also indirectly alters the x-confinement ($\hbar\omega_x$). When the S3 gate voltage decreases, the chemical potential in the local region surrounding S3 increases. In order to maintain the same tunnel coupling and electron occupation of the P2 dot, the surrounding barrier and plunger gates must be increased, which reduces $\hbar\omega_x$.

Fig. 5.8 demonstrates orbital anisotropy tuning in two distinct regimes: at low S3 gate voltages (strong $\hbar\omega_y$) and high S3 gate voltages (weak $\hbar\omega_y$). In the strong $\hbar\omega_y$ regime, shown in Fig. 5.8(a, b), $\hbar\omega_x$ is found to be lower in energy than $\hbar\omega_y$. For S3 = 270 mV, $\omega_x/\omega_y = 0.83$, shown in Fig. 5.8(a). Increasing the S3 gate voltage by 10 mV to 280 mV reduces $\hbar\omega_y$ and the compensating voltages on B2 and B3 increases $\hbar\omega_x$ considerably, resulting in Fig. 5.8(b) where $\omega_x/\omega_y = 0.94$. *In situ* anisotropy tuning where $\hbar\omega_y$ is weak is shown in Fig. 5.8(c, d). When S3 = 430 mV, the quantum dot is found to be highly anisotropic, where $\omega_x/\omega_y = 1.96$ (Fig. 5.8(c)). Decreasing the S3

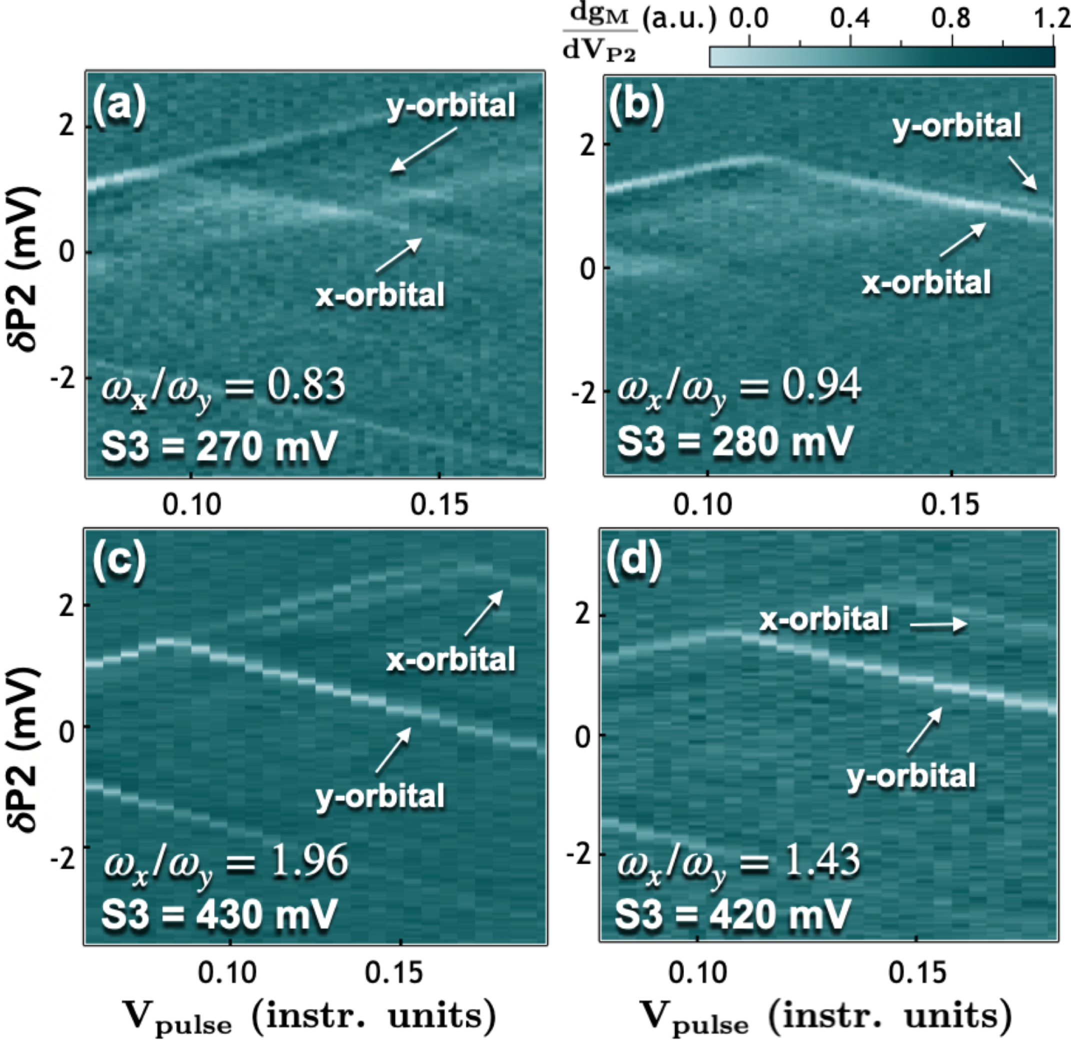

Figure 5.8: Anisotropy tuning of a quantum dot. (a, b) The P2 quantum dot is tightly confined in the y-dimension due to the low S3 gate voltage. Upon raising the voltage from 270 mV to 280 mV, the dot goes from anisotropic to nearly isotropic. The x-confinement rises due to compensating voltages on neighboring plunger and barrier gates. (c, d) The P2 quantum dot is weakly confined in the y-dimension due to the high S3 gate voltage. In (c), the 2DEG is nearly accumulated beneath the S3 gate, causing the dot to be highly anisotropic, where the ratio of the orbital splitting in the x-dimension to the y-dimension is $\omega_x/\omega_y = 1.96$. A large change in anisotropy is induced by only changing S3 by 10 mV, where in (d) the ratio is $\omega_x/\omega_y = 1.43$.

gate voltage to 420 mV increases $\hbar\omega_y$ and the compensating gate voltages on B2 and B3 decreases $\hbar\omega_x$ (shown in Fig. 5.8(d)), where $\omega_x/\omega_y = 1.43$. We note that E_{ST} and E_{val} were not reported for S3 = 430 mV because they were too small to be measured accurately.

Despite strong *in situ* orbital anisotropy tuning at the extreme ends of V_{S3}, the dot is highly isotropic over the vast majority of the electrostatic configurations S3 = 260–420 mV where E_{ST} and E_{val} were measured. Above S3 = 420 mV, E_{ST} and E_{val} are too small to measure with low error, and below S3 = 260 mV, the tunnel rate into the dot became too low to detect the E_{val} using pulsed-gate spectroscopy.

Chapter 6

Outlook: Si/SiGe quantum dots

In this thesis, a number of findings were presented that helped advance Si/SiGe quantum dot devices and valley-orbit qubits. In Chapter 3, a fabrication process for overlapping aluminum gate devices was presented with an emphasis on yield engineering. Using a plasma ash or UV-ozone process, the thickness of the inter-gate AlO_x was increased, resulting in higher yield devices. The thicker inter-gate oxide helps to prevent damage from ESD and gate-to-gate leakage. Additionally, TEM failure analysis was performed to identify key failure mechanisms stemming from aluminum gate morphology of the multi-layer architecture. In Chapter 4, the QDscript scripting environment was presented. QDscript helps enable a more flexible data acquisition environment, extending the functionality of commercial software Labber. In Chapter 5, experimental results on one- and two-electron valley-orbit states were shown. In this experiment, the valley, singlet-triplet and orbital splittings were measured at various electrostatic configurations. Gate voltages were tuned such that the position of a quantum dot changed approximately 25–30 nm. Through tight binding and FCI calculations, it was determined that atomic steps in the quantum well interface play an important role in

determining the relationship between valley-orbit states.

Silicon quantum dots as a platform for quantum information has progressed steadily in the last decade. Key demonstrations of the DiVencenzo criteria at fault tolerant levels have been made at the one- and two-qubit level. Several experiments have also taken significant steps in scaling silicon qubits toward much larger one- and two-dimensional arrays. The most important next steps for silicon qubits are improving the yield of fabrication processes and continuing to investigate new architectures that will allow for high fidelity one- and two-qubits gates that scale well in two dimensions. The most attractive aspect of silicon qubits is that logical bases can be encoded in spin-degrees of freedom, which are highly insensitive to charge noise, leading to long relaxation and coherence times. However, the complex architectures necessary for defining quantum dots within a semiconductor remain a technological hurdle that will require careful planning and consideration moving forward.

Appendix A

Fabrication recipes

A.1 Photolithography

A.1.1 Lithography Process (SPR-955)

1. Spin rinse dry wafer.
2. Set hot plate to 100°C and wait for temperature to equilibrate.
3. Dehydration bake at 100°C for 60 sec.
4. Spin SPR-955 at 5500 RPM for 30 sec.
5. Bake at 100°C for 90 sec.
6. Expose on stepper for 250 ms.
7. Post-exposure bake at 110°C for 90 sec.
8. Soak sample in MF-24A or CD-26 (200 mL) for 60 sec. Use strong agitation.
9. Quench in DI water (400 mL) for 30 sec. Use gentle agitation.
10. Rinse under DI water faucet for ∼10 sec.
11. Blow dry with N_2, at least 30 sec past visual dry.

A.1.2 Lithography Process (AZ-5214)

1. Spin rinse dry wafer.
2. Set hot plate to 95°C and wait for temperature to equilibrate.
3. If doing an HF dip with this pattern, use HMDS Primer Oven at 105°C to increase adhesion of the resist to substrate.
4. Spin AZ-5214 at 4000 RPM for 30 sec.
5. Bake at 95°C for 60 sec.
6. Expose on stepper for 75 ms.
7. *Wait 5 min for photoresist to outgas before continuing.
8. Post-exposure bake at 110°C for 60 sec.
9. *Wait 5 min for photoresist to outgas before continuing.
10. Flood expose on MA6 Aligner for 60 sec to reverse polarity of PR.
11. Soak sample in MF-24A (200 mL) for 45 sec. Use strong agitation.
12. Quench in DI water (400 mL) for 30 sec. Use gentle agitation.
13. Rinse under DI water faucet for $\sim$ 10 sec.
14. Blow dry with N_2, at least 30 sec past visual dry.

A.1.3 Lithography Process (S1813)

1. Spin rinse dry wafer.
2. Set hot plate to 90°C and wait for temperature to equilibrate.
3. If doing an HF dip with this pattern, use HMDS Primer Oven at 105°C to increase adhesion of the resist to substrate.
4. Spin S1813 at 4000 RPM for 30 sec.
5. Bake at 90°C for 180 sec.
6. Expose on stepper for 550 ms.
7. Soak sample in MF-321 for 60 sec. Use strong agitation.

8. Quench in DI water (400 mL) for 30 sec. Use gentle agitation.
9. Rinse under DI water faucet for ∼ 10 sec.
10. Blow dry with N_2, at least 30 sec past visual dry.
11. If doing an HF dip with this pattern, bake at 120°C for 180 sec.

A.1.4 Post-Ion Implantation Clean

1. Sonicate sample in acetone (300 mL or more), 10 min.
2. Soak in 1165 remover (100 mL) (held between 75-80°C using a hot plate) for 60 min.
3. Sonicate in acetone (200 mL), 10 min.
4. Sonicate in isopropanol (200 mL), 1 min.
5. Submerge in DI water (400 mL) and agitate, several seconds.
6. Spray wafer with DI water gun, ∼ 10 sec.
7. Blow dry aggressively with N_2, at least 30 sec past visual dry.
8. Check wafer for residue. If any is visible, or even if none is visible and the following step is a deposition step (e.g., field oxide, metal feature), clean with O_2 plasma, 250W, for at least 10 min.

A.1.5 Stripping Photoresist

1. 5 min soak in hot 1165 (held between 75-80°C using a hot plate) with manual agitation.
2. Transfer to fresh 1165.
3. Sonicate 1 min at 20% in 1165.
4. Move beaker to 110°C hotplate for 10 min soak.
5. Sonicate 1 min at 20% in 1165.
6. Transfer to acetone. Sonicate 5 min at 20% in acetone.
7. Transfer to IPA. Sonicate 1 min at 20% in IPA.

8. Transfer to DI water for quick (10sec) dip.
9. Blow dry aggressively with N_2, at least 30 sec past visual dry.
10. Descum for 10 min in YES Asher at 250 W and 80 sccm.

A.1.6 Liftoff

1. Soak sample in 1165 remover (200 mL) on 110°C hot plate for 30 min.
2. Every 5-10 min, gently agitate.
3. Transfer wafer to a beaker of fresh 1165 remover (200 mL).
4. Sonicate wafer in second beaker of 1165 remover for 10 min at 20% power.
5. Sonicate in acetone (200 mL), 10 min at 20% power.
6. Sonicate in isopropanol (200 mL), 1 min at 20% power.
7. Submerge in DI water (400 mL) and agitate, several seconds.
8. Blow dry aggressively with N_2, at least 30 sec past visual dry.
9. Perform a 5 min clean in the YES Asher (250 W, 80 sccm).

A.1.7 Stripping Photoresist (Quick)

1. Spin sample at 5500 RPMs for 30 sec.
2. During the spin, spray acetone and isopropanol (together) on sample for $\sim$ 10 sec.
3. While finishing acetone rinse, continue spraying isopropanol for $\sim$ 10 sec.
4. Blow dry with N_2 while spin is still going on.
5. Check under microscope that all photoresist is gone.
6. If photoresist is still present, follow the longer stripping procedure found in Section A.1.5.

A.1.8 Stripping Native Oxide

1. Dip sample in 20:1 BOE for 30 sec, using light agitation.

2. Rinse sample in 3 successive beakers of DI water.

3. Blow dry with N_2.

4. Must bring sample(s) to vacuum within 15 min of their last rinse to prevent oxide re-growth.

A.2 Nanolithography

A.2.1 Lithography Process (PMMA-495 A4)

1. Confirm the chip is as clean as possible before spinning resist. This may include sonication in various solvents as well as plasma cleans.

2. 'Dehydrate' chip by placing on 180 C hot plate for at least 1 minute. This is to ensure good adhesion of the resist to the surface. Allow chip to cool completely before spinning.

3. Apply 1-2 drops of PMMA 495 A4 to the chip depending on chip size. Spin at 5500 RPM for 45 seconds.

4. Bake at 180 C for 90 seconds.

5. Load chip into EBL. It is suggested to wedge the chips against chip mount corners with the clips, not to place the clips on top of the chips.

6. Use an acceleration voltage of 100 kV. If the tool was used at a different voltage recently, set to 100 kVand wait $> 1/2$ hour.

7. Write the pattern desired. Suggested settings are 250 μm field, 10^6 dots, and 5 nm pitch. Dose is typically in the range of 10

8. For room temperature development, decant a sufficient amount of MIBK:IPA 1:3 into a glass beaker. Decant IPA into a separate beaker. For cold development, both the MIBK:IPA and beaker should have been stored in the refrigerator beforehand.

9. Dip the chip into the development solution and agitate in a consistent manner for 60 seconds.

10. Quickly transfer the chip to IPA. Agitate for 10 seconds. This step helps remove the MIBK, but be aware that pure IPA develops faster than the development mixture.

11. Quickly remove from IPA and blow dry with N_2 gas for > 30 seconds.

12. Descum: Run the barrel O_2 asher in 1148 for 5 minutes to clean the chamber. Place chip on glass slide in the chamber. Pump down and flow 4 sccm O_2 for 1 minute. Turn power on 'high' for 10 seconds. This step helps remove any small bits of resist that were incompletely developed results in a cleaner metal deposition.

13. Vent the evaporator. Place the chip(s) on the chip holder stage and load into evaporator. Use the heat shields meant for chips, which have a smaller aperture to reduce unnecessary heating.

14. Once the chamber pressure is sufficiently low ($<$ 2E-6 torr), evaporate aluminum at a rate of 0.3-0.4 Å/s. To contact over the interconnects, deposit 15-20 nm more metal than the thickness of the interconnects. For each additional layer, add 15-20 nm more metal. For example, for 20 nm interconnects, deposit 35, 55, and 70 nm of aluminum for layers 1, 2, and 3, respectively. For thicker depositions, test that liftoff is possible and repeatable.

A.2.2 Liftoff

1. Heat 1165 or PG remover in a glass beaker on a 110-120 C hot plate to achieve a solution temperature of 75 C. Place chip(s) in liquid and soak for 30 minutes. Meanwhile, heat a second glass beaker of 1165/PG.

2. Remove beaker from hotplate and use the 'pipette method' to blow off the majority of metal. This requires some skill and practice, but is critical to remove stubborn edges of metal.

3. Transfer chip(s) to second beaker of hot solution. This helps drop the amount of metal floating around in the beaker and results in a cleaner surface. Soak for another 30 minutes.

4. Transfer chip(s) to beaker of acetone and sonicate for 5 minutes. If in the NFC, use power 0; if in Chamberlin, use 20% power. 1165/PG is very sticky and good solvent rinsing/sonication is important to help remove it.

5. Transfer chip(s) to beaker of IPA and sonicate again for at least 1 minute. Use the same sonication powers as before.

6. Quickly remove chip(s) from IPA and blow dry with N_2 gas for > 30 seconds.

A.2.3 Plasma Clean and Oxidation

1. Load the sample into the YES plasma asher in the NFC.
2. Run a 250 W, 80 sccm, 4.5-4.6 Torr O_2 plasma clean for 10 minutes.
3. Unload chip(s) and inspect your hard work.

Bibliography

[1] https://www.r-ccs.riken.jp/en/fugaku/about/.

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, and *et. al.*, Nature **574**, 505 (2019).

[3] `https://quantumalgorithmzoo.org/`.

[4] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, USA, 2011).

[6] K. R. Brown, A. C. Wilson, Y. Colombe, C. Ospelkaus, A. M. Meier, E. Knill, D. Leibfried, and D. J. Wineland, Phys. Rev. A **84**, 030303 (2011).

[7] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[8] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **76**, 042319 (2007).

[9] L. P. Kouwenhoven, S. Jauhar, J. Orenstein, P. L. McEuen, Y. Nagamune, J. Motohisa, and H. Sakaki, Phys. Rev. Lett. **73**, 3443 (1994).

[10] T. Fujisawa, D. G. Austing, Y. Tokura, Y. Hirayama, and S. Tarucha, Nature **419**, 278 (2002).

[11] J. M. Elzerman, R. Hanson, J. S. Greidanus, L. H. Willems van Beveren, S. De Franceschi, L. M. K. Vandersypen, S. Tarucha, and L. P. Kouwenhoven, Phys. Rev. B **67**, 161308 (2003).

[12] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, L. M. K. Vandersypen, and L. P. Kouwenhoven, Appl. Phys. Lett. **84**, 4617 (2004).

[13] J. R. Petta, A. C. Johnson, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **93**, 186802 (2004).

[14] J. R. Petta, A. C. Johnson, A. Yacoby, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. B **72**, 161301 (2005).

[15] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven, Nature **430**, 431 (2004).

[16] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Science **309**, 2180 (2005).

[17] F. H. L. Koppens, C. Buizert, K. J. Tielrooij, I. T. Vink, K. C. Nowack, T. Meunier, L. P. Kouwenhoven, and L. M. K. Vandersypen, Nature **442**, 766 (2006).

[18] K. C. Nowack, F. H. L. Koppens, Y. V. Nazarov, and L. M. K. Vandersypen, Science **318**, 1430 (2007).

[19] M. Pioro-Ladrière, T. Obata, Y. Tokura, Y.-S. Shin, T. Kubo, K. Yoshida, T. Taniyama, and S. Tarucha, Nat. Phys. **4**, 776 (2008).

[20] D. J. Reilly, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Appl Phys Lett **91**, 162101 (2007).

[21] J. I. Colless, A. C. Mahoney, J. M. Hornibrook, A. C. Doherty, H. Lu, A. C. Gossard, and D. J. Reilly, Phys. Rev. Lett. **110**, 046805 (2013).

[22] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).

[23] B. M. Maune, M. G. Borselli, B. Huang, T. D. Ladd, P. W. Deelman, K. S. Holabird, A. A. Kiselev, I. Alvarado-Rodriguez, R. S. Ross, A. E. Schmitz, M. Sokolich, C. A. Watson, M. F. Gyure, and A. T. Hunter, Nature **481**, 344 (2012).

[24] D. Kim, Z. Shi, C. B. Simmons, D. R. Ward, J. R. Prance, T. S. Koh, J. K. Gamble, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Nature **511**, 70 (2014).

[25] Z. Shi, C. B. Simmons, D. R. Ward, J. R. Prance, X. Wu, T. S. Koh, J. K. Gamble, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Nat. Commun. **5**, 3020 (2014).

[26] B. Thorgrimsson, D. Kim, Y.-C. Yang, L. W. Smith, C. B. Simmons, D. R. Ward, R. H. Foote, J. Corrigan, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, npj Quantum Inf. **3**, 32 (2017).

[27] E. Kawakami, P. Scarlino, D. R. Ward, F. R. Braakman, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nat. Nanotechnol. **9**, 666 (2014).

[28] J. Yoneda, T. Otsuka, T. Nakajima, T. Takakura, T. Obata, M. Pioro-Ladrière, H. Lu, C. J. Palmstrøm, A. C. Gossard, and S. Tarucha, Physical Review Letters **113**, 267601 (2014).

[29] X. Wu, D. R. Ward, J. R. Prance, D. Kim, J. K. Gamble, R. T. Mohr, Z. Shi, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, PNAS **111**, 11938 (2014).

[30] D. Kim, D. R. Ward, C. B. Simmons, J. K. Gamble, R. Blume-Kohout, E. Nielsen, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Nat. Nanotechnol. **10**, 243 (2015).

[31] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, Nat. Nanotechnol. **9**, 981 (2014).

[32] J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, Nature Nanotechnol. **13**, 102 (2018).

[33] M. Veldhorst, C. H. Yang, J. C. C. Hwang, W. Huang, J. P. Dehollain, J. T. Muhonen, S. Simmons, A. Laucht, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, Nature **526**, 410 (2015).

[34] T. F. Watson, S. G. J. Philips, E. Kawakami, D. R. Ward, P. Scarlino, M. Veldhorst, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, M. A. Eriksson, and L. M. K. Vandersypen, Nature **555**, 633 (2018).

[35] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta, Science **359**, 439 (2018).

[36] X. Xue, T. F. Watson, J. Helsen, D. R. Ward, D. E. Savage, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, S. Wehner, and L. M. K. Vandersypen, Phys. Rev. X **9**, 021011 (2019).

[37] R. W. Andrews, C. Jones, M. D. Reed, A. M. Jones, S. D. Ha, M. P. Jura, J. Kerckhoff, M. Levendorf, S. Meenehan, S. T. Merkel, A. Smith, B. Sun, A. J. Weinstein, M. T. Rakher, T. D. Ladd, and M. G. Borselli, Nature Nanotechnology **14**, 747 (2019).

[38] M. J. Curry, T. D. England, N. C. Bishop, G. Ten-Eyck, J. R. Wendt, T. Pluym, M. P. Lilly, S. M. Carr, and M. S. Carroll, Appl. Phys. Lett. **106**, 203505 (2015).

[39] L. A. Tracy, E. P. Nordberg, R. W. Young, C. B. Pinilla, H. L. Stalford, G. A. T. Eyck, K. Eng, K. D. Childs, J. R. Wendt, R. K. Grubbs, J. Stevens, M. P. Lilly, M. A. Eriksson, and M. S. Carroll, Appl. Phys. Lett. **97**, 192110 (2010).

[40] A. Rossi, R. Zhao, A. S. Dzurak, and M. F. Gonzalez-Zalba, Applied Physics Letters **110**, 212101 (2017).

[41] A. West, B. Hensen, A. Jouan, T. Tanttu, C.-H. Yang, A. Rossi, M. F. Gonzalez-Zalba, F. Hudson, A. Morello, D. J. Reilly, and A. S. Dzurak, Nature Nanotech. https://doi.org/10.1038/s41565-019-0400-7 (2019).

[42] G. Zheng, N. Samkharadze, M. L. Noordam, N. Kalhor, D. Brousse, A. Sammak, G. Scappucci, and L. M. K. Vandersypen, Nature Nanotechnology **14**, 742 (2019).

[43] Y. Y. Liu, S. G. J. Philips, L. A. Orona, N. Samkharadze, T. McJunkin, E. R. MacQuarrie, M. A. Eriksson, L. M. K. Vandersypen, and A. Yacoby, Radio frequency reflectometry in silicon-based quantum dots (2021), arXiv:2012.14560 .

[44] A. Sigillito, J. Loy, D. Zajac, M. Gullans, L. Edge, and J. Petta, Phys. Rev. Applied **11**, 061006 (2019).

[45] D. M. Zajac, T. M. Hazard, X. Mi, E. Nielsen, and J. R. Petta, Phys. Rev. Appl. **6**, 054013 (2016).

[46] J. P. Zwolak, T. McJunkin, S. S. Kalantre, J. Dodson, E. MacQuarrie, D. Savage, M. Lagally, S. Coppersmith, M. A. Eriksson, and J. M. Taylor, Phys. Rev. Applied **13**, 034075 (2020).

[47] S. Nadj-Perge, S. M. Frolov, E. P. A. M. Bakkers, and L. P. Kouwenhoven, Nature **468**, 1084 (2010).

[48] B. M. Freeman, J. S. Schoenfield, and H. Jiang, Applied Physics Letters **108**, 253108 (2016).

[49] F. Schäffler, Semicond. Sci. Tech. **12**, 1515 (1997).

[50] M. Friesen, S. Chutia, C. Tahan, and S. N. Coppersmith, Phys. Rev. B **75**, 115318 (2007).

[51] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Rev. Mod. Phys. **85**, 961 (2013).

[52] N. Shaji, C. B. Simmons, M. Thalakulam, L. J. Klein, H. Qin, H. Luo, D. E. Savage, M. G. Lagally, A. J. Rimberg, R. Joynt, M. Friesen, R. H. Blick, S. N. Coppersmith, and M. A. Eriksson, Nat. Phys. **4**, 540 (2008).

[53] C. B. Simmons, T. S. Koh, N. Shaji, M. Thalakulam, L. J. Klein, H. Qin, H. Luo, D. E. Savage, M. G. Lagally, A. J. Rimberg, R. Joynt, R. Blick, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Phys. Rev. B **82**, 245312 (2010).

[54] Z. Shi, C. B. Simmons, J. Prance, J. K. Gamble, M. Friesen, D. E. Savage, M. G. Lagally, S. N. Coppersmith, and M. A. Eriksson, Appl. Phys. Lett. **99**, 233108 (2011).

[55] M. G. Borselli, R. S. Ross, A. A. Kiselev, E. T. Croke, K. S. Holabird, P. W. Deelman, L. D. Warren, I. Alvarado-Rodriguez, I. Milosavljevic, F. C. Ku, W. S. Wong, A. E. Schmitz, M. Sokolich, M. F. Gyure, and A. T. Hunter, Appl. Phys. Lett. **98**, 123118 (2011).

[56] J. S. Schoenfield, B. M. Freeman, and H. Jiang, Nature Commun. **8**, 64 (2017).

[57] X. Mi, C. G. Péterfalvi, G. Burkard, and J. R. Petta, Phys. Rev. Lett. **119**, 176803 (2017).

[58] A. Jones, E. Pritchett, E. Chen, T. Keating, R. Andrews, J. Blumoff, L. De Lorenzo, K. Eng, S. Ha, A. Kiselev, S. Meenehan, S. Merkel, J. Wright, L. Edge, R. RS, R. MT, B. MG, and A. Hunter, Physical Review Applied **12**, 014026 (2019).

[59] A. Hollmann, T. Struck, V. Langrock, A. Schmidbauer, F. Schauer, T. Leonhardt, K. Sawano, H. Riemann, N. V. Abrosimov, D. Bougeard, and L. R. Schreiber, Phys. Rev. Applied **13**, 034068 (2020).

[60] E. H. Chen, K. Raach, A. Pan, A. A. Kiselev, E. Acuna, J. Z. Blumoff, T. Brecht, M. Choi, W. Ha, D. Hulbert, M. P. Jura, T. Keating, R. Noah, B. Sun, B. J. Thomas, M. Borselli, C. A. C. Jackson, M. T. Rakher, and R. S. Ross, Detuning axis pulsed spectroscopy of valley-orbital states in Si/SiGe quantum dots (2020), arXiv:2010.04818 .

[61] Y. Hada and M. Eto, Phys. Rev. B **68**, 155322 (2003).

[62] M. Friesen and S. N. Coppersmith, Phys. Rev. B **81**, 115324 (2010).

[63] S. Pecker, F. Kuemmeth, A. Secchi, M. Rontani, D. C. Ralph, P. L. McEuen, and S. Ilani, Nature Physics **9**, 576 (2013).

[64] J. P. Dodson, H. E. Ercan, J. Corrigan, M. Losert, N. Holman, T. McJunkin, L. F. Edge, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, How valley-orbit states in silicon quantum dots probe quantum well interfaces (2021), arXiv:2103.14702 .

[65] Z. Shi, C. B. Simmons, J. R. Prance, J. K. Gamble, T. S. Koh, Y.-P. Shim, X. Hu, D. E. Savage, M. G. Lagally, M. A. Eriksson, M. Friesen, and S. N. Coppersmith, Phys. Rev. Lett. **108**, 140503 (2012).

[66] T. S. Koh, J. K. Gamble, M. Friesen, M. A. Eriksson, and S. N. Coppersmith, Phys. Rev. Lett. **109**, 250503 (2012).

[67] N. E. Penthorn, J. S. Schoenfield, L. F. Edge, and H. Jiang, Phys. Rev. Applied **14**, 054015 (2020).

[68] G. M. Xia, J. L. Hoyt, and M. Canonico, Journal of Applied Physics **101**, 044901 (2007).

[69] L. DiCarlo, H. J. Lynch, A. C. Johnson, L. I. Childress, K. Crockett, and C. M. Marcus, Phys. Rev. Lett. **92**, 226801 (2004).

[70] R. Dingle, H. Stormer, A. Gossard, and W. Wiegmann, Appl. Phys. Lett. **33**, 665 (1978).

[71] R. People, J. C. Bean, D. V. Lang, A. M. Sergent, H. L. Störmer, K. W. Wecht, R. T. Lynch, and K. Baldwin, Applied Physics Letters **45**, 1231 (1984).

[72] M. R. Sakr, H. W. Jiang, E. Yablonovitch, and E. T. Croke, Appl. Phys. Lett. **87**, 223104 (2005).

[73] C. B. Simmons, M. Thalakulam, N. Shaji, L. J. Klein, H. Qin, R. H. Blick, D. E. Savage, M. G. Lagally, S. N. Coppersmith, and M. A. Eriksson, Appl. Phys. Lett. **91**, 213103 (2007).

[74] R. R. Hayes, A. A. Kiselev, M. G. Borselli, S. S. Bui, E. T. C. III, P. W. Deelman, B. M. Maune, I. Milosavljevic, J.-S. Moon, R. S. Ross, A. E. Schmitz, M. F. Gyure, and A. T. Hunter, Lifetime measurements (T1) of electron spins in Si/SiGe quantum dots (2009), arXiv:0908.0173 .

[75] C. B. Simmons, J. R. Prance, B. J. Van Bael, T. S. Koh, Z. Shi, D. E. Savage, M. G. Lagally, R. Joynt, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Phys. Rev. Lett. **106**, 156804 (2011).

[76] J. R. Prance, Z. Shi, C. B. Simmons, D. E. Savage, M. G. Lagally, L. R. Schreiber, L. M. K. Vandersypen, M. Friesen, R. Joynt, S. N. Coppersmith, and M. A. Eriksson, Phys. Rev. Lett. **108**, 046808 (2012).

[77] Z. Shi, C. B. Simmons, D. R. Ward, J. R. Prance, T. S. Koh, J. K. Gamble, X. Wu, D. E. Savage, M. G. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, Phys. Rev. B **88**, 075416 (2013).

[78] M. G. Borselli, K. Eng, R. S. Ross, T. M. Hazard, K. S. Holabird, B. Huang, A. A. Kiselev, P. W. Deelman, L. D. Warren, I. Milosavljevic, A. E. Schmitz, M. Sokolich, M. F. Gyure, and A. T. Hunter, Nanotechnology **26**, 375202 (2015).

[79] S. J. Angus, A. J. Ferguson, A. S. Dzurak, and R. G. Clark, Nano Lett. **7**, 2051 (2007).

[80] P. C. Spruijtenburg, S. V. Amitonov, W. G. van der Wiel, and F. A. Zwanenburg, Nanotechnology **29**, 143001 (2018).

[81] K. R. Williams, K. Gupta, and M. Wasilik, Journal of Microelectromechanical Systems **12**, 761 (2003).

[82] H. Jansen, M. de Boer, R. Legtenberg, and M. Elwenspoek, Journal of Micromechanics and Microengineering **5**, 115 (1995).

[83] `http://www.srim.org/`.

[84] T. F. Rosenbaum, R. F. Milligan, M. A. Paalanen, G. A. Thomas, and R. N. Bhatt, Phys Rev B **27**, 7509 (1983).

[85] J. P. Dodson, N. Holman, B. Thorgrimsson, S. F. Neyens, E. R. MacQuarrie, T. McJunkin, R. H. Foote, L. F. Edge, S. N. Coppersmith, and M. A. Eriksson, Nanotechnology **31**, 505001 (2020).

[86] E. J. Connors, J. Nelson, H. Qiao, L. F. Edge, and J. M. Nichol, Phys. Rev. B **100**, 165305 (2019).

[87] M. Ciorga, A. S. Sachrajda, P. Hawrylak, C. Gould, P. Zawadzki, S. Jullian, Y. Feng, and Z. Wasilewski, Phys. Rev. B **61**, R16315 (2000).

[88] E. A. Laird, J. M. Taylor, D. P. DiVincenzo, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. B **82**, 075403 (2010).

[89] W. A. Coish and D. Loss, Phys. Rev. B **70**, 195340 (2004).

[90] M. G. Borselli, K. Eng, E. T. Croke, B. M. Maune, B. Huang, R. S. Ross, A. A. Kiselev, P. W. Deelman, I. Alvarado-Rodriguez, A. E. Schmitz, M. Sokolich, K. S. Holabird, T. M. Hazard, M. F. Gyure, and A. T. Hunter, Appl. Phys. Lett. **99**, 063109 (2011).

[91] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, A. Morello, A. Laucht, and A. S. Dzurak, Nature **569**, 532 (2019).

[92] D. M. Zajac, T. M. Hazard, X. Mi, K. Wang, and J. R. Petta, Appl. Phys. Lett. **106**, 223507 (2015).

[93] S. F. Neyens, E. MacQuarrie, J. Dodson, J. Corrigan, N. Holman, B. Thorgrimsson, M. Palma, T. McJunkin, L. Edge, M. Friesen, S. Coppersmith, and M. Eriksson, Phys. Rev. Applied **12**, 064049 (2019).

[94] M. Field, C. G. Smith, M. Pepper, D. A. Ritchie, J. E. F. Frost, G. A. C. Jones, and D. G. Hasko, Phys. Rev. Lett. **70**, 1311 (1993).

[95] A. Quade, H. Wulff, H. Steffen, T. Tun, and R. Hippler, Thin Solid Films **377**, 626 (2000).

[96] S. Gupta, S. Hannah, C. Watson, P. Šutta, R. Pedersen, N. Gadegaard, and H. Gleskova, Organic Electronics **21**, 132 (2015).

[97] F. P. Fehlner, *Low temperature oxidation, the role of vitreous oxides* (John Wiley & Sons, Inc., 1986).

[98] L. Nguyen, T. Hashimoto, D. N. Zakharov, E. A. Stach, A. P. Rooney, B. Berkels, G. E. Thompson, S. J. Haigh, and T. L. Burnett, ACS Applied Materials & Interfaces **10**, 2230 (2018).

[99] J. Evertsson, F. Bertram, F. Zhang, L. Rullik, L. Merte, M. Shipilin, M. Soldemo, S. Ahmadi, N. Vinogradov, F. Carlà, J. Weissenrieder, M. Göthelid, J. Pan, A. Mikkelsen, J.-O. Nilsson, and E. Lundgren, Applied Surface Science **349**, 826 (2015).

[100] B. R. Strohmeier, Surface and Interface Analysis **15**, 51 (1990).

[101] I. Iatsunskyi, M. Kempiński, M. Jancelewicz, K. Załęski, S. Jurga, and V. Smyntyna, Vacuum **113**, 52 (2015).

[102] R. Y. Khosa, E. B. Thorsteinsson, M. Winters, N. Rorsman, R. Karhu, J. Hassan, and E. O. Sveinbjörnsson, AIP Advances **8**, 025304 (2018).

[103] C. M. Tanner, Y.-C. Perng, C. Frewin, S. E. Saddow, and J. P. Chang, Applied Physics Letters **91**, 203510 (2007).

[104] J. Yota, H. Shen, and R. Ramanathan, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **31**, 01A134 (2013).

[105] W. H. Lim, H. Huebl, L. H. Willems van Beveren, S. Rubanov, P. G. Spizzirri, S. J. Angus, R. G. Clark, and A. S. Dzurak, Appl. Phys. Lett. **94**, 173502 (2009).

[106] M. Brauns, S. V. Amitonov, P.-C. Spruijtenburg, and F. A. Zwanenburg, Scientific Reports **8** (2018).

[107] H. C. Lin, P. D. Ye, and G. D. Wilk, Applied Physics Letters **87**, 182904 (2005).

[108] Y. Q. Wu, H. C. Lin, P. D. Ye, and G. D. Wilk, Applied Physics Letters **90**, 072105 (2007).

[109] F. Müller, *Single-charge tunneling in ambipolar silicon quantum dots*, Ph.D. thesis, University of Twente, Netherlands (2015).

[110] P. Nealey, M. Stoykovich, K. Yoshimoto, and H. Cao, in *Encyclopedia of Materials: Science and Technology* (Elsevier, 2003) pp. 1–9.

[111] M. Friesen, M. A. Eriksson, and S. N. Coppersmith, Appl. Phys. Lett. **89**, 202106 (2006).

[112] E. A. Irene, Journal of The Electrochemical Society **125**, 1708 (1978).

[113] K. Yoneda, Journal of The Electrochemical Society **142**, 4304 (1995).

[114] S. Wolf and R. Tauber, *Silicon Processing for the VLSI Era: Process Technology*, 2nd ed., Vol. 1 (Lattice Press, 2000).

[115] P. C. Spruijtenburg, S. V. Amitonov, F. Mueller, W. G. van der Wiel, and F. A. Zwanenburg, Scientific Reports **6** (2016).

[116] K. Hinode, I. Asano, and Y. Homma, IEEE Transactions on Electron Devices **36**, 1050 (1989).

[117] E. Galli, G. Majni, C. Nobili, and G. Ottaviani, ElectroComponent Science and Technology **6**, 147 (1980).

[118] J. M. Poate, Gold Bulletin **14**, 2 (1981).

[119] U. Köster, P. Ho, and M. Ron, Thin Solid Films **67**, 35 (1980).

[120] S. Ramanathan, D. Chi, P. C. McIntyre, C. J. Wetteland, and J. R. Tesmer, Journal of The Electrochemical Society **150**, F110 (2003).

[121] P. Heij, *Single-charge transport in coupled nanostructures*, Ph.D. thesis, Delft University of Technology (2019).

[122] L. Vandersypen, H. Bluhm, J. Clarke, A. Dzurak, R. Ishihara, A. Morello, D. Reilly, L. Schreiber, and M. Veldhorst, npj Quantum Information **3**, 1 (2017).

[123] M. Veldhorst, H. Eenink, C.-H. Yang, and A. S. Dzurak, Nature Communications **8**, 1 (2017).

[124] X. Mi, M. Benito, S. Putz, D. M. Zajac, J. M. Taylor, G. Burkard, and J. R. Petta, Nature **555**, 599 (2018).

[125] A. Mills, D. Zajac, M. Gullans, F. Schupp, T. Hazard, and J. Petta, Nature Communications **10**, 1 (2019).

[126] N. Holman, J. Dodson, L. Edge, S. Coppersmith, M. Friesen, R. McDermott, and M. Eriksson, Appl. Phys. Lett. **117**, 083502 (2020).

[127] K. Takeda, A. Noiri, T. Nakajima, J. Yoneda, T. Kobayashi, and S. Tarucha, Quantum tomography of an entangled three-spin state in silicon (2020), arXiv:2010.10316 .

[128] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[129] T. B. Boykin, G. Klimeck, M. A. Eriksson, M. Friesen, S. N. Coppersmith, P. von Allmen, F. Oyafuso, and S. Lee, Appl. Phys. Lett. **84**, 115 (2004).

[130] N. Kharche, M. Prada, T. B. Boykin, and G. Klimeck, Appl. Phys. Lett. **90**, 092109 (2007).

[131] D. Culcer, X. Hu, and S. Das Sarma, Phys. Rev. B **82**, 205315 (2010).

[132] J. K. Gamble, M. A. Eriksson, S. N. Coppersmith, and M. Friesen, Phys. Rev. B **88**, 035310 (2013).

[133] P. Boross, G. Széchenyi, D. Culcer, and A. Pályi, Physical Review B **94**, 035438 (2016).

[134] A. Hosseinkhani and G. Burkard, Phys. Rev. Research **2**, 043180 (2020).

[135] A. Szabo, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory (Dover Books on Chemistry)* (Dover Publications, 1996).

[136] T. B. Boykin, G. Klimeck, M. Friesen, S. N. Coppersmith, P. vonAllmen, F. Oyafuso, and S. Lee, Phys. Rev. B **70**, 165325 (2004).

[137] C. H. Yang, W. H. Lim, N. S. Lai, A. Rossi, A. Morello, and A. S. Dzurak, Phys. Rev. B **86**, 115319 (2012).

[138] J. K. Gamble, P. Harvey-Collard, N. T. Jacobson, A. D. Baczewski, E. Nielsen, L. Maurer, I. Montaño, M. Rudolph, M. Carroll, C. Yang, A. Rossi, A. Dzurak, and R. P. Muller, Appl. Phys. Lett. **109**, 253101 (2016).

[139] W. H. Lim, F. A. Zwanenburg, H. Huebl, M. Möttönen, K. W. Chan, A. Morello, and A. S. Dzurak, Appl. Phys. Lett. **95**, 242102 (2009).

[140] W. H. Lim, C. H. Yang, F. A. Zwanenburg, and A. S. Dzurak, Nanotechnology **22**, 335704 (2011).

[141] K. MacLean, S. Amasha, I. P. Radu, D. M. Zumbühl, M. A. Kastner, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **98**, 036802 (2007).

[142] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, Rev. Mod. Phys. **79**, 1217 (2007).

[143] W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Rev. Mod. Phys. **75**, 1 (2003).

[144] C. Kittel, H. Charles Kittel, K. Charles, H. Kroemer, and K. Herbert, *Thermal Physics* (W. H. Freeman, 1980).

[145] M. Stopa, Phys. Rev. B **54**, 13767 (1996).

[146] B. P. Wuetz, M. P. Losert, A. Tosato, M. Lodari, P. L. Bavdaz, L. Stehouwer, P. Amin, J. S. Clarke, S. N. Coppersmith, A. Sammak, M. Veldhorst, M. Friesen, and G. Scappucci, Phys. Rev. Lett. **125**, 186801 (2020).

[147] H. E. Ercan, S. N. Coppersmith, and M. A. Eriksson (2020), unpublished.