ABSTRACT

Title of Dissertation: DEFECTS AND STRAIN IN SILICON METAL-OXIDE-SEMICONDUCTOR (MOS) QUANTUM DOTS

Ryan Matthew Stein
Doctor of Philosophy, 2021

Dissertation Directed by: Professor John Cumings, Department of Materials Science and Engineering

Dr. Michael D. Stewart, Jr.
National Institute of Standards and Technology

Silicon-based single electron devices (SEDs), fabricated using gate-defined quantum dots are some of the world's most sensitive devices. Local charge fluctuations and disorder caused by defects in the oxide or substrate impurities can profoundly affect device operation. While most workers consider the above when fabricating SEDs in the Si MOS system, they do not typically consider strain. The fabrication process and gate material deposition usually results in a thin film under a significant amount of stress, which locally modulates the silicon conduction band. Additionally, the coefficient of thermal expansion mismatch between typical MOS gate materials, such as aluminum, and the underlying silicon substrate also produces strain, which further modifies the conduction band. For quantum dot devices measured at cryogenic temperatures, this local modification of the conduction band is strong enough to lead to the formation of unintentional quantum dots and to af-

fect the tunnel coupling between dots. To realize the potential of quantum devices, gate-induced strain must be understood so as to be mitigated or exploited.

In this work, we investigate the role of gate-induced strain in quantum dot devices by comparing measurements of the 4-terminal $I(V)$ characteristics of tunnel barrier devices at cryogenic temperatures. From this, we demonstrate a new electrical measurement of gate-induced strain using tunnel junctions (TJs). Our COMSOL simulations of these devices show that the gate-induced strain will modify the barrier height, depending on both the magnitude and sign of inhomogeneous stress. We fabricate MOS devices on bulk silicon wafers with a variety of gate electrodes, including aluminum and titanium. By comparing nearly identical tunnel junction devices fabricated with two different gate materials, Al and Ti, we measure a relative strain difference consistent with our experimentally measured coefficients of thermal expansion. Our results show that the commonly used bulk parameters for simulating strain effects in silicon QDs do not work well in practice.

Additionally, we present measurements of oxide defect densities (fixed charge and interface trap density) as a function of forming gas anneal temperature for three different gate metals: Al, Ti/Pd, and Ti/Pt. We also investigate the effect of these anneals on the mechanical properties of the gate material, such as the intrinsic film stress and coefficient of thermal expansion. The combination of our charge defect and mechanical measurements show that there is no way to simultaneously minimize the effects of both using the forming gas anneal. This result puts tension on designing fabrication processes for MOS QDs where one must choose between setting the anneal such that defects are minimized or the strain-induced modulation

of the conduction band is minimized. Additionally, we find that our measured values of the coefficient of thermal expansion deviate significantly from the expected bulk values. This suggests that the common material parameters used to simulate gate-induced strain in MOS QD are not accurate.

Building towards the goal of controlling non-idealities in silicon MOS QDs requires methods of measuring strain under relevant conditions while also finding ways to adjust processing to minimize the impact of other non-idealities. The work in thesis represents a significant step towards that goal. The devices presented easily lend themselves to future work exploring deposition parameters and anneals to manipulate inhomogeneous strain. Our method for measuring relative strain satisfies the sensitivity, spatial resolution and low-temperature requirements relevant for MOS QDs. Moreover, the fabrication and measurements are similar to those for QDs so that this method is directly relevant for QD devices. Our data provide an important step forward in assessing gate-induced strain in QD devices in-situ while highlighting the need for further experimental work and a greater theoretical understanding of the electrostatics and strain behavior.

DEFECTS AND STRAIN IN SILICON METAL-OXIDE-SEMICONDUCTOR (MOS) QUANTUM DOTS

by

Ryan Matthew Stein

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor John Cumings, Chair/Advisor
Dr. Michael D. Stewart Jr., Co-Advisor
Professor Ichiro Takeuchi
Professor Lourdes G. Salamanca-Riba
Professor Martin C. Peckerar

Foreword

Chapter 4 is partially reproduced from: "The effect of strain on tunnel barrier height in silicon quantum devices", Journal of Applied Physics, 128(2):024303, July 2020. with the permission of AIP publishing. (I am the first author of this paper).

Chapter 5 is partially reproduced from: "Alternatives to aluminum gates for silicon quantum devices: defects and strain" currently under review by the Journal of Applied Physics. (I am the first author of this paper).

Acknowledgments

During last six years as a PhD student working at NIST, I have had the privilege interacting with many people around me who have enabled me to thrive during my PhD career. First, I would like to thank my advisor at NIST, Dr. Michael Stewart, Jr., for giving me an invaluable opportunity to work at NIST and guiding me through my research the last six years. He has always made himself available when I needed help or advice even in the probably too frequent occurrences where I wandered into his office to discuss some random idea or problem I had. I would also like to thank Professor John Cumings, my advisor at College Park, for his helpful guidance during my PhD career.

I would also like to thank all of the group members I have crossed paths with over the last six years. I am grateful to Roy Murray and Binhui Hu, both of whom were post-docs early on during my time at NIST, for their help when I was first learning how to fabricate devices in the cleanroom. I would also like to thank the PIs at NIST: Dr. Neil Zimmerman, Dr. Josh Pomeroy, and Dr. Garnett Bryant, for providing feedback and support on my research particularly during our group meetings and journal clubs. I would also like the other group members whom I have worked closely with during my time at NIST: Dr. Aruna Ramanayaka, Zac Barcikowski, Yanxue Hong, Dr. Ke Tang, Dr. Hyun Soo Kim, and Dr. Kevin Dwyer.

Finally, I would to thank all of my committee members: Professor Ichiro Takeuchi, Professor Lourdes G. Salamanca-Riba and Professor Martin Peckerar for reviewing and helping me improve this dissertation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Silicon-based single electron devices (SEDs), fabricated using gate-defined quantum dots are some of the world's most sensitive devices. The sensitivity and single-electron nature of SEDs leads to a wide array of potential applications including nanoscale electrometry [3, 4] and thermometry [5], low power logic and memory [6, 7], electrical metrology as a quantum current standard [8, 9, 10], and solid-state quantum computing [11, 12, 13]. Local inhomogeneity and disorder caused by defects in the oxide, substrate impurities, or strain can profoundly affect device operation. Characterizing and controlling the sources of inhomogeneity is key to the full realization of any of the applications of silicon-based SEDs.

This chapter provides an overview of the fundamental physics and limitations of SEDs in silicon. The building block of the silicon SEDs we are interested in this work are quantum dots (QDs). Although most of the results presented in this work are not directly from QDs, all of results will be relevant towards their fabrication and design. In the first section, we discuss the major applications of interest specifically for quantum computing and quantum current standards. Next, we will provide some background information on the basic properties of generic QDs. Using this background; we will then focus our discussion towards QDs in silicon and how they

are impacted by non-idealities, such as charge defects and inhomogeneous strain. The final portion of this chapter will lay the groundwork for understanding the interplay between these non-idealities and QDs.

1.0.1 Motivation

Quantum dots (QDs) are conducting regions where charge can be isolated using electrostatic potentials. The confinement must be strong enough in 3-dimensions so that the energy levels associated with system are no longer degenerate. In this way, the discrete states of the QD are similar to that found in atoms so that QDs are sometimes referred to as artificial atoms. In reality, a QD is composed of a large number of atoms with a comparable number of electrons and holes bound to the nuclei. The properties of QDs, such as charge transport and bias dependence, are ultimately set by a much smaller number, from one to a few hundred, of free carriers isolated on the dot. In general, a QD can be occupied by either electrons or holes. For simplicity, the discussion in the following sections will focus on electrons, but similar physical arguments could be easily applied to holes.

1.0.2 Quantum computing

A quantum computer has the potential to efficiently solve certain computational problems, which have no efficient solution on a classical computer. Quantum computation works with quantum bits, or qubits. In contrast to a classical bit that can be described as either 0 or 1, a qubit can be in a superposition of quantum

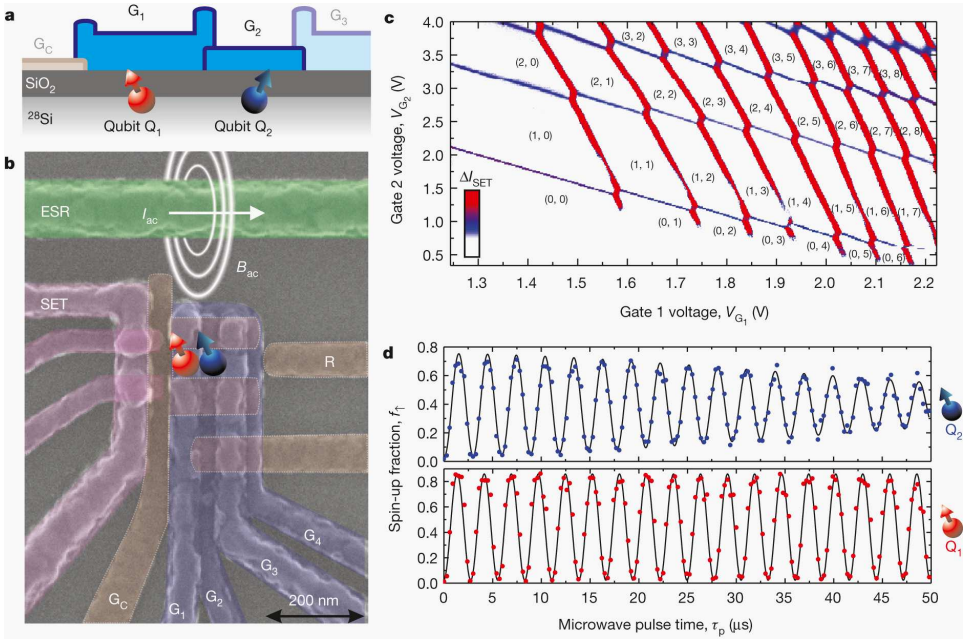

Figure 1.1: (a) Cross section schematic of a silicon MOS two qubit device. (b) SEM image of the two-qubit device. (c) Charge stability diagram of the two qubit device where gates G_1 and G_2 both have a single electron confined under them. (d) Individual Rabi oscillations for each qubit performed via ESR. Reprinted from Veldhorst, M., Yang, C., Hwang, J. et al. A two-qubit logic gate in silicon. Nature 526, 410–414 (2015). https://doi.org/10.1038/nature15263 with the permission of Spring Nature [11].

states $|0>$ or $|1>$. Any two level quantum system with two well-defined quantum states that can be coupled could act as a qubit. Here, well-defined refers to true a two level system with no other states for the system to leak into. In practice, the requirements for building a quantum computer are much more complicated. In 2000, DiVincenzo identified several other important requirements in what is now referred to as "DiVincenzo Criteria" [14]:

1. A scalable physical system with well characterized qubits.

2. The ability to initialize the state of the qubits.

3. Coherence times long compared to the gate operation time.

4. The ability to perform one and two-qubit gates.

5. The ability to read-out the qubit state.

Over the past few decades, a wide array of different systems have been demonstrated and/or proposed as architectures for building a quantum computer. This includes trapped ions [15, 16],superconducting circuits [17, 18, 19, 20], neutral atoms in a optical lattice [21, 22], photonic devices [23, 24], liquid-state nuclear magnetic resonance [25, 26], topological qubits [27, 28], optical color centers in diamond [29, 30] and semiconducting devices [13, 31, 32]. The semiconductor spins approach to quantum computation offers a high degree of control over the device design, fabrication, and operation. In 1998, Loss and DiVincenzo proposed the use of electron spins in semiconductor QDs as qubits [13]. Here, the spin of the electron, which can be $\pm\frac{1}{2}$, is used to encode the qubit state with the two basis states being the spin-up and spin-down states. This proposal was built upon the usage of QDs in semiconducting structures, where a qubit is composed of a single QD that has a single confined electron. Fig. 1.1 shows an example of a silicon MOS based two-qubit device. Here, two QDs each occupied with one electron are formed under the gates labeled G_1 and G_2 in Fig. 1.1 (a) and (b). It is important to note that even though there are only two qubits in the device in Fig. 1.1 that gate layout is already complex. In addition to the gates G_1 and G_2 for controlling the qubits, there are layers for controlling the size of the QD (G_C), moving electrons from a reservoir onto or off the QDs (R), and an extra QD used as a charge sensor for the qubits (labeled SET

for single electron transistor). This level of complexity for only a two qubit device means that understanding the limitations and improving the MOS QD fabrication process and device design is key to achieving the scalability necessary for quantum computing.

1.0.3 Quantum current standard

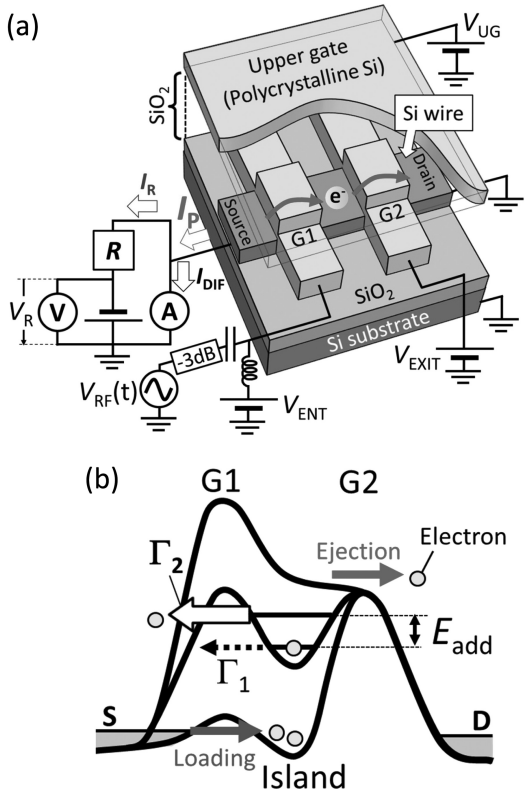

Figure 1.2: (a) Schematic of silicon MOS QD device used as single electron pump. (b) Energy diagram in the silicon device showing the potential modulation used to pump electrons. Reprinted from Gento Yamahata, Stephen P. Giblin, Masaya Kataoka, Takeshi Karasawa, and Akira Fujiwara , "Gigahertz single-electron pumping in silicon with an accuracy better than 9.2 parts in 10^7", Applied Physics Letters 109, 013101 (2016) https://doi.org/10.1063/1.4953872 with the permission of AIP Publishing [33].

The ampere, which is the SI base unit for current, has no good standard for use in metrology. A previous standard defined the ampere as a "constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to $2 \times 10^{-7} N$ per meter of length" [34]. From a practical perspective, this current standard is difficult to achieve. In 2019, the SI base units were redefined in terms of fundamental physical constants [35]. Thus, the ampere was defined in terms of the elementary charge of the electron e and the hyperfine transition frequency of cesium-133, $\Delta\nu_{cs}$. While the ampere is a base unit of the SI, as a practical matter it is treated as a derived unit from the voltage and resistance standards. In the case of other electrical SI units, such as resistance (R) and voltage (V), the standard is based on a physical system derived from quantum properties and ultimately derived from fundamental physical constants such as elementary charge (e) and Planck's constant (h). For the resistance standard, the quantum standard for the Ohm is based on the Quantum Hall effect [36]. Here, the quantum hall resistance is given as $R_{QHR} = \frac{h}{e^2}$. Similarly, the quantum voltage standard is based on the Josephson effect [37] and given as $V_{JV} = \frac{nhf}{2e}$. Based on the quantum resistance and voltage standards, a much simpler and practical current standard would be based on the charge of an electron. A single electron pump [38] is a device that acts as a quantum standard analogous to resistance and voltage, but for current. Similar to applications in quantum computing with semiconducting devices, the building block of single electron pump is the QD. An example of a silicon based single electron pump is shown in Fig. 1.2. Here, a QD is induced between the

gates G_1 and G_2 in a silicon nanowire device. By applying an AC voltage to G_1, electrons can be shuttled through the device from the source (S) to the QD then off to the drain (D). In the ideal case, this process will eject a single electron from the QD to the drain once every cycle of the AC signal on G_1. If the AC signal has a frequency, f, then the current (I_P) through the pump will be given by: $I_P = ef$. For a typical pumping frequency around 1 GHz, the single electron pump will generate a current around 160 pA.

1.1 Quantum dots in Silicon

1.1.1 Why silicon?

Many of the early demonstrations of SEDs in semiconductor devices were in GaAs/AlGaAs heterostructures [39, 40, 41, 42, 43]. In fact much of the foundational work on quantum computing with semiconductor spins was performed in SEDs in GaAs including coherent manipulation of a single spin [12, 44], electrical single shot readout of a spin state [45], and the implementation of two-qubit gate operations [46]. From the perspective of quantum computing, one of the major problems with making qubits in III-V heterostructures, such as the GaAs/AlGaAs, is the relative abundance of non-zero nuclear spins. Any qubits fabricated in such systems will couple to this nuclear spin bath and experience decoherence [47], which refers to how long a quantum state can be held in a superposition before losing information to its environment as described by the third DiVincenzo criteria discussed above. Silicon has a significant advantage in this area since its most common nuclear spin isotope

(^{28}Si, 92% in natural abundance) [48] has zero spin and the non-zero spin isotope (^{29}Si) is only around 4.7% abundance in natural silicon. Additionally, methods exist to isotopically purify it such that the nuclear spin bath no longer dominates the sources of qubit decoherence [49].

For charge pumps, devices have also been fabricated in GaAs/AlGaAs [10, 50] and silicon [9, 33]. One of the major differences between charge pumps in the two material systems is that GaAs/AlGaAs devices typically have used large magnetic fields to achieve their reported accuracies and performance [51]. Silicon devices don't appear to need such large magnetic fields to achieve the same current and uncertainty. These large magnetic fields are not compatible with the quantum resistance and voltage standards. Since one of the desired goals of any quantum current standard it to perform the experiments using all three standards, silicon-based charge pumps have an advantage over GaAs.

For applications in quantum computing and a quantum current standard, the scalability of the devices is a key component in the choice of the material system. For quantum computing, if one qubit is formed by a single QD we will need millions of QDs to build a functional quantum computer. Similarly for a quantum current standard to make it comparable to the voltage and resistance standards, it would be advantageous to have a quantized current on the order of 1 μA [52]. This would necessitate running many charge pumps in parallel. Silicon-based devices have a significant advantage from the scalability and fabrication perspective due to ubiquitous usage of silicon in industrial transistors. The mature fabrication history and infrastructure of the silicon industry could offer a promising path towards scaling for

large qubit systems and charge pumps. For this reason, silicon is one on the more attractive semiconductor material systems for quantum computing and quantum current standards.

1.1.2 Material systems

Within the subset of silicon SEDs there are three main materials systems of interest: metal-oxide-semiconductor (MOS) heterostructures, Si/SiGe heterostructures, and individual donors. Fig. 1.3 shows an overview of these different material systems. The Si MOS and Si/SiGe devices (the bottom row in Fig. 1.3) are very similar to each other. Both systems rely on a 2-dimensional electron gas (2DEG) in silicon confined at an interface between silicon and a different material: silicon dioxide in the MOS case and silicon germanium alloy in the Si/SiGe case. The formation of QDs in both systems is achieved by using gates to used further confine the 2DEG as shown in the third column of Fig. 1.3. Individual donors are quite different as the donor nuclear potential acts as the source of the 3-dimensional confinement needed for a QD and the gate is used to move electrons on or off the donor. A single donor will also act as a QD with three available states: ionized (D^+), neutral (D^0), and negatively charged (D^-) as shown in the second column of Fig. 1.3. Meanwhile, QDs in silicon MOS and Si/SiGe will show a ladder of many different electron occupations. These different material systems have all shown promising demonstrations of their capabilities. All three systems have demonstrated devices with single electron occupations [53, 54, 55] and coherent manipulation of single

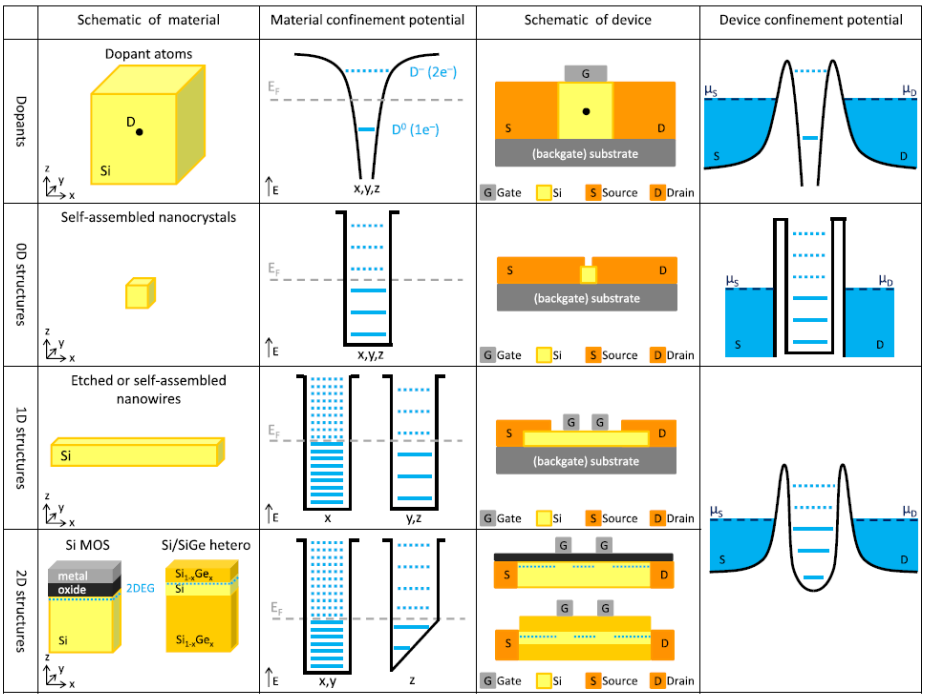

Figure 1.3: Schematic overview of the material stacks and confining potentials used in silicon based SEDs. The first column shows the primary materials used including individual dopants, Si 0D and 1D nanostructures, and 2d electrostatically defined devices in silicon metal-oxide-semiconductor (MOS) and Si/SiGe heterostructures. The second column shows the confining potential used to make the SED. States occupied by electrons are shown as solid lines and unoccupied states are shown as dashed lines. The third column shows a schematic of the typical device structure used to make the SED including the gate used to control the QD potential and the source/drain that act as reservoirs of electrons for the QD. Fourth column shows the potential landscape seen by electron in the QD devices from the third column. Reprinted with permission from Floris A. Zwanenburg et al. Silicon quantum electronics. Reviews of Modern Physics 10.1103/RevModPhys.85.961. Copyright 2012 by the American Physical Society.

spins [11, 56, 57]. In this work, we will focus on SEDs in the silicon MOS system.

1.2 Physics of Quantum Dots (QDs)

In order to understand how non-idealities will affect the potential applications of SEDs discussed above, we will first cover the basics of how QDs operate. In this section, we will define the requirements for making QDs and their ideal electrical properties. This foundation of the ideal behavior will provide the context for how non-idealities ultimately present themselves on the device performance.

1.2.1 Coulomb Blockade

The foundation of quantum dot physics and operation is the effect known as Coulomb blockade. In the case of QDs, we will see that Coulomb blockade will dominate transport and lead to oscillations in the measured current. To understand Coulomb blockade, consider the toy model of an isolated conducting disk. Initially, if the disk is uncharged we can add an electron to it without any energetic cost. Now if we try to add an additional electron, then we must overcome the Coulomb repulsion between the two electrons. The energy needed to overcome the repulsive Coulomb forces and add an additional electron onto the disk is defined as the charging energy (E_{ch}). E_{ch} is simply related to the energy stored in a capacitor with capacitance C as $E_{ch} = \frac{q^2}{2C}$. In the case of this isolated disk, C is the self-capacitance given as $C = 8\pi\epsilon_r r$, where r is the radius of the disk and ϵ_r is the relative permittivity. Importantly, there is an additional constraint we need to satisfy in order to be able

to see the individual charging events due to blockade. Consider a metallic disk with a radius of 100 nm. This gives a self-capacitance of 7 aF. Here, the charging energy would be roughly 11 meV. At room temperature, the thermal energy ($k_bT \approx 26$ meV) would exceed the charging energy so that any single electron charging events would be thermally smeared out. This means in order to observe the single electron charging events signifying blockade we would need to move towards much lower temperatures. For example, at 1 K the thermal energy is roughly 86 μeV or more than two orders of magnitude smaller than the charging energy. In this way, we can now define our first criteria for operation of QDs: $E_{ch} >> k_bT$. From this, we can also infer that we will need QD with small dimensions so that the charging energy stays reasonably large.

1.2.2 Confinement

QDs in silicon are most often formed electrostatically through the application of voltages to gates arranged in specific geometries to confine electrons. Fig. 1.4 shows a generic layout of gate structure for a silicon MOS QD. The operation of the gates in these QD devices is similar to a conventional MOS field effect transistor (MOSFET). In the case of a conventional MOSFET, the gate is used to control the conductivity of the channel formed between two heavily doped regions, referred to as the source and drain. The conducting channel is confined in the direction perpendicular to the oxide-silicon interface (z-direction in Fig. 1.4) by the electric field from gate, forming a 2-dimensional electron gas (2DEG). With the application

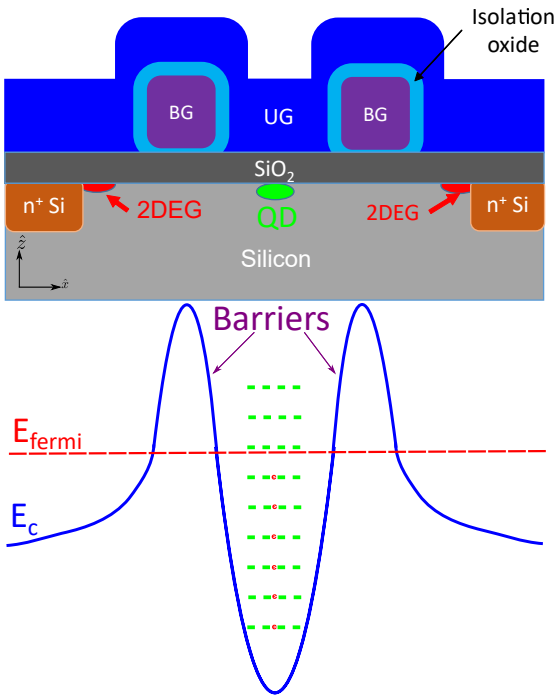

Figure 1.4: Schematic of the formation of QD in silicon MOS device. In this device, there are three gates total: two gates that are used to form barriers(purple) and one gate that used turn on conduction from the source to drain and control the dot chemical potential (blue). The bottom portion of the diagram shows the ideal conduction band profile for such a QD device. The dashed green lines indicate the dot levels and the red circles indicate when the level is occupied.

of a small bias voltage difference between the source and drain (V_{SD}) current will flow between through the device (I_{SD}). When the gate voltage is below a certain value, known as the threshold voltage (V_t), the conduction band in the channel (E_c) is above the fermi level (E_f) in the source/drain regions forming a electrostatic barrier. In this case, no current flows and the MOSFET is "off". When the gate voltage is above V_t, E_c is now less than E_f in the channel, the 2DEG is formed, and the MOSFET is "on". The operation of an QD will follow the same principles as the MOSFET, but in the case of a QD the gate is split up into multiple regions so as to

create a confining potential. The generic QD layout in Fig. 1.4 has a total of three gates, where two of the gates (purple) are used to form tunnel barriers or raise (E_c) above (E_f) selectively in different portions of the channel. The other gate (blue) is used for two purposes: 1) to turn on conduction outside of the barrier regions similar to the normal MOSFET and 2) to control the potential of the region where the QD is formed. In Fig. 1.4, we have drawn a ladder of evenly spaced states in the QD region where the spacing equals the charging energy of the dot. Dot levels that fall below E_f will be occupied (shown as red dots in Fig. 1.4) by electrons that tunnel on to the QD from source/drain.

In order to add or remove electrons, we need to couple the QD to some reservoir(s) of electrons. This coupling comes from the tunnel barriers discussed above and shown in Fig. 1.4. The tunnel barrier must be resistive enough that the tunneling lifetime, or the RC time constant of the QD, and charging energy of the dot satisfy the Heisenberg uncertainty principle, $\Delta E \Delta t > \frac{\hbar}{2}$, where $\Delta E = \frac{e^2}{2C}$ and $\Delta t = RC$. In more practical terms, this means the tunnel barrier resistance (R) must be greater than the resistance quantum ($R_Q = \frac{h}{e^2}$) or 25.8 $k\Omega$. This forms the second criteria we need in order to see single charging events in QD devices. Importantly, this only sets the lower limit of the barrier resistance. The barrier resistance must also not be so high such that the current is not measurable. This means that ultimately the barrier resistance will need to be less than $G\Omega$.

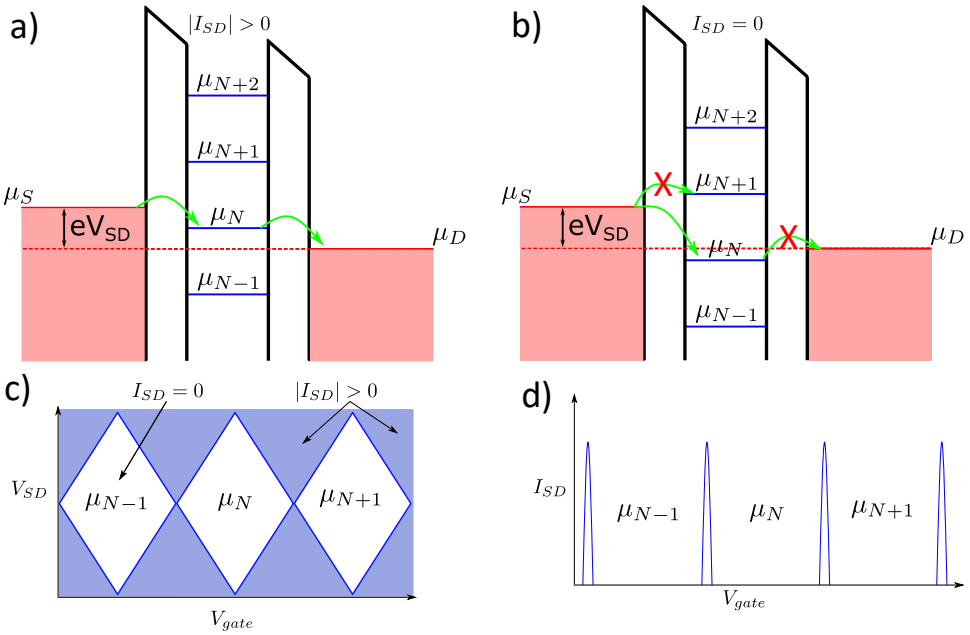

Figure 1.5: (a) Energy diagram of a quantum dot when there is current through the device ($|I_{SD}| > 0$), where μ_i refers to the chemical potential of the ith level. Here, electrons from the source (μ_S) are able to sequentially tunnel onto the dot level (μ_N) and onto the drain (μ_D). This is because μ_N lies between the source and drain chemical potentials, which are separated by a small source-drain bias voltage (V_{SD}). (b) Energy diagram of a quantum dot when there is no current ($I_{SD} = 0$) and the device is the Coulomb blockade regime. Here, electrons from the source (μ_S) are able to tunnel onto the dot level (μ_N) but it is not energetically favorable to tunnel onto the drain since μ_N is no longer in between μ_S and μ_D. (c) 2d transport diagram of an ideal quantum dot, where the x-axis is the gate voltage used to control the dot chemical potential and y-axis is the source-drain bias. Inside the diamonds, the device is blockaded as shown in (b) and there is no current. Outside of the diamonds, there is transport through the device as shown in (a) and there is a measurable current. (d) Schematic showing the expected behavior of current through the quantum dot with changes in gate voltage at constant source-drain bias.

1.2.3 Transport characteristics

Transport through the QDs will only occur when at least one of the discrete dot chemical potential levels (μ) is within the bias window formed by the source-drain bias voltage (ie $\mu_S > \mu_N > \mu_D$), as illustrated in Fig. 1.5(a). When the source-drain bias is small enough ($eV_{SD} < E_{ch}$), and the temperature is low enough, there may be only one state in the transport window and current through the device will be measured. In Fig 1.5(a), the chemical potential level μ_{N-1} is occupied by an electron that has tunneled on to the dot from the source since $\mu_{N-1} < \mu_S$, but the electron is unable to tunnel off to the drain since $\mu_{N-1} < \mu_D$. Fig 1.5(b) shows the case of Coulomb blockade, where there are no states available that can allow an electron to tunnel from source to dot to drain (ie $\mu_S > \mu_D > \mu_N$). Electrons will fill any states that are below the source or drain chemical potentials trapping the electron(s) on the dot. At larger source-drain bias (V_{SD}), the situation in Fig 1.5(a) changes as eventually multiple states on the dot will be available for transport and blockade will not occur. This happens once eV_{SD} exceeds the charging energy of the dot. The gate voltage can be used to move the chemical potential of the dot relative to the level of the source and drain. Therefore, transport through the QDs is allowed only for specific combinations of gate voltage and source-drain bias. Hence, the current through the SED will oscillate with gate voltage. This effect can be seen in the bias spectroscopy for QD shown in Fig 1.5(c). Here the no current regions (white) correspond to regions of Coulomb blockade and higher current regions (blue) correspond to electron transport through the dot. The shape of the blockade regions

are typically referred to as the Coulomb Diamonds, where the height of the diamonds on the y-axis (V_{SD}) gives the charging energy of the dot. Fig 1.5(d) shows a line cut through (c) showing the effect of sweeping only the gate voltage at low V_{SD}. Here, we can identify the signature of SED transport, an oscillating current as a function of gate voltage with a single period.

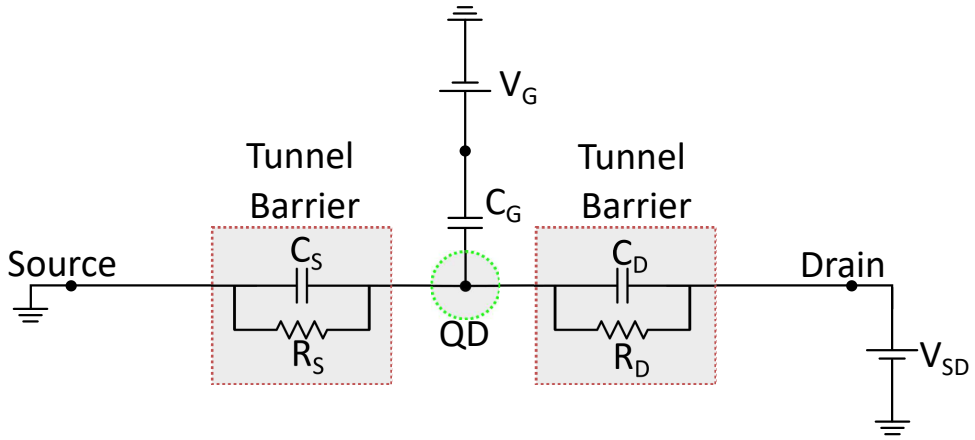

Figure 1.6: Circuit diagram for a single electron device. Tunnel barriers are represented as a parallel combination of a capacitor and resistor.

The period of the current oscillations and the slopes of the diamonds in Fig. 1.5 are determined by the gate (C_G), source (C_S), and drain (C_D) capacitances to the QD. A simplified circuit diagram for an SED is shown in Fig. 1.6. In this model, the tunnel barriers are treated as a parallel combination of a capacitor and resistor between the QD and either the source or drain. Meanwhile, the gate control of QD chemical potentials is purely capacitive. This means that the current oscillations as a function of gate voltage in Fig. 1.5 are inversely proportional to this gate-to-dot capacitance ($\Delta V = \frac{e}{C_G}$). It is important to note in an ideal QD when sweeping this gate voltage, there should be oscillations with only a single period. Any behavior

in a 1d gate voltage sweep or in the Coulomb diamond maps that shows something different from this suggests that the transport is due to more than one QD in device.

1.3 Non-idealities in silicon quantum dots

The discussion surrounding QDs so far has assumed that the local environment is ideal. In reality, this is not the case. The tremendous advantage of electrostatically defined QDs is that aspects such as the tunnel rates through barriers, the occupancy of electrons on the dot, and depth of the electrostatic potential well can all be controlled through gate voltages. This is high degree of tunability is one of the major reasons electrostatically defined QDs are considered good candidates for the potential applications previously mentioned. While electrostatically controlled tunnel barriers form the basis for QDs, their nature creates additional complexity through the fact that the tunnel barrier properties are highly sensitive to non-idealities in the local environment. Fig. 1.7 shows Coulomb Diamond maps for two MOS QDs fabricated on silicon-on-insulator (SOI) substrates by Binhui Hu at NIST. The data here are taken by sweeping the upper gate voltage (V_{UG} as shown in the inset of Fig. 1.7) and keeping the barrier gate voltages (V_{BG}) fixed. Importantly, both of these devices are nominally identical. They were fabricated on the same 1 cm^2 chip and the measurements were taken during the same cryostat cooldown. Under these conditions, we would expect that transport measurements through the two devices would also be identical. It is easy to see in Fig. 1.7 that this is not the case. Fig. 1.7(a) shows nearly ideal behavior for a single QD: many

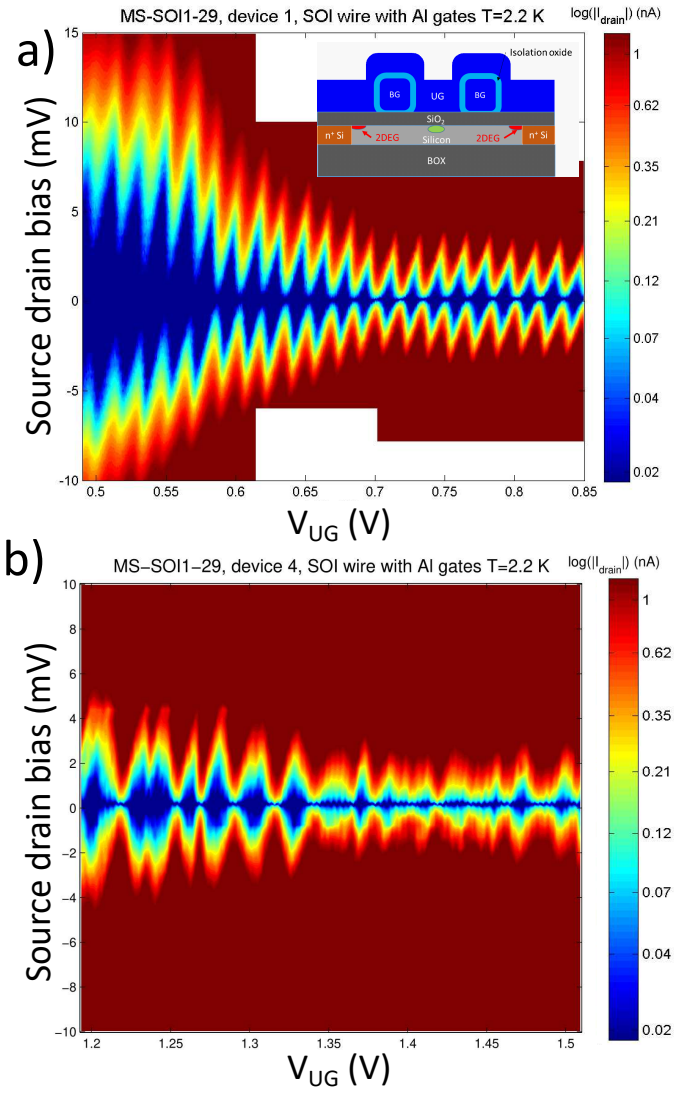

Figure 1.7: Bias spectroscopy of silicon MOS silicon-on-insulator (SOI) SEDs: (a) device that shows nearly ideal QD behavior with only a single gate period. Inset: schematic diagram of the SED. (b) Device that shows the unexpected behavior of multiple overlapping sets of Coulomb diamonds. This suggests that there are several unintentional QDs in the device. The devices in (a) and (b) are nominally identical, fabricated on the same chip, and measured in the same cryostat cooldown. Devices fabricated by Binhui Hu at NIST. Measurements performed by Ryan Stein.

diamonds that occur at a single period with V_{UG}. The opposite behavior is shown in the device in Fig. 1.7(b), where we see multiple sets of coulomb diamonds with different periods. This suggests that in this device there is not a single QD coupled to the gate. Instead, the fact that we see multiple sets of oscillations with different periods means that there are multiple QDs participating in transport. Since the devices in Fig. 1.7 are identical by design, we can only conclude that inhomogeneity in the local environment has led to the formation of QDs not consistent with the design and the intended electrostatics of the device.

Many groups use electrostatic simulations to guide fabrication and device design, but in general these simulations generally do not consider such non-idealities. These non-idealities typically show their impact in the form of QDs appearing in regions of the device where they should not be based on the gate layout, usually referred to as unintentional quantum dots (UQD). For both applications in quantum computing and the quantum current standard, we will require large numbers of well-defined and reproducible QDs. For instance in quantum computing, the best case scenario where one qubit is formed by a single QD we will still need millions of qubits to be a functional quantum computer [58]. Likewise for a quantum current standard to produce a practical level of current (1 μA) [52] it will be necessary to run many SED charge pumps in parallel. This again would require many well-defined and reproducible QDs. In both instances, the presence of UQDs hampers any applications that rely on reproducibility.

Anything that can modulate the potential in a silicon device on the nanometer length scales relevant for QDs can ultimately be a source of UQDs. In MOS devices,

this includes charge defects in the oxide [8, 59], impurities [60, 61], and inhomogeneous strain [62, 63]. In this work, we focus on characterizing charge defects and inhomogeneous strain.

1.3.1 Charge defects

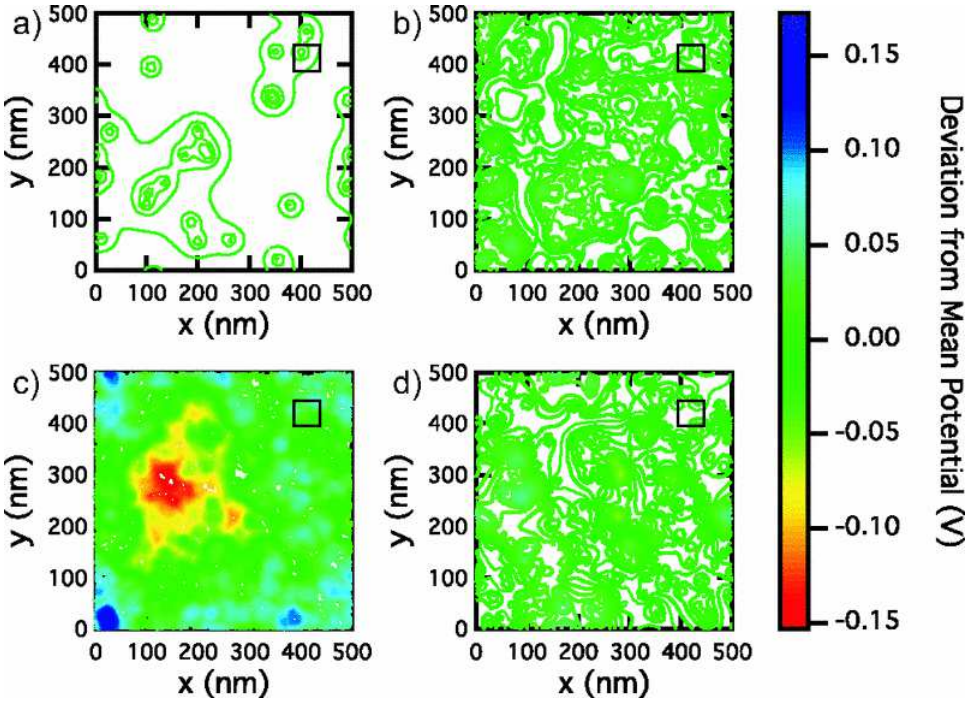

Figure 1.8: Contour plot from simulations of the potential modulations in the silicon caused by different random densities of charge defects in the oxide of a MOS device. Each plot represents the mean modulation from 20 different random spatial distributions of charge defects for a given density of (a) $+10^{10}cm^{-2}$, (b) $+10^{11}cm^{-2}$, (c) $+10^{12}cm^{-2}$, (d) $+5 \times 10^{10}cm^{-2}$ and $-5 \times 10^{10}cm^{-2}$ for net charge value of zero. Reprinted with permission from E. P. Nordberg et al. Enhancement-mode double-top-gated metal-oxide-semiconductor nanostructures with tunable lateral geometry, Phys. Rev. B 80, 115331 – Published 29 September 2009. Copyright 2009 by the American Physical Society.

Due the fact that silicon MOS structures are ubiquitous in industrial transistor fabrication, the presence and effects of charge defects in silicon dioxide have been

thoroughly studied. In QDs, these charge defects become important because they can induce potential modulations in the silicon that can lead to the formation of UQDs, noise, and be QDs themselves at the interface. In Ref. [59], Nordberg et al. calculated the potential modulations associated with different 2d charge distributions in the oxide. The contour plots from these simulations are shown in Fig. 1.8. Since the spatial distribution of charges in the oxide is random, the contour plots show the average modulation from 20 different random distributions for the same areal densities. Here, we can see that as the density increases from $+10^{10}cm^{-2}$ in (a), to $+10^{11}cm^{-2}$ in (b), and to $+10^{12}cm^{-2}$ in (c) that the average potential modulation increases significantly in the contour plots suggesting an increased probability of forming UQDs. Importantly, we note that typical densities found in MOS SEDs are in the $+10^{11}cm^{-2}$ range so that defect induced UQDs are likely to be common. The impact of charge defects in the oxide have also been demonstrated in the operation of single electron pumps. In Ref [33], Yamahata et al. showed pumping through a silicon MOS device that was dominated by an interface trap, or charge defect associated with a dangling bond at the $Si - SiO_2$ interface. In this work, they used the interface trap to their advantage, but as a practical matter, this work showed how a single defect could profoundly change the characteristics of an SED. We will discuss the nature of the oxide charge defects and methods for characterizing them in more detail in Chapter 3.

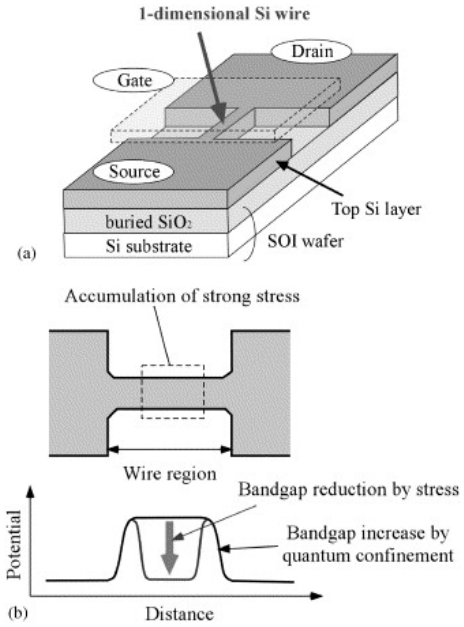

Figure 1.9: Schematic of the Pattern Dependent OXidation(PADOX) formation of inhomogeneous strain induced QDs. (a) 3d view of the silicon-on-insulator (SOI) nanowire device. (b) Top view of the nanowire pattern showing how the fan out of the nanowire region leads to high stress after oxidation. (c) Conduction band profile (E_c) along the nanowire showing a modulation profile similar that used to form QDs electrostatically. Reprinted from Yasuo Takahashi, Yukinori Ono, Akira Fujiwara, Hiroshi Inokawa, Development of silicon single-electron devices, Physica E: Low-dimensional Systems and Nanostructures,Volume 19, Issues 1–2,2003,Pages 95-101,ISSN 1386-9477, https://doi.org/10.1016/S1386-9477(03)00314-X, with permission from Elsevier.

1.3.2 Inhomogenous strain

In contrast to the expected random nature of UQDs due to charge defects, what is sometimes seen in SEDs is UQDs appearing the same locations in different devices. These reproducible UQDs suggest a source that is dictated by the device design and fabrication. A source that is consistent with the production of UQDs in this manner could be strain induced from dissimilar materials used in the device fabrication [62, 63, 64]. In Ref. [64], Takahashi et al. demonstrated the formation of

QDs in silicon nanowires devices where only a single gate was present in the device. Schematics of the silicon-on-insulator nanowire devices are shown in Fig. 1.9(a) and (b). Based purely on the electrostatics discussed earlier, these devices should have been unable to create the potential profile shown in Fig. 1.4 in order to form a QD because they lack the necessary barrier gates. Takahashi et al. found that, based on strain induced from the thermal oxidation of the silicon nanowire, tunnel barriers would form at the edges of the nanowire region in Fig. 1.9(b). Here, strain from the thermal oxidation lowers the conduction band energy near the center of the nanowire but strain changes near the edges of the nanowire as the device pattern widens as shown in Fig. 1.9(c). Since this oxidation induced strain, profile is dependent on the patterning of the silicon nanowire. They referred to this effect as Pattern Dependent OXidation or PADOX.

The strain induced QDs from PADOX are not the focus of this work. Here, we are focused on studying inhomogeneous strain induced from the gate used in QD devices. Gate-induced strain as a potential source of UQDs was suggested by Ted Thorbeck and Neil M. Zimmerman [62]. In that work, they used COMSOL to simulate the strain induced due to the coefficient of thermal expansion mismatch between the common materials used in SEDs. Fig. 1.10 shows a simulation of the overlapping aluminum gate structure similar to the devices used in Ref. [65]. In Fig. 1.10 (d), we can see the resulting modulation of the silicon conduction band (E_c) due to the strain induced in the device from cooling from 293 K to 1 K. The strain here is due solely the coefficient of thermal expansion (α) mismatch of materials in the device, where the silicon and silicon dioxide have relatively small α

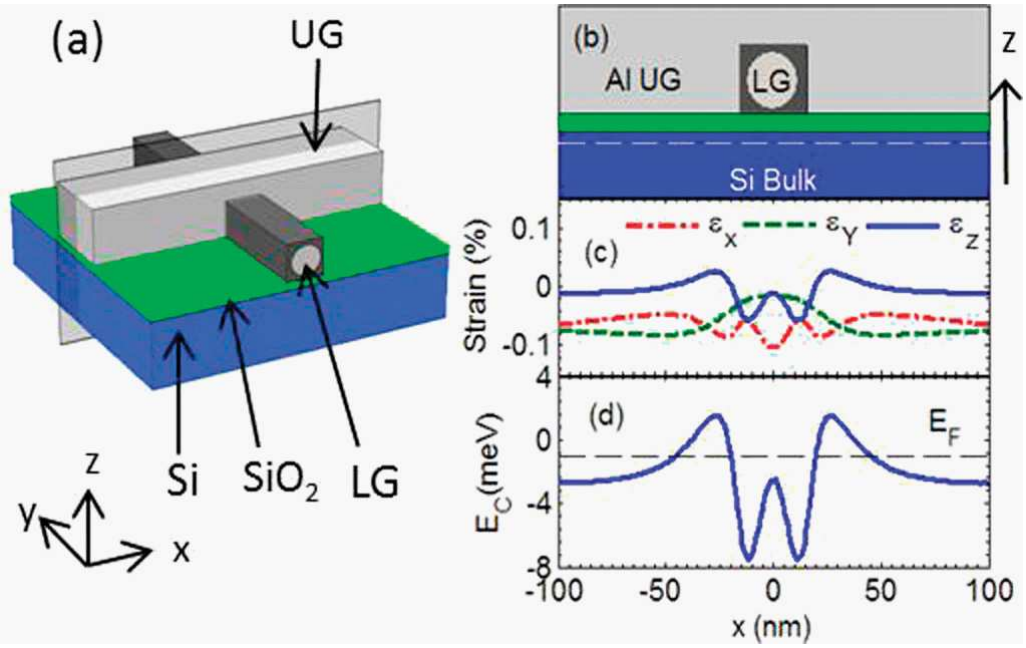

Figure 1.10: (a)3d view of the model used to simulate gate-induced strain in COMSOL. Here, the lower gate (LG) and upper gate (UG) are aluminum(grey) and are separated from each other with 3 nm of aluminum oxide (dark grey). The silicon (blue) and 10 nm of silicon oxide (green) typical combination seen in MOS QDs on bulk silicon wafers. (b) Schematic of the cut-plane through the gate overlap region in (a). (c) Strains calculated in the silicon (dashed line in (b)) due the coefficient of thermal expansion in the material in cooling the device from 293 K down to 1 K. (d)Modulation of the conduction band (E_c) due to the strains in (c) showing the formation of a QD under the lower gate. Reprinted from Ted Thorbeck; Neil M. Zimmerman; AIP Advances 5, 087107 (2015)DOI: 10.1063/1.4928320, with permission from AIP Publishing.

and the aluminum gates have a significantly larger α. Therefore, during cooling the aluminum gates want to contract significantly more than any of the other materials in the device. This leads to a buildup of strain locally at the edge of the gates. Since strain can modulate the silicon band structure, the resulting E_c modulation in Fig. 1.10 (d) follows the lithographic pattern of the lower gate (labeled LG) and leads to the formation of an UQD directly underneath LG. This overlapping gate structure

is ubiquitous in silicon MOS SEDs used in applications for quantum computing and current standards, understanding and controlling the formation of these strain-induced UQDs is extremely important for improving device performance. We will review how strain affects the silicon band structure in more detail in Chapter 2.

1.3.3 Materials considerations

The discussion of gate-induced strain and defects in the preceding section leads to an important aspect of the fabrication and design of silicon MOS SEDs: the choice of gate material. In the ideal case, the choice of gate material does not affect the operation or reproducibility of a SED outside of the value of voltage applied to achieve the desired effects. In reality, the choice is extremely important. Due to the desire for smaller QD sizes ($< 50nm$) and larger E_{ch}, most SEDs need to be fabricated with electron beam (e-beam) lithography on the university production scale or ultraviolet immersion lithography [66] on the industrial production scale. These techniques put constraints on lithography and materials used. For e-beam lithography, the most common process outside of industry is positive tone lithography and liftoff using electron beam or thermally evaporated metals. Although significant work has also been done with negative tone lithography and dry etching of chemical vapor deposited gate materials, such as doped poly-silicon [67, 68, 69]. Additionally, industrial MOS SEDs exclusively use negative tone processes for fabrication. The downside to this negative tone process is the damage done to the silicon dioxide during the dry etch process [70, 71]. As mentioned previously, the

most commonly used silicon MOS SED gate material is aluminum [53, 65]. One of the primary reasons for this is that the effectiveness of the native aluminum oxide to act as isolation oxide in the multilayered gate structures ubiquitous in MOS SEDs. From the above discussion of gate-induced strain, we can see that Al has a significant downside in this area because of the large α mismatch with Si substrate. The simplest way to avoid this large α mismatch is to use a more thermally matched gate material, such as doped poly-silicon. This switch comes with the trade off the possibility for increased charge defects in the silicon dioxide from the dry etching process.

More recent work in MOS SEDs has moved from devices made with Al gates and native aluminum oxide isolation to devices with Ti/Pd gates and atomic layer deposition (ALD) aluminum oxide. The real driving factor behind this change is improving the e-beam lithography, where typical grain sizes in Pd films are significantly smaller than that obtained from Al films. We note that a lack of measurements of the physical properties of these gate materials, such as the coefficient of thermal expansion, intrinsic film stress, and charge defects induced in the oxide via the gate processing, is a significant obstacle towards making optimal decisions about MOS SED fabrication processes and design.

1.4 Outline of thesis

The goal of this thesis is to elucidate the effects of defects and gate-induced strain on silicon metal-oxide-semiconductor (MOS) quantum dot (QD) devices. This

will provide guidance to the community on the commonly used MOS QD design and fabrication choices with respect to different levels of defects and strain realized in the final device. We perform our study not directly in the MOS QDs, but via the use of test devices that allow us to see the individual contributions more clearly.

In Chapter 2, we show that strain alters the band structure of silicon shifting the conduction band (E_c) edge. We will use COMSOL to simulate strain effects in silicon MOS QDs material stacks. In later chapters, we will use these simulations to develop a device design to electrically measure strain from the gates at cryogenic temperatures. We also discuss experimental measurements of thin film stress on the wafer scale. These techniques will also be used in later chapters to probe changes in intrinsic film stress (σ) and the coefficient of thermal expansion (α) as a function of processing conditions. The goal of this chapter is to provide a review of elastic mechanics relevant for silicon MOS devices and setup the concepts needed for understanding the results presented later.

In Chapter 3, we will cover the ideal properties of a MOS device and will show that deviations from the ideal model are primarily due to several types of oxide charge defects. The different defect densities are inferred from capacitance as function of voltage (CV) measurements on MOS capacitors (MOSCAPs). These charge defect measurements are crucial to understanding the impact that common MOS QD fabrication process can have on the electrical properties of the device. Later chapters will use this foundation to study the changes in defect densities for different gate materials and anneals common to MOS QD fabrication.

In Chapter 4, we will use the groundwork laid out in Chapter 2 to develop

a new device to electrically measure strain under a similar fabrication process and operating conditions to that seen in QD devices. The goal of this chapter is to present a comparison between simulations and measurements of the effect of strain on the tunnel barrier height of MOS tunnel junction devices. Here, we infer relative differences in strain between tunnel junction devices with Al and Ti gates from relative differences in the measured barrier heights. We find that our tunnel junction measurement of the strain difference agrees with simulations provided we use our experimentally measured values of α.

In Chapter 5, we address the interplay of defects and strain in MOS devices by presenting a comparison of oxide defect densities (fixed charge, Q_f, and interface trap density, D_{it}), σ, and α for Ti/Pd, Ti/Pt, and Al as a function of forming gas anneal temperature and hydrogen concentration. We vary the anneal temperature from 200 °C to 425 °C, using both 5 % and 10 % mixtures in 30 minute anneals. We show that Ti/Pd and Ti/Pt have larger D_{it} than Al when optimally annealed and that the magnitude of Q_f is larger for Ti/Pd and Ti/Pt than for Al, with Al showing a net negative charge while Ti/Pd and Ti/Pt display net positive charge. Additionally, we show that both α and σ increase with increasing anneal temperature. Moreover, these results show that due primarily to intrinsic strain, Pd-gated devices have larger strain-induced modulation of the conduction band than their Al-gated counterparts, directly contradicting expectations based on the bulk α alone [72]. Finally, and most importantly, we find no anneal which simultaneously minimizes defects and the effects of strain in any of the materials studied. Thus, a tension arises in designing fabrication processes for MOS QDs

where one must choose between setting the anneal such that defects are minimized or the strain-induced modulation of the conduction band is minimized.

In Chapter 6, leveraging the optimized forming gas anneal, we develop and fabricate a new set of MOS tunnel junction devices. In these new TJ devices, we fabricate two different metals, Al and Ti/Pd, on either side of the barrier making the TJ asymmetric. In this way, we project our measurement of strain onto the barrier asymmetry from the BRD model rather than the barrier height as in Chapter 4. The advantage of this asymmetric TJ design over the previously used symmetric TJs is that we can potentially extract the strain difference from a single tunnel junction rather than comparing relative differences between different tunnel junctions measured on different chips and cooldowns.

Finally, in Chapter 7, we conclude by discussing the outlook of future work on characterizing defects and strain in MOS QDs. On the subject of measuring defect densities, we propose a low temperature measurement of the interface trap density using MOS transistors with the goal of probing closer to the band edges, which is potentially more relevant for MOS QDs performance. Additionally, we will propose alternative device designs and measurements for the goal of electrically measuring strain.

Chapter 2: Stress and strain in the silicon MOS system

The goal of this chapter is to provide a review of elastic mechanics relevant for silicon MOS devices and setup the concepts needed for understanding the results presented later. As discussed in Chapter 1, strain induced from the gates has been suggested as a source of unintentional QDs [73, 74]. The strain can come from the intrinsic film stress (σ_0), built in during the deposition process and subsequent fabrication process, or from the coefficient of thermal expansion (α_{film}) mismatch between the gate material and silicon substrate (α_{si}) , built in as the device is cooled to cryogenic temperatures. In this chapter, we show that strain alters the band structure of silicon shifting the conduction band (E_c) edge. We will use COMSOL to simulate strain effects in silicon MOS QDs material stacks. In later chapters, we will use these simulations to develop a device design to electrically measure strain from the gates at cryogenic temperatures. Finally, in this chapter, we will discuss experimental measurements of thin film stress on the wafer scale. These techniques will also be used in later chapters to probe changes in σ_0 and α_{film} as a function of processing conditions.

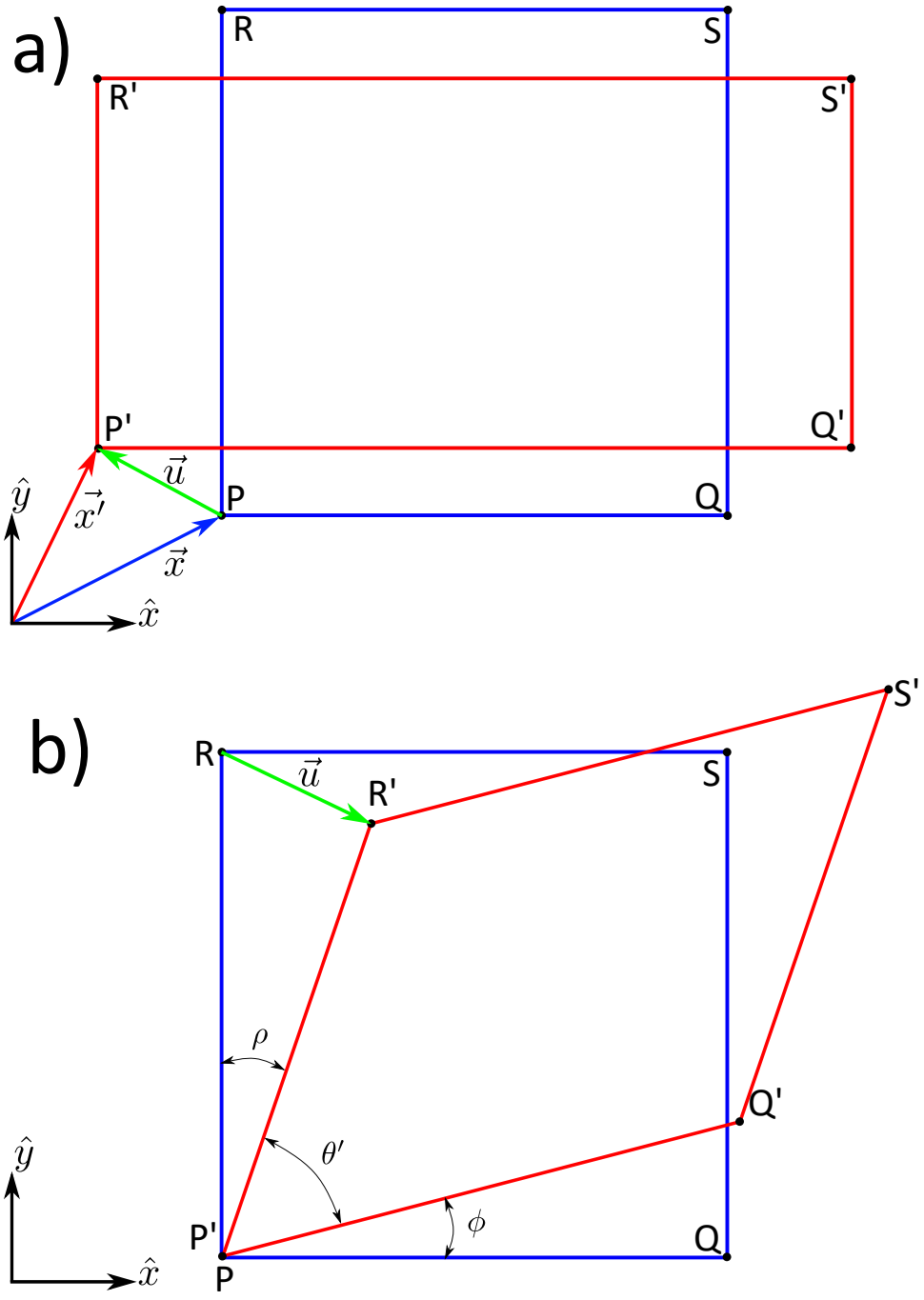

Figure 2.1: Schematic showing the deformation of a 2d rectangle where the blue rectangle is the relaxed state and the red rectangle is deformed state. (a) Example of normal strain, where tension is applied along the x direction changing the length of the sides of the rectangle. (b) Example of shear strain, where the deformation changes the angle between the sides of the rectangle.

2.1 Overview of general stress and strain

2.1.1 Strain and displacement

Strain describes the deformation of an object from its nominal or relaxed state, shown in the diagram in Fig. 2.1 (a), where a rectangle has been deformed in the x and y directions. In order to mathematically describe strain, we will use the concept of the displacement vector, $\vec{u}(x, y)$. Following the derviations in Ref. [75], consider a simple 2-dimensional deformation shown in Fig. 2.1(a), where the blue rectangle is in the original or relaxed state and the red rectangle is the deformed state. Here, tension has been applied along the x-direction increasing the length of the sides along parallel to the x-axis. The strains associated with this type of deformation are referred to as normal strains. The relaxed rectangle has sides of length dx and dy and reference points of P, Q, R, and S. In the deformed rectangle, the reference points (P',Q',R',and S') are related to their counterparts in the relaxed rectangle by the displacement vectors. In Fig. 2.1(a), we define the displacement vector: $\vec{u} = \vec{x'} - \vec{x}$, where $\vec{x}$ and $\vec{x'}$ are vectors describing the positions of P and P' from the origin respectively. We note that this choice of an origin is arbitrary as the displacement vector is only affected by changes in the lengths of sides of the rectangle and not uniform shifts. The change from P to P' is a shift of $u_x(x, y)$ in the x-direction and a shift of $u_y(x, y)$. If we consider only small deformations, then

the strain in the x-direction, ϵ_{xx}, is related to the displacement vectors by [75]:

$$\epsilon_{xx} = \frac{\delta u_x}{\delta x} \tag{2.1}$$

Similarly for the y-direction normal strain is:

$$\epsilon_{yy} = \frac{\delta u_y}{\delta y} \tag{2.2}$$

In addition to the normal strains derived above, Fig. 2.1(b) also shows a rotational deformation. We use the angles ϕ and ρ to describe the rotation of the sides of the rectangle PQ and PR respectively. Here, it is important to note that PQ and PR are orthogonal in the relaxed state, but are not orthogonal in deformed state. Thus, this is not just a pure rotation of the rectangle. The strain associated with this change in angle between the sides of the two rectangles is referred to as shear strain [75] denoted γ_{ij}, where i≠j and $i, j = x, y,$ or z. Using the diagram in Fig. 2.1, γ_{xy} is given by:

$$\gamma_{xy} = \theta - \theta' = \frac{\pi}{2} - \theta' = \phi + \rho \tag{2.3}$$

If we again only consider small deformations, then the angles ϕ and ρ are small and we can use the small angle approximation ($\phi \approx tan(\phi)$). This gives a definition of γ_{xy} in terms of the displacement vectors [75]:

$$\gamma_{xy} = \frac{\delta u_x}{\delta y} + \frac{\delta u_y}{\delta x} \tag{2.4}$$

We can write the above equations for both normal and shear strains in a generalized tensor form:

$$\boldsymbol{\epsilon_{ij}} = \frac{1}{2}[\frac{\delta u_{ij}}{\delta i} + \frac{\delta u_{ij}}{\delta j}] \tag{2.5}$$

where in the rest of this work we will denote tensor quantities in bold font. It is important to note that the shear strain defined by Eq. 2.5 differs from those defined by Eq. 2.4 by a factor of 2. The relation in Eq. 2.5 is usually called the tensor or true strain while Eq. 2.4 is called the engineering strain [75]. We will use the tensor strain definitions for the rest of this work as it more convenient for the finite element modeling (FEM) done later in the chapter. Eq. 2.4 describes strain as a second-order tensor that is symmetric such that the off-diagonal strains are equal ($\epsilon_{xy} = \epsilon_{yx}$). Consider a simple example of a 1D free-standing block of length L_0, that undergoes a change in length of $L - L_0$. The strain in this case would be $\epsilon = \frac{L-L_0}{L_0}$. If we set $L_0 = 1.0\ \mu m$ and $L = 0.001\ \mu m$, then the strain would 0.001. Since strain is related to the fractional change in size of an object, it is typically reported as a percentage, or a 0.1% strain in this case. As mentioned before the derivations above only apply to the case of small deformations, this is known as the linear elastic model [1]. Another important component of linear elasticity is that all deformations are reversible; such that once we remove the source of the deformation, the object will return to its original state.

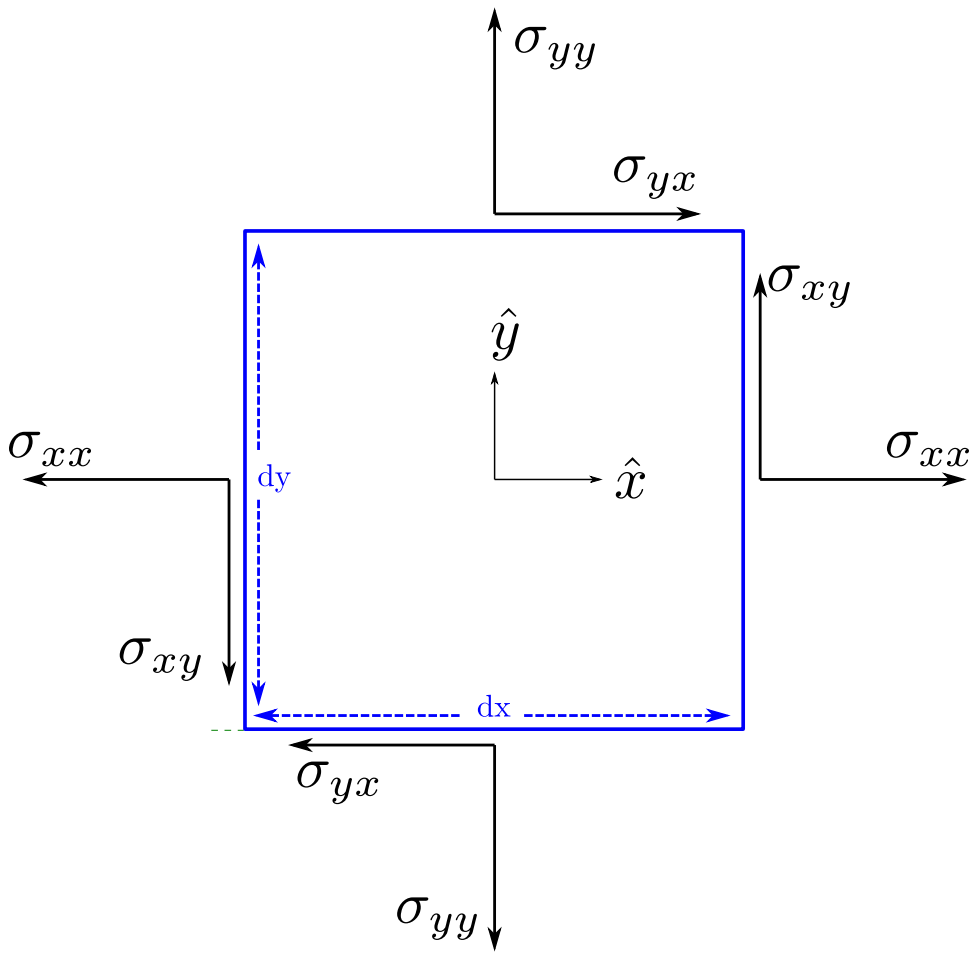

Figure 2.2: Schematic showing the stress components for a 2d rectangle.

2.1.2 Stress

Stress describes the internal forces acting on a particle within a solid from the rest of particles that make up the solid. Stress at a particular position is considered as the forces $\vec{F}$ acting on an small area A, such that $\vec{\sigma} = \frac{\vec{F}}{A}$. If we consider an infinitesimally small area, we can decompose the stress into three components: a perpendicular or normal component and in-plane or shear components. We show this decomposition in Fig. 2.2 for stress on a rectangle, where similar to the naming

conventions used for strain the normal components are labeled σ_{ii} and the shear components are label σ_{ij} with $i \neq j$. The normal stress components are stress perpendicular to the surface, while the shear stress are stress parallel to the surface. The total stress acting on any portion of the cube can be written in tensor form as [75]:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \tag{2.6}$$

where the off diagonal terms are symmetric ($\sigma_{xy} = \sigma_{yx}$, $\sigma_{yz} = \sigma_{zy}$, and $\sigma_{xz} = \sigma_{zx}$). Here, we treat tensile stress or elongation as a positive value and compressive stress as negative. For mechanics problems related to the MOS system, we always consider the case of equilibrium stress where the sum of all the forces acting on four sides of the rectangle in Fig. 2.2 are zero. Otherwise, the rectangle would be accelerating. In mathematical form, we express this as:

$$\begin{aligned} \frac{\delta \sigma_{xx}}{\delta x} + \frac{\delta \sigma_{xy}}{\delta y} = 0 \\ \frac{\delta \sigma_{yx}}{\delta x} + \frac{\delta \sigma_{yy}}{\delta y} = 0 \end{aligned} \tag{2.7}$$

These differential equations of equilibrium are one part of the equations used in solving elastic mechanics problems. We can extend these differential equations in 3-dimensions easily and more generally:

$$\nabla \cdot \boldsymbol{\sigma} = 0 \tag{2.8}$$

Since the tensors for strain and stress typically only have six distinct elements due to symmetry, they are more compactly written as six element vectors:

$$\vec{\epsilon} = \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{xy} \\ \epsilon_{yz} \\ \epsilon_{xz} \end{pmatrix} \tag{2.9}$$

$$\vec{\sigma} = \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{xy} \\ \sigma_{yz} \\ \sigma_{xz} \end{pmatrix} \tag{2.10}$$

We will use the above forms for the remainder of this work. The second equation we need solve to is Hooke's Law, which gives the connection between stress and strain.

2.1.3 Hooke's Law

Hooke's law gives the relationship between the amount an object deforms and the applied stress under the conditions discussed above for linear elasticity. Here

the elastic strain ($\vec{\epsilon_{el}}$) is related linearly to the total stress ($\vec{\sigma}$):

$$\vec{\epsilon_{el}} = \boldsymbol{S}\vec{\sigma} \tag{2.11}$$

where $\boldsymbol{S}$ is the stiffness matrix in units of $\frac{1}{Pressure}$. Alternately, we can write $\vec{\sigma} = \boldsymbol{C}\vec{\epsilon_{el}}$ where $\boldsymbol{C} = \boldsymbol{S}^{-1}$ is the compliance matrix. Since $\vec{\sigma}$ and $\vec{\epsilon_{el}}$ are both contain 6 distinct elements, then $\boldsymbol{C}$ and $\boldsymbol{S}$ will be 6×6 tensors:

$$\boldsymbol{S} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{pmatrix} \tag{2.12}$$

Where we can write a similar tensor for $\boldsymbol{C}$. For cubic crystals, the symmetry reduces some components to zero:

$$\boldsymbol{S} = \begin{pmatrix} s_{11} & s_{12} & s_{31} & 0 & 0 & 0 \\ s_{12} & s_{22} & s_{23} & 0 & 0 & 0 \\ s_{31} & s_{23} & s_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66} \end{pmatrix} \tag{2.13}$$

For the simplest case of isotropic materials with a cubic crystal structure [1], we have $s_{11} = s_{22} = s_{33}$, $s_{12} = s_{23} = s_{31}$, and $s_{44} = s_{55} = s_{66}$. Thus, $\boldsymbol{S}$ and $\boldsymbol{C}$ can be written in terms of two different material constants: Young's Modulus, E, and Poisson's ratio, ν. Mathematically, E is the slope of the stress-strain curve under uniaxial deformation [1], but the physical representation is more accurately the stiffness of a solid. The larger the value of E, the stiffer the material. Since strain is dimensionless, E will have the same units as stress with most materials relevant to this work on the order of GPa. ν describes the ratio of lateral to longitudinal strain under uniaxial stress, e.g. stretching a material causes a lateral contraction. ν is dimensionless with $0 < \nu < 0.5$ for common materials and for most metals $\nu \approx 0.3$. It is possible in some special materials to have $\nu < 0$ [1], but this is not possible for the materials studied in this work.

$$\boldsymbol{S} = \frac{1}{E}\begin{pmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 1+\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+\nu \end{pmatrix} \tag{2.14}$$

Since the constants s_{44}, s_{55}, and s_{66} are related to the shear strain and stress component, they are commonly replaced with a shear modulus constant G where $s_{44} = s_{55} = s_{66} = {}^{1}/_{G}$. Similarly to E, G describes the ability of material to resist changes due to shear deformations and typically $G < E$. In the case of an isotropic

Material	c_{11}	c_{44}	c_{12}	s_{11}	s_{44}	s_{12}	E (GPa)	ν	G (GPa)
Al	107.3	28.3	60.9	15.82	35.36	-5.73	63.20	0.36	28.30
Pd	227.1	71.7	176.0	13.63	13.94	-5.95	73.41	0.44	71.70
Pt	346.7	76.5	250.7	7.34	13.07	-3.08	136.29	0.42	76.50

Table 2.1: List of room temperature bulk elastic constants for FCC metals commonly used as MOS gate materials. c_{ii} are in units of GPa and s_{ii} are in units of 10^{-12} Pa^{-1}. From Ref [1]

material, $G = {}^{E}/_{1+\nu}$. In general, most materials of interest are not strictly isotropic and as the crystal symmetry moves away from the simple cubic symmetry, we will require more distinct elastic constants to characterize a material. Face-centered cubic (FCC) metals [1], such as Al, Pt, and Pd which are common MOS gate materials, are close to isotropic with the exception of the shear terms in $\boldsymbol{S}$ (s_{44},s_{55}, and s_{66} are not equal to ${}^{1+\nu}/_{E}$ as in the perfectly isotropic case). However, for these metals, we still have $s_{44} = s_{55} = s_{66}$. Table 2.1 shows the elastic constants for some FCC metals relevant for later chapters in this work.

For other common crystal structures, such as hexagonal close-packed (HCP) and diamond cubic, the change in symmetry forces more complexity in $\boldsymbol{S}$ and $\boldsymbol{C}$. For these materials, we will use the case of a general orthotropic material, where we need only assume that the material has three mutually perpendicular symmetry planes leading to a maximum of nine independent constants. In this case, we can

Material	E_x	E_y	E_z	ν_{yx}	ν_{zx}	ν_{zy}	G_{yz}	G_{zx}	G_{xy}	Ref
Ti	104.37	104.37	143.27	0.48	0.20	0.20	46.70	46.70	35.20	[1]
Si (100)	169	169	130	0.064	0.28	0.36	79.6	79.6	50.9	[76]

Table 2.2: List of room temperature elastic constants for orthotropic materials used in MOS QDs. E and G are in units of GPa. Si (100) refers to a silicon (100) wafer where the z direction is aligned along the (001) crystal axis.

Material	c_{11}	c_{22}	c_{33}	c_{12}	c_{13}	c_{23}	c_{44}	c_{55}	c_{66}	Ref
Ti	162.4	162.4	180.7	92	69	69	46.70	46.70	35.20	[1]
Si (100)	194.5	194.5	165.7	35.7	64.1	64.1	79.6	79.6	50.9	[76]

Table 2.3: List of room temperature compliance matrix for orthotropic materials used in MOS QDs in units of GPa.

rewrite Eq. 2.13 as:

$$\boldsymbol{S} = \begin{pmatrix} {}^{1}/_{E_x} & -{}^{\nu_{yx}}/_{E_y} & -{}^{\nu_{zx}}/_{E_z} & 0 & 0 & 0 \\ -{}^{\nu_{yx}}/_{E_x} & {}^{1}/_{E_y} & -{}^{\nu_{zy}}/_{E_z} & 0 & 0 & 0 \\ -{}^{\nu_{zx}}/_{E_x} & -{}^{\nu_{zy}}/_{E_y} & {}^{1}/_{E_z} & 0 & 0 & 0 \\ 0 & 0 & 0 & {}^{1}/_{G_{yz}} & 0 & 0 \\ 0 & 0 & 0 & 0 & {}^{1}/_{G_{zx}} & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^{1}/_{G_{xy}} \end{pmatrix} \tag{2.15}$$

For HCP metals, such as Ti which is commonly used as an adhesion layer for MOS gate materials, we can reduce this further because of rotational symmetry such that $s_{11} = s_{22}$, $s_{12} = s_{23}$, and $s_{44} = s_{55}$. Here we have $E_x = E_y$ and $G_{yz} = G_{zx}$ meaning a total of seven constants shown in Tables 2.2 and 2.3.

One of the more important set of elastic constants is for the silicon substrate used in MOS QDs. Silicon has a diamond cubic crystal structure, but is an orthotropic elastic material like Ti. Tables 2.2 and 2.3 show the elastic constants

relevant for silicon wafers used in this work, where Si(100) refers to the crystal direction perpendicular to the wafer surface. Additionally the constants in Table 2.3 are derived such that the in-plane directions are the $< 110 >$ crystal directions. The final remaining material in the MOS system that we have not yet covered is the SiO_2 gate oxide. Since SiO_2 is amorphous, we treat this material as a perfectly isotropic material with $E_{SiO_2} = 75GPa$ and $\nu_{SiO_2} = 0.17$ [77].

Typically, in this work, we are concerned with the case of a thin film deposited on a silicon substrate. Here, the thin film stress is biaxial in nature. Thus it is convenient to introduce another modulus term, the biaxial modulus, B. Similar to the other moduli discussed above, B describes the ability of a material to resist biaxial deformations. In the case of a isotropic material, $B = \frac{E}{1-\nu}$. A more general relationship that applies for orthotropic materials is written in terms of the compliance matrix,$\boldsymbol{C}$ [78]:

$$B = c_{11} + c_{12} - \frac{2c_{12}^2}{c_{11}} \tag{2.16}$$

We have discussed above the elastic strain caused by the applied stress $\vec{\sigma}$ however in practice the total strain, $\vec{\epsilon_{tot}}$, is a combination of a few other components. We will consider additional strain components: the thermal strain, $\vec{\epsilon_{th}}$, and intrinsic strain, $\vec{\epsilon_0}$. Thus, $\vec{\epsilon_{tot}} = \vec{\epsilon_{el}} + \vec{\epsilon_{th}} + \vec{\epsilon_0}$ and Eq. 2.11 becomes $\boldsymbol{S}\vec{\sigma} = \vec{\epsilon_{tot}} - \vec{\epsilon_{th}} - \vec{\epsilon_0} = \vec{\epsilon_{el}}$. It is important to note the reason for this distinction between $\vec{\epsilon_{tot}}$, $\vec{\epsilon_{th}}$, and $\vec{\epsilon_{el}}$. $\vec{\epsilon_{el}}$ describes the elastic strain caused by stress. $\vec{\epsilon_{th}}$ differs in the sense that thermal strains do not result from stress, but from the deformation with temperature. Consider a freestanding block of material that undergoes a temperature change of ΔT.

Material	α (10^{-6})	Ref
SiO_2	0.49	[77]
Si	2.6	[77]
Ti	8.6	[79]
Pt	8.8	[79]
Pd	11.8	[79]
Al	23.0	[79]

Table 2.4: List of bulk room temperature coefficients of thermal expansion for commonly used materials in MOS QD.

Here, the strain is $\alpha\Delta T$, where α is the coefficient of thermal expansion which is material dependent. In this case, the object has deformed but there is no stress as there is no force acting on the object. In practice we do not have free standing blocks of material, we have stacks of different materials with different α. In this way, each material is no longer allowed to deform as it would in the freestanding case and this leads to thermal stresses. For all of the materials we have discussed above, we can consider the thermal strain to be isotropic and the shear components are all zero because the thermal expansion is volumetric. $\vec{\epsilon_{th}}$ describes the deformation in an object when the temperature changes from T_i to T_f:

$$\vec{\epsilon_{th}} = \alpha(T_i - T_f) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.17}$$

The other additional strain component, $\vec{\epsilon_0}$, is similar in the sense that it is a strain

not tied to an external force. We will use this term to describe strains related to growth or deposition of different materials. For instance when thermally oxidizing silicon, the SiO_2 expands from the normal silicon lattice spacing to incorporate oxygen atoms. The amount of expansion for the SiO_2 film is limited by silicon leading to compressive stress in the film, typically on the order of -200 to -300 MPa for oxides grown in pure O_2 [80]. $\vec{\epsilon_0}$ is the strain due to this stress. $\vec{\epsilon_{el}}$ is ultimately the quantity of interest in proceeding sections, but $\vec{\epsilon_{th}}$ and $\vec{\epsilon_0}$ represent important physical phenomena that must be included. Table 2.4 shows the values of α in the bulk for the materials used in this work.

2.1.4 Simple analytical example

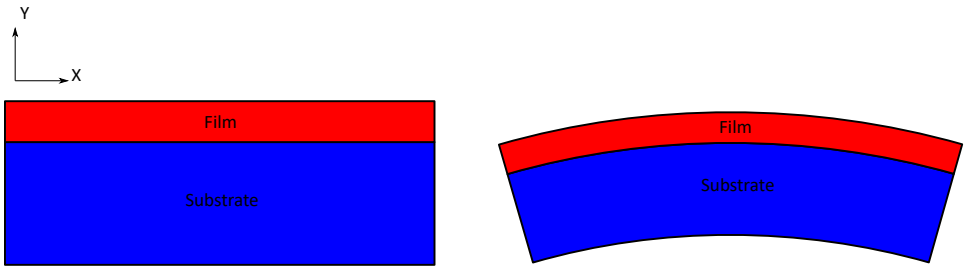

Figure 2.3: Schematic of the bimetallic strip. Left image shows the strip at the initial or relaxed temperature. The right image shows the strip at the final temperature where the buildup of stress has changed the curvature.

In order provide some exposure to typical elastic mechanics solutions; we will turn to the classic problem of the bimetallic strip. This system was originally solved in 1925 by Timoshenko [81]. The bimetallic strip is a simple structure composed of two materials adhered together with different elastic constants and α. Similar

to the situation discussed in the preceding section, stress builds up in materials as they undergo a temperature change of ΔT. Here, we will assume that our stack, a bi-material stack, is a silicon substrate of thickness t_{si} with SiO_2 thin film of thickness t_{ox}. We show this case in Fig. 2.4. For the purposes of this problem we will treat both materials as isotropic with elastic constants of E_{si} and E_{ox} and coefficient of thermal expansion of α_{si} and α_{ox}. Starting with the analytical solution from Timoshenko [81], we can write the curvature,κ, of the strip as:

$$\kappa = \epsilon_{th} \frac{6E_{si}E_{ox}t_{si}t_{ox}(t_{si}+t_{ox})}{E_{si}^2 t_{si}^4 + 4E_{si}E_{ox}t_{si}^3 t_{ox} + 6E_{si}E_{ox}t_{si}^2 t_{ox}^2 + E_{ox}t_{ox}^4} \tag{2.18}$$

Where $\epsilon_{th} = (\alpha_{si} - \alpha_{ox})\Delta T$ is the thermal strain . Eq. 2.18 assumes that the length of the strip, L, is much larger than the thickness of both materials. Additionally since this is a 2d problem, we will use what is known as the plane stress approximation which means that there are no shear strains and the out plane stress in Fig. 2.3 is zero ($\sigma_{zz} = 0$). The fact that $L >> t_{si}, t_{ox}$ also means that σ_{yy} is also 0. From Timoshenko, the in plane stress, σ_{xx}, as a function of position y in the strip is related to the curvature by:

$$\begin{aligned} \sigma_{xx,ox}(y) &= \kappa \left[E_{ox}(y - \frac{t_{ox}}{2}) - \frac{E_{si}t_{si}^3 + E_{ox}t_{ox}^3}{6t_{si}(t_{si}+t_{ox})} \right] \\ \sigma_{xx,si}(y) &= \kappa \left[E_{si}(y + \frac{t_{si}}{2}) + \frac{E_{si}t_{si}^3 + E_{ox}t_{ox}^3}{6t_{si}(t_{si}+t_{ox})} \right] \end{aligned} \tag{2.19}$$

Where $y = 0$ is the oxide-silicon interface with $y < 0$ is in the silicon and $y > 0$ is

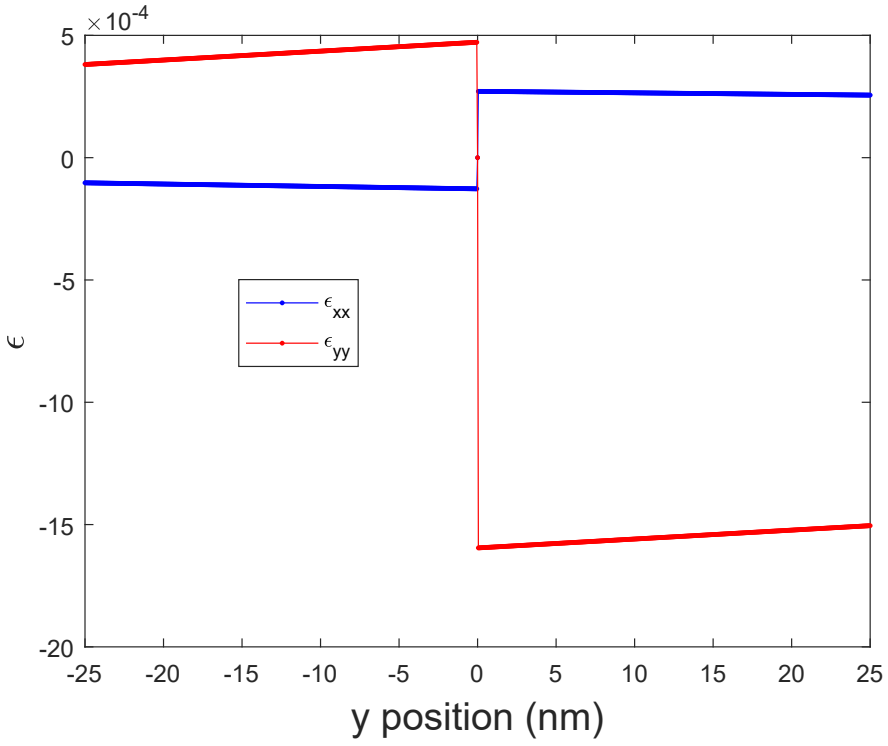

Figure 2.4: Plot of the strain in a SiO_2-Si bimetallic strip using Eq. 2.20. $y = 0$ is the oxide-silicon interface with $y < 0$ is in the silicon and $y > 0$ is the silicon dioxide. The silicon dioxide is 25 nm thick.

the silicon dioxide. In the plane stress approximation, we can write the strain as:

$$\begin{aligned} \epsilon_{xx} &= \frac{1}{E}(\sigma_{xx} - \nu\sigma_{yy}) = \frac{\sigma_{xx}}{E} \\ \epsilon_{yy} &= \frac{1}{E}(\sigma_{yy} - \nu\sigma_{xx}) = \frac{-\nu\sigma_{xx}}{E} \end{aligned} \tag{2.20}$$

Where Eq. 2.20 can used to calculate strain as a function of y by substituting the appropriate material constants for the oxide and silicon. Fig. 2.4 shows the stress strain calculated vs position for the strip cooled from 1000°C down to room temperature and the silicon dioxide is 25 nm thick. Here, we see that the strain is relatively small, about 0.01%, even though the temperature change is large. This is

because of the relatively small difference in α between the oxide and silicon.

2.2 Strain and the silicon band structure

Silicon is an indirect band gap semiconductor with anisotropy in both its electron and hole effective masses and differing symmetries in E_c and E_v. The conduction and valence bands will be behave differently under applied strain. Since later in this work we are only concerned with devices operating with electrons, we will cover only the E_c and strain relationship.

2.2.1 Conduction band

At the minimum of E_c in Silicon band structure; there are six equivalent points, or valleys, in the Brillouin zone. Additionally at this minimum, the energy dependence is roughly parabolic leading to the effective electron mass. The dispersion relations for each of the valleys can be written as [82]:

$$\begin{aligned} E_{c,\pm k_x}(k) &= \frac{\hbar^2}{2}\left(\frac{(k_x \mp k_0)^2}{m_l} + \frac{k_y^2}{m_t} + \frac{k_z^2}{m_t}\right) \\ E_{c,\pm k_y}(k) &= \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_t} + \frac{(k_y \mp k_0)^2}{m_l} + \frac{k_z^2}{m_t}\right) \\ E_{c,\pm k_z}(k) &= \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_t} + \frac{k_y^2}{m_t} + \frac{(k_z \mp k_0)^2}{m_l}\right) \end{aligned} \tag{2.21}$$

where the longitudinal effective mass is $m_l = 0.98m_e$, the transverse effective mass is $m_t = 0.19m_e$, and m_e is the free electron mass. $k_0 = 0.85\frac{2\pi}{a_{si}}$ is the wave vector offset of the conduction band minimum. In bulk silicon at k_0, Eq. 2.21 is six-fold

degenerate, but this degeneracy can be broken by confinement due to the presence of the SiO_2-Si interface and by strain and electric fields.

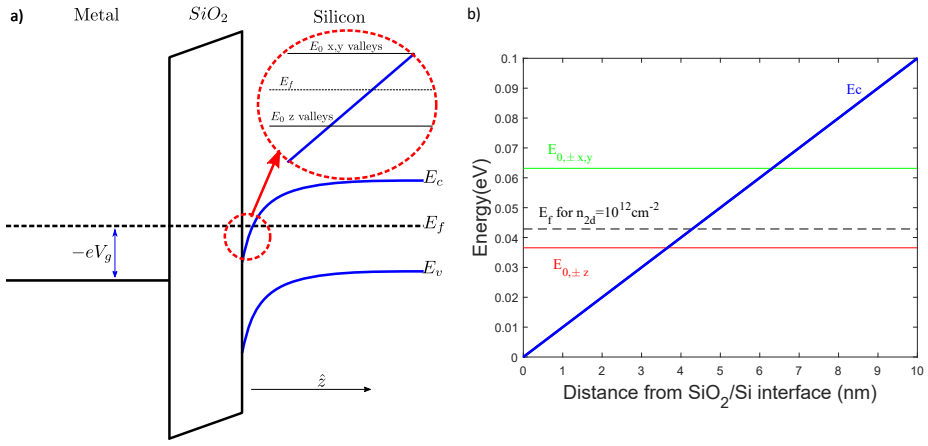

Figure 2.5: (a) Band diagram of the MOS system when the silicon is in inversion. The inset shows the valley structure due to confinement at the SiO_2-Si interface. (b) Plot of the energy levels using an electric field of $10^5 \frac{V}{cm}$. Here, the lowest subband is the z-valleys with the x and y valleys being much higher in energy.

The confinement at the SiO_2-Si interface causes a splitting of the 6-fold valley degeneracy due the effective mass anisotropy described above. Fig. 2.5(a) shows an energy band diagram when an inversion layer forms in a MOS device. Here, when we apply a positive voltage to the gate, the positive charges at the metal-SiO_2 interface pull electrons towards the SiO_2-Si interface. In this case, the bands are bent downward at the SiO_2-Si interface. The inset of Fig. 2.5(a) shows a zoomed in view of E_c near the interface. Here, we can approximate E_c by a triangular potential well for electrons confined near the interface. If we treat the SiO_2 as a infinitely high barrier and the potential in the silicon as $e\mathscr{E}z$, where $\mathscr{E}$ is the electric field and z is the distance from the interface. The eigenenergies for the triangular potential

well are given by [83]:

$$E_n = |a_n| \left(\frac{(e\mathscr{E})^2 \hbar^2}{2m_z} \right)^{1/3} \tag{2.22}$$

Where a_n are the zeros of the Airy function with $a_0 = -2.33$. m_z is the inversion layer effective mass, which is $m_l = 0.98$ for the z-valleys and $m_t = 0.19$ for the x and y-valleys. If we use a lower estimate for the electric field ($\mathscr{E} = 10^5 \ \frac{V}{cm}$), Eq. 2.22 gives $E_0 = 37 \ meV$ for the z-valleys and $E_0 = 63 \ meV$ for the x and y-valleys. For reference, $kT = 25.8 \ meV$ at 300 K and $kT = 172 \ \mu eV$ at 2 K. Thus at cryogenic temperatures, the confinement from the triangular potential at the SiO_2-Si interface results in electrons being primarily in the z-valleys since they are many kT lower than the other valleys. We show this situation numerically in Fig. 2.5(b). As mentioned previously, this electrostatic confinement is not the only way to break the valley degeneracy. Strain will affect the valleys differently depending on the sign and direction of the strain.

Deformation potential theory was developed by Bardeen and Shockley [84] and applied by Herring and Vogt [85] in 1955 in their transport studies of semiconductors. Here, they used a set of energies, Ξ, to label shifts in the conduction band valley minima. For normal strains, we have two constants referred to as deformation constants. Ξ_u and Ξ_d, are the uniaxial and dilation deformation constants, respectively. For bulk silicon, experimental values of Ξ_u range from 8.7 eV [86, 87, 88] to 9.6 eV [89] with theoretical values in the range of 8-10.5 eV [90]. In contrast, Ξ_d ranges from 1.1 eV [86, 89] to 5.0 [90]. In this work, we use the experimental values from ref [86] ($\Xi_u = 8.7$ eV, $\Xi_d = 1.1$ eV). The change in energy for the pairs of

valleys is:

$$\begin{aligned}\Delta E_{c,k_x} &= \Xi_u \epsilon_x + \Xi_d(\epsilon_x + \epsilon_y + \epsilon_z) \\ \Delta E_{c,k_y} &= \Xi_u \epsilon_y + \Xi_d(\epsilon_x + \epsilon_y + \epsilon_z) \\ \Delta E_{c,k_z} &= \Xi_u \epsilon_z + \Xi_d(\epsilon_x + \epsilon_y + \epsilon_z)\end{aligned} \tag{2.23}$$

As an example of the energy scale involved, consider a strain only in the z-direction of $\epsilon_z = 10^{-4}, \epsilon_x = \epsilon_y = 0$, corresponds to approximately 1 meV. This level of ΔE_c is comparable to the intended electrostatic modulations in MOS QD devices. It should also be noted that it is possible for strains in the silicon MOS system to work against the valley splitting due to confinement, e.g. $\epsilon_z > 0$,$\epsilon_x < 0$, and $\epsilon_y < 0$. In practice, this situation does not occur in any of the devices in this work. Consider an oxidized bulk silicon wafer discussed in the preceding sections, where the oxide is under compressive biaxial stress $\sigma_{xx} = \sigma_{yy} = -\sigma_b$ and uniform across the entire wafer. If the oxide is under compressive stress than the silicon must be under tensile stress $\sigma_{xx} = \sigma_{yy} = \sigma_b$. Solving for the strains in the silicon yields: $\epsilon_{xx} = \epsilon_{yy} = \frac{1-\nu_{xy}}{E_x}\sigma_b$ and $\epsilon_{zz} = -2\nu_{xy}\sigma_b$. Thus $\epsilon_{xx} > 0$ and $\epsilon_{yy} > 0$ while $\epsilon_{zz} < 0$. This means the z-valleys will move to lower energy while the x and y valleys will move to higher energy. Thus, the biaxial stress from the oxidation of the silicon will actually enhance the splitting between z and x/y valleys due to confinement.

The above discussion has only considered the modulation of E_c due to normal strain components as that is the dominate strain in the MOS QD devices studied in this work. Shear strains can also modulate E_c[91]. The modulation of E_c with

in-plane shear strain (ϵ_{xy}) is given by:

$$\Delta E_{c,shear} = \begin{cases} \frac{-\epsilon_{xy}^2}{4\kappa^2} & |\epsilon_{xy}| < \kappa \\ \frac{-(|\frac{2\epsilon_{xy}}{\kappa}|-1)\Delta}{4} & |\epsilon_{xy}| > \kappa \end{cases} \tag{2.24}$$

Where $\Delta = 0.53\ eV$ and $\kappa = \frac{\Delta}{4\Xi_{u'}}$ and the shear deformation potential $\Xi_{u'} = 7.0\ eV$ [91]. In contrast to modulation with normal strains, the modulation of E_c does not depend on the sign of the strain.

2.3 Thin film stress measurements

An important aspect that this thesis contributes to studying mechanical effects in Si MOS QDs is the measurement of thin film stresses for common gate materials and typical fabrication processing conditions. We measure the thin film stress using the Tencor FLX-2320-S film stress measurement system. Images of the FLX-2320-S are shown in Fig. 2.6 (a) and (b). The FLX-2320-S measures stress by using a laser interferometer setup to measure the radius of curvature of a wafer. This setup is shown in Fig. 2.6 (c), where the system uses a laser and a series of mirrors to scan across the wafer. The measured deflection angle vs scan position is used to calculate the radius of curvature along that scan direction. Typical radius of curvature values for the films in this work range from few 100 m to 5 km, but the tool is capable of measuring from 2 m to 33 km. The FLX-2320-S is capable of performing these scans in 15° increments as shown in Fig. 2.6 (d), but for all the data presented in this work we use 30° increments. The tool is also capable of heating the substrate

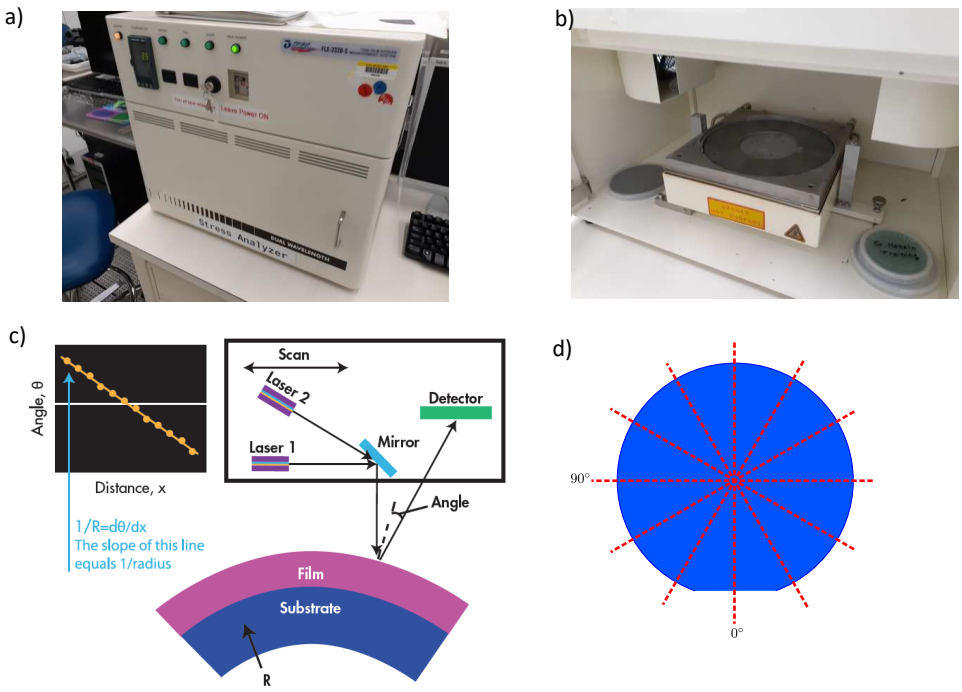

Figure 2.6: (a) Image of the Tencor FLX-2320-S film stress measurement system in the NIST cleanroom. (b) Image of the inside of the FLX-2320-S showing the substrate chuck. (c) Schematic of the measurement setup from Toho Technology.

up to 500 °C. The Stoney equation [92, 93] gives a relationship between the radius of curvature without the film (R_{bare}), the radius of curvature with the film (R_{film}), and the thin film stress (σ_{film}):

$$\sigma_{film} = \frac{E_{sub}}{1-\nu_{sub}} \frac{t_{sub}^2}{6t_{film}} \left(\frac{1}{R_{film}} - \frac{1}{R_{bare}} \right). \tag{2.25}$$

Where subscripts with "sub" refer to substrate values and "film" refer to film values. In order to calculate σ_{film} from Eq. 2.25, we need to measure both the substrate and the thin film thicknesses and we will need to use an accepted value for the biaxial modulus ($\frac{E_{sub}}{1-\nu_{sub}}$) of the substrate. It is assumed in the derivation of Eq. 2.25 that

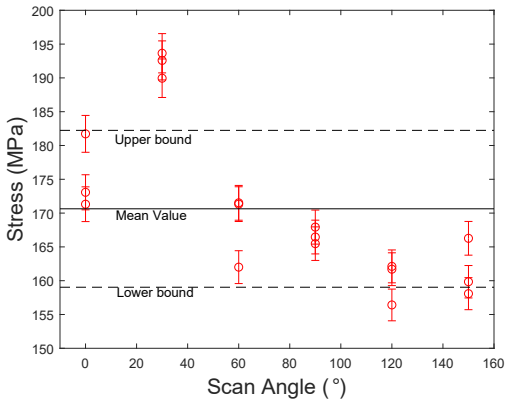

Figure 2.7: Angular dependence of stress measurements on 100 nm thick Al film.

$t_{film} << t_{sub}$ and that both the film and substrate are homogeneous and isotropic along the scan direction [94]. As mentioned before, all of the substrates used in this work are (100) silicon wafers. Thus, we use a biaxial modulus of $1.805 \times 10^{11}\ Pa$. We measure both the thickness of the substrate and film using a thickness gauge and profilometry respectively. The film thickness will vary depending on the process, with the most common thickness being in the range of 50-60 nm. The substrate thickness is fairly consistent and taken as $675 \pm 25\ \mu m$ for all calculations using Eq. 2.25. We will also use Eq. 2.25 to compare changes in film stress for other processing steps other than deposition, specifically after annealing. It is important to note that the stress described by Eq. 2.25 is equilbiaxial in nature ($\sigma_{xx} = \sigma_{yy}$, $\sigma_{zz} = 0$, and no shear stress) and assumes that the film thickness is uniform across the wafer [94]. Fig. 2.8 shows the sign conventions used in wafer curvature measurements. A negative change in curvature ($R_{bare} < R_{film}$) corresponds to the situation that the film is under compressive stress. Likewise the opposite change in

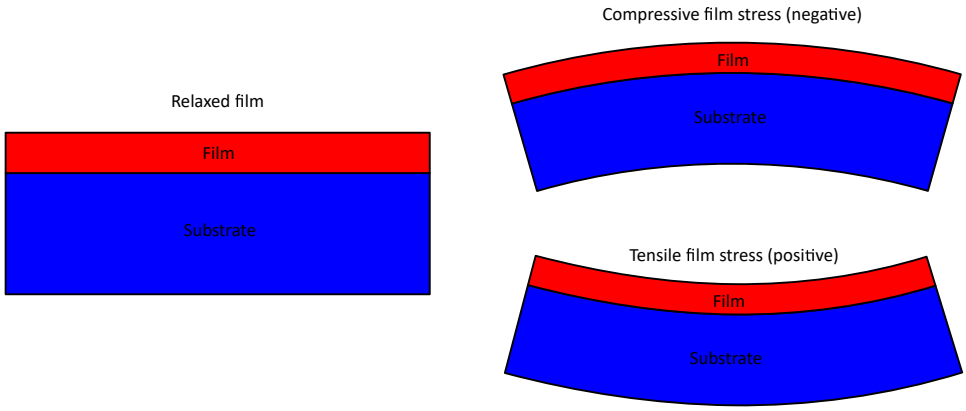

Figure 2.8: Diagram of the sign conventions used for wafer curvature measurements. The curvature is positive for tensile film stress and negative for compressive film stress.

curvature ($R_{bare} > R_{film}$) means that the film is under tensile stress. Additionally while we have only discussed thin film stresses resulting purely from the deposition process in later chapters we will consider the changes in the film stress with different fabrication processes after deposition, particularly annealing. In those situations, we will still use R_{bare}, but instead use the curvature measured post-annealing in order to calculate the new film stress. Fig. 2.7 shows an example of the angular dependence of the measured stress for an Al film. For any stress values reported later in this work, we use the average value of three measurements taken at each angle. The above discussion has focused on using the FLX-2320-S to measure σ_{film}, but in order to quantify mechanical effects in silicon MOS QDs we need to also measure the thin film coefficient of thermal expansion α_{film}. To measure α_{film}, we step the substrate temperature while measuring σ. We fit the resulting data to

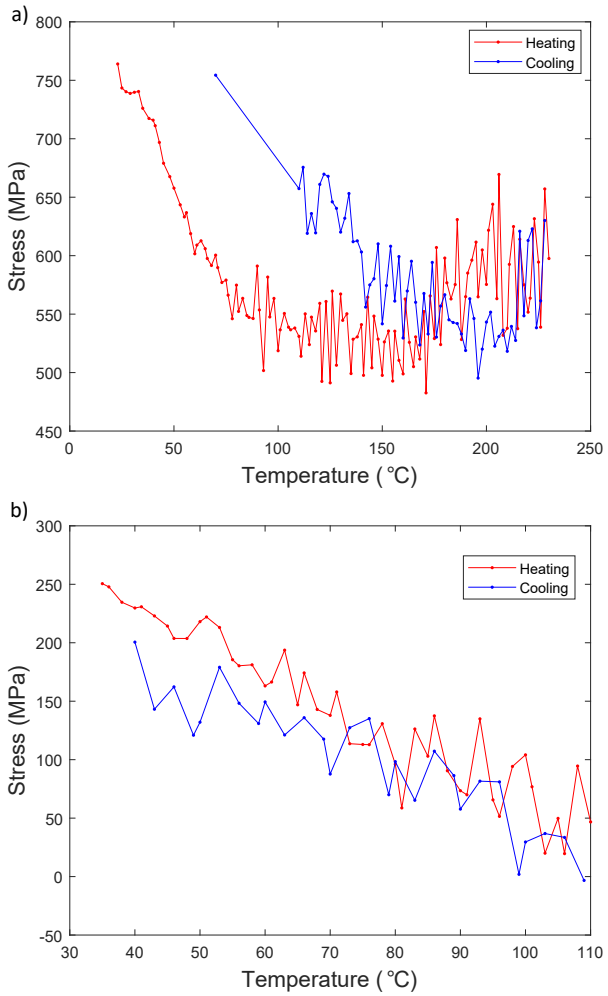

Figure 2.9: Plots of measurements of α_{film} by measuring stress as a function of temperature for Al films. (a) Case where the temperature step is high enough to cause non-elastic behavior. (b) Case where only elastic behavior is observed.

a line and extract α_{film} through:

$$\frac{\delta\sigma}{\delta T} = \frac{E_{film}}{1-\nu_{film}}(\alpha_{sub} - \alpha_{film}) \tag{2.26}$$

Where α_{sub} is the coefficient of thermal expansion of the silicon substrate and E_{film}, and ν_{film} are the Young's modulus, and Poisson ratio of the film, respectively. We have used the bulk values [79] for these three quantities in our calculation of α_{film}. It is important to note that the linearity of stress vs temperature does not hold for all temperatures. This is a deviation from elastic behavior and is not useful in the determination of α. Fig. 2.9 shows the raw data used in measuring α for Al films on silicon wafers. In Fig. 2.9(a), the substrate is heated to a high enough temperature where the film undergoes non-elastic deformations. Fig. 2.9(b) shows the case where the substrate temperature is limited such that we never heat it high enough to deviate from elastic behavior, denoted by the cooling curve returning close to the original state. For the materials in this work, we limit the measurements of α to temperatures less than 110 °C to avoid this issue.

2.4 Simulations in COMSOL

The complex nature of the material stacks used in MOS QD devices makes solving elastic mechanics problems discussed earlier in this chapter difficult without numerical methods. Finite element modeling (FEM) is commonly used to simulate such continuum mechanics problems. In this work, we have used COMSOL Multiphysics for all finite element simulations. The premise behind FEM is that any

object can be broken up into small discrete elements and those quantities, such as stress and strain, are described by the displacements of the individual elements. For the derivation here, we will assume a 2-dimensional picture in order to understand the FEM process following examples in Ref. [1]. Similar to Sect. 2.1.4, we will again use the plane stress approximation ($\sigma_{zz} = \sigma_{xz} = \sigma_{yz} = 0$). In this case Hooke's law for an isotropic material becomes:

$$\begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{xy} \\ \epsilon_{xz} \\ \epsilon_{yz} \end{pmatrix} = \frac{1}{E} \begin{pmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 1+\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+\nu \end{pmatrix} \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{xy} \\ \sigma_{xz} \\ \sigma_{yz} \end{pmatrix} \tag{2.27}$$

Consider the arbitrary shaped object shown in Fig. 2.10, where we have drawn a triangular element inside. The corners of the triangle are referred to as nodes with node number i having positions of (x_i, y_i). Under applied forces $(F_{x,i}, F_{y,i})$, the nodes will undergo a displacements of $(u_{x,i}, u_{y,i})$. We can write the strain in the triangular element solely in terms of the displacement of nodes, similar to the situation described in Fig. 2.1. By calculating the displacement in each node, we can in turn calculate the strain in the triangular element [95]. This discretization of the displacement field is the essence of FEM in solid mechanics problems. The

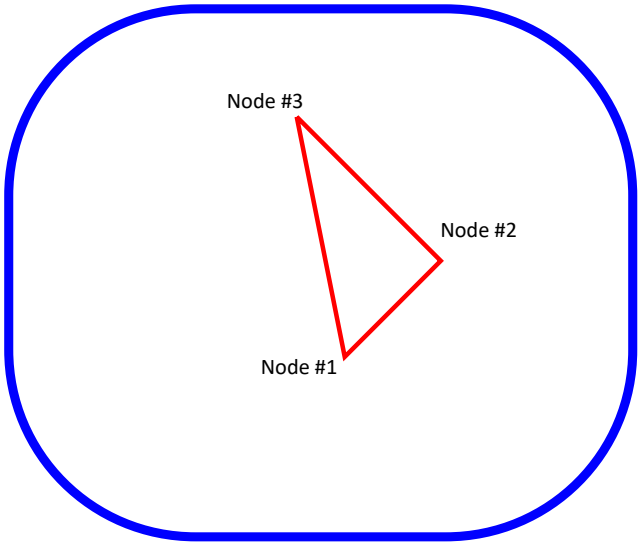

Figure 2.10: An arbitrary sized object with a triangular node used to build a FEM model.

strain and displacements for the three nodes in Fig. 2.10 are related by [95]:

$$\begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{xy} \end{pmatrix} = \boldsymbol{B} \begin{pmatrix} u_{1,x} \\ u_{1,y} \\ u_{2,x} \\ u_{2,y} \\ u_{3,x} \\ u_{3,y} \end{pmatrix} \tag{2.28}$$

Where $\boldsymbol{B}$ is given by in terms of the area of the triangle A:

$$\boldsymbol{B} = \frac{1}{2A}\begin{pmatrix} y_2 - y_3 & 0 & y_3 - y_1 & 0 & y_1 - y_2 & 0 \\ 0 & x_3 - x_2 & 0 & x_1 - x_3 & 0 & x_2 - x_1 \\ x_3 - x_2 & y_2 - y_3 & x_1 - x_3 & y_3 - y_1 & x_2 - x_1 & y_1 - y_2 \end{pmatrix} \tag{2.29}$$

If we meshed the entire object in Fig. 2.10 with different triangles, we would have a different $\boldsymbol{B}$ for each one. For an object with N nodes, we can write Hooke's Law in terms of forces, F, and displacements, u, at each node: $\vec{F} = -\boldsymbol{K}\vec{u}$. Here, F and u will be $2N \times 1$ matrices and K will be $2N \times 2N$. $\boldsymbol{K}$ is known as the global stiffness matrix. For any node that is on the exterior edge of the object, we must specify boundary conditions in order to solve the problem. This requires we either specify the force on the node or the displacement of the node. For the models used in this work, we always specify the displacement to be zero at specific points. This is to avoid uniform displacements of any object. For a 3d model of a MOS device, we apply this boundary condition at the bottom of the silicon substrate. Additionally for nodes interior to the object, the sum of the forces acting on the node must be zero in the same way as when we discussed stress in the preceding sections. From the above, we always know either the force or displacement at every node and therefore we can solve Hooke's Law as long as we know $\boldsymbol{K}$. $\boldsymbol{K}$ is assembled using the stiffness matrix, $\boldsymbol{k}$, at each node [95]:

$$\boldsymbol{k} = A\boldsymbol{B}^{\boldsymbol{T}}\boldsymbol{C}\boldsymbol{B} \tag{2.30}$$

Where A is the element area, $\boldsymbol{C}$ is the compliance matrix, and $\boldsymbol{T}$ represents the transpose. Once $\boldsymbol{K}$ is known, we can use Hooke's law to calculate the displacement at every node and from that calculate the strain at every node using Eq. 2.29. Finally with the strain at every node, we can then calculate the stress using Eq. 2.27. This is a simplified version of what COMSOL is doing in order to solve elastic mechanics problems.

2.4.1 Bimetallic strip in COMSOL

To test setting up models in COMSOL, we can return to the bimetallic strip problem discussed in Sect. 2.1.4. The bimetallic strip is a good test as it is a relatively simple model with a purely analytical solution. Fig. 2.11(a) shows the model setup in COMSOL to simulate the bimetallic strip with a thin film 10 nm thick and substrate 200 nm thick. For this 2d problem, we have two boundary conditions we need to specify. First, we set all displacements in lower left corner to zero. Second, we set only the y-displacements in the lower right corner to zero. This prevents the strip from rotating around the point where we specified the other boundary condition. We also made the strip 2 μm long to satisfy the assumptions made in the derivation of analytical solution for the bimetallic strip. Using the elastic constants and α discussed in the preceding sections, we can calculate the stress and strain induced in cooling the strip down from room temperature to 2 K. We show this in Fig. 2.11(b), which shows the interpretation of the deformation of the strip and the color scale represents the stress. We can now compare the values

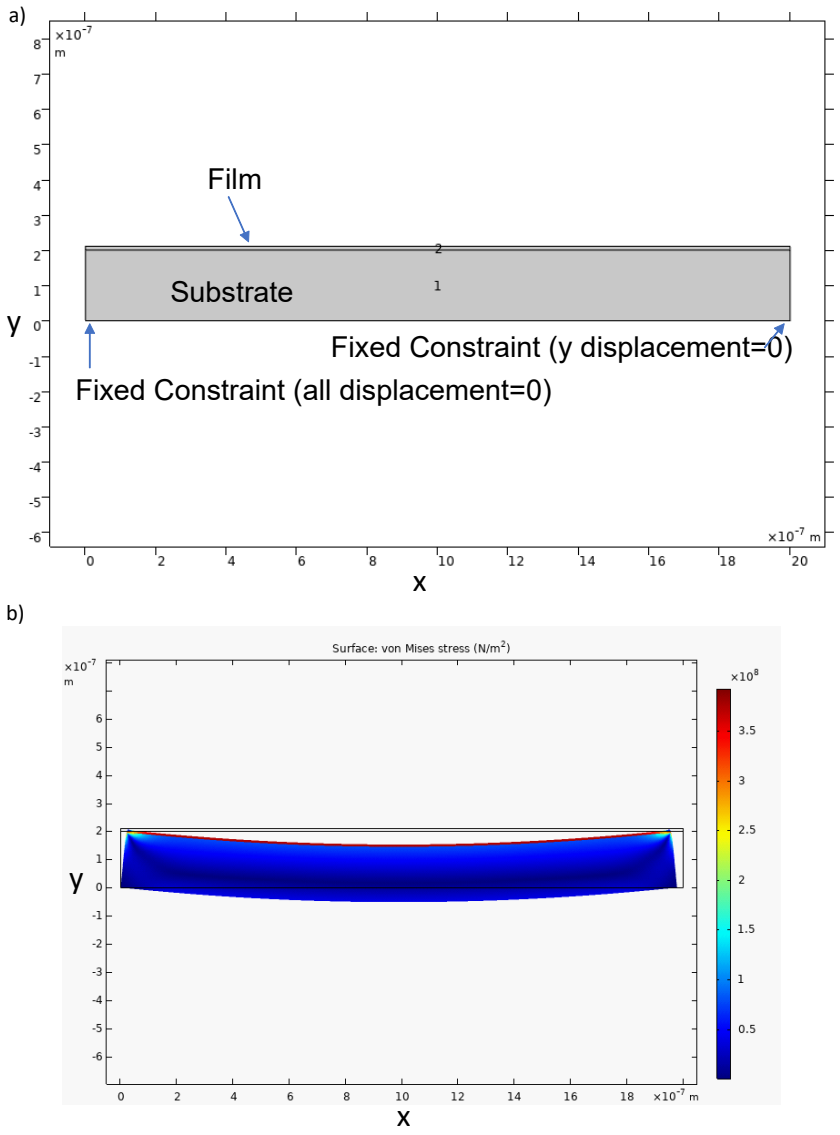

Figure 2.11: (a) Diagram of the bimetallic strip model in COMSOL. (b)Example of an output for stress in bimetallic strip cooled down from room temperature to 2 K. The black outline represents the strip at room temperature.

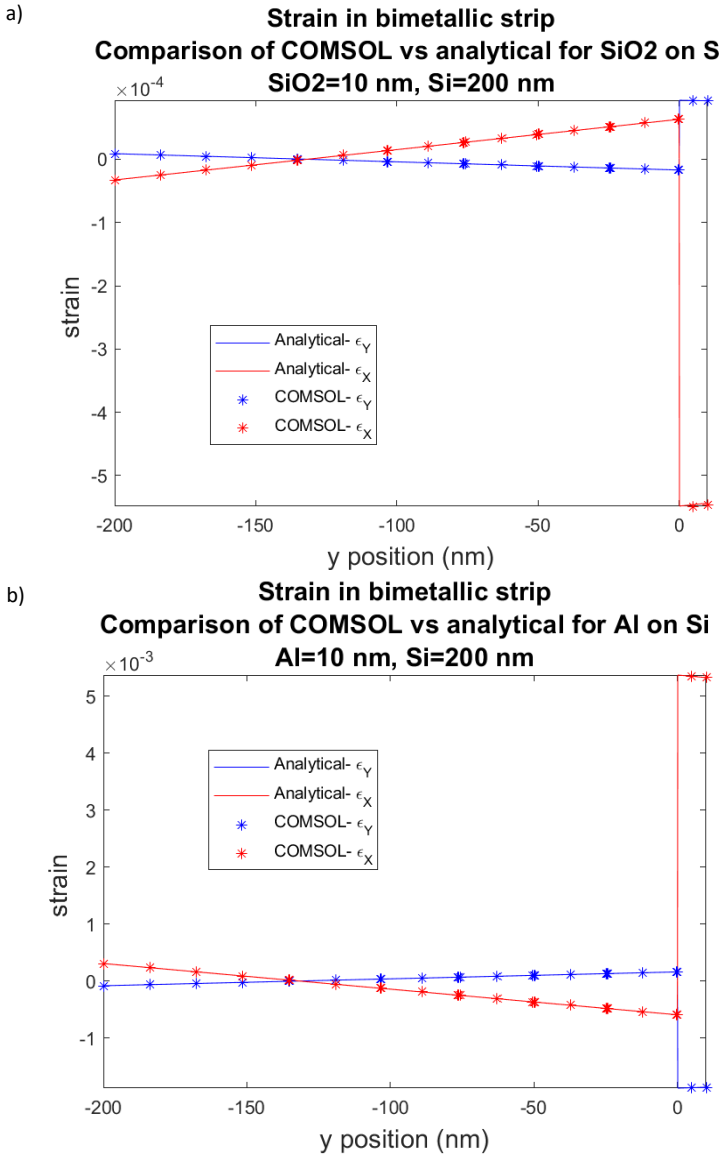

Figure 2.12: (a) Comparison of analytical and COMSOL solutions for a bimetallic strip made of a thin silicon dioxide film on Si. (b) Comparison of analytical and COMSOL solutions for a bimetallic strip made of a thin Al film on Si. Profiles are taken along the middle of the strip in the x-direction. $y = 0$ is the film and silicon interface, where $y > 0$ is in the film and $y < 0$ is in the silicon substrate.

calculated from COMSOL for this problem against the analytical solution from Sect. 2.1.4. We show the results for the strain in Fig. 4.3 for two different films: SiO_2 film in (a) and an Al film in (b). In both cases, the film is 10 nm thick. We can see that we achieve good agreement between COMSOL and the analytical solutions in both cases. This allows us to extend the principles used in setting up this bimetallic strip model into the slightly more complex systems necessary to simulate strain effects in MOS QD devices.

Chapter 3: Charge Defects in the MOS system

3.1 Introduction to the MOS system

The primary building block of the silicon MOS QD architecture is the silicon MOSFET. The principles used for understanding basic MOSFET operation act as a foundation to understand the more complicated MOS QDs. For instance, the formation of tunnel barriers in QD is analogous to a MOSFET in its off state or the formation of isolated regions of charge is analogous to a MOSFET in its on state. This chapter will cover the ideal properties of a MOS device and will show that deviations from the ideal model are primarily due to several types of oxide charge defects. The different defect densities are inferred from capacitance as function of voltage (CV) measurements on MOS capacitors (MOSCAPs). These charge defect measurements are crucial to understanding the impact that common MOS QD fabrication processes can have on the electrical properties of the device. Later chapters will use this foundation to study the changes in defect densities for different gate materials and anneals common to MOS QD fabrication.

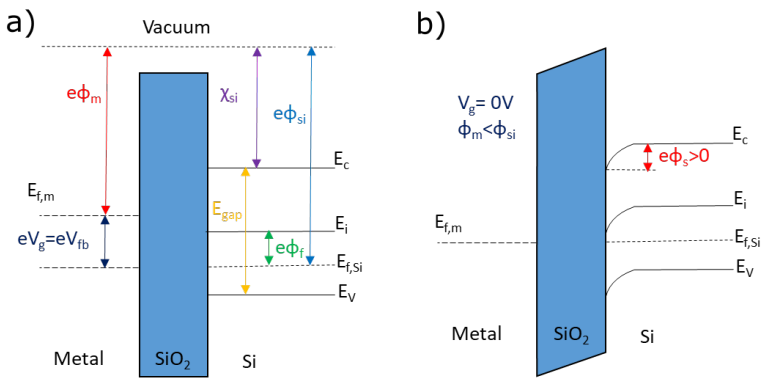

Figure 3.1: Band diagrams for various gate voltage configurations for the MOS system where the silicon is a p-type substrate. (a) Flatband condition where V_g=V_{fb}. (b) Bands at V_g=0 for the case of $\phi_m < \phi_{si}$ which is true most for elemental metal gates. Here, the bands bend downward at the interface by an amount depending on the difference between ϕ_m and ϕ_s.

3.1.1 MOS physics

The operating characteristics of any MOS device can understood by considering the band diagrams shown in Fig. 3.1. The diagrams show the potential energy perpendicular to SiO_2-Si interface and we have drawn them such that lower energy is more favorable to electrons and higher on the diagram is more favorable for holes. Fig. 3.1(a) shows a band diagram for a condition known as "flatband", so-called since the bands in the silicon near the SiO_2-Si interface are flat into the bulk far from the interface. Flatband is a special condition that occurs at a specific gate voltage, V_{fb}, in a MOS device. At flatband, the total charge density, Q_s, near the interface is zero, meaning no net electrons or holes and no ionized dopants or area near the interface is identical to silicon bulk from the interface. In an ideal case, V_{fb} is entirely dependent on the work function difference between the gate material and

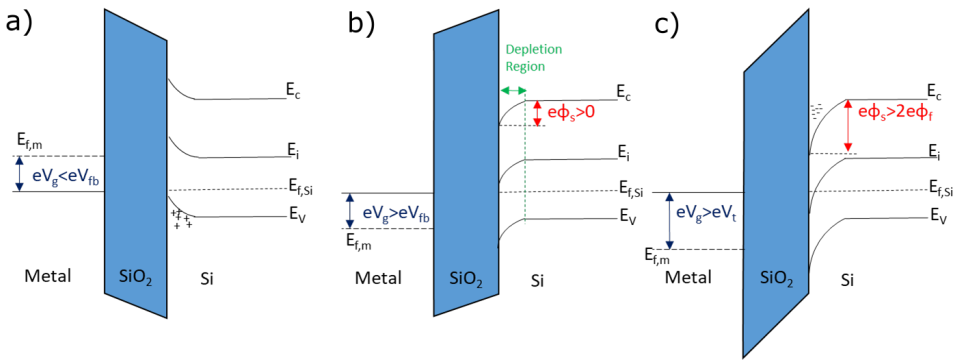

Figure 3.2: Band diagrams for various for space charge regions for p-type Si MOS. (a) Band diagram for accumulation when $V_g < V_{fb}$. (b) Band diagram for depletion when $V_t > V_g > V_{fb}$. (c) Band diagram for inversion when $V_g > V_t$.

the silicon substrate, labeled ϕ_m and ϕ_{si} in Fig. 3.1(a) respectively. Here, the bands in the silicon are E_c and E_v for the conduction and valence bands respectively. E_i is referred to as the intrinsic level since it is where the Fermi level would be in undoped or intrinsic silicon. Since the work function refers to the energy required to take an electron from the Fermi level of the material and promote it to the vacuum energy level, Fig. 3.1a shows the relative energy levels of the metal and silicon work functions relative to the vacuum energy level. For the case of a p-type silicon substrate, the metal-semiconductor work function difference (ϕ_{MS}) is:

$$\phi_{MS} = \phi_M - \phi_{Si} = \phi_M - \chi_{Si} - \frac{E_c - E_i}{e} - \phi_f \approx \phi_M - \chi_{Si} - \frac{1}{2e}E_{gap} - \frac{k_b T}{e} ln(\frac{N_a}{n_i}) \quad (3.1)$$

where the last three terms on the right hand side of the equation all describe the p-type silicon work-function and ϕ_f is the Fermi or bulk potential which gives the position of the silicon Fermi level due to doping relative to the intrinsic level given

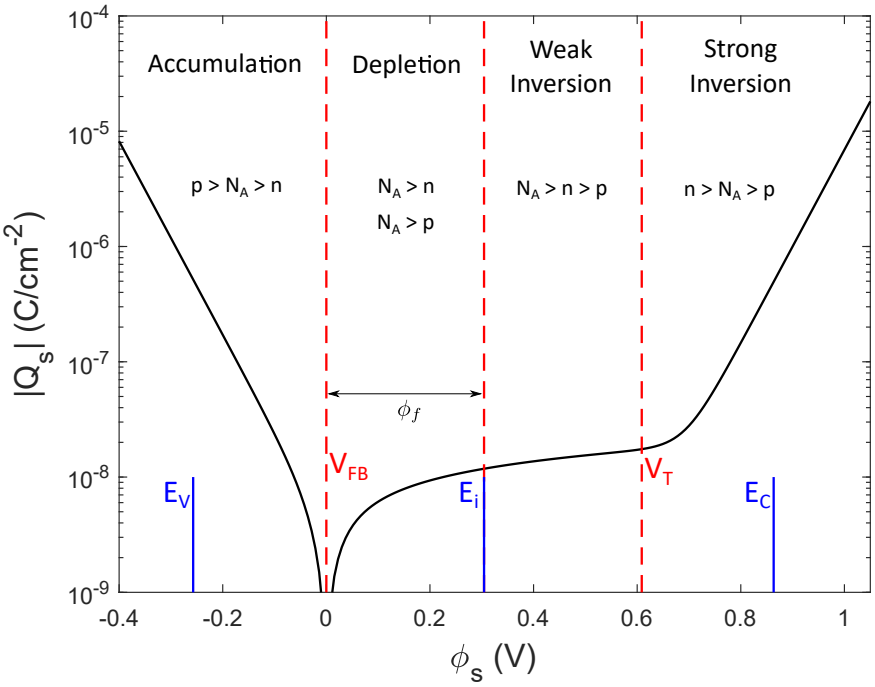

Figure 3.3: Plot of the space charge density, Q_s, in the silicon as function of ϕ_s. Here, we used conditions common to MOSCAPs used in this work with T=300 K and N_a= 1.5×10^{15} cm^{-3}. n and p refer to the surface charge density of electrons and holes respectively. Data is produced based on equations presented in Ref. [96].

by $\frac{k_bT}{e}ln(\frac{N_a}{n_i})$. In eq. 3.1, χ_{Si} is the electron affinity of Si, E_{gap} is the silicon band gap energy, and N_A and n_i are the acceptor doping concentration and intrinsic carrier concentration respectively and it is assumed that $N_a >> n_i$. Additionally, we have used the approximation that $E_c - E_i \approx \frac{E_{gap}}{2}$, which is not strictly true due the differences in the conduction and valence band density of states [97]. For the data in this thesis, the acceptor dopant is boron with a doping density in the range of $1-2\times10^{15}$ cm^{-3}. Using typical values for the substrates from this work, $\chi_{Si} = 4.05$ eV [96], $E_{gap} = 1.12$ eV [97], $N_A = 1.5\times10^{15}$ cm^{-3}, $n_i = 1.07\times10^{10}$ cm^{-3} [97], and $\phi_{M,Al} = 4.28$ eV [98], Eq. 3.1 gives $\phi_{MS} = -0.63$ V. It is important to

note that the bands near the interface in the silicon are not flat at V_g=0 as shown in Fig. 3.1(b) when $\phi_m < \phi_{si}$, which is typical for most gates made of elemental metals. Here, the bands are bent down near the interface and $Q_s \neq 0$.

To describe Q_s in the MOS device as a function of V_g, the surface potential, ϕ_s, is used. As shown in Fig. 3.2, ϕ_s refers the difference in energy at the SiO_2-Si interface relative to the bulk. At V_g=V_{fb} (Fig. 3.1a), ϕ_s is zero by definition. Fig. 3.2(a) shows the case of $\phi_s < 0$ (or $V_g < V_{fb}$) for p-type Si. Here, the bands are bent upward at the interface favoring an increase in the hole concentration (labeled p in units of cm^{-3}) relative to the bulk. As V_g is tuned more negative than V_{fb}, a highly conducting layer of holes will form at the surface, such that $p > N_a$ as shown in Fig. 3.3 where holes are the dominate contribution to Q_s. Since the surface now has a large concentration of majority carriers, this is the state known as accumulation for a MOS device. The $\phi_s > 0$ (or $V_g > V_{fb}$) case is slightly more complex. Initially as ϕ_s becomes positive, the bands begin bending down at the interface as shown in Fig. 3.2(b). This forces a decrease in the majority carrier concentration at the surface, such that $p < N_a$ as shown in Fig. 3.3. This is referred to as depletion since the hole concentration at the surface has decreased relative to the bulk. For values of ϕ_s between 0 and ϕ_f the now ionized dopants will be the dominant contribution to Q_s within the "depletion approximation". The size of this depletion region of ionized dopants will depend on ϕ_s as:

$$W_{dep} = \sqrt{\frac{2\epsilon_{si}\phi_s}{eN_a}} \tag{3.2}$$

With the charge in the depletion region given as:

$$Q_{dep} = eN_a W_{dep} = \sqrt{2\epsilon_{si}\phi_s eN_a} \tag{3.3}$$

Where ϵ_{si} is the permittivity of silicon. As ϕ_s increases further, the gate voltage pulls thermally generated electrons towards the interface. Since the surface now has a large concentration of minority carriers, this state of the MOS device is known as inversion. Typically, inversion is divided into two different regions based on the surface potential referred to as weak inversion, when $\phi_f < \phi_s < 2\phi_f$, and strong inversion, when $\phi_s > 2\phi_f$. For weak inversion, the electron concentration near the surface exceeds the hole concentration but is still less than N_a, as shown in Fig. 3.3. The tipping point into strong inversion occurs when ϕ_s is twice the Fermi potential, ϕ_f. Here, the depletion region reaches it maximum extent ($W_{dep,max}$) and charge ($Q_{dep,max}$) so that further increases in the gate voltage only serve to increase the electron density at the surface causing $n > N_a$. The gate voltage where the MOS device reaches strong inversion is called the threshold voltage, V_t:

$$V_t = \phi_{MS} + 2\phi_f + \frac{Q_{dep,max}}{C_{ox}} = \phi_{MS} + 2\phi_f + \frac{\sqrt{4\epsilon_{si}\phi_f eN_a}}{C_{ox}} \tag{3.4}$$

Beyond this point, the gate voltage dependence of the MOS system behaves similarly to accumulation but now with a highly conducting layer of electrons at the surface. The MOS physics described above covers the ideal case where there are no defects or charge in the oxide. In practice, these non-idealities can have a large impact on MOS

device operations, but importantly the band bending concepts described above can be used to characterize these defects. For instance, adding positive charge to the oxide, or increasing ϕ_s, will have the same effect as putting a slightly more positive voltage on the gate. In this way, non-idealities in MOS voltage characteristics allow the study of charge defects in the oxide. From this shift in potential and effects detailed below, there are four main charge defects in SiO_2 [99]: fixed oxide charge (Q_f), interface trap density (D_{it}), mobile ion charge (Q_m), and trapped oxide charge (Q_{ot}). We can describe the shift in V_{fb} due the contributions of these defects as [100]:

$$V_{fb} = \phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{Q_{it}(\phi_s)}{C_{ox}} - \frac{Q_m}{C_{ox}} - \frac{Q_{ot}}{C_{ox}} \tag{3.5}$$

It is important to note the differences in how the V_{fb} shift incorporates affects from each defect type, where Q_m, Q_{ot}, Q_f, and D_{it} are areal densities. These differences arise solely from the specific nature of each defect, which we will now discuss in detail.

3.1.2 Interface trap density, D_{it}

The interface trap density (D_{it}) describes a set of dangling bonds at the SiO_2-Si interface that are able to interact with the 2DEG in the silicon in two ways. First, since their charge state will depend on their occupancy they will contribute to the overall V_{fb} shift like all other oxide charges. There are two types of traps, which differ by their charge states: acceptor-like and donor-like. The charge states will be negative and neutral when occupied and unoccupied respectively for an acceptor-like

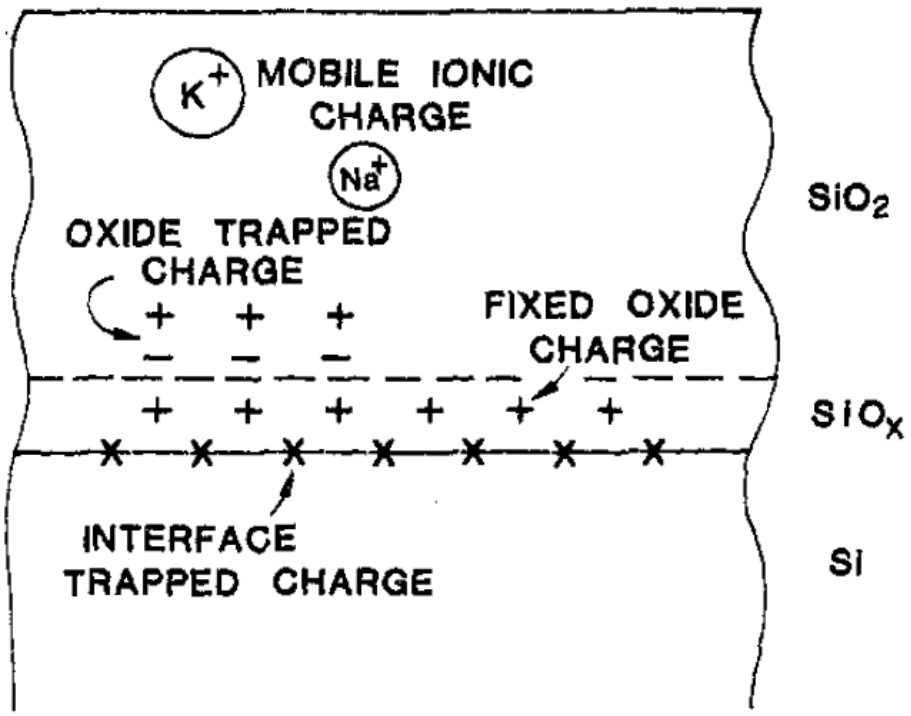

Figure 3.4: Schematic of the four main types of charge defects in SiO_2identified by the Deal committee in 1979. From B.E. Deal,"Standardized Terminology for Oxide Charges Associated with Thermally Oxidized Silicon", J. Electrochem. Soc. Vol 127, No. 4 (April 1980) pp. 979-981)

trap [101]. Similarly, the charge states will be neutral and positive when occupied and unoccupied respectively for a donor-like trap. An electron occupies a donor-like trap when the trap level is below the Fermi level in the band diagrams in Fig. 3.1. From this picture of D_{it} one expects that at V_{fb} the contribution of D_{it} to Eq. 3.5 would be positive [100, 101] on p-type Si, since donor-like traps dominate on the valence band side of the band gap [102]. This means at V_{fb} the donor traps between the Fermi level and midgap will be positively charge, while any acceptor traps on the conduction band portion of the band gap remain neutral. The second

way interface traps can interact with the 2DEG is by the capture and emission of carriers. The capture and emission of carriers from interface traps varies depending on the nature of the trap (donor or acceptor) and its energy position in the silicon bandgap. Therefore, the impact of D_{it} on the MOS system will vary with gate voltage. This contrasts with the three other types of defects, which will induce simple shifts in V_{fb}. D_{it} will cause nonlinear shifts in the gate voltage dependence, essentially stretching out the nominal behavior. Given this energy dependence, D_{it} is typically reported in units of $eV^{-1}cm^{-2}$ rather than purely spatial density.

Interface states are associated with broken bonds and structural defects at the SiO_2-Si interface, referred to as P_b centers. The density of P_b centers has been shown to be correlated with the density of interface states [103, 104]. For Si(100), there are two types of dangling bonds: P_{b0} and P_{b1}. P_{b0} consists of a silicon atom back bonded to three other silicon atoms with a dangling bond, while a P_{b1} consists of a silicon atom back bonded to two silicon atoms and one oxygen atom with a dangling bond [105]. These dangling bonds naturally form during the thermal oxidation process so typically the density is reduced by annealing MOS devices in forming gas (N_2 and H_2) [102, 106]. We will see in Chapter 5 that forming gas anneals can be used to reduce D_{it} below the 10^{10} $eV^{-1}cm^{-2}$.

3.1.3 Mobile Ion Charge, Q_m

The contamination of MOS oxides by metal ions with high mobility in SiO_2 is the primary source of Q_m. Since these ions possess high mobility in SiO_2, their

position in the oxide can drift with applied fields and elevated temperatures. The most common problematic ions seen in MOS devices are the alkali metals, such as Na^+, K^+, and Li^+ [107, 108]. The mobile ion density is predominately positive in nature and Q_m is distributed across the thickness of the oxide. The instability of the position of the mobile ions with applied fields means that their effect on the MOS system can vary with time, where mobile ions closer to the SiO_2-Si interface will have a stronger impact than those that are closer to the SiO_2-gate interface. The mobile nature of Q_m allows it to be easily distinguished from the other three oxide defects. They can easily drift with gate voltage at elevated temperatures, but would be effectively immobile at lower temperatures [107]. Q_m arises from contamination during the fabrication process; particularly its presence is tied to the cleanliness of the pre-oxidation cleaning process and the oxidation chamber [109]. Standard RCA cleaning procedures are generally sufficient at removing any contamination from the wafer surface prior to oxidation. It possible for ions to be introduced after oxidation, such as in Ref. [110] where K^+ ion contamination was traced to tungsten filaments used for evaporation for gate materials. For optimal operation of MOS devices, it desirable for the density to be below 10^{10} cm^{-2} level.

3.1.4 Trapped Oxide Charge, Q_{ot}

The another oxide charge in the silicon MOS system that is distributed the oxide thickness is the oxide-trapped charge (Q_{ot}). Q_{ot} is similar to Q_m in that under normal devices operating conditions its charge state is fixed, but differs in that Q_{ot}

may be either positive or negative. These traps, similar to D_{it} are broken bonds in the oxide that act as trapping center for electrons or holes [111]. Q_{ot} is also the most difficult to measure independently, because these electron or hole traps are unable to exchange charge with the silicon unlike D_{it}. Damaging processes, such as ionizing radiation [112], avalanche injection [113], and Fowler-Nordheim tunneling [114], can produce Q_{ot}. Similar to D_{it}, the density of these traps can reduced with relatively low temperature (<450 °C) anneals [111]. E-beam lithography processes, a key component of MOS QD fabrication, can lead to significant increases in the Q_{ot} density [115].

3.1.5 Fixed Oxide Charge, Q_f

The fixed charge density (Q_f) is the charge directly at the SiO_2-Si interface, within 2.5 nm [99] as shown in Fig. 6.9. This predominately-positive charge, unlike D_{it}, is unable to exchange charge with the 2DEG in the silicon so its charge remains fixed during device operation. The magnitude of Q_f is a function of the properties of the oxide growth such as oxidation temperature, post oxidation annealing, and ramping conditions [116]. Q_f will also depend on the oxidation ambient with oxides grown via a dry process (only O_2) resulting in significantly lower densities than oxides grown via a wet process (O_2 and H_2) [116]. The orientation of the silicon wafer is also important to minimizing the fixed oxide density during oxide growth. The (100) silicon wafer yields roughly a factor of 3 lower oxide charge density than the (111) surface for the same oxidation conditions at 920 °C[116]. This is because

the number of bonds through this crystal plane at the SiO_2-Si interface is minimized in the (100) case.

As we see in Sect.3.2.3, there is some difficulty in measuring Q_f at the interface as described above. This is due the fact that three other main types of oxide charge will contribute to the total effective fixed charge in the oxide. In Eq. 3.5, Q_f and D_{it} are the only terms without a dependence on the oxide thickness as they represent a true areal density. In practice, the Q_f value extracted using the CV methods in this thesis will correspond to the total charge, projected to an areal density at the interface.

The four types of oxide charge we have described above are not necessarily the full picture of the SiO_2-Si system. Fleetwood [117] suggested that it necessary to add an additional type of defect called border traps(Q_{bt}) shown in Fig. 3.5. Border traps technically would fit into the category of Q_{ot} based on Ref. [99]. These border traps are within a few nanometers of the SiO_2-Si interface so it is possible for them to capture/emit carriers like D_{it}, but unlike D_{it} are not due to P_b related structures instead they are most commonly identified with oxygen vacancy related defects, called E' defects [118, 119]. It is difficult to write a firm distinction between Q_{ot} and Q_{bt} outside of characteristic switching times for each. Here,"fast" traps are associated with D_{it} while Q_{bt} would be "slow" traps since their distance from the interface limits the exchange. D_{it} would occupy a range of switching times from roughly 10^{-6} to 10^{-3} s and Q_{bt} would extend to any times greater than 10^{-3} s [120]. With this new distinction, Q_{ot} would describe the so-called fixed states since they are unable to exchange charge with the silicon under normal circumstances. For the

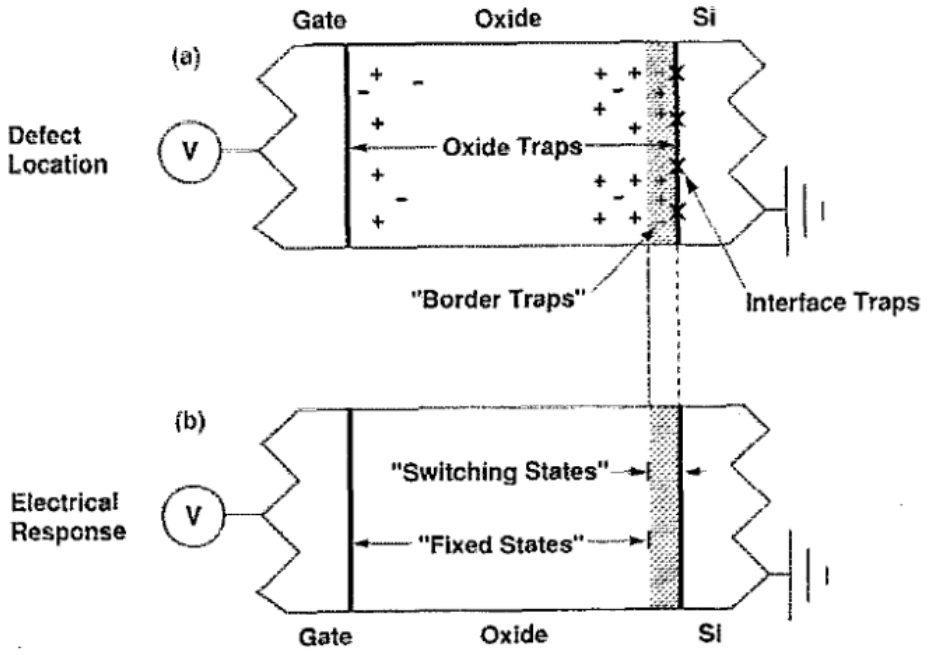

Figure 3.5: Schematic showing the differences between D_{it}, Q_{ot}, and Q_{bt} in SiO_2. Reprinted from D. M. Fleetwood, P. S. Winokur, R. A. Reber Jr., T. L. Meisenheimer, J. R. Schwank, M. R. Shaneyfelt, and L. C. Riewe, "Effects of oxide traps, interface traps, and "border traps" on metal-oxide-semiconductor devices" , Journal of Applied Physics 73 , 5058-5074 (1993) https://doi.org/10.1063/1.353777, with the permission of AIP Publishing.)

purposes of this work, we will treat Q_{bt} as just another contribution to Q_{ot}. We note that it is potentially important to consider the impacts of Q_{bt} for silicon MOS quantum dots. The slow trap states that make up Q_{bt} have been attributed as one of sources of $^1/_f$ or $^1/_{f^2}$ noise in MOS devices [120]. This noise and the slow switching times could lead to similar issues with instability in MOS QDs.

3.2 Capacitance-Voltage (CV) measurements

The previous section described the basic physical principles for the MOS system from the perspective of the ideal case and the impact of the non-idealities. This section will focus on the techniques used in this dissertation to measure the properties of the MOS system. Here, we show how to use measurements of a MOS capacitator to extract various oxide defect densities. There are a variety of techniques available, such as high-frequency CV, Conductance, and Quasi-static CV, all of which make use of the frequency dependence of the MOS capacitor impedance.

3.2.1 Frequency dependence of CV measurements

Fig. 3.6(a) shows a comparison of measured high frequency CV data taken at 1 MHz and low frequency data taken via the quasi-static method, with an approximate frequency ranging from 100-500 mHz derived from the slow DC voltage ramp. To understand this data we model the system as in Fig. 3.6(c). Here C_s is the semiconductor capacitance which changes as a function of ϕ_s , C_{ox} is the oxide capacitance, and C_{it} is the capacitance associated with D_{it}. In accumulation, or gate voltages less than -2 V, the HF and QS curves are identical and equal to C_{ox} as C_s becomes significantly larger such that the total capacitance in Fig. 3.6(b) and (c) reduces to approximately C_{ox}. In this range of voltage, the negative charge building on the gate pulls holes towards the SiO_2-Si interface. The MOS device is essentially a simple parallel plate capacitor with a SiO_2 dielectric. The HF and QS curves begin to diverge from each other near 0 V on the gate, corresponding to

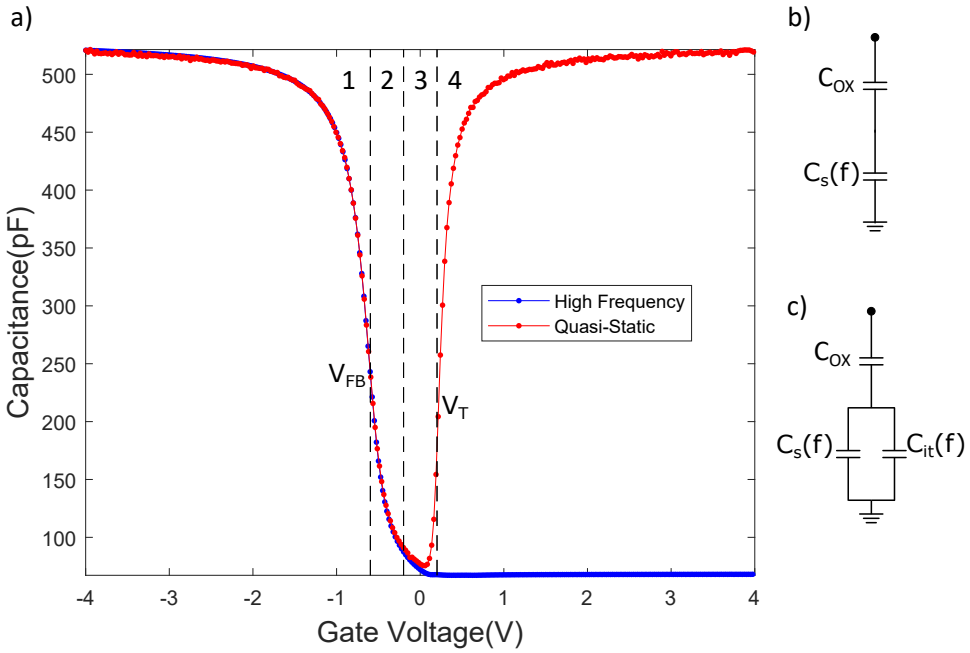

Figure 3.6: (a) Example of typical measured high frequency (blue data taken at 1 MHz using a LCR meter) and quasi-static (red data taken with a parameter analyzer using a 2 sec integration time) CV curves. This data is from a p-type substrate so the accumulation portion of the CV curve occurs at gate voltages less than zero, where the HF and QS curves are nominally identical. The HF and QS curves diverge at voltages greater than zero as the minority carriers are unable to follow the HF signal. The numbered regions are as: 1) Accumulation, 2) Depletion, 3) Weak Inversion 4) Strong Inversion. (b) Circuit diagram for the HF measurement. (c) Circuit diagram for the QS measurement, which includes a contribution due to the presence of D_{it}. Other details: e-beam evaporated Al for the gate material, 25 nm thick gate oxide grown at 1000 °C, and annealed at 350 °C for 30 mins in 10 % forming gas.

the depletion/weak inversion region of the CV curve. We will see in the proceeding sections that this separation is in fact due the frequency response of D_{it}, shown in the circuit diagrams in Fig. 3.6(b) and (c). As the gate voltage moves toward values that are more positive, the device enters weak inversion and the HF and QS curves diverge significantly. This change is now due to the frequency dependence of the semiconductor capacitance, C_s. In the HF curve, the electrons are unable to follow the 1 MHz AC excitation and build an inversion layer at the SiO_2-Si interface. The flat capacitance in this portion of HF curve is solely due the depletion region reaching its maximum equilibrium value [100]. Conversely, in the QS curves, the electrons are able to follow the slower signal and form an inversion layer. In this case, C_s is large again, the total capacitance reduces to C_{ox} and we are back to a simple parallel plate capacitor.

3.2.2 Measurement setup

Fig. 3.7 shows a schematic diagram of the typical MOS capacitor measurement. We make electrical connections to the MOS gate material via flexible tungsten probe tips and micro positioning stages. For high frequency measurements, we measure capacitance via a BNC cable connection between the metal probe tip and the Agilent E4980 LCR meter. The LCR meter applies both a DC voltage and an AC voltage with frequencies up to 2 MHz to the MOSCAP gate. For quasi-static measurements, we make this connection with triax cables in order to use the Agilent 4156C parameter analyzer to measure currents. This measurement is made with

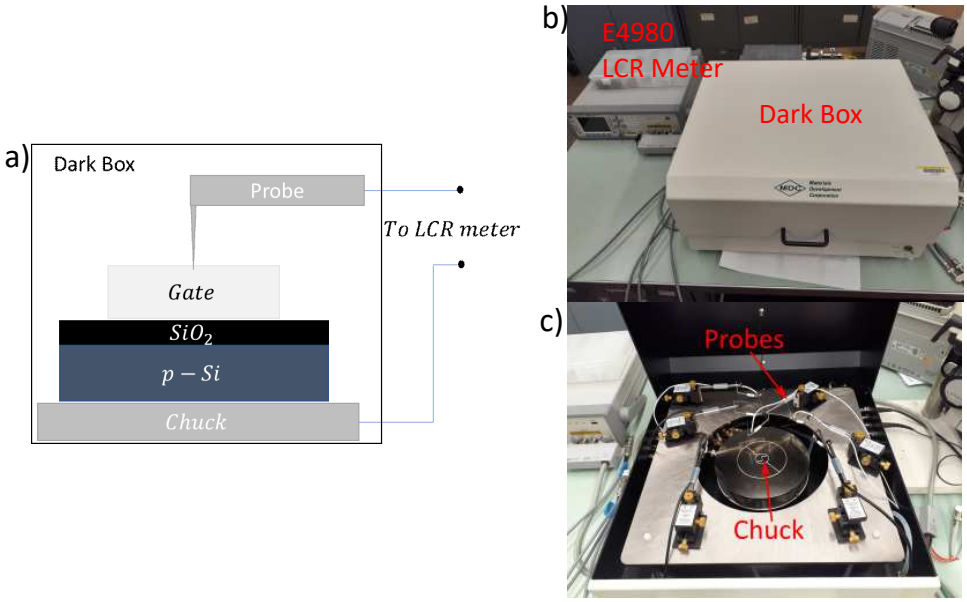

Figure 3.7: (a)Schematic drawing of the CV measurement setup. To avoid complications from photo-excited carriers, the dark box enclosure blocks any stray light from hitting the samples. (b) Image of the dark box with lid closed and the E4980 LCR meter used for taking high frequency CV measurements. (c) Image with dark box open showing the vacuum chuck and measurement probes.

slow dc voltage ramps rather than a AC excitation, with typical ramp time values ranging from 1-3 seconds for a 25 mV voltage step. For both measurements, we make electrical connection to the substrate via a metal vacuum chuck and metallize the back of the silicon wafer to reduce the parasitic resistance. The measurement setup includes a dark enclosure to avoid light affecting the carrier concentrations and response time. The chuck can be heated up to 200 °C, which is necessary for measuring Q_m.

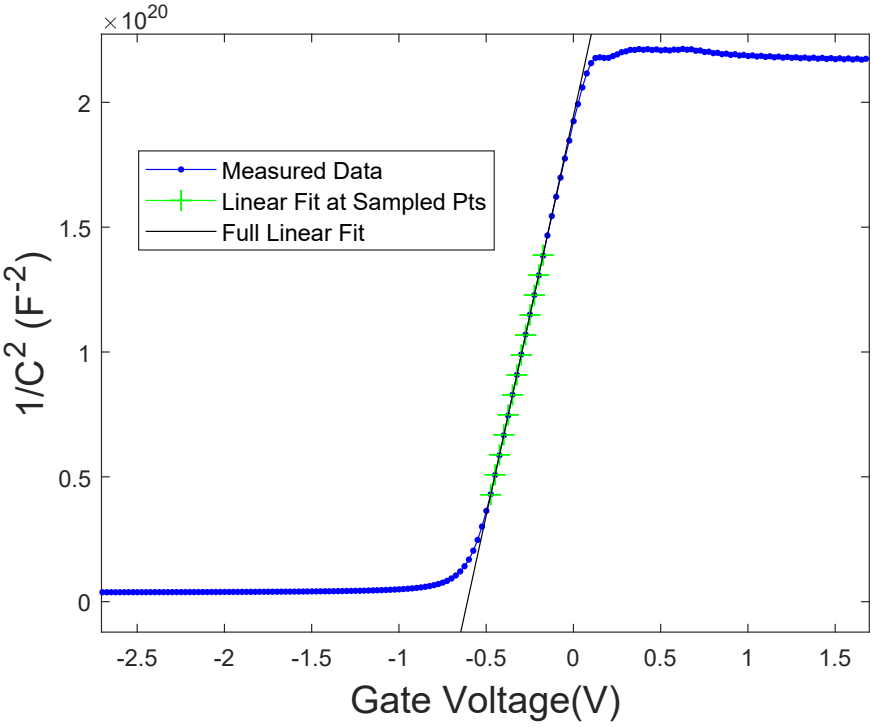

Figure 3.8: Example plot used for extracting flatband voltage. The raw data here are the high frequency data from Fig 3.6. The green crosses represent the subset of full data set used in the linear fit representing the depletion portion of the MOS characteristics. We have chosen this region manually by eye for each individual MOSCAP.

3.2.3 Measurement of Q_f: Flatband voltage extraction

The flatband voltage, V_{fb}, is calculated based on the linear fit of the $^1/_{C^2}$ vs V_g plot in depletion region of the CV curve. We extract the V_{fb} by finding when the linear fit (with slope = m and y-axis intercept = b) reaches a value of $^1/_{C_{OX}^2}$ [100]:

$$V_{fb} = \frac{\frac{1}{C_{OX}^2} - b}{m}. \tag{3.6}$$

Fig. 3.8 shows an example of a typical $^1/_{C^2}$ vs Gate Voltage plot for one of Al MOSCaps in this thesis. For this data, $V_{fb} = -0.64$ V.

All of the charge defects listed in Eq. 3.5 can contribute to a shift in flatband voltage. For the purposes of analyzing the CV data in the rest of this thesis, we combine all of the contributions to V_{fb} into a single effective fixed charge density, Q_f^{eff}:

$$Q_f^{eff} = Q_f + Q_{it}(\phi_s = 0) + \int_0^{t_{ox}} \frac{x}{t_{ox}} \rho_m(x) dx - \int_0^{t_{ox}} \frac{x}{t_{ox}} \rho_{ot}(x) dx \tag{3.7}$$

where Q_f^{eff} is related the measured V_{fb} and the oxide capacitance per unit area, C_{ox}, by:

$$Q_f^{eff} = \frac{Cox}{e}(\phi_{MS} - V_{fb}) \tag{3.8}$$

The individual contributions to a flatband shift from D_{it} or Q_m would need to be calculated based on the defect densities obtained from other methods and removed from this effective fixed charge density to obtain a better estimate of the true fixed charge at the interface. There is some contribution from Q_{ot} in the oxide but isolating it from the total requires Q_f as a function of oxide thickness or other processing parameters such as annealing [101]. Ideally, we would measure Q_m, D_{it}, and Q_{ot} independently. In practice, we generally focus on measuring Q_f^{eff}, using the method described above, and D_{it}, using methods in proceeding sections. We find this is sufficient because both Q_m and Q_{ot} are not intrinsic to the device. Q_m relates to cleanliness and we avoid operating devices in such a way to substantially change Q_{ot}. Q_m is routinely checked to ensure nominal processing conditions, but we do not measure it for each device to avoid heating and causing a reaction between the

vacuum chuck and the metallized back contacts necessary for D_{it} measurements on our dark box chuck.

3.2.4 Measurement of D_{it}: High Frequency-Quasi-static (HF-QS) method

Using the frequency dependence discussed in Section 3.2.1, we calculated D_{it} as function of gate voltages based on the measured high frequency CV data (C_{HF}) and the measured quasi-static CV data (C_{QS}) [100, 121]. Here, we make use of the differences in the high and low frequency circuit models shown in Fig. 3.6(b) and (c) respectively. In the high frequency model, we have only two capacitances: the voltage independent C_{ox} and the voltage dependent semiconductor capacitance, C_s. The measured high frequency capacitance, C_{hf}, is just the series combination of C_{ox} and C_s:

$$C_{hf} = \left[\frac{1}{C_{ox}} + \frac{1}{C_s}\right]^{-1} \tag{3.9}$$

In this way, C_{hf} provides us with measurement of C_s in depletion as long as the measurement frequency is high enough to avoid contribution from D_{it}. Conversely, the measured quasi-static capacitance, C_{qs}, has the additional contribution from interface traps with capacitance C_{it}[100, 121]. Here, C_{ox} is in series with the parallel combination C_s and C_{it} as shown in Fig. 3.6 and we can write C_{qs} as:

$$C_{qs} = \left[\frac{1}{C_{ox}} + \frac{1}{C_s + C_{it}}\right]^{-1} \tag{3.10}$$

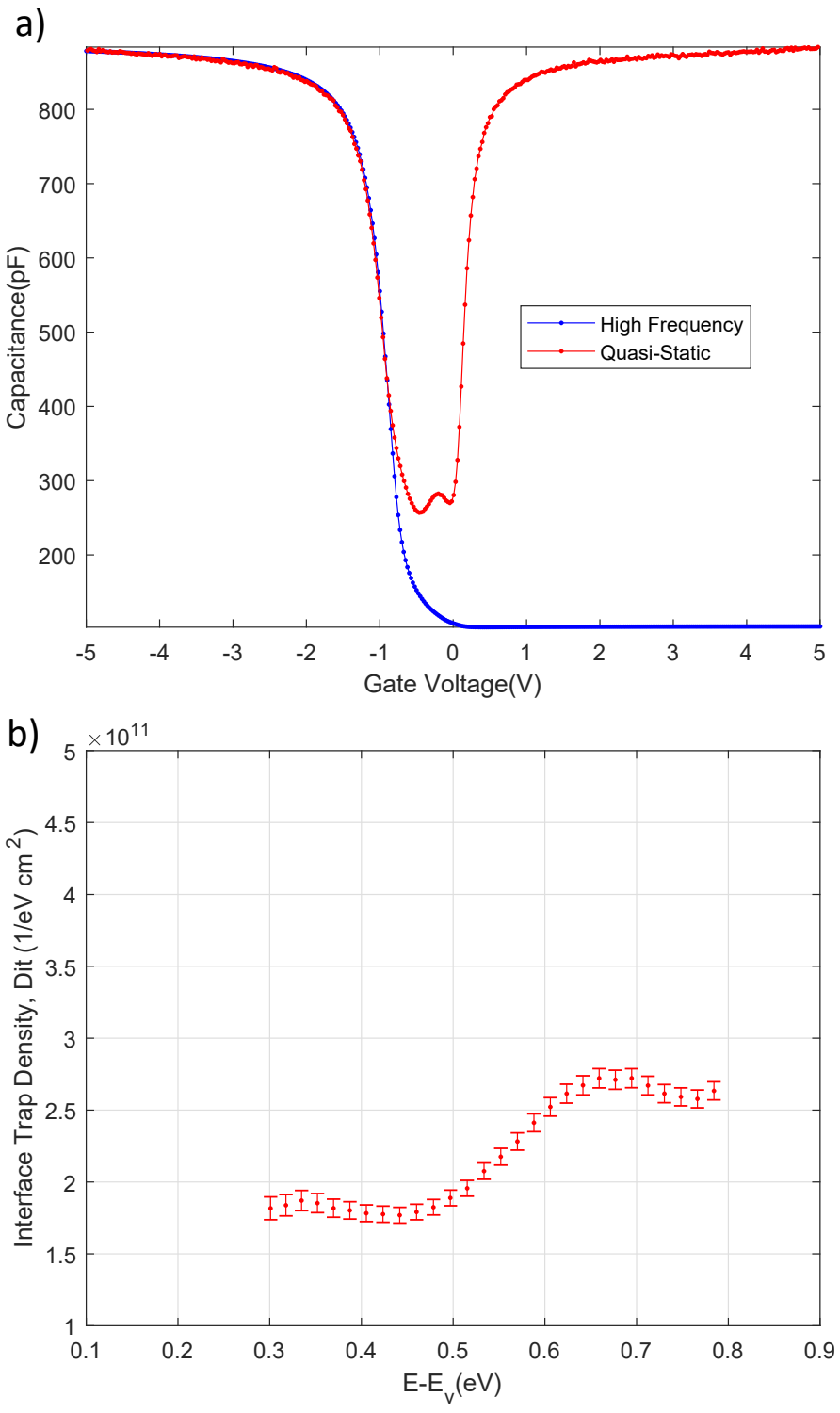

Figure 3.9: Example of a HFCV and QSCV measurements for a Al gated MOSCAP with a high level of D_{it}: (a) Measured HFCV and QSCV data and (b) Plot of D_{it} calculated using Eq. 3.11, where the x-axis has been generated using the methods in Sect.3.2.6.

Using both Eqs. 3.9 and 3.10, we can solve for C_{it} in terms of the measurable quantities C_{hf}, C_{qs}, and C_{ox}. Since $C_{it} = e^2 D_{it}$, D_{it} can be measured by:

$$D_{it} = \frac{C_{ox}}{e}\left[\frac{C_{QS}}{C_{ox} - C_{QS}} - \frac{C_{HF}}{C_{ox} - C_{HF}}\right] = \frac{C_{ox}}{e}\left[\frac{C_{ox}(C_{QS} - C_{HF})}{(C_{ox} - C_{QS})(C_{ox} - C_{HF})}\right] \quad (3.11)$$

The main assumption made in 3.11 is that we have taken the HFCV data at a frequency high enough such that the interface traps are unable to charge or discharge and that the QSCV data is slow enough that all interface traps do respond. Typical minority carrier generation rates are around 0.1-10 s meaning that minority carriers can't usually respond to AC signals in excess of 100 Hz [96]. For the HFCV data in this thesis, we used an AC frequency of 1 MHz, but frequencies as low 10 kHz -20 kHz can also be sufficient for measuring D_{it} near midgap in silicon. Higher frequencies push the measurement of D_{it} further away from midgap towards flatband [101]. For the HFCV data, a flat minimum capacitance response in strong inversion indicates the frequency is high enough because it shows a lack of free carrier response and a stable depletion width. On the other hand, if the QSCV data recovers C_{ox} in strong inversion, we view the assumption that the ramp is able to charge and discharge appropriately as satisfied. Fig. 3.9 shows an example of the data typically extracted using the HFQS method. For the MOSCAP shown, the D_{it} is relatively large displaying significant changes in the QSCV compared to a sample with low D_{it}(see Fig 3.6), where D_{it} is roughly two orders of magnitude lower.

3.2.5 Measurement of D_{it}: Conductance method

The previous section discussed CV measurements to determine D_{it}. A disadvantage of that method is the time is takes to measure and the need for two separate measurements with different equipment. Here, we discuss an alternative method to extract D_{it} using the measured conductance of the MOS capacitor in depletion [108]. This method is more sensitive than the HFQS method described in Sect. 3.2.4, capable of measuring densities as low as 10^9 $eV^{-1}cm^{-2}$. This method makes use of the parallel capacitance-conductance model used when performing HFCV measurements to extract the interface trap density. Here, we assume that the capture and emission of carriers from traps is the dominant process contributing to the conductance component in depletion for the MOSCAP. Measurements in weak inversion are also possible, but the modeling is more complicated since the minority carrier response to the AC signal becomes a significant contribution to the measured conductance. Here we will consider three main cases: 1) traps at a single energy level, 2) a continuum of trap levels, and 3) a statistical model that accounts for fluctuations in the surface potential due the presence of charge defects. Fig. 3.10 shows the equivalent circuit diagram necessary for measuring D_{it} via the conductance method. As mentioned previously, we perform all high frequency measurements with an LCR meter, where the measured admittance is modeled as a parallel capacitance and conductance circuit, a diagram for which is shown in Fig. 3.10(a). To calculate D_{it}, we need to use a circuit model similar to Fig. 3.10(d) where $G_p(\omega)$ is the desired parallel conductance. We can write G_p in terms of C_{ox},the measured parallel capacitance

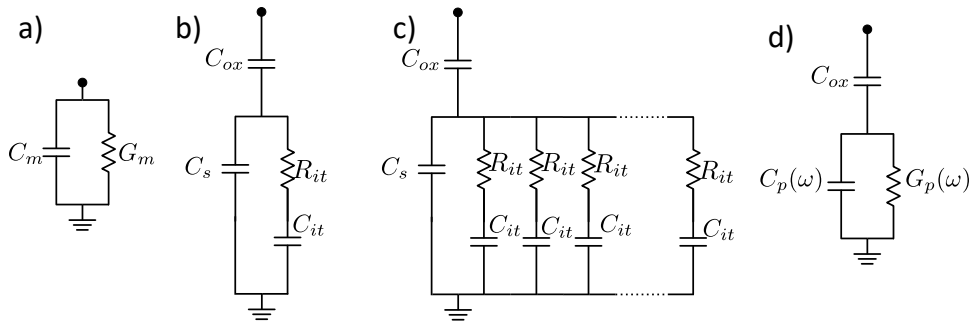

Figure 3.10: a) Circuit diagram for the measured admittance using a LCR meter. (b) Circuit diagram for the single-level trap model in depletion. C_{it} is the interface trap capacitance and R_{it} is the majority carrier capture resistance. Here, we have assumed that there is no generation-recombination via the bulk and D_{it} dominates the measured loss. (c) Circuit diagram for the continuum and statistical trap models. The traps act as a distributed network that is equivalent to a series of branches of single-level traps. (d) Simplified versions of (c) used for calculating D_{it}, where C_p and G_p are functions of V_g and frequency.

(C_m), the measured parallel conductance (G_m), and the angular frequency (ω) by comparing the measured admittances from Figs. 3.10(a) and (d) and solving for G_p [101]:

$$\frac{G_p}{\omega} = \frac{\omega^2 C_{ox} G_m}{G_m^2 + \omega^2 (C_{ox} - C_m)^2} \tag{3.12}$$

It is important to note that when calculating D_{it} via any conductance model G_p, calculated with Eq. 3.12, must be used and not G_m. With an understanding of how to turn measured data into the necessary quantities, we can now turn our attention to the different models for D_{it} conductance. First, we consider the simplest model of traps composed of a single energy level. This is the so-called single-level model and an equivalent circuit diagram is shown in Fig. 3.10(b). Using this equivalent circuit for the semiconductor admittance, which is the admittance of the circuits in

Fig. 3.10(b) without C_{ox}, yields:

$$Y_s = j\omega C_s + \frac{j\omega C_{it} G_{it}}{G_{it} + j\omega C_{it}} = j\omega \left[C_s + \frac{C_{it} G_{it}^2}{G_{it}^2 + \omega^2 C_{it}^2} \right] + \frac{\omega^2 C_{it}^2 G_{it}}{G_{it}^2 + \omega^2 C_{it}^2} \tag{3.13}$$

where C_{it} is the interface trap capacitance, $j = \sqrt{-1}$, and $G_{it} = {}^1/_{R_{it}}$ is the carrier capture conductance. Since we are concerned with the interface trap conductance, we need only focus on the real part of Eq. 3.13. Using Shockley-Read statistics [122], we can write C_{it} and G_{it} in terms of the density of interface states (D_{it}), the 3d carrier density (p), the majority carrier capture rate (c_p), and the Fermi distribution function (f_{FD}) [108]:

$$G_{it} = \frac{e^2 p D_{it} c_p (1 - f_{FD})}{k_b T} \tag{3.14}$$

$$C_{it} = \frac{e^2 D_{it} f_{FD} (1 - f_{FD})}{k_b T} \tag{3.15}$$

Substituting Eqs. 3.14 and 3.15 into the real part of Eq. 3.13 to obtain [100]:

$$\frac{G_p}{\omega} = \frac{e\omega\tau_{it} D_{it}}{1 + (\omega\tau_{it})^2} \tag{3.16}$$

Where we have defined a characteristic time, τ_{it}, associated interface traps as [101]:

$$\tau_{it} = \frac{C_{it}}{G_{it}} = \frac{f_{FD}}{p c_p} \tag{3.17}$$

Eq. 3.16 gives the conductance from interface traps, with density D_{it}, at a single energy in the silicon band gap with a single characteristic time, τ_{it}. In practice, the interface traps will have a broad distribution of closely spaced energy levels. This means discrimination of individual levels is not possible. To account for this we consider a continuum of states with energy in the silicon band gap, which suggests a circuit composed of a network of interface traps. Fig. 3.10(c) shows the equivalent circuit for this case. For this continuum of states, any interface states located within a few k_bT of the Fermi level can capture majority carriers. τ_{it} will now describe a mean characteristic time of this distributed network. The admittance in this case is similar to that in Eq. 3.13, but now we must integrate over the series of states [123, 124]:

$$Y_s = j\omega C_s + \int \frac{j\omega C_{it} G_{it}}{G_{it} + j\omega C_{it}} d\phi_s = j\omega C_s + j\omega \frac{e^2}{k_b T} \int \frac{D_{it} f_{FD}(1 - f_{FD})}{1 + j\omega \frac{f_{FD}}{pc_p}} d\phi_s \quad (3.18)$$

Where we have used Eqs. 3.14 and 3.15 to write the integral in terms of f_{FD}. In order to evaluate this integral, we will assume that D_{it} and p do not vary significantly over a range of a few k_bT in ϕ_s. Using a transformation of $f_{FD}(1 - f_{FD}) = \frac{k_b T}{e} \frac{df_{FD}}{d\phi_s}$, we can write Eq. 3.18 in terms of f_{FD} instead ϕ_s. Evaluating this integral yields [108]:

$$Y_s = j[\omega C_s + \frac{eD_{it}}{\tau_{it}} arctan(\omega \tau_{it})] + \frac{eD_{it}}{2\tau_{it}} ln[1 + \omega^2 \tau_{it}^2] \quad (3.19)$$

Where in this model we now have $\tau_{it} = \frac{1}{pc_p}$. Similar to the derivation of Eq. 3.16, we only need to focus on the real part of Eq. 3.19 to obtain the conductance for a

continuum of states is given by [108]:

$$\frac{G_p}{\omega} = \frac{eD_{it}}{2\omega\tau_{it}} ln(1 + (\omega\tau_{it})^2) \tag{3.20}$$

In comparing the Eqs. 3.16 and 3.20 for the conductance, we can see that both will have peaks as a function of ω. The major difference between the models is the dispersion in τ_{it} in the continuum model. We can still easily compare the two models at the $\frac{G_p}{\omega}$ peak. For the single level model (Eq. 3.16), the peak occurs at $\omega = \frac{1}{\tau_{it}}$. For the continuum model (Eq. 3.20), the peak occurs at $\omega = \frac{1.98}{\tau_{it}}$. This leads to two different expressions for D_{it} derived from the maximum of $\frac{G_p}{\omega}$: $D_{it} = \frac{2}{e}(\frac{G_p}{\omega})_{max}$ and $D_{it} = \frac{2.5}{e}(\frac{G_p}{\omega})_{max}$ for the single level and continuum models respectively. The experimental determination of D_{it} from conductance is performed by measuring G_p as a function of frequency at given gate voltage in depletion.

When compared with experimentally measured G_p vs ω data, Eqs. 3.16 and 3.20 are unable to accurately fit the shape of the curve at any bias range as shown in Fig. 3.11. The experimental width of the curve is wider than predicted [108]. The disagreement is due to presence of non-uniformly distributed charge in the oxide, the ϕ_s being probed at a given V_g will vary across the area of the MOSCAP. This means that across the MOSCAP a broader range of trap energies and characteristic times are accessible than predicted by the continuum model, because ϕ_s samples more than one value. Usually the distribution of surface potentials is assumed to

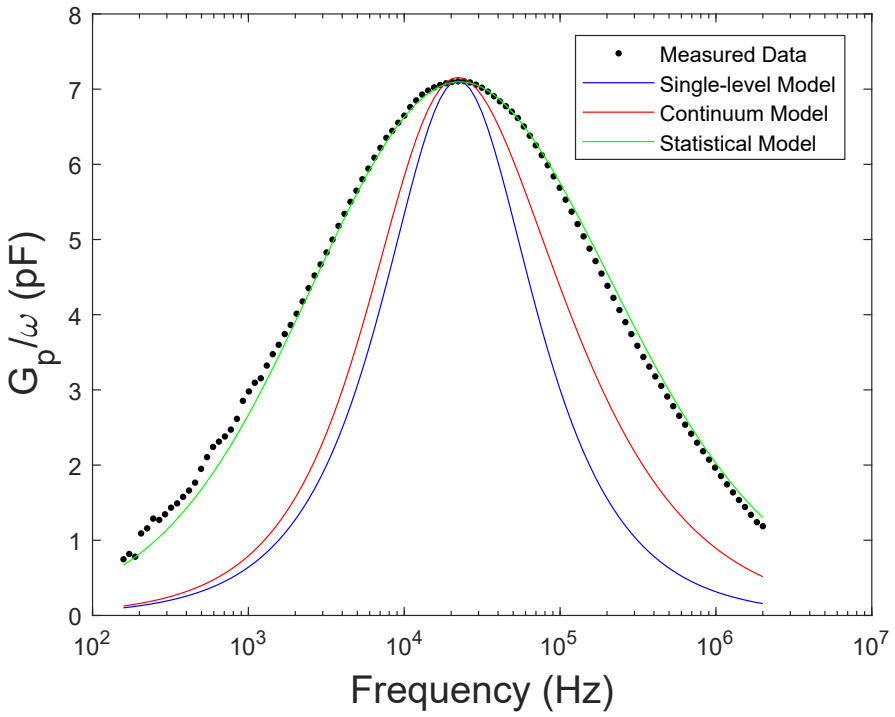

Figure 3.11: Examples of the three models for measured conductance response from D_{it} compared with measured data from an Al-gated MOSCAP. The statistical model is the only one capable of reproducing the width of the experimental curve. For the data shown here, D_{it} values of $(2.31\pm0.11)\times10^{10}$ cm^{-2}, $(2.88\pm0.11)\times10^{10}$ cm^{-2}, and $(4.29\pm0.21)\times10^{10}$ cm^{-2} for the single-level, continuum, and statistical models respectively. There is roughly a 30-40 % error in D_{it} by neglecting the surface potential fluctuations used in the statistical model.

be Gaussian [108]:

$$P(\phi_s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} exp[-\frac{(\phi_s - <\phi_s>)^2}{2\sigma_s^2}] \tag{3.21}$$

Where $<\phi_s>$ is the average surface potential and σ_s^2 is the variance of the surface potential fluctuations and $P(\phi_s)$ is the probability of obtaining ϕ_s. The distributed network in this model is similar to that used in the derivation of the continuum

model, with the exception that admittance discussed above is now replaced by an average value corresponding to the value of $< \phi_s >$. Multiplying Eq. 3.20 by the probability defined in Eq. 3.21 and integrating over ϕ_s, we can write trap conductance in depletion as:

$$\frac{G_p}{\omega} = \frac{eD_{it}}{2} \int_{-\infty}^{\infty} \frac{1}{\omega\tau_{it}} ln(1 + (\omega\tau_{it})^2) P(\phi_s) d\phi_s \tag{3.22}$$

At the peak in $\frac{G_p}{\omega}$ data, Eq. 3.22 can be rewritten to solve for D_{it} resulting in:

$$D_{it} = \frac{1}{ef_D(\sigma_s)} (\frac{G_p}{\omega})_{max} \tag{3.23}$$

Where $f_D(\sigma_s)$ is given by:

$$F_d(\sigma_s) = \frac{1}{2\xi_p\sqrt{2\pi\sigma_s^2}} \int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2}) exp(-\eta) ln(1 + \xi_p^2 exp(2\eta)) d\eta \tag{3.24}$$

where $\eta = \phi_s - < \phi_s >$ and $\xi_p = \omega\tau$. In order to calculate D_{it}, we need to determine the values of both σ_s and ξ in order to calculate F_d. Using the condition from [101] that at the peak:

$$\frac{d}{d\xi} \left(\frac{G_p}{\omega} \right)_{f_p} = 0 \tag{3.25}$$

From Eq. 3.22, this condition yields an integral that allows ξ to be determined as a function of σ_s:

$$\int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2}) exp(-\eta) (\frac{2\xi_p^2 exp(2\eta)}{1 + \xi_p^2 exp(2\eta)} - ln(1 + \xi_p^2 exp(2\eta)) d\eta = 0 \tag{3.26}$$

In practice without a value σ_s, evaluating the integrals involved in Eqs. 3.21 and 3.22 is cumbersome when trying to fit measured data, however we follow the procedures outlined in [101, 124] to determine both σ_s and D_{it} by exploiting the peaked nature of $\frac{G_p}{\omega}$. Since we need to solve for two unknowns (σ_s and ξ_p) in order to calculate D_{it} we require two simultaneous equations. Here, we can use Eq. 3.22 taken at two different values of ω. For instance, at the frequency peak (f_p) and at multiplies of that frequency (nf_p and $\frac{f_p}{n}$), the ratio of $\frac{G_p}{\omega}$ at these points relative to the peak frequency can be written as:

$$\frac{(G_p/\omega)_{f_p/n}}{(G_p/\omega)_{f_p}} = \frac{1}{n}\frac{\int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2})exp(-\eta)ln(^{1+\xi_p^2 exp(2\eta)}/_{n^2})d\eta}{\int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2})exp(-\eta)ln(1+\xi_p^2 exp(2\eta))d\eta} \tag{3.27}$$

For the point at $\frac{f_p}{n}$ and similarly from the point at nf_p

$$\frac{(G_p/\omega)_{nf_p}}{(G_p/\omega)_{f_p}} = n\frac{\int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2})exp(-\eta)ln(1+\xi_p^2 exp(2\eta)n^2)d\eta}{\int_{-\infty}^{\infty} exp(-\frac{\eta^2}{2\sigma_s^2})exp(-\eta)ln(1+\xi_p^2 exp(2\eta))d\eta} \tag{3.28}$$

Where in both Eqs. 3.27 and 3.28 $\eta = \phi_s - <\phi_s>$ and $\xi = \omega\tau$. We determine σ_s by: 1) picking off the values of $\frac{(G_p/\omega)_{f_p/n}}{(G_p/\omega)_{f_p}}$ and $\frac{(G_p/\omega)_{nf_p}}{(G_p/\omega)_{f_p}}$ from the experimental data with $n = 5$, 2) use Eq. 3.26 to give ξ as a function of σ_s, and 3) numerically solve for the value σ_s using fsolve in Matlab. With a value of σ_s determined, we can now turn back to Eqs. 3.24 and 3.23 and calculate values for F_d and D_{it} respectively. It is important to note at low values of σ_s it is possible that Eqs. 3.27 and 3.28 do not provide a single value of σ_s [101, 108] since $\frac{G_p}{\omega}$ is not symmetric about its maximum. For the purposes of the data in this thesis we have taken the average

value obtained from the two methods and have propagated that uncertainty in σ_s into the uncertainty in D_{it}.

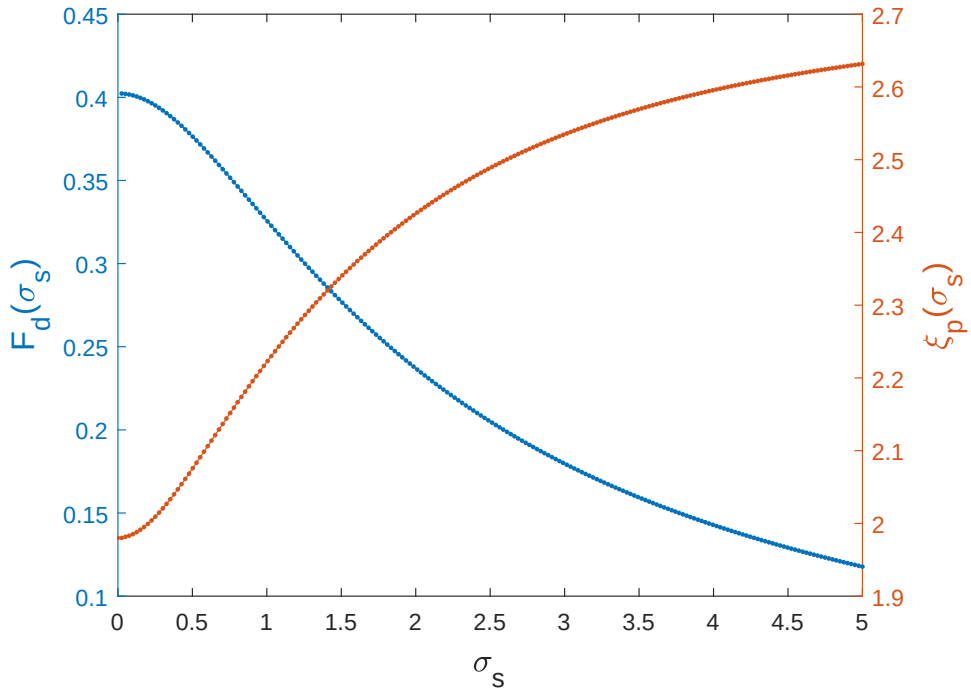

Figure 3.12: Plots of the statistical model factors F_d and ξ_p from Eqs. 3.24 and 3.26 respectively as a function of σ_s.

As a consistency check of the preceding equations, we can set $\sigma_s = 0$ and then evaluate Eq. 3.26. This yields a value of $\xi_p = 1.98$ as shown in Fig. 3.12, which is the same as the value of $\omega\tau$ at the conductance peak found in the continuum model without any surface potential fluctuations.

3.2.6 Converting Gate Voltage to Energy

The D_{it} data extracted following the methods in Sect. 3.2.4 and 3.2.5 is in the form of D_{it} vs V_g but it is desirable to convert the V_g scale into energy scale

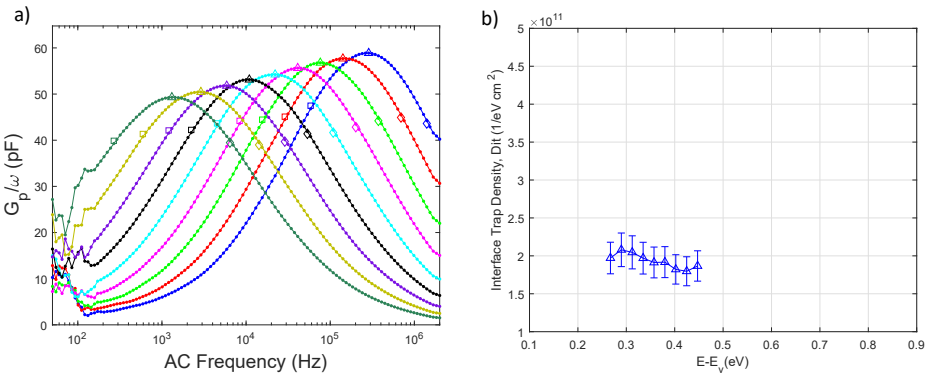

Figure 3.13: Example of conductance measurements for a Al gated MOSCAP with a high level of D_{it}: (a) $\frac{G_p}{\omega}$ data corrected for series resistance effects following the methods of Sect. 3.2.8. The triangle data points indicate the peak position (f_p) while the squares shows the points of $5f_p$ and $\frac{f_p}{5}$ used to compute σ_s. For this data set, the extracted parameters for the statistical model are: $\sigma_s = 1.693 \pm 0.181$, $\xi_p = 2.375 \pm 0.013$, and $f_D = 0.261 \pm 0.015$ (b) Plot of D_{it} calculated from the data in (a), where the x-axis has been generated using the methods in Sect.3.2.6.

that relates directly to the trap energy in the silicon bandgap. The first step in this conversion is to convert V_g into ϕ_s using the Berglund method described in Ref. [125]. The QSCV data can be used to calculate ϕ_s as a function of the applied gate voltage:

$$\phi_s = \int_{V_{fb}}^{V_{G2}} (1 - \frac{C_{QS}}{C_{Ox}}) dV_G \tag{3.29}$$

Where we have chosen to use V_{fb} as one of the limits of integration. This means that the any integration constant is equal to zero since the surface potential should be zero at the flatband voltage. Using this method, there will need to be two separate integrations performed: one from accumulation to flatband and one from flatband into inversion. Fig. 3.14 gives an example of the typical ϕ_s vs V_g data from MOSCAPs in this thesis. The blue/red coloring of the curve delineates the two dif-

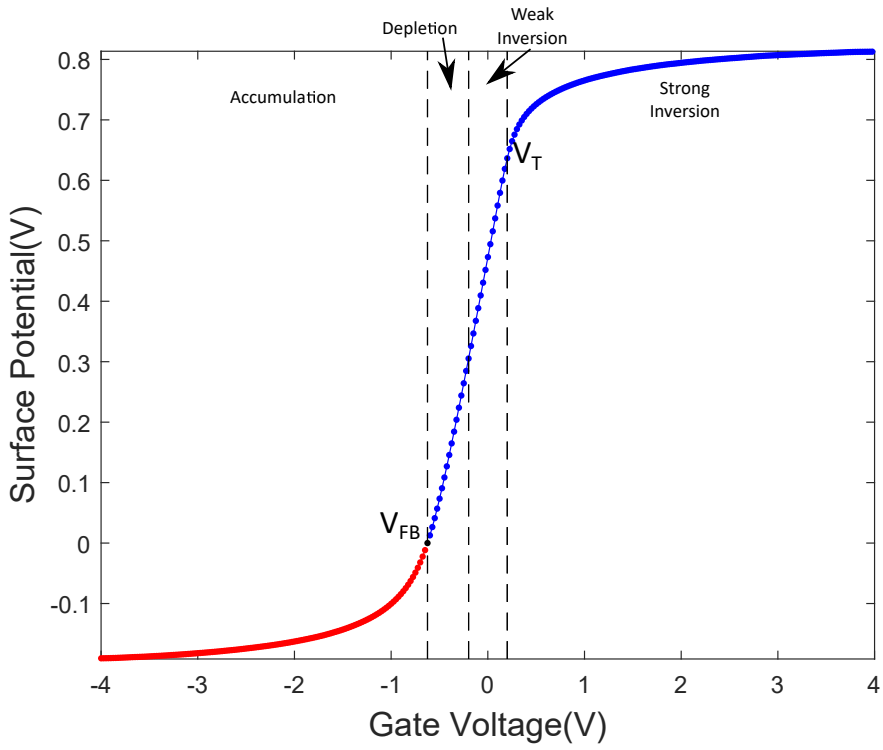

Figure 3.14: Example of the surface potential vs gate voltage relationship obtained from equation 3.29. The colors in plot show the splitting of the integral into regions with the blue data for $V_g < V_{fb}$ and the red data for $V_g > V_{fb}$. The dashed lines denote the different MOSCAP charge density regimes.

ferent integrations performed from the total QSCV curve. Note at large magnitudes of V_g ($|V_g| > 2V$), the changes in ϕ_s become asymptotically flat relative to changes in V_g signifying saturation that occurs in the accumulation and inversion regions. In depletion, ϕ_s shows an approximately linear relationship with V_g. Importantly, this ϕ_s vs V_g data doesn't relate the measured D_{it} to the energy position of the traps in the silicon band gap by itself. To achieve that, it is necessary to relate a given value of ϕ_s to either the silicon valence or conduction band edge. In this thesis, we performed the majority of CV measurements using MOSCAPs fabricated on p-type substrates, it is more convenient to relate the trap energy levels to valence band edge, but similar analysis using midgap or the conduction band edge as the reference point could also be done. In this case, the depletion region corresponds to energies on the valence band side of the band gap. On p-type substrates, traps on the conduction band side are not as easily measured. Typically, to measure D_{it} over a more complete range in the bandgap, measurements on n-type substrates with similar doping concentration are used. The surface potential, ϕ_s, is related to the trap energy level from the valence band, $E - E_V$ for a p-type silicon substrate by [101]:

$$E - E_v = \phi_s + \frac{k_b T}{e} \ln \frac{N_v}{N_a} \tag{3.30}$$

N_v is the effective density of states in the valence band given by $N_v = 2(\frac{2\pi m_h^* k_b T}{h^2})^{\frac{3}{2}} = 3.19 \times 10^{19}\ cm^{-2}$ [97] where h is Planck's Constant and $m_h^* = 0.386 m_e$ is the effective mass (density of states) for holes in silicon [96]. Eq. 3.30 allows for the D_{it} values extracted in either Sect. 3.2.4 or 3.2.5 to be plotted as a function of trap energy, E,

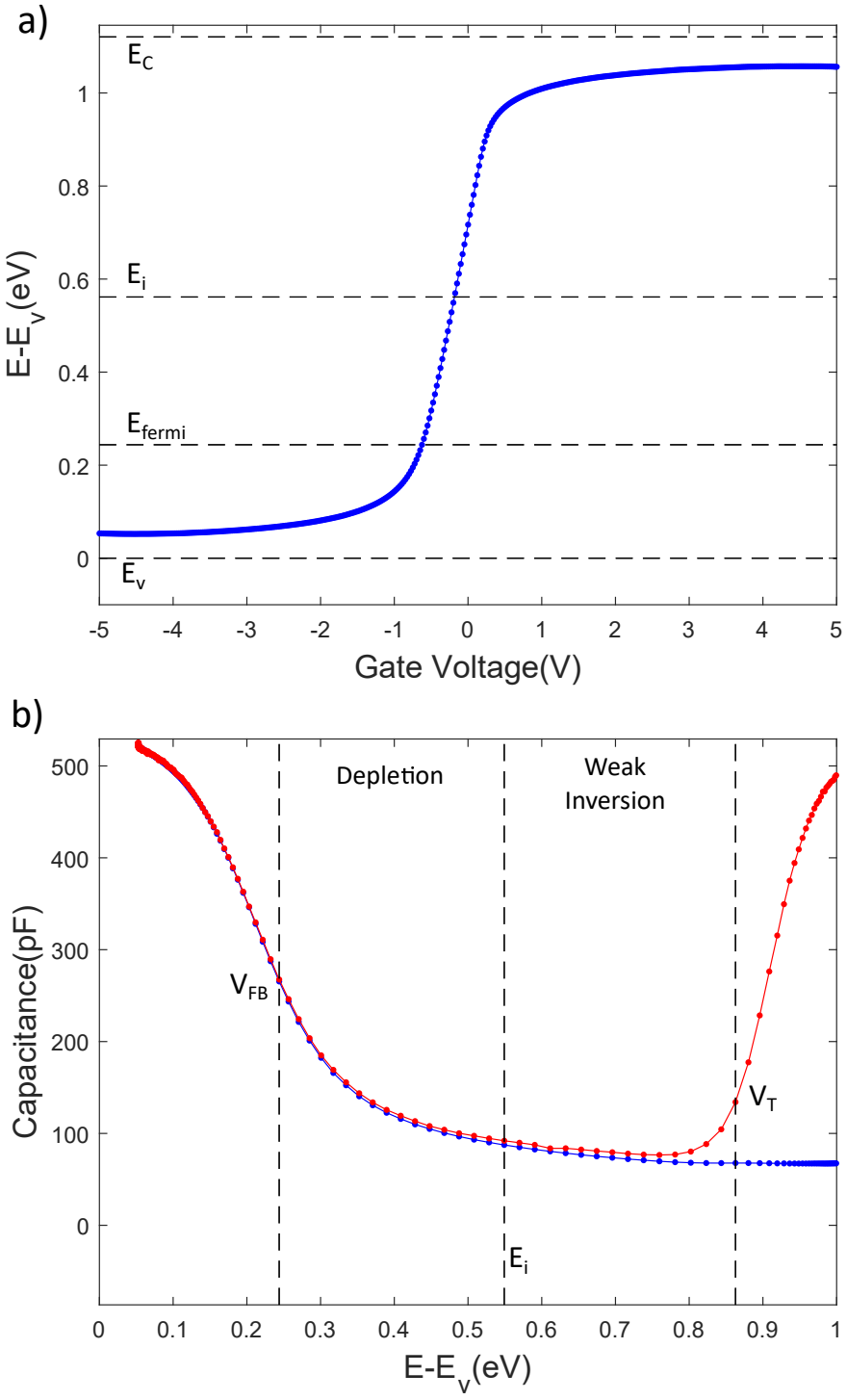

Figure 3.15: (a) Example of the $E - E_v$ vs Gate Voltage relationship obtained from equation 3.30. The dashes lines are used to denote the band edges and Fermi level relative to V_g. (b) Plot of the CV data from Fig. 3.6 plotted as a function of $E - E_v$ instead of V_g.

in the silicon band gap. Fig. 3.15 (a) shows the relationship between $E - E_v$ and V_g using the data from Fig. 3.14 with the above equation. Using this value of $E - E_v$, we can then plot any CV data on an energy axis instead V_g. This is shown in Fig. 3.15 (b) using CV data from Fig. 3.6. Here, we can see that depletion region of MOS operation extends from roughly 0.244 eV to 0.549 eV in $E - E_v$ and weak inversion goes from 0.549 eV to 0.863 eV. Here, the dividing point of $E - E_v = 0.549\ eV$ is actually equivalent to E_i.

Additionally in weak inversion, the assumption of D_{it} dominated conductance response is no longer valid. In this region, the conductance is a combination of the D_{it} response and the conductance due the minority carriers [101]. Given the minimum frequencies used in this thesis of around 100 Hz, D_{it} measurements are limited to $E - E_v < 0.49\ eV$. Similarly, the HFQS method (Sect. 3.2.4) the minority carrier response limits the range by causing the separation between the QS and HF data. The valid limit can be determined by comparing a QSCV curve to the measured HFCV curve, shown in Fig. 3.15(b). Here, we can see that the QSCV and HFCV naturally begin to separate for $E - E_v > 0.8\ eV$. This signifies when the minority carrier response is dominating the QSCV data. For the MOSCAPs in this thesis, we choose to limit D_{it} extracted from the HFQS method to $E - E_v < 0.75\ eV$.

3.2.7 Measurement of Q_m: Triangular voltage sweep (TVS)

We extract Q_m from the difference between the QSCV and HFCV curves at a sufficiently high temperature. This is known as the triangular voltage sweep (TVS)

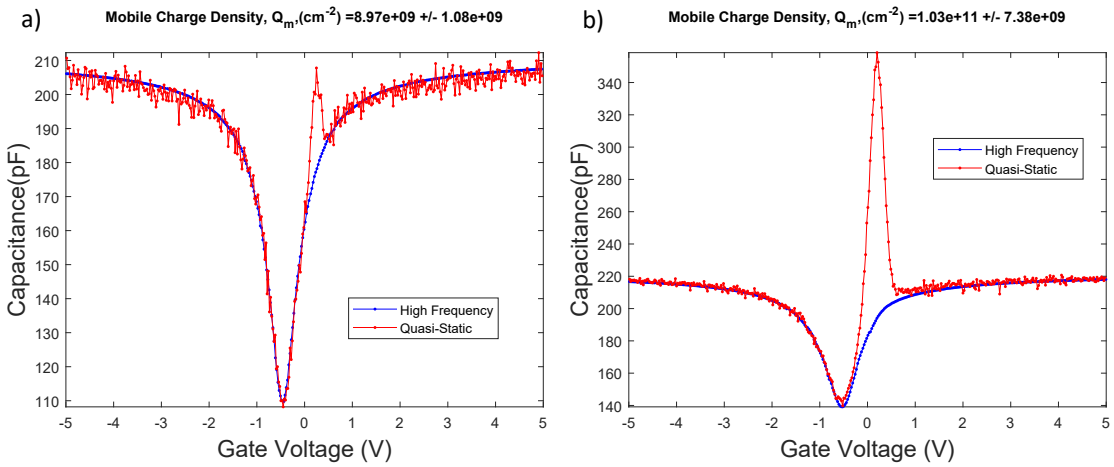

Figure 3.16: Examples of a TVS of the mobile ion density in thermally grown SiO_2 for: (a) An oxide with a low contamination level and (b) A heavily contaminated oxide.

method [110, 126]. Here, sufficiently high refers to the minority carrier response displayed when the high frequency curve recovers the oxide capacitance in inversion, i.e. a temperature high enough that the minority carrier density is able to respond to the AC frequency. In this case, the HF and QS curves should be identical in an ideal MOS system. The TVS method follows a similar idea to that of HFQS method for measuring D_{it}, where the expectation is that the mobile ions are able to effectively follow the slow dc ramp in the QS measurement and unable to move easily in response to the high frequency voltage modulation even at elevated temperatures. In this way, the QS curve, which should be identical to the HF curve in the ideal case, will measure an additional capacitive component from the mobile ions moving through the oxide:

$$Q_M = \frac{1}{eA}\int_{V_{G1}}^{V_{G2}} (C_{QS} - C_{HF}) dV_G \tag{3.31}$$

The calculation of Q_m from TVS data requires numerical integration of the data in order to calculate the difference in area between the QS and HF curves. The TVS method additionally has an advantage that it should be sensitive to the type of ionic contaminant in the oxide [100], where ions with different mobility will show peaks at different gate voltages along the QSCV curve. To satisfy normal operating requirements, Q_m should typically be below 10^{10} cm^{-2}. The TVS data shown in Fig. 3.16a are for an oxide with a contamination level below 10^{10} cm^{-2} measurable by TVS, while Fig. 3.16b shows TVS data for a heavily contaminated sample with Q_m on the order of 10^{12} cm^{-2}. We attribute the difference between these two data sets to contamination within the tube furnace used for the oxide growth. For data in these figures, there is only one peak visible in the QSCV data suggesting that the contamination is due to a single type of ion in the oxide, but from this data alone we cannot identify which ion specifically.

3.2.8 Series Resistance (R_{series}) Corrections

$$R_{series} = \frac{G_{meas,acc}}{G_{meas,acc}^2 + \omega^2 C_{meas,acc}^2} \tag{3.32}$$

For a given set of measured capacitance, $C_{meas,acc}$, and conductance, $G_{meas,acc}$, in strong accumulation (and using a parallel C/G impedance model for the measurement) at an AC frequency of $f = \frac{\omega}{2\pi}$, the value of the series resistance can be estimated by equation 3.32. This value for the stray series resistance component can be then used to calculate corrected values of the capacitance and conductance,

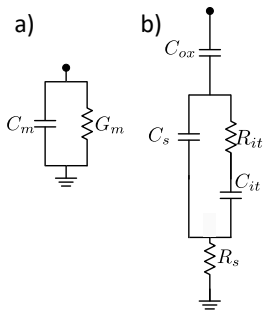

Figure 3.17: Circuit models used in the series resistance corrections. (a) Measured parallel combination of capacitance and conductance by the LCR meter. (b) Circuit model for D_{it} measurements including the parasitic series resistance, R_s, from the silicon wafer back side contact. The circuit model shown here is for a single trap level, but an analogous circuit can also be drawn for the statistical trap model discussed earlier.

C_c and G_c respectively.

$$C_{corrected} = \frac{(G_{meas}^2 + \omega^2 C_{meas}^2) C_{meas}}{[G_{meas} - (G_{meas}^2 + \omega^2 C_{meas}^2) R_{series}]^2 + \omega^2 C_{meas}^2} \tag{3.33}$$

$$G_{corrected} = \frac{(G_{meas}^2 + \omega^2 C_{meas}^2)[G_{meas} - (G_{meas}^2 + \omega^2 C_{meas}^2) R_{series}]}{[G_{meas} - (G_{meas}^2 + \omega^2 C_{meas}^2) R_{series}]^2 + \omega^2 C_{meas}^2} \tag{3.34}$$

These values for C_c and G_c are necessary for the analysis described in the preceding sections when using any HFCV data. In particular, for the conductance method, we apply this correction across the measured frequency range. For instance, the conductance component for D_{it} in Eq. 3.22 requires performing the series resistance correction so that D_{it} conductance can be accurately measured. Fig. 3.18(a) shows an example of the measured frequency response in accumulation for a Al gate MOSCAP with Ti/Au back contact. We use this data as the input for Eq. 3.32 in order to correct the raw data using Eqs. 3.33 and 3.34. Fig. 3.18(b) compares

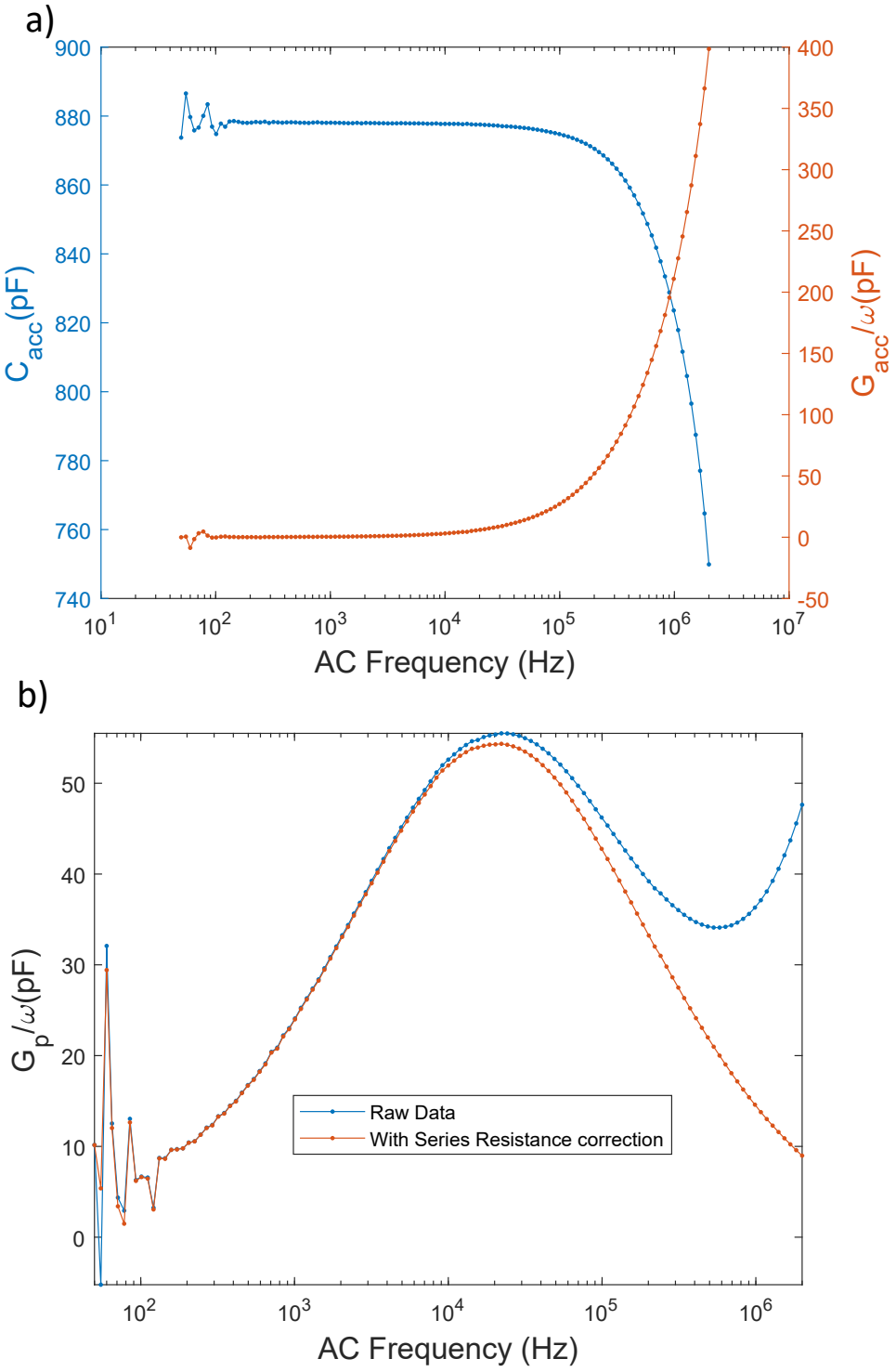

Figure 3.18: (a) Plot of the measured frequency dependence of the capacitance and conductance in accumulation used for correcting for series resistance effects. (b) Example plot showing the effect of performing the series resistance correction on data used for the conductance method from Sect. 3.2.5.

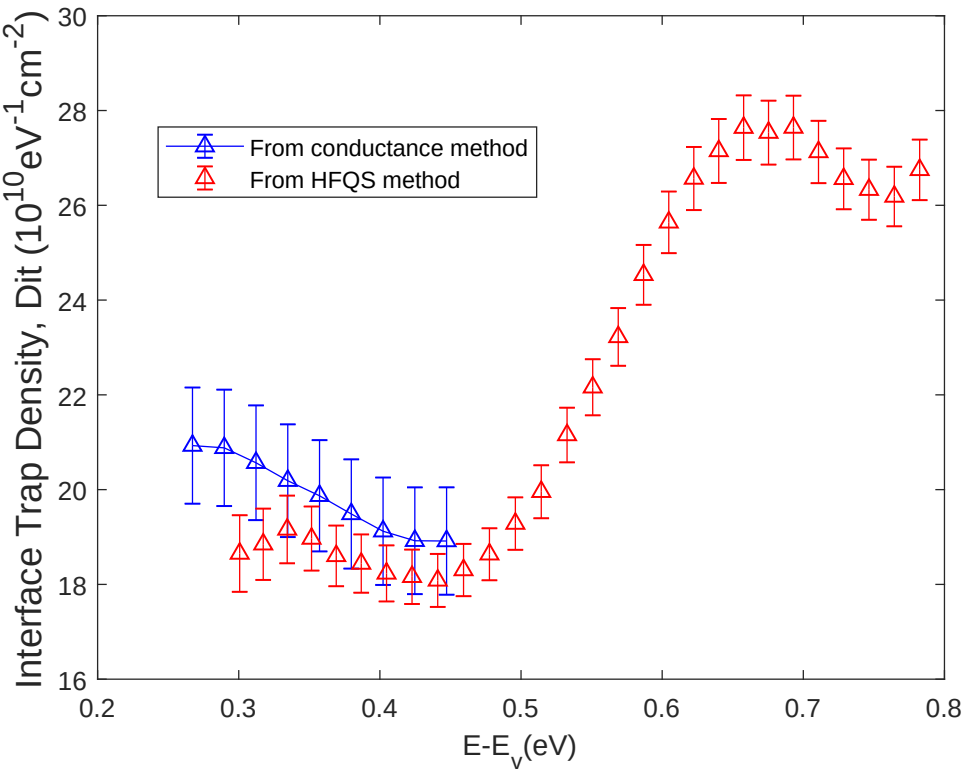

Figure 3.19: Comparison of D_{it} extracted from the HFQS (red) and conductance methods (blue) for the same Al-gated MOSCAP from Figures 3.9 and 3.13 respectively. Both methods produce similar values for D_{it} with the only major differences arising from the more limited bandgap range for the conductance method due the lower frequency limit of around 100 Hz.

an example of the raw and corrected data for calculating D_{it} using the conductance method in Sect. 3.2.5. We show here that even with a metallized back contact it is still necessary to correct for the series resistance, especially as the peak in conductance moves to higher frequencies.

3.2.8.1 Comparison of Dit measured via Conductance and HF-QS methods

In the preceding sections of this chapter, we have discussed two methods for measuring D_{it} using a MOS capacitor: the HFQS method (Sect. 3.2.4) and the conductance method (Sect. 3.2.5). Fig. 3.19 shows a comparison of D_{it} extracted via these methods for the same Al gated device. As mentioned previously, one of the major differences between these methods is the range of the bandgap that we can measure D_{it}. Here, the conductance method is limited by the low end of our measurement frequency range of 100 Hz, which restricts the measurement range to less than midgap. This limit is different for the HFQS data, which probes some portion of the conduction band side. In the region on the valence band side where both methods are valid, they show reasonable agreement in the extracted D_{it}. Additionally, our use of the statistical D_{it} conductance model is validated based on Fig. 3.19, as this agreement would not be achieved using the single level or pure continuum models.

Chapter 4: Measurement of strain using MOS tunnel junctions

As discussed in chapter 1, strain induced from the gates has been suggested as a source of unintentional QDs [73, 74]. The strain landscape in a silicon QD depends heavily on the operating conditions and fabrication process. This suggests the most applicable measurement of strain is one that can be performed under the same operating conditions ($T \leq 10$ K) and adhering to the same fabrication constraints. The inhomogeneous strain is typically $\frac{\Delta x}{x} \approx 10^{-4}$ and varies over 10's of nanometers. It is challenging to find a method for measuring strain that satisfies the necessary sensitivity and spatial resolution along with low temperature capability. For instance, transmission electron microscope (TEM)-based methods [127] can meet the spatial and sensitivity requirements but are typically not performed at low temperature, destroy the sample, and may alter the strain through sample preparation [128]. High resolution electron back-scatter detection [128, 129] is a non-destructive method that could be used to meet the spatial and sensitivity requirements, but similar to TEM-based techniques, is not typically performed at cryogenic temperatures. X-ray diffraction (XRD) [130] and Raman [131] techniques can perform a non-destructive measurement but have difficulty achieving the necessary spatial resolution while also not approaching cryogenic temperatures. Electrical measurements of strain are

advantageous because we can achieve the necessary resolution at cryogenic temperatures. Piezo resistive sensors [132, 133] have been demonstrated to meet both of these requirements but only in micron scale devices. Ref [134] measured strain via a shift in the electron spin resonance frequency of Bi donors at T=20 mK with a sensitivity of 10^{-7} but the results cannot be easily translated to gate-defined QDs.

In this chapter, we will use the groundwork laid out in chapter 2 for simulating stress and strain in the MOS system in COMSOL to develop a new device to electrically measure strain under a similar fabrication process and operating conditions to that seen in QD devices. The goal of this chapter is to present a comparison between simulations and measurements of the effect of strain on the tunnel barrier height of MOS tunnel junction devices.

Our strategy is to first perform transport measurements on separate tunnel junction devices made with aluminum and titanium gates. A tunnel barrier is formed in the gap between the gates where, for a range of gate voltages, inversion layers form at the Si-SiO_2 interface under the gates but not in the gap between them (see Fig. 4.1). We then fit the conductance as a function of bias voltage and extract the barrier height, ϕ_{tot}, as a function of gate voltage. When properly controlled, the difference between these barrier heights gives a measure of the change in strain due to the change in gate material in otherwise identical devices. We further characterize the metal films by measuring the coefficient of thermal expansion, α, at room temperature. Then, using the measured geometry for the device, we simulate the α-induced strain difference using bulk values of α, and our experimentally measured values of α. We find that our tunnel junction measurement of the strain difference

agrees with simulations provided we use our experimentally measured values of α. Some portions of the work in this chapter are previously published in the Journal of Applied Physics in Ref. [135].

4.1 Impact of strain on a tunnel junction

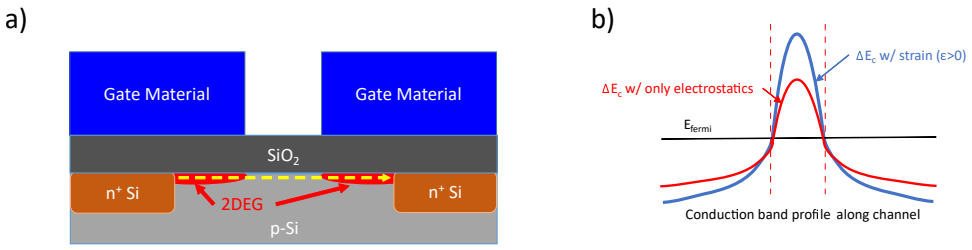

Figure 4.1: (a) Schematic cross-section of the metal-oxide-semiconductor (MOS) tunnel barriers used in this work. The barrier is formed by modulation of the conduction band in the gap between the gates. (b) Sketch of the expected modulation of the conduction band (ΔE_c) for the cases of electrostatics only (blue) and the combination of electrostatics and strain from the gates (red). The dashed line indicates where the edges of the gate lie.

To study the impact of gate-induced strain, we use a tunnel junction (TJ) device defined by two 2DEGS separated by a narrow potential barrier. We show a model of this device in Fig. 4.1(a). The expectation for this device is that strain will modify the intrinsic electrostatic barrier as shown in Fig. 4.1(b). Here we can predict the modulation of the barrier from the deformation potential equations given in chapter 2, where tensile (compressive) strains in the gap will raise (lower) the height of the barrier provided the barrier is narrow enough.

One of the major advantages of this TJ device is that the fabrication closely mirrors that used to form QDs, ensuring the results will be applicable to QDs.

The strain induced by the gates will modify the built-in electrostatic barrier height and shape, which we will extract from fits to the four-terminal conductance vs bias voltage measurements. This requires each device to have six contacts: two separate gate contacts and four ohmic contacts. The relative simplicity of this device design also allows us a large amount of freedom in terms of the fabrication process. Since the test device is composed of a single layer of gates, we can use gate materials that do not have native dielectrics if we can tailor a suitable deposition and the e-beam lithography patterning process for it. Later in this work, we will discuss alternative device designs for measuring strain. These methods have the disadvantage that they require overlapping gate structures separated by an isolation oxide. This isolation oxide presents a challenge in terms of the fabrication. We would need a deposited oxide compatible with many different gate materials, and in the strain landscape of the device, where it is difficult to characterize the mechanical properties of such an oxide using the methods described in chapter 2. Ultimately, for any of our designs, the details of the fabrication process are important, as we need to minimize device-to-device variations to be able to compare tunnel junctions across different material systems. Therefore, we have decided to design the fabrication process for our devices presented in this chapter to be as similar as possible. The method still allows a significant amount of freedom to modify this process for future devices.

The strain induced modulation of the tunnel barrier in our test devices can be simulated using finite-element modeling (FEM) in COMSOL, which we discussed in chapter 2. Using the linear elastic model in COMSOL, we can simulate the strain in bulk silicon induced from cooling the device to 2 K and the effects of the intrinsic

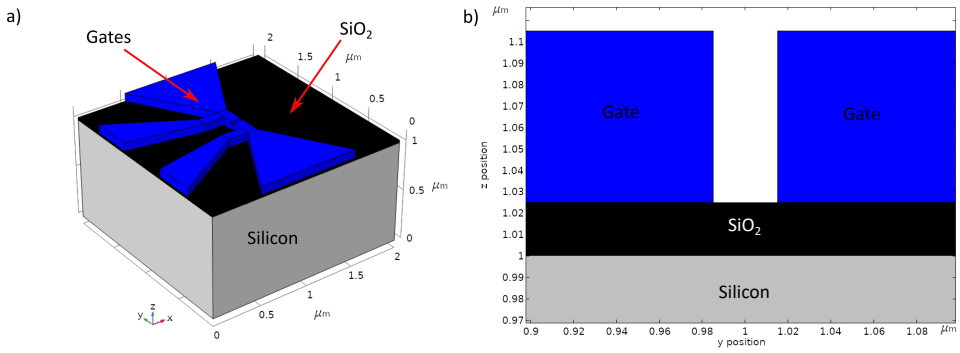

Figure 4.2: (a) Image of the 3d tunnel junction device model used for simulating strain in COMSOL. The device is shown here with a 25 nm thick gate oxide, 30 nm wide tunnel gap, and 100 nm wide gates. (b) 2d slice of the model in (a) along the center of the tunnel junction.

stress of material stack that makes the device. We use calculated strains in the 3-dimensional device structure with the deformation potential to determine the local modulation in the silicon band energies. Using the material parameters shown in chapter 2, we simulate the resulting inhomogeneous strain at cryogenic temperatures in our tunnel junctions. Fig. 4.2(a) shows an image of the 3d model used in our COMSOL simulations. For the device model in Fig. 4.2(a), we have fixed the bottom of the silicon block with a zero displacement boundary condition such that all strain components are zero at that interface for all temperatures. The crystal axes are set using the orthotropic elasticity matrix for silicon applicable to silicon (100) wafers used in our devices, where $x = [110]$, $y = [\bar{1}10]$,$z = [001]$. For these simulations, we have used the room temperature value of α for all materials. For the materials used in this work, α varies with temperature and goes to zero at cryogenic temperatures [79]. Therefore, we expect that the use of the room temperature α in this simulations means that we will overestimate the change in barrier height.

This choice is made because measurements yield a room temperature value of α and we do not have a measure of the temperature dependence down to cryogenic temperatures. For the model, we use the design file from the e-beam lithography to create the gate layer for the TJ device. Fig. 4.2(b) is a 2d cut along the center of the 3d TJ model and shows the gap region where the barrier will be formed.

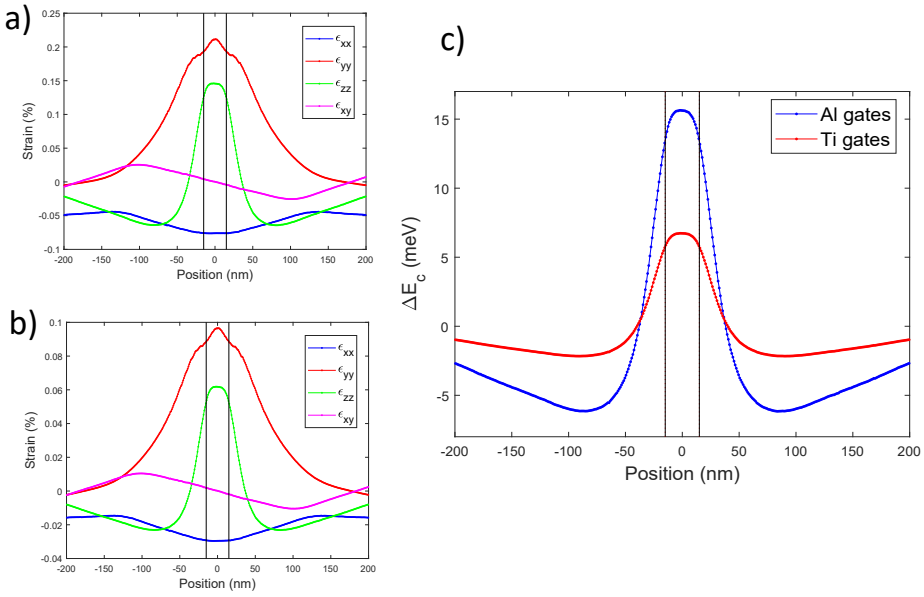

Figure 4.3: Plots of the α-induced strain from cooling a TJ device from 300 K down to 2 K for: (a) Al gates and (b) Ti gates. (c) Modulation of the conduction band (ΔE_c) calculated using the deformation potentials given in chapter 2. The solid black lines in all plots indicate the edges of the gates in the device.

We show the results of the COMSOL simulation in Fig. 4.3 for Al and Ti gates. The results shown here only consider the effects of thermal strain from the CTE mismatch of the different materials; we have not introduced any intrinsic film stress. Additionally, we have assumed that the silicon dioxide between the gate and silicon substrate is uniform and contributes a homogenous stress on the length

scale of the tunnel junctions. In Fig. 4.3 (a) and (b), we show the difference in the various strain components for Al and Ti gates respectively. In comparing Fig. 4.3(c) to Fig. 4.3(a) and (b) we see that the profile of ΔE_c vs position very closely follows that of ϵ_{zz}. For these results, we can expect that ϵ_{zz} will dominate the conduction band modulation in the TJ device based on the confinement and deformation potential discussed in chapter 2. Here, we can note a few observations: 1) the shear strain (ϵ_{xy}) is negligible in our devices, 2) The x-direction normal strain (ϵ_{xx}), which is perpendicular to the transport direction, is the opposite sign from the other components in the gap region, and 3) ϵ_{yy} and ϵ_{zz} are both positive in the gap region. The modulation of the conduction band (ΔE_c) for Al and Ti devices is shown in Fig. 4.3(c) where, because of the differences in the bulk α, ΔE_c is larger in Al compared to Ti. The results of the COMSOL simulation are strongly dependent on the choice of parameters such as the tunnel junction dimensions and oxide thickness. The results in the Fig. 4.3 are calculated for parameters most relevant to the experimental data in this work (25 nm thick oxide and 30 nm wide gap length). Based purely on the bulk α values, the expected difference in the barrier height in the gap is about 11 meV.

4.2 Device fabrication and characterization

We designed the process flow for our TJ devices with the goal of keeping the processing steps as similar as possible between Ti and Al gated devices that are the focus of this chapter. We have shown this process as a series of diagrams in

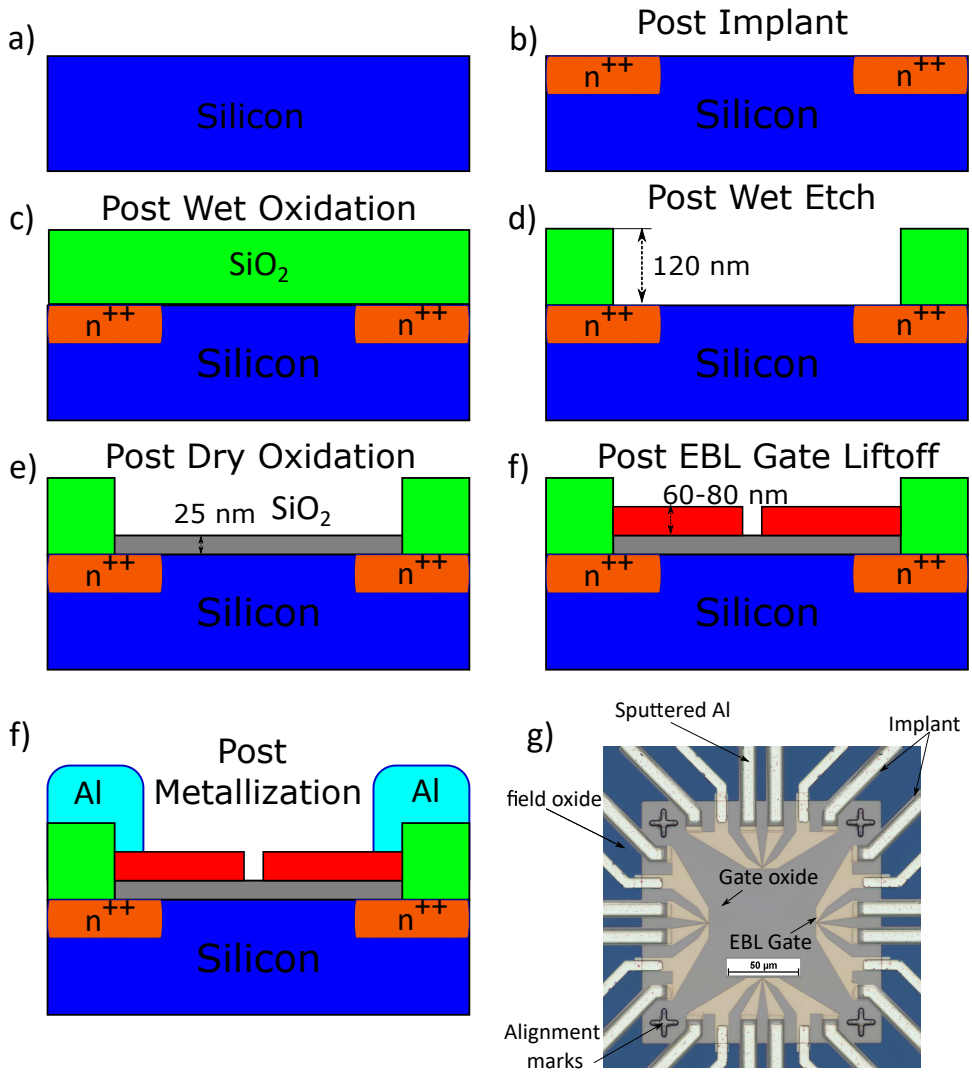

Figure 4.4: Schematic of our TJ device fabrication process. (a) Bulk Si wafer starting point. (b) After phosphorous ion implantation for source/drain regions. (c) After wet oxidation to grow a field oxide for electrical isolation. (d) After using a wet etch process to remove the field oxide from the region where we will do the EBL patterning. (e) After dry oxidation to grow 25 nm of silicon dioxide. (f) After the EBL patterning and liftoff of the e-beam evaporated gate metal. (g) After final metallization using sputter deposited Al. (f) Optical image of finished Ti gated TJ device.

Fig. 4.4 (a) through (f). We start with lightly boron-doped bulk silicon wafers with a resistivity of 5-10 Ω·cm. We use a phosphorous ion implant for the current and voltage probe ohmics. We then grow a 120nm thick wet field oxide at 900°C for electrical isolation when wire bonding the device for low temperature measurements. This oxide is removed from the fine regions where the e-beam lithography will be performed via wet etching in buffered oxide etch (BOE). The gate oxide is 25 nm thick and is grown using a dry oxidation in a Cl ambient at 950°C. The processing then diverges for the fine area lithography used for the gate layout of the devices. For Al gates, we use a positive tone e-beam lithography (EBL) liftoff process and the e-beam evaporator to deposit about 80 nm of Al. The process is similar for Ti gates, but the thickness is 60 nm. This was an unintentional difference introduced via an error with the e-beam evaporator. We believe this thickness difference is not a significant factor in the results as in both cases the film thickness is significantly larger than the tunnel gap length. For our simulations with those film thicknesses and gap lengths, we do not identify any thickness dependence of the barrier height.

We choose to compare Ti and Al in our devices for two reasons. First, Al is the most commonly used silicon MOS QD gate material [136]. Second, Ti is commonly used an adhesion layer for other commonly used gate materials, such as Pd [137, 138], and so we wanted to begin with studying the properties of Ti first since it will likely appear in any future devices. Additionally, Al and Ti have very similar bulk work functions and thus we can expect similar electrostatic properties, such as threshold voltage, in the absence of charge defects. In both cases, we use a bilayer PMMA stack composed of 180 nm of PMMA 950 A4 on top of 180 nm of

PMMA 495 A4 for liftoff. This PMMA stack leads to very narrow, less than 15 nm, of undercut in the under layer that is ideal for liftoff with the critical dimensions needed for our TJ devices. We then sputter deposit Al for the ohmic contacts and metal lines. Both types of devices are finished with an anneal in 10 % forming gas (H_2/N_2) at 425 °C for 30 minutes.

Another gate material that has been used in silicon MOS QDs [67, 68, 69] and possesses a significantly lower coefficient of thermal expansion than most metallic gates [77] is doped poly-silicon. We also fabricated TJ devices analogous to the Al- and Ti-gated devices discussed above, but with doped poly-silicon. The majority of the fabrication process for these poly-silicon-gated devices is identical to that listed above with the exception of the gate deposition and patterning. The doped poly-silicon is deposited via low-pressure chemical vapor deposition in a tube furnace at 625 °C at pressure of 200 mT. Here, a mixture of silane and phosphine gases are used to deposit an in-situ phosphorous doped (n-type) poly-silicon film. The EBL process is significantly different from the PMMA bilayer liftoff used for Al and Ti gates. For poly-silicon, we use a negative tone EBL resist (XR-1514-6%, 150 nm thick) to pattern the TJ region. To remove the poly-silicon, we perform a dry etching process in inductively coupled plasma (ICP) etcher with process gases of Cl_2 and O_2 with flow rates of 50 sccm and 20 sccm respectively. As a final step of the EBL process, we remove the remaining cured XR-1514-6% in 100:1 hydrofluoric (HF) acid. The combination of the dry etching process and this 100:1 HF etch lead to damage and unintentional etching of the gate oxide between the gap in the gates in the TJ.

The poly-silicon films also have a significantly larger magnitude of as-deposited stress than the room temperature evaporated Al and Ti films, where we measured an as deposited stress (σ) for doped poly-silicon of $\sigma_{polysi} = -319 \pm 34$ MPa as compared to $\sigma_{Al} = -83 \pm 10$ MPa and $\sigma_{Ti} = -69 \pm 22$ MPa for Al and Ti respectively. A potential upside of the high deposition temperature of the doped poly-silicon is that it does not change after our normal forming gas anneals, where the anneal temperature is always significantly less than the deposition temperature so no plastic deformation will occur. In terms of designing devices to set a particular stress state, this is advantageous because it allows us to tailor the forming gas anneal as needed for other uses such as reducing oxide charge defect densities. For the purposes of our TJ measurements, which requires a relative comparison between different devices, the poly-silicon TJ device fabrication is too different from the fabrication of the Al and Ti devices. We find that the poly-silicon devices are unable to fit our electrostatic criteria discussed later in this chapter and we attribute this difference to the damage done to gate oxide from EBL processing.

For all of the data presented in this chapter, we have used the same low temperature measurement setup. After performing some room temperature measurements in the dark box described in chapter 3 to screen devices, we cleave the wafer into chips for low temperature measurements. We mount and wire bond the chips to headers that fit into the cryostat shown in Fig. 4.5, where (a) shows the system with the vacuum can and thermal shielding in place as it would be during a cooldown and (b) shows the sample stage where were mount the chip header. This closed cycle cryostat reaches a base temperature of ≈ 2 K. Once the sample is at 2 K, we will

proceed with measuring each TJ device on the chip. For all the measurements in this chapter, we use an Agilent 4156C parameter analyzer. To measure currents and apply the bias and gate voltages, we use the four source measurement units (SMU) on the 4156C. To measure voltages when doing 4-terminal IV measurements, we use the two voltage measurement units (VMU).

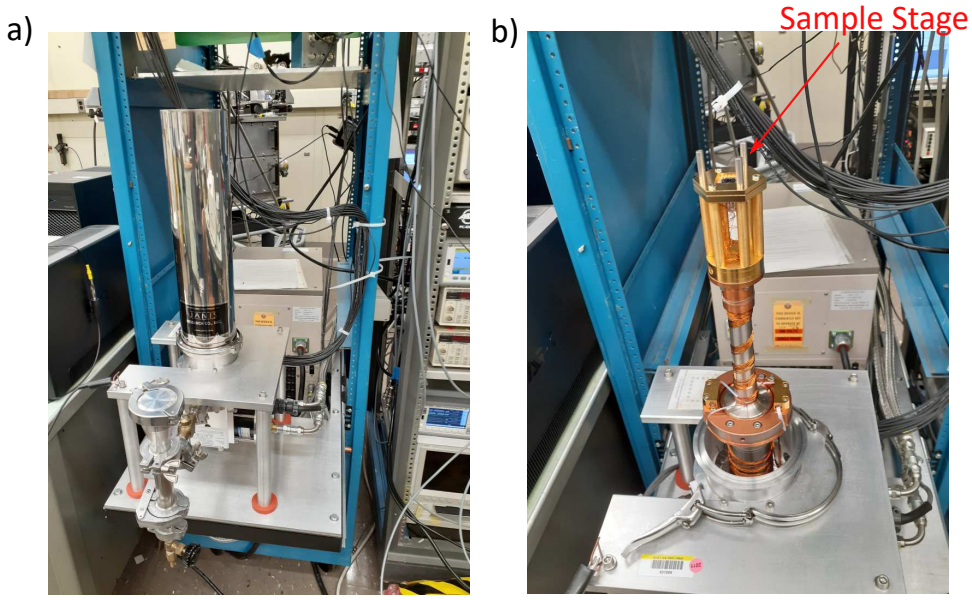

Figure 4.5: Images of the 2 K closed cycle cryostat used for low temperature measurements of the devices in this work. (a) Image of the system with vacuum can and thermal shield in place. (b) Image of the system without the vacuum can and thermal shield showing the sample stage and wiring.

We have designed the device layout to enable four-terminal measurements and independent tuning of the electron density on either side of the barrier. As a consequence of the four-terminal design, the left (right) gate, source (drain) and one of the voltage probes can be used as a transistor to measure threshold, V_T, on either side of the barrier. An optical image of an Al gated TJ junction is shown in Fig. 4.6(b). The different colored paths correspond to the data colors in Fig. 4.6(a),

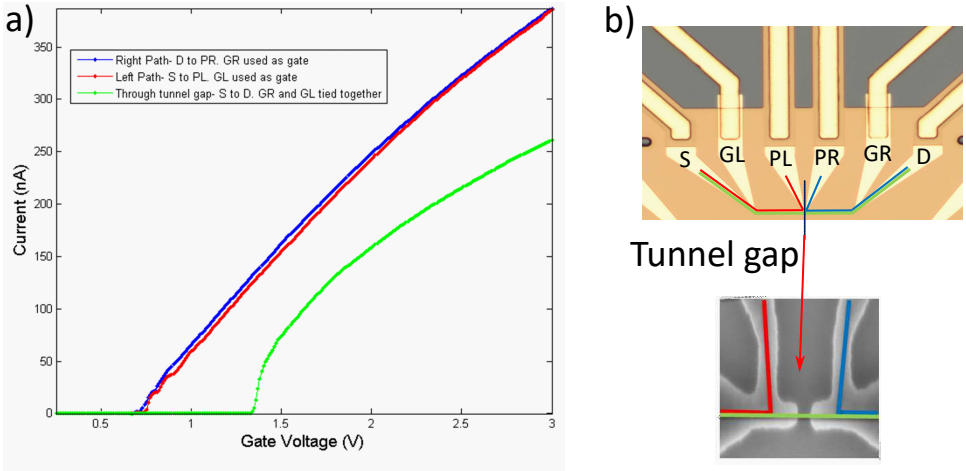

Figure 4.6: (a)Turn on for a TJ device at 2 K using different combinations of ohmics and gates. (b) Optical and SEM images of a TJ junction device showing the different colored paths corresponding to the turn-on data in (a). Here GL and GR are gates and S, D, PL, and PR are ohmics. The red (blue) path measures current from the S (D) ohmic to the PL (PR) ohmic and avoids any transport through the tunnel gap. The green path measures current from the S to D, which is transport through the TJ.

where, for example, the red path is measuring current from the S to PL ohmics while sweeping the GL gate voltage. Importantly the red and blue paths do not measure any current through the TJ itself. This gives us an independent measure of V_T for each gate, signifying when a strong inversion layer is formed under each. The green path represents transport through the barrier. In Fig. 4.6(a), we can see that turn on through the green path is significantly delayed from the red and blue paths. This shows that the formation of an inversion layer in the gap region requires a significantly higher gate voltage, as we would expect for our TJ devices.

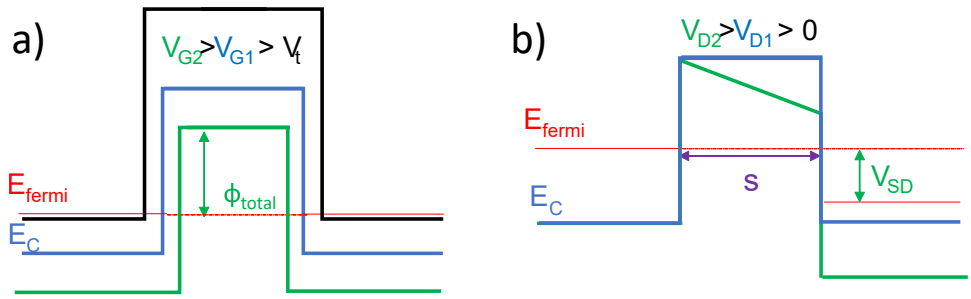

Figure 4.7: Schematic of the expected electrostatic dependence in the trapezoidal barrier model, where the barrier height (labeled ϕ_{tot}) and width (labeled s) will both decrease with increasing gate voltage shown in (a) and (b) the source-drain bias has the effect of tilting the barrier.

4.3 Barrier modeling

After performing measurements to map out the gate voltage and drain bias dependence of the tunnel junction, we can identify a set of gate voltages where the transport characteristics show non-linear $I(V_D)$ behavior. For our 4-terminal $I(V_D)$ measurements, we connect SMUs on the 4156C parameter analyzer to the source and drain contacts and the left and right gate voltages. Additionally, we connect the VMUs on the parameter analyzer to the voltage probe ohmics. Following reference [139] and as discussed below, we fit the differential conductance to extract the barrier parameters from numerically differentiated 4-terminal DC-IV measurements using a sampling method. The process is as follows: 1) Set the gate voltage on both gates, 2) Set the drain bias, 3) Average some number of current and voltage samples, 4) Repeat 2 and 3 on the next drain bias step. Using this averaging method allows us to minimize the noise in the numerical derivative of the IV data.

The modeling of tunneling in nanoscale devices is complicated and still an

intensive area of research [140]. Most models require assumptions to be made that can place limitations on the systems where the model will be valid. In the case of the tunable voltage tunnel barriers under study in this project, the situation has increased complexity due to effects such as the fringing electric fields, which deform the tunnel barrier. Here we assume the trapezoidal barrier model, because it leads to tunneling models with compact closed form solutions. Within that model we assume that the barrier height and width both will decrease with increasing gate voltage. In this way, we can calculate tunneling conductance vs drain bias for tunnel barriers at a specific gate voltage.

Directly measuring absolute strain using our devices is not possible at present because it requires an accurate electrostatic model in order to quantify strain with low enough uncertainty. However, we can measure a relative strain using a simplified electrostatic model [2]. One of the typical behaviors seen in MOS tunnel junctions is that the barrier height shows a linear dependence over some range of gate voltages. We can extend this simple model to include a strain component where strain is independent of gate voltage. If we have two sets of devices with similar electrostatics, we can extract a strain difference by comparing the relative difference in barrier heights.

We model the total tunnel barrier height, ϕ_{tot} in a single device at zero bias as,

$$\phi_{tot} = \phi_\epsilon + \phi_0 + \phi_{ES}(V_G - V_T) \tag{4.1}$$

Where ϕ_ϵ is the strain-induced portion of the barrier, ϕ_0 is the electrostatic portion

of the barrier at threshold, V_T, $\phi_{ES}(V_G - V_T)$ describes the gate voltage dependence of ϕ_{tot}, and V_G is the gate voltage. To extract the absolute value of ϕ_ϵ in a single device requires a model which predicts both ϕ_0 and $\phi_{ES}(V_G - V_T)$ from the geometry, semiconductor physics, and defect charge densities. Our attempts to model ϕ_0 and $\phi_{ES}(V_G - V_T)$ using COMSOL to solve the Poisson and drift-diffusion equations fail to produce a tunnel barrier over any appreciable range of gate voltage above threshold for the leads, contradicting the experimental data. We speculate this is due to a larger density of states, which overestimates the charge density in the barrier. We, therefore, do not extract an absolute value of ϕ_ϵ. We can, however, extract changes in ϕ_ϵ between devices with different gate materials, if $\phi_0^1 \approx \phi_0^2$ and $\phi_{ES}^1(V_G - V_T) \approx \phi_{ES}^2(V_G - V_T)$ so that $\phi_{tot}^1 - \phi_{tot}^2 \approx \phi_\epsilon^1 - \phi_\epsilon^2$ where the superscripts 1 and 2 refer to different materials. ϕ_0 is determined by the metal semiconductor work-function difference and defect charge densities. Controlling ϕ_0 requires we reproducibly minimize the unwanted charge density at the interface and in the oxide, particularly the defects discussed in chapter 3. To control for the inevitable work function difference in our analysis, we will compare ϕ_{tot} from different devices on a $V_G - V_T$ axis. In addition to charge density and the work-function difference $\phi_{ES}(V_G - V_T)$ is also determined by the geometry (gate and gap dimensions). We control for this effect by comparing devices with similar geometry. Thus, our analysis assumes 1) that the work function difference between the two materials is accounted for by subtracting off the threshold voltage; 2) using standard fabrication methods, variations in the amount of charge in the gate oxide have been reduced to a negligible level; and 3) any effect other than strain which would produce a difference in barrier

height in nominally identical devices, save for the gate materials, is negligible. We examine whether our experiment satisfies these assumptions later.

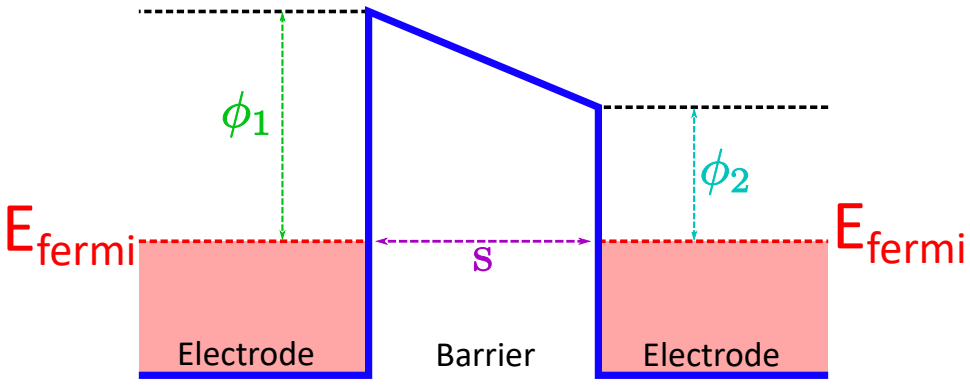

Figure 4.8: Diagram of the trapezoidal barrier model used in the Brinkman-Dynes-Rowell (BDR) model [139] for asymmetric electrodes ($\phi_1 \neq \phi_2$) at zero bias. In contrast to the barrier model shown in Fig. 4.7, the barrier drawn here is tilted at zero bias due to the barrier asymmetry. The width of the insulating region is defined as s.

To achieve the goal of extracting properties of the barrier from transport measurements, we use the Brinkman-Dynes-Rowell (BDR) model [139] which gives the current density, J, through a trapezoidal barrier shown in Fig. 4.8 as:

$$J = \frac{2e}{h}\sum_{k_i}\int_{-\infty}^{\infty} exp\left(-\frac{2}{\hbar}\int_0^s \sqrt{2m[\phi(x,V)-E_x]}dx\right)[f(E)-f(E-eV)]dE_x \quad (4.2)$$

where $\phi(x,V) = \phi_1 + \frac{x}{s}(\phi_2 - eV - \phi_1)$, k_i are the transverse momentum, and $f(E)$ is the fermi distribution function. Eq. 4.2 assumes the WKB approximation for the barrier, where the band profile must vary slowly relative to the electron-wave function. We use a low voltage approximation of Eq. 4.2 which gives the

conductance through the tunnel barrier as a function of bias voltage, G(V), as:

$$\frac{G(V_D)}{G_0} = 1 - \frac{\sqrt{2m}s\Delta\phi}{12\hbar\phi_{tot}^{3/2}}eV_D + \frac{ms^2}{4\phi_{tot}}(eV_D)^2 \tag{4.3}$$

Where s, ϕ_{tot}, $\Delta\phi$, and m are the barrier width, barrier height, barrier asymmetry, and effective carrier mass respectively. In the context of Fig. 4.8, $\phi_{tot} = \frac{\phi_1+\phi_2}{2}$ and $\Delta\phi = \phi_2 - \phi_1$. Here, G_0 is the zero bias conductance of the junction given by:

$$G_0 = \frac{eW_g t_{inv}}{h}\frac{\sqrt{2m_e^* e\phi}}{s} \tag{4.4}$$

In this equation, we have included the terms W_g and t_{inv} that are representative of the area of our TJs. Here, W_g is the width of the gates, typically between 100 nm and 500 nm, and $t_{inv} \approx 4$ nm is the inversion layer thickness [83]. It should be noted that due the exponential dependence of the conductance on the barrier parameters, the choices of the physical values for W_g and t_{inv} do not strongly impact the extracted barrier height or width. As noted above, Eq. 4.3 is a low voltage approximation of the numerical solution of Eq. 4.2 at zero temperature and is within 10 % when barrier asymmetry is low ($\frac{\Delta\phi}{\phi_{tot}} \leq 1$) and the barrier width is larger than 1 nm [139]. For the MOS tunnel barriers studied in this work, we find that these conditions are always met for the junctions presented later. In TJ devices that show non-ideal transport properties, such as Coulomb blockade through unintentional dots, we do see situations where the barrier asymmetry is large. In this case, we have rejected these devices from any further analysis with respect to strain. Due to the

nature of these MOS device and as has been already discussed; the extracted barrier parameters and tunneling conductance will all be a function of gate voltage.

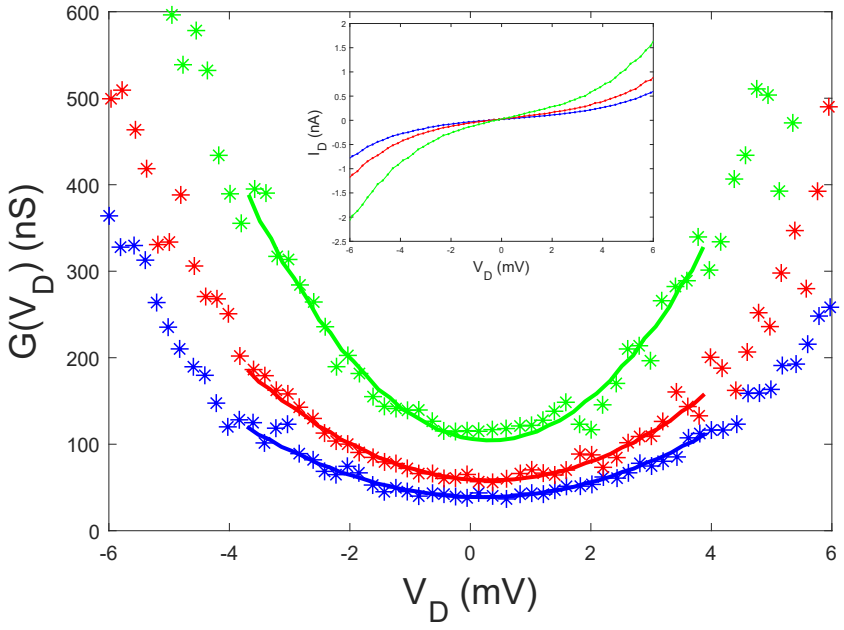

Figure 4.9: 4-terminal DC transport data for a Ti gated tunnel junction device. The inset shows the measured $I(V_D)$ used to obtain the conductance (dI/dV_D) through numerical differentiation that is plotted in the main panel. We have taken the blue, red, and green curves at gate voltages of 0.87 V, 0.88 V, and 0.89 V respectively. The lines are quadratic fits to equation 4.3 [139](see text). All data are taken at $T = 2$ K.

Fig. 4.9 shows an example of differential conductance vs drain bias data obtained for one of our tunnel junction devices. This gate voltage dependence of tunnel barrier parameters is an inherent property of the electrostatics of the system. Specifically, it will be highly dependent on the choice of gate oxide thickness and lithographic dimensions of the gate. Therefore, we can use the gate voltage dependence of the barrier height and width as a means of assessing whether we have sufficiently controlled the electrostatic environment, which is essential for our

approach to measuring induced strain from the barrier parameters.

4.4 Extracted barrier parameters

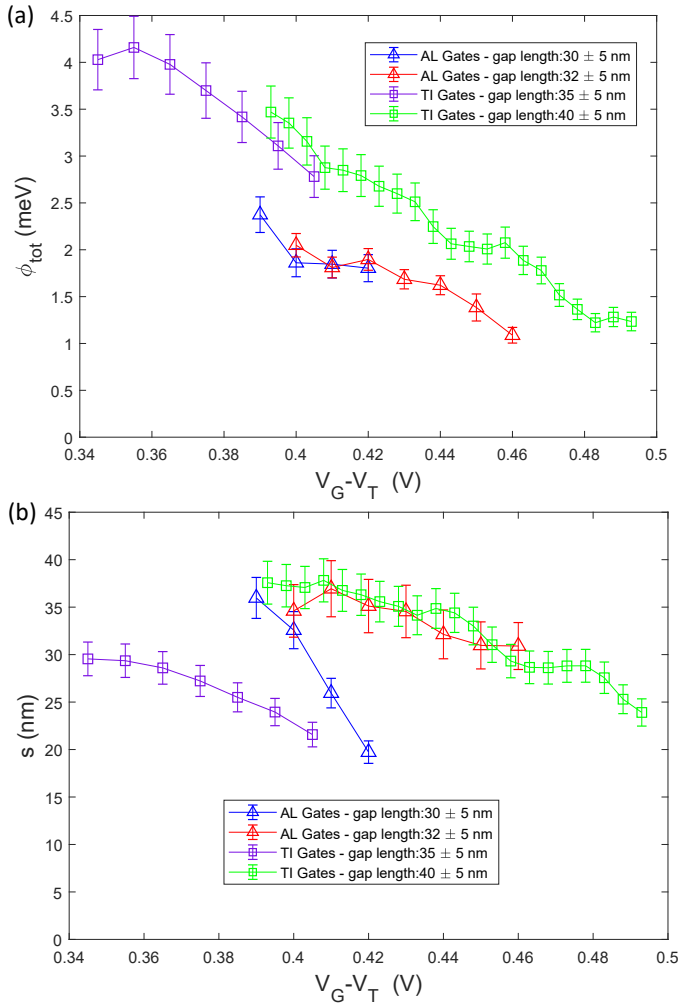

Figure 4.10: (a) Barrier height and (b) barrier width as a function of gate voltage for different MOS tunnel junctions. The barrier heights and widths for metal devices both show a consistent trend of decreasing height. The uncertainty on the barrier parameters represents a statistical uncertainty for a 95% confidence interval.

In our model, the barrier width, s, and height, ϕ_{tot}, are a function of V_G. We expect each to decrease with increasing gate voltage due to fringing fields. This

Material	SEM gap length (nm)	Maximum s (nm)	Threshold voltage (V)
Al	30 ± 5	35.2	0.62
Al	32 ± 5	33.5	0.61
Ti	35 ± 5	29.5	0.52
Ti	40 ± 5	37.9	0.60

Table 4.1: Comparison of the extracted barrier widths (s), measured lithographic sizes using a SEM, and threshold voltage for the four TJ devices presented in this work.

relationship is seen in all of the tunnel junction data shown, in Fig. 4.10(a) and (b). The barrier width does not directly affect our measurement of strain but it serves as a consistency check for the electrostatic behavior of the tunnel junctions. The maximum $s(V_G)$ agrees well with the lithographic widths measured in a FE-SEM (dimensions listed in Table 4.1). We use devices in our analysis which show consistent behavior in the barrier parameters as a function of $V_G - V_T$. Here, this means we exclude the 30 nm wide Al device when extracting relative strain because the slope of the $s(V_G - V_T)$ deviates significantly from the others.

4.5 Validity of electrostatic model

Material	β from 2D $G(V_D, V_G)$	β from $\phi_{tot}(V_G)$
Al	0.067 ± 0.02	0.014 ± 0.005
Ti	0.073 ± 0.02	0.022 ± 0.002

Table 4.2: Comparison of the capacitive lever arms, β, obtained from 2D conductance data by performing a linear fit at constant conductance following Ref [2] (column 1) and by calculating the slope of the data sets in Fig. 4.10(a) (column 2).

As discussed earlier, a detailed electrostatic model to predict ϕ_0 and $\phi_{ES}(V_G - V_T)$ is beyond the scope of this work. We investigate whether the simple model mentioned earlier can predict the slope of ϕ_{tot} in Fig. 4.10(a). Motivated by the linear

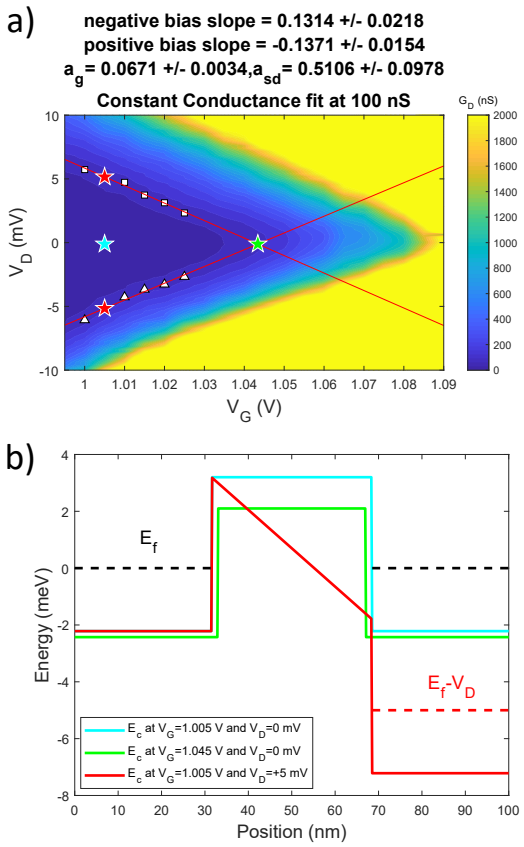

Figure 4.11: (a) Determination of gate and source-drain capacitive couplings to the barrier region from 2-dimensional conductance map for a titanium-gated device at 2 K based on the work in ref [2]. The red lines are a fit to constant conductance points (white squares and triangles for positive and negative source-drain bias respectively) at 100 nS. The choice of the conductance value does not significantly affect the results for the slopes. (b) Trapezoidal barrier profile based on equation 3 of the main text and values for conductance in (a), where the color of the star corresponds to color of band profile ($E_{Fermi} = 0$). The band profile illustrates the qualitative difference in the tunneling between the case of a trapezoidal barrier (fits to equation 3 in the main text) and Fowler-Nordheim tunneling (2-d conductance data) which may impact our determination of β.

dependence of ϕ_{tot} on V_G, we apply the linear gate voltage model from reference [2] to $\phi_{ES}(V_G - V_T)$ from equation 4.1 as $\phi_{ES}(V_G - V_T) = -e\beta(V_G - V_T)$. Here, e is the elementary charge and β is the lever arm of the gate on the barrier. We can now compare the value of β determined in two different ways: 1) the slope of the data in Fig. 4.10(a), and 2) linear fits to 2D conductance data (see Fig. 4.11).

We use the linear model from ref [2] to extract gate and source-drain capacitive couplings to the barrier region from 2-d conductance maps for tunnel barrier devices. An example of one such set of conductance data and linear fits is shown in Fig. 4.11 (a) for a titanium gated device. The red and green stars occur at the same conductance $G(V_D) \approx 100$ nS. From Fig. 4.11 (b), we can see that for the trapezoidal barrier these two points represent qualitatively different physical pictures for the barrier. In the case of the green star, the transport is only via direct tunneling (since $V_D = 0$) with a width and height of 34.4 ± 2.1 nm and 2.1 ± 0.2 meV respectively from the BRD model fits. In the case of the red star, the barrier has tilted such that the right edge of the barrier has dropped below the Fermi level on the left side. This is known as Fowler-Nordheim tunneling and significantly decreases the effective width of the barrier. The determination of β from 2D conductance plots corresponds to V_G and V_D on the line connecting the red and green stars. This line represents a change in both ϕ_{tot} and s such that the product $s\sqrt{\phi_{tot}}$ results in constant $G(V_D)$ over the range of V_G and V_D. In contrast, the determination of β from fits to Eq. 4.3 corresponds to $V_D = 0$ and V_G connecting the blue and green stars. Therefore, this line corresponds to the change in ϕ_{tot} purely due to V_G and a non-constant $s\sqrt{\phi_{tot}}$ product. These differing pictures likely limit the agreement of

β as determined by these two methods in Table 4.2. The result of this comparison is shown in Table 4.2. The values obtained from the 2D conductance data agree to within a factor of five with those determined by the slope of ϕ_{tot}. Considering the simplicity of the model and that the range of V_D considered for the 2D conductance value of β corresponds to Fowler-Nordhiem tunneling, while β from $\phi_{tot}(V_G - V_T)$ is at $V_D = 0$, we believe the agreement is reasonable to prove that our assumption of similar electrostatics between different devices is valid.

4.6 Determination of Relative Strain from Tunnel Junction Measurements with Strain

We calculate the difference in strain between Ti and Al-gated devices from the data in Fig. 4.10(a) as $\phi_\epsilon^{Ti} - \phi_\epsilon^{Al} = \phi_{tot}^{Ti}(V_G - V_T) - \phi_{tot}^{Al}(V_G - V_T)$, where superscripts Al and Ti refer to the different gate materials. $\phi_{tot}^{Ti}(V_G - V_T) - \phi_{tot}^{Al}(V_G - V_T)$ is averaged over $0.4 \leq V_G - V_T \leq 0.46$ and appears as the right-most data point in Figure 4.12. Based on bulk α values of the gate materials, we would expect $\phi_\epsilon^{Al} > \phi_\epsilon^{Ti}$, however, our data show that $\phi_\epsilon^{Ti} > \phi_\epsilon^{Al}$. We can make this comparison more quantitative by performing COMSOL simulations of the mechanical effects using the bulk values of α for each gate material ($\alpha_{Ti} = 8.9 \pm 0.1 \times 10^{-6} K^{-1}$, $\alpha_{Al} = 23.0 \pm 1.0 \times 10^{-6}\, K^{-1}$ [141]). This value appears as the leftmost data point in Fig. 4.12 and strongly disagrees with our data. To resolve this disagreement we measured the value of α for each metal film. We found for our films: $\alpha_{Ti} = 16.2 \pm 2.0 \times 10^{-6} K^{-1}$, $\alpha_{Al} = 23.0 \pm 2.8 \times 10^{-6}\, K^{-1}$. While there is good agreement between our measured α_{Al} and the bulk value, our

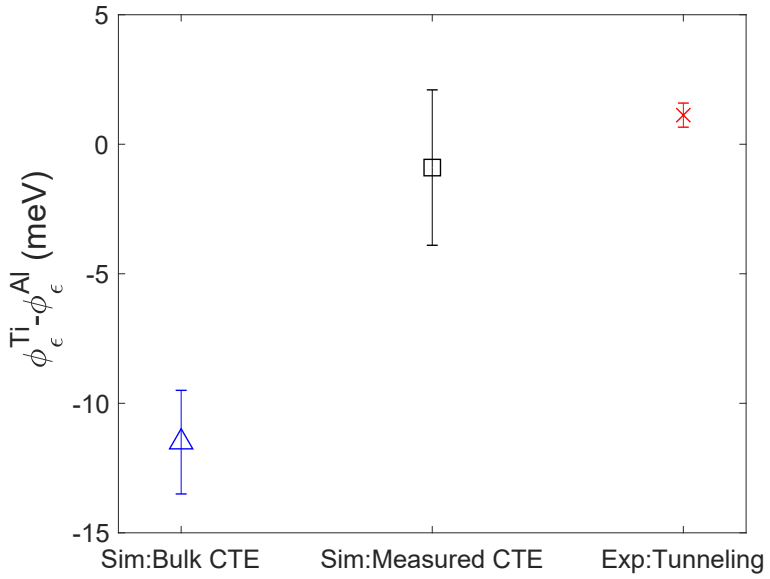

Figure 4.12: Comparison of the barrier height difference between Ti and Al devices using the data from Fig. 4.10(a) and the expected barrier height due solely to strain from COMSOL simulations. The experimental data point is calculated from the average difference over $0.4 \leq V_G - V_T \leq 0.46$ V. The uncertainty in bulk simulations corresponds to the range of differences obtained by assuming an uncertainty of one in the last digit of the values of α in Ref [141]. The uncertainty in the measured α simulations corresponds to the 1σ uncertainty in our measurements of α. The uncertainty in the tunneling data corresponds to the propagated uncertainties in Fig. 4.10.

measured α_{Ti} is significantly larger than the bulk value. This is likely the result of the deposition process which impacts the film morphology so that $\alpha_{film} \neq \alpha_{bulk}$ [142]. The α_i are measured from the slope of film stress, $\sigma(T)$, while stepping temperature, T, of blanket films processed in the same deposition and anneal as the tunnel junction devices using the wafer curvature measurement discussed in chapter 2. The result of simulations using these experimental values as inputs appears as the middle data point in Fig.4.12 and agrees with our experimentally measured value to within our uncertainties.

It is important to note that the simulations only consider strain due to the α mismatch between the materials generated by cooling to $T = 2K$, and treat α as a constant equal to its room temperature value. Since α decreases toward zero with decreasing temperature [143], the simulated barrier height is likely an upper bound on $\phi_\epsilon^{Ti} - \phi_\epsilon^{Al}$. Importantly, while our data agree with continuum mechanics simulations, it is unclear why we reach agreement while neglecting the intrinsic stress of the gate. For the films in this chapter, we measured $\sigma_{Al} = 317 \pm 35$ MPa and $\sigma_{Ti} = 35 \pm 23$ MPa after the forming gas anneal. For this we would expect that the inclusion of the intrinsic film stress would only make $\phi_\epsilon^{Ti} - \phi_\epsilon^{Al}$ even more negative. This would be an even larger disagreement than using the bulk values of α and thus is also inconsistent with our results. It is important to note that it should be expected that α for a thin film will differ from the bulk value. Since α for a material is determined by the anharmonicity of the interatomic potential, anything that modifies this potential will shift the value of α [144]. In thin films, this will include effects such as: residual stress in the film [145], surface interactions between the substrate and film [144], defects and impurities in the film [142], and from the presence of grain boundaries in poly-crystalline films [146, 147].

As mentioned previously, the devices presented in this chapter easily lend itself to future work exploring deposition parameters and anneals to manipulate inhomogeneous strain. Our method for measuring relative strain satisfies the sensitivity, spatial resolution and low-temperature requirements noted in chapter 1. Moreover, the fabrication and measurements are similar to those for QDs so that this method is directly relevant for QD devices. Our data provide an important step forward in

assessing gate-induced strain in QD devices in situ while highlighting the need for further experimental work and a greater theoretical understanding of the electrostatics.

Chapter 5: Controlling the defect densities and strain via annealing

In chapter 1, we discussed that recent work in the field of silicon MOS QDs has transitioned to using Ti/Pd gates from the more tradition all Al gates [137, 138, 148]. This transition has been driven by the ability to achieve smaller critical dimensions in the nanoscale lithography due the smaller grain sizes achieved with evaporated Ti/Pd films as opposed to Al. Additionally, based on the bulk coefficients of thermal expansion (α) [79], the expectation is that QDs fabricated with Ti/Pd gates will experience a lower magnitude of strain-induced modulation than Al. Based on the results presented in the previous chapter, we know that assumptions based on the bulk values of α are not necessarily valid. This shift in gate material likely also changes the charge density, such as the fixed charge density (Q_f) and interface trap density (D_{it}) discussed in chapter 3, in the gate oxide. It is unclear from previous work in the silicon MOS QD community how much of an impact that all of these changes have had in device performance.

In this chapter, we address these topics by presenting a comparison of Q_f, D_{it}, σ, and α for Ti/Pd, Ti/Pt, and Al as a function of forming gas anneal temperature and hydrogen concentration. We vary the anneal temperature from 200 °C to 425 °C, using both 5 % and 10 % mixtures in 30 minute anneals. We show that

Ti/Pd and Ti/Pt have larger D_{it} than Al when optimally annealed and that the magnitude of Q_f is larger for Ti/Pd and Ti/Pt than for Al, with Al showing a net negative charge while Ti/Pd and Ti/Pt display net positive charge. Additionally, we show that both α and σ increase with increasing anneal temperature. Moreover, these results show that, due primarily to intrinsic strain, Pd-gated devices have larger strain-induced modulation of the conduction band than their Al-gated counterparts, directly contradicting expectations based on the bulk α alone [72]. Finally, and most importantly, we find no anneal which simultaneously minimizes defects and the effects of strain in any of the materials studied. Thus, a tension arises in designing fabrication processes for MOS QDs where one must choose between setting the anneal such that defects are minimized or the strain-induced modulation of the conduction band is minimized. Some portions of the work in this chapter are currently under peer review for the Journal of Applied Physics under the title: "Alternatives to aluminum gates for silicon quantum devices: defects and strain".

5.1 Methodology

In chapter 4, we made tunnel junctions (TJs) to measure strain from Al and Ti gates using a nominal fabrication process with no effort made to optimize any steps to control for either strain or defects. In any MOS QD fabrication process, we have a wide array of potential variables to use in this optimization. For example, just purely from the perspective of the gate material deposition, we could change all of the following to modify the mechanical properties or induced defect density:

Gate Material	Anneal Temperature (°C)	Time (min)	H_2%	Ref
Poly-Silicon	425	30	10	[69]
Ti/Pd	400	25	5	[148]
Ti/Pd	400	30	5	[137]
Ti/Pd	400	30	5	[138]
Poly-Silicon	400	30	10	[149]
Al, Poly-silicon	400	30	5	[150]
Al	245	60	5	[151]

Table 5.1: List of forming gas anneals commonly used in the fabrication of silicon MOS QDs.

material, deposition method (evaporation/sputtering/CVD), deposition rate, deposition temperature, and deposition pressure. This large phase space is complicated by the fact that any changes in the gate deposition must not interfere with nanoscale lithography process. In this work, we have chosen to focus on tuning the gate material and the forming gas anneal. Using different gate materials is a simple method that allows us to draw comparisons between the metals commonly used in the MOS QD community. We chose to use the forming gas anneal as a tuning method because it minimizes the impact on the electron beam (e-beam) lithography since it is typically performed near the end of the fabrication process after all e-beam lithography patterning as been completed. Additionally, the anneal is specifically used in silicon MOS devices to control defects such as D_{it} so it is good starting point for any optimization as any attempts to optimize the process for strain must not be in conflict with the typical defect density reduction processes.

Table 5.1 shows forming gas anneals used commonly by various groups with working silicon MOS QDs. We can see that the most commonly used anneal is at 400°C for 30 minutes in 5 % H_2/N_2. Here, we have chosen to cover a temperature

range shown in the table from 200 °C to 425 °C. We will also compare results for both 5 % H_2/N_2 and 10 % H_2/N_2 ambient anneals. Additionally, we have narrowed the phase space by performing the anneals for only 30 minutes. In this way, we are only considering the effects of changing the temperature, but it is important to note that any trends will likely also depend on the choice of anneal time. For instance, we could achieve similar levels of D_{it} via anneals at 425 °C and 200°C by simply shortening and lengthening the anneal times, respectively [106].

In terms of gate materials, we will study results for both Al and Ti/Pd since they are the most frequently used, but we will also add Ti/Pt into our study. Here, we have added Ti/Pt, even though it has not been used as a MOS QD gate in the literature, because it meets two interesting requirements. First, Pt has a bulk α even lower than Pd meaning there is a possibility for lowering the strain-induced impact. Second, we should be able to evaporate Pt with similar grain sizes as Pd, meaning that all of the benefits to the e-beam lithography seen in Pd devices could also be achieved in Pt.

5.2 Changes in defect densities with anneal

To measure the oxide defect densities due to gate metal deposition, we fabricate metal-oxide-semiconductor (MOS) capacitors with each gate material for capacitance-voltage (CV) and conductance-voltage (GV) measurements. The fabrication process steps are similar to those shown for the TJ devices in chapter 4, but here we do not need to perform e-beam lithography to pattern the MOS capacitors.

The wafers are boron-doped silicon <100> wafers with a resistivity of 5 Ω·cm to 10 Ω·cm. The wafers are cleaned using standard RCA clean procedures immediately prior to growing the gate oxide. A roughly 25 nm thick gate oxide is grown in a dry oxidation furnace at 1000 °C for 22 minutes with a 10 min post oxidation anneal performed in N_2 at the oxidation temperature. This is the same thickness as that used in the TJ measurements presented in the previous chapter, but we changed the recipe because the NIST nanofab acquired a new furnace stack. The oxide in the TJ devices was grown with Cl in the ambient, but the oxides presented in this chapter were grown in pure O_2 as the new furnaces are unable to run Cl processes. MOS capacitor gates are patterned using negative tone resist (maN-1410) and liftoff in solvent. We deposited all of the metals in this work with electron beam (e-beam) evaporation to mimic common QD device fabrication. In e-beam evaporation, the target material is placed in a crucible and is bombarded with an electron beam from a filament source to evaporate and convert the target material to a gaseous state for deposition on the substrate in a high vacuum chamber. For each material, we have used the same deposition rate and pressure: 0.1 nm/s and 3×10^{-6} Torr, respectively. Following deposition and liftoff, we perform isochronal forming gas anneals in an AnnealSys model AS-Master rapid thermal annealer (RTA). Here, we have chosen to use the RTA for the anneals to reduce the effect of the additional time it takes for the tube furnace systems to ramp to and from the set anneal temperature. The typical ramp-up and ramp-down time ranged from 1 to 3 minutes in the RTA with both steps performed in N_2. After annealing, the oxide on the backside of the wafers is removed via a 6:1 buffered oxide etch (BOE) etch and sputtered

with Ti/Au to form a low resistance back contact for measurements (see the series resistance corrections in chapter 3).

We performed CV and GV measurements using the methods discussed in chapter 3. We show an example of the CV and GV curves extracted for the Pd-gated MOS capacitors in this work in Fig. 5.1(a) and (b) respectively. In Fig. 5.1(a), we can see the shift in the CV curves for Pd MOS capacitors with different anneal temperatures with the overall trend showing the flatband voltage shifting towards more positive values of gate voltage as the anneal temperature increases. From this, we can expect a consistent trend of Q_f shifting with anneal temperature alone. Similarly in Fig. 5.1(b), we can see the same shift by focusing on the peak in G for gate voltages less than zero, which corresponds to the interface trap response in depletion. Additionally, we see that the magnitude of the height of the peak in G is changing with anneal temperature. From this, we can infer that D_{it} is changing with the anneal temperature since the MOS capacitors are the same area and the measurements are taken at the same 1 MHz AC frequency. The interface trap density, D_{it}, is calculated from the peak in the conductance vs frequency data using the statistical model from ref [152] (see chapter 3). We extract the flatband voltage (V_{fb}) from CV curves taken at 1 MHz using the $1/C^2$ fitting method [100]. The fixed charge density is calculated using: $Q_f = \frac{C_{ox}}{eA}(\phi_{MS} - V_{fb})$, where ϕ_{MS} is the metal-semiconductor work function difference. The semiconductor work function is calculated using physical constants described in chapter 3, such as the electron affinity and band gap energy for silicon, and the measured substrate doping implied from the slope of the $1/C^2$ plot. The thin film metal work function is measured sep-

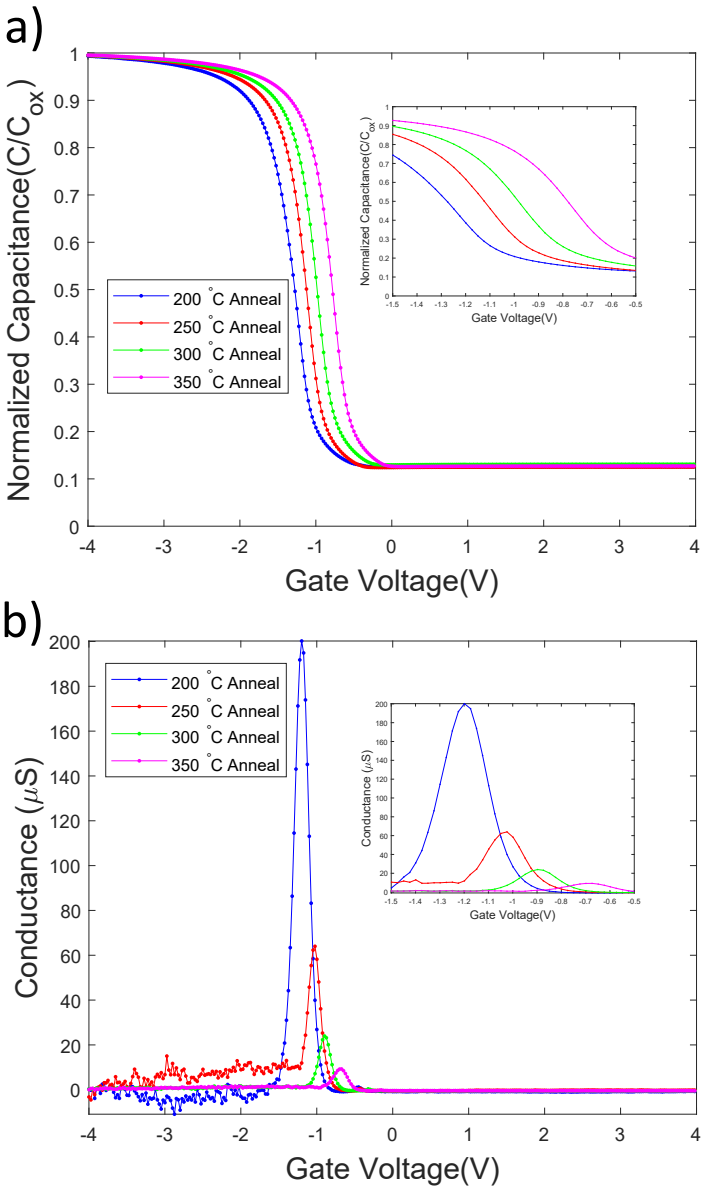

Figure 5.1: (a) capacitance and (b) conductance vs Gate voltage for Pd-gated MOS capacitors for different anneal temperatures in 10 % forming gas. The shift in capacitance shows how flatband voltage and fixed charge density change with anneal temperature. We also show the same shift in the peak in conductance as the position of the peak near flatband moves. Additionally, we can see that the magnitude of the peak changes with anneal temperature showing that the interface trap density is also changing. The insets in both plots show a zoom of the gate voltage region near flatband from V_G =-1.5 V to V_G =-0.5 V.

arately using ultraviolet photoelectron spectroscopy (UPS) rather than assuming the bulk values [153]. Here, we have performed UPS measurements on representative samples and averaged the resulting work functions from measurements of three different spots for each metal. The UPS measurements were performed by Zachary Barcikowski at NIST on samples fabricated by the author.

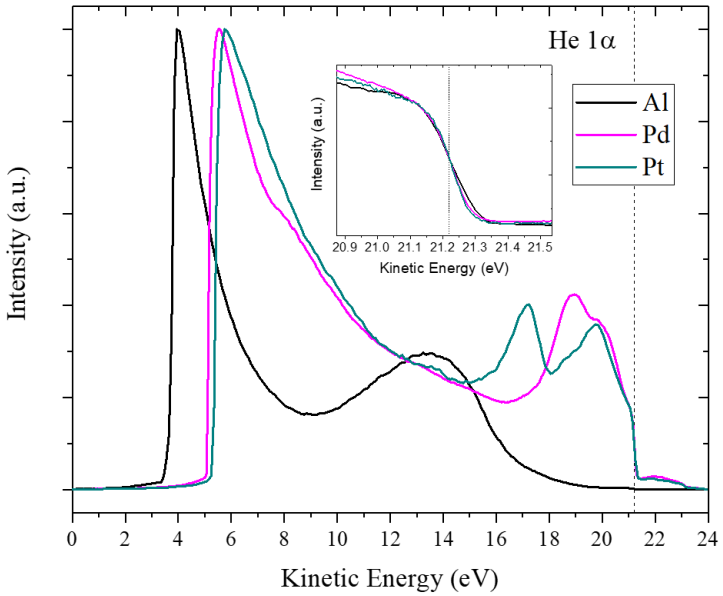

Figure 5.2: Representative Ultraviolet photoelectron spectroscopy (UPS) spectra of 60 nm Al, Pd, and Pt thin films on SiO_2/Si wafers. We normalize the intensity of each spectrum to their peak value. The horizontal axis has been shifted such that $h\nu$= 21.22 eV represents zero binding energy. The inset shows normalized high-resolution scans of the Fermi edge cutoff for each film. Measurements performed and data provided by Zachary Barcikowski and Joshua Pomeroy at NIST

Ultraviolet photoelectron spectroscopy (UPS) operates on the principle of using ionizing radiation to induce a photoelectric effect in the material of interest. Here, ultraviolet photons are typically generated using a gas discharge lamp filled with helium. Recall from chapter 3 that we defined the work function as the differ-

ence between the Fermi level in the material and the vacuum energy level. In the UPS measurements, we measure the work function spectroscopically by measuring the difference between the Fermi level, found at the higher kinetic energies of the spectrum, and the lower kinetic energy cutoff. UPS was performed using the helium 1α excitation from a SPECS UVS 300 high intensity VUV source and a SPECS PHOIBOS wide angle lens hemispherical energy analyzer in ultrahigh vacuum. For the measurement, the samples were biased at -10 V, the electron take-off angle was the sample surface normal, and the pass energy was set to 1 eV. Here, the pass energy refers to potential energy difference between analyzer hemispheres and controls how many electrons can pass through the analyzer. Prior to any measurements, we perform an argon ion sputter clean on the sample, particularly to remove any native oxides present on metals such as Al. To ensure good electrical contacts, we use Cu tape between the thin films and the sample holder. We display the UPS spectra on a kinetic energy scale corrected for the applied sample bias and analyzer work function, where we show an example of this corrected spectra in Fig. 5.2. We determine the effective work function by the low energy secondary electron cutoff, which is located at the left side of the plot in Fig. 5.2. The effective metal work function values, ϕ_M, used in the calculation of Q_f are an average of three measurements performed for each film in different spots with ϕ_{Al}=3.89 $\pm$ 0.16 eV, ϕ_{Pt}=5.33 $\pm$ 0.10 eV, and ϕ_{Pd}=5.13 $\pm$ 0.10 eV. The uncertainty in the work functions is the quadrature sum of the standard deviation between the three measurement spots and the experimental resolution as determined by the full-width at half-maximum of the Fermi edge, denoted on the right side of the plot of Fig. 5.2.

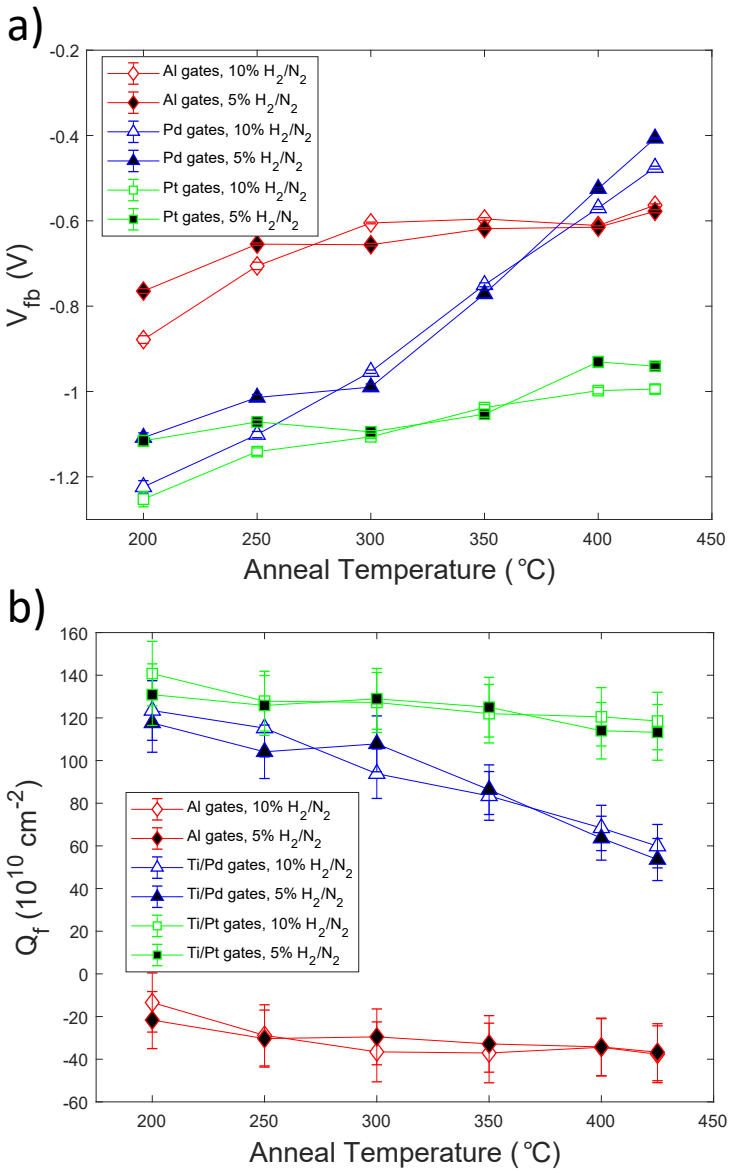

Figure 5.3: (a) Flatband voltage, V_{fb}, and (b) Fixed charge density, Q_f,as a function of anneal temperature for Al, Ti/Pd, and Ti/Pt gates in 10 % and 5 % forming gas. Q_f is calculated using the flatband voltage, V_{fb}, extracted from 1 MHz CV curve (not shown). For the V_{fb} data, each point is an average of at least three different MOS capacitors and the standard deviation is propagated into Q_f based on the equation in the main text. The uncertainty in Q_f is dominated by the uncertainty from the UPS measured work functions.

We show the results for the flatband voltage and fixed charge density as a function of anneal temperature for Al, Ti/Pd, and Ti/Pt gates in Fig.5.3 (a) and (b) respectively. The V_{fb} data versus anneal temperature shows similar trends for all three metals with all of the metals showing increasing positive values of V_{fb} with increasing anneal temperature. We see a difference in the rate at which V_{fb} changes with anneal temperature. Here, Pd shows a larger total shift than both Al and Pt over the same temperature range. These results for V_{fb} are combined with the UPS results to produce the Q_f plots in Fig.5.3(b). There are three important observations to be made from the Q_f data in Fig.5.3(b). First, there is an order of magnitude difference in Q_f between the three different gate materials, with Pt being the largest overall. This is likely due to differences in the e-beam evaporation process between the different metals. For instance, to produce the same deposition rate, Ti/Pt requires a significantly higher applied power than both Pd and Al, which can increase damage to the oxide [154]. Second, Al shows an overall negative net charge value for all anneals while in contrast Ti/Pd and Ti/Pt show net positive charge values. Importantly, the magnitude of Q_f is still smaller in Al than Ti/Pt and Ti/Pd. Third, there is also a difference in the rate at which Q_f decreases with increasing anneal temperature between the three materials. We note that this behavior suggests that there are differences in defects created in the depositions [111, 115], but we are unable to confirm this with the present data.

We show the results for the interface trap density as a function of anneal temperature for Al, Ti/Pd, and Ti/Pt gates in Fig.5.4. The values of D_{it} presented here are taken as the average value across the range of $E - E_v$= 0.34 eV to 0.45 eV

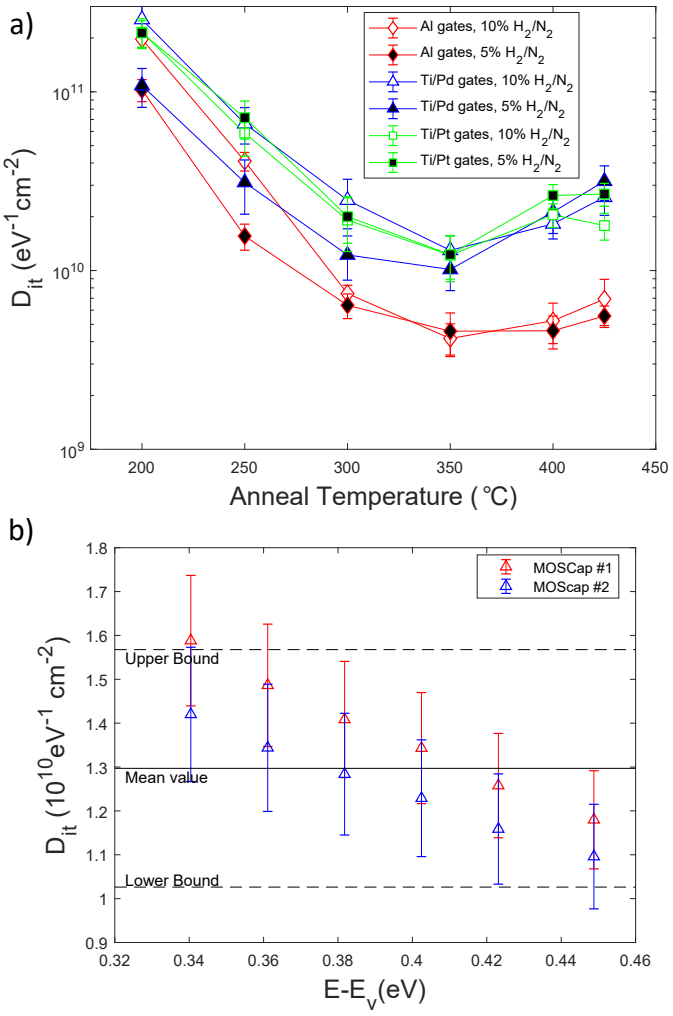

Figure 5.4: (a) Interface trap density, D_{it}, vs anneal temperature for Al, Ti/Pd, and Ti/Pt gates in 10 % and 5 % forming gas. D_{it} is measured using the conductance method and the values reported here are the weighted average of D_{it} measured for an energy range in the band gap of $E - E_v$= 0.34 eV to 0.45 eV. The minimum D_{it} is reached at 350 °C for all metals. The 10 % and 5 % forming gas anneals shows qualitatively similar behavior with temperature and material. (b) Example of the measured D_{it} as a function of energy from the valence band edge ($E - E_v$) for one of the Pd films in (a). The mean value is computed from the weighted average value across this energy range for all MOS capacitators measured. The uncertainties on the individual data points are propagated uncertainties from the measured conductance and the statistical model parameters from chapter 3 and Ref [152]. Those uncertainties are used as weights in calculating the mean and standard deviation. The upper and lower bounds (dashed lines) represent the error bars shown in (a) in the main text, which represent twice the weighted standard deviation of D_{it} about the mean value in the band gap.

in the silicon band gap. For all the metals, D_{it} (Fig. 5.4) reaches a minimum value at 350 °C. Above 350 °C, D_{it} begins to increase with increasing anneal temperature. This "reverse anneal" behavior is a well-known effect activated with long anneals and high temperatures [106]. Most importantly, Al-gated devices reach a lower $D_{it} \approx 3 \times 10^9\ eV^{-1}cm^{-2}$ than both Ti/Pd and Ti/Pt where $D_{it} \approx 9 \times 10^9\ eV^{-1}cm^{-2}$. In comparison to Table 5.1, where the most common forming gas anneal was 400 °C for 30 minutes, we find that this anneal is not optimal for D_{it} and has likely increased D_{it} from its minimum possible value. Although we note that D_{it} in Al devices remain significantly lower (below $10^{10}\ eV^{-1}cm^{-2}$) than Pd and Pt even at this elevated temperature. Here, it is important to note that choosing a different anneal time may affect the value of D_{it} obtained, however, differences between gate materials themselves are expected to persist [106]. For both Q_f and D_{it}, we do not see a large difference between anneals performed in 5 % or 10 % forming gas which agrees well with literature results for metal gates [155].

5.3 Changes in intrinsic film stress with anneal

Fig. 5.5 shows the change in σ in the as-deposited state and following forming gas anneals at various temperatures. Using the wafer curvature methods discussed in chapter 2, each stress value in Fig. 5.5 is the average of six different scans across the wafer 30° apart. For all metals, σ becomes increasingly more tensile with increasing anneal temperature. Ti/Pd and Ti/Pt show similar levels of as-deposited stress, between 160-190 MPa. Both Ti/Pd and Ti/Pt experience a large increase (700-800

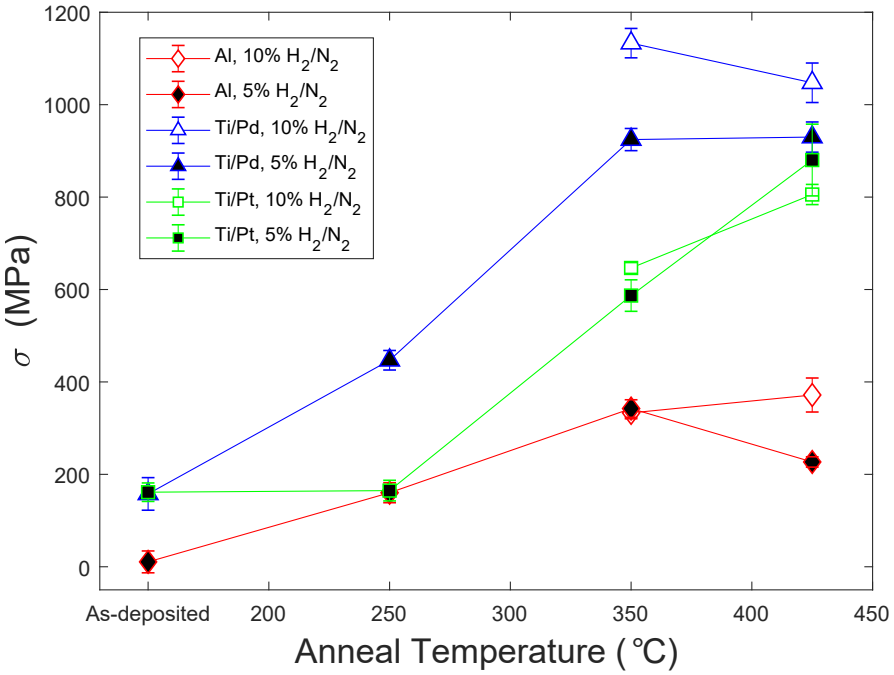

Figure 5.5: Film stress (σ) vs anneal temperature. The intrinsic film stress represents the average measured stress across a 150 mm wafer in 30° increments and the uncertainty is the standard deviation. All films are 50 nm to 60 nm thick and deposited using e-beam evaporation at a rate of roughly 0.1 nm/s with a chamber pressure of 3×10^{-6} Torr

MPa) in stress from their as-deposited values to the highest temperature anneal at 425 °C. This increase is significantly more than the roughly 300 MPa increase observed in Al over the same set of anneals. As with the results in preceding section, we do not see a consistent difference in σ between anneals in 5 % or 10 % forming gas.

5.4 Changes in coefficient of thermal expansion with anneal

Fig. 5.6 shows the change in the room temperature α_{film} in the as-deposited state and following forming gas anneals at various temperatures. Here, we have

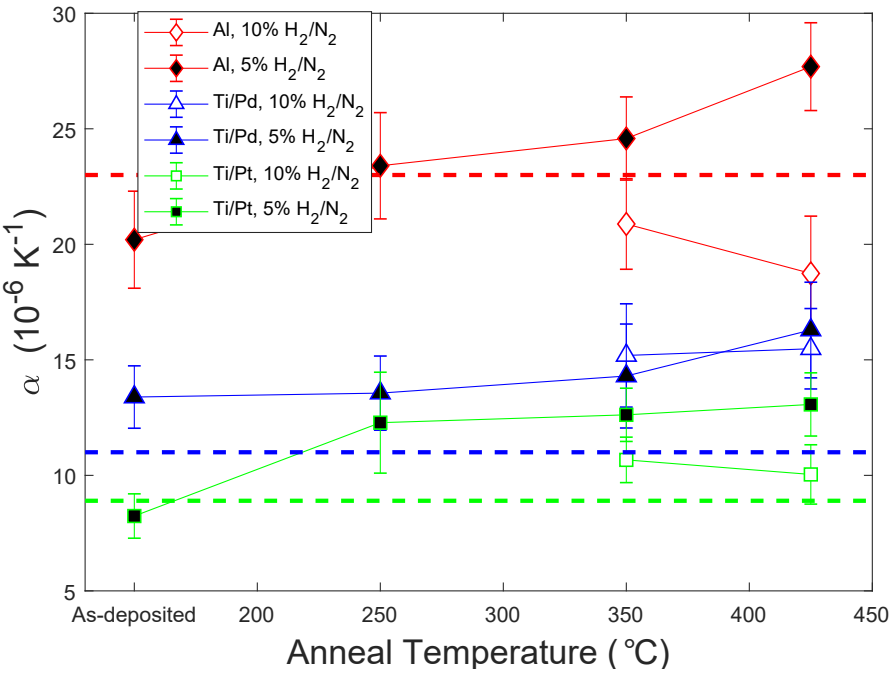

Figure 5.6: Measured coefficient of thermal expansion (α) vs anneal temperature for Al, Ti/Pd, and Ti/Pt films. α is measured from the linear fit of stress vs temperature data from 40°C to 110°C. We equate the value obtained over this range to the room temperature value of α in each case. The uncertainty represents the 95 % confidence interval from the linear fit. The dashed lines indicate the expected bulk α [79].

measured α_{film} using the wafer curvature methods outlined in chapter 2. In this figure, we have shown the expected value of the coefficient of thermal expansion in the bulk state. As with the results in chapter 4, we can see that α_{film} is generally larger than the bulk value and increases with increasing anneal temperature. Since bulk values are typically used to simulate the impact of strain on MOS QD devices [73, 148], these α measurements indicate that such simulations do not fully capture the strain in the device [156]. This is especially true given that the most common forming gas anneal temperature is around 400°C for Ti/Pd-gated QD devices [137, 148].

5.5 Microstructure anaylsis

In the preceding sections, we have shown that for our films σ and α_{film} depend on the anneal and, in the case of α_{film}, deviate significantly from the bulk value. We expect such deviations in elastic properties are due to the microstructure of the thin film. To investigate if this is true for our films, we have compared the changes in α and σ with the average grain size of the films. Here, we have computed the average grain size in ImageJ from an SEM image. We have used 'Analyze Particles' tool to identify the edges of the metal grains. Using the average diameter of the fitted ellipse for each grain, we calculate a lognormal distribution [157] for each material data set for a given anneal. Appendix B goes through this image analysis process step-by-step and shows some additional comparisons of the lognormal distribution compared to other distributions. We show an example of the SEM images and measured grain size distributions calculated from the diameters of the ellipses from the ImageJ analysis in Fig. 5.7 for Ti/Pd films, Fig. 5.8 for Al films and Fig. 5.9 for Ti/Pt films. From the histograms shown in these figures, we can see that in the cases of the Ti/Pt and Ti/Pd films there are changes in both mean and standard deviation of the lognormal distribution. Here, the mean and standard deviation are both increasing with increasing anneal temperature ((a) is as deposited, (b) after 250 °C anneal, (c) after 350 °C anneal, and (b) after 425 °C anneal). These changes can be visually seen in the SEM images as well. Conversely, the Al films show no significant changes in the distributions. To investigate these trends further with we will focus on the average grain size as it trends with other measured properties.

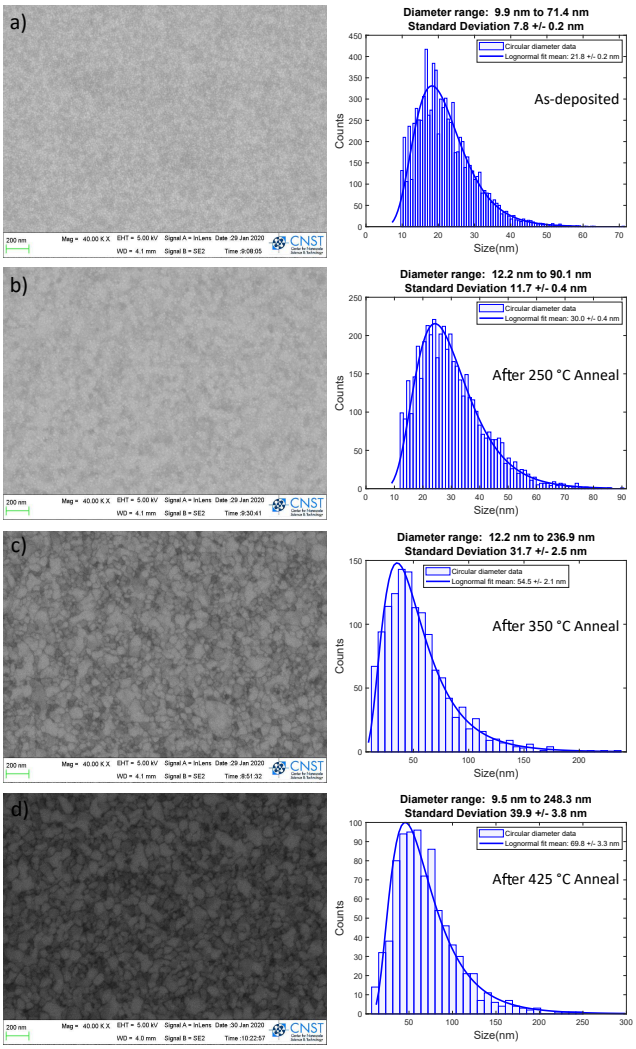

Figure 5.7: SEM images of Ti/Pd films presented in the results of this chapter. Each image represents: (a) after deposition, (b) after a 30 min anneal at 250 °C, (c) after a 30 min anneal at 350 °C, and (d) after a 30 min anneal at 425 °C. The side plots in each SEM image are the resulting grain size distribution for each film. The grain size is obtained by fitting the individual grains to ellipses and using the area of that ellipse to obtain the diameter of grain.

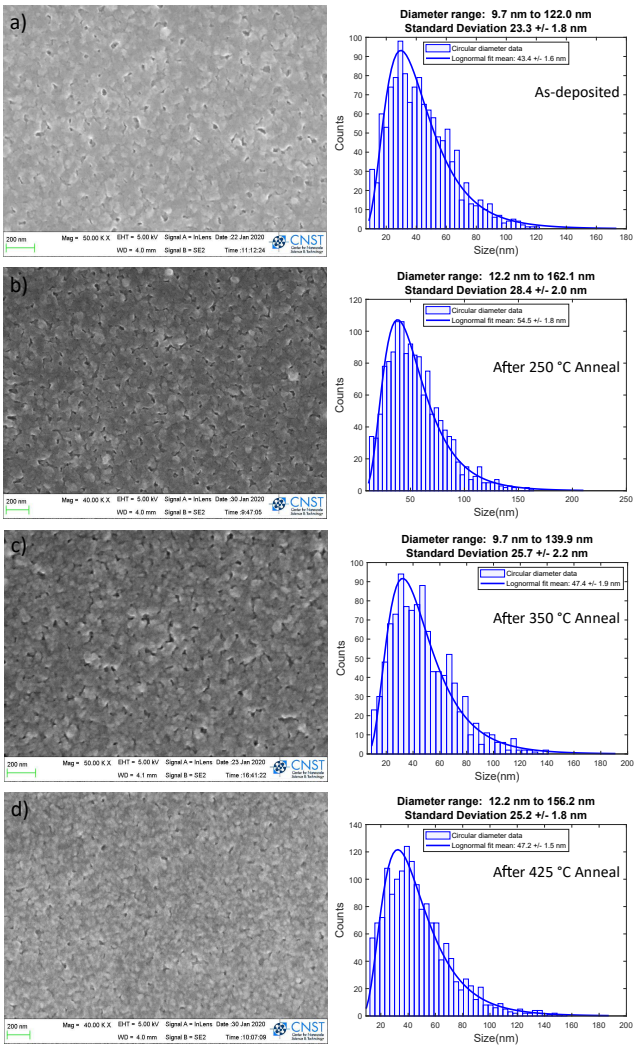

Figure 5.8: SEM images of Al films presented in the results of this chapter. Each image represents: (a) after deposition, (b) after a 30 min anneal at 250 °C, (c) after a 30 min anneal at 350 °C, and (d) after a 30 min anneal at 425 °C. The side plots in each SEM image are the resulting grain size distribution for each film. The grain size is obtained by fitting the individual grains to ellipses and using the area of that ellipse to obtain the diameter of grain.

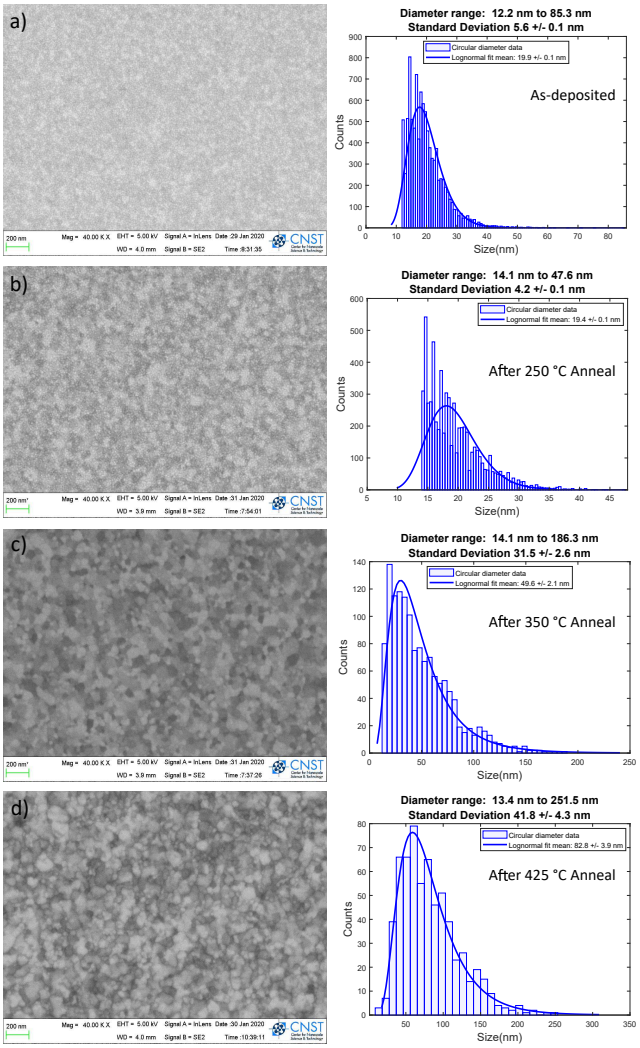

Figure 5.9: SEM images of Ti/Pt films presented in the results of this chapter. Each image represents: (a) after deposition, (b) after a 30 min anneal at 250 °C, (c) after a 30 min anneal at 350 °C, and (d) after a 30 min anneal at 425 °C. The side plots in each SEM image are the resulting grain size distribution for each film. The grain size is obtained by fitting the individual grains to ellipses and using the area of that ellipse to obtain the diameter of grain.

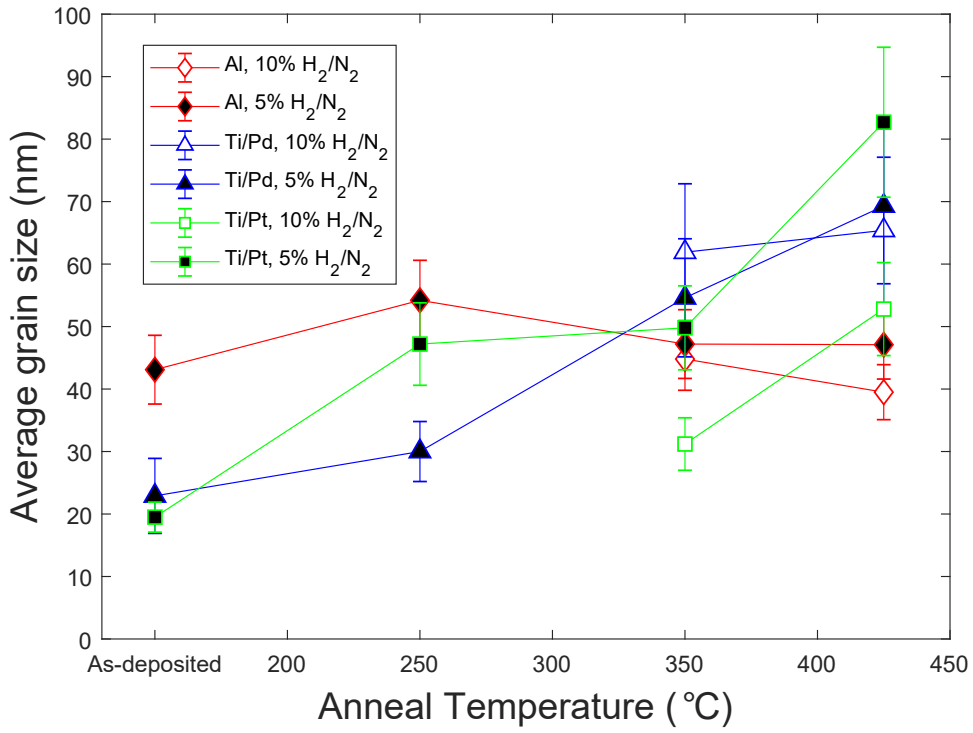

Figure 5.10: Average grain size as a function of anneal temperature. The increasing anneal temperature in the Pd and Pt tends to follow with an increase in the average grain size. For the Al films, the average grain size is relatively constant.

Using the analysis from the SEM images, we can plot the average grain size as a function of the anneal temperature for each of the metal films. This is shown in Fig. 5.10. In general, the trend for grain size with anneal temperature is rather weak in the Al films, but the Ti/Pd and Ti/Pt films show a trend of increasing average grain size with increasing anneal temperature. Fig. 5.11 and Fig. 5.12 shows the change in α and σ respectively with the average grain size. In general, the Al films show similar average grain sizes across all films and anneals. Thus for Al, the changes in α and σ have no discernable morphological trends. Conversely, the Pt and Pd films show rather dramatic changes in morphology depending on

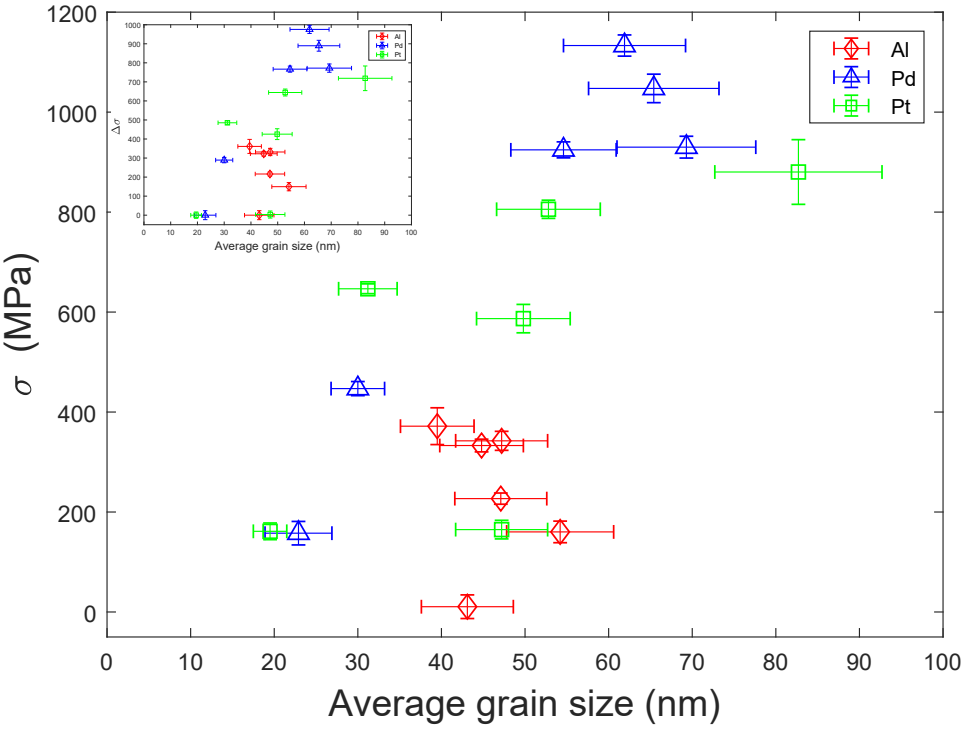

Figure 5.11: Film stress vs average grain size. The increasing tensile stress in the Pd and Pt appears to tend following with the increasing grain size. For the Al films, the average grain size is relatively constant for all films.

the anneal as shown in Fig.5.10. We find that the increasing trends in α and σ for Pt and Pd with increasing anneal temperature tend to follow an increasing trend in average grain size, but the overall trend is rather weak. We note that in this study we have only seen these morphological changes in blanket films used for wafer curvature measurements and such morphological transition do not occur for our e-beam lithography patterned features.

We note that in this work we have chosen an approach of simply using a basic set of deposition parameters for all the films, such as a constant deposition rate of 0.1 nm/s and deposition at room temperature. These choices affect the morphology

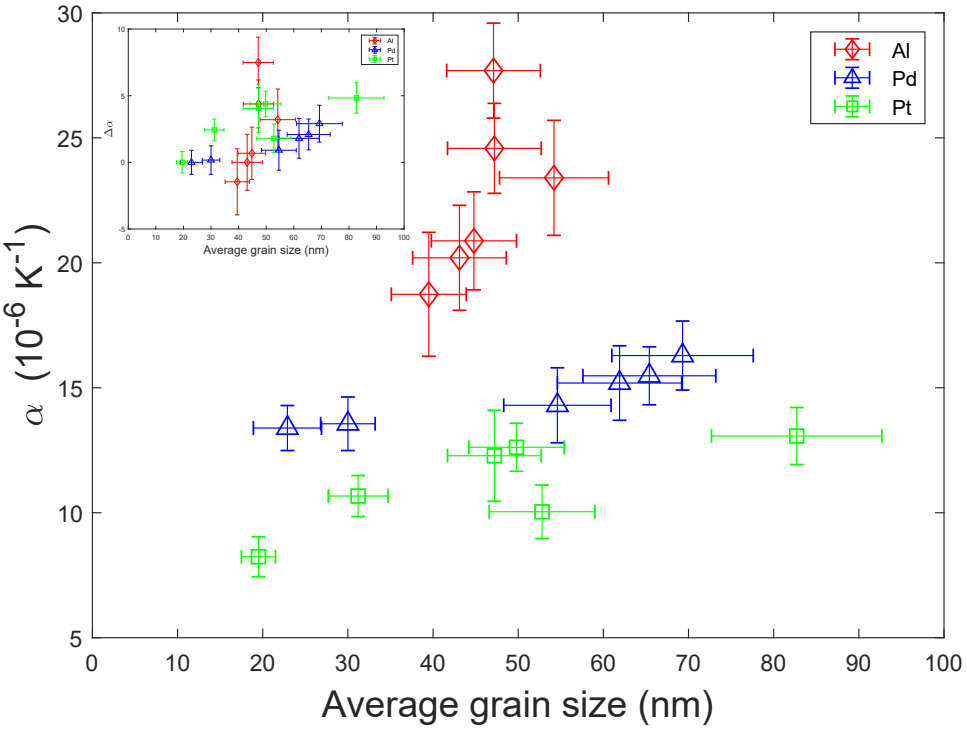

Figure 5.12: The change in the coefficient of thermal expansion with average grain size. The change in α (inset) shows that the Pd and Pt films tend to show an increase average grain size as α increases.

of the film, which could in turn affect the as-deposited values of the mechanical properties. We are not aware of a consistently reported set of deposition parameters commonly used by groups fabricating MOS QDs. Additionally, we note that in this work, we have chosen to study films roughly a factor of two thicker than the average grain size. The films are roughly 60 nm thick and the typical as-deposited grain size is less than 30 nm. This suggests that our analysis via the SEM imaging may not fully be capturing the magnitude of the grain structure changes. This may explain why we see little to no change in the average grain size in Al films, but we see significant changes in the mechanical properties of the film.

5.6 Simulations of strain induced modulation

The significant film stresses present in the gate metals combined with the deviations of α from the bulk value with annealing lead to strain induced modulation of the conduction band that differs significantly from the expectation based on bulk properties. To illustrate the impact of the observed behavior in σ and α_{film} with anneal temperature, we perform finite-element modeling of the mechanical impact on the conduction band of silicon using COMSOL, similar to the simulations outlined in chapters 2 and 4. Here, we simulate a single 100 nm wide, 60 nm thick gate on a 25 nm thick SiO_2 layer to mirror the stack for the measurements presented except for the lateral (100 nm) width. For Ti/Pd and Ti/Pt gates, we simulate gates composed of only Pd and Pt respectively as the Ti layer is too thin to measure using wafer curvature methods. Given the relative thicknesses of the Ti layer to Pd and Pt

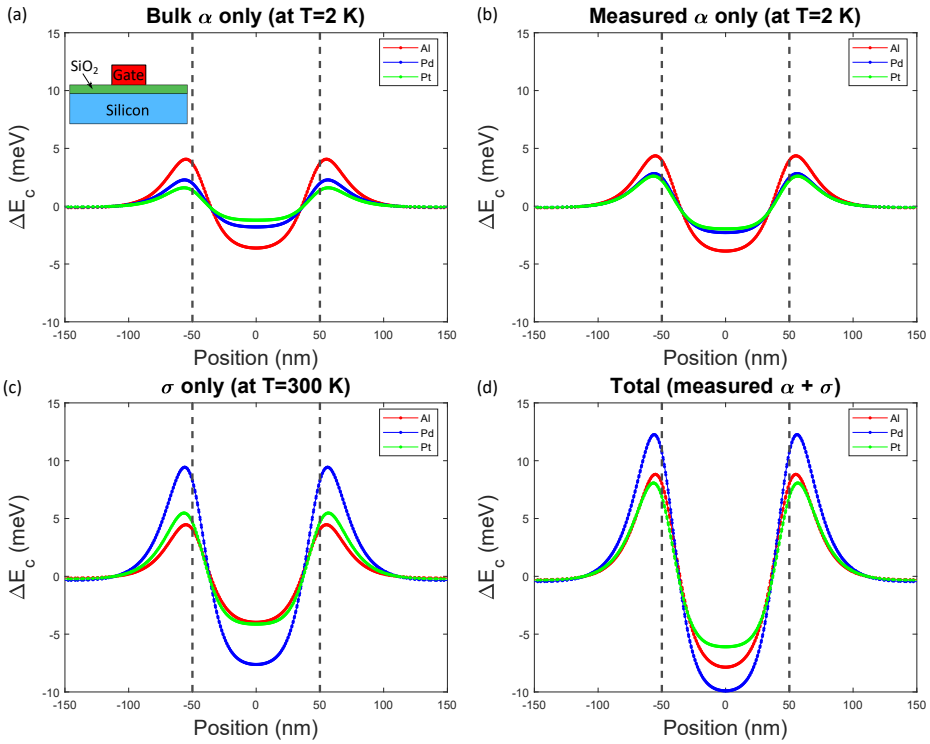

Figure 5.13: Simulated conduction band modulation (ΔE_c) due to strain for 100 nm wide metal gate on top of 25 nm SiO_2 on a silicon substrate. (a) ΔE_c due only to the strain generated from the coefficient of thermal expansion (α) mismatch of the gate materials in cooling down from 300 K to 2 K based on the bulk values for each film. The inset shows a diagram of the simulated structure. (b) is the same as (a) except we have instead used the measured α values from this manuscript. For each metal, the α used in the simulation is the measured value after a forming gas anneal at 350 °C in 5 % H_2/N_2. (c) ΔE_c due only to intrinsic film stress, σ, for each gate material after a forming gas anneal at 350 °C in 5 % H_2/N_2. (d) Total ΔE_c due to both α and σ. The dashed lines represent the edges of the 100 nm wide metal gate. Due primarily to the significant difference in intrinsic film stress, the Pd films show larger potential modulation than Al and Pt devices.

layers, it is reasonable to expect that the measured α_{film} and σ are dominated by the Pd and Pt layers. We note that the shape and magnitude of the strain profile is dominated by the edges of the gate and that dimensions simulated here are sufficient for comparing different metals as changing the width or thickness of the gate will not change which material has the lowest ΔE_c. Fig. 5.13 (a) and (b) show the change in the conduction band, ΔE_c, generated in the silicon substrate 2.5 nm from the SiO_2-Si interface from strain originating from α-mismatch between the metal gate and the silicon when cooled to 2 K. For Fig. 5.13 (a), we have used the bulk value of α_{film} and in Fig. 5.13 (b) we use the value of α_{film} measured after the 350 °C anneal in 5 % forming gas. Al shows the largest ΔE_c in agreement with the magnitude of α_{film}. Fig. 5.13 (c) shows ΔE_c, due only to σ after the same anneal. In this case, Pd gives the largest ΔE_c, reflecting the much larger value of σ. The combined effect of α_{film} and σ is shown in Fig. 5.13 (d), with Pd showing the largest ΔE_c followed by Al and Pt. Thus, the expectation that moving to Ti/Pd gates from Al gates will reduce the mechanically induced ΔE_c is not supported in this work.

5.7 Implications for QD design and fabrication

The results from the annealing dependence of the oxide charge defects suggest two important considerations for MOS QD fabrication. First, the optimal anneal appears to be around 350 °C based on the minimization of D_{it}. Second, from the annealing dependence of Q_f, we can see that the contributing defects are not due solely to D_{it} since the change in Q_f is much greater than the change in D_{it}.

Additionally we can see that Pd shows a different annealing dependence than both Al and Pt, while all three show relatively similar behaviors in D_{it} with annealing. Based on the discussion in chapter 3, we can infer that the annealing dependence of Q_f is likely to be dominated by trapped oxide charge, Q_{ot}, created by ionizing radiation during the e-beam evaporation. The difference between the three metals in terms of Q_f suggests that Q_{ot} in the oxide must be different in some significant way for each metal. Consider the case of the Al MOS capacitors. Recall from chapter 3, the Q_f value that we measure here is a net charge value. Here, we can say that Al leads to a lower net charge in the oxide than Pd and Pt, but the trend of the Al Q_f becoming increasingly negative with increasing anneal means that the total Q_f must be larger in Al. This is an important consideration in comparing the performance of Al gated MOS QDs versus Pd-gated since a higher total number of charges is more likely to lead to unintentional QD formation than a higher net charge [158]. Unfortunately, from the present data alone, we cannot confirm the above speculation further.

Conversely, the results from the annealing dependence of the mechanical properties do not show any sort of minimization behavior like that seen for D_{it}. This suggests that tuning the annealing process is not a good method for minimizing the impact of strain in MOS QD devices, since the minimization actually occurs in the as-deposited state. Additionally, we find that the deviations of the thin film coefficient of thermal expansion from the bulk seen in chapter 4 are not due to the forming gas anneal. This means our best path towards reducing such deviations would be to explore adjusting the deposition process instead. If minimization of

the strain impact is not the goal, then the results presented here open more possibilities. Our results show that, in general, σ is a strong function of the anneal temperature for all the metals. If the goal of the QD device design was to intentionally increase the stress induced by gates with the intention of forming strain-induced QDs, then adjusting the forming gas anneal temperature could be a potential optimization method. However, we must consider the results above in the context of those from chapter 4, where we achieved agreement between measurements and simulation without considering σ in the tunnel junctions. At face value, this means that the value of σ measured via wafer curvature methods is not translating well down to nanoscale patterned features or that the stress evolution in the film with annealing is different for pattern structures opposed to blanket films.

The above results show that we must consider the fabrication process of silicon MOS QDs holistically. The choice of gate material, deposition parameters, and anneal parameters impacts at least lithographic fidelity, threshold voltage, defect densities, strain and perhaps more properties of QD devices. The evolution with anneal temperature we have presented here makes it clear that the negative effects of charge defects and mechanical effects cannot be simultaneously minimized. This forces a significant choice for researchers to make in fabrication process design. This choice is displayed most clearly in the observation that, of the gate materials studied, Pt has both the smallest ΔE_c and the largest overall Q_f when annealed to minimize D_{it}. Of the materials studied, Al appears to be best with respect to a minimized D_{it}, a smaller magnitude of Q_f, and very nearly so considering ΔE_c from strain. It also affords a convenient inter-gate dielectric. However, the native oxidation of

Al gates may itself lead to a negative impact on QD performance with increased noise [159, 160] and from the distorted shape of the gate [138] as well as its direct mechanical-induced effects on the conduction band, which have not been studied. Were it not for the larger grain size in the deposition it could remain the clear choice among the materials studied.

The magnitude of charge defects, α-induced and σ-induced stress are highly dependent on the choice of deposition process, anneal temperature, and material. These results indicate three potential paths in MOS QD fabrication processes to move beyond the issues discussed above. First, e-beam deposition parameters should be evaluated with respect to D_{it}, Q_f, α, and σ similarly to the present study. Second, workers should explore alternative metal deposition techniques other than thermal or e-beam evaporation. For instance, we could use low energy sputtering techniques, like ion beam deposition, to produce films with different grain structures, which could modify mechanical effects while minimizing deposition-induced defects. Third, the observation above that Pt has the lowest strain modulation contrasted with the highest charge defect density motivates the exploration of gate materials outside of the three we presented here and those typically used in other work. This might include metals like TiN, which has recently been used in QD fabrication [161, 162]. It is clear from our work that the choice of material itself is a significant factor in determining the final mechanical and electrostatic properties and that the mechanical impact is not immediately obvious based solely on commonly used bulk mechanical values.

Chapter 6: Measurement of asymmetric tunnel junctions

In chapter 4, we demonstrated a measurement of strain using tunnel junctions (TJ) fabricated with aluminum and titanium gates. There, the measurement of strain was a relative difference in strain *between devices* derived from the average difference in the measured barrier heights. Using this measurement technique, we were able to demonstrate a relative difference in barrier heights that was consistent with the measured coefficient of thermal expansion mismatch in the Al and Ti films. In this chapter, we build upon the framework from those results to develop an alternative method for measuring strain from our TJ devices. In the previous measurements, all of the TJ were fabricated with same gate material on each side of the barrier. We refer to these as symmetric TJs. In these new TJ devices, we intentionally make the barrier asymmetric by putting different gate materials on either side. The goal in this new design is to extract a strain difference between the gates using the barrier asymmetry rather than the barrier height. Here, the major advantage is that we can avoid comparing separate devices and extract the strain difference from a single device. This significantly reduces the burden of minimizing device-to-device variations that hampers the relative barrier height measurements and allows measurement in a single cool-down. In this chapter, we will design,

fabricate, and measure asymmetric TJs with gates composed of Al and Ti/Pd. Using the results of the forming gas anneal optimization from chapter 5, we adjust the TJ fabrication procedure in order to reduce the role oxide charge defects play in the device transport. In contrast to the TJ results presented chapter 4, we are unable to identify any new TJs that are free from unintentional quantum dots (UQD) and are thus unable to proceed with extracting a strain measurement. While the asymmetric TJ measurements we show in this chapter were impeded by UQDs, we are still confident that the technique is a viable path towards the electrical measurement of strain assuming we can identify and reduce the sources of UQDs in future devices.

6.1 Simulations

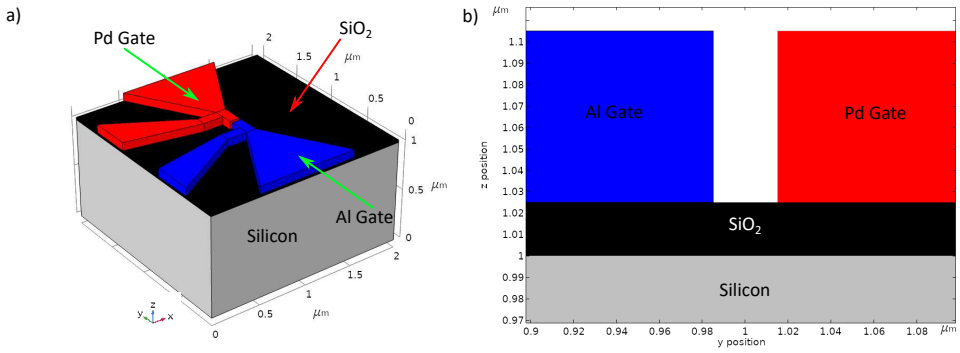

Figure 6.1: (a) Image of the 3d tunnel junction device model used for simulating strain in COMSOL. The device is shown here with a 25 nm thick gate oxide, 30 nm wide tunnel gap, and 100 nm wide gates. Here, the right gate is Pd and the left gate is Al. (b) 2d slice of the model in (a) along the center of the tunnel junction.

The strain induced modulation of the asymmetric TJ devices can be simulated

using finite-element modeling (FEM) in COMSOL in the same way as in chapter 4. The simulations here are identical to those performed in chapter 4, with the only major difference being that the gate materials on either side of junction are now different. Fig. 6.1(a) shows an image of the 3d model used in our COMSOL simulations where the red gate represents Pd and the blue gate Al. For these simulations, we have used a room temperature measured value of α for all materials. For the materials used in this work, α rapidly goes to zero at temperatures below 30 K [79]. Therefore, we expect that the use of a room temperature α in this simulations means that we will overestimate the change in barrier height. This choice is made because our method for measuring α takes place near room temperature and we do not have a method to measure α at cryogenic temperatures. Fig. 6.1(b) is a 2d cut along the center of the 3d TJ model and shows the gap region where the barrier will be formed. The TJ shown has a 30 nm wide gap.

We show the results of the COMSOL simulations in Fig. 6.2 for two different sources of strain in the silicon: 1) only the coefficient of thermal expansion (α)-induced strain (Fig. 6.2(a)) from cooling the device down to 2 K, 2) from intrinsic film stress (σ)-induced strain only (Fig. 6.2(b)). Fig. 6.2(c) shows the total combination of α and σ. For the values of α and σ, we have used the measured values from chapter 5 for the films after 350 °C for 30 minutes in 5% H_2/N_2, which we used when fabricating the asymmetric TJs. These results show the same qualitative behavior similar to symmetric TJs presented in chapter 2 with the major exception that the strain profiles are no longer symmetric about the center of the gap between the gates.

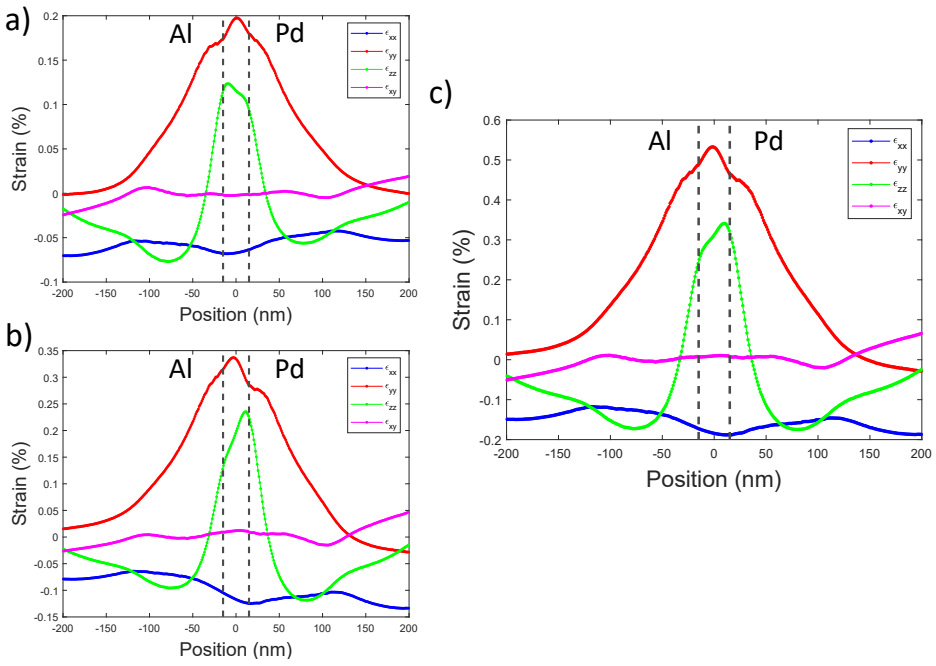

Figure 6.2: Plots of various strain components for Al/Pd asymmetric TJ device with 30 nm wide gap. The dashed lines in each plot denote the edges of the gates in the device. (a) Plot of the coefficient of thermal expansion mismatch (α)-induced strain from cooling a TJ device from 300 K down to 2 K for using the α measured for Pd and Al after a 350 °C anneal. (b) Plot of the intrinsic stress (σ)-induced strain using the σ measured for Pd and Al after a 350 °C anneal. (c) Combined effects of α and σ from (a) and (b) respectively.

Fig. 6.3 shows the resulting conduction band modulation (ΔE_c) calculated using the deformation potentials from chapter 2 and the strains presented in Fig. 6.2. Here, again, we show three cases from the COMSOL simulations: ΔE_c based solely on the α-induced strain (blue data), ΔE_c from σ-induced strain only (red data), and ΔE_c from the total combination of α and σ (green), where we have used the measured values from the films in chapter 5. In comparison with the strains in Fig. 6.2, we can see that ΔE_c is dominated by the strain perpendicular to the MOS interface (ϵ_{zz}) consistent with the results in chapter 4. We also identify a shift in

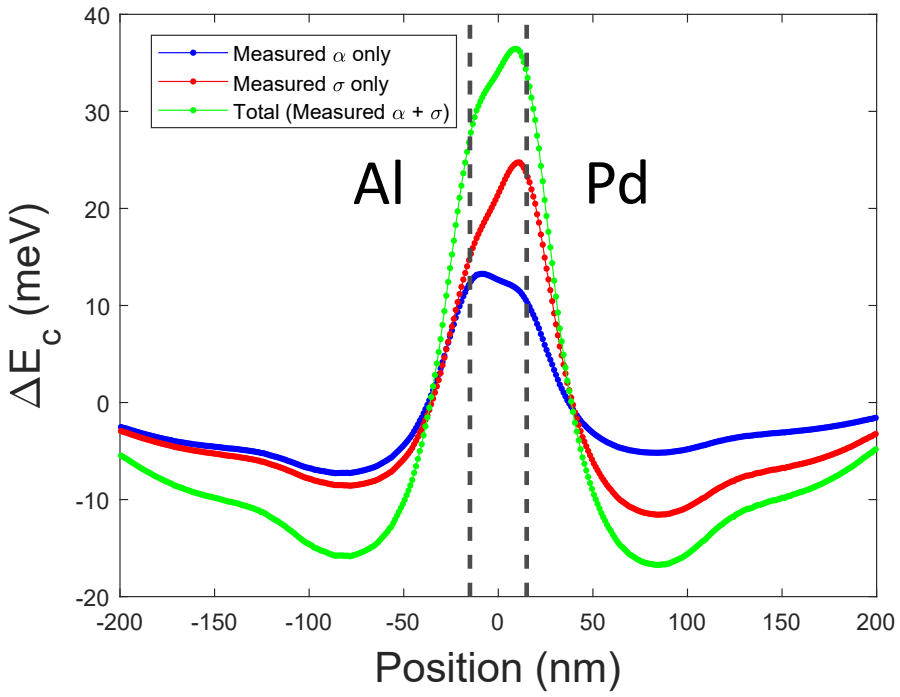

Figure 6.3: Modulation of the conduction band (ΔE_c) calculated using the deformation potentials given in chapter 2. ΔE_c is calculated for three different cases from Fig. 6.2. Blue data from Fig. 6.2(a) for the coefficient of thermal expansion (α)-induced strain. Red data from Fig. 6.2(b) for the intrinsic film stress (σ)-induced strain. The green data for the combined effect of α and σ from Fig. 6.2(c). The dashed lines in all plots indicate the edges of the gates in the device.

the asymmetry from Fig. 6.3 between the different data sets shown. The blue data only considers α-induced strain so the gate-induced strain barrier height is higher on the Al side because $\alpha_{Al} > \alpha_{Pd}$. The red data only considers σ-induced strain so now the asymmetry has flipped with the barrier height now higher on the Pd side since $\sigma_{Al} < \sigma_{Pd}$. The green data, which considers both contributions, shows the same asymmetry as the σ-only data. This is due to the large σ value in the Pd films after annealing in chapter 5. Recall from chapter 4 that in order to explain the relative barrier height difference between the Al and Ti devices, we only considered

the α-induced component. From the above, we can draw a simple expectation for the behavior for the asymmetric TJs. In the event that σ does not matter again in our measurements, we will expect to have an asymmetry such that the Al side of device has a higher barrier height by roughly 1.8 meV. In the context of the BRD model [139], this corresponds to a $\Delta\phi \approx +1.8$ meV if the ohmic on the Al side is positively biased. Alternatively, we note that if we consider the effect of σ in these new devices then we expect to see a reversal in the sign of asymmetry. In fact, the red data in Fig. 6.3 shows the barrier height on the Pd side is roughly 5.6 meV higher than the Al side. This corresponds to a $\Delta\phi \approx -5.6$ meV if the ohmic on the Al side is positively biased. In both cases, the levels of asymmetry should be measurable as typical limits on the uncertainty of the fitted values of $\Delta\phi$ are in the few 100 μeV range.

6.2 Device Fabrication

In chapter 4, we established a baseline TJ device fabrication process that we used measure a strain difference between Al- and Ti-gated devices. For the asymmetric TJs that are the focus of this chapter, we have adjusted the fabrication process slightly. Here, we only detail the steps that have changed and how we proceeded with the e-beam lithography for these asymmetric TJs, which requires that we have good layer-to-layer alignment in order to achieve the necessary sub 40 nm gap lengths.

6.2.1 Changes to fabrication from previous run

For the asymmetric TJs presented here, we have made three major changes to the fabrication process: 1) The thick field oxide used for electrical isolation has been replaced by making the Al bond pads around 1.3 μm thick, 2) The gate oxidation process has changed due to the furnaces at NIST being replaced, and 3) We have changed the forming gas anneal from 425 °C to the lower temperature anneal at 350 °C found from the work in chapter 5.

The first change to remove the field oxide and make the metal from bond pads thicker was driven by some failures with the field oxide. The major driving force for this change was to allow improvements to the electron beam lithography, where the presence of the field oxide causes the PMMA 495 A4 underlayer to be thicker than intended in the TJ patterning region limiting the gap lengths. Additionally in previous wafers, we had cases were the 120 nm thick field oxide was beginning to leak (>1 nA at 1 V on the gate when the device was at $T \approx 2\ K$) after wire bonding. This leakage, while not frequent enough to ruin all of our attempted measurements, was large enough to impede any measurements of the TJ properties because typical zero bias currents are less <500 pA. We found that making the Al bond pads in excess of 1 μm provides a similar level of leakage protection without the complicating factors of having to grow and etch an oxide on the wafer prior to the gate oxide growth and leaving the surface of the wafer flat for the electron beam lithography patterning.

The second change that was made in for the asymmetric TJs was to the gate

oxidation process. We have kept the same 25 nm thickness that was used in the TJ measurements in chapter 4 and the MOS capacitor measurements made in chapter 5. Similar to the the MOS capacitor measurements in chapter 5, we grow a roughly 25 nm thick gate oxide in a dry oxidation furnace at 1000 °C for 22 minutes with a 10 min post oxidation anneal performed in N_2 at the oxidation temperature, whereas the symmetric TJs were grown at 950 °C. We made this change not for any specific scientific reason but due to the cleanroom at NIST changing to a new furnace stack between our initial TJ device fabrications and the asymmetric ones. In addition to the temperature change, the oxides in the asymmetric TJ devices were grown in a pure O_2 ambient, while the oxides presented in chapter 4 were grown in an ambient with Cl (the new furnaces are unable to run Cl processes). Oxidation with Cl in ambient is typically used for gettering of impurities, but it also has the effect of increasing the oxide growth rate. The third and final change was the shift in the forming gas anneals process. For the TJ measured previously, we used a forming gas anneal at 425 °C for 30 minutes in 10 % H_2/N_2. Based on the results from chapter 5, we found that the interface trap density, D_{it}, was minimized using an anneal at 350 °C for 30 minutes. For the asymmetric TJs, we used this anneal in 5 % H_2/N_2 in order to minimize the impact of defects on the device. As discussed in chapter 5, this comes with the tradeoff that the intrinsic film stress is increased significantly in the gate materials. Nominally, this is not a problem for our devices, as more strain in the device should only increase the size of the measured asymmetry so we consider attempting to lower the contribution of D_{it} as the optimal path for these devices.

6.2.2 Alignment to control gap length

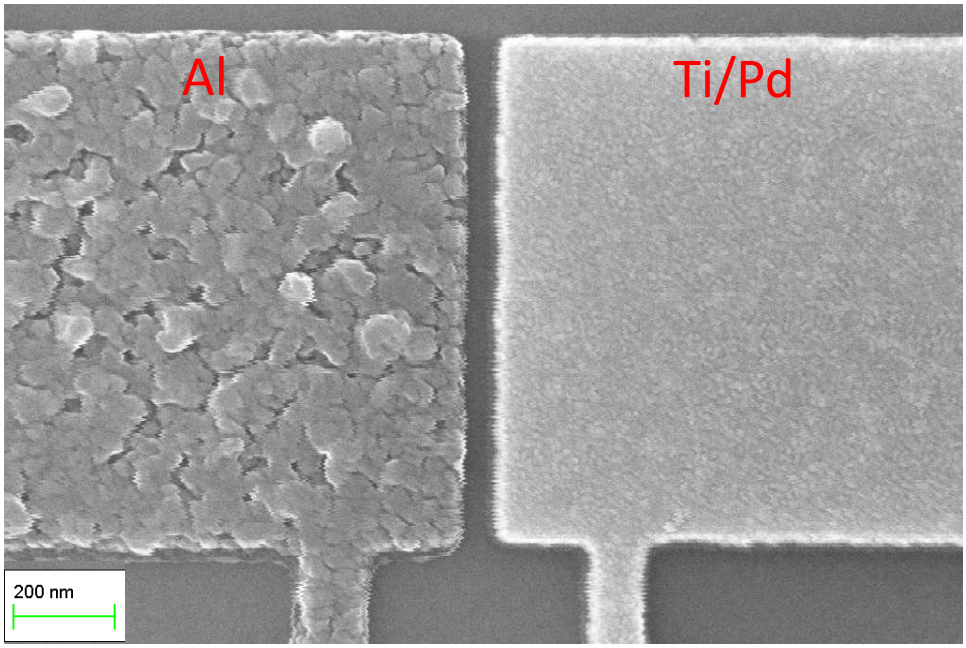

Figure 6.4: SEM image of an asymmetric TJ device. The left gate is Al and the right gate is Ti/Pd. Both gates are roughly 60 nm thick. The width of the gates in this device 1 μm.

In the previous TJ devices, we patterned both sides of the TJ in a single electron beam (e-beam) lithography write and liftoff step. The size of the gate width and gap length was set purely by the dosing and proximity effects in the resist stack. In the asymmetric TJs, we are not able to write both sides in a single step because we need to deposit two different metals. Therefore we must rely on the ability to align two different layers to set the tunnel gap length. The Jeol e-beam lithography system (6300-FS) typically achieves about 20 nm alignment accuracy between layers. Since the accuracy is random and the average value is a sizeable fraction of the gap length, it presents a problem in fabricating these devices at the

desired gap lengths (< 40 nm). This method has an advantage in that we should be able achieve smaller gap lengths (< 15 nm) [163] than before because we are no longer limited by a narrow piece of resist surviving the deposition and liftoff process intact. The tradeoff is that the gap length will have a random distribution of sizes for a given designed gap length. Therefore, we ultimately have less control and reproducibility over the feature sizes we write using this method.

Fig. 6.4 shows an SEM image of one of the asymmetric TJs patterned in this way. Here, the gate on the left side is Al and the right gate is Ti/Pd. As with all the gates in previous devices, all of the metals have been electron-beam evaporated with total film thicknesses around 60 nm. Additionally, the device in this SEM image has gates that are roughly 1 μm wide. This highlights another advantage of this new method for TJ patterning. In the previous method, we could not write any gates with widths greater than 500 nm and still achieve tunnel gap lengths less than 40 nm.

In order to assess how misalignment errors affect what TJ gap lengths we can achieve, we wrote a series of test devices with different designed gap lengths ranging from 0 to 70 nm. After patterning and liftoff of both metal gates, we use a SEM to measure the actual gap length in the TJ device. Fig. 6.5 shows the resulting measured gap lengths for our asymmetric TJs. In Fig. 6.5(a), we plot the statistics of the difference between the measured and designed gap lengths. Here, we can see there is a fairly wide spread in the distribution but the peak is roughly around -20 nm meaning on average any designed gap length is most likely to come out 20 nm smaller. In Fig. 6.5(b), we plot the measured gap length against the designed

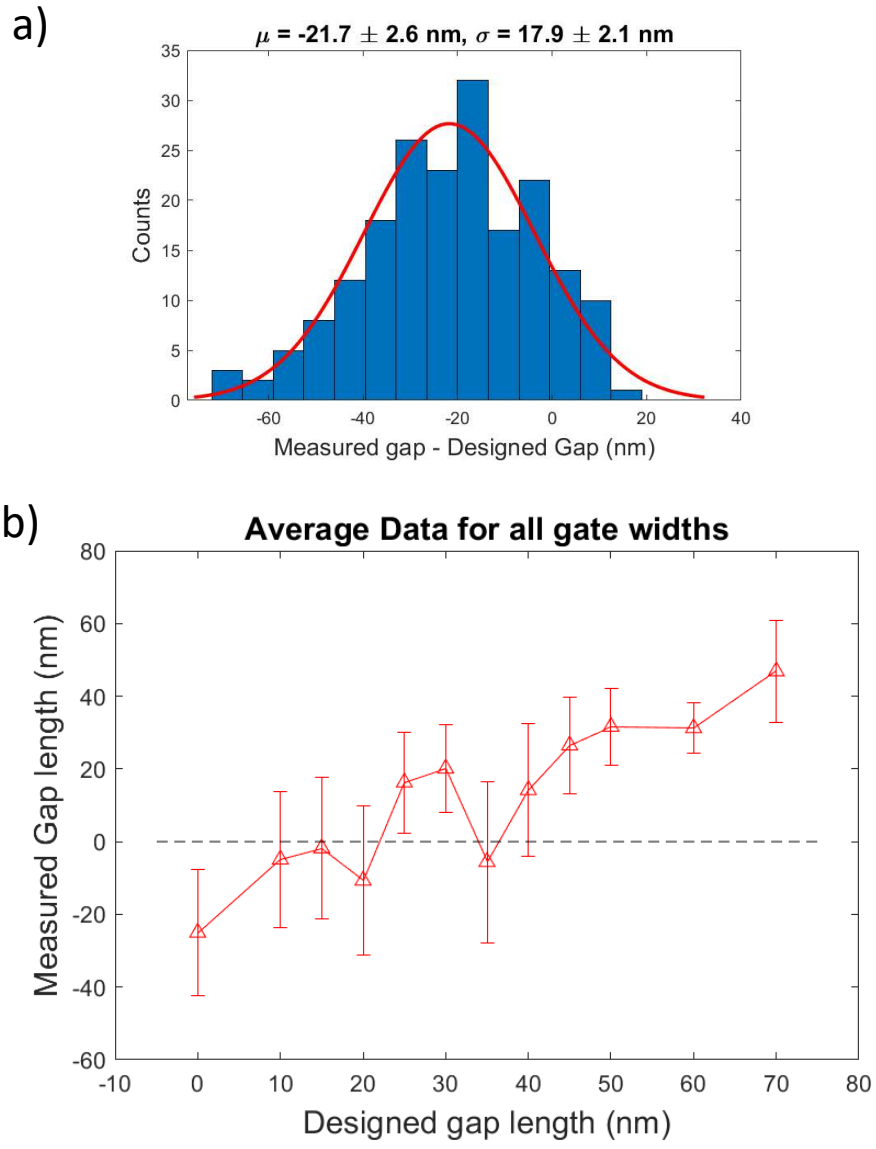

Figure 6.5: Measurements of the tunnel gap widths obtained via alignment in a Jeol 6300-FS electron beam lithography tool. (a)Statistics of the difference between the measured gap length and the designed gap length. (b)Plot of the measured gap length vs the designed gap length.

length. From this, we can identify the optimal range to write our asymmetric TJs, which corresponds to designed gaps in 30-45 nm range.

6.3 Low temperature electrical measurements

For all the low temperature measurements presented here, we use the same 2 K closed cycle cryostat and electrical setup for DC measurements as in chapter 4. Additionally, while we made changes to device design and fabrication, we did not make any changes to the 4-terminal layout of the devices. As with previous TJs, we have 6 contacts per device: 4 ohmics (2 current contacts, 2 voltage probes) and 2 gates which are now distinguished by being either Al or Ti/Pd. In this way, we can still use these contacts to measure threshold voltage (V_T) on either side of the junction by using current paths involving the voltage probe ohmics. We show an example of the turn-on behavior in Fig. 6.6. The different colored paths correspond to the data colors in Fig. 6.6(a), where, for example, the blue path is measuring current from the S to PL ohmics while sweeping the Al gate voltage. Importantly, the red and blue paths do not measure any current through the TJ itself just the adjacent paths. This gives us an independent measure of V_T for each gate material, signifying when a strong inversion layer is formed under each gate. As can be seen in Fig. 6.6(a), there is a significant difference in V_T of roughly 0.445 V between the Al($V_T = 0.514 \pm 0.003$ V) and Ti/Pd($V_T = 0.959 \pm 0.008$ V) gated sides of the device.

The green path represents transport through the barrier. In Fig. 6.6(a),

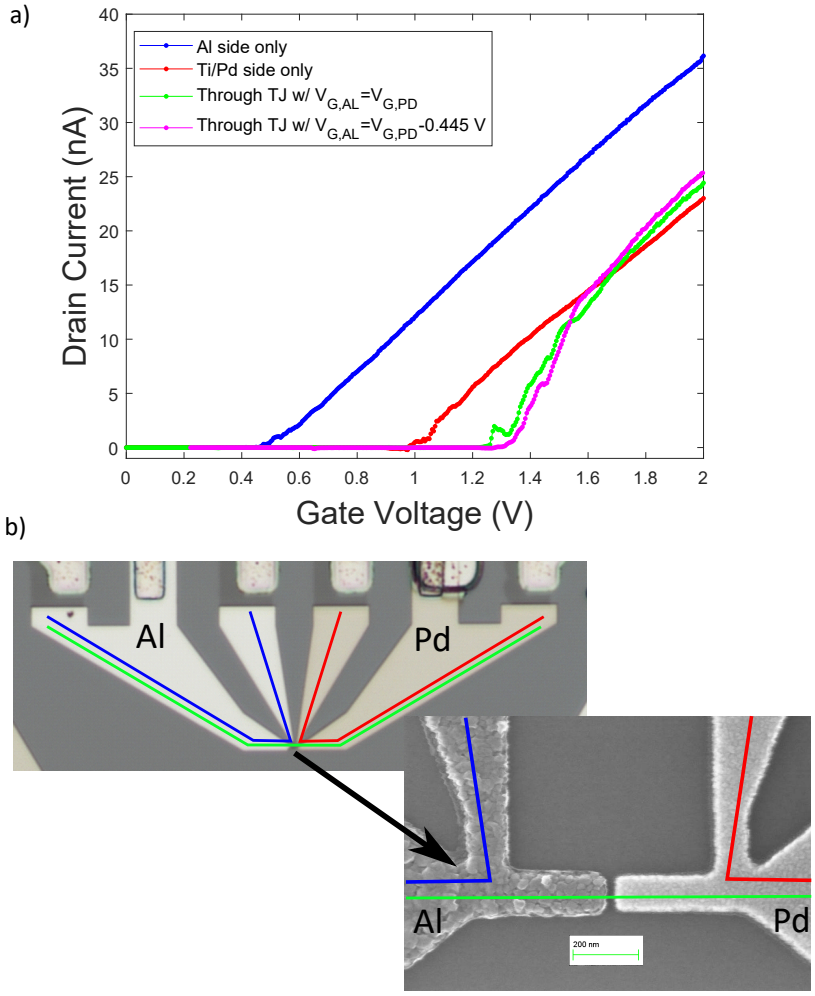

Figure 6.6: (a) DC Turn on for a asymmetric TJ device at 2 K using different combinations of ohmics and gates taken with a constant source-drain bias (V_{SD}) of +500 μV applied to the drain contact. (b) Optical and SEM images of a TJ junction device showing the different colored paths corresponding to the turn-on data in (a). Here Al and Pd are gates and S, D, PL, and PR are ohmics. The red(blue) path measures current from the S (D) ohmic to the PL (PR) ohmic using the Al (Pd) gate for turn on and avoids any transport through the tunnel gap. The green path measures current from the S to D, which is transport through the TJ. For the green data in (a), the gate voltages for the Al and Pd gates were set to be equal ($V_{G,AL} = V_{G,Pd}$). For the pink data in (a), we are again measuring current through the green path but in this case the gate voltages for the Al and Pd gates were not equal and corrected for the measured differences in threshold voltage ($V_{G,AL} = V_{G,Pd} - 0.445$ V). For the green and pink data, the x-axis is the average of the two gate voltages applied.

the green data represents transport through the tunnel junction when $V_{G,AL} = V_{G,PD}$. From this we can see the delay in conduction through the green path that we indetified in chapter 4 as the presence of the tunnel barrier. It is important to note that this $V_{G,AL} = V_{G,PD}$ condition ignores the difference in threshold voltage expected for transistors made with different gate materials due, at least, to work function differences. Each gate should be swept such that $V_{G,AL} - V_{T,AL} = V_{G,PD} - V_{T,PD}$ is maintained. This condition is satisfied in the pink data in Fig. 6.6(a), where we are measuring transport through the green TJ path with $V_{G,AL} = V_{G,PD} - 0.445$ V.

Turn-on behavior discussed above for the asymmetric TJs is qualitatively similar to the TJs presented in chapter 4 once we account for the differences in threshold voltage between the Al and Ti/Pd. Recall also from chapter 4, that we need to screen our TJ devices to ensure that there is little to no evidence of any transport signatures of unintentional quantum dots (UQDs) in the device. We can see in Fig. 6.6(a) that there are slight hints of oscillations in the current as function of gate voltage. This non-monotonic behavior is inconsistent with a simple tunnel barrier and suggests the present of UQDs in the device. In order to fully investigate the presence of UQDs, we take 2d transport data, current as a function of gate voltage (x-axis) and source-drain bias (y-axis). Fig. 6.7 shows some representative samples of the asymmetric TJs measured in this work. Fig. 6.7(a) is the 2d transport data for the device in Fig. 6.6. Here, we note the slight presence of coulomb diamonds on top of the funnel-like shape we normally expect in the current through the TJ. This confirms the presence of at least one UQD in the device that is contributing

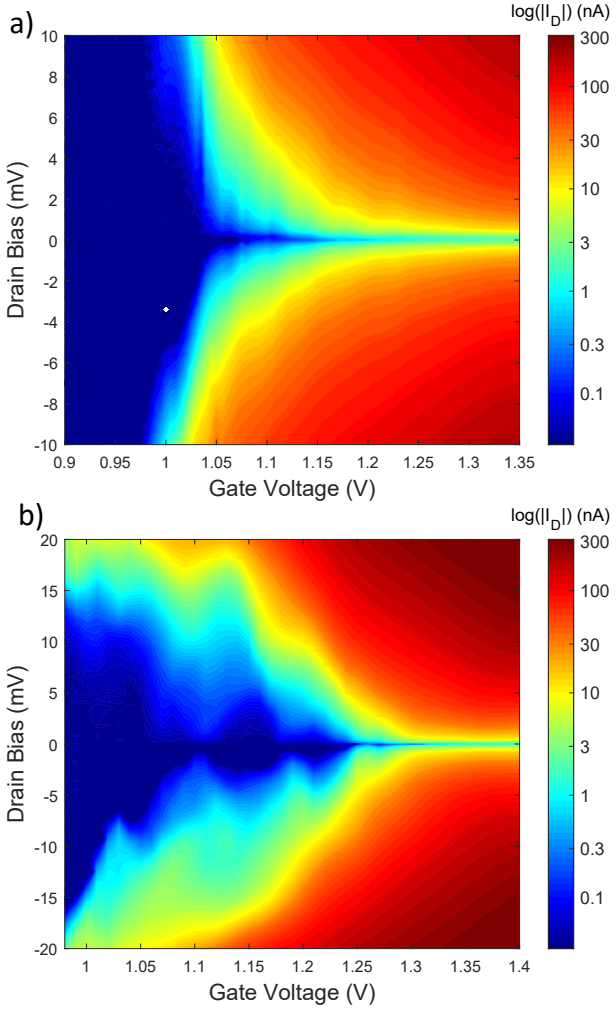

Figure 6.7: Examples of 2d DC transport data for representative asymmetric TJs. (a) 2d data from the device that shows the least disorder of all the measured asymmetric TJs. This is also the same device whose turn-on data is shown in Fig. 6.6. The width of the gates for this device is roughly 100 nm. (b) Data from a TJ device showing the presence of the multiple unintentional quantum dots (UQDs). This data is representative and consistent with issues faced with many of the other asymmetric TJs measured in this work. The width of the gates for this TJ device is roughly 1μm.

to the transport in addition to TJ. Importantly, devices which exhibit UQDs break our assumptions of a single barrier through which transport occurs (there must be two barriers and a well to see Coulomb blockade oscillations). We, therefore, opt to not proceed with barrier parameter extraction through the BRD model. We note that this device is actually shows the least amount of disorder related behavior out of all the devices from this fabrication run. Fig. 6.7(b) shows a different device that is more typical of the other devices we measured in this work. In this device, the blockade from UQDs is significantly more obvious and it is likely that there is more than one dot in the device.

This increased frequency of UQDs in the asymmetric TJs compared to previous results suggests two possible issues that need to be considered: 1) The strain is behaving differently then we expected in these TJs and the UQDs are an unavoidable feature of strain and 2) fabrication process in the new TJs has lead to more defects in the oxide, which in turn has caused more UQDs to form. We will discuss these issues in the proceeding sections. Importantly, the conclusion from this run of TJ measurements is that we are unable to extract any barrier asymmetry in our devices due the presence of UQDs.

6.4 MOS capacitor measurements

To investigate if oxide charge defects in our asymmetric TJs are the cause of the UQDs mentioned above, we fabricated MOS capacitors concurrently with the TJ devices. We designed these MOS capacitors to act as process monitors for the

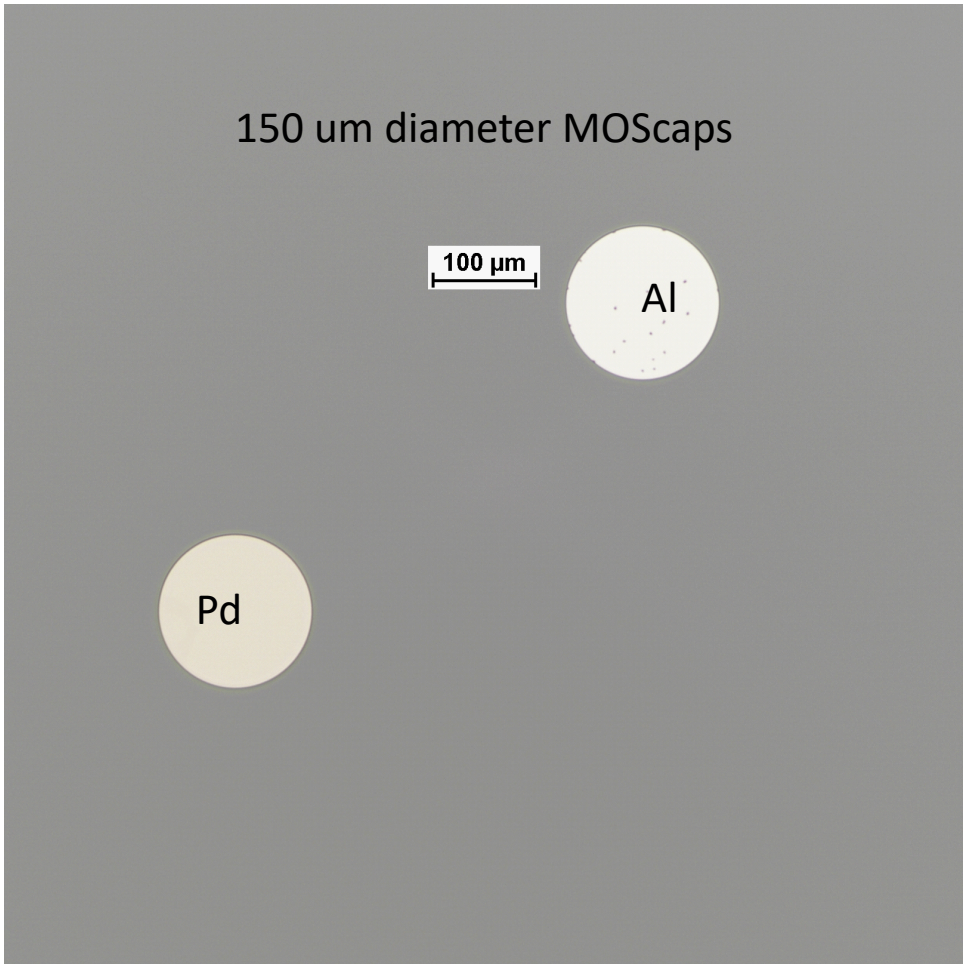

Figure 6.8: Optical image of diagnostic MOS capacitors fabricated on chip with tunnel junction devices. These MOS capacitors are patterned in the same Electron Beam Lithography (EBL) step as the gates for the tunnel junctions in this work.

TJ device fabrication process, so they were patterned using the same electron beam lithography (EBL) step as the TJs rather than using optical lithography as we did with all of our previous MOS capacitors. This gives a more accurate picture of the charge defects in the devices since we have now included the damage done to the oxide by the EBL [115] rather than just the damage from the metal deposition [164]. These MOS capacitors are written at electron beam doses typically used in the TJ device lithography which ranges anywhere from 1300 μC/cm^2 to 2100 μC/cm^2 . For this reason, these MOS capacitors are fabricated on the same 1 cm^2 chips as the TJ junctions with two MOS capacitors total: one Al gated and one with Ti/Pd gates. An optical image of one these sets of 150 μm diameter MOS capacitors is shown in Fig. 6.8.

We perform capacitance vs voltage (CV) measurements on the EBL patterned MOS capacitors following the methods outlined in chapter 3. The average defect densities for the fixed charge density, Q_f, and the interface trap density, D_{it}, extracted are shown in Fig. 6.9 where 'Al EBL' and 'Ti/Pd EBL' refer to EBL patterned MOS capacitors with Al and Ti/Pd gates respectively. Additionally, we compare these results against the MOS capacitor results from chapter 5 where the major difference between these different sets is that the devices from chapter 5 were patterned with optical lithography. These results are labeled 'Al Optical' and 'Ti/Pd optical' in Fig. 6.9 for Al and Ti/Pd gates respectively. Importantly, the MOS capacitors we are comparing here were fabricated using the same oxidation and forming gas anneal processes. This means any deviations should be due the extra damage done by the EBL processing assuming all of our other processes are

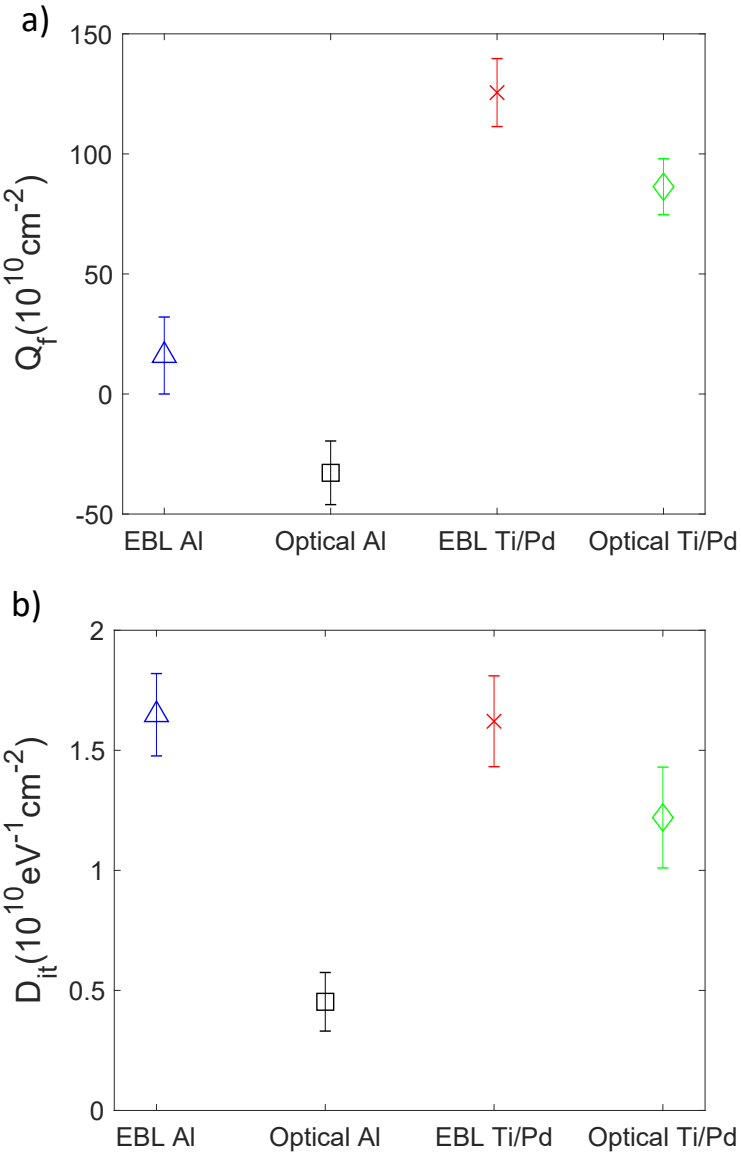

Figure 6.9: Comparison of defect densities: (a) measured fixed charge density, Q_f and (b) average interface trap density, D_{it}, measured on diagnostic electron beam lithography (EBL) MOS capacitors fabricated alongside the asymmetric TJs. Labels 'Al EBL' and 'Ti/Pd EBL' are for MOS capacitors with Al and Ti/Pd gates respectively. Also shown are defect densities measured on the optical lithography patterned MOS capacitors from the study in chapter 5 with otherwise identical processing, labeled 'AL Optical' and 'Ti/Pd Optical' for MOS capacitors with Al and Ti/Pd gates respectively. We write the EBL MOS capacitors with areal dosages comparable with those used in the TJ device patterning. These doses ranged from 1300 μC/cm^2 to 2100 μC/cm^2 .

nominal. Fig. 6.9(a) shows a comparison of the average Q_f for these MOS capacitors. Interestingly, we observe nearly identical shifts in Q_f from optical to EBL MOS capacitors for both metals, where $\Delta Q_{f,Al} = +(48.9 \pm 16.5) \times 10^{10} cm^{-2}$ and $\Delta Q_{f,Pd} = +(39.2 \pm 17.6) \times 10^{10} cm^{-2}$. This positive shift in Q_f is expected for the damage done to silicon dioxide from electron beam exposure [165]. This suggests two important results. First, it is reasonable to assume that the shift is solely due to the EBL induced damage and all other processes are nominal. Second, the damage done by EBL is still significant even after a forming gas anneal at 350 °C. This fact is also shown in the D_{it} results in Fig. 6.9(b). Here, we find that both the Al and Ti/Pd EBL MOS capacitors show nearly identical levels of D_{it}. Conversely, and as was shown in chapter 5, our optically patterned MOS capacitors show different levels of D_{it} with Al gates reaching significantly lower values than Ti/Pd gates. This suggests that the EBL induced damage is increasing D_{it} significantly and possibly that the minimization of D_{it} from EBL damage might not occur under the same conditions as that from metal deposition damage. Nevertheless, these deviations alone do not allow us to conclude that the apparent increased frequency of UQDs in our asymmetric TJs as compared to symmetric TJs is due to an increased defect density.

6.5 Discussion of disorder in tunnel junction devices

As mentioned in the preceding sections, we were unable to to find any asymmetric TJ devices that were free enough from the presence of UQDs. This means

Wafer	Gate Material	Tunnel junction device type	Date of Fabrication	Yield
RS-SD1-11	Al+ Ti/Pd	asymmetric	August 2020	0/12
	Al	symmetric	August 2020	0/5
	Ti/Pd	symmetric	August 2020	0/4
RS-SD1-10	Al	symmetric	October 2019	2/9
RS-SD1-8	Ti	symmetric	February 2018	2/6
RS-SD1-1	doped poly-silicon	symmetric	July 2017	0/5
RS-SD1-4	Al	symmetric	April 2017	2/6

Table 6.1: Summary of the yield of all tunnel junctions fabricated and measured at cryogenic temperatures in this work. Here, we define the TJ yield as the number of TJ measured that showed little to no evidence of UQDs or disorder divided by the total number of TJ devices.

we are unable to proceed with any type of barrier fitting for these devices and attempt to extract strain differences. The frequency of the UQDs in these devices are much higher than those seen in past devices. Based on factors such as the charging energies, locations, and voltage ranges, we can infer that these UQDs are not from a consistent source or more correctly, their sources are essentially random. This suggests that there is no unforeseen strain effect in the device [73], but likely the sources are due to either charge defects in the silicon dioxide [158] or impurities in the substrate [60]. In order to compare the asymmetric TJs to past devices, we will define the TJ yield as the number of TJ measured that showed little to no evidence of UQDs or disorder divided by the total number of TJ devices. Table 6.1 summarizes the yield of the all tunnel junctions fabricated and measured during this thesis. In total, we have measured 12 different TJ that had no device failures, such as gate leakage or non-functional contacts. None of these 12 devices shows behavior that seems free from UQDs. Additionally as a consistency check, we made symmetric

TJs similar to our previous TJs on the same wafer with the asymmetric TJs. We measured nine total (5 Al, 4 Ti/Pd) of these new symmetric TJs and again we found that none of them met the standard of TJ free from disorder. In a separate device run fabricated a year prior to this asymmetric TJ run, we made a set of symmetric Al gated TJs for the purpose of testing new methods for electrical isolation. The fabrication for this wafer and the asymmetrical tunnel junction wafer were nearly identical, particularly we use the same oxidation and forming gas anneal processes. On this wafer, we measured 9 total tunnel junctions and found 2 of those that did not show UQD behavior, a roughly 22 % yield. For all the wafers fabricated using the processes described in chapter 4, we measured 19 tunnel junctions in total and found 4 good tunnel junctions or a roughly 21 % yield. This count includes devices made with Al, Ti, and doped poly-silicon gates. Thus, prior to the fabrication of the asymmetric TJs, we were seeing a reasonably consistent yield even across different processes. In all cases, our statistics are too low to draw too many conclusions on the question of whether there is a significant difference in the yield between asymmetric and symmetric devices on the asymmetric TJ wafer versus previous wafers. Such a study would require a significantly higher throughput on low temperature measurements then we have previously been able to achieve.

Another possibility is that our MOS capacitor measurements are not capturing the full picture of the defect landscape important for the operation of TJs. The discrepancy could arise due the different operating regimes between the MOS capacitors and TJs, particularly the temperature difference. Consider for instance the measurement of D_{it}. For MOS capacitors, we measured D_{it} at room temper-

ature and thus only probe interface traps in a narrow range in the silicon band gap. We use p-type substrates so we probe the valence band side of gap, typically from $E - E_v$= 0.34 eV to 0.45 eV. For TJs, we measure at 2 K and thus the only electrically active traps will be very near the conduction band edge [166, 167]. D_{it} is typically significantly (several orders of magnitude) lower near midgap than near the band edges. The major assumption that we make by using MOS capacitors is that traps near midgap and the band edges will anneal the same, e.g. a reduction in the midgap trap density coincides with a similar reduction in the band edge trap density. There is lack of research on if such an assumption is valid, particularly if there any differences in the annealing character of the band edge traps. If this assumption was invalid, it would certainly be possible that our asymmetric TJs are suffering from larger impact from D_{it} than we expect based on the MOS capacitors. Therefore from this and previous discussion of the defect densities, we conclude that there is not a clear reason why we have been unable to find good asymmetric TJ devices.

Chapter 7: Conclusions and Future Work

In this last chapter, we summarize the results obtained in this work and propose future directions. Specifically, we will motivate potential future electrical measurements of strain in an alternative device and propose an alternative method to expand on the defect measurements performed in this work.

7.1 Conclusions

In this work, we have laid the foundation for quantitatively studying the effect of strain on silicon metal-oxide-semiconductor (MOS) quantum dots (QDs). In chapter 4, we demonstrated a new electrical measurement of gate-induced strain using tunnel junctions (TJs). First, we developed the device design by simulating strain in the TJ devices for different metals in COMSOL, where we showed that strain induced from the gates will modify the electrostatic barrier height. We present measurements of Al- and Ti-gated TJs and found an average barrier height difference of $\phi_{Ti} - \phi_{Al} \approx 1.12\ meV$. This result strongly contradicts the expected barrier height difference based the bulk values of the coefficient of thermal expansion of the metals of $\phi_{Ti} - \phi_{Al} \approx -11.2\ meV$. To address this discrepancy, we measured the thin film coefficient of thermal expansion using the wafer curvature measurements techniques

from chapter 2 and found that while the Al coefficient of thermal expansion was close to its bulk value ($\alpha_{Al,bulk} = 23.0 \pm 1.0 \times 10^{-6}\, K^{-1}$ [141],$\alpha_{Al,meas} = 23.0 \pm 2.8 \times 10^{-6}\, K^{-1}$) the measured Ti coefficient of thermal was nearly a factor of 2 larger than its bulk value ($\alpha_{Ti,bulk} = 8.9 \pm 0.1 \times 10^{-6} K^{-1}$ [141],$\alpha_{Ti,meas} = 16.2 \pm 2.0 \times 10^{-6} K^{-1}$). Using these measured values for the metal films, we found that our simulations and experimental results from the TJs were consistent.

Also shown in this work, the interplay between oxide charge defects and strain is complex and extremely important for reproducible quantum dot formation. This interplay was studied in detail via the forming gas anneal results from chapter 5. In that chapter, we measured the fixed charge density (Q_f), interface trap density (D_{it}), coefficients of thermal expansion (α), and intrinsic film stress (σ) for three metals used in MOS QD fabrication: Al, Ti/Pd, and Ti/Pt. We studied those quantities in MOS devices by varying the forming gas anneal temperature from 200°C to 425°C and showed that Ti/Pd and Ti/Pt have larger D_{it} than Al when optimally annealed. Moreover, the magnitude of Q_f is larger for Ti/Pd and Ti/Pt than for Al, with Al showing a net negative charge while Ti/Pd and Ti/Pt display net positive charge. Additionally, we showed that both α and σ increase with increasing anneal temperature and that, due primarily to intrinsic strain, Pd-gated devices have larger strain-induced modulation of the conduction band than their Al-gated counterparts, directly contradicting expectations based on the bulk α alone [72]. Finally, and most importantly, we found no anneal which simultaneously minimizes defects and the effects of strain in any of the materials studied. This result puts tension on the design and fabrication of silicon MOS QDs, where a choice must be made between

setting the anneal such that defects are minimized or the strain-induced modulation of the conduction band is minimized. This tension was not previously known to the quantum dot community.

Finally, in chapter 6, we showed our efforts to improve our TJ measurements from chapter 4. We developed and fabricated asymmetrical TJ devices with the goal of removing the impact of device-to-device variations on the strain measurement. This new design had an advantage in that the differences in strain could be measured directly from the barrier asymmetry rather than by comparing barrier heights. The goal was to remove the burden of the high yield necessary to measure well-behaved TJs in each gate material. Unfortunately, we were unable to obtain any asymmetrical TJs that were free of non-idealities to extract strain from the barrier parameter fits. To understand this difference in the performance of the two types of TJ devices, we analyzed MOS capacitors. We concluded from this analysis that the lack of good TJ devices could not be traced to any particular issue with the fabrication and, therefore, that asymmetrical TJs may still be a viable device for the measurement of strain. Answering this question fully will require fabricating new devices and obtaining better TJs.

7.2 Future Work

The results discussed in the preceding section revealed a few opportunities for improvement. First, we discuss some additional material characterization techniques to attempt to identify issues with gate materials in our TJ devices, including possible

interdiffusion of the materials of the MOS material stack and possible relaxation of the gate material with thermal cycling. Secondly, we discuss an alternative to our measurements of strain using TJs with the goal of avoiding the negative impacts of the unintentional quantum dots (UQDs) that hampered the yield of usable TJ devices. Finally, to address the limitations of our MOS defect measurements, we present a method to measure the interface trap density (D_{it}) at low temperatures and in a more appropriate portion of the silicon bandgap than the MOS capacitor measurements.

7.2.1 Additional materials charaterization

7.2.1.1 TEM analysis of MOS material interfaces

Metals such as Pd and Pt used in this work readily form silicides with silicon over the temperature range used in the forming gas anneals in chapter 5 [168]. In Ref. [169], TEM was used to analyze the interdiffusion of a copper gate material in MOS structure into the silicon dioxide and the results were used the confirm the hystersis present in MOS capacitor measurements after annealing. A similar study would be a benefit for our devices as to identify possible sources of device failure such as those found in chapter 6. We note that the metals used in our work have significantly lower diffusivities into silicon dioxide than Cu [170]. Additionally, TEM techniques can be used to measure strain in the MOS structures used in our TJ devices [171]. This would provide an alternative to the wafer curvature measurements with the advantage that we would be able to measure strain on a length scale more revelant

for our TJ devices. The downside of strain measurements with TEM is preparation of the sample lamella will destory the TJ device and possibly lead to shifts in the measured strain value [172].

7.2.1.2 AFM analysis of gate relaxation with temperature

As noted in chapter 5, the morphological changes in blanket films used for wafer curvature measurements do not occur for our e-beam lithography patterned features. This suggests that the patterned films in the TJ devices may relax differently than the blanket films during annealing or cooling processes [173]. Atomic force microscopy (AFM) could be used to indentify the relaxation of the gate material directly in the TJ devices. Here, we would measure the TJ device using AFM before and after any thermal cycling process and by comparing the measured gate widths and grain sizes in the film we could indentify if any relaxation has occured. Here, we would be able to indentify changes in the microstructure from changes in the surface roughness as well[174]. Additionally similar to TEM techniques above, the AFM can be combined with optical techniques to measure strain at the nanoscale in silicon. Scanning near-field optical microscopy (SNOM) [175, 176, 177] has been shown to be able measure strain due to defects in SiC and strained silicon using IR active phonons [178]. This technique has a significant advantage over the TEM-based method discussed above because it avoids the destruction of the sample that occurs in TEM sample preparation.

7.2.2 Measurement of strain using magnetoresistance oscillations

In chapter 4 and 6, we used TJs to electrically measure gate-induced strain in silicon MOS devices. As was discussed in those chapters although we demonstrated results consistent with our simulations, we found that the presence of unintentional quantum dots (UQDs) significantly hampered the yield of the devices from which we could extract strain. This yield is likely around 20 %. However, given our low statistics, this number could be either much higher or lower. Increasing the measurement throughput would aid in avoiding this problem somewhat. Perhaps, a better path forward is to increase the yield with an alternative device that could be measured and fabricated under similar conditions to the TJ devices and while also being more resistant to the formation of UQDs. Ye et. al [179] measured magnetoresistance oscillations in gated Hall bar devices in GaAs-AlGaAs heterostructures due to strain induced from a periodic array of Nb or Ni gates patterned on top of the Hall bar device. The array of gates creates a weak, approximately sinusoidal, periodic potential which leads to oscillations in the longitudinal magnetoresistance of the Hall bar as long as the gate spacing is much less than the electron mean free path. The relative shift in the magnetoresistance from the zero-field resistivity is ultimately proportional to the amplitude of the modulation potential. Thus, we could measure strain from this method by comparing the shifts in magnetoresistance from zero-field between devices with arrays made of different gate materials. This hall bar has a significant advantage over our TJ measurements in that it is less sensitive to the presence of UQDs. However, the lower mobility [180] and larger effective

mass in silicon MOS as compared to GaAs-AlGaAs, makes the measurements more difficult. For a silicon inversion layers at cryogenic temperatures, the mean free path is a function of the mobility (μ) and electron density (n_s) [181]:

$$l_{mfp} = 116.7\ [nm]\left[\frac{\mu}{10^4 cm^2 V^{-1} s^{-1}}\right]\left[\frac{n_s}{10^{12} cm^{-2}}\right]^{\frac{1}{2}} \tag{7.1}$$

Using values from silicon MOS hall bar with similar fabrication processes ($\mu = 6070\ cm^2V^{-1}s^{-1}$ and $n_s = 1.2 \times 10^{12}\ cm^{-2}$) [180], we obtain a mean free path of around 78 nm. If we use one the highest non-industrial MOS mobility reported in QD community ($\mu = 14,000\ cm^2V^{-1}s^{-1}$ and $n_s = 6.1 \times 10^{11}\ cm^{-2}$) [182], then we get a mean free path of 127 nm. First, this means that the period of the gate array would need to be significantly smaller than the periods of 500 nm and 950 nm used in Ref. [179], where the mean free path was on order of 1 μm. The oscillations in magnetoresistance occur at when the gate period, a, satisfies: $2R_c = (\lambda + \frac{1}{2})a$ where λ is an integer and R_c is cyclotron radius. R_c is dependent on the magnetic field, B, as [181]:

$$R_c = \sqrt{2n+1}\left[\frac{\hbar}{eB}\right]^{\frac{1}{2}} = \sqrt{2n+1}\left[\frac{25.66\ [nm]}{\sqrt{B}}\right] \tag{7.2}$$

Where B is in units of T and n is the integer for the Landau levels. For an optimistic gate period of $a = 30\ nm$, we expect to see oscillations at magnetic fields to occur at 1.06 T, 0.83 T, 0.69 T, and 0.62 T for the lowest Landau level ($n = 0$). These fields and oscillations should be observable assuming we could achieve the necessary gate pitch. This would put some burden on the electron beam lithography used

to fabricate the gate array, as the pitch would need to be less than 40 nm. In practice, this is difficult especially with positive liftoff processes where we have struggled at times getting consistent features especially for as dense a pattern as the gate array necessary for this device. The lithography constraints are the major drawback for this hall bar technique, but overall it appears to be a viable alternative or complement to TJ measurements. Additionally, it may be easier to achieve the necessary gate period with negative tone and dry etching process, but this would require significant changes to our current device process flows. Another potential issue with these hall bar devices is based on the mean free path calculations above; we need to be able to achieve very high mobilities in order to perform the measurement. This is a non-trivial problem in silicon MOS and would need to be addressed in the fabrication process. It is reasonable to assume that any work to significantly increase the mobilities of any silicon MOS devices we fabricated would also lead to significant reductions in the appearance of UQDs in the TJ devices as well. Thus, we believe this hall bar measurement of strain is best used as a compliment to the TJ measurements rather than a replacement. Fabricating both types of devices simultaneously on the same wafers or chips would be an advantageous two-pronged measurement of strain.

7.2.3 Low temperature measurement of the interface trap density

In chapters 3 and 5, we used MOS capacitors as a proxy for measuring the oxide charge defect densities, such as fixed charge density (Q_f) and interface trap

density (D_{it}). These measurements suffer from the disadvantage that they were performed at low temperatures where QDs operate (<10 K). At low temperatures, the only electrically active traps will be very near the conduction band edge [166, 167]. D_{it} measured at room temperature only probe interface traps in a narrow range of energy away from the band edge (typically from $E - E_v$= 0.34 eV to 0.45 eV). D_{it} is typically several orders of magnitude lower near midgap than near the band edges. There are several alternatives to MOS capacitors that could be used to measure D_{it} at temperatures $< 10\ K$ and all involve the use of MOSFETs. Hafez et. al [166] demonstrated measurements of D_{it} at room temperature, 77 K, and 4.2 K using MOSFETs with D_{it} in the range of $10^{13} - 10^{14}\ eV^{-1}cm^{-2}$ at 4.2 K. They extracted D_{it} using two methods: the subthreshold slope and the dynamic transconductance. The subthreshold slope relies on measuring source-drain current as a function of gate voltage using a semiconductor parameter analyzer and fitting the temperature dependent slope of the curve in the weak inversion regime. The dynamic transconductance technique operates similarly to the conductance method with MOS capacitors we presented in chapter 3 and used in chapter 5 to measure D_{it}. The D_{it} measurement proceeds in the same way as the parallel capacitance and conductance of the MOSFET as seen by the gate is measured as a function of frequency and gate voltage. Appropriate circuit models are then used to convert the measured conductance to an equivalent parallel conductance due to D_{it}. The models used in transconductance method require short channel ($< 20\ um$) MOSFETs, so we would be unable to directly perform these measurements on QD devices or our TJs without modifications. We have routinely fabricated short channel MOSFETs

as diagnostic devices on the same wafers as TJs and QDs, but have not typically measured them at low temperature. For both methods, we have the necessary equipment (parameter analyzer and LCR meter) to cover most of the desired frequency range. In Ref. [166], the measurement was performed from 1 kHz up 10 MHz, while we have the ability to measure from 50 Hz to 2 MHz. Additionally, we have the capability to make the necessary devices, where the only design changes would be to remove some of the stray capacitance due the layout of our metal lines on the chip to connect to the MOSFET. The measurement of D_{it} discussed above would provide a significant contribution to the silicon MOS QD community as a whole in terms of determining how defects are truly affecting the device at low temperature. This technique would complement the few previous measurements of the band tail states for MOS QDs using ESR techniques [182, 183], while also having a higher throughput allowing for quicker feedback on fabrication processes.

7.2.4 Simulation improvements

Throughout this thesis, we used COMSOL to simulate the strain effects in our TJ devices and did not present any simulations of the electrostatics. Our attempts to model the electrostatics of the TJs using COMSOL using the Poisson and drift-diffusion equations failed to produce a tunnel barrier over any appreciable range of gate voltage above threshold for the leads. This result clearly contradicted the experimental data. The quantum mechanical effects of a MOS device like our TJ are not accurately captured via these simulations particularly because of the 3d

density of states used in the model. More accurate modeling could be achieved with a Poisson-Schrodinger solver [184], which can be setup such that the density of states is 2d and quantum mechanical effects are included. The framework for performing these simulations for MOS devices now exists fully in the most recent version of COMSOL 5.6 [185]. Additionally, these Poisson-Schrodinger simulations can be setup to include the effects of strain in the device. This would allow us to self consistently solve for strain and electrostatic components together in a single simulation. These combined electrostatic and strain simulations would be a key piece in understanding the transport behavior and lead to significant improvements to the measurement of strain.

It is important to note that while Poisson-Schrodinger solvers tend to incorporate quantum effects on device behavior better than the drift-diffusion models, they fail to correctly model important aspects of QD devices, such as tunnel couplings between QDs [184] and barrier capacitances [186]. Tight binding models incorporating disorder in QD devices [187] have shown the ability to account for the order of magnitude changes on tunnel couplings that charge defects can cause. Applying such electrostatic modeling to our tunnel junction devices would be a significant benefit to the MOS QD field as our TJ would act as a more direct probe of these behaviors without having to account for the more complicated device structure of the QD. This could allow for iterating on both the modeling and our TJ devices to improve on QD device design as a whole, which is essential for achieving the full potential for applications with quantum computing and current standards.

7.2.5 Outlook

The work in this thesis focused on working towards controlling inhomogeneity in silicon MOS QDs with the goal of improving the reproducibility for applications in quantum computing and quantum current standards. In the case of strain, the control aspect is especially important. Strain need not be viewed as a detriment. The possibly of controlling strain could lead to design changes in MOS QDs, where it could be used to form QDs rather than relying solely on electrostatics. The advantage of such a strategy can be seen in previous QD work [73], where strain induced QDs should form under a single gate. This would be a significant improvement over the current state of the art for scaling QD devices since the large number of gates necessary for each QD is an impediment to scaling. Building towards the goal of manipulating strain in silicon MOS QDs requires methods of measuring strain under relevant conditions while also finding ways to adjust processing to minimize the impact of non-idealities. The work in thesis represents a significant step towards that goal. The devices presented easily lend themselves to future work exploring deposition parameters and anneals to manipulate inhomogeneous strain. Our method for measuring relative strain satisfies the sensitivity, spatial resolution and low-temperature requirements relevant for MOS QDs. Moreover, the fabrication and measurements are similar to those for QDs so that this method is directly relevant for QD devices. Our data provide an important step forward in assessing gate-induced strain in QD devices in-situ while highlighting the need for further experimental work and a greater theoretical understanding of the electrostatics and

strain behavior.

Appendix A: Tunneling Barrier Fitting

This appendix provides the neccessary files to reproduce the tunnel barrier fits for one of the Ti-gated tunnel junctions in Chapter 4.

A.1 Matlab files for Brinkman-Dynes-Rowell (BDR) model

The following section contains all the matlab related files needed to produce Brinkman-Dynes-Rowell (BDR) model fits for the blue data chapter 4, which is for a Ti-gated tunnel junction at a gate voltage of 0.87 V. The raw sampling data is from the file '`Feb15_636.txt`' stored on the NIST network drive under `\SET_data\QUITS_I\RS-SD1-8\Die13\Feb15_636.txt`.

A.1.1 Converting DC sampling for use in fitting

```
%%RSampLoop - convert raw sample into an averaged variable and
   calculate different conductance and
%% plot resistance.
%%Labeling scheme averaged values muDay_File# -> columns: Vdrain,
   Idrain,Isource,Vsource,VprobeL,VprobeR)
%%  2-pt sliding average -> muDay_File#S,  3-pt sliding average ->
   muDay_File#S3,   delVDay_File# -> voltage difference between
```

```
    probes, gDay_File#-> differential conductance

Feb15_636 = importdata('Feb15_636.txt','\t',14);
strVG=char(Feb15_636.textdata(11,1));
strS=char(Feb15_636.textdata(12,1));
Vgl=str2num(strVG(22:28)); %VGL

Vgr=str2num(strVG(36:42)); %VGR
numS=str2num(strS(16:21));
numD=(length(Feb15_636.data(:,1))./numS);
str15_636=strcat('Feb15_636,',strVG(22:28)); %String to be used by
    BRDfit.m
mu15_636= zeros(numD,4);
std = zeros(numD,4);

DiffAmp=0;

%%Average raw sample data
i=1;
for j=1:numD
    n=(numS*(j-1))+1;
    mu15_636(i,1)=mean(Feb15_636.data(n:n+numS-1,2));
    mu15_636(i,2)=mean(Feb15_636.data(n:n+numS-1,3));
    mu15_636(i,3)=mean(Feb15_636.data(n:n+numS-1,4));
    mu15_636(i,4)=mean(Feb15_636.data(n:n+numS-1,5));
    mu15_636(i,5)=mean(Feb15_636.data(n:n+numS-1,6));
    mu15_636(i,6)=mean(Feb15_636.data(n:n+numS-1,7));
```

```
    i =i+1;
end


%%Two point smoothing
mu15_636S=movmean(mu15_636,6,1);
%%Three point smoothing
mu15_636S3=movmean(mu15_636,12,1);


%%Rescaled differential bias arrays
if DiffAmp==1
    mu15_636(:,7)=mu15_636(:,5)./100;
    mu15_636S(:,7)=mu15_636S(:,5)./100;
    mu15_636S3(:,7)=mu15_636S3(:,5)./100;
else
    mu15_636(:,7)=mu15_636(:,6)-mu15_636(:,5);
    mu15_636S(:,7)=mu15_636S(:,6)-mu15_636S(:,5);
    mu15_636S3(:,7)=mu15_636S3(:,6)-mu15_636S3(:,5);
end


j=1;
for i =1:(length(mu15_636)-1)
delV15_636(j,:)=(mu15_636(i+1,7)+mu15_636(i,7))/2;
j=j+1;
end
```

```
57
58 j=1;
59 for i =1:(length(mu15_636S)-1)
60 delV15_636S(j,:)=(mu15_636S(i+1,7)+mu15_636S(i,7))/2;
61 j=j+1;
62 end
63
64 j=1;
65 for i =1:(length(mu15_636S3)-1)
66 delV15_636S3(j,:)=(mu15_636S3(i+1,7)+mu15_636S3(i,7))/2;
67 j=j+1;
68 end
69
70
71 [R4T15_636,Bint,fit15_636] = FuncFitR4T(movmean(mu15_636(:,7),3),
      movmean(mu15_636(:,2),3),1000E-6);
72 R4T15_636
73
74 g15_636 = diff(mu15_636(:,2))./diff(mu15_636(:,7));
75 g15_636S = diff(mu15_636S(:,2))./diff(mu15_636S(:,7));
76 g15_636S3 = diff(mu15_636S3(:,2))./diff(mu15_636S3(:,7));
77
78 gE = diff(mu15_636S(:,2)+mu15_636S(:,3))./diff(mu15_636S(:,7)); %gE
       should be zero if Is(column 3)=Id(column 4), nonzero when diff
      amp impedance is a problem
79 Fitstr=sprintf('R_{0}= %.3f MOhm, ',R4T15_636./(1e6));
80 figure(1);
```

```
81 plot(mu15_636S(:,7)./(1e-3),movmean(mu15_636(:,2),3)./(1E-9),'.','
      DisplayName','I_{DC} data');
82 hold on;
83 plot(mu15_636S(:,7)./(1e-3),fit15_636./(1e-9),'-','DisplayName',
      Fitstr,'LineWidth',1);
84 xlabel('4term Bias voltage (mV)','FontSize',14);
85 ylabel('DC Current (nA)','FontSize',14);
86 % title({'DC current vs Bias';str(12:28);'Ti gated tunnel junction
      device at 2 K';Constr;strVg},'FontSize',14);
87 legend 'show'
88
89
90 figure(2);
91 plot(delV15_636./(1E-3),(g15_636)./(1E-9),'b*-','DisplayName','no
      smoothing');
92 hold on;
93 plot(delV15_636S./(1E-3),(g15_636S)./(1E-9),'r*-','DisplayName','2
      pt smoothing');
94 plot(delV15_636S3./(1E-3),(g15_636S3)./(1E-9),'g*-','DisplayName','
      3 pt smoothing');
95 xlabel('4term Bias voltage (mV)','FontSize',14);
96 ylabel('Diff Conductance (nS)','FontSize',14);
97 axis tight;
98 legend 'show'
```

```
99 set(gca,'yscale','log')
```

Listing A.1: Script file for averaging DC sampling data from tunnel junction measurements

The m-file '`RSampLoop.m`' is used to take the raw sampling data and convert to an averaged data set of $I(V_D)$. The m-file for this example is stored under `\SET_team\Ryan\Thesis\thesistex\AppendixA\RSampLoop.m`.

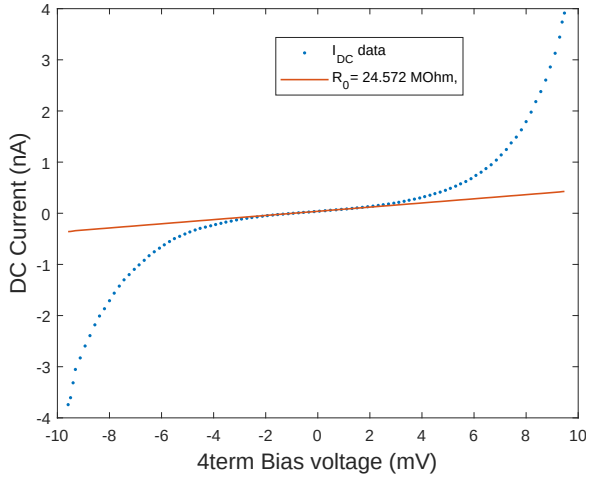

Figure A.1: I(V_D) plot and zero-bias resistance fit produced by `RSampLoop.m`

Example of the output plots from '`RSampLoop.m`' are shown in Fig. A.1 and Fig. A.2. Fig. A.1 shows the averaged $I(V_D)$ data along with a linear fit performed to obtain the zero-bias resistance. Fig. A.2 shows the differential conductance obtained from the $G(V_D)$ data that is used later for fitting. The BRD fitting files later will use the $I(V_D)$ as input over a narrow voltage range choosen by hand.

A.1.2 BRD function file

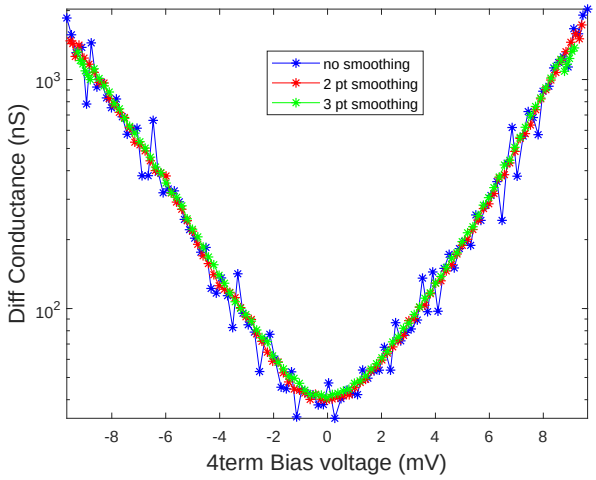

Figure A.2: Differential conductance plot produced by `RSampLoop.m`

```
function [JBRD,GBRD,Gm,Vmin]=BRDFunc(Wg,tinv,asym)
%%Function that generates J, G, and zero-bias G(Gm) symbolic
    functions
%%a -> barrier height in V, b-> barrier width in nm, c -> barrier
    asymmetry
%%in volts,Gm -> minimum conductance
syms J A C B a x b a1 a2 d Gm Ao
hbar=1.05E-34;
m=9.109384E-31;
me=.19.*m; %reduced mass for silicon kz
q=1.602176E-19;
h=6.626076E-34;
Ao=4*sqrt(2*me*q)*(b*(1E-9))/(3*hbar); %A0 constant from BRD paper
%BRD zero-bias conductance
Gm=Wg.*tinv.*((q/h)^2)*(sqrt(2*me*q*a)/(b.*(1e-9)))*exp(-(2*(b.*(1e
    -9))/hbar)*sqrt(2*me*q*a));
```

```
14 if asym == 1
15     syms c
16     JBRD=(Gm.*(x-((((b.*(1e-9))*sqrt(2*me*q)*c)/(24*hbar*(a^(3/2)))
   )*x^2)+((((b.*(1e-9))^2*me*q)/(12*a*hbar^2))*x^3)))+d; %BRD
   Current Func with asymmetry
17     GBRD=real(Gm.*(1-((((b.*(1e-9))*sqrt(2*me*q)*c)/(12*hbar*(a
   ^(3/2))))*x)+((((b.*(1e-9))^2*me*q)/(4*a*hbar^2))*x^2))); %BRD
   Conductance Func with asymmetry
18     Vmin=0.649*c/(sqrt(me/m).*10*b*sqrt(a));
19 else
20     syms c
21     JBRD=(Gm.*(x-((((b.*(1e-9))*sqrt(2*me*q)*0)/(24*hbar*(a^(3/2)))
   )*x^2)+((((b.*(1e-9))^2*me*q)/(12*a*hbar^2))*x^3)))+c; %BRD
   Current Func without asymmetry
22     GBRD=real(Gm.*(1-((((b.*(1e-9))*sqrt(2*me*q)*0)/(12*hbar*(a
   ^(3/2))))*x)+((((b.*(1e-9))^2*me*q)/(4*a*hbar^2))*x^2))); %BRD
   Conductance Func without asymmetry
23     Vmin=0;
24 end
```

Listing A.2: Function file for BRD model

The function file '`BRDFunc.m`' used for fitting data using the BRD model. It has three inputs: the inversion layer thickness (inv), the gate width (Wg), and specifier for using the asymmetric ($asym = 1$) or symmetric ($asym = 0$) BRD model functions. It outputs symbolic functions for the BRD current (J), conductance (G), minimum

conductance (Gm), and the voltage position of conductance minimum ($Vmin$).

A.1.3 BRD fitting m-file

```
function [fitresult,gof,fitresult3,gof3,fitOut] = BRDFitAllQ(x, z,
    Vgl,Vgr)

%% Fit:
[xData, yData] = prepareCurveData( movmean(x,1), movmean(z,1) );
dg=diff(movmean(z,1))./(diff(movmean(x,1)));
% dg=movmean(dg,3);
[m,n]=min(dg);
j=1;
for i =2:(length(x))
dx(j)=(x(i-1)+x(i))./2;
j=j+1;
end
[xData2, yData2] = prepareCurveData( dx, dg' );

% Set up fittype and options.
[JBRD,GBRD,Gm]=BRDFunc(100E-9,4E-9,1); %Thrid input 1= asymmetrical
     function, 0=symmetrical function
[JS,GS,GmS]=BRDFunc(100E-9,4E-9,0); %Thrid input 1= asymmetrical
    function, 0=symmetrical function
GBRD=simplify(GBRD);
syms s phi
%Fit asymmetric J model
y=matlabFunction(real(JBRD));
```

```
ft = fittype(y, 'independent', 'x', 'dependent', 'y' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Algorithm = 'Levenberg-Marquardt';
opts.DiffMinChange = 1e-155;
opts.Display = 'Off';
opts.MaxFunEvals = 9000;
opts.MaxIter = 9000;
opts.Robust = 'Bisquare';
opts.StartPoint = [0.004 5 0 0];
opts.TolFun = 1e-155;
opts.TolX = 1e-155;
% Fit model to data.
[fitresult, gof] = fit( xData, yData, ft, opts );

%Fit symmetric J model
% y=matlabFunction(JS);
% ft2 = fittype(y, 'independent', 'x', 'dependent', 'y','
    coefficients',{'a','b','c'} );
% opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
% opts.Algorithm = 'Levenberg-Marquardt';
% opts.DiffMinChange = 1e-155;
% opts.Display = 'Off';
% opts.MaxFunEvals = 5000;
% opts.MaxIter = 5000;
% opts.Robust = 'Bisquare';
% opts.StartPoint = [0.001 10 0];
% opts.TolFun = 1e-155;
```

```
% opts.TolX = 1e-155;
% % Fit model to data.
% [fitresult2, gof2] = fit( xData, yData, ft2, opts );

%Fit asymmetric G model
y=matlabFunction(real(GBRD));
ft3 = fittype(y, 'independent', 'x', 'dependent', 'y' );
opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
opts.Algorithm = 'Levenberg-Marquardt';
opts.DiffMinChange = 1e-155;
opts.Display = 'Off';
opts.MaxFunEvals = 9000;
opts.MaxIter = 9000;
opts.Robust = 'Bisquare';
opts.StartPoint = [0.004 20 0];
opts.TolFun = 1e-155;
opts.TolX = 1e-155;
% Fit model to data.
[fitresult3, gof3] = fit( xData2, yData2, ft3, opts );

%Fit symmetric G model
% y=matlabFunction(GS);
% ft4 = fittype(y, 'independent', 'x', 'dependent', 'y','
    coefficients',{'a','b'} );
% opts = fitoptions( 'Method', 'NonlinearLeastSquares' );
% opts.Algorithm = 'Levenberg-Marquardt';
% opts.DiffMinChange = 1e-155;
```

```
% opts.Display = 'Off';
% opts.MaxFunEvals = 5000;
% opts.MaxIter = 5000;
% opts.Robust = 'Bisquare';
% opts.StartPoint = [0.001 20 ];
% opts.TolFun = 1e-155;
% opts.TolX = 1e-155;
% % Fit model to data.
% [fitresult4, gof4] = fit( xData2, yData2, ft4, opts );
%Get all important fit values
BparamJ=coeffvalues(fitresult);BPJCI=confint(fitresult);
BparamG=coeffvalues(fitresult3);BPGCI=confint(fitresult3);
phiJ=BparamJ(1);phiJL=BPGCI(1,1);phiJH=BPJCI(2,1);
sJ=BparamJ(2);sJL=BPJCI(1,2);sJH=BPJCI(2,2);
dphiJ=BparamJ(3);dphiJL=BPJCI(1,3);dphiJH=BPJCI(2,3);
phiG=BparamG(1);phiGL=BPGCI(1,1);phiGH=BPGCI(2,1);
sG=BparamG(2);sGL=BPGCI(1,2);sGH=BPGCI(2,2);
dphiG=BparamG(3);dphiGL=BPGCI(1,3);dphiGH=BPGCI(2,3);
[Gmin,Gm_pos]=min(fitresult3(xData2))
Vmin=xData2(Gm_pos);
fitOut=[Vgl,Vgr,Vmin./(1E-3),Gmin./(1E-9),phiG./(1E-3),phiGL./(1E
    -3),phiGH./(1E-3),sG,sGL,sGH,dphiG./(1E-3),dphiGL./(1E-3),dphiGH
    ./(1E-3),gof3.rsquare,BparamJ(4)./(1E-12),phiJ./(1E-3),phiJL./(1
    E-3),phiJH./(1E-3),sJ,sJL,sJH,dphiJ./(1E-3),dphiJL./(1E-3),
    dphiJH./(1E-3),gof.rsquare];
% Plot fit with data.
plot( fitresult, xData, yData);
```

```
% plot( fitresult2, xData, yData);
figure(2)
plot( fitresult3, xData2, yData2 );
% plot( fitresult4, xData2, yData2 );
```

Listing A.3: Script file for fitting data to BRD model

The file '`BRDFitAllQ.m`' used for fitting data using the BRD model and uses the function file '`BRDFunc.m`' in the fitting. It uses the matlab fittype function with the 'nonlinearleastsquares' option. It has four inputs the V_D voltage array (x), the I_D current array (z), the left gate voltage (Vgl) and the right gate voltage (Vgr). It has 6 outputs: the output fit current data and goodness of fit data ($fitresult$ and gof), the output fit conductance data and goodness of fit data ($fitresult3$ and $gof3$), and an output string containing all the information that I typically dump into a txt file such as the barrier parameters and the confidence intervals.

A.2 Example fits

This section will show how to use the m-files discussed above to produce fit data similar to that from Chapter 4.

```
>>[fitJ15_636,gofJ15_636,fitG15_636,gofG15_636,fitout15_636] =
   BRDFitAllQ(mu15_636(32:70,7),mu15_636(32:70,2),Vgl,Vgr)

Gmin =

   3.8666e-08

```

```
Gm_pos =

    20


fitJ15_636 =

     General model:
     fitJ15_636(x) = real((exp(b.*sqrt(a.*5.54603783944192e-50)
   .*(-1.904761904761905e25)
                    ).*sqrt(a.*5.54603783944192e-50).*(x+(b.^2.*x
   .^3.*2.096008253757339e-
                    1)./a-1.0./a.^(3.0./2.0).*b.*c.*x
   .^2.*9.345248912676919e-
                    2))./b).*2.338662809557086e22+real(d)
     Coefficients (with 95% confidence bounds):
       a =    0.002683  (0.002599, 0.002768)
       b =        24.6  (24.26, 24.94)
       c =   0.0004842  (0.0002512, 0.0007171)
       d =   3.555e-11  (3.455e-11, 3.656e-11)

gofJ15_636 =

  struct with fields:

           sse: 1.4786e-22
```

```
       rsquare: 0.9998
           dfe: 35
    adjrsquare: 0.9997
          rmse: 2.0554e-12


fitG15_636 =

     General model:
     fitG15_636(x) = real((7.642420785014145e7.*exp
   (7.642420785014145e7.*sqrt(a)
                    .*b.*(-5.869500782894433e-8)).*(a.^(3.0./2.0)
   .*2.054215522980229e47-
                    b.*c.*x.*3.839431076507008e46+sqrt(a).*b.^2.*x
   .^2.*1.291695807348903e47)
                    )./(a.*b)).*3.508181636728192e-58
     Coefficients (with 95% confidence bounds):
       a =    0.002782  (0.002656, 0.002909)
       b =       24.24  (23.71, 24.78)
       c =   5.332e-05  (-0.0003781, 0.0004848)

gofG15_636 =

  struct with fields:

           sse: 4.1884e-16
       rsquare: 0.9815
```

```
           dfe: 35
    adjrsquare: 0.9805
          rmse: 3.4593e-09


fitout15_636 =

  Columns 1 through 12

    0.8700    0.8700    0.0376   38.6660    2.7820    2.6555
    2.9085   24.2449   23.7089   24.7809    0.0533   -0.3781

  Columns 13 through 24

    0.4848    0.9815   35.5547    2.6831    2.6555    2.7675
   24.6018   24.2616   24.9419    0.4842    0.2512    0.7171

  Column 25

    0.9998
```

Listing A.4: Example input and output from `BRDFitAllQ.m`

The file '`FitOutput.m`' shows an example of the input and output used for running a fit for the data '`Feb15_636.txt`' using '`BRDFitAllQ.m`'. Examples of the output plots from '`FitOutput.m`' are shown in Fig. A.3 and Fig. A.4. Fig. A.3 shows the fit of the current from the BRD model obtained from the $I(V_D)$ data from

'BRDFitAllQ.m'. Fig. A.4 shows the fit of the differential conductance obtained from the $G(V_D)$ data from 'BRDFitAllQ.m'.

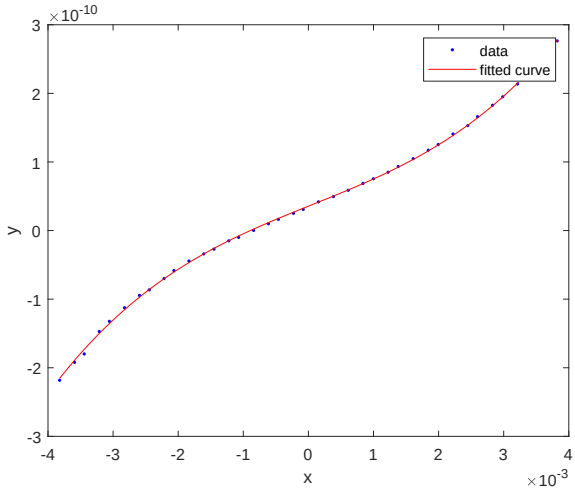

Figure A.3: Example fit output plot for the $I(V_D)$ data produced by BRDFitAllQ.m

```
figure(1)
plot(delV15_636./(1E-3),g15_636./(1E-9),'b*','MarkerSize',8);
hold on;
plot(delV15_636(32:70)./(1E-3),fitG15_636(delV15_636(32:70))./(1E
    -9),'b-','LineWidth',2);
xlim([-6,6]);
ylim([0,600]);
xlabel('V_{D} (mV)','FontSize',18);
ylabel('G(V_{D}) (nS)','FontSize',18);
```

Listing A.5: Example plot setup to produce $G(V_D)$ figures from Chapter 4.

The file 'Compplot.m' is used for producing a $G(V_D)$ plot similar to those seen in Chapter 4 of the main text. The result of running 'Compplot.m' is shown in Fig. A.5.

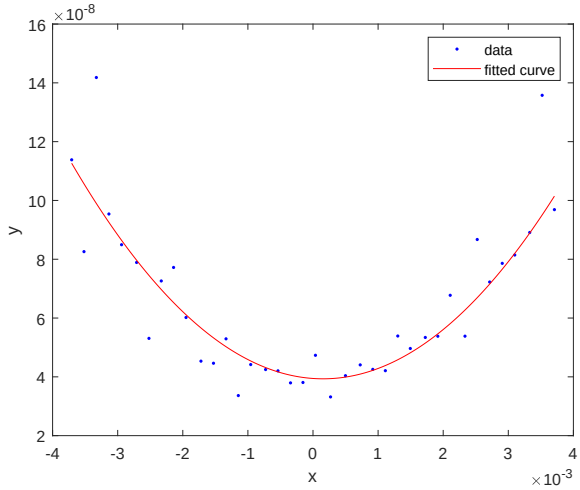

Figure A.4: Example fit output plot for the $G(V_D)$ data produced by `BRDFitAllQ.m`

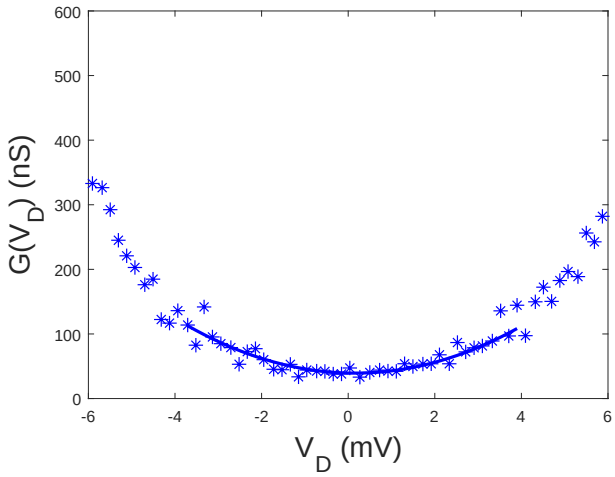

Figure A.5: Example fit for $G(V_D)$ data produced by `Compplot.m` analogous to plots from Chatper 4

Appendix B: Image Processing for Grain Size analysis

This appendix gives the image processing steps used to analyze the grain structures from the films in chapter 5.

B.1 Image processing steps

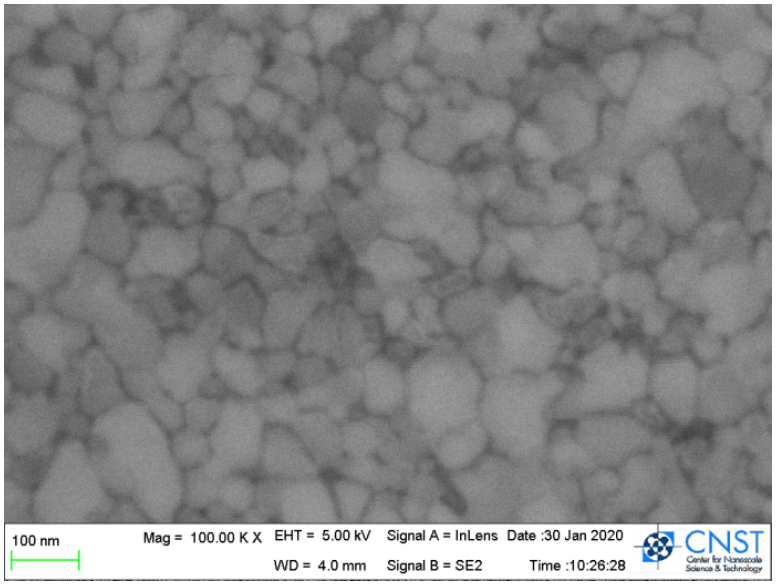

Figure B.1: Original SEM image for a Ti/Pd film after a forming gas anneal in 5 %H2/N2 at 350°C. The image was taken using the ZEISS FESEM in the CNST cleanroom. Scale bar is 0.9 pixels per nm.

Figures B.1 through B.11 outline the image processing in ImageJ used to analyze the average grain sizes of our metal films from chapter 5. The steps are as

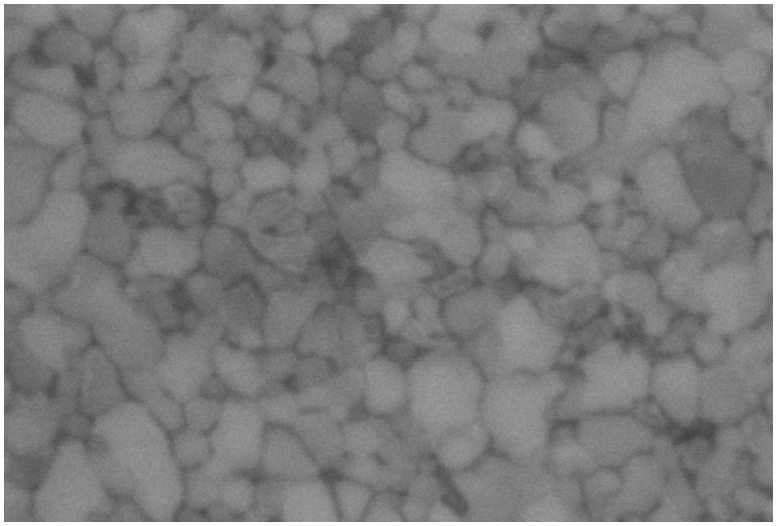

Figure B.2: SEM image after cropping out the information bar at the bottom in ImageJ.

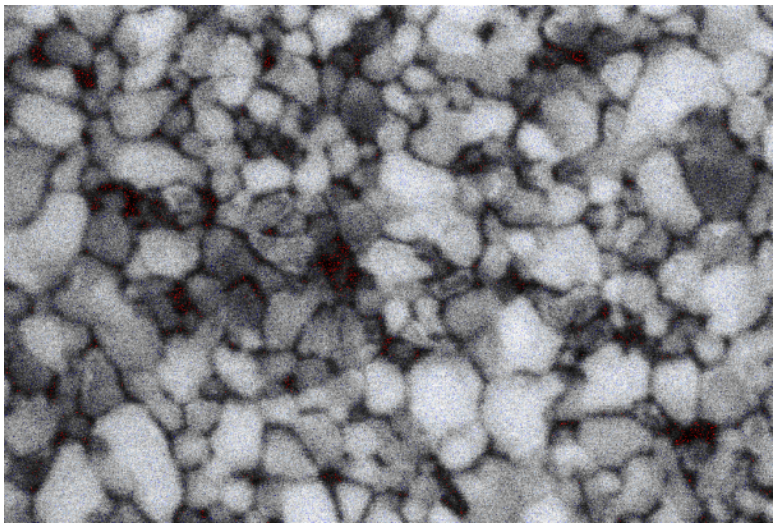

Figure B.3: SEM image after using 'Enhance contrast' function in ImageJ.

follows:

1. Starting point is original SEM image taken using the ZEISS FESEM in CNST cleanroom. Shown in Fig. B.1.

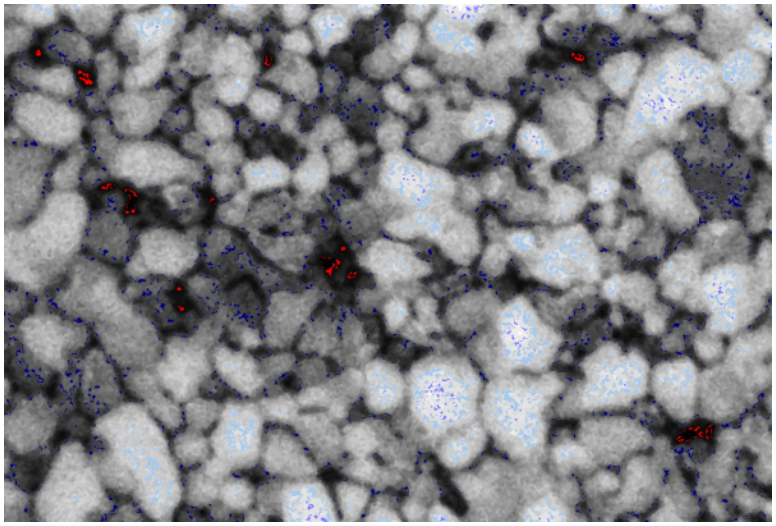

Figure B.4: SEM image after using 'Despeckle' (Procces− >Noise− >Despeckle) function in ImageJ twice.

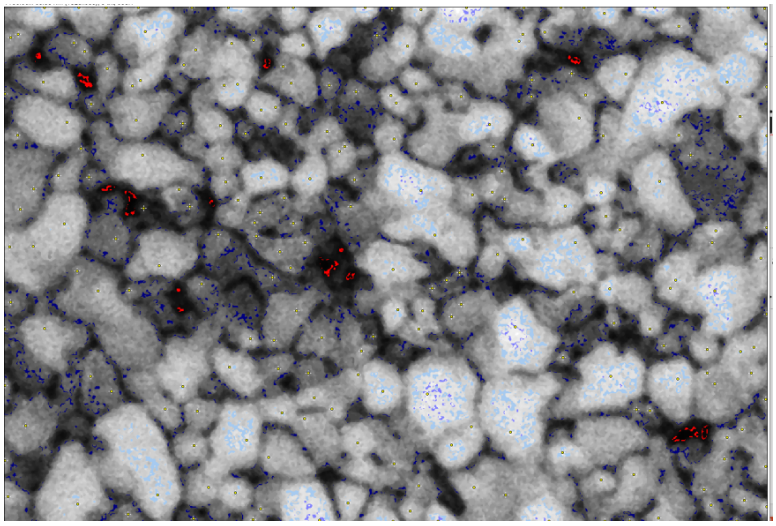

Figure B.5: SEM image after using 'Find Maxima' (Procces− >Find Maxmia) function in ImageJ .

2. After setting the pixel per nm scale, we crop out the information bar at the bottom of the image. Shown in Fig. B.2.

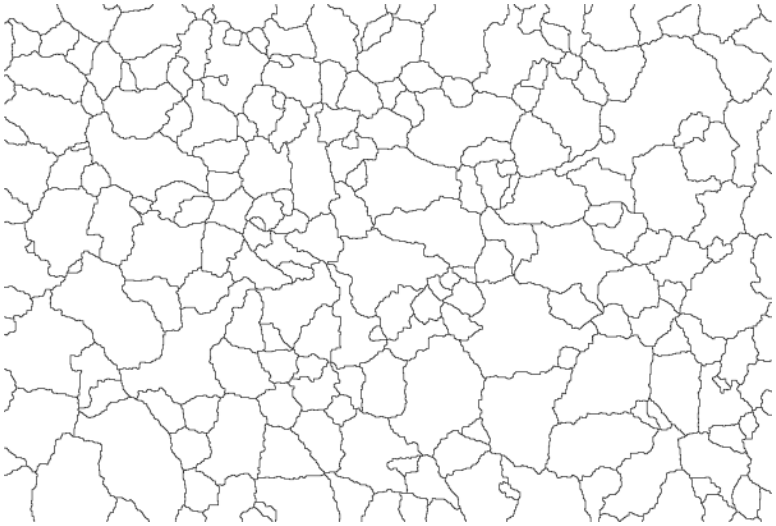

Figure B.6: SEM image after using 'Segmented Particles' (Procces− >Find Maxmia− >Segmented particles) function in ImageJ.

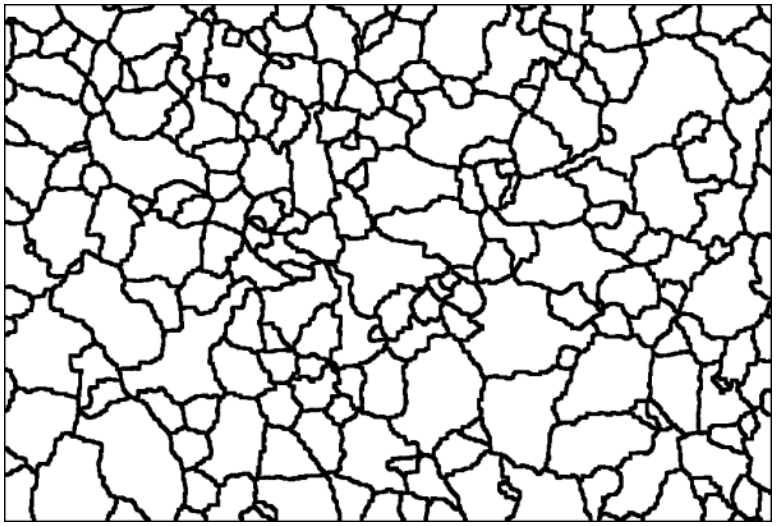

Figure B.7: SEM image after using 'Erode' on segmented image (Process− > Binary− > Erode) function in ImageJ.

3. We use the 'ehance contrast' function to highlight the grain boundaries more.

   Shown in Fig. B.3.

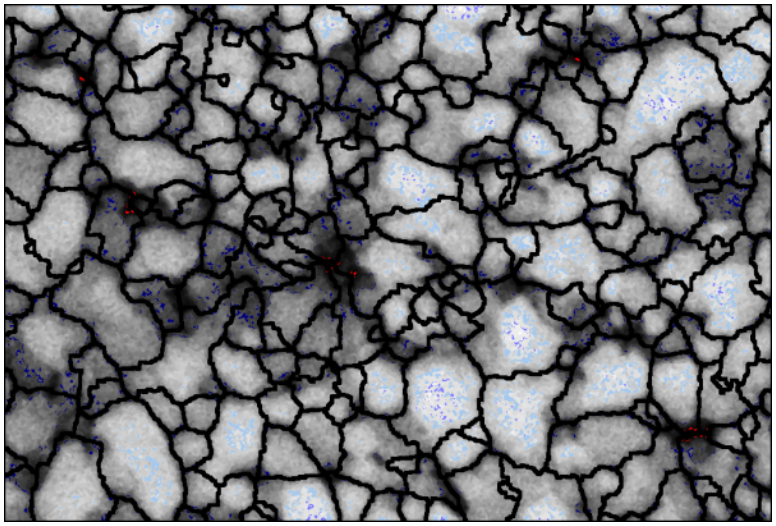

Figure B.8: SEM image after using 'AND' segmented image and despeckled image (Process− >Image calculator− > AND) in ImageJ.

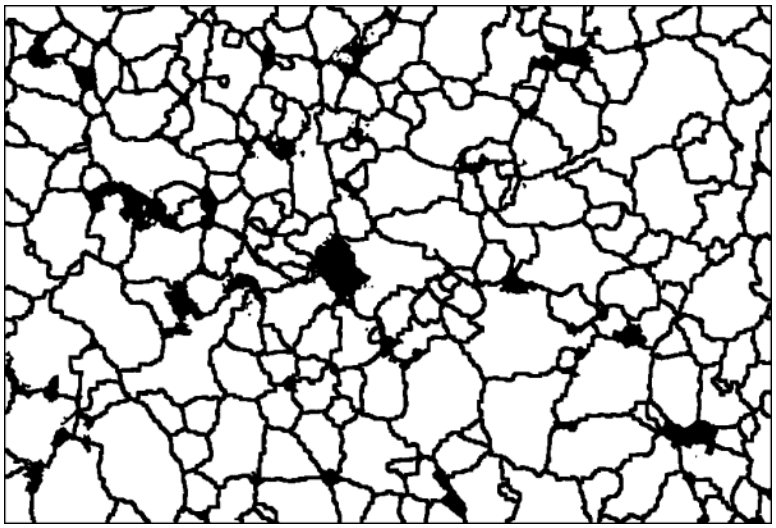

Figure B.9: SEM image after using 'Threshold' on AND image(Image− >Adjust− >Threshold) in ImageJ.

4. Using the image after 'ehance contrast', we use the despeckle function twice to clean up some of the noise in the image. Shown in Fig. B.4.

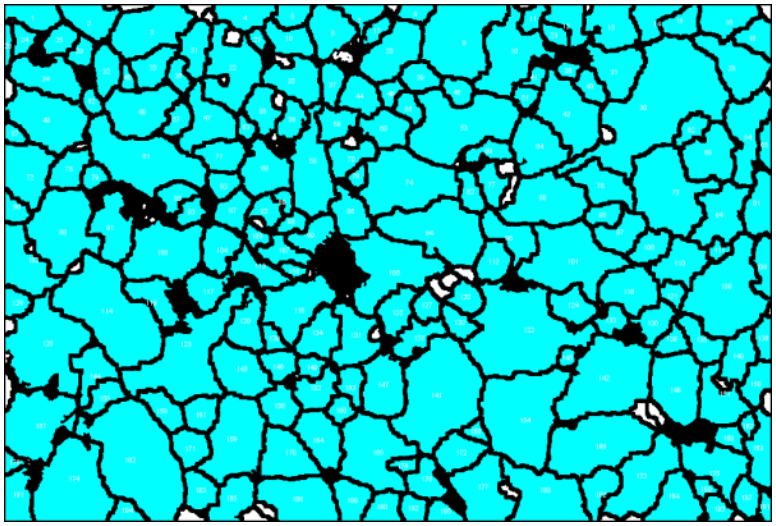

Figure B.10: Image showing grain regions selected by 'Analyze Particles' (Analyze− >Analyze Particles) in ImageJ.

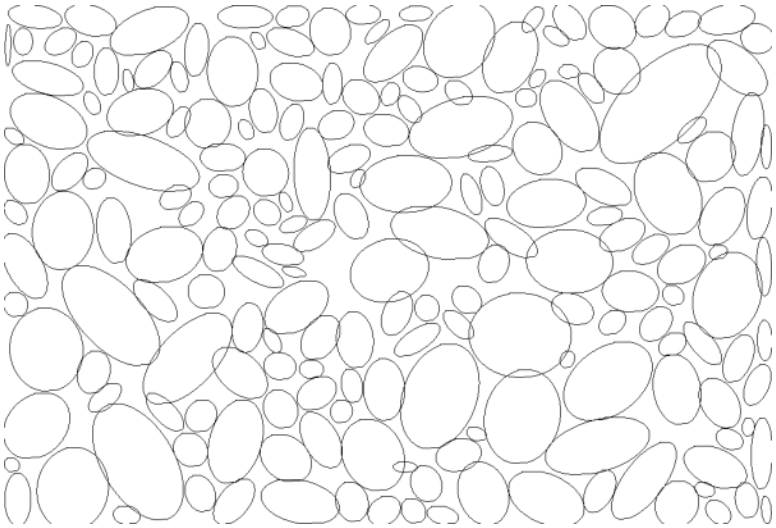

Figure B.11: Image showing fit ellipses generated by 'Analyze Particles' (Analyze− >Analyze Particles) in ImageJ.

5. Using the despeckled image, we use the 'Find Maxima' function to highlight (yellow data points) the peak in each grain. Shown in Fig. B.5.

6. Using the despeckled image, we use the 'Find Maxima' function to create a segmented image using the peaks identified previously. Shown in Fig. B.6.

7. We combine the segmented image and original despeckled image using the 'AND' function. Shown in Fig. B.7.

8. Using the combined AND image, we use the 'threshold' function to make the image binary. Shown in Fig. B.9.

9. Using the binary image, we can use the 'Analyze Particles' function to fit the areas to ellipses. The grains identifiedy by this function are shown in Fig. B.10 and the resulting fit ellipses are shown in Fig. B.11.

10. The distributions of the fit ellipse parameter are exported to a text file for future analysis.

B.2 Distributions for grain size analysis

In this section we give some additional details on the choice of distribution used in the grain size analysis. We chose to use the lognormal distribution in our analysis

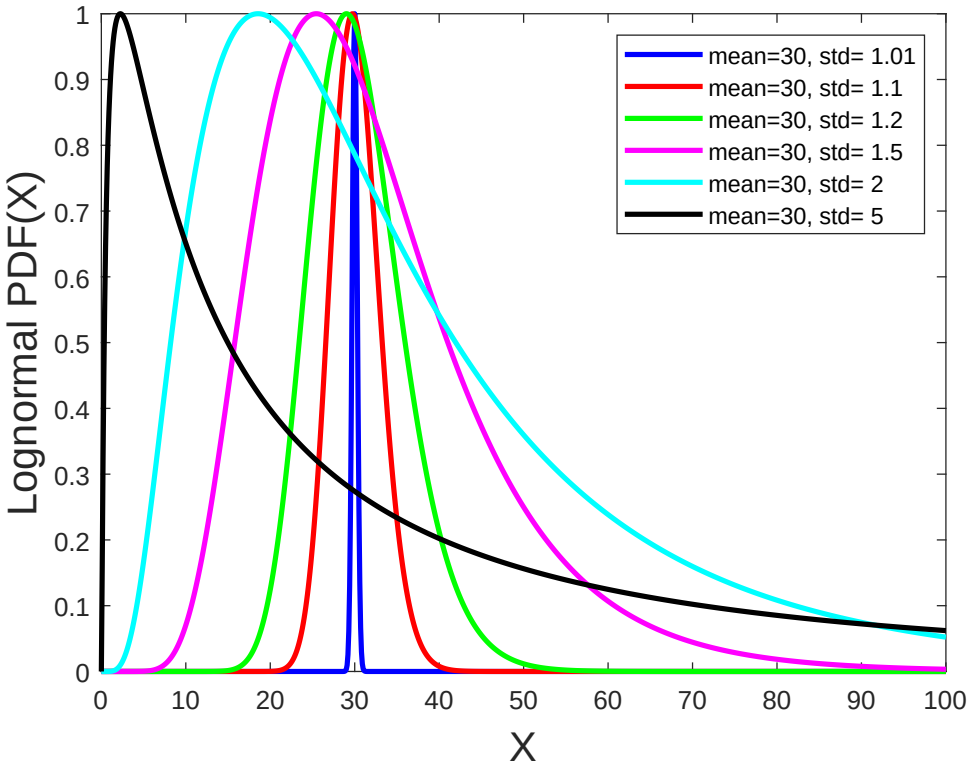

Figure B.12: Lognormal PDF for different mean and standard deviations (std).

of the grain sizes of the films from chapter 5. The lognormal distribution is similar to a normal distribution in that if a random variable, X, is lognormally distributed with mean, μ, and standard deviation, σ, then the random variable $ln(X)$ is normally distributed with with mean, $ln(\mu)$, and standard deviation, $log(\sigma)$. The probability

density function (PDF) for lognormal distributed random variable X is given as:

$$PDF(X,\mu,\sigma) = \frac{1}{X\sigma\sqrt{2\pi}} exp\left(\frac{-(ln(x)-\mu)^2}{2\sigma^2}\right) \tag{B.1}$$

In Fig. B.12, we show the dependence of the lognormal probability density function on the mean and standard deviation (labeled 'std'). Here, we can see that the standard deviation signficantly affects how much of a tail there is in the distribution. For low standard deviations, the distribution is almost symmetric with the mean occuring near the peak in the distribution. For larger standard deviations, tail of the distribution gets significantly larger causing the mean to move higher X values than where the peak occurs.

In Fig. B.13, we show a comparison of distributions similar to the lognormal used in the analysis from chapter 5. Here, we compare the lognormal to the weibull and gamma distributions for one of the Pd films from chapter 5. We can see from this comparison that the lognormal distribution results in the best fit to the histogram data. Additionally, we can that all three distributions give very similar results for their mean values suggesting our analysis is not strongly dependant on the choice of distribution.

B.3 Additional grain size analysis for other films

In this section we will show results for the grain size analysis for all of the films used in the results from chapter 5.

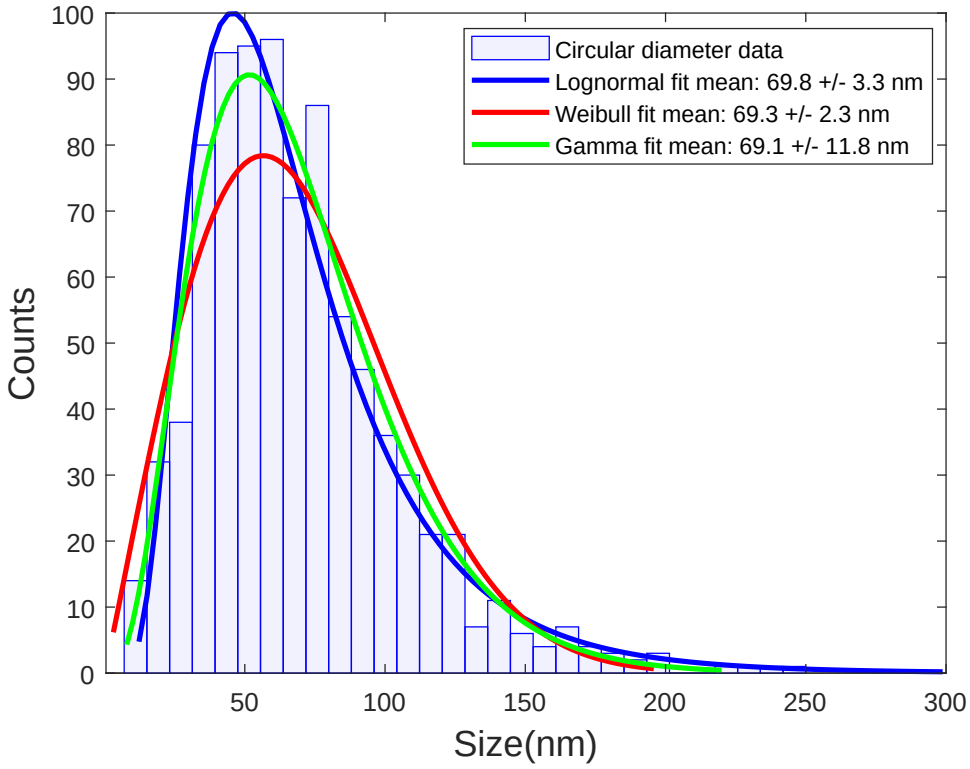

Figure B.13: Comparison of different distributions used in the grain size analysis.

B.3.1 Summary of average grains sizes from mean of lognormal distribution fits

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-EBAL-11	Al	N/A	43.0 ± 1.5	
RS-EBAL-9	Al	250°C with 5 %H2/N2		54.5 ± 1.8
RS-EBAL-7	Al	350°C with 5 %H2/N2	43.0 ± 1.5	47.4 ± 1.9
RS-DOWDW-6	Al	350°C with 10%H2/N2		44.8 ± 9.0
RS-EBAL-8	Al	425°C with 5 %H2/N2		47.1 ± 1.6
RS-DOWDW-7	Al	425°C with 10%H2/N2		39.5 ± 6.8

Table B.1: Average grain size summary for Al films from chapter 5.

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-NFDRY-16	Pd	250°C with 5 %H2/N2		30.0 ± 4.0
RS-NFDRY-14	Pd	350°C with 5 %H2/N2	24.1 ± 2.3	54.6 ± 2.1
RS-DOWDW-4	Pd	350°C with 10%H2/N2		61.9 ± 2.0
RS-NFDRY-15	Pd	425°C with 5 %H2/N2	21.8 ± 2.1	69.7 ± 3.2
RS-DOWDW-5	Pd	425°C with 10%H2/N2		65.4 ± 2.7

Table B.2: Average grain size summary for Ti/Pd films from chapter 5.

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-DOXDW-1	Pt	CTE measurement to 110C		20.6 ± 1.3
RS-NFDRY-21	Pt	250°C with 5 %H2/N2		19.4 ± 1.2
RS-NFDRY-19	Pt	350°C with 5 %H2/N2		49.6 ± 2.1
RS-DOXDW-2	Pt	350°C with 10%H2/N2		31.2 ± 5.5
RS-NFDRY-20	Pt	425°C with 5 %H2/N2	19.8 ± 1.3	82.7 ± 3.9
RS-DOXDW-3	Pt	425°C with 10%H2/N2		52.7 ± 1.6

Table B.3: Average grain size summary for Ti/Pt films from chapter 5.

B.3.2 Summary of grains size analysis from standard deviation of lognormal distribution fits

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-EBAL-11	Al	N/A	21.6 ± 1.5	
RS-EBAL-9	Al	250°C with 5 %H2/N2		28.3 ± 2.0
RS-EBAL-7	Al	350°C with 5 %H2/N2	23.3 ± 1.8	25.7 ± 2.2
RS-DOWDW-6	Al	350°C with 10%H2/N2		18.2 ± 1.0
RS-EBAL-8	Al	425°C with 5 %H2/N2		25.2 ± 1.8
RS-DOWDW-7	Al	425°C with 10%H2/N2		15.5 ± 0.7

Table B.4: Lognormal distribution standard deviations for Al films from chapter 5.

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-NFDRY-16	Pd	250°C with 5 %H2/N2		11.7 ± 4.1
RS-NFDRY-14	Pd	350°C with 5 %H2/N2	8.3 ± 2.2	31.7 ± 2.5
RS-DOWDW-4	Pd	350°C with 10%H2/N2		28.8 ± 2.1
RS-NFDRY-15	Pd	425°C with 5 %H2/N2	7.8 ± 1.9	39.9 ± 3.8
RS-DOWDW-5	Pd	425°C with 10%H2/N2		34.8 ± 3.0

Table B.5: Lognormal distribution standard deviations for Ti/Pd films from chapter 5.

Wafer	Material	Forming gas anneal	As-deposited grain size (nm)	Post anneal grain size (nm)
RS-DOXDW-1	Pt	CTE measurement to 110C		5.5 ± 0.2
RS-NFDRY-21	Pt	250°C with 5 %H2/N2		4.2 ± 0.2
RS-NFDRY-19	Pt	350°C with 5 %H2/N2		31.5 ± 2.6
RS-DOXDW-2	Pt	350°C with 10%H2/N2		14.3 ± 0.5
RS-NFDRY-20	Pt	425°C with 5 %H2/N2	5.6 ± 0.2	41.8 ± 0.2
RS-DOXDW-3	Pt	425°C with 10%H2/N2		25.9 ± 1.7

Table B.6: Lognormal distribution standard deviations for Ti/Pt films from chapter 5.

Appendix C: Setting up COMSOL simulations

This appendix provides the steps neccessary for reproducing a COMSOL model similar to one used to simulate tunnel junction (TJ) devices in Chapters 4 and 6.

C.1 Initial model setup

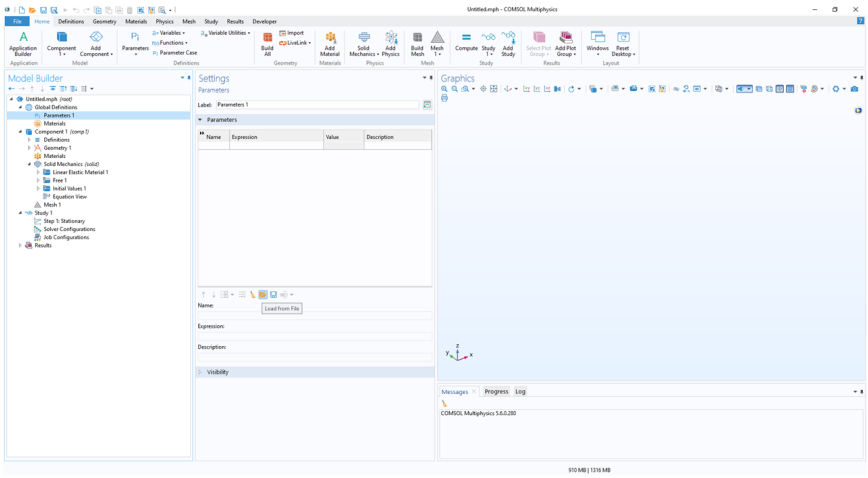

Figure C.1: Blank model produced by COMSOL model builder.

1. Open a blank model using the 'Model Wizard'.

2. Select a '3D' space dimension.

3. Select 'Structural Mechanics', then 'Solid Mechanics(solid)' for Physics.

4. Add 'Solid Mechanics(solid)' and then Hit 'Study'.

5. Select 'Stationary' for the study and then hit 'Done'. Screen should now look like Fig. C.1.

C.2 Building the model geometry

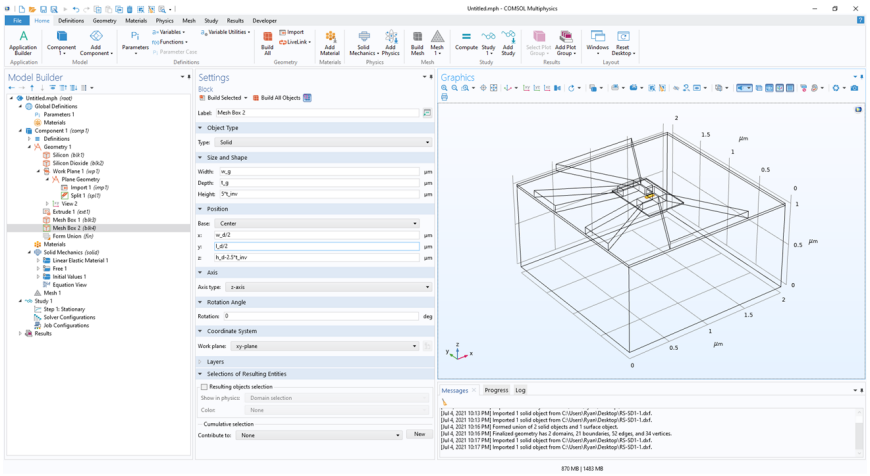

Figure C.2: Geometry of TJ device.

1. Select 'Parameters 1' under 'Global Definitions'.

2. Load the file 'TJparams.txt' `\SET_team\Ryan\COMSOL\`.

3. Click on 'Geometry 1', change 'Length unit' to μm.

4. Right click on 'Geometry 1', and add a 'Block'.

5. Adjust Settings for 'Block 1', which is the silicon substrate block. Enter 'w_d' for width, 'l_d' for length, and 'h_d' for height. Hit build.

6. Add another block. Settings for 'Block 2', which is the silicon dioxide block. Enter 'w_d' for width, 'l_d' for length, and 't_ox' for height and change the position to be (0,0,h_d) from the corner. Hit build.

7. Right click on 'Geometry 1', and select 'Work Plane'.

8. Adjust settings for the 'Work Plane 1'. Set z-coordinate as 'h_d+t_ox', xw-displacement as 'w_d/2', and yw-displacement as 'l_d/2'.

9. Right click on 'Plane Geometry', and select 'Import'.

10. Import the TJ gate lithography file 'RS-SD1-1.dxf' stored under `\SET_team\Ryan\COMSOL\` Set the relative repair tolerance to '1E-8' to make sure the scaling is correct.

11. Right click on 'Plane Geometry', and select 'Split'.

12. Settings for 'Split': select the gate layer ('imp1') as the input object. This will segment the gate layer into individual blocks which is useful for modifying the structure in COMSOL rather than importing a new file.

13. Right click on 'Work Plane', and select 'Extrude'. Settings for 'Extrude': set the distance as 't_Al' to make the plane the gate thickness.

14. Add a new block for meshing using settings: 'w_d/2-.55[um]' for width, 'l_d/2-.35[um]' for length, and '5*t_inv' for height and change the position to be (w_d/2-0.025,l_d/2,h_d-2.5*t_inv) from the center.

15. Add another new block for meshing TJ gap region using settings: 'w_g' for width, 't_g' for length, and '5*t_inv' for height and change the position to be (w_d/25,l_d/2,h_d-2.5*t_inv) from the center. Screen should now look like Fig. C.2.

C.3 Adding materials to the model

For the purposes of this example, we will use the materials library built into COMSOL. It is important to note that these values are not neccessarily correct and the user should manually adjust the parameters when needed.

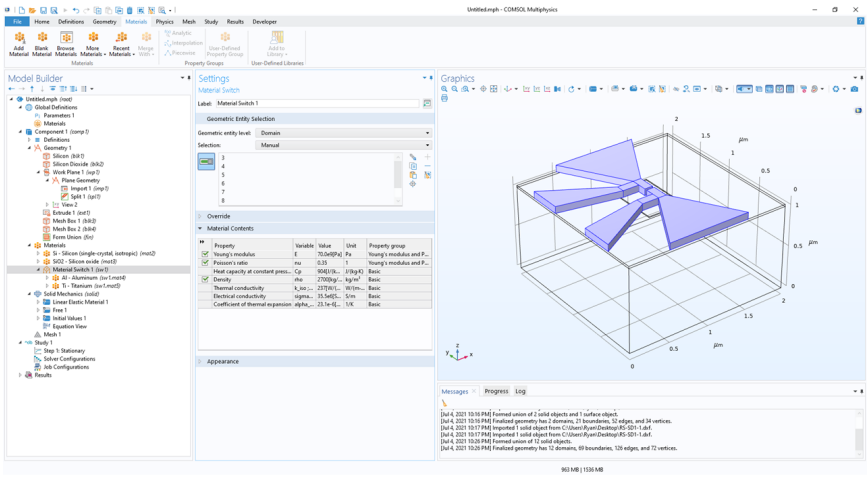

Figure C.3: 'Materials' section after adding all materials for TJ model.

1. Right click on 'Materials', and select 'Add Material from Library'.

2. Add 'Si-Silicon(single-crystal,isotropic)' from 'MEMS-Semiconductors' on the 'Add Material' window.

3. Add 'SiO2-silicon oxide' from 'MEMS-Insulators' on the 'Add Material' window.

4. Add 'Al-Aluminum' from 'MEMS-Metals' on the 'Add Material' window.

5. Add 'Ti-Titanium' from 'MEMS-Metals' on the 'Add Material' window.

6. Add 'Silicon' to Domains 1,9, and 11 from the settings window.

7. Add 'SiO2' to Domain 2 from the settings window.

8. Click on 'More Materials' at the top of COMSOL window and select 'Material Switch' so we can setup to sweep the gate material.

9. Add all the gate layer domains to the 'Material Switch'. Drag and drop 'Al' and 'Ti' into the 'Material Switch'. Screen should now look like Fig. C.3.

C.4 Setting up the Solid Mechanic module

1. Default settings for 'Solid Mechanics' are fine for this example. No changes needed.

2. Right Click on 'Solid Mechanics', then select 'Linear Elastic Material' from the 'Material Models' menu.

3. 'Linear Elastic Material 2' settings shown for silicon: set the model as 'Orthotropic' with ordering of 'Standard(11,22,33,12,23,13)'. Set E as '169E9,169E9,130E9', ν as '0.064,0.36,0.28', and G as '50.9E9,79.6E9,79.6E9'.

4. Right click on 'Linear Elastic Material 1' and add 'Thermal Expansion'.

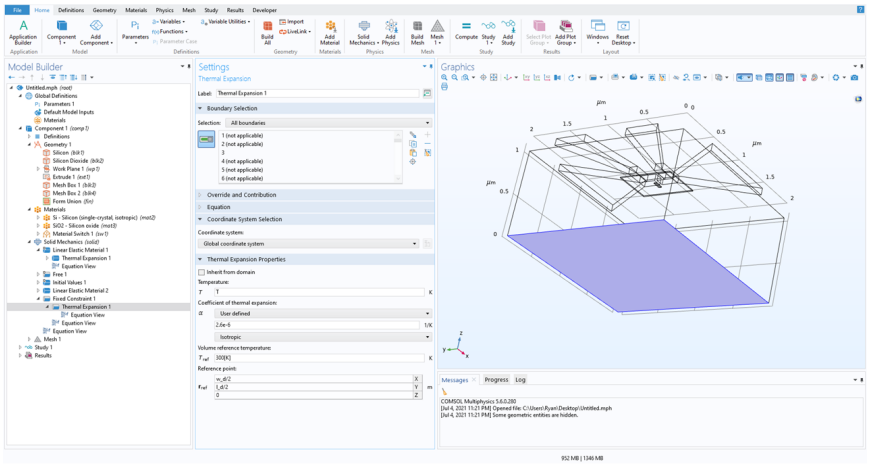

Figure C.4: Solid mechanics module after setting up Fixed constraint.

5. Add all the gate and silicon dioxide domains. Settings for 'Linear Elastic Material 1-Thermal Expansion': Set 'Tref' to 300[K] and 'Temperature' to T.

6. Add 'Thermal expansion' to 'Linear Elastic Material 2' but only for silicon domains (1,9,11). Settings: Set 'Tref' to 300[K] and 'Temperature' to T.

7. Right click on 'Solid Mechanics' and add 'Fixed Constraint (boundary)'.

8. Add the boundary for the bottom of the silicon substrate (boundary 3) to 'Fixed Constraint'.

9. Right click on 'Fixed Constraint' and add 'Thermal Expansion'.

10. Settings for 'Fixed Constraint-Thermal Expansion': set α as '2.6E-6', 'Tref' to 300[K], 'Temperature' to T, and the reference point as (w_d/2,l_d/2,0). This is neccessary to keep the displacement zero at this contact for all temperatures, otherwise it would only be zero at the reference temperature. Screen should

now look like Fig. C.4.

C.5 Meshing the model

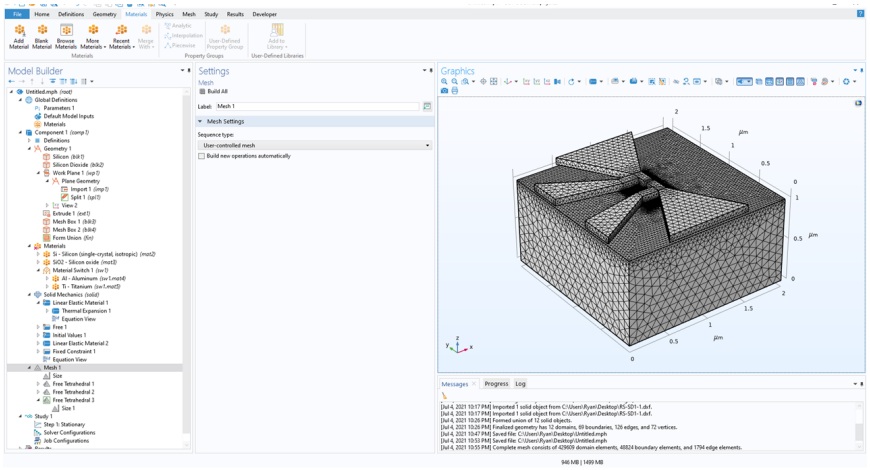

Figure C.5: COMSOL model after meshing.

1. Right click on 'Mesh 1' and add 'Free Tetrahedral' to Domain 9.

2. Right click on 'Free Tetrahedral 1',add 'Size', and use manual settings: 'Maximum element size: 0.01','Minimum element size: 1e-5', and 'Resolution: 1.2'.

3. Right click on 'Mesh 1' and add 'Free Tetrahedral' to Domain 11.

4. Right click on 'Free Tetrahedral 2',add 'Size', and use manual settings: 'Maximum element size: 0.002','Minimum element size: 2e-6', and 'Resolution: 1.2'.

5. Right click on 'Mesh 1' and add 'Free Tetrahedral' to remaining domains

6. Right click on 'Free Tetrahedral 3', Add 'Size', and used pre-build 'finer' setting.

7. RHit 'Build all' to mesh the model. Screen should now look like Fig. C.5

C.6 Running the computation

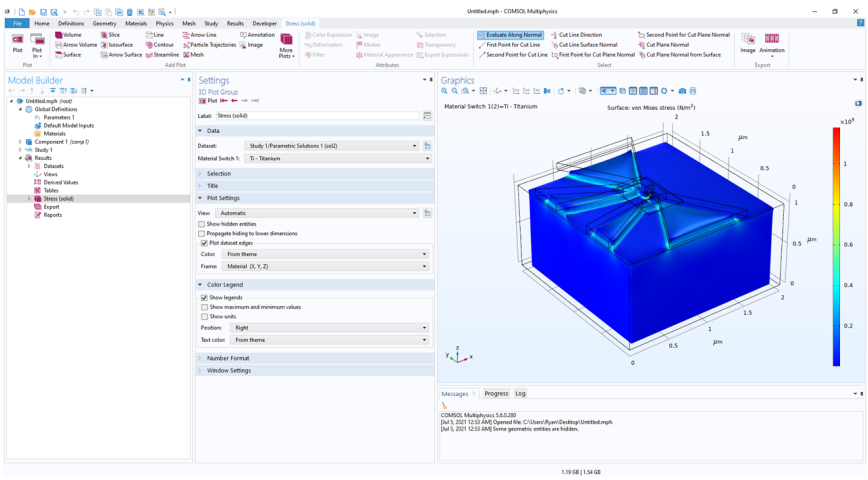

Figure C.6: Default 3d stress plot in the 'Results' section.

1. Use default settings for 'Study 1'. No changes needed.

2. Right click on 'Study 1' and add 'Material Sweep'.

3. Settings for 'Material Sweep': make sure 'Cases' is set to 'All'.

4. Use default settings for 'Stationary 1'. No changes needed.

5. Run the study by hitting 'Compute'.

6. Default 3d stress plot should up in the 'Results' section after the computation finishes. Screen should now look like Fig. C.6

C.7 Plotting 1d linecuts of ΔE_c

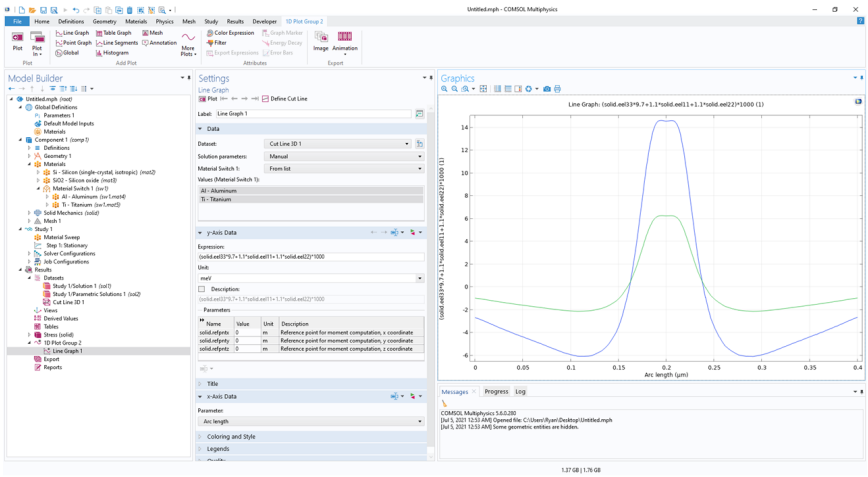

Figure C.7: 1d cuts of ΔE_C through the tunnel junction for Al and Ti gates.

1. Right click on 'Datasets' and add 'Cut Line 3D'.

2. Setting for 'Cut Line 3D': Set the dataset as 'Study1/Parametric Solutions 1 (sol2)', 'Point 1' = (w_d/2, l_d-200[nm], h_d-t_inv) and 'Point 2' = (w_d/2, l_d+200[nm], h_d-t_inv)

3. Right click on 'Results' and add '1D Plot Group'.

4. Right click on '1D Plot Group' and add 'Line Graph'.

5. Setting for 'Cut Line 3D': Set the expression as '(solid.eel33*9.7+ 1.1*solid.eel11+ 1.1*solid.ee22) *100', set the units as 'meV'. Hit plot. Screen should now look like Fig. C.7

Appendix D: AC tunnel junction measurements

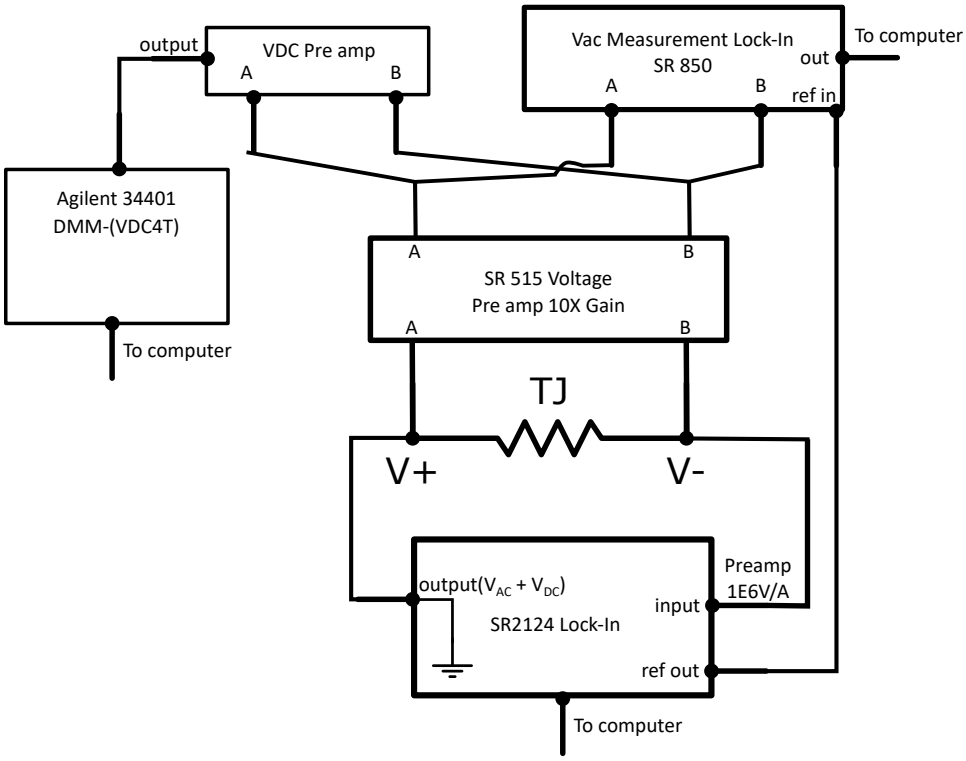

Figure D.1: Circuit diagram for AC measurements of TJ devices.

In this final appendix, we will discuss an alternative measurement setup that we worked on in parallel to the asymmetric TJ measurements. All of the TJ measurements presented previously were measured using standard DC measurement setup of an Agilent 4156C parameter analyzer, where to measure currents we use the source measurement units (SMU) and to measure voltages we use voltage measurement

units (VMU). In fitting data using the BRD model, we typically want to use the conductance through the TJ as a function of bias $G(V)$, we need to take numerical derivatives of the raw DC $I(V)$ data. This comes with a downside of significant noise on the resulting conductance data. In previous measurements, we employed sampling techniques to try to average out some of this noise on $G(V)$, which significantly increases the measurement time. An alternative to the DC measurements typically used in our TJ measurements, are low-frequency AC measurements using lock-in amplifier techniques. We decided to implement an AC measurement setup for our TJs based on a few factors. First, using a lock-in amplifier allows for a direct measurement of the conductance without the need for numerical derivatives necessary for the DC measurements. This will lead to a significant improvement in the signal to noise ratio of the $G(V)$ data. Second, our lock-in setup allows us to bypass some limitations placed on the voltage source by our use of the 4156C parameter analyzer. The 4156C is limited to voltage steps of 200 μV. This is problematic for doing TJ fitting because for small barriers this voltage step often leaves us with only a few points to fit. The lock-in amplifier is capable of applying significantly smaller DC voltage steps (during our testing we typically used 10 μV). The third and final reason is to avoid some of the low frequency noise in our measurement setup. Since the lock-in amplifier is operating at a specific frequency, we will be able to tune the frequency in attempt to reject some of the lower frequency components. For example in our testing, we found using AC frequencies between 31-300 Hz leads to $G(V)$ data with significantly better signal to noise than from our DC measurements. We tested several different versions of the AC measurement setup and the one that

we found had the best performance is shown in Fig. D.1. In this circuit, we use two different lock-in amplifiers: one to measure the AC current (I_{AC}) and one to measure the 4-terminal AC voltage drop across the TJ (V_{AC}). We use an SRS SR2124 analog lock-in amplifier to apply DC bias, the AC excitation, and measure I_{AC} and use the SRS SR850 digital lock-in to measure V_{AC}. For voltage measurements, we use two different voltage pre-amplifiers. The SRS SR515 is connected directly to the TJ voltage probes because it has input impedance of 1TΩ, which is importantly the large resistances of our TJs. We use an additional SRS SR570 voltage pre-amplifier to provide some additional filtering and gain on the DC voltage measurement, since the SR515 is limited to a 10X gain. Due to the UQD issues mentioned in chapter 6 with our TJs, we have not as of yet used this AC measurement setup for any fitting or analysis of TJ device data. We note that in any future work with TJs, this AC measurement will be useful for obtaining $G(V)$ data with significantly reduced noise than the differential $G(V)$ data from the DC measurements.

Bibliography

[1] Allan F. Bower. *Applied Mechanics of Solids*. CRC Press, October 2009.

[2] A Shirkhorshidian, N C Bishop, J Dominguez, R K Grubbs, J R Wendt, M P Lilly, and M S Carroll. Transport spectroscopy of low disorder silicon tunnel barriers with and without Sb implants. *Nanotechnology*, 26(20):205703, May 2015.

[3] Neil M. Zimmerman and Mark W. Keller. Electrical metrology with single electrons. *Measurement Science and Technology*, 14(8):1237–1242, July 2003.

[4] David Berman, Nikolai B. Zhitenev, Raymond C. Ashoori, Henry I. Smith, and Michael R. Melloch. Single-electron transistor as a charge sensor for semiconductor applications. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, 15(6):2844–2847, November 1997.

[5] Tobias Bergsten, Tord Claeson, and Per Delsing. A fast, primary Coulomb blockade thermometer. *Applied Physics Letters*, 78(9):1264–1266, February 2001.

[6] Vladimir Bubanja, Kazuhiko Matsumoto, and Yoshitaka Gotoh. Single Electron Memory at Room Temperature: Experiment and Simulation. *Japanese Journal of Applied Physics*, 40(1R):87, January 2001.

[7] Kosuke Maeda, Norio Okabayashi, Shinya Kano, Shuhei Takeshita, Daisuke Tanaka, Masanori Sakamoto, Toshiharu Teranishi, and Yutaka Majima. Logic Operations of Chemically Assembled Single-Electron Transistor. *ACS Nano*, 6(3):2798–2803, March 2012.

[8] Gento Yamahata, Katsuhiko Nishiguchi, and Akira Fujiwara. Gigahertz single-trap electron pumps in silicon. *Nature Communications*, 5(1):5038, December 2014.

[9] Alessandro Rossi, Tuomo Tanttu, Kuan Yen Tan, Ilkka Iisakka, Ruichen Zhao, Kok Wai Chan, Giuseppe C. Tettamanzi, Sven Rogge, Andrew S. Dzurak,

and Mikko Möttönen. An Accurate Single-Electron Pump Based on a Highly Tunable Silicon Quantum Dot. *Nano Letters*, 14(6):3405–3411, June 2014.

[10] S. P. Giblin, M. Kataoka, J. D. Fletcher, P. See, T. J. B. M. Janssen, J. P. Griffiths, G. a. C. Jones, I. Farrer, and D. A. Ritchie. Towards a quantum representation of the ampere using single electron pumps. *Nature Communications*, 3(1):930, July 2012.

[11] 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. A two-qubit logic gate in silicon. *Nature*, 526(7573):410–414, October 2015.

[12] 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. Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots. *Science*, 309(5744):2180–2184, September 2005.

[13] Daniel Loss and David P. DiVincenzo. Quantum computation with quantum dots. *Physical Review A*, 57(1):120–126, January 1998.

[14] David P. DiVincenzo. The Physical Implementation of Quantum Computation. *Fortschritte der Physik*, 48(9-11):771–783, 2000.

[15] J. I. Cirac and P. Zoller. Quantum Computations with Cold Trapped Ions. *Physical Review Letters*, 74(20):4091–4094, May 1995.

[16] Rainer Blatt and David Wineland. Entangled states of trapped atomic ions. *Nature*, 453(7198):1008–1015, June 2008.

[17] M.F. Bocko, A.M. Herr, and M.J. Feldman. Prospects for quantum coherent computation using superconducting electronics. *IEEE Transactions on Applied Superconductivity*, 7(2):3638–3641, June 1997.

[18] Y. Nakamura, Yu A. Pashkin, and J. S. Tsai. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398(6730):786–788, April 1999.

[19] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret. Manipulating the Quantum State of an Electrical Circuit. *Science*, 296(5569):886–889, May 2002.

[20] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij. Coherent Quantum Dynamics of a Superconducting Flux Qubit. *Science*, 299(5614):1869–1871, March 2003.

[21] Ivan H. Deutsch, Gavin K. Brennen, and Poul S. Jessen. Quantum Computing with Neutral Atoms in an Optical Lattice. *Fortschritte der Physik*, 48(9-11):925–943, 2000.

[22] Oliver Morsch and Markus Oberthaler. Dynamics of Bose-Einstein condensates in optical lattices. *Reviews of Modern Physics*, 78(1):179–215, February 2006.

[23] Alberto Politi, Jonathan C. F. Matthews, and Jeremy L. O'Brien. Shor's Quantum Factoring Algorithm on a Photonic Chip. *Science*, 325(5945):1221–1221, September 2009.

[24] Jonathan C. F. Matthews, Alberto Politi, André Stefanov, and Jeremy L. O'Brien. Manipulation of multiphoton entanglement in waveguide quantum circuits. *Nature Photonics*, 3(6):346–350, June 2009.

[25] David G. Cory, Amr F. Fahmy, and Timothy F. Havel. Ensemble quantum computing by NMR spectroscopy. *Proceedings of the National Academy of Sciences*, 94(5):1634–1639, March 1997.

[26] Neil A. Gershenfeld and Isaac L. Chuang. Bulk Spin-Resonance Quantum Computation. *Science*, 275(5298):350–356, January 1997.

[27] A. Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, January 2003.

[28] Michael H. Freedman. Quantum Computation and the Localization of Modular Functors. *Foundations of Computational Mathematics*, 1(2):183–204, January 2001.

[29] J. Harrison, M. J. Sellars, and N. B. Manson. Measurement of the optically induced spin polarisation of N-V centres in diamond. *Diamond and Related Materials*, 15(4):586–588, April 2006.

[30] M. V. Gurudev Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin. Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond. *Science*, 316(5829):1312–1316, June 2007.

[31] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen. Spins in few-electron quantum dots. *Reviews of Modern Physics*, 79(4):1217–1265, October 2007.

[32] B. E. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393(6681):133–137, May 1998.

[33] Gento Yamahata, Stephen P. Giblin, Masaya Kataoka, Takeshi Karasawa, and Akira Fujiwara. Gigahertz single-electron pumping in silicon with an accuracy better than 9.2 parts in 107. *Applied Physics Letters*, 109(1):013101, July 2016.

[34] Neil M. Zimmerman. A primer on electrical units in the Système International. *American Journal of Physics*, 66(4):324–331, April 1998.

[35] Michael Stock, Richard Davis, Estefanía de Mirandés, and Martin J T Milton. The revision of the SI—the result of three decades of progress in metrology. *Metrologia*, 56(2):022001, April 2019.

[36] A. Hartland. The Quantum Hall Effect and Resistance Standards. *Metrologia*, 29(2):175–190, January 1992.

[37] C.A. Hamilton, C.J. Burroughs, and S.P. Benz. Josephson voltage standard-a review. *IEEE Transactions on Applied Superconductivity*, 7(2):3756–3761, June 1997.

[38] Nobu-Hisa Kaneko, Shuji Nakamura, and Yuma Okazaki. A review of the quantum current standard. *Measurement Science and Technology*, 27(3):032001, February 2016.

[39] Leo P. Kouwenhoven, Charles M. Marcus, Paul L. McEuen, Seigo Tarucha, Robert M. Westervelt, and Ned S. Wingreen. Electron Transport in Quantum Dots. In Lydia L. Sohn, Leo P. Kouwenhoven, and Gerd Schön, editors, *Mesoscopic Electron Transport*, NATO ASI Series, pages 105–214. Springer Netherlands, Dordrecht, 1997.

[40] E. B. Foxman, P. L. McEuen, U. Meirav, Ned S. Wingreen, Yigal Meir, Paul A. Belk, N. R. Belk, M. A. Kastner, and S. J. Wind. Effects of quantum levels on transport through a Coulomb island. *Physical Review B*, 47(15):10020–10023, April 1993.

[41] C. Livermore, C. H. Crouch, R. M. Westervelt, K. L. Campman, and A. C. Gossard. The Coulomb Blockade in Coupled Quantum Dots. *Science*, 274(5291):1332–1335, November 1996.

[42] F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Single-Electron Charging in Double and Triple Quantum Dots with Tunable Coupling. *Physical Review Letters*, 75(4):705–708, July 1995.

[43] F. R. Waugh, M. J. Berry, C. H. Crouch, C. Livermore, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Measuring interactions between tunnel-coupled quantum dots. *Physical Review B*, 53(3):1413–1420, January 1996.

[44] 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. Driven coherent oscillations of a single electron spin in a quantum dot. *Nature*, 442(7104):766–771, August 2006.

[45] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven. Single-shot read-out of an individual electron spin in a quantum dot. *Nature*, 430(6998):431–435, July 2004.

[46] R. Brunner, Y.-S. Shin, T. Obata, M. Pioro-Ladrière, T. Kubo, K. Yoshida, T. Taniyama, Y. Tokura, and S. Tarucha. Two-Qubit Gate of Combined Single-Spin Rotation and Interdot Spin Exchange in a Double Quantum Dot. *Physical Review Letters*, 107(14):146801, September 2011.

[47] D. J. Reilly, J. M. Taylor, E. A. Laird, J. R. Petta, C. M. Marcus, M. P. Hanson, and A. C. Gossard. Measurement of Temporal Correlations of the Overhauser Field in a Double Quantum Dot. *Physical Review Letters*, 101(23):236803, December 2008.

[48] Floris A. Zwanenburg, Andrew S. Dzurak, Andrea Morello, Michelle Y. Simmons, Lloyd C. L. Hollenberg, Gerhard Klimeck, Sven Rogge, Susan N. Coppersmith, and Mark A. Eriksson. Silicon quantum electronics. *Reviews of Modern Physics*, 85(3):961–1019, July 2013.

[49] K. J. Dwyer, J. M. Pomeroy, D. S. Simons, K. L. Steffens, and J. W. Lau. Enriching28Si beyond 99.9998 % for semiconductor quantum computing. *Journal of Physics D: Applied Physics*, 47(34):345105, August 2014.

[50] M. D. Blumenthal, B. Kaestner, L. Li, S. Giblin, T. J. B. M. Janssen, M. Pepper, D. Anderson, G. Jones, and D. A. Ritchie. Gigahertz quantized charge pumping. *Nature Physics*, 3(5):343–347, May 2007.

[51] Stephen P. Giblin, Akira Fujiwara, Gento Yamahata, Myung-Ho Bae, Nam Kim, Alessandro Rossi, Mikko Möttönen, and Masaya Kataoka. Evidence for universality of tunable-barrier electron pumps. *Metrologia*, 56(4):044004, July 2019.

[52] M. Stewart and Neil Zimmerman. Stability of Single Electron Devices: Charge Offset Drift. *Applied Sciences*, 6(7):187, June 2016.

[53] 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. Electrostatically defined few-electron double quantum dot in silicon. *Applied Physics Letters*, 94(17):173502, April 2009.

[54] C. B. Simmons, Madhu Thalakulam, Nakul Shaji, Levente J. Klein, Hua Qin, R. H. Blick, D. E. Savage, M. G. Lagally, S. N. Coppersmith, and M. A. Eriksson. Single-electron quantum dot in Si/SiGe with integrated charge sensing. *Applied Physics Letters*, 91(21):213103, November 2007.

[55] Martin Fuechsle, Jill A. Miwa, Suddhasatta Mahapatra, Hoon Ryu, Sunhee Lee, Oliver Warschkow, Lloyd C. L. Hollenberg, Gerhard Klimeck, and Michelle Y. Simmons. A single-atom transistor. *Nature Nanotechnology*, 7(4):242–246, April 2012.

[56] D. M. Zajac, A. J. Sigillito, M. Russ, F. Borjans, J. M. Taylor, G. Burkard, and J. R. Petta. Resonantly driven CNOT gate for electron spins. *Science*, 359(6374):439–442, January 2018.

[57] Jarryd J. Pla, Kuan Y. Tan, Juan P. Dehollain, Wee H. Lim, John J. L. Morton, David N. Jamieson, Andrew S. Dzurak, and Andrea Morello. A single-atom electron spin qubit in silicon. *Nature*, 489(7417):541–545, September 2012.

[58] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, September 2012.

[59] E. P. Nordberg, G. A. Ten Eyck, H. L. Stalford, R. P. Muller, R. W. Young, K. Eng, L. A. Tracy, K. D. Childs, J. R. Wendt, R. K. Grubbs, J. Stevens, M. P. Lilly, M. A. Eriksson, and M. S. Carroll. Enhancement-mode double-top-gated metal-oxide-semiconductor nanostructures with tunable lateral geometry. *Physical Review B*, 80(11):115331, September 2009.

[60] Y. Ono, K. Nishiguchi, A. Fujiwara, H. Yamaguchi, H. Inokawa, and Y. Takahashi. Conductance modulation by individual acceptors in Si nanoscale field-effect transistors. *Applied Physics Letters*, 90(10):102106, March 2007.

[61] Adnan Afiff, Arup Samanta, Arief Udhiarto, Harry Sudibyo, Masahiro Hori, Yukinori Ono, Michiharu Tabe, and Daniel Moraru. Coulomb-blockade transport in selectively-doped Si nano-transistors. *Applied Physics Express*, 12(8):085004, August 2019.

[62] Ted Thorbeck and Neil M. Zimmerman. Formation of strain-induced quantum dots in gated semiconductor nanostructures. *AIP Advances*, 5(8):087107, August 2015.

[63] J. Park, Y. Ahn, J. A. Tilka, K. C. Sampson, D. E. Savage, J. R. Prance, C. B. Simmons, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, M. V. Holt, and P. G. Evans. Electrode-stress-induced nanoscale disorder in Si quantum electronic devices. *APL Materials*, 4(6):066102, June 2016.

[64] Yasuo Takahashi, Yukinori Ono, Akira Fujiwara, and Hiroshi Inokawa. Development of silicon single-electron devices. *Physica E: Low-dimensional Systems and Nanostructures*, 19(1):95–101, July 2003.

[65] Susan J. Angus, Andrew J. Ferguson, Andrew S. Dzurak, and Robert G. Clark. Gate-Defined Quantum Dots in Intrinsic Silicon. *Nano Letters*, 7(7):2051–2055, July 2007.

[66] A. M. J. Zwerver, T. Krähenmann, T. F. Watson, L. Lampert, H. C. George, R. Pillarisetty, S. A. Bojarski, P. Amin, S. V. Amitonov, J. M. Boter, R. Caudillo, D. Corras-Serrano, J. P. Dehollain, G. Droulers, E. M. Henry, R. Kotlyar, M. Lodari, F. Luthi, D. J. Michalak, B. K. Mueller, S. Neyens, J. Roberts, N. Samkharadze, G. Zheng, O. K. Zietz, G. Scappucci, M. Veldhorst, L. M. K. Vandersypen, and J. S. Clarke. Qubits made by advanced

semiconductor manufacturing. *arXiv:2101.12650 [cond-mat, physics:quant-ph]*, January 2021.

[67] L. A. Tracy, E. P. Nordberg, R. W. Young, C. Borrás Pinilla, H. L. Stalford, G. A. Ten Eyck, K. Eng, K. D. Childs, J. R. Wendt, R. K. Grubbs, J. Stevens, M. P. Lilly, M. A. Eriksson, and M. S. Carroll. Double quantum dot with tunable coupling in an enhancement-mode silicon metal-oxide semiconductor device with lateral geometry. *Applied Physics Letters*, 97(19):192110, November 2010.

[68] S. Rochette, M. Rudolph, A.-M. Roy, M. J. Curry, G. A. Ten Eyck, R. P. Manginell, J. R. Wendt, T. Pluym, S. M. Carr, D. R. Ward, M. P. Lilly, M. S. Carroll, and M. Pioro-Ladrière. Quantum dots with split enhancement gate tunnel barrier control. *Applied Physics Letters*, 114(8):083101, February 2019.

[69] Binhui Hu, Erick D. Ochoa, Daniel Sanchez, Justin K. Perron, Neil M. Zimmerman, and M. D. Stewart. Effect of device design on charge offset drift in Si/SiO_2 single electron devices. *Journal of Applied Physics*, 124(14):144302, October 2018.

[70] D.M. Fleetwood, P.S. Winokur, M.R. Shaneyfelt, L.C. Riewe, O. Flament, P. Paillet, and J.L. Leray. Effects of isochronal annealing and irradiation temperature on radiation-induced trapped charge. *IEEE Transactions on Nuclear Science*, 45(6):2366–2374, December 1998.

[71] Stephen J. Fonash. An Overview of Dry Etching Damage and Contamination Effects. *Journal of The Electrochemical Society*, 137(12):3885, December 1990.

[72] Brian Chi Ho Mooy, Kuan Yen Tan, and Nai Shyan Lai. Comparison of Strain Effect between Aluminum and Palladium Gated MOS Quantum Dot Systems. *Universe*, 6(4):51, April 2020.

[73] Ted Thorbeck and Neil M. Zimmerman. Formation of Strain-Induced Quantum Dots in Gated Semiconductor Nanostructures. *AIP Advances*, 5(8):087107, August 2015.

[74] J. Park, Y. Ahn, J. A. Tilka, K. C. Sampson, D. E. Savage, J. R. Prance, C. B. Simmons, M. G. Lagally, S. N. Coppersmith, M. A. Eriksson, M. V. Holt, and P. G. Evans. Electrode-Stress-Induced Nanoscale Disorder in Si Quantum Electronic Devices. *APL Materials*, 4(6):066102, June 2016.

[75] Martin H. Sadd. *Elasticity: Theory, Applications, and Numerics*. Academic Press, February 2009.

[76] M. A. Hopcroft, W. D. Nix, and T. W. Kenny. What is the Young's Modulus of Silicon? *Journal of Microelectromechanical Systems*, 19(2):229–238, April 2010.

[77] Y. S. Touloukian, R. K. Kirby, E. R. Taylor, and T. Y. R. Lee. Thermophysical Properties of Matter - the TPRC Data Series. Volume 13. Thermal Expansion - Nonmetallic Solids. Technical report, THERMOPHYSICAL AND ELECTRONIC PROPERTIES INFORMATION ANALYSIS CENTER LAFAYETTE IN, January 1977.

[78] Kevin M. Knowles. The Biaxial Moduli of Cubic Materials Subjected to an Equi-biaxial Elastic Strain. *Journal of Elasticity*, 124(1):1–25, June 2016.

[79] Y. S Touloukian. *Thermophysical Properties of Matter. Metallic Elements and Alloys Vol. 12, Vol. 12,*. IFI/Plenum, New York, 1975.

[80] E. Kobeda and E. A. Irene. A measurement of intrinsic SiO2 film stress resulting from low temperature thermal oxidation of Si. *Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena*, 4(3):720–722, May 1986.

[81] S. Timoshenko. Analysis of Bi-Metal Thermostats. *JOSA*, 11(3):233–255, September 1925.

[82] Peter YU and Manuel Cardona. *Fundamentals of Semiconductors: Physics and Materials Properties*. Springer Science & Business Media, April 2010.

[83] Frank Stern and W. E. Howard. Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit. *Physical Review*, 163(3):816–835, November 1967.

[84] J. Bardeen and W. Shockley. Deformation Potentials and Mobilities in Non-Polar Crystals. *Physical Review*, 80(1):72–80, October 1950.

[85] Conyers Herring and Erich Vogt. Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering. *Physical Review*, 101(3):944–961, February 1956.

[86] I. Balslev. The influence of high uniaxial stress on the indirect absorption edge in silicon. *Solid State Communications*, 3(8):213–218, August 1965.

[87] Sven Riskaer. Determination of Shear Deformation Potentials from the Free-Carrier Piezobirefringence in Germanium and Silicon. *Physical Review*, 152(2):845–849, December 1966.

[88] J. E. Aubrey, W. Gubler, T. Henningsen, and S. H. Koenig. Piezoresistance and Piezo-Hall-Effect in n -Type Silicon. *Physical Review*, 130(5):1667–1670, June 1963.

[89] Ji-Song Lim, Xiaodong Yang, Toshikazu Nishida, and Scott E. Thompson. Measurement of conduction band deformation potential constants using gate direct tunneling current in n-type metal oxide semiconductor field effect transistors under mechanical stress. *Applied Physics Letters*, 89(7):073509, August 2006.

[90] M. V. Fischetti and S. E. Laux. Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys. *Journal of Applied Physics*, 80(4):2234–2252, August 1996.

[91] E. Ungersboeck, S. Dhar, G. Karlowatz, V. Sverdlov, H. Kosina, and S. Selberherr. The Effect of General Strain on the Band Structure and Electron Mobility of Silicon. *IEEE Transactions on Electron Devices*, 54(9):2183–2190, September 2007.

[92] G. Gerald Stoney. The Tension of Metallic Films Deposited by Electrolysis. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 82(553):172–175, 1909.

[93] G. C. A. M. Janssen, M. M. Abdalla, F. van Keulen, B. R. Pujada, and B. van Venrooy. Celebrating the 100th anniversary of the Stoney equation for film stress: Developments from polycrystalline steel strips to single crystal silicon wafers. *Thin Solid Films*, 517(6):1858–1867, January 2009.

[94] X. Feng, Y. Huang, and A. J. Rosakis. On the Stoney Formula for a Thin Film/Substrate System With Nonuniform Substrate Thickness. *Journal of Applied Mechanics*, 74(6):1276–1281, January 2007.

[95] Daryl L. Logan. *A First Course in the Finite Element Method*. Cengage Learning, January 2011.

[96] Yuan Taur and Tak H. Ning. *Fundamentals of Modern VLSI Devices*. Cambridge University Press, Cambridge, second edition, 2009.

[97] Martin A. Green. Intrinsic concentration, effective densities of states, and effective mass in silicon. *Journal of Applied Physics*, 67(6):2944–2954, March 1990.

[98] Herbert B. Michaelson. The work function of the elements and its periodicity. *Journal of Applied Physics*, 48(11):4729–4733, November 1977.

[99] B. E. Deal. Standardized terminology for oxide charges associated with thermally oxidized silicon. *IEEE Transactions on Electron Devices*, 27(3):606–608, March 1980.

[100] Dieter K. Schroder. *Semiconductor Material and Device Characterization*. John Wiley & Sons, June 2015.

[101] Edward H. Nicollian and John R. Brews. *MOS (Metal Oxide Semiconductor) Physics and Technology*, volume 1987. Wiley New York, 1982.

[102] P. L. Castro and B. E. Deal. Low-Temperature Reduction of Fast Surface States Associated with Thermally Oxidized Silicon. *Journal of The Electrochemical Society*, 118(2):280, February 1971.

[103] Gary J. Gerardi, Edward H. Poindexter, Philip J. Caplan, and Noble M. Johnson. Interface traps and P_b centers in oxidized (100) silicon wafers. *Applied Physics Letters*, 49(6):348–350, August 1986.

[104] E. H. Poindexter, G. J. Gerardi, M.-E. Rueckel, P. J. Caplan, N. M. Johnson, and D. K. Biegelsen. Electronic traps and Pb centers at the Si/SiO2 interface: Band-gap energy distribution. *Journal of Applied Physics*, 56(10):2844–2849, November 1984.

[105] Gary J. Gerardi, Edward H. Poindexter, Philip J. Caplan, and Noble M. Johnson. Interface traps and Pb centers in oxidized (100) silicon wafers. *Applied Physics Letters*, 49(6):348–350, August 1986.

[106] Michael L. Reed and James D. Plummer. Chemistry of Si-SiO2 interface trap annealing. *Journal of Applied Physics*, 63(12):5776–5793, June 1988.

[107] G. Greeuw and J. F. Verwey. The mobility of Na+, Li+, and K+ ions in thermally grown SiO2 films. *Journal of Applied Physics*, 56(8):2218–2224, October 1984.

[108] E. H. Nicollian and A. Goetzberger. The Si-SiO2 Interface — Electrical Properties as Determined by the Metal-Insulator-Silicon Conductance Technique. *Bell System Technical Journal*, 46(6):1055–1133, 1967.

[109] Bernard Yurash and Bruce E. Deal. A Method for Determining Sodium Content of Semiconductor Processing Materials. *Journal of The Electrochemical Society*, 115(11):1191, November 1968.

[110] G. F. Derbenwick. Mobile ions in SiO2: Potassium. *Journal of Applied Physics*, 48(3):1127–1130, March 1977.

[111] C. W. Gwyn. Model for Radiation-Induced Charge Trapping and Annealing in the Oxide Layer of MOS Devices. *Journal of Applied Physics*, 40(12):4886–4892, November 1969.

[112] Genda J. Hu and Walter C. Johnson. Relationship between x-ray-produced holes and interface states in metal-oxide-semiconductor capacitors. *Journal of Applied Physics*, 54(3):1441–1444, March 1983.

[113] S. K. Lai and D. R. Young. Effects of avalanche injection of electrons into silicon dioxide—generation of fast and slow interface states. *Journal of Applied Physics*, 52(10):6231–6240, October 1981.

[114] Piyas Samanta and C. K Sarkar. Analysis of positive charge trapping in silicon dioxide of MOS capacitors during Fowler-Nordheim stress. *Solid-State Electronics*, 41(3):459–464, March 1997.

[115] J. M. Aitken. Radiation-induced trapping centers in thin silicon dioxide films. *Journal of Non-Crystalline Solids*, 40(1):31–47, July 1980.

[116] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow. Characteristics of the Surface-State Charge (Qss) of Thermally Oxidized Silicon. *Journal of The Electrochemical Society*, 114(3):266, 1967.

[117] D. M. Fleetwood. 'Border traps' in MOS devices. *IEEE Transactions on Nuclear Science*, 39(2):269–271, April 1992.

[118] W. L. Warren, D. M. Fleetwood, M. R. Shaneyfelt, P. S. Winokur, and R. A. B. Devine. Defect-defect hole transfer and the identity of border traps in ${\mathrm{SiO}}_{2}$ films. *Physical Review B*, 50(19):14710–14713, November 1994.

[119] W. L. Warren, M. R. Shaneyfelt, D. M. Fleetwood, J. R. Schwank, P. S. Winokur, and R. A. B. Devine. Microscopic nature of border traps in MOS oxides. *IEEE Transactions on Nuclear Science*, 41(6):1817–1827, December 1994.

[120] D. M. Fleetwood, M. R. Shaneyfelt, W. L. Warren, J. R. Schwank, T. L. Meisenheimer, and P. S. Winokur. Border traps: Issues for MOS radiation response and long-term reliability. *Microelectronics Reliability*, 35(3):403–428, March 1995.

[121] R. Castagné and A. Vapaille. Description of the SiO2Si interface properties by means of very low frequency MOS capacitance measurements. *Surface Science*, 28(1):157–193, November 1971.

[122] W. Shockley and W. T. Read. Statistics of the Recombinations of Holes and Electrons. *Physical Review*, 87(5):835–842, September 1952.

[123] K. Lehovec. Frequency dependence of the impedance of distributed surface states in mos structures. *Applied Physics Letters*, 8(2):48–50, January 1966.

[124] E. H. Nicollian, A. Goetzberger, and A. D. Lopez. Expedient method of obtaining interface state properties from MIS conductance measurements. *Solid-State Electronics*, 12(12):937–944, December 1969.

[125] C. N. Berglund. Surface states at steam-grown silicon-silicon dioxide interfaces. *IEEE Transactions on Electron Devices*, ED-13(10):701–705, October 1966.

[126] N. J. Chou. Application of Triangular Voltage Sweep Method to Mobile Charge Studies in MOS Structures. *Journal of The Electrochemical Society*, 118(4):601, April 1971.

[127] David Cooper, Thibaud Denneulin, Nicolas Bernier, Armand Béché, and Jean-Luc Rouvière. Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope. *Micron*, 80:145–165, January 2016.

[128] William Osborn, Lawrence H. Friedman, and Mark Vaudin. Strain measurement of 3D structured nanodevices by EBSD. *Ultramicroscopy*, 184:88–93, January 2018.

[129] M. D. Vaudin, W. A. Osborn, L. H. Friedman, J. M. Gorham, V. Vartanian, and R. F. Cook. Designing a standard for strain mapping: HR-EBSD analysis of SiGe thin film structures on Si. *Ultramicroscopy*, 148:94–104, January 2015.

[130] S. Di Fonzo, W. Jark, S. Lagomarsino, C. Giannini, L. De Caro, A. Cedola, and M. Müller. Non-destructive determination of local strain with 100-Nanometre spatial resolution. *Nature*, 403(6770):638–640, February 2000.

[131] Peter Hermann, Michael Hecker, Dmytro Chumakov, Martin Weisheit, Jochen Rinderknecht, Artem Shelaev, Pavel Dorozhkin, and Lukas M. Eng. Imaging and strain analysis of nano-scale SiGe structures by tip-enhanced Raman spectroscopy. *Ultramicroscopy*, 111(11):1630–1635, November 2011.

[132] Rongrui He and Peidong Yang. Giant Piezoresistance Effect in Silicon Nanowires. *Nature Nanotechnology*, 1(1):42–46, October 2006.

[133] Di Gao, Zhenyu Yang, Lingli Zheng, and Kun Zheng. Piezoresistive Effect of N-Type ⟨111⟩-Oriented Si Nanowires under Large Tension/Compression. *Nanotechnology*, 28(9):095702, 2017.

[134] J. Mansir, P. Conti, Z. Zeng, J. J. Pla, P. Bertet, M. W. Swift, C. G. Van de Walle, M. L. W. Thewalt, B. Sklenard, Y. M. Niquet, and J. J. L. Morton. Linear Hyperfine Tuning of Donor Spins in Silicon Using Hydrostatic Strain. *Physical Review Letters*, 120(16):167701, April 2018.

[135] Ryan M. Stein and M. D. Stewart. The effect of strain on tunnel barrier height in silicon quantum devices. *Journal of Applied Physics*, 128(2):024303, July 2020.

[136] 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. A two-qubit logic gate in silicon. *Nature*, 526(7573):410–414, October 2015.

[137] W. I. L. Lawrie, H. G. J. Eenink, N. W. Hendrickx, J. M. Boter, L. Petit, S. V. Amitonov, M. Lodari, B. Paquelet Wuetz, C. Volk, S. G. J. Philips, G. Droulers, N. Kalhor, F. van Riggelen, D. Brousse, A. Sammak, L. M. K. Vandersypen, G. Scappucci, and M. Veldhorst. Quantum dot arrays in silicon and germanium. *Applied Physics Letters*, 116(8):080501, February 2020.

[138] Matthias Brauns, Sergey V. Amitonov, Paul-Christiaan Spruijtenburg, and Floris A. Zwanenburg. Palladium gates for reproducible quantum dots in silicon. *Scientific Reports*, 8(1):1–8, April 2018.

[139] W. F. Brinkman, R. C. Dynes, and J. M. Rowell. Tunneling Conductance of Asymmetrical Barriers. *Journal of Applied Physics*, 41(5):1915–1921, April 1970.

[140] Amir Shirkhorshidian, John King Gamble, Leon Maurer, Stephen M. Carr, Jason Dominguez, Gregory A. Ten Eyck, Joel R. Wendt, Erik Nielsen, Noah Tobias Jacobson, Michael P. Lilly, and Malcolm S. Carroll. Spectroscopy of Multielectrode Tunnel Barriers. *Physical Review Applied*, 10(4):044003, October 2018.

[141] Bo Zhang, Xiaobo Li, and Dan Li. Assessment of thermal expansion coefficient for pure metals. *Calphad*, 43:7–17, December 2013.

[142] Weileun Fang and Chun-Yen Lo. On the thermal expansion coefficients of thin films. *Sensors and Actuators A: Physical*, 84(3):310–314, September 2000.

[143] F. C. Nix and D. MacNair. The Thermal Expansion of Pure Metals: Copper, Gold, Aluminum, Nickel, and Iron. *Phys. Rev.*, 60(8):597–605, October 1941.

[144] Y. Kuru, M. Wohlschlögel, U. Welzel, and E. J. Mittemeijer. Coefficients of thermal expansion of thin metal films investigated by non-ambient X-ray diffraction stress analysis. *Surface and Coatings Technology*, 202(11):2306–2309, February 2008.

[145] M. M. de Lima, R. G. Lacerda, J. Vilcarromero, and F. C. Marques. Coefficient of thermal expansion and elastic modulus of thin films. *Journal of Applied Physics*, 86(9):4936–4942, November 1999.

[146] Hans J. Fecht. Intrinsic instability and entropy stabilization of grain boundaries. *Phys. Rev. Lett.*, 65(5):610–613, July 1990.

[147] Michael Wagner. Structure and thermodynamic properties of nanocrystalline metals. *Phys. Rev. B*, 45(2):635–639, January 1992.

[148] R. Zhao, T. Tanttu, K. Y. Tan, B. Hensen, K. W. Chan, J. C. C. Hwang, R. C. C. Leon, C. H. Yang, W. Gilbert, F. E. Hudson, K. M. Itoh, A. A. Kiselev, T. D. Ladd, A. Morello, A. Laucht, and A. S. Dzurak. Single-spin qubits in isotopically enriched silicon at low magnetic field. *Nature Communications*, 10(1):1–9, December 2019.

[149] M. Rudolph, B. Sarabi, R. Murray, M. S. Carroll, and Neil M. Zimmerman. Long-term drift of Si-MOS quantum dots with intentional donor implants. *Scientific Reports*, 9(1):7656, May 2019.

[150] J.-S. Kim, T. M. Hazard, A. A. Houck, and S. A. Lyon. A low-disorder metal-oxide-silicon double quantum dot. *Applied Physics Letters*, 114(4):043501, January 2019.

[151] Eduardo B. Ramirez, Francois Sfigakis, Sukanya Kudva, and Jonathan Baugh. Few-electrode design for silicon MOS quantum dots. *Semiconductor Science and Technology*, 35(1):015002, November 2019.

[152] E. H. Nicollian, A. Goetzberger, and A. D. Lopez. Expedient method of obtaining interface state properties from MIS conductance measurements. *Solid-State Electronics*, 12(12):937–944, December 1969.

[153] Herbert B. Michaelson. The work function of the elements and its periodicity. *Journal of Applied Physics*, 48(11):4729–4733, November 1977.

[154] S. Mayo, K. F. Galloway, and T. F. Leedy. Radiation Dose Due to Electron-Gun Metallization Systems. *IEEE Transactions on Nuclear Science*, 23(6):1875–1880, December 1976.

[155] Reda R. Razouk and Bruce E. Deal. Dependence of Interface State Density on Silicon Thermal Oxidation Process Variables. *Journal of The Electrochemical Society*, 126(9):1573, September 1979.

[156] Ryan M. Stein and M. D. Stewart. The effect of strain on tunnel barrier height in silicon quantum devices. *Journal of Applied Physics*, 128(2):024303, July 2020.

[157] P. Feltham. Grain growth in metals. *Acta Metallurgica*, 5(2):97–105, February 1957.

[158] E. P. Nordberg, G. A. Ten Eyck, H. L. Stalford, R. P. Muller, R. W. Young, K. Eng, L. A. Tracy, K. D. Childs, J. R. Wendt, R. K. Grubbs, J. Stevens, M. P. Lilly, M. A. Eriksson, and M. S. Carroll. Enhancement-Mode Double-Top-Gated Metal-Oxide-Semiconductor Nanostructures with Tunable Lateral Geometry. *Physical Review B*, 80(11):115331, September 2009.

[159] Elliot J. Connors, JJ Nelson, Haifeng Qiao, Lisa F. Edge, and John M. Nichol. Low-frequency charge noise in Si/SiGe quantum dots. *Physical Review B*, 100(16):165305, October 2019.

[160] Neil M. Zimmerman, Chih-Hwan Yang, Nai Shyan Lai, Wee Han Lim, and Andrew S. Dzurak. Charge offset stability in Si single electron devices with Al gates. *Nanotechnology*, 25(40):405201, September 2014.

[161] Fabio Ansaloni, Anasua Chatterjee, Heorhii Bohuslavskyi, Benoit Bertrand, Louis Hutin, Maud Vinet, and Ferdinand Kuemmeth. Single-electron operations in a foundry-fabricated array of quantum dots. *Nature Communications*, 11(1):6399, December 2020.

[162] Simon Geyer, Leon C. Camenzind, Lukas Czornomaz, Veeresh Deshpande, Andreas Fuhrer, Richard J. Warburton, Dominik M. Zumbühl, and Andreas V. Kuhlmann. Self-aligned gates for scalable silicon quantum computing. *Applied Physics Letters*, 118(10):104004, March 2021.

[163] Michael A. Guillorn, Dustin W. Carr, Richard C. Tiberio, Elias Greenbaum, and Michael L. Simpson. Fabrication of dissimilar metal electrodes with nanometer interelectrode distance for molecular electronic device characterization. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, 18(3):1177–1181, May 2000.

[164] Tak H. Ning. Electron trapping in SiO2 due to electron-beam deposition of aluminum. *Journal of Applied Physics*, 49(7):4077–4082, July 1978.

[165] Masakazu Shimaya, Noboru Shiono, Osaake Nakajima, Chisato Hashimoto, and Yutaka Sakakibara. Electron Beam Induced Damage in Poly-Si Gate MOS Structures and Its Effect on Long-Term Stability. *Journal of The Electrochemical Society*, 130(4):945, April 1983.

[166] I. M. Hafez, G. Ghibaudo, and F. Balestra. Assessment of interface state density in silicon metal-oxide-semiconductor transistors at room, liquid-nitrogen, and liquid-helium temperatures. *Journal of Applied Physics*, 67(4):1950–1952, February 1990.

[167] J.-S. Kim, A. M. Tyryshkin, and S. A. Lyon. Annealing shallow Si/SiO2 interface traps in electron-beam irradiated high-mobility metal-oxide-silicon transistors. *Applied Physics Letters*, 110(12):123505, March 2017.

[168] Scott W. Schmucker, Pradeep N. Namboodiri, Ranjit Kashid, Xiqiao Wang, Binhui Hu, Jonathan E. Wyrick, Alline F. Myers, Joshua D. Schumacher, Richard M. Silver, and M. D. Stewart. Low-Resistance, High-Yield Electrical Contacts to Atom Scale Si:P Devices Using Palladium Silicide. *Physical Review Applied*, 11(3):034071, March 2019.

[169] C. Byrne, B. Brennan, R. Lundy, J. Bogan, A. Brady, Y. Y. Gomeniuk, S. Monaghan, P. K. Hurley, and G. Hughes. Physical, chemical and electrical characterisation of the diffusion of copper in silicon dioxide and prevention via a CuAl alloy barrier layer system. *Materials Science in Semiconductor Processing*, 63:227–236, June 2017.

[170] J. D. McBrayer, R. M. Swanson, and T. W. Sigmon. Diffusion of Metals in Silicon Dioxide. *Journal of The Electrochemical Society*, 133(6):1242, June 1986.

[171] David Cooper, Thibaud Denneulin, Nicolas Bernier, Armand Béché, and Jean-Luc Rouvière. Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope. *Micron*, 80:145–165, January 2016.

[172] L. Clément, R. Pantel, L. F. Tz. Kwakman, and J. L. Rouvière. Strain measurements by convergent-beam electron diffraction: The importance of stress

relaxation in lamella preparations. *Applied Physics Letters*, 85(4):651–653, July 2004.

[173] P. J. McNally, J. Kanatharana, B. H. W. Toh, D. W. McNeill, A. N. Danilewsky, T. Tuomi, L. Knuuttila, J. Riikonen, J. Toivonen, and R. Simon. Geometric linewidth and the impact of thermal processing on the stress regimes induced by electroless copper metallization for Si integrated circuit interconnect technology. *Journal of Applied Physics*, 96(12):7596–7602, December 2004.

[174] Attila Bonyár. AFM characterization of the shape of surface structures with localization factor. *Micron*, 87:1–9, August 2016.

[175] A. J. Huber, D. Kazantsev, F. Keilmann, J. Wittborn, and R. Hillenbrand. Simultaneous IR Material Recognition and Conductivity Mapping by Nanoscale Near-Field Microscopy. *Advanced Materials*, 19(17):2209–2212, 2007.

[176] Bosai Lyu, Hongyuan Li, Lili Jiang, Wanfei Shan, Cheng Hu, Aolin Deng, Zhe Ying, Lele Wang, Yiran Zhang, Hans A. Bechtel, Michael C. Martin, Takashi Taniguchi, Kenji Watanabe, Weidong Luo, Feng Wang, and Zhiwen Shi. Phonon Polariton-assisted Infrared Nanoimaging of Local Strain in Hexagonal Boron Nitride. *Nano Letters*, 19(3):1982–1989, March 2019.

[177] Qi Zhang, Guangchong Hu, Gabriele G. de Boo, Miloš Rančić, Brett C. Johnson, Jeffrey C. McCallum, Jiangfeng Du, Matthew J. Sellars, Chunming Yin, and Sven Rogge. Single Rare-Earth Ions as Atomic-Scale Probes in Ultrascaled Transistors. *Nano Letters*, 19(8):5025–5030, August 2019.

[178] A. J. Huber, A. Ziegler, T. Köck, and R. Hillenbrand. Infrared nanoscopy of strained semiconductors. *Nature Nanotechnology*, 4(3):153–157, March 2009.

[179] P. D. Ye, D. Weiss, R. R. Gerhardts, K. von Klitzing, K. Eberl, H. Nickel, and C. T. Foxon. Strain-induced magnetoresistance oscillations in GaAs-AlGaAs heterojunctions with ferromagnetic and superconducting submicrometre gratings. *Semiconductor Science and Technology*, 10(5):715–719, May 1995.

[180] A. N. Ramanayaka, Ke Tang, J. A. Hagmann, Hyun-Soo Kim, D. S. Simons, C. A. Richter, and J. M. Pomeroy. Use of quantum effects as potential qualifying metrics for "quantum grade silicon". *AIP Advances*, 9(12):125153, December 2019.

[181] Tsuneya Ando, Alan B. Fowler, and Frank Stern. Electronic properties of two-dimensional systems. *Reviews of Modern Physics*, 54(2):437–672, April 1982.

[182] J.-S. Kim, A. M. Tyryshkin, and S. A. Lyon. Annealing shallow Si/SiO2 interface traps in electron-beam irradiated high-mobility metal-oxide-silicon transistors. *Applied Physics Letters*, 110(12):123505, March 2017.

[183] R. M. Jock, S. Shankar, A. M. Tyryshkin, Jianhua He, K. Eng, K. D. Childs, L. A. Tracy, M. P. Lilly, M. S. Carroll, and S. A. Lyon. Probing band-tail states in silicon metal-oxide-semiconductor heterostructures with electron spin resonance. *Applied Physics Letters*, 100(2):023503, January 2012.

[184] X. Gao, E. Nielsen, R. P. Muller, R. W. Young, A. G. Salinger, N. C. Bishop, and M. S. Carroll. The QCAD framework for quantum device modeling. In *2012 15th International Workshop on Computational Electronics*, pages 1–4, May 2012.

[185] COMSOL® Software Version 5.6 Release Highlights. https://www.comsol.com/release/5.6.

[186] Ted Thorbeck, Akira Fujiwara, and Neil M. Zimmerman. Simulating Capacitances to Silicon Quantum Dots: Breakdown of the Parallel Plate Capacitor Model. *IEEE Transactions on Nanotechnology*, 11(5):975–978, September 2012.

[187] Jan Klos, Fabian Hassler, Pascal Cerfontaine, Hendrik Bluhm, and Lars R. Schreiber. Calculation of tunnel couplings in open gate-defined disordered quantum dot systems. *Physical Review B*, 98(15):155320, October 2018.