Charge Detection in Semiconductor Nanostructures

by

Kenneth MacLean

B.S. in Physics and Mathematics, Yale University, 2002

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2010

Author . . . .
Department of Physics
November 4, 2009

Certified by. . . . .
Marc A. Kastner
Donner Professor of Science and Dean of the School of Science
Thesis Supervisor

Accepted by . . . .
Krishna Rajagopal
Professor of Physics and Associate Department Head for Education

Charge Detection in Semiconductor Nanostructures

by

Kenneth MacLean

Abstract

In this thesis nanometer scale charge sensors are used to study charge transport in two solid state systems: Lateral GaAs quantum dots and hydrogenated amorphous silicon (a-Si:H). In both of these experiments we use time-resolved charge sensing to study electron transport in regimes that are not accessible to traditional transport measurements.

For the lateral GaAs quantum dot experiments, we use a GaAs quantum point contact integrated with the dot as a charge sensor. We use this sensor to observe single electrons hopping on and off the dot in real time. By measuring the time intervals for which the dot contains one and zero electrons, we probe the rate Γ at which electrons tunnel on and off the dot from the leads. We measure Γ as a function of the drain source bias V_{ds} and gate voltages V_g applied to the dot. At zero magnetic field, we show that the dependencies of Γ on V_{ds} and V_g can be understood in terms of a simple quantum mechanical model which takes into account variations in the electron energy relative to the top of the tunnel barriers separating the dot from the leads. We also show that the tunneling is dominated by elastic processes. At high magnetic fields, we show that tunneling into the excited spin state of the dot can be completely suppressed relative to tunneling into the ground spin state. The extent of the suppression depends on the shape of the electrostatic potential defining the quantum dot.

For the a-Si:H experiments, we pattern a nanometer scale strip of a-Si:H adjacent to a narrow silicon MOSFET (metal-oxide-semiconductor field-effect transistor), which serves as an integrated charge sensor. We show that the MOSFET can be used to detect charging of the a-Si:H strip. By performing time-resolved measurements of this charging, we are able to measure extremely high resistances ($\sim 10^{17}$ Ω) for the a-Si:H strip at T ≈ 100 K. At higher temperatures, where the resistance of the a-Si:H strip is not too large, we show that the resistances obtained from our charge detection method agree with those obtained by measuring current. Our device geometry allows us to probe a variety of electron transport phenomena for the a-Si:H, including the field effect and dispersive transport, using charge detection. We extract the density of localized states at the Fermi level for the a-Si:H and obtain consistent results. We

discuss the effect of screening by the substrate on the sensitivity of the MOSFET to charge in the a-Si:H, and show that the MOSFET can detect switching noise in the a-Si:H.

Thesis Supervisor: Marc A. Kastner
Title: Donner Professor of Science and Dean of the School of Science

Acknowledgments

When I was about to start graduate school, I pictured spending long hours working alone in a laboratory in some basement, and worried I would feel extremely isolated. My experience in graduate school was nothing like this, and looking back, what I remember most about my time at MIT are the people I was surrounded by, to many of whom I owe a great deal of gratitude.

I would first like to thank my research adviser, Marc Kastner. Before I started at MIT, I do not think I realized how crucial my choice of research adviser would be to my experience in graduate school. Looking back now, my experience working with Marc was so positive that it is hard for me to know how to begin thanking him. Marc is an excellent leader, and a truly inspiring person to work with. I will be forever indebted to him for teaching me how to think about physics and providing an ideal environment in which to do creative research.

All of the work I did at MIT was done in collaboration with a number of other scientists. When I first got to MIT I started working on experiments involving lateral GaAs quantum dots with Sami Amasha, who was a few years into graduate school at that time. I can't say enough about Sami. He is an extremely talented physicist and a great guy, both of which facts are obvious to anyone who has worked with him for an hour or two. Early on, when I did not have much experience working in a laboratory, Sami took a huge amount of time to show me how to design and execute experiments. As time went on our collaboration became very productive. We worked on a number of interesting experiments, all of which drew heavily on ideas we had been bouncing back and fourth, usually while transferring Helium or performing similar routine laboratory tasks. I also worked with two other members of the Kastner group on the GaAs quantum dot portion of my thesis, Dominik Zumbühl and Iuliana Radu, who fabricated the samples we used for our experiments. Dominik was a post doc in our group at this time, and was a great source of guidance. Iuliana is a great physicist and a really fun person to work with. Throughout graduate school, I found myself continually asking for her advice, as she is someone who can provide a

good perspective on virtually anything. I also want to thank Micah Hanson ant Art Gossard for growing the excellent heterostructures we used for our devices, without which none of our GaAs quantum dot experiments would have been possible.

As my work on GaAs quantum dots began to come to a close, I started working with Tamar Mentzel, and Moungi Bawendi, Scott Geyer, and Venda Porter, our collaborators in the chemistry department, on arrays of semiconducting nanocrystals, and I really enjoyed participating in this facet of our group's research. Following this work, Tamar and I collaborated closely, working on silicon charge sensors. This collaboration was extremely rewarding. Tamar is an outstanding physicist, and working together we were able to learn silicon fabrication and a number of other techniques that were new to both of us in a short amount of time. This was facilitated in large part by Tamar's creativity, which kept our project moving along quickly despite the number of obstacles we ran into along the way. Tamar is also a great friend. I also want to thank the staff of the MTL for their help with our MOSFET fabrication.

There are a number of other members of the Kastner group who helped me during the course of my Ph.D. Andrei Kogan and Ghislain Granger were members of the group during the beginning of my Ph.D. and they both taught me a lot about research in our laboratory. Ian Gelfand started as a graduate student around the same time I did. We shared an office together and I enjoyed discussing science and other things with him over the years. Colin Dillard started a few years after me and has been a great person to work with. Within the last two years, our group has been fortunate to get two excellent post-docs, Jingshi Hu and Xi Lin. I been lucky to work with Jingshi a bit on his silicon germanium work. His work is very impressive, and he has moved a very difficult project ahead a lot in a short time frame. I have enjoyed getting to know Xi over the past year, he is a great person to ask for advice, and I am looking forward to seeing whether we can combine some of the techniques Tamar and I developed with the systems he and Colin are investigating. Our group has also been fortunate to have two new graduate students, Nirat Ray and Andrew Lai, who I am looking forward to getting to know. I also want to thank the undergraduate students who worked with Sami and I on the GaAs quantum dot project: Anat Burger, Jessica

6

Schirmer, and Vladimir Rosenhaus.

There are a number of scientists outside of our group who helped me a lot with my research during my time at MIT. I frequently consulted Ray Ashoori and Leonid Levitov, my thesis committee members, about my research and received excellent insight and advice. I also want to thank Juan Climente, Andrea Bertoni, Guido Goldoni, Vitaly Golovach, Daniel Loss, Emmanuel Rashba, Jeff Miller, Oliver Dial, Jae-kyu Lee, Bertrand Halperin, Mark Rudner, Helmut Fritzsche, and Jim Kakalios, for fruitful discussions about the work presented in this thesis.

I want to thank all of my friends and family. I can't thank all of my friends sufficiently in writing, but they know who they are and how I feel about them. I want to thank my fiance Abby, who is the best thing to happen to me in graduate school, or in my life for that matter, and I want to also thank Abby's family. I want to thank my parents, Bec, Lizzie and Brendan, Kevin, all of my nieces and nephews, Atara and Megan and their families. Lastly I want to thank my sister Mary. I miss her very much.

Contents

List of Figures

Chapter 1

Introduction

One of the most important properties of semiconducting materials is their sensitivity to relatively small changes in charge density. For a pure semiconductor crystal, the Fermi level lies in the middle of the band gap, so that the charge density in the conduction and valence bands is very small, and the conductivity of the material is very low. What distinguishes such a material from an insulator is that for a semiconductor, a modest change in the electrostatic environment of the crystal, in the form of an applied electric field, a small number of impurity atoms introduced into the crystal, or exposure to electromagnetic radiation, creates a change in the charge density in the conduction or valence band that has a dramatic effect on the conductivity. This sensitivity leads directly to the wide variety of semiconductor transistors, diodes, and photovoltaic devices that, since the invention of the solid state transistor [1], have come to dominate modern electronics [2].

As semiconductor technology has matured and miniaturized, techniques for sculpting materials on smaller and smaller length scales have developed. Empowered by this technology, physicists have explored the behavior of electrons confined to nanometer scale semiconductor devices. Electrons confined to small solid state structures, referred to as quantum dots, behave much like electrons occupying the orbitals of single atoms [3]. A variety of experiments have explored aspects of this atomic behavior [4, 5, 6, 7, 8]. Electrons occupying quantum dots can be coupled through tunnel barriers to a Fermi sea, and this has lead to the observation of other interesting physical

phenomena, including the single electron transistor functionality of quantum dots [9], and many-body electron interactions involving electrons on the quantum dot and in the leads [10, 11]. In addition to quantum dots, where the electronic wavefunction is confined in all three dimensions, structures that confine electrons in only two dimensions have also been studied extensively, revealing a number of interesting effects to solid state physicists [12, 13].

The conductivity through all nanoscale semiconductor devices is enormously sensitive to the electrostatic environment of the device. This fact was reported almost as soon as physicists began fabricating semiconducting devices of a sufficiently small size [14] and is fairly obvious to any experimentalist who has observed the "switching" behavior evident to some extent in all of these devices [15, 16]. In fact, the electric field generated by a single electron in the immediate vicinity of a nanoscale semiconductor device creates a measurable change in the device conductivity. For example, for a nanoscale MOSFET (metal-oxide-semiconductor field-effect transistor), the conductivity is sensitive to the occupation of single defects in the oxide near the oxide-semiconductor interface. This sensitivity can be used to study electrons confined to these defects, and has been utilized to observe a wide range of interesting effects, from many body electron-electron interactions [17] to the spin resonance of a single electron spin [18].

As nanometer semiconductor devices were further developed, it was demonstrated that they could function as ultra-sensitive electrometers, capable of detecting not only single electron fluctuations within the device itself, but also in structures adjacent to the device [19]. Nanometer scale semiconductor devices functioning as electrometers, patterned adjacent to a structure of interest, are called integrated charge sensors. In the work reported in this thesis, we use integrated charge sensors to study two very different systems: Lateral GaAs quantum dots and hydrogenated amorphous silicon (a-Si:H). Specifically, in Chapters 2 and 3, we report measurements of single electron tunneling in GaAs quantum dots obtained using a GaAs quantum point contact as an integrated charge sensor, and in Chapters 4 and 5 we report measurements of electron transport in a-Si:H obtained using a nanometer scale MOSFET as an

integrated charge sensor.

Lateral GaAs quantum dots are one of the most well understood nanometer scale semiconducting devices [20]. Integrated charge sensors were first demonstrated for this system [19], in the form of a nanometer scale conducting channel, or quantum point contact (QPC), positioned adjacent to the dot. Recently, some time after the initial report of the integrated charge sensor, a number of research efforts have demonstrated that integrated charge sensors can be used to study aspects of the quantum mechanical behavior of electrons confined to GaAs quantum dots that are impossible to observe by directly measuring the current flowing through the dot [21, 22, 23, 24, 25]. The work reported in Chapters 2 and 3 constitute part of this initiative. Specifically, we use integrated charge sensors to reveal fundamental aspects of the way in which electrons tunnel on and off a GaAs quantum dot.

In Chapter 2, we discuss the techniques we have developed for using integrated charge sensors to study GaAs quantum dots, and use them to study the tunneling process at zero magnetic field. We demonstrate that the tunneling is dominated by elastic processes, and that variations in the tunnel rate as the drain-source-bias and gate voltages applied to the quantum dot are changed can be described by a simple quantum mechanical model that takes into account variations in the electron energy relative to the tunnel barrier potential.

In Chapter 3, we examine the tunneling process in the presence of a large magnetic field. Surprisingly, we find that as the magnetic field is increased, tunneling into the excited spin state can be completely suppressed relative to tunneling into the ground spin state. The extent of this suppression varies as the shape of the quantum dot is changed. While the phenomenology of this effect is clear, these results have not as of yet been explained theoretically. The experiments presented in Chapters 2 and 3, enabled by the integrated charge sensing technique, contribute to a full understanding of the physics of tunneling in quantum dots, a process that is central to all applications of these devices. Having seen the substantial impact of the integrated charge sensor on transport studies of quantum dots, we sought to use these sensors to study electron transport in other solid state systems.

One aspect of the integrated charge sensing technique that stands out in contrast to the measurement of current is its ability to measure extremely slow electron dynamics. Current, a measure of charge per unit time, becomes immeasurably small for systems where electrons move slowly. Replacing the measurement of current with a time-resolved measurement of charge, one obtains a probe of electron transport that is ideally suited for the study of slow electron dynamical processes. It is in part this aspect of the integrated charge sensing technique that has resulted in its large impact on the study of GaAs quantum dots, as it allows the experimenter to study transport in dots that are almost completely isolated from the Fermi sea in their leads.

In Chapters 4 and 5 we exploit the ability of integrated charge sensors to measure slow electron dynamics, and use this technique to probe electron transport in highly resistive materials for which traditional transport techniques fail. As a resistive material, we use a nanometer scale strip of a-Si:H, cooled to a temperature where the resistance is very high. Having demonstrated the application of integrated charge sensing to the study of a-Si:H, a material for which the electronic structure has been studied extensively [26], we anticipate that our work will be extended to characterize the electronic properties of other technologically significant materials for which a high electrical resistance inhibits the measurement of current.

Chapter 4 is the most technical chapter of this thesis. It covers the fabrication, instrumentation, and electrical characteristics of the narrow MOSFET sensor and a-Si:H strip. Since over the past five years many studies of electron transport in GaAs quantum dots utilizing a QPC as an integrated charge sensor have been reported, and technical aspects of such experiments are now well documented [27, 28, 29]. However, our work with a-Si:H is the first utilization of integrated charge sensors for the study of resistive materials, so a presentation of the technical aspects of these experiments is well warranted.

In Chapter 5, we present our measurements of transport in a-Si:H, obtained using a narrow MOSFET as an integrated charge sensor. We are able to measure incredibly high resistances ($\sim 10^{17}$ Ω) for the a-Si:H strip, using only moderate voltages (~ 1 volt). Our technique allows us to probe a variety of transport phenomena, including

the temperature, electric field, and time dependence of the transport, using charge sensing. We extract the density of states at the Fermi level for the a-Si:H and obtain consistent results. Our measurements are also consistent with the standard model for electron transport in a-Si:H [26]. The work presented in this chapter constitutes a fundamentally new way of measuring electrical resistance that is ideally suited to the study of electron transport in highly resistive thin film materials.

In Chapter 6, we discuss the broader implications of the work presented in this thesis, and present ideas for future experiments.

Chapter 2

Energy Dependent Tunneling at Zero Magnetic Field

In this chapter we discuss time-resolved measurements of single electron tunneling in a GaAs quantum dot at zero magnetic field. In Section 2-1, we give an introduction to lateral GaAs quantum dots. In Section 2.2 we describe how we measure the charge on the dot using a quantum point contact (QPC) charge sensor, and show how we use time-resolved charge detection to measure the tunneling rate Γ between the dot and its leads. In Section 2.3, we measure the dependence of Γ on drain-source bias V_{ds}, and show that this dependence can be explained in terms of elastic tunneling at a rate set by the difference between the electron energy and the height of the tunnel barrier, a process we will refer to as energy dependent tunneling. In Section 2.4, we measure the dependence of Γ on plunger gate voltage V_g, and show that this dependence can also be understood in terms of energy dependent tunneling. Parts of the results presented here appear also in K. MacLean *et al.* [30].

2.1 Introduction: GaAs Quantum Dots

A quantum dot is a nanometer scale region of semiconductor to which a small number of electrons are localized, and for which both the Coulomb energy E_C required to add an additional electron, and the energy splitting between different quantum orbital

states E_Q play an important role. Because $E_C = e^2/C$, where C is the capacitance of the dot, and $E_Q \sim \hbar^2/2m^*L^2$, where m^* is the electron effective mass and L is the size of the dot, E_Q and E_C become larger the smaller L and C, which are both reduced by making the dot smaller. Quantum dots are therefore constructed using nanofabrication techniques.

The variety of ways in which quantum dots can be fabricated can roughly be divided into two categories. For the first category, small chunks of semiconductor of various sizes and shapes are grown using chemical processes. For the second category, the quantum dot is patterned using nanolithographic techniques. These two categories can overlap. For instance, a number of groups have studied surface gated carbon nanotubes [31, 32]: For these devices, the carbon nanotube is grown by chemical vapor deposition, and the gates are then patterned around the nanotube using electron beam lithography.

Very small nanostructures can be made with chemical processes, and the energy scales E_C and E_Q for quantum dots made with these processes can be quite large. The effect of quantum confinement can be observed at room temperature and can play an important role in how these dots interact with electromagnetic radiation in the optical and infrared frequency ranges [33, 34]. Quantum dots fabricated with nanolithography are generally larger, so that E_C and E_Q are smaller, and the effects of quantum confinement and Coulomb repulsion are observable only at cryogenic temperatures. Transitions between the quantum energy levels of these dots are usually driven with lower frequency radiation, in the microwave range [35, 36].

For quantum dots fabricated with nanolithography, surface gates patterned adjacent to the dot can be used to tune the properties of the quantum dot. This tunablity has enabled a wide variety of experiments. For instance, for lateral GaAs quantum dots, one can control in-situ the rate of tunneling Γ coupling the dot to its leads: Γ can be adjusted over more than ten orders of magnitude by changing the voltages applied to the gates that define the tunnel barriers of the quantum dot. Because of this tunability, this type of quantum dot has been used to study correlated-electron physics [10, 11], which becomes important when the coupling of the dot to its leads

is strong, and also the coherence of electron charge and spin states [7, 22, 37], which can be maintained only when the coupling to the leads is weak.

The experiments discussed in this chapter are performed with a lateral GaAs quantum dot (Fig. 2-1) fabricated with electron beam lithography. To make such a device, we start with a heterostructure grown by molecular beam epitaxy consisting of GaAs topped with a thin (110 nm) layer of AlGaAs (more precisely, $Al_{0.3}Ga_{0.7}As$), as depicted in Fig. 2-1(a). Within the AlGaAs, a distance $d_d = 60$ nm from the GaAs/AlGaAs interface, there are a small number of silicon atoms (less than one atomic layer). The silicon acts as an electron donor in GaAs. Because the conduction band of AlGaAs lies above the conduction band of GaAs by an amount $eV_{bo} \approx 240$ meV, a number of electrons donated by the silicon atoms will move into the GaAs. The positively charged ionized silicon donors left behind in the AlGaAs create a strong electric field that pulls these electrons to the GaAs/AlGaAs interface. These electrons, confined to within ~ 5 nm of the GaAs/AlGaAs interface, form a two dimensional electron gas (2DEG). The number of electrons per unit area n_e in the 2DEG is given by the amount of charge required to charge the capacitance C_d between the silicon donor layer and the GaAs/AlGaAs interface up to a voltage equal and opposite to the potential difference between the AlGaAs and GaAs conduction bands:

$$en_e = C_dV_{bo} \tag{2.1}$$

Here $C_d = \epsilon_0\kappa_A/d_d$, where here κ_A is the dielectric constant of AlGaAs. For the heterostructure used in our experiments, $n_e = 2.2 \times 10^{11}$ cm^{-2}.[1] The conductance per square of the 2DEG is given by:

$$\sigma = en_e\mu \tag{2.2}$$

Here μ is the mobility of the 2DEG, which for our sample is equal to 6.4×10^5 cm^2/Vs [38]. A high mobility reflects a small amount of scattering for electrons in the

---

[1]This value was obtained from Hall effect data by G. Granger [38]. The value calculated from Eqn. 2.1 differs by only 25 %.

2DEG, which leads to a high conductance. The GaAs/AlGaAs structures discussed here have mobilities that are quite high compared to 2DEGs in other semiconductor structures (Chapter 4). This is a result of the fact that AlGaAs and GaAs are lattice matched, so that there are very few dangling bonds, dislocations, or other defects near the AlGaAs/GaAs interface, and the fact that the silicon dopants are physically separated from the 2DEG.[2] There are therefore few potential fluctuations off of which to scatter. Because of their high mobility, these type of heterostructures are used in high electron mobility transistors (HEMT) and other devices designed to operate at high frequencies for which a high conductivity is desired [40, 41].

The heterostructure provides one dimension of confinement. To confine the electrons in the other two dimensions, we pattern nanometer scale metallic gates on the surface of the heterostructure.[3] A micrograph of the gate structure is shown in Fig. 2-1(b). We apply negative voltages to the top five gates (SG1, LP1, PL, LP2, and SG2) and to the gate OG. This depletes the electrons in the 2DEG underneath these gates. As the voltages are made more negative, the electrons in the regions between OG and SG1 and between OG and SG2 are depleted, forming two tunnel barriers. The region between the two tunnel barriers, indicated by the blue oval in Fig. 2-1(b), constitutes the quantum dot.

The potential as a function of position taken along the dashed green curve in Fig. 2-1(b) is shown in Fig. 2-1(c). There are two tunnel barriers where the potential rises above the Fermi level in the 2DEG. These barriers separate the potential well in between them from the two Fermi sea leads, forming the quantum dot confining potential. The potential at the bottom of the well can be shifted relative to the Fermi level in the leads by adjusting the voltages applied to the three gates LP1, PL, and LP2, so that the number of electrons on the dot can be controlled. We refer to these three gates as the plunger gates, and we apply approximately the same gate voltage V_g to each of them for all of the experiments discussed here. This is done so that there is only one potential minimum between the two tunnel barriers.

---

[2]This type of doping is referred to as modulation doping [39].
[3]The nanofabrication for our devices was performed by I. P. Radu and D. M. Zumbühl [29].
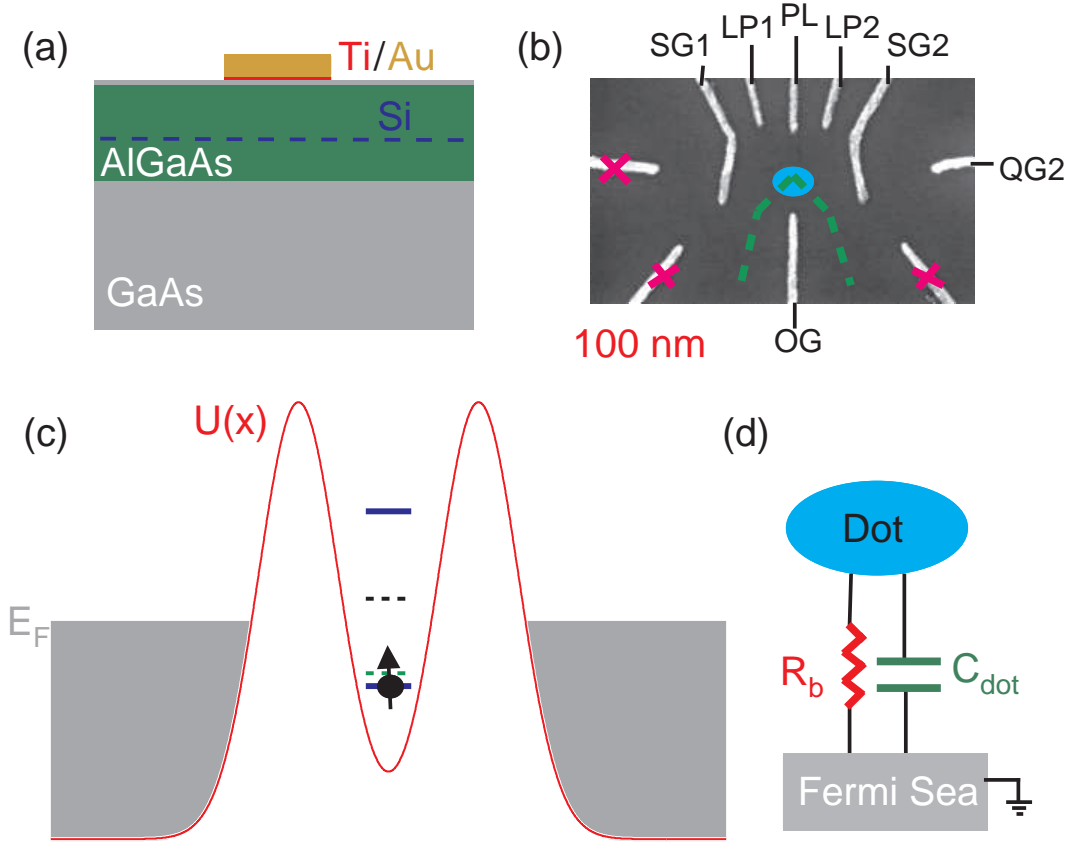
Figure 2-1: (a) Cross section of the GaAs/AlGaAs heterostructure and a nanopatterned metallic gate, as discussed in the main text. The top of the heterostructure has a thin 10 nm capping layer of GaAs. The gates consist of gold on top of a thin titanium layer, the latter of which adheres well to the GaAs cap [29]. (b) Electron micrograph of our lateral GaAs quantum dot gate structure. The gates with pink x's over them are kept grounded for all of the experiments discussed here and can be ignored. Negative voltages are applied the gates OG, SG1, LP1,PL, LP2, and SG2, forming a quantum dot at the position of the blue oval. A negative voltage is also applied to the gate QG2 to create a quantum point contact charge sensor as discussed in Section 2.2. (c) Sketch of electrostatic potential $U(x)$ in the 2DEG along the dashed green curve in (b). The Fermi seas in the two leads are separated from the quantum dot by two tunnel barriers. The one and two electron states of the quantum dot are indicated by the lower and upper solid blue lines. The dashed black line indicates a single electron excited orbital state. The dashed green line indicates the single electron excited spin state in the presence of a magnetic field. (d) Simple model for localization of electrons on a quantum dot, as discussed in the main text.

23

The relevant energy scales for the quantum dot system are sketched in Fig. 2-1(c). The largest energy scale is the Fermi level E_F relative to the bottom of the conduction band, which can be calculated from the electron density in the 2DEG, and is approximately 7 meV. Next largest is the Coulomb energy $E_C \approx 4$ meV required to add an additional electron to the quantum dot. This energy is given by the difference in energy between the one and two electron states, sketched as the solid blue lines in Fig. 2-1(c). For all of the experiments discussed here, the two electron state is above E_F, so that there is either one or zero electrons on the quantum dot.

The next largest energy scale is the quantum confinement E_Q, which is given by the difference between the energy of an excited orbital state (dashed black line in Fig. 2-1(c)) and the ground state. Of course, there are a number excited orbital states. For lateral quantum dots, the confinement created by the heterostructure along the direction perpendicular to the GaAs/AlGaAs interface is typically much stronger than the confinement created by the gates. For low energy excitations, one can therefore model the quantum dot confining potential as a two dimensional harmonic oscillator potential well. For our device, the excited state of this harmonic oscillator with the smallest energy is ~ 2 meV above the ground state, as we will see in the following sections.

For the experiments discussed in the following chapter, we apply a magnetic field B in the plane of the 2DEG. For these experiments, where typically B is a few Telsa, there is an excited spin state ~ 100 μeV above the ground spin state. This state is sketched as the dashed green line in Fig. 2-1(c). Finally the smallest energy scale is kT. This energy scale determines the broadening of the Fermi distribution function in the leads (not shown in Fig. 2-1(c)). All of the experiments discussed here are performed in an Oxford 75 μW dilution refrigerator, at an electron temperature T $\approx$ 120 mK so that $kT \approx 10$ μV is much smaller than any of the other relevant energy scales.

In the following sections, we use time resolved single electron charge sensing to study the tunneling process by which electrons move on and off the quantum dot. For these experiments, the tunnel barrier resistances R_b are very large. Before turning to

a discussion of these experiments, we discuss a fundamental question: How resistive do the tunnel barriers need to be in order for charge to be localized on the quantum dot in the first place? To provide an intuitive answer to this question, we consider a quantum dot connected to a grounded, Fermi Sea lead through one tunnel barrier of resistance R_b (Fig. 2-1(d)). The time it takes to charge the quantum dot is given by $\tau \sim R_b C$, where here C is the capacitance of the dot to ground. In order for localization to occur, the Coulomb charging energy of the dot, $E_C = e^2/C$, must be larger than the time-energy uncertainty $\delta E \sim h/\tau \sim h/R_b C$. From this we obtain the minimum barrier resistance necessary for localization to be $R_b \sim h/e^2$, the inverse of the quantum conductance.

2.2 Charge Detection Measurement

For the experiments discussed in the following sections, we utilize time-resolved single electron charge detection. Our measurement circuit is shown Fig. 2-2(a). Applying a negative voltage to the gate QG2, we form a quantum point contact (QPC) between the gates QG2 and SG2. The device characteristics of quantum point contacts will be discussed in Section 3.3. For now, we regard the QPC simply as a narrow conduction path, which, because of its nanoscale dimensions, is extremely sensitive to its electrostatic environment. In particular, the QPC is sensitive to the charge on the quantum dot: Adding an electron to the quantum dot has the same effect on the QPC resistance as making the voltage on either QG2 or SG2 slightly more negative. We can therefore use the resistance of the QPC as a measure of the charge on the quantum dot. To measure the resistance of the QPC, we source a current $I = 1$ nA through the QPC and measure the voltage V_{qpc} across the QPC.

An example of time-resolved charge sensing data is shown in Fig. 2-2(c). For these data, we apply a negative voltage bias V_{ds} to one of the two leads coupled to the quantum dot so that the one electron state is between the Fermi levels of the two leads, as sketched in Fig 2-2(b). We refer to the two Fermi seas as lead 1 and lead 2, as labeled in Fig. 2-2(a). Leads 1 and 2 are connected to the quantum dot through
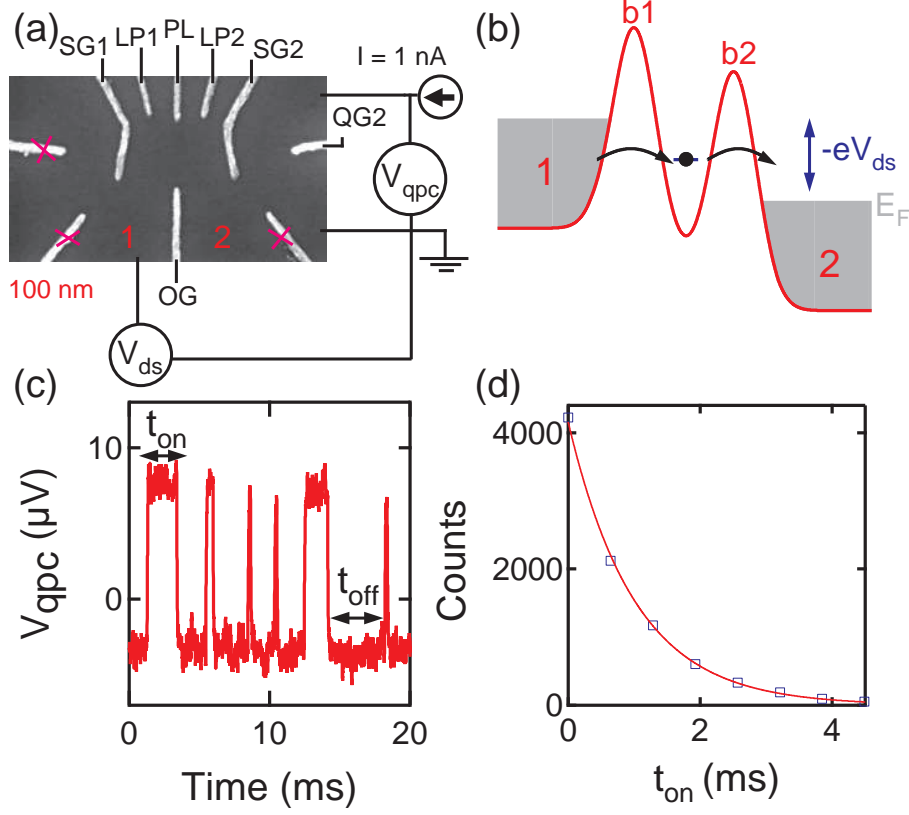
Figure 2-2: (a) Electron micrograph of the gate geometry and schematic of the measurement circuit. Gates with pink x's are grounded and can be ignored. We apply a negative voltage to the gates QG2 and SG2, forming a quantum point contact (QPC) between these two gates. The resistance R_{qpc} of the QPC is sensitive to the charge on the quantum dot, as discussed in the main text. We measure R_{qpc} by sourcing a current through the QPC and monitoring the voltage V_{qpc} across the QPC. (b) When a voltage bias V_{ds} is applied across the quantum dot a small current flows and the number of electrons on the dot fluctuates between 0 and 1. For the negative voltage bias shown here, electrons from lead 1 tunnel onto the dot through b1 and then off of the dot and into lead 2 through b2. (c) As the electrons hop on and off the dot, V_{qpc} jumps up and down. We measure the time intervals t_{on} (t_{off}) that the electron is on (off) the dot using the automated triggering system described in the main text and in Fig. 2-3. The offset in the trace is caused by the AC coupling of the voltage preamplifier. (d) Histogram of t_{on} times from data such as in (c). Fitting this histogram to an exponential yields Γ_{off} as described in the main text.

the two tunnel barriers, which we will henceforth refer to as b1 and b2, respectively. For all of the experiments discussed here, lead 2 is kept grounded, and V_{ds} is applied to lead 1, as shown in Fig. 2-2(a). For negative V_{ds}, electrons tunnel onto the dot from lead 1 and off of the dot into lead 2. As this happens, the number of electrons on the quantum dot fluctuates between 0 and 1, causing the QPC resistance and thus V_{qpc} to fluctuate between a low and a high value. We measure V_{qpc} as a function of time, and from these time series we measure the times that the electron is on (t_{on}) and off (t_{off}) the dot, as shown in Fig 2-2(c).

From the statistics of the time intervals t_{on} and t_{off}, we can measure the rates Γ_{off} and Γ_{on} at which electrons tunnel off and on the dot, respectively. To measure Γ_{off}, we histogram the times t_{on} that the electron spends on the dot as shown in Fig. 2-2(d). Because tunneling is a Poisson process, these time intervals are distributed exponentially, and we fit the histogram to an exponential $Ae^{-\Gamma_{off}t_{on}}$ to obtain the tunnel rate Γ_{off} [42]. We obtain Γ_{on} from the time intervals t_{off} in the same manner.

In the following sections, we use the charge detection technique presented here to measure Γ_{on} and Γ_{off} as a function of drain-source bias V_{ds} and plunger gate voltage V_g in order to probe the underlying physics of the tunneling process. Before discussing these measurements, we give an overview of the more important details of our measurement circuit and technique. A sketch of the circuit used for our measurements is shown in Fig. 2-3(a). We amplify the voltage across the QPC with a Signal Recovery 5184 voltage preamplifier. This amplifier, which sits at room temperature near the top of the dilution refrigerator dewar, is connected to the QPC through a coaxial cable of capacitance $C \approx 500$ pF. When the QPC resistance R_{qpc} changes by a small amount, the voltage across the QPC V_{qpc} takes a time $\tau = R_{qpc}C \approx 50$ μs to respond, where here we have used a typical value for the QPC resistance $R_{qpc} = 100$ kΩ. The time intervals t_{on} and t_{off} must be longer than τ in order for them to be measurable. For all of the data reported here, we have used simulations to check that the finite bandwidth of our measurement does not substantially affect our results [43]. The few small effects that the finite measurement bandwidth has will be noted below. In Section 2.4 and in the following chapter, we quickly modulate the energy of the one
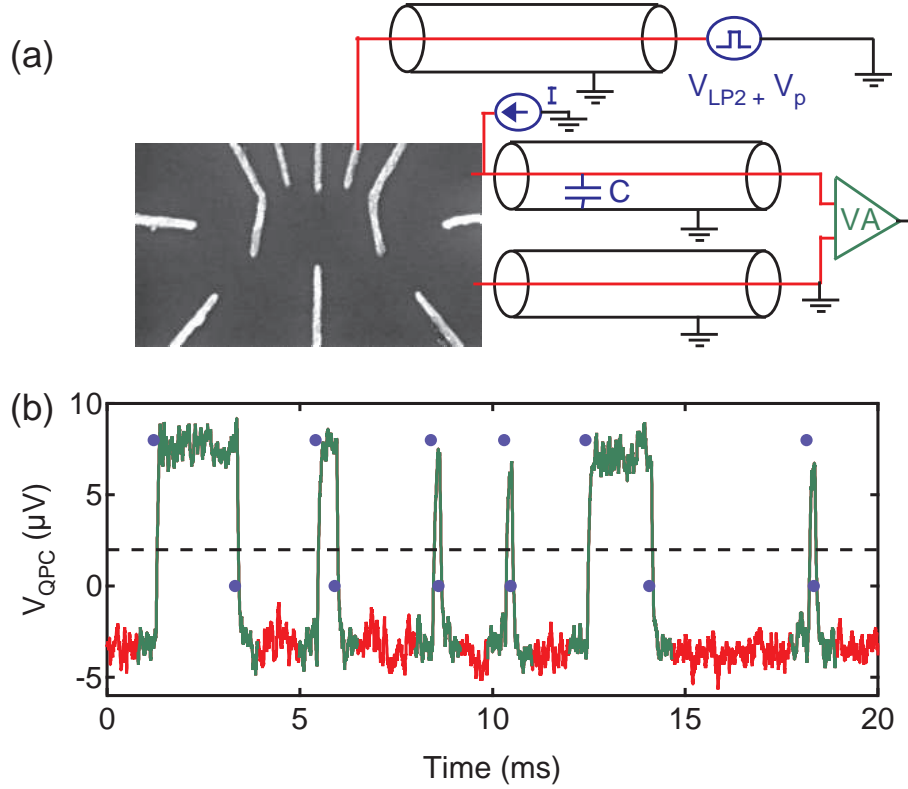
Figure 2-3: (a) Sketch of circuit used for performing time-resolve measurements of the charge on the quantum dot. A current $I = 1$ nA is sourced through the QPC resistance R_{qpc} and the voltage across the QPC V_{qpq} is amplified with a Signal Recovery 5184 voltage preamplifier. The bandwidth for this measurement is limited by the combination of the point contact resistance R_{qpc} and the capacitance C of the coaxial cable connecting the input of the preamplifier to the QPC, as discussed in the main text. To quickly modify the energy of the one electron state relative to the Fermi level, we add a small, high speed voltage V_p to the voltage applied to the gate LP2. (b) Example of automated triggering software. Starting with a voltage vs. time series such as the one shown in Fig. 2-2(c), we first determine where the data exceeds a specified threshold, shown here as a dashed black line, and retain data only in the immediate vicinity of where the threshold is exceeded (green sections of the time series). We then determine where the charge transitions from 0 to 1 (1 to 0) electrons are within these subsections of the data by determining where the derivative of the time series has a large positive (negative) value. We save only the times when the electron tunnels on and off the dot. These times are indicated by the blue circles: The upper value corresponds to an electron tunneling on, the lower value corresponds to an electron tunneling off.

electron state relative to the Fermi level. This is done by adding a small, high speed voltage V_p to the voltage applied to the gate LP2. V_p can be varied on a microsecond time scale.

To measure the rates Γ_{on} and Γ_{off} accurately, we must measure many time intervals t_{off} and t_{on}. For instance, for the histogram shown in Fig. 2-2(d), we have measured almost 10,000 time intervals. The time series shown in Fig. 2-2(c), which contains 10,000 data points, contains only 6 such time intervals. Therefore, the number of data points required for this measurement is approximately 20,000,000, requiring about 0.2 GB of memory. While performing one such measurement does not require too much memory, performing a large number of these measurements results in an unmanageable amount of data.

In order to reduce the amount of data acquired for our measurements, we developed an automated triggering and acquisition system (Fig. 2-3(b)). For each time series that we acquire, our data is passed through a series of triggering algorithms before it is stored. The first algorithm determines when V_{qpc} exceeds a specified threshold. Only data close to where the threshold is exceeded are processed further (green regions in Fig. 2-3(b)). Following the threshold trigger, the data is passed through a second algorithm which finds the times at which V_{qpc} jumps up or down suddenly. This is done by taking the derivative of the time series. Our derivative algorithm is modified so that it works properly for noisy data sets. We take the difference between the average of two sets of data points separated by a specified delay time, and determine when this difference is sufficiently positive or negative, which tells us when an electron tunnels on or off the dot, respectively. The only data that we save are the time and type (tunneling on or off) of each of these charge transitions. Thus, the number of data points we actually save for a 10,000 point trace like the one shown in Fig. 2-3(b) is only 24. Our algorithms work fairly quickly so that the data acquisition time is not dominated by computation. In particular, the threshold trigger is implemented for the purpose of increased speed: It quickly cuts down on the amount of data that must be processed by the more time consuming differentiation algorithm. Our data acquisition system randomly saves a few voltage vs. time series

29

during the coarse of an experiment so that we can check that the triggering is working properly.

2.3 Energy Dependent Tunneling: Drain-Source Bias Dependence

Using the charge sensing technique described in the previous section, we measure the rate for tunneling on Γ_{on} and off Γ_{off} the dot as a function of drain-source bias V_{ds} (Fig. 2-4(a)). These data are taken such that at $V_{ds} = 0$, the one electron state is just above the Fermi level in the leads. In the top panel of Fig. 2-4(a), we see that as V_{ds} is made more negative, Γ_{off} increases exponentially.

To understand the exponential dependence of Γ_{off} on V_{ds}, we start with the WKB formula [44] for an electron tunneling at an energy E through a potential barrier $U(x)$:

$$\Gamma = f_0 e^{-\frac{2}{\hbar} \int \sqrt{2m^*(U(x) - E)} dx} \tag{2.3}$$

Here m^* is the effective mass, f_0 can be regarded semiclassically as an attempt rate, and the integral in the exponent is taken over the classically forbidden region of length w where $U(x) > E$. We linearize this formula for a small a perturbation δE to the electron energy E and a small deviation δU of the tunnel barrier potential $U(x)$:

$$\Gamma = \Gamma_0 e^{-\kappa(\delta U - \delta E)} \tag{2.4}$$

Here κ and Γ_0 depend on the details of the barrier potential. For a quantum dot, for small perturbations to the plunger gate voltage or drain-source bias δV_g and δV_{ds} about arbitrary but fixed V_{ds} and V_g values, the energy states on the dot vary linearly as $\delta E = -e\alpha_{dsE}\delta V_{ds} - e\alpha_{gE}\delta V_g$.[4] Similarly, δU varies linearly as $\delta U = -e\alpha_{dsU}\delta V_{ds} - e\alpha_{gU}\delta V_g$, where α_{dsU} and α_{gU} give the coupling of V_{ds} and V_g to the

---

[4]Here, following the standard capacitor model for a quantum dot [38], α_{dsE} is the ratio of the drain-source capacitance to the total dot capacitance, and likewise α_{gE} is the capacitance ratio for the three plunger gates.

barrier potential. There will, of course, be different parameters α_{dsU}, α_{gU}, κ, and Γ_0 for the two barriers b1 and b2. Note that Γ depends exponentially on $\delta U - \delta E$, and therefore depends exponentially on V_{ds} and V_g: One can show that this holds independent of the particular shape $U(x)$ of the barrier potential, or the shape of the perturbation to the potential induced by the change δV_g or δV_{ds}.

Using this linearization, we can write down equations for the V_{ds} dependence of Γ_{off} and Γ_{on}, considering only the one electron ground state of the dot. We include the Fermi statistics in the leads by assuming that Γ_{on} (Γ_{off}) is the sum of two terms, proportional to the number of electrons (holes) in leads 1 and 2 at the ground state energy E. From the arguments given above, each term must also depend exponentially on V_{ds}, so that we obtain:

$$\Gamma_{off} = \Gamma_{2,0}e^{-\beta_2\delta V_{ds}}(1 - f_2(E)) \tag{2.5}$$
$$+\Gamma_{1,0}e^{\beta_1\delta V_{ds}}(1 - f_1(E))$$
$$\Gamma_{on} = \eta\Gamma_{2,0}e^{-\beta_2\delta V_{ds}}f_2(E) \tag{2.6}$$
$$+\eta\Gamma_{1,0}e^{\beta_1\delta V_{ds}}f_1(E)$$

Here $\beta_{1,2} = \kappa_{1,2}|\alpha_{dsU_{1,2}} - \alpha_{dsE}|$. The energy of the ground state relative to the Fermi level E_F is given by $E = -e\alpha_{dsE}V_{ds} - e\alpha_{gE}\Delta V_g$, where ΔV_g is the shift in V_g from the 0 to 1 electron transition. f_2 and f_1 are the Fermi distribution functions for the two leads $f_2(E) = f(E)$, $f_1(E) = f(E + eV_{ds})$, and η is the ratio between the tunnel rate onto and off of the dot for a given lead when the one-electron ground state is aligned with the Fermi level in that lead. We expect that $\eta = 2$ because of spin degeneracy [24, 45], and use this value in the calculations below.

To understand Fig. 2-4(a), we note that whether Γ increases or decreases with V_{ds} depends on whether lead 1 is better coupled to the barrier or the dot, that is, whether α_{dsU} or α_{dsE} is larger. Since b1 is closer to lead 1 than the dot, and b2 is farther from lead 1 than the dot, it follows that $\alpha_{dsU_1} > \alpha_{dsE} > \alpha_{dsU_2}$ (see Fig. 2-4(c)) [46]. Therefore, tunneling through b1 (b2) increases (decreases) exponentially with increasing V_{ds}. This is reflected in the signs of the exponentials appearing in Eqn.
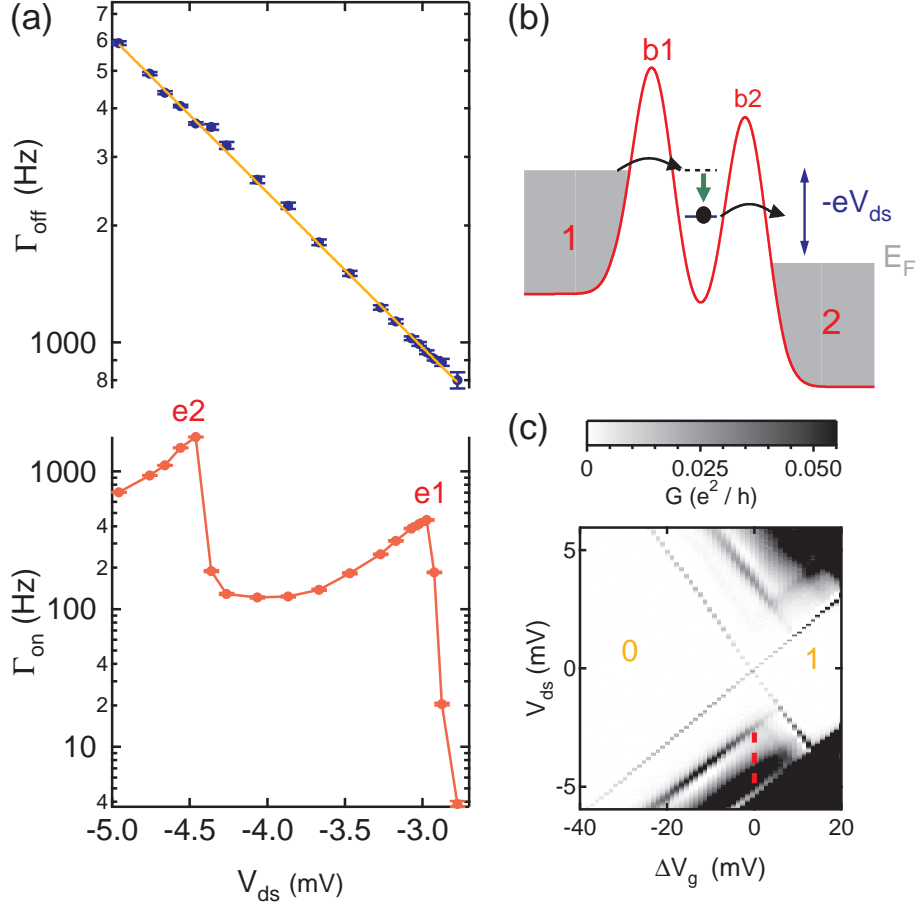
Figure 2-4: (a). Γ_{on} and Γ_{off} as a function of V_{ds} for large negative V_{ds}. The solid line in the upper panel is based on a theoretical fit to the data discussed in the main text. e1 and e2 indicate where the Fermi level in lead 1 is aligned with the 1st and 2nd orbitally excited states, respectively. (b) Sketch of the energy configuration of the quantum dot at the position e1 noted in the lower panel of (a). The solid blue line indicates the ground state of the quantum dot, and the dashed black line indicates the orbitally excited state e1. After tunneling into the orbitally excited state e1, the electron quickly relaxes to the ground state (green arrow), as discussed in the main text. (c) Differential conductance vs. V_{ds} and V_g, showing the 0 to 1 electron transition. The tunnel rates for this case are made large enough so that the differential conductance can be measured using standard transport techniques. The data shown in (a) are taken at the position of the dashed line: At $V_{ds} = 0$, the 1 electron state is close to the Fermi level in the leads. The zero and one electron Coulomb blockaded regions are noted.

2.5 and Eqn. 2.6.

The solid line through the the Γ_{off} data in Figure 5-4(a) is a fit to Eqn. 2.5, which for large negative bias reduces to $\Gamma_{2,0}e^{-\beta_2 \delta V_{ds}}$, since electrons only tunnel onto the dot from lead 2. The rate Γ_{on} is shown as a function of V_{ds} in the lower panel of Fig 2-4(a). At the two points marked e1 and e2 in the figure Γ_{on} increases rapidly as the Fermi energy in lead 1 is aligned with an excited orbital state of the dot [47], as depicted in Fig. 2-4(b). The higher-energy orbital states are better coupled to the leads and thus Γ_{on} rises rapidly when these states become available. From the positions of these rises and a measurement of α_{dsE}, we find the energies of the 1st and 2nd excited orbital states to be 1.9 and 2.9 meV, respectively, above the ground state. These energies can also be measured, with larger tunneling rates through b1 and b2, using standard transport techniques [20] (Fig. 2-4(c)), and the results are consistent.[5]

We note that Γ_{off} does not show any special features at the points e1 and e2.[6] This is because the electron decays rapidly out of the excited orbital states via emission of acoustic phonons [6, 48], and subsequently tunnels off the dot from the ground state [49]. We can therefore continue to use Eqn. 2.5 when there are multiple orbital states in the transport window.

In the regions between the points e1 and e2, Γ_{on} is seen to decrease exponentially as V_{ds} is made more negative, as expected from Eqn. 2.6. Note that this decrease in Γ_{on}, with increasingly negative V_{ds}, occurs even though the number of electrons that could tunnel onto the dot inelastically from lead 1 is growing. This is strong evidence that the tunneling is predominantly elastic, dominated by states very close to the dot energy E. There is, however, an apparent flattening of Γ_{on} above the extrapolated

---

[5]We will not discuss the standard model for Coulomb blockade diamonds as it is covered extensively elsewhere [20, 38]. From these diamonds, measured either with charge detection or with traditional transport techniques, one can measure α_{gE} and α_{dsE}.

[6]There is a very small kink in the Γ_{off} data at the position of e2. This is caused by the finite measurement bandwidth discussed in Section 2.2. When Γ_{on} suddenly becomes very large, some time intervals t_{off} will become too short to be measured. When a time interval t_{off} is missed entirely, the two adjacent time intervals t_{on} are measured as a single interval, which has the effect of decreasing the measured rate Γ_{off} relative to its actual value. We have performed simulations of this effect, and find that it accounts for the observed kink in the Γ_{off} data.

exponential decrease near V_{ds} = -4 mV, close to the second excited orbital state. We find that this line shape is consistent with broadening of the second excited state by a Lorentzian of full-width at half-maximum $\gamma \sim 10$ μeV. Calculated line shapes are shown for the broadening of the first excited state in Fig. 2-5 and are discussed below.

If a square tunnel barrier is assumed, one can compute an effective barrier height U_2 and width w_2 for b2 from the fit in Fig. 5-4(a). For a square barrier $\Gamma = f_0 e^{-2w_2\sqrt{2m^*(U_2-E)/\hbar^2}}$ [44]. Linearizing the square root in the exponential and estimating $\alpha_{dsE} - \alpha_{dsU_2} \sim \alpha_{dsE}$ and $f_0 \sim E_Q/h \sim 1$ THz, where E_Q is the level spacing of the dot, we obtain $w_2 \approx 130$ nm and $U_2 - E_F \approx 5$ meV at $V_{ds} = V_0$. These values are only logarithmically sensitive to f_0 and thus depend very little on our estimate of this parameter. If the barrier is assumed to be a different shape, for instance parabolic (as drawn in Fig. 2-4(b)), similar values are obtained (7 meV and 120 nm for the height and width of the barrier, respectively). The voltages we apply to the gates to create the tunnel barriers are the same order of magnitude as the voltages at which the 2DEG depletes and thus it is reasonable that $U_2 - E_F$ is found to be comparable to the Fermi energy ($E_F \approx 7.7$ meV). The value for w_2 is also reasonable given the dimensions of our gate pattern and heterostructure.

We next examine, in Fig. 2-5, the dependence of Γ_{on} and Γ_{off} on V_{ds} for both positive and negative V_{ds}. The data are taken with V_g adjusted so that E is $\sim kT$ away from the 0 to 1 electron transition at $V_{ds} = 0$. The solid blue and red lines in Fig. 2-5 are calculated from Eqn. 2.5 and Eqn. 2.6 and are in good agreement with the data. Eqn. 2.5 agrees with the Γ_{off} for all V_{ds}, while Eqn. 2.6 agrees with the Γ_{on} data only for small V_{ds}. This is to be expected because Equations 2.5 and 2.6 only take into account the ground orbital state of the quantum dot. As shown above, the orbitally excited states into which electrons can tunnel when V_{ds} is made sufficiently positive or negative increase Γ_{on} but do not affect Γ_{off}. We note that the value of β_1 is very close to the value of β_2. If the height and width of b1 and b2 are comparable one can show that $\beta_1 \sim \frac{\alpha_{dsU_1}-\alpha_{dsE}}{\alpha_{dsE}-\alpha_{dsU_2}}\beta_2$, and it is therefore expected that $\beta_1 \sim \beta_2$.

At $V_{ds} \approx$ -3 mV, Γ_{on} rapidly increases as electrons can tunnel into the first excited
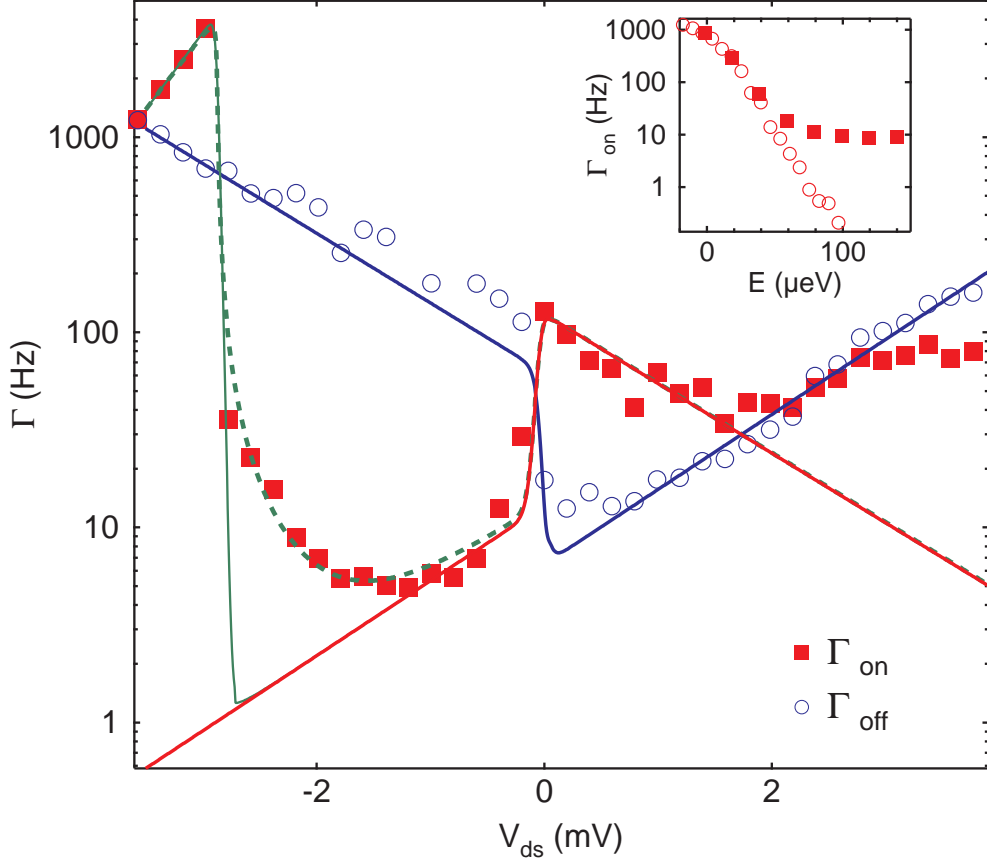
Figure 2-5: Γ_{on} (closed squares) and Γ_{off} (open circles) as a function of V_{ds}. The solid red and blue curves are calculations of Γ_{on} and Γ_{off} that include only the contribution of the ground state to the tunneling as described in the main text. The dashed and solid green lines are calculations of Γ_{on} including the contribution of one of the orbitally excited states to the tunneling for $V_{ds} < 0$ with and without Lorentzian broadening, as discussed in the main text. The step features near $V_{ds} = 0$ result from the Fermi distribution. (Inset) Rate of tunneling into the ground state (open circles) and orbitally excited state e1 (closed squares) as a function of the energy E of the state relative to the Fermi level in the lead from which the electron is tunneling. The orbitally excited state data are taken from the drain-source bias dependence shown in the main part of this figure. The ground state data are taken by modulating the ground state energy relative to the Fermi level using the gate voltages as is discussed in the following section. The data are scaled vertically so that they agree for $E < 0$.

state e1 of the quantum dot from lead 1. In order to extend our model to account for this we add a term $\eta \hat{\Gamma}_{1,0} e^{\hat{\beta}_1 \delta V_{ds}} f_1(E + E_1)$ to Eqn. 2.6. Here E_1 is the energy of e1 relative to the ground state. Eqn. 2.6 with this addition is plotted as the solid green line in Fig. 2-5. We see that, for V_{ds} slightly more positive than $V_{ds} \approx$ -3 mV, where e1 is just above the Fermi level in lead 1, the Γ_{on} data is substantially larger than the solid green line. In order to account for this enhancement of Γ_{on} when e1 is just above the Fermi level, we included broadening of the state e1 by a Lorentzian of full-width at half-maximum $\gamma = 13$ μeV. This calculation is plotted as a dashed green line in Eqn. 2.6, and agrees well with the data. Γ_{on} also deviates from the solid curve for $V_{ds} \gtrsim 2$ mV: This deviation may be caused by broadening of the first excited state as well. Though the lineshape including Lorentzian broadening of e1 agrees quite well with the data, it is not clear what this broadening could come from. Relaxation from e1 to the ground state via acoustic phonon emission leads to energy broadening of e1. However, from the time-energy uncertainty principle, the energy broadening observed here corresponds to a lifetime of $\tau_{e1} = 50 \pm 30$ ps, and while emission of acoustic phonons can lead to lifetimes ~ 100 ps, for our device parameters we expect much slower relaxation from this mechanism [6, 48, 50]. The tunneling process, which as shown above is much slower than the acoustic phonon relaxation, leads to a negligible amount of energy broadening. The enhancement of Γ_{on} could alternatively be caused by inelastic processes, which might begin to contribute significantly to Γ_{on} when the ground state of the dot is sufficiently far below the Fermi energy.

An obvious question with regards to this apparent broadening of the excited state e1 is whether the same broadening is observed for the ground state. As is shown in the inset to Fig. 2-5 the answer appears to be no. Using the data in the main plot, we plot Γ_{on} as a function of the difference between the energy of e1 and the Fermi level in lead 1. In a separate experiment, we vary the energy E of the ground state relative to the Fermi level using the gates (as discussed in the following section), and we plot Γ_{on} vs. E for this data set in the inset at well. As the ground state is brought above the Fermi level, Γ_{on} drops exponentially over almost four orders of magnitude as the number of electrons in the leads at the energy of the ground state drops. In contrast,

for the excited state e1, Γ_{on} drops exponentially by about two orders of magnitude but then decreases only very slowly as e1 is brought further above the Fermi level in lead 1. Thus, while there is an apparent excess of tunneling into the dot when the excited state e1 is just above the Fermi level, there is no excess tunneling into the dot when the ground state is just above the Fermi level.

2.4 Gate Voltage Dependence

We now turn to the dependence of the rates Γ_{on} and Γ_{off} on the plunger gate voltage V_g applied to LP1, PL, LP2. For these measurements, the voltages applied to the gates defining the tunnel barriers (OG, SG1 and SG2) are tuned so that the tunneling through b1 is negligible. The dot can therefore be regarded as a localized state coupled to a single Fermi sea in lead 2 through the tunnel barrier b2. In the region near $\Delta V_g = 0$ (where the energy of the one electron state is close to the Fermi level, so that $E \sim$ kT), the electron hops on and off the dot spontaneously because there are both electrons and holes in the lead at these energies, as depicted in Fig. 2-6(a). In this region we measure the rates Γ_{on} and Γ_{off} in the same way as for the drain-source bias dependence.

A pulsed technique (Fig. 2-6(b) and Fig. 2-6(c)) is used to measure Γ_{off} when ΔV_g is made sufficiently negative that the one electron state is far above the Fermi level and thermally assisted electron tunneling ceases. We begin with the electron energy, determined by ΔV_g, well above E_F. We then apply a positive voltage pulse to the gate LP2 to bring the ground state near E_F, so that an electron can hop onto the dot. A short time after the pulse the electron will hop off the dot because it is above E_F. We record the time at which this occurs, Δt, measured relative to the end of the gate pulse. This process is repeated and we make a histogram of the number of tunnel-off events vs. Δt (Fig. 2-6(d)); an exponential fit to this histogram yields Γ_{off}. An analogous technique is used to measure Γ_{on} when the one-electron state is well below E_F.

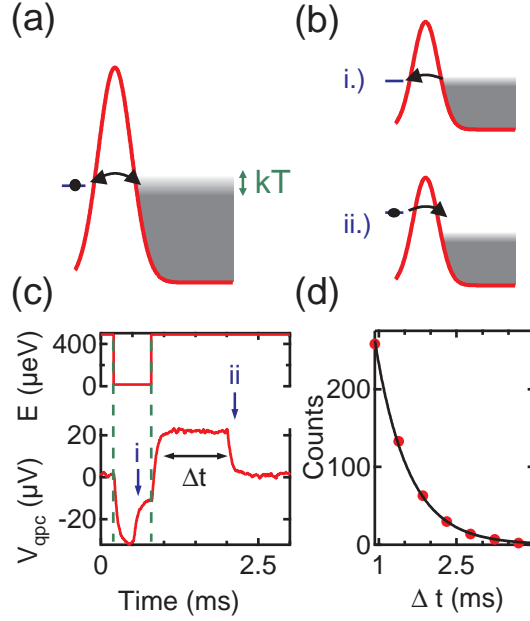Using these techniques, we measure Γ_{on} and Γ_{off} as functions of ΔV_g (Fig. 2-7).

37

Figure 2-6: (a) When the one electron state is within kT of the Fermi level, electrons spontaneously hop on and off the dot, and the tunnel rates Γ_{on} and Γ_{off} can be measured in the same way as illustrated in Fig. 2-2 and Fig. 2-3 . (b) When the one electron state is far above the Fermi level, we use a pulsed technique to measure Γ_{off}. We quickly bring the one electron state close to or below the Fermi level so that an electron can tunnel on (i). We then quickly change the gate voltage applied to LP2 so that the one electron state is above the Fermi level, and watch for when the electron tunnels off (ii). (c) Pulsed technique used to measure Γ_{off} as sketched in (b). The top panel shows the pulsed modulation of the one-electron state energy E relative to the Fermi level. The bottom panel shows a sample time trace. The dashed vertical lines indicate when the gate pulse begins and ends: The QPC responds to the gate pulse because of direct capacitive coupling to LP2. When the electron energy level is brought near E_F an electron tunnels onto the device (indicated by a i). When the electron level is brought back above E_F the electron tunnels off the device (indicated by a ii) We measure the time interval Δt between when the the pulse ends and when the electron tunnels off of the dot. The fact that the charge transitions i and ii appear to be rounded in time rather than instantaneous is a result of the finite measurement bandwidth discussed in Section 2.2.(d) We measure a large number of time intervals Δt using our automated triggering system and histogram them. We fit to an exponential (solid black line) to obtain Γ_{off}, as described in the main text.

38

To understand the V_g dependence of Γ, we note that $\delta U_2 = \alpha_{gU_2}\delta V_g$, and $\alpha_{gU_2} < \alpha_{gE}$ because the three plunger gates are closer to the dot than they are to b2. Starting at the far left of Fig. 2-7, we see that Γ_{off} decreases exponentially as ΔV_g is made less negative: This happens because the electron energy is being moved farther from the top of the tunnel barrier b2. Γ_{off} decreases rapidly at the 0 to 1 electron transition as the ground state is brought below E_F.[7] Γ_{on} increases rapidly as the ground state is brought below the Fermi level, but then as ΔV_g is made more positive, Γ_{on} decreases exponentially as the one electron state is brought farther below the top of the tunnel barrier. This decrease in Γ_{on} is further evidence that the tunneling is elastic because as the one-electron state is brought farther in energy below E_F there are more electrons that could tunnel onto the dot inelastically at energies closer to the top of the barrier. Γ_{on}, however, decreases because elastic tunneling onto the dot happens at a lower energy relative to the barrier height.

We can model the data in Fig. 2-7 by writing down equation similar to Equations 2.6 and 2.5:

$$\Gamma_{off} = \Gamma_0 e^{-\beta\Delta V_g}(1 - f(E)) \tag{2.7}$$

$$\Gamma_{on} = \eta\Gamma_0 e^{-\beta\Delta V_g}f(E) \tag{2.8}$$

The solid lines in Fig. 2-7 are fits to these equations and describe the data well. Here again we take $\eta = 2$, although the fit is improved with a smaller value for η [24]. We note that the value for β obtained from the ΔV_g dependence is smaller than the value for β_2 obtained from the V_{ds} dependence: This is expected because $\alpha_{gE} < \alpha_{dsE}$.

In the inset to Fig. 2-7(a) we plot the probabilities that the electron is on and off the dot, given by $p_{on} = \Gamma_{on}/(\Gamma_{on} + \Gamma_{off})$ and $p_{off} = \Gamma_{off}/(\Gamma_{on} + \Gamma_{off})$, respectively. We compare p_{on} and p_{off} to the expected Fermi statistics $f(E, \eta) = \frac{1}{1 + \frac{1}{\eta}e^{E/kT}}$ and

---

[7]For some experiments, our data acquisition system periodically measures the rate Γ_{off} with the one electron state just below the Fermi level (or Γ_{on} when the one electron state is just above the Fermi level). One of the three plunger gates is automatically adjusted to keep the measured rate constant. Because of the exponential sensitivity of Γ_{off} to the energy of the ground state below the Fermi level, this feedback system very accurately compensates for any drift or switches caused by electrostatic fluctuations in the heterostructure [27].
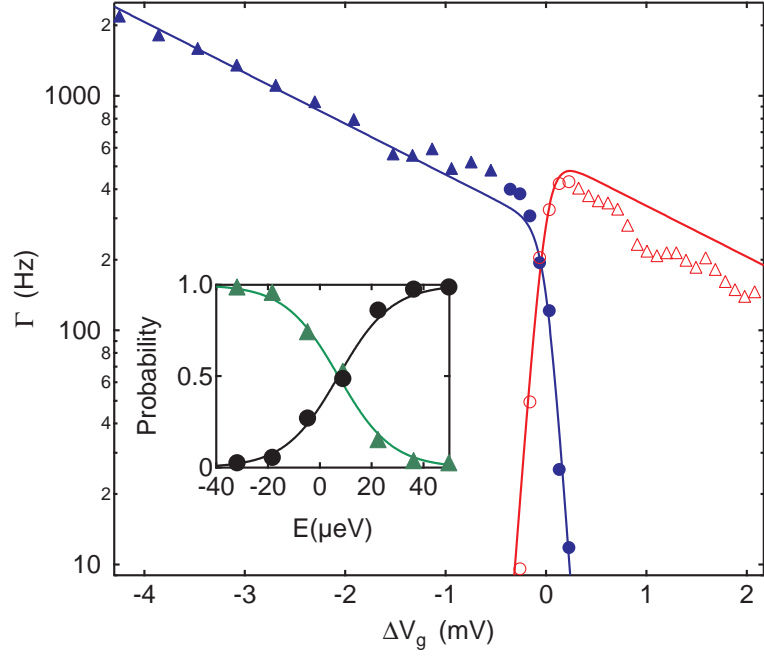
Figure 2-7: Γ_{on} and Γ_{off} as a function of ΔV_g. Closed (open) circles are Γ_{off} (Γ_{on}) measured by observing spontaneous hopping caused by thermal broadening in the leads as depicted in Fig. 2-6(a) and discussed in the main text. Closed (open) triangles are Γ_{off} (Γ_{on}) measured using the pulsed gate technique depicted in Fig. 2-6(b-d) and discussed in the main text. The solid lines are calculations described in the main text. (Inset) p_{on} (triangles) and p_{off} (circles) compared to $f(E, \eta)$ and $1 - f(E, \eta)$ (solid lines), respectively, as described in the text.

$1 - f(E, \eta)$, respectively, and find good agreement [49]. These fits are consistent with an electron temperature of T = 120 mK.

Thus we see that almost all aspects of tunneling at zero magnetic field can be described in terms of our simple model for energy dependent tunneling. In the following chapter, we extend this model to investigate tunneling in large magnetic fields, a situation in which, surprisingly, we find that the tunneling is highly spin dependent.

Chapter 3

Spin Dependent Tunneling in Large Magnetic Fields

In this chapter we extend the measurement techniques and theoretical models presented in the previous chapter to study tunneling in large magnetic fields. In Section 3.1 we show how we use our tunneling rate measurements to extract the spin dependent tunneling parameter χ, defined as the ratio of the rate for tunneling into the excited spin state of the dot to the rate of tunneling into the ground state when both states are below the Fermi level. We then show that, for large magnetic fields, χ can be very close to zero, so that tunneling into the excited spin state is almost completely suppressed. In Section 3.2, we show that at large magnetic fields χ depends very strongly on the shape of the potential landscape defining the quantum dot. We then end our discussion of single electron tunneling, and, in Section 3.3, we discuss charge detection with GaAs quantum point contacts more generally, in order to motivate the chapters that follow. Parts of the work presented in this chapter appear also in Amasha *et al.* [25, 51].

3.1 Measurement and Magnetic Field Dependence of χ

Using the same techniques described in Section 2.4, we measure the rate Γ_{on} of tunneling into a lateral GaAs quantum dot containing zero electrons as a function of the voltage ΔV_{LP2} applied to the plunger gate LP2, in the presence of a magnetic field applied in the plane of the 2DEG.[1] Here ΔV_{LP2} is the voltage relative to the value for which the ground spin state is aligned with the Fermi level. As in Section 2.4, for all of the experiments discussed in this chapter, tunneling through the barrier b1 has been made to be negligibly small compared to the tunneling through b2.

In Fig. 3-1(c), we measure Γ_{on} as a function of ΔV_{LP2} at $B = 5$ T. As ΔV_{LP2} is made more positive, Γ_{on} increases at two points labeled by $\uparrow$ and $\downarrow$, which correspond to when the ground spin state is aligned with E_F (Fig. 3-1(a)) and when the excited spin state is aligned with E_F (Fig. 3-1(b)), respectively. Other than at these two points, Γ_{on} decreases exponentially with increasing ΔV_{LP2} as the two spin states are brought farther below the top of the tunnel barrier b2, as discussed in Section 2.4. To model this tunneling lineshape quantitatively, we add a term corresponding to tunneling into the excited spin state to Eqn. 2.8:

$$\Gamma_{on} = \Gamma_0 e^{-\beta \Delta V_{LP2}} (f(E) + \chi \cdot f(E + \Delta)) \tag{3.1}$$

Here the energy splitting between the excited and ground spin states is given by $\Delta = g\mu_B B$, where g is the g-factor for electrons in the 2DEG and μ_B is the Bohr magneton. The factor χ describes any possible spin dependence to the tunneling. Specifically, χ is the ratio of the rate of tunneling into the excited spin state to the rate of tunneling into the ground spin state when both spin states are far below E_F. A fit to Eqn. 3.1 is given by the solid line in Fig. 3-1(c), and agrees well with the data. From fits such as this one, and a measurement of the coupling of LP2 to the dot energy α_{LP2}, we obtain a value for g consistent with previous reports [52]. The

---

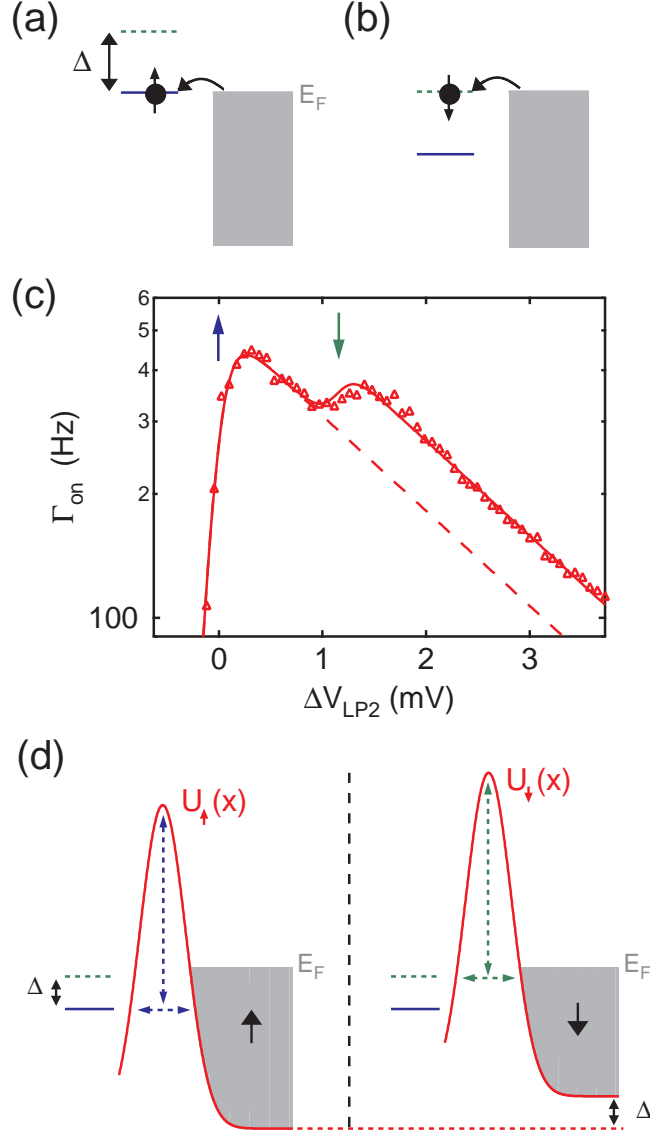[1]Specifically, referring to Fig 3-4(a) below, the magnetic field points along the y direction.

Figure 3-1: Measurement of χ. We measure Γ_{on} as a function of the ΔV_{LP2}, the voltage applied to LP2 (relative to a much larger offset), in the presence of a magnetic field B applied in the plane of the 2DEG. (a) As ΔV_{LP2} is made more positive, Γ_{on} grows as the one electron ground state is brought below the Fermi level. Here, the solid blue line indicates the energy of the ground spin state of the dot, and the dashed green line indicates the energy of the excited spin state of the dot. (b) Γ_{on} grows again when the excited spin state is brought below the Fermi level. (c) Measurement of Γ_{on} vs. ΔV_{LP2} at B = 5 T. The solid and dashed lines are theoretical calculations described in the main text. The ΔV_{LP2} values for which the ground and excited spin states are aligned with the Fermi level are indicated by $\uparrow$ and $\downarrow$, respectively. (d) Potential landscape as seen by the ground (left) and excited (right) spin states. The Fermi level is the same for the two spin states, but the potential is shifted, as discussed in the main text. The height and width of the tunnel barrier for each spin state are indicated by the dashed vertical and horizontal lines with arrows at the ends.

same calculation, but with χ set to zero, is plotted as the dashed line in Fig. 3-1(c). By fitting tunneling lineshapes like the one shown in Fig. 3-1(c) to Eqn. 3.1 we can extract the spin dependent tunneling parameter χ. Graphically ,the value of χ extracted from the fit is given by the ratio of the solid line to the dashed line at a fixed value of ΔV_{LP2} more positive than the value indicated by $\downarrow$.

Based on simple considerations, we expect $\chi = 1$. The reasoning behind this expectation is not immediately obvious. Because the excited spin state is higher in energy, one might expect faster tunneling into the excited spin state than into the ground spin state based on the arguments in Chapter 2. However, a more careful consideration of the problem reveals that this is not the case, as is depicted in Fig. 3-1(d). Here we plot the potentials $U_\uparrow(x)$ and $U_\downarrow(x)$ for the ground and excited spin states, respectively. They are not the same: The potential for the excited spin state is shifted relative to the potential for the ground spin state, so that $U_\uparrow(x) = U_\downarrow(x) + \Delta$. This shift in the conduction band potential is the same shift that is responsible for Pauli paramagnetism [53] in metals. Pauli paramagnetism can be seen explicitly in Fig. 3-1(d): Because the Fermi level is the same for each spin state, but the potential for the excited spin state is larger, there are more ground state spins than excited state spins in the Fermi sea lead. Because of this potential shift, the excited spin state sees a tunnel barrier which is shifted up by Δ relative to the ground state tunnel barrier. This shift cancels exactly the increased energy Δ of the excited spin state. The effective heights and widths of the tunnel barriers for elastic tunneling into the two spin states are indicated in Fig. 3-1(d): They are exactly the same, and thus we expect $\chi = 1$.

In Fig 3-2 we plot χ as a function of B. For low fields $B \approx 3$ T, we see that as expected $\chi \approx 1$. However, as the field is increased, χ gets smaller, and at B = 7.5 T we have $\chi = 0$ to within the resolution of the experiment, so that tunneling into the excited spin state is completely suppressed relative to tunneling into the ground spin state. This spin dependent tunneling effect is very clear from looking at the Γ_{on} vs. ΔV_{LP2} traces, as is shown in the insets to Fig 3-2. At $B = 3$ T, there is a clear rise in Γ_{on} when the excited spin state is brought below the Fermi level, whereas at $B = $
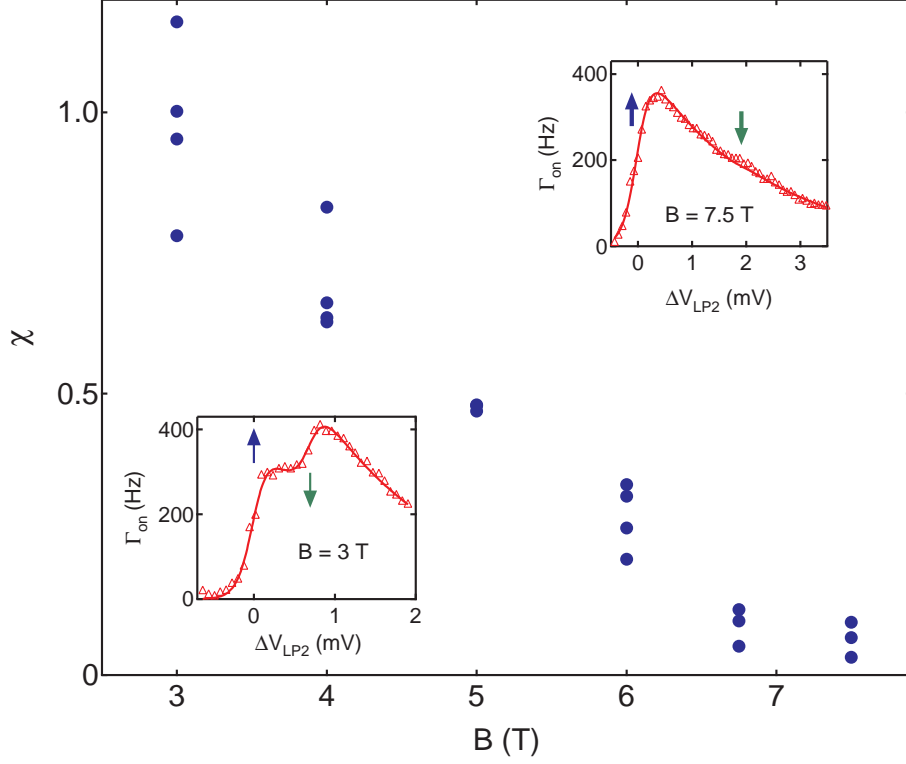
46

Figure 3-2: Measurement of χ as a function of B, as described in the main text. The multiple values for χ at a single value of B are obtained by measuring the lineshape and fitting to obtain χ multiple times. The data from which χ is extracted for $B = 3$ T and $B = 7.5$ T are shown in the two insets. For each of these insets, the ΔV_{LP2} values for which the ground and excited spin states are aligned with the Fermi level are indicated by $\uparrow$ and $\downarrow$, respectively. At $B = 3$ T there is a clear rise in Γ_{on} when the excited state is aligned with the Fermi level. However, at $B = 7.5$ T, there is no feature in the tunneling lineshape when the excited state is aligned with the Fermi level, showing that tunneling into the excited spin state is completely suppressed. For these data, $V_{sg1} \approx$ -1350 mV, as is discussed in the main text.

7.5 T there is no such rise.

Our tunneling lineshapes allow us to extract the ratio χ of tunneling into the excited spin state to tunneling into the ground spin state. However, from these lineshapes, as χ approaches zero at large magnetic fields, we cannot determine whether tunneling into the excited state is getting slower or whether tunneling into the ground state is getting faster. This is because for each magnetic field we have adjusted the voltages applied to SG2 and OG so that for each data set the value of Γ_0 is roughly the same. This is done for two reasons. First, though Γ_0 is usually stable over the coarse of a given lineshape measurement, it sometimes drifts or switches suddenly over the course of many days, in particular during Helium transfers. The data for each of these lineshape takes ~ 4 hours to acquire, and generally we acquired a number of lineshapes at each field in order to check that the lineshape is stable, and not corrupted by any possible switching events during the course of the measurement. Therefore, the data shown in Fig. 3-2 takes a number of days to acquire, during which time Γ_0 cannot be expected to remain stable.

The second reason we retuned Γ_0 at each field is that Γ_0 decreases with increasing field, and performing the lineshape measurements with very slow tunneling rates is very time consuming. This decrease in Γ_0 may be caused by the effect of the magnetic field on the orbital degrees of freedom of the quantum dot. Though the magnetic field is aligned to be in the plane of the 2DEG,[2] it is possible that at high magnetic fields there is a small component of the magnetic field perpendicular to the 2DEG caused by misalignment, which would tend to increase the strength of the confinement of the electron in the dot [54]. We expect that Γ_0 would be exponentially sensitive to such an effect. However, we do not expect that a perpendicular field could be responsible for the decrease in the spin dependent tunneling ratio χ. The picture presented in Fig. 3-1 can be extended to include the effects of a perpendicular magnetic field, including Landau quantization in the lead, and one would still expect equal tunnel barrier heights and widths for the two spin states in the presence of a perpendicular magnetic field. Thus, while any observed dependence of Γ_0 on magnetic field could

---

[2]To within 5 degrees [51]

simply be caused by field misalignment, the observed decrease in χ defies such a simple explanation.

Currently there is no established theoretical model for this unexpected spin dependent tunneling effect. In the following section, we present more of the phenomenology of spin dependent tunneling, and then speculate as to what could be possible causes of this effect.

3.2   Shape Dependence of Spin Dependent Tunneling

In addition to depending strongly on magnetic field as was shown in the previous section, the spin dependent tunneling parameter χ depends strongly on the shape of the potential defining the quantum dot. In order to probe the shape of the confining potential, we measure the energy level structure of the quantum dot. Specifically, we measure the energy of the first two excited states relative to the ground state using the technique shown in Fig. 3-3. In order to measure these energies, we simply measure Γ_{on} vs. ΔV_{LP2} at $B = 0$. The results are shown in Fig. 3-3(c): We see two clear rises in Γ_{on} with increasing ΔV_{LP2} labeled e1 and e2, which occur when the 1st and 2nd excited states are aligned with the Fermi level, respectively. Other than at these two points, Γ_{on} decreases with increasing ΔV_{LP2} as the energy levels are brought further below the top of the tunnel barrier. The positions of these two rises, combined with a measurement of α_{LP2}, give us the energies of e1 and e2 relative to the ground state. However, because the tunnel rates shown in Fig. 3-3(c) are well above the bandwidth of our measurement circuit (Section 2.2), we cannot use the techniques presented in Section 2.4 to obtain this data.

To measure the lineshape shown in Fig. 3-3(c), we use the technique illustrated in Fig. 3-3(a). We start with the ground state of the dot so that it is above the Fermi level, so that an electron tunnels off, ionizing the dot. We then pulse the ground state below the Fermi level, for a time t_{load}. During this time period, an electron can
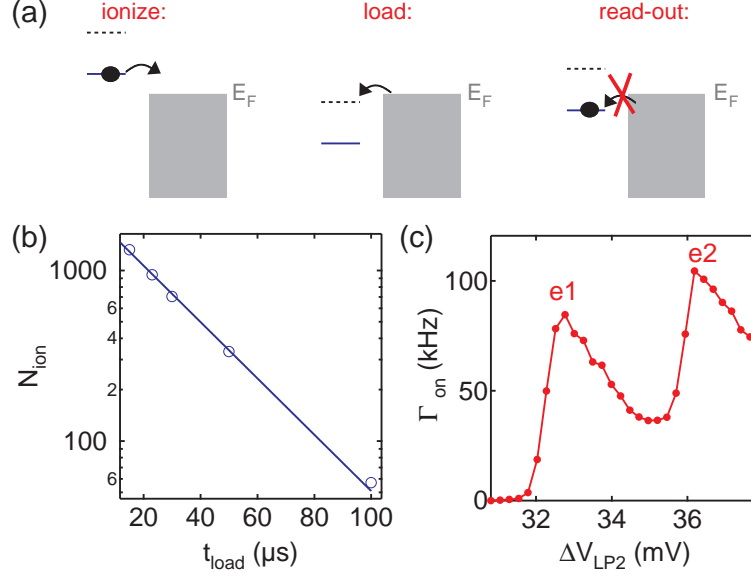
49

Figure 3-3: Measurement of the excited orbital state energies at zero field. Tunneling into the excited orbital states can be very fast, so we use a pulsed gate technique. (a) We bring the ground state of the dot above the Fermi level to ionize the dot. Here the solid blue line is the ground state and the dashed black line is an excited orbital state. (b) We pulse the gate voltage ΔV_{LP2} so that the ground state is brought below the Fermi level for a time t_{load}, and an electron hops on the dot. (c) We bring the ground state just below the Fermi level. If an electron hopped on the dot in step (b), another one cannot hop on. If the dot is still ionized, an electron will hop on, and we detect this with our QPC. (c) Histogram of the number of times N_{ion} the dot is ionized at the start of the read-out state vs. t_{load}. The solid line is a theoretical fit from which we extract Γ_{on} as described in the main text. (c) Γ_{on} as a function of ΔV_{LP2}. The values of ΔV_{LP2} where the first and second excited states are aligned with the Fermi level in the lead are indicated. The data shown in these examples are taken with V_{SG1} = -856 mV

tunnel either directly into the ground state of the dot, or, if an excited state is below the Fermi level, into an excited state from which it relaxes rapidly to ground state. After the pulse, we align the ground state just below the Fermi level. If an electron tunneled onto the device during the loading pulse, another electron cannot tunnel on. However, if the dot remained ionized during the loading pulse, an electron can tunnel onto the dot, and we detect this with the QPC. We iterate this sequence a large number of times, and count the number of times the dot remains ionized N_{ion} after a loading pulse of length t_{load}. We repeat this experiment for different values of t_{load}, and plot N_{ion} as a function of t_{load} (Fig. 3-3(b)). For small t_{load}, the pulse is not long enough for an electron to tunnel onto the dot, and the dot is usually ionized after the loading pulse. For large t_{load}, an electron usually tunnels onto the dot during the loading pulse so that N_{ion} is small. We fit the N_{ion} vs. t_{load} data to an exponential $N_{ion} = Ae^{-\Gamma_{on}t_{load}}$ to obtain the tunnel rate Γ_{on}. The maximum rate Γ_{max} that can be measured using this pulsed gate method is limited by the minimum pulse time t_{load}, which for our circuit is a few microseconds, so that $\Gamma_{max} \sim 100$ kHz. This is considerably faster than the maximum rate that can be measured using the technique presented in Section 2.4. The method presented in Section 2.4 is limited by the RC time constant of our measurement circuit $\tau \approx 50$ μs (Section 2.2), and cannot be used to measure rates faster than about 10 kHz.

In order to change the shape of the dot, we vary the voltages applied to the three plunger gates (LP1,PL, and LP2) and SG1. Starting with SG1 very negative, and the plunger gates less negative, we expect that the confining potential is stronger along the x direction as shown in Fig. 3-4(a). If we model the confining potential as a two dimensional harmonic oscillator $U(x,y) = \frac{1}{2}m^*\omega_x x^2 + \frac{1}{2}m^*\omega_y y^2$, the lowest energy excited state corresponds to an orbital excitation along the y direction having an energy $E_y = \hbar\omega_y$ relative to the ground state. Next, we make the voltage applied to SG1 less negative, and the voltages applied to the three plunger gates more negative, so that the energy of the ground state relative to the Fermi level is left unchanged. For this case, we expect that the lower energy excited state is an orbital excitation along x of energy $E_x = \hbar\omega_x$. If V_{SG1} is made slightly more positive than the most

51

positive value used in these experiments $V_{SG1} \approx$ -690 mV, then the tunneling rate through the barrier b1 starts to become appreciable: It is necessary to stretch the electron wavefunction out along the x direction in order for the tunnel rates through b1 and b2 to be comparable.

We measure the energies of the first two excited states using the methods discussed above as a function of the voltage applied to SG1 while simultaneously changing the voltages applied to the three plunger gates, and the results are shown in the top panel of Fig. 3-3(b). We see that one of the excited state energies increases and the other one decreases as SG1 is made less negative, and we therefore determine these states to be the y and x orbital excitations, respectively.[3] For the most positive voltage applied to SG1, we can make a small change in V_{SG1} so that tunneling through b1 is appreciable, and determine the excited state energies from a drain-source bias dependence (Section 2.3). The results agree with the energies obtained from the pulsed gate method as expected. We note that in order to change the energy level spectrum of the quantum dot appreciably, we need to change the voltage applied to SG1 as well as to the three plunger gates by $\sim$ 200 mV. The changes in the voltage applied to LP2 ΔV_{LP2} we use to perform tunneling lineshape measurements are much smaller than this, so that for each tunneling lineshape, the energy level spectrum is well defined.

At each value of V_{SG1}, in addition to measuring the energy level spectrum of the quantum dot, we measure the spin dependent tunneling factor χ at $B = 7.5$ T, and this is plotted in the bottom panel of Fig. 3-4(b).[4] For the most negative value of SG1, the regime in which the data in Fig. 3-2 were taken, $\chi = 0$. However, as V_{SG1} is made more negative, χ grows, peaking at $\chi \approx 1$ when the confining potential is

---

[3]Further evidence that this interpretation is correct is given in Amasha *et al.* [25]. In this paper we measure relaxation from the excited spin state to the ground spin state, which for the direction of the magnetic field used in our experiment, is theoretically predicted to be sensitive to the energy $\hbar\omega_y$, but not to the energy $\hbar\omega_x$.

[4]Here, as in Section 3.1, we measure only the ratio of the rates for tunneling into the two spin states from the tunneling lineshapes, and not the absolute rates. Γ_0 changes drastically as we manipulate the shape of the dot by making large changes in the voltages applied to SG1 and the plunger gates, and we therefore retune Γ_0 to roughly the same value at each V_{SG1} by adjusting the voltages applied to OG and SG2.
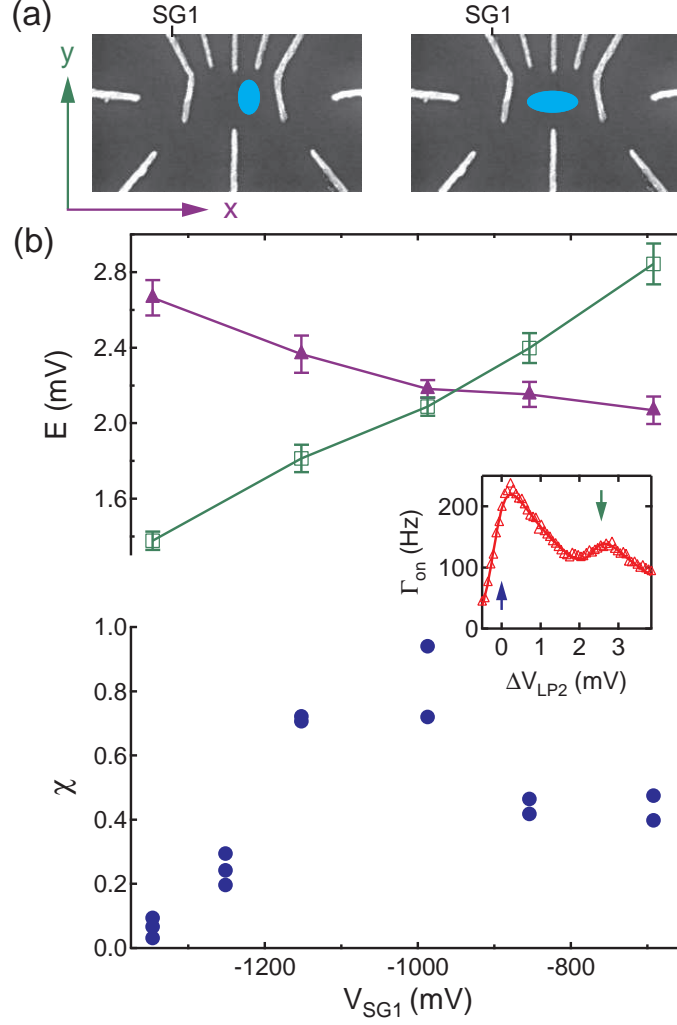
Figure 3-4: (a) When V_{SG1} is made very negative (left hand side), and the plunger gates are made more positive, the dot is confined more strongly along the x direction, so that the lower energy excited state is E_y. The shape of the electron wavefunction is indicated by the blue oval. When V_{SG1} is made more positive (right hand side) and the voltages applied to the plunger gates are made more negative, the dot is confined more strongly along the y direction, so that the lower energy excited state is E_x. The shape of the electron wavefunction, indicated by the blue oval, is now more extended along the x direction than along the y direction (b) Excited state energies and χ as a function of the shape of the dot at $B = 7.5$ T. The multiple values for χ at a single value of V_{SG1} are obtained by measuring the lineshape and fitting to obtain χ multiple times. The green squares and purple triangles in the upper plot indicate E_y and E_x, respectively. The inset shows the data from which we obtain $\chi \approx 1$ at V_{SG1} = -987 mV, the solid line is the theoretical fit from which we extract χ as discussed in the main text. The ΔV_{LP2} values for which the ground and excited spin states are aligned with the Fermi level are indicated by ↑ and ↓, respectively. There is a clear rise in Γ_{on} as the excited state is brought below the Fermi level.

approximately circular ($\hbar\omega_x \approx \hbar\omega_y$). An example of the tunneling lineshape in this regime is shown in the inset to Fig. 3-4(b): The rise in Γ_{on} when the excited spin state is brought below the Fermi level is clearly visible. As SG1 is made more positive, χ decreases. Thus, we see that χ depends strongly on the shape of the potential landscape defined by the voltages applied to the gates. It is important to note that it is likely that we are changing aspects of the potential landscape other than the energy level spectrum as we vary the gate voltages. For instance, as V_{SG1} is made more positive the position of the minimum in the electrostatic potential moves in the negative x direction. We observe this directly: α_{LP2} decreases as V_{SG1} is made more positive, indicating that the average position of the electron wavefunction is being moved farther away from the gate LP2. It is possible that this change, or other changes which we cannot directly measure, is responsible for the variations in χ, rather than the changes in the energy level spectrum. However, the fact that there is a peak of $\chi \approx 1$ when the dot is circular does seem to suggest that the spin dependent tunneling is correlated directly with the energy level spectrum.

The mechanism of the spin dependent tunneling observed in our experiments is not clear. We have left a number of effects out of the simple picture which lead us to expect $\chi = 1$ (Fig. 3-1d)), for instance, spin-orbit effects and electron-electron interactions. We find it unlikely that the spin-orbit interaction plays an important role in the spin-dependent tunneling we observe. We would expect spin-orbit effects to become more important as the strength of the spin-orbit interaction is made larger in comparison to Δ. However, the spin dependent tunneling we observe becomes important at large Δ rather than at small Δ. It is possible that electron-electron interactions, in particular the exchange interaction which is sensitive to spin [55], play some role, but a microscopic picture of the mechanism is not clear.

3.3 Closing Remarks on Charge Detection with GaAs Nanostructures

The results presented in this chapter and in the previous chapter could not be obtained by measuring current. In order for an appreciable current to flow through a quantum dot, the rates for tunneling in and out of the dot must be very fast. For most current amplifiers, the current noise over a 1 Hz bandwidth is ~ 100 fA.[5] The smallest tunneling rate for which there is a measurable current is therefore $\Gamma \sim 100$ fA/e ~ 1 MHz. This is only ~ 100 times slower than the rate of relaxation between different orbital states [6]. We therefore expect that for the measurement of current, an interpretation of variations in the tunneling rates would be complicated by the fact that the tunneling rate is not always slower than the other time scales of the system. Furthermore, accurate measurements of such small currents are not trivial, and if the current is made significantly larger ~ 1 nA, a number effects, such as lifetime broadening of the energy states on the dot [56] as well as cotunneling processes [57], can start to become important. In contrast, the measurement of charge allows us to easily probe the quantum dot when the tunnel rates are extremely slow compared to other processes.[6]

Charge detection also allows separate measurements of the tunneling rates through barriers b1 and b2. Because the drain-source bias has a different effect on b1 than on b2 (Section 2.3), the ability to measure the tunnel rates through b1 and b2 separately is a significant advantage. It is possible to extract the tunnel rates for individual tunnel barriers from measurements of current for three terminal quantum dots [58]. However, these measurements are significantly more complicated than those presented here, and for the usual two terminal devices, only the series combination of the tunnel rates is accessible. Furthermore, charge detection in combination with pulsed gate techniques (Section 2.4) enables the measurement of the tunnel rate when the

---

[5]Current amplifier noise is discussed more extensively in Section 4.4.

[6]One exception is the relaxation rate from the excited spin state to the ground spin state, which can be very slow [25]. This makes an investigation of the dependence of Γ_{off} on V_g and V_{ds} in the presence of a magnetic field very complicated because the electron can tunnel out of either the ground spin state or the excited spin state.

quantum dot is Coulomb blockaded, a regime that is inaccessible when measuring of current.

Having seen the significant impact charge detection has had on the study of quantum dots, in the following chapters we use this technique to probe a very different solid state system, hydrogenated amorphous silicon (a-Si:H). Before beginning our discussion of these experiments, it is important that we determine under what conditions the charge sensing technique actually works. In particular, for the experiments presented in Chapters 2 and 3, we use as a charge sensor a QPC fabricated from a high mobility 2DEG and cooled to very low temperatures. We now question which aspects of these experiments contribute to the high sensitivity of our charge detector.

The following arguments lead us to believe that while the high mobility material and low temperature used in our experiments are necessary conditions for the study of GaAs quantum dots (Section 2-1), they are not necessary conditions for the operation of a sensitive charge detector. Consider a field effect transistor (FET), where the distance between the gate and the channel is 100 nm. As an example, consider a particular variety of FET, a high electron mobility transistor (HEMT). A HEMT can be fabricated by patterning a metallic gate on top of the heterostructure shown in Fig. 2-1, and measuring the conductance through the 2DEG between two contacts located on either side of the gate. If we change the voltage applied to the gate by $\Delta V_g \approx 60$ mV, this is equivalent to changing the charge density on the gate by 1 electron per 50 nm square. Now, a 60 mV change in gate voltage will have a measurable effect on the conductivity of a HEMT, or on the conductivity of a wide variety of lower mobility transistors. If we reduce the size of the HEMT gate to a 50 nm square, this means that one electron added to the gate of the HEMT will have a measurable effect on its conductivity. From this argument, we expect that any FET for which a change in gate voltage $\Delta V_g \approx 60$ mV has a measurable effect should have single electron sensitivity if it is sufficiently small. Of course, we would expect devices with a stronger gate effect to be more sensitive, but the point here is that sensitivity to a ~ 60 mV change in gate voltage is not a very stringent requirement for an FET.

Next we consider whether or not temperature should strongly effect the sensitivity

of a charge detector. For this discussion, it is useful to specialize to a particular detector. We choose a QPC created by confining a 2DEG along one dimension, which we will call x. For electrons in the QPC, the contribution to the energy from motion along the x direction is quantized, and current flows along the y direction. We consider a QPC for which only the lowest energy wavefunction along the x direction is energetically accessible, and has energy E_0 relative to the Fermi level. We apply a bias V_{ds} symmetrically between two Fermi seas to which the QPC is connected at either end, and measure the current I flowing through the QPC. We write the standard result for the current flowing through this one dimensional channel [59]:

$$I = \frac{2e^2}{h} \int_{E_0}^{\infty} (f(E - eV_{ds}/2) - f(E + eV_{ds}/2))dE \qquad (3.2)$$

Here we integrate over E, the energy of an electron in the QPC, and the two Fermi functions give the electron occupation factors for the two Fermi seas. The minimum energy of an electron in the QPC is E_0, and $E - E_0$ is the kinetic energy of an electron in the QPC along y, the direction of current flow. Higher energy electrons travel faster, so one might expect them to be weighted more heavily in their contribution to I. However a decrease in the density of states with increasing E cancels exactly with this increase in velocity.

If we differentiate Eqn. 3.2 with respect to V_{ds}, we find that at zero temperature, the conductance $G(E_0) = dI/dV_{ds}$ is a step function that jumps from $2e^2/h$ to 0 at $E_0 = 0$.[7] At finite temperature, the step is not infinitely sharp but instead has a width $\sim kT$. E_0 varies linearly with the gate voltages V_g applied to the QPC, so this step can be observed by measuring the G vs. V_g characteristics of the QPC. Because the width of the step decreases with decreasing temperature, the conductance G is more sensitive to V_g the lower the temperature. However, this does not imply that the quantity that we actually measure, in this case the current I, is more sensitive to V_g as the temperature is reduced. To see this, we compute the transconductance of

---

[7]This result can be generalized to include multiple quantum wavefunctions along the x direction, from which follows the quantization of the conductance of the QPC in units of $2e^2/h$ [59].

the QPC, defined as the derivative of I with respect to V_g, or equivalently E_0:

$$\frac{\partial I}{\partial E_0} = \frac{2e^2}{h}(f(E_0 + eV_{ds}/2) - f(E_0 - eV_{ds}/2))) \tag{3.3}$$

This function decreases monotonically with V_{ds}, and no matter what the value of kT or E_0, approaches $\pm 2e^2/h$ when $|V_{ds}| >> kT$, where here the sign is opposite to the sign of V_{ds}. Therefore the maximum sensitivity of the current to changes in V_g is the same no matter what the temperature.

To understand this result, we note that although $\partial G/\partial V_g \sim G/kT$ gets larger with decreasing temperature, the maximum bias one can apply to the QPC before the I vs. V_{ds} characteristics become sub-linear is $V_{max} \sim kT$. The transconductance, which is of order the product of $\partial G/\partial V_g$ and V_{max}, is therefore temperature independent. Here we have limited our discussion to a QPC containing a single mode, but we expect that this argument should hold more generally: As temperature is decreased, the conductance of a charge sensor may become more sensitive to gate voltage, but the maximum bias one can apply in order to measure the conductance is reduced, so that the sensitivity of the current to gate voltage does not depend strongly on temperature.

From these arguments, we expect that using a QPC cooled to very low temperatures as a charge sensor is probably not necessary. In fact, the characteristics of the QPC used in our experiments departed significantly from the quantized behavior observed for ideal QPCs, as did the characteristics of QPCs used in similar experiments conducted by other groups [60], so it is clear that an ideal QPC is not necessary for a high charge sensitivity. Furthermore, in a large number of studies of narrow MOS-FETs (metal-oxide-semiconductor field-effect transistors), discrete telegraph noise, associated with single electron tunneling into and out of traps, is observed [14, 15], even at room temperature. The origin of this switching noise is within the MOSFET itself, and its presence does not guarantee that the MOSFET has a high sensitivity to charge located in a nearby structure. However, in the following chapters, we will see that we can indeed use a narrow MOSFET as a sensitive charge detector.

Chapter 4

Charge Detection Using a Narrow MOSFET

As we have seen in Chapters 2 and 3, charge detection is a powerful technique for studying electrons confined in lateral GaAs quantum dots. In this chapter we discuss technology we have developed for using an integrated charge sensor, in the form of a narrow MOSFET, to study charge transport in a highly resistive material, hydrogenated amorphous silicon (a-Si:H). In Section 4.1 we give a brief introduction to the physics of MOSFETs. In Section 4.2 we give an overview of the fabrication of our charge detector. In Section 4.3 we discuss our procedures for patterning a-Si:H adjacent to our detector. Details for both of these fabrication procedures can be found in Appendix A. In Section 4.4, we discuss the instrumentation we use for our charge sensing experiments. In Section 4.5 and 4.6 we give an overview of the electrical characteristics and noise characteristics of our narrow MOSFET detectors. In Section 4.7, we briefly compare our MOSFETs to GaAs quantum point contacts. In Chapter 5, we apply or MOSFET integrated charge sensor to the study of electron transport in a-Si:H.

4.1   Introduction: MOSFET Physics

In this section we give a very brief introduction to the physics of MOSFETs. When designing, fabricating, and measuring these devices, a much more detailed knowledge of semiconductor device fundamentals is required than what is presented here. A more thorough introduction to semiconductor devices can be found in Streetman's textbook [2] and many more details are reviewed in Sze's textbook [40]. An overview of the electronic properties of MOSFET inversion layers can be found in Ando *et al.* [61].

A sketch of an n-channel MOSFET is shown in Fig 4-1(a). The conductance of the MOSFET (G_M) is measured between the two $n+$ regions. The p-type substrate is kept grounded, and G_M is measured as a function of the voltage V_G applied to the metallic gate, which is electrically insulated from the $n+$ regions and p-type substrate by the silicon dioxide, of thickness d_{ox}. With $V_G = 0$, the p-type silicon extends all the way up to the silicon-oxide interface below the gate. Therefore G_M is very small, as the device consists effectively of two pn junction diodes with opposing polarities in series. However, if a positive voltage is applied to the gate, the conduction band at the Si/SiO$_2$ interface under the gate is pulled downward toward the Fermi level as shown in Fig 4-1(b). For smaller positive gate voltages, this has the effect of depleting the p-type carriers under the gate. The depth L_d to which this depletion region extends from the silicon-oxide interface grows as the gate voltage is made more positive as $\delta L_d N_A e C_s = \delta V_G$. Here N_A is the acceptor density and C_s is the series capacitance of the gate oxide $C_{ox} = \kappa_{ox}\epsilon_0/d_{ox}$ and depletion layer $C_{Si} = \kappa_{Si}\epsilon_0/L_d$, where here $\kappa_{ox} \approx 3.9$ and $\kappa_{Si} \approx 11.9$ are the oxide and silicon dielectric constants, respectively. However, once the gate voltage is made sufficiently positive so that the Fermi level is close to the conduction band, the depletion region width saturates, and when the voltage is made still more positive an inversion layer, a very thin (~ 5 nm [2]) layer of negative charge at the silicon-oxide interface, is formed. Because the inversion layer is n-type, it provides a conduction path between the two $n+$ regions, and G_M grows. Further increases in V_G above the threshold voltage for inversion (V_T) increase

the charge density per unit area n_{inv} in the inversion layer as $n_{inv} = C_{ox}(V_G - V_T)/e$. Expressions for threshold voltage V_T can be found in the textbooks [40]. It depends on the metal-semiconductor work function difference and the amount of trapped charge in the oxide. The conductance of the device for $V_G > V_T$ is given approximately by:

$$G_M = (w/L)e\mu n_{inv} = (w/L)\mu C_{ox}(V_G - V_T) \tag{4.1}$$

Here L and w are the width and length of the inversion layer channel, respectively, and μ is the mobility of the inversion layer. Just like the electrons at the AlGaAs/GaAs interface discussed in Chapter 2, the electrons in the inversion layer constitute a two dimensional electron gas [62, 61]. Using electron beam lithography, we can create a narrow constriction in the inversion layer to make a charge sensor, just like the QPC discussed in Chapter 2. The design of the gates for this narrow MOSFET are different than for the QPC, because for the MOSFET the gate defines where there *are electrons* in the 2DEG, where as for the QPC the gates define where there *are no electrons* in the 2DEG, but each device consists essentially of a narrow constriction in a 2DEG.

Using a narrow MOSFET as opposed to a GaAs quantum point contact as a charge sensor has advantages and disadvantages. Though the semiconductor fabrication processes utilized to make GaAs quantum point contacts is well understood [29, 63], CMOS fabrication procedures are undoubtably the most well developed processes in the semiconductor industry [64]. Furthermore, while fabrication of quantum point contacts requires a high quality AlGaAs/GaAs heterostructure, the fabrication of silicon MOSFETs requires only a silicon substrate. Utilizing state of the art CMOS fabrication facilities, one can therefore reproducibly fabricate large numbers of silicon MOSFETs. There are of course potential disadvantages to using MOSFET charge detectors. Though the development of CMOS technology was driven in part by the high electrical quality of the silicon-oxide interface, the inversion layer formed at the silicon-oxide interface generally has a far lower mobility μ than the 2DEG formed at the interface of a AlGaAs/GaAs heterostructure. For example, for our MOSFETs
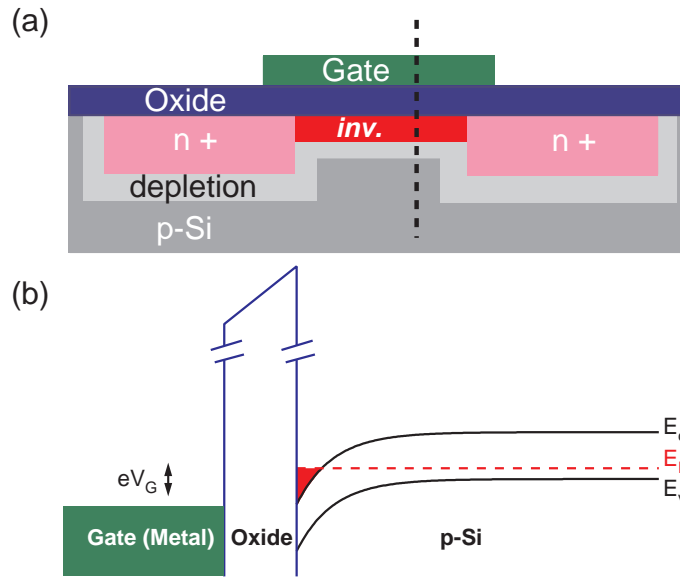
Figure 4-1: (a) Vertical sketch of an *n*-channel MOSFET, where here a positive voltage applied to the gate forms an inversion layer at the silicon-oxide interface underneath (*inv.*). The device is not drawn to scale. (b) Band diagram for the MOSFET, taken along the dashed black line in (a), showing the valence band energy (E_v), conduction band energy (E_v), and Fermi level (E_F). The red shaded region represents the inversion layer.

$\mu \approx 1200$ cm^{2}V/s at T = 4.2 K, comparable to what is typically found for n channel MOSFETs [12, 17], as compared to $\mu \approx 6.4 \times 10^5$ cm^{2}V/s, the mobility of the GaAs 2DEG used in the single electron tunneling experiments discussed in Chapters 2 and 3. However, as was discussed in Section 3.3, it is not clear that a very large mobility is necessary for a high charge sensitivity. In Section 4.7 we compare our GaAs and MOSFET charge sensors, and find that while it does seem that the GaAs sensors have a somewhat higher charge sensitivity, the difference between the two is not so dramatic, and the narrow MOSFET has a sufficiently high charge sensitivity for our purposes.

4.2 MOSFET Fabrication

In this section we discuss our procedure for fabricating narrow n-channel MOSFETs. In this discussion we will only outline the fabrication process. Details of the fabrication steps can be found in Appendix A. We start with a p-type silicon wafer, doped with Boron $N_B \approx 3 \times 10^{15}$ cm^{-3}. We grow a thick (≈ 650 nm) layer of SiO$_2$ on the wafer, using a wet thermal oxidation process.[1] This oxide serves as the field oxide for the MOSFET. We then etch holes in the SiO$_2$ using photolithography followed by an HF based etch, forming what will become the active region of the device. After another photolithography step, the wafers are sent out for phosphorous ion implantation, defining the $n+$ regions. Following ion implantation, we grow 100 nm of SiO$_2$ using a dry thermal oxidation process, which serves as our gate oxide. This step also causes the phosphorous implants to diffuse $\sim$0.5 μm into the silicon substrate. We then deposit ≈ 80 nm of $n+$ polysilicon using LPCVD (low pressure chemical vapor deposition), which serves as the gate material of our narrow MOSFET. Following the LPCVD deposition, we anneal the wafers, which increases the conductivity of the polysilicon. A sketch of the device following this step is shown in Fig. 4-2.

The next step is to nano-pattern our gate. We start by spinning a thin film of

---

[1]For wet thermal oxidation, steam is introduced into the furnace. Relative to dry thermal oxidation processes, where only oxygen is present, wet oxidation grows oxide of a lower electrically quality, but at a much faster rate [64].
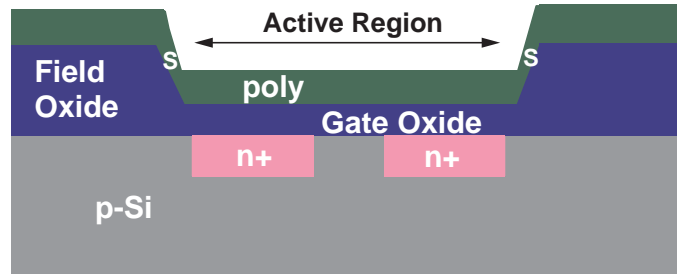
Figure 4-2: Vertical sketch of device following polysilicon deposition and anneal (not drawn to scale). The regions marked "S" are where "stringers" can form, as discussed in the main text.

HSQ [65], a negative ebeam resist, onto our wafer. We pattern this film using electron beam lithography and develop it in a TMAH based developer. After this step, the HSQ remains only where it has been exposed to the electron beam: These regions will serve as a mask for the small features of our gate. After another photolithography step to define the larger parts of the gate, we remove the exposed polysilicon using reactive ion etching with a Cl_2/HBr chemistry, thus creating the gate of the MOSFET. The etch is anisotropic, which is important because any undercut could destroy the narrowest portion of the gate. We then perform another photolithography step that covers the entire MOSFET gate in photoresist, and reactive ion etch the wafers with an HBr chemistry at low RF power. This step ensures that there are no so called "stringers," residual polysilicon left at the edge of the active region (Fig. 4-2). Since the polysilicon is vertically thicker here, it takes a longer time to remove with an anisotropic etch. Because the Cl_2/HBr etch will etch HSQ and SiO_2 somewhat, one cannot overetch too much when defining the gate. The low power HBr etch is much more selective, and so can be used to safely remove any residual polysilicon without damaging the device. Following this etch step, we grow a thin layer of oxide on the wafer using a short dry thermal oxidation process, which has the effect of growing ~ 8 nm of oxide on the polysilicon gate.[2]

Next we make electrical contact to the gate and inversion layer of the MOSFET.

---

[2]This step was intended to insulate the gate from any material that is touching it, which would in principle have allowed us to deposit the material we wished to study with our integrated charge sensor right on top of the gate. We found however that this thin oxide leaked a little bit, so that this intended functionality was not realized.

First, we cut holes in the gate oxide above the $n+$ regions and in the thin oxide on top of the polysilicon gate, using a photolithography step followed by an HF based etch. We then sputter 350 nm of aluminum over the entire sample, and pattern the aluminum with photolithography followed by a PAN wet etch. The aluminum contacts are then sintered in forming gas.

The narrow MOSFET is now complete. Before dicing up the wafer, we add a pair of gold contacts, which serve as contacts for the resistive material we wish to probe with our charge sensor. This is accomplished by a photolithography step using a negative resist, followed by electron beam evaporation of Ti/Au and lift off in acetone. An electron micrograph and optical micrograph of the resulting structure are shown in Fig. 4-3. Following lift-off, we spin a protective layer of photoresist on the wafer, and cut out individual devices using a diesaw. We then mount, bond, and test some of the narrow MOSFETs to obtain their electrical characteristics. To perform our charge sensing experiment, we need to pattern an amorphous material in close proximity to the narrowest portion of the MOSFET. We tried patterning both doped hydrogenated amorphous silicon (a-Si:H) and amorphous germanium (aGe) adjacent to our charge sensor. Because the measurements discussed in the following chapter were performed with an a-Si:H device, we describe the fabrication procedure for patterning a-Si:H adjacent to our MOSFET in the following section.

4.3 Patterning Hydrogenated Amorphous Silicon

The electronic properties of a-Si:H are introduce in detail in Chapter 5. Here we discuss only the practical issues involved in depositing and patterning these films. A step by step fabrication procedure for patterning a-Si:H is given in Appendix A.2.

At a first glance, patterning hydrogenated amorphous silicon (a-Si:H) adjacent to our narrow MOSFET seems straightforward, using electron beam lithography combined with either reactive ion etching or a lift-off procedure. There are, however, a number of issues which make it a challenging task. Reactive ion etching is not appropriate for our process, because one would need to reactive ion etch a-Si:H with a
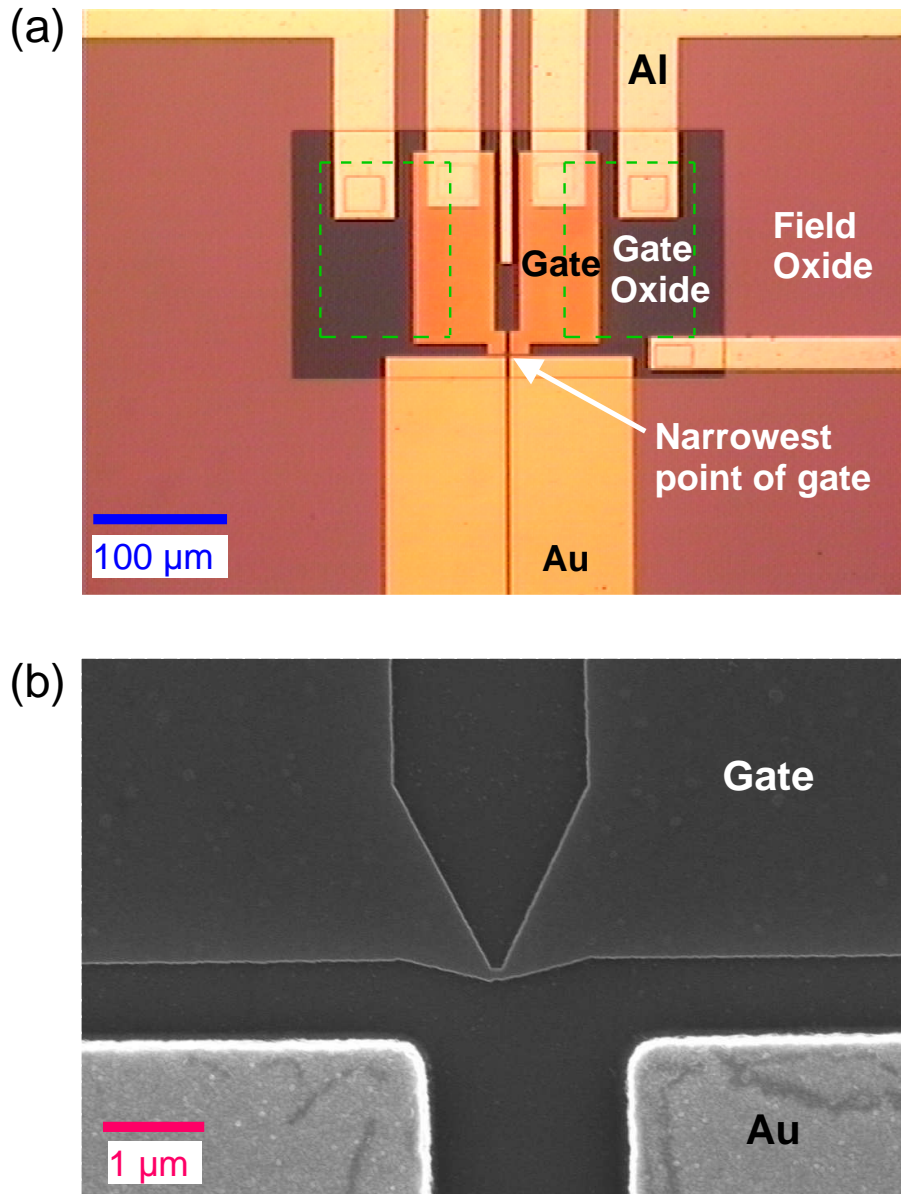
Figure 4-3: (a) Optical micrograph of device following lift-off. The dashed green squares indicate the $n+$ regions (where the silicon substrate is phosphorous doped). Electrical contact is made to the inversion layer and gate with the aluminum contacts through holes etched in the oxide. The aluminum contact coming from the right side of the figure contacts the p-substrate, but we instead made contact to the p-substrate through the back of the chip and left this contact floated as discussed in the text. The central aluminum contact of the five extending from the top of the figure is kept grounded, and does not affect any of the experiments discussed in this thesis. (b) An electron micrograph of the narrowest portion of the MOSFET gate and two gold contacts.

very high selectivity to all of the exposed materials on the surface of the MOSFET structure shown in Fig. 4-3, in order not to damage the MOSFET with the etch. As this would be a daunting task, we chose instead to develop a lift-off process.

In order to deposit a-Si:H films of decent quality, one must use plasma enhanced chemical vapor deposition (PEDVD) of silane [26], as sputtered or electron beam evaporated samples typically have much higher defect densities. Furthermore, the deposition must be performed with the substrate heated, ideally with 200 C $< T_s <$ 300 C, where here T_s is the substrate temperature. The heat is required in order to promote reconstruction of the silicon-hydrogen bonding network during growth, which reduces the defect density in the resulting film [26]. The elevated T_s required for a-Si:H deposition is problematic because the glass transition temperature of PMMA, the most commonly used positive ebeam resist, is $T_g \approx$ 120 C. We found that films of doped hydrogenated amorphous silicon deposited below this temperature, at temperatures $T_s \approx$ 100 C, had much lower room temperature conductivities ($\sim 10^{-8}$ $\Omega \cdot$cm) than films deposited under the same conditions at higher temperatures ($\sim 10^{-4}$ $\Omega \cdot$cm). Presumably, this is because the films deposited at lower temperatures have such a high defect density that the doping efficiency is becomes very low, as all of the electrons added by the donor atoms are trapped in defect states (Chapter 5). We tried various annealing procedures [66] at temperatures greater than 200 C to increase the conductivity of films deposited at $T_s \approx$ 100 C, but were unsuccessful. In contrast to the a-Si:H films we deposited at low temperatures, the films we deposited at higher temperatures had electron transport properties similar to those commonly reported in the literature. For instance, for films deposited at $T_s \approx$ 180 C, it was trivial to dope the a-Si:H by adding phosphine gas to the deposition chamber [67]. Our results are shown in Fig. 4-4.

In order to pattern films deposited at $T_s \approx$ 200 C, we used an alternative electron beam resist, PMGI. This resist, produced by Mirochem and typically used for various lift-off processes, has a high glass transition temperature $T_g \approx$ 190 C. Furthermore, it has recently been shown to be sensitive to electron beam exposure, using either solvent [68] or basic developers [69]. We chose the former, as we found that the base TMAH
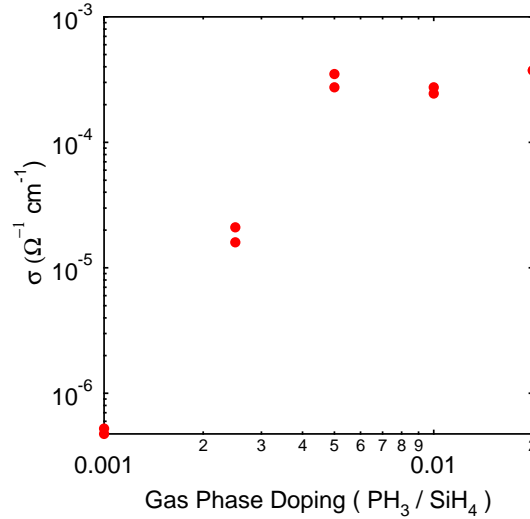
Figure 4-4: Doping effect for ≈ 100 nm thick a-Si:H films deposited at $T_s \approx 180$ C. For each film, the silane flow rate is 80 Sccm, the pressure is 300 mtorr, the RF power is 20 W (resulting in a deposition rate of ≈ 10 nm/minute), and the deposition time is 10 minutes. For each film, a variable flow rate of a mixture of phosphine and hydrogen gas ($PH_3/H_2 = 0.02$) is added to the deposition chamber, and the room temperature conductivity of the resulting films are plotted as a function of their phosphine to silane gas phase doping ratio, PH_3/SiH_4.

in many developers gradually corroded the metalization of our MOSFET. We found that increasing T_s to 200 C did not seem to affect the PMGI despite being slightly above its nominal glass transition temperature, so we performed some depositions at 200 C.

Having found a solution to the problem of having to use an elevated deposition temperature, the next challenge is to implement a lift off step for a film that has been deposited isotropically, as are films deposited with PECVD. Because of the isotropic deposition process, the lift-off must be aided with ultrasonic agitation, which serves to tear off pieces of a-Si:H that will inevitably be stuck to the edges of the pattern[3]. However, if the film is too thick, the ultrasound will simply rip the pattern off the substrate rather than tear the film at the edges of the pattern. Therefore, one must deposit fairly thin films for this lift-off process to work. We used a film thickness

---

[3]This agitation is normally not necessary for lift off processes using, for instance, electron beam evaporation, because the the anisotropy of the deposition process combined with a slight under-cut in the resist profile produces a break in the film at the edge of the pattern.
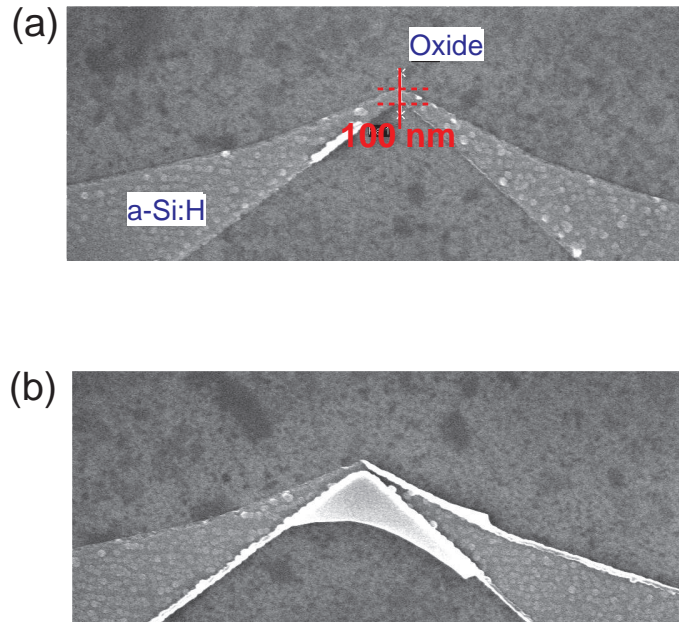
Figure 4-5: (a) a-Si:H test pattern. The pattern tapers down to a minimum width of 100 nm (b) Test pattern with a defect

of $\approx$ 50 nm, below which we found the film conductivity drops rapidly, presumably because of surface effects [70]. The pattern will also be ripped off the substrate if there are any adhesion problems, which can arise if there is any organic material on the substrate surface prior to deposition. To ensure that this does not happen, the sample is cleaned in O_2 plasma immediately before the film is deposited.

After the PECVD deposition the lift-off procedure is completed as follows: We transfer the chip into a heated bath of Microchem's PG resist remover. The chip is left for about 1 day in the remover. At this stage, viewed under a microscope, the film will look puffy, as there is liquid between the film and the substrate. The beaker of PG remover is then transfered to an ultrasonic bath. It is important not to take the chip out of the remover and let it dry during this process. The ultrasonic agitation of the resist remover underneath the film is what tears it from the pattern most efficiently. If the chip is allowed to dry, the remover flows out from underneath the film, making it very difficult to remove.

Some a-Si:H test patterns fabricated in this manner are shown in Fig. 4-5. As is shown in the figure, for some devices there are defects consisting of extra flakes
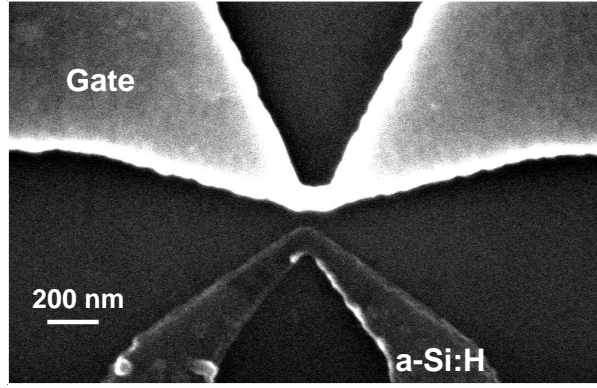
Figure 4-6: Electron micrograph of MOSFET gate and a-Si:H strip. Either side of the a-Si:H strip overlays a gold contact not shown in this micrograph. The narrowest portions of the MOSFET gate and a-Si:H strip are separated by $\approx$ 70 nm.

of a-Si:H that are not torn off of the pattern, but by and large the process works well for our purposes, and can be used to pattern a-Si:H with feature sizes $\sim$ 100 nm. When patterning the a-Si:H adjacent to a charge sensor, it is important to have very good alignment between the gate of the MOSFET and the a-Si:H pattern. The Raith is perfectly capable of overlaying patterns with a tolerance < 50 nm, but the procedure required to actually accomplish this is somewhat involved, and is discussed in Appendix A. Using our a-Si:H patterning procedure, we pattern a strip of a-Si:H adjacent to a MOSFET charge sensor. An electron micrograph of such a structure is shown in Fig 4-6. The narrowest portion of the strip is located $\approx$ 70 nm from the narrowest portion of the MOSFET. The application of this structure to the study of electron transport in a-Si:H will be discussed in Chapter 5: In the following section, we discuss the instrumentation used to measure these devices.

4.4   Instrumentation

Following the fabrication sequence described in the Section 4.2, we clean the device in acetone and methanol, glue it to a chip carrier with silver epoxy, and wire bond to it. To contact the p-type substrate, we wire bond to the back of the chip carrier, which contacts the substrate through the conductive silver epoxy. We then load the device

70

into a Janis He^4 flow cryostat. For this cryostat, the sample space is a sealed cylinder filled with He^4 exchange gas. A small flow of Liquid helium from a pressurized dewar goes through a constriction, and the resulting gas flows out to atmosphere through an annulus that is concentric with with the sample space. This cold gas cools the walls of the sample space, which cools the sample through the exchange gas. Two silicon diodes, one mounted near the constriction, and the other mounted next to the sample, are used to monitor the temperature. The temperature is controlled using a resistive heater near the constriction and a Lakeshore 331 Temperature Controller. The sample and constriction thermometers are generally in agreement, the sample thermometer of course taking longer to respond to the heater than the constriction thermometer. As they do however sometimes sit at slightly different temperatures, we quote the sample thermometer in all temperature dependencies discussed here. Using this cryostat, and by pumping on the He^4 exhaust to cool the sample below 4.2 K, we can stabilize the temperature of the sample anywhere between 300 K and 2.4 K, and the time required to go from one extreme of the temperature range to the other is only ~ 20 minutes. The cryostat is retrofitted with 12 coaxial cables [4]. All electrical contact to the sample is made through these cables, which ensures that there is very little cross-capacitance between the electrical connections to the sample. This is important when implementing the pulse sequences and fast read-outs that will be discussed the following chapter.

The circuit used for measuring our a-Si:H charge sensing device is shown in Fig. 4-7. A positive voltage V_G is applied to both of the MOSFET gate contacts, forming an inversion layer underneath. The MOSFET (inversion layer) resistance is measured by applying a small bias voltage V_b to one of the $n+$ contacts, and measuring the current flowing from the other $n+$ contact using a Femto 400 kHz bandwidth amplifier. This small amplifier, as well as the voltage dividers/filters for the bias voltage, is mounted only a few feet away from the coaxial connections at the top of the flow cryostat in order to minimize the size of the measurement circuit loop, and thus minimize the noise, as well as to reduce the shunt capacitance at the current amplifier input

---

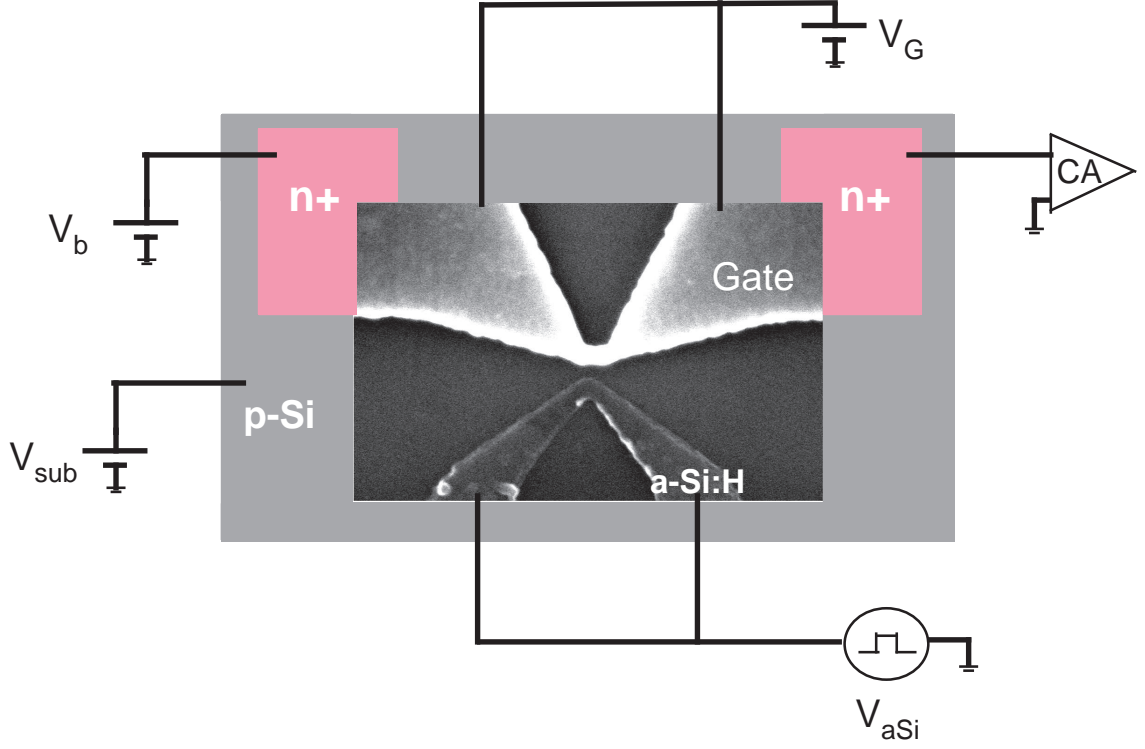[4]This retrofitting was done by Tamar Mentzel

Figure 4-7: Measurement set up

for reasons that will be discussed below. With our circuit, the measurement of the MOSFET conductance can be performed using either an AC or DC excitation. For high speed measurements of the MOSFET resistance, we used a DC excitation. For a DC voltage bias, we supply/monitor a voltage using NI's 6703/6251 output/input voltage cards (this voltage is divided and filtered next to the current amp as mentioned above). The output of the current amplifier, after being further amplified and filtered if necessary with an Ithaco 1201 voltage pre-amplifier, is fed into a high speed voltage card, the NI-6110, which can read in time series with a minimum time step of 200 ns. For AC measurements, we add an AC excitation to the bias voltage through a transformer, and the output of the current amplifier is monitored with an Stanford Research 830 lock-in amplifier.

For the pulsing measurements described in Chapter 5 we needed to pulse the voltage V_{aSi} (Fig. 4-7) applied to the a-Si:H contacts. To accomplish this, we combined the output of the 6110 voltage card with the voltage supplied by a Yokagawa 7651
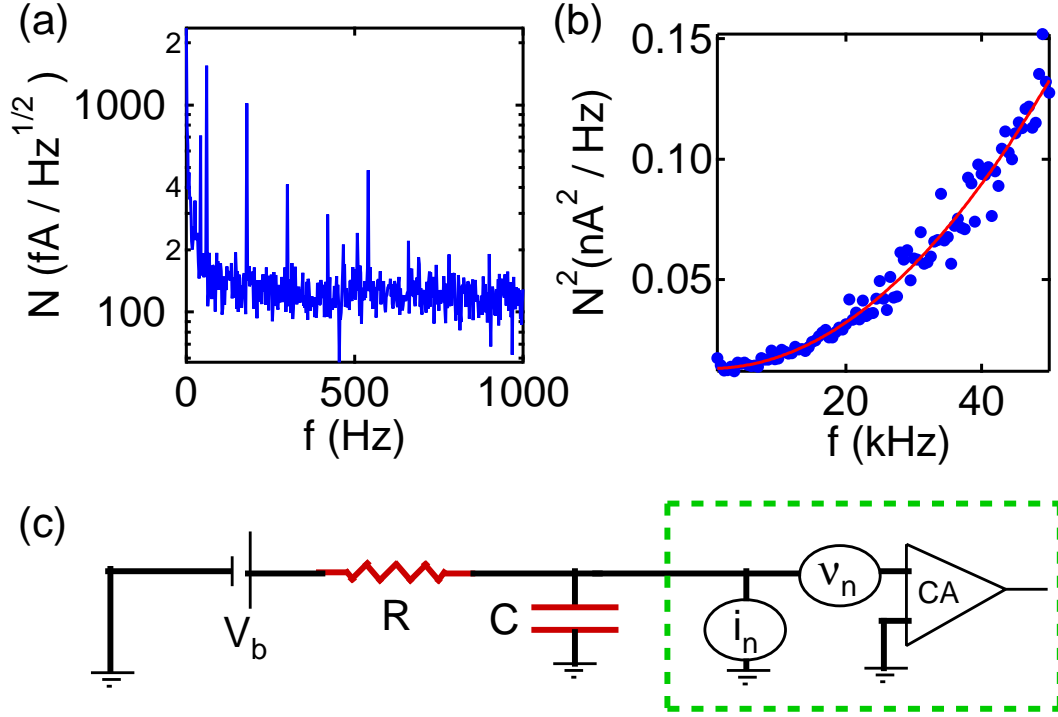
Figure 4-8: Noise Characterization (a) Low frequency noise (b) Increase in noise at higher frequencies caused by voltage noise at the current amplifier input driving currents through the shunt capacitance, as described in the main text. The solid line is a theoretical fit described in the main text. (c) Standard model for amplifier noise. The noise sources i_n and v_n in the dashed green box are intrinsic to the current amplifier.

voltage source with a $\pm$ 30 V range. This allowed us to add small high speed voltage pulses ($\approx$ 5 μs rise time, after low pass-filtering) to a large offset. The voltage applied to the p-type substrate and MOSFET gate were supplied by either a Yokagawa or voltage card, divided and filtered as necessary.

For high speed measurements of the MOSFET conductance, we need to measure the current from our MOSFET over a large bandwidth, and it is therefore important to understand sources of current noise in our measurement set-up. To accomplish this we replace the MOSFET in Fig. 4-7 with an 80 kΩ resistor, cooled to 4.2 K to eliminate Johnson noise. We monitor the current as a function of time, and digitally fast Fourier transform this time series to obtain the noise spectrum. The results are shown in Fig. 4-8. For $f < 1$ kHz (Fig. 4-8(a)), the current noise $N(f)$ is

approximately constant at ≈ 120 fA/Hz$^{1/2}$, with a few spikes at various frequencies caused by coherent noise sources (at 60 Hz, for instance). At higher frequencies $f \sim 30$ kHz (Fig. 4-8(b)), N grows with f: Specifically, N^2 grows quadratically with f. This effect can be understood in terms of the standard model for noise in current amplifiers [71], a cartoon for which is shown in Fig. 4-7(c). The current amplifier is a source of both current noise i_n and voltage noise v_n. The current noise i_n always flows into the amplifier CA, and is therefore always present on the amplifier output, where it is measured (multiplied by the gain). The voltage noise, on the other hand, drives a current into to amplifier CA that depends inversely on the impedance Z to ground of whatever is connected to the current amplifier input. In our case, this impedance is given by the resistor $R = 80$ kΩ, in parallel with the shunt capacitance C of the coaxial cable between the current amplifier input and the resistor. We have therefore $1/Z = j2\pi fC + 1/R$. Provided i_n and v_n are incoherent and thus add in quadrature, we have then for the total noise current N flowing into the amplifier input:

$$N^2 = i_n^2 + (v_n/R)^2 + (v_n 2\pi fC)^2 \tag{4.2}$$

Thus we have that N^2 grows quadratically with f, as more current driven by v_n flows through the shunt capacitance and into the amplifier input at higher frequencies. We fit the data in 4-8(b) to Eqn. 4.2, and, estimating $C \approx 150$ pF from the length and specifications of the cables between the resistor and the current amplifier, obtain $v_n \approx 7$ nV/Hz$^{1/2}$ and $i_n \approx 70$ fA/Hz$^{1/2}$, which are roughly consistent with the manufacturers specifications of 5 nV/Hz$^{1/2}$ and 65 fA/Hz$^{1/2}$, respectively. As mentioned above, we minimized the shunt capacitance C, by mounting the current amplifier very close to the cryostat, in order to reduce the current noise at higher frequencies. These measurements fully characterize the noise background of our instrumentation, the telegraph noise inherent to our MOSFET charge sensors themselves is investigated in Section 4.6.

4.5   Electrical Characteristics

In this section, we give an overview of the electrical characteristics of our MOSFET charge sensors. In general, our narrow MOSFETs behaved similarly to those studied in previous reports [12, 14], so in this section, we will focus on the aspects of their performance that are most pertinent to their use as charge sensors. The following chapter will discuss the application of our MOSFET charge sensors to the study of charge transport in amorphous semiconductors.

Typically, as soon as a narrow device was loaded into our cryostat, we checked to make sure that the heavily doped polysilicon gate was continuous, in particular because we judged that this was the most likely part of the device to be damaged by electrostatic discharge during sample preparation. This was accomplished by measuring the resistance between to two contacts to the polysilicon gate shown in Fig. 4-3. This resistance was typically ≈ 7 kΩ. Using the polysilicon film resistivity $\rho \approx$ 1.5 m$\Omega\cdot$ cm, measured with a four point probe, and the number of squares in series calculated for our gate geometry, we would expect the gate resistance to be a bit smaller ≈ 2 kΩ, but the result is fairly close to what we expect. We found that, taking reasonable precautions, the MOSFET gate was fairly robust against electrostatic discharge[5]

On each wafer, in addition to the narrow MOSFET charge sensors, we fabricated simultaneously a variety of test structures, including structures for measuring the contact resistance between the aluminum electrodes and the phosphorous implanted regions as well as the polysilicon gate, diodes, and transistors with wide (10 μm) gates. For a few wafers, we did not successfully make contact to the phosphorous implanted regions, a result which may have been caused by surface damage of the silicon substrate during the ion implantation process (Appendix A), but all of the devices studied here came from wafers with negligible contact resistances.

In Fig. 4-9, we show the G_M vs. V_G (MOSFET conductance vs. gate voltage)

---

[5]We tried depositing films of PbS nanocrystals on some of our devices in a nitrogen glove box. In this very dry environment, we found that we had to be much more careful with the grounding of the device, or the narrow gate was damaged.
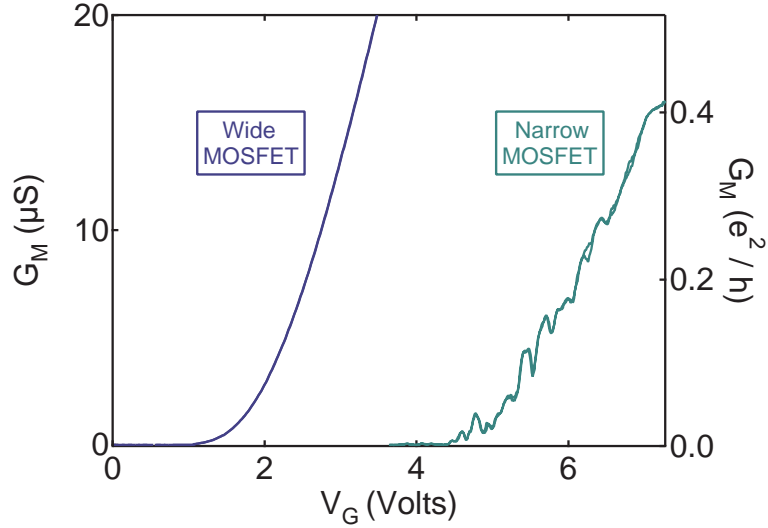
Figure 4-9: Conductance vs. gate voltage characteristics of wide and narrow MOS-FETs. Note that for the narrow MOSFET, the V_G sweeps in both the positive and negative direction are both shown, and are nearly indistinguishable

characteristics of a wide and narrow channel MOSFET from the same wafer at T = 4.2 K. For the wide MOSFET, from the slope of the G_M vs. V_G curve, along with the capacitance of the gate oxide and length and width of the gate and thus inversion layer, we extract the 4.2 K mobility of our inversion layer $\mu \approx 1200 \text{ cm}^2\text{V/s}$ using Eqn. 4.1, a typical value for an n-channel MOSFET [12, 17]. For the narrow MOSFET, we see two obvious changes in the G_M vs. V_G curve from the wide gate case. The first is that the narrow MOSFET turns on at a larger voltage than does the wide MOSFET. The increase in threshold voltage is caused by the fact that the gate tapers down to a width (≈ 70 nm) that is less than the thickness of the gate oxide ($d_{ox} = 100$ nm). At its narrowest point therefore, the gate capacitance is less than the parallel plate value $C = \kappa_{ox}\epsilon_0/d_{ox}$, so that a larger voltage must be applied to create the inversion layer [72]. The second change is that for the narrow MOSFET G_M oscillates with V_G. Note that traces for sweeping V_G both in the positive and negative direction are both shown on the graph: They are nearly indistinguishable, indicating that these oscillations are highly reproducible. These oscillations vanish as the temperature is increased to $T \sim$ 20 K. Conductance oscillations in narrow MOSFETs have been studied extensively [73, 72]. They are caused by localization effects induced by disorder. When its

conductance is made $< e^2/h$, the 1 dimensional inversion layer is effectively made into a series of quantum dots by random fluctuations in the local potential, so that the G_M vs V_G characteristics reveal complicated oscillations. Overall, the slope of the G_M vs. V_g curve for the narrow MOSFET is only a factor of ≈ 2 less than that of the wide MOSFET. This means that the number of squares in series N_{sq} for the Narrow MOSFET inversion layer is only about twice as large as that of the wide gate inversion layer, from which we obtain for our Narrow MOSFET $N_{sq} \approx 5$. This number agrees roughly with the number of squares in series of our narrow MOSFET gate pattern, which makes sense. Our gate pattern was intentionally designed to have a relatively low value for N_{sq}, which is accomplished by having the gate taper rapidly to its narrowest point, as can be seen in Fig. 4-3. Because of the low value for N_{sq}, the narrowest part of the inversion layer contributes substantially to its resistance, so that changes in the local electrostatic environment of the narrowest part of the transistor create measurable changes in G_M.

One problem we encountered with our MOSFETs was parasitic surface leakage. We will describe this effect in some detail here, because it is a problem that, as will be discussed below, is somewhat specific to our device geometry, and so is not frequently mentioned in the literature. To characterize this effect, we studied devices where the gate was discontinuous, so that without any surface leakage, the transistor should always be off. For these devices, we found that there was a substantial room temperature conductivity (~ 100 kΩ) between the $n+$ contacts. This leakage was not present at low temperatures, and could also be eliminated by applying a negative voltage bias V_{sub} to the p-type substrate relative to the $n+$ contacts. This effect can be understood in terms of surface leakage as follows. The SiO$_2$ of a MOSFET typically has some trapped charge per unit area within it n_{ox}. The charge may have a number of components, for instance mobile ionic charge from alkali impurities or other defects, but the sign of the net charge is typically positive [40]. For an n-channel device, this positive charge lowers the threshold for inversion [40], and if there is enough of it, can create an inversion layer at the silicon-oxide interface even where there is no gate. In order for this to happen, n_{ox} must be at least as large as the amount of charge in the

depletion layer of the transistor in inversion: $n_{ox} > n_{min} = N_B L_{inv}$, where N_B is the Boron acceptor density and L_{inv} is the width of the depletion layer at inversion, given by $L_{inv} = 2\sqrt{\epsilon_0 \kappa_{ox} kT ln(N_B/N_i)/(e^2 N_B)}$, where N_i is the intrinsic carrier density (number of charges activated from the valence band to the conduction band) of silicon [40]. We therefore see that $n_{min} \propto \sqrt{N_B}$ (neglecting the weak logarithmic dependence on N_B), so the higher the doping level in the p-type substrate, the higher the charge required to form an inversion layer. Because our substrate doping level was low ($N_B \approx 3 \times 10^{15}$ cm^{-3}), for our wafers $n_{min} \approx 1.5 \times 10^{11}$ cm^{-2} which is not an unusual amount of trapped charge for oxide grown under typical conditions [72], and so for our devices we had parasitic surface leakage.[6]

Fortunately, the parasitic surface leakage was easily eliminated by applying a reverse bias to the substrate $V_{sub} \approx$ -3 V. Applying a reverse bias makes the depletion region width L_D larger than L_{inv} which increases n_{min} [40], and thus eliminates the parasitic conduction path. To verify this, we applied a reverse substrate bias to a device with a discontinuous gate, and measured a very low room temperature conductance between the $n+$ contacts. The parasitic conduction path was also not present at low temperatures, as evidently for the parasitic inversion paths in our devices the Fermi level was below the conduction band, so that the carriers could be frozen out. All of the data shown in this thesis were taken in a regime where the surface conduction was negligible.

At a first glance, it seems unusual that the problem of parasitic surface leakage is not frequently encountered with commercial MOSFET devices. The reason it is not is that for a typical MOSFET geometry, there are straightforward isolation methods that ensure there is no parasitic surface leakage. Because the gate extends all of the way across the active region, dividing it in half, one only needs to eliminate parasitic surface conduction under the field oxide. This can easily be accomplished using a LOCOS (Local Oxidation of Silicon) process, where the silicon under the field oxide is heavily doped $p+$, or one of the more modern trench isolation techniques

---

[6]We tried to measure n_{ox} directly for oxides grown under the same conditions as we grew our field and gate oxides using a surface charge analyzer [74], and the results seemed to indicate that indeed for our oxides $n_{ox} > n_{min}$ though further work would be required to verify this more precisely.

[64]. Trapped charge in the gate oxide can shift the threshold of the transistor, but it cannot form a parallel conduction path. For our devices however, the gate does not divide the active region in half (Fig. 4-3(a)), so that parasitic conduction paths under the gate oxide become possible. The only way to eliminate these in the fabrication sequence, other than carefully analyzing the sources of trapped charge in the oxide and eliminating them, is to start with a more heavily doped wafer, which is somewhat undesirable because this will tend to increase screening from the substrate and therefore reduce the charge sensitivity of the device (Chapter 5). In any case, using a negative substrate bias allowed us to circumvent having to fix this problem.

4.6   Noise Properties

Having seen that the electrical characteristics of our transistors are what we expect, we now give an overview of the MOSFET noise properties. Ralls and collaborators first observed telegraph noise in the conductance of a nanometer scale MOSFET, which they attributed to electrostatic fluctuations involving a single electron charge. [14]. The $1/f$ noise observed in larger devices was attributed to a large ensemble of these fluctuators. Many others have observed similar effects since this report [15], and, as discussed in Chapter 1, have used this phenomenon to investigate a variety of interesting physical effects [17, 18].

The standard model for this phenomenon is very similar to the model for the telegraph noise we studied in Chapter 2, and is presented in Fig. 4-10(a). By applying a positive voltage V_G to the metallic gate of the MOSFET, an inversion layer is formed at the silicon-oxide interface. We assume there is an electron trap in the oxide located some distance x_t away from the silicon-oxide interface. When the energy of the trap E_t is above the Fermi level as is shown in the figure, the trap is empty. However, if the gate voltage is made slightly more positive, E_t will be brought below the Fermi level, and the trap will fill with an electron, which will decrease the conductance through the inversion layer G_M. If the gate voltage is adjusted so that E_t is brought to within kT of the Fermi level, the electron will spontaneously hop on and off of the trap,
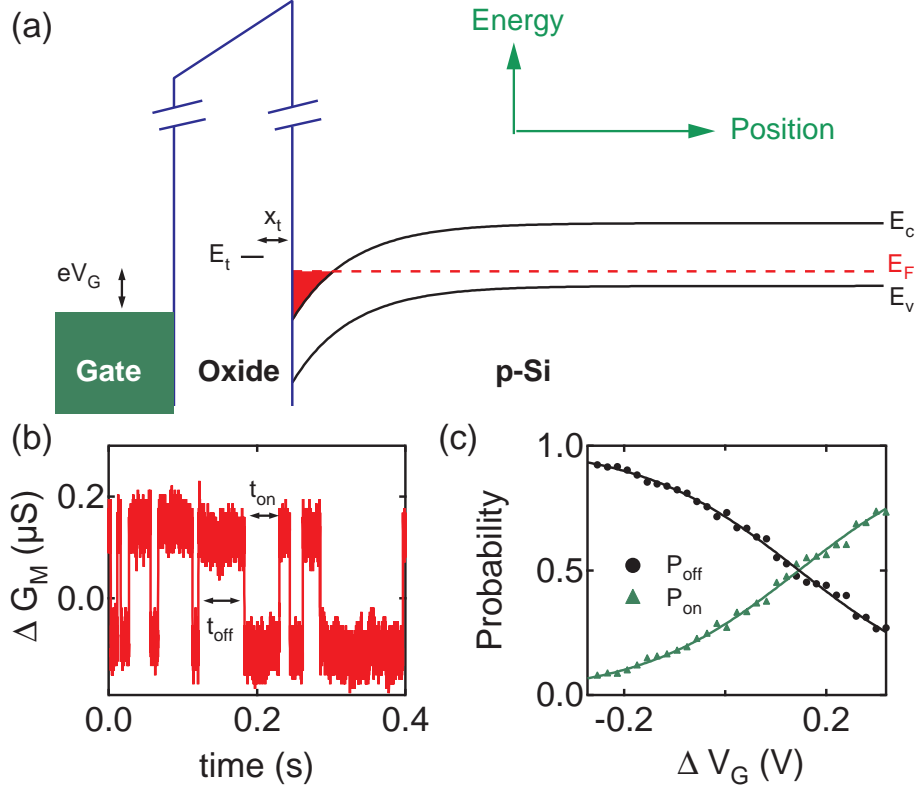
Figure 4-10: (a) Sketch of the standard model for telegraph noise in MOSFETs. The red shaded region represents the inversion layer formed at the silicon-oxide interface. (b) Telegraph noise observed for one of our devices at T = 11 K with a transistor bias $V_b \approx 1$ mV (c) Measurement of P_{on} and P_{off} as a function of ΔV_G obtained from measurements of telegraph noise as in (b) at T = 11 K with a transistor bias $V_b \approx 1$ mV. The solid lines are theoretical fits described in the main text. The range of ΔV_G at which P_{on} and P_{off} can be measured is limited by variations in the telegraph switch signal size as well as other switches which appear as ΔV_G is varied and interfere with the measurement.

producing telegraph noise. An example of such telegraph noise measured for one of our devices is shown in Fig. 4-10(b). This noise can be analyzed using exactly the same techniques discussed in Chapter 2. Using our automated triggering software, we measure a large number of intervals t_{on} (t_{off}) for which an electron is on (off) of the trap (Fig. 4-10(b)), and from these intervals determine the rates Γ_{off} (Γ_{on}). From these rates we can determine the probability that an electron is on or off of the trap $p_{off} = \Gamma_{off}/(\Gamma_{off} + \Gamma_{on})$, $p_{on} = \Gamma_{on}/(\Gamma_{off} + \Gamma_{on})$. These quantities are plotted in Fig. 4-10(c). As one would expect, the probabilities obey Fermi statistics: If we set $E = E_t - E_F$ then we have $P_{on}(E) = f(E)$, where here $f(E) = 1/(e^{E/kT} + 1)$ is a Fermi function, so we can write $P_{on}(V_G) = f(-e\alpha\Delta V_G)$, $P_{off}(V_G) = 1 - f(-e\alpha\Delta V_G)$, where α is the conversion factor gate voltage and energy (Here ΔV_G is the gate voltage measured relative to an arbitrary offset) The solid lines in Fig. 4-10(c) are fits to these equations.

The conversion factor α is determined by x_t, the distance between the trap and the inversion layer [14]. The difference between the trap energy E_t and the the conduction band energy E_c at the silicon-oxide interface is given by the product of the electric field in the oxide and the trap depth, $E_t - E_c = -ex_t F = -ex_t V_G/d_{ox}$. To determine $\partial E/\partial V_G = \partial(E_t - E_F)/\partial V_G$, we note that $E_F - E_c$ is proportional to the inversion layer electron density which is proportional to the gate oxide capacitance and the gate voltage, from which we obtain:

$$\frac{\partial(E_F - E_c)}{\partial V_G} = \frac{2\pi C\hbar^2}{e\varphi m_{eff}} \approx 1.3 \text{ meV/Volt} \tag{4.3}$$

Here φ and m_{eff} are the degeneracy and effective mass, respectively, for electrons in the inversion layer. We can therefore write for α:

$$\alpha = x_t/d_{ox} + 1.3 \times 10^{-3} \tag{4.4}$$

From the fits shown in Fig. 4-10(c) we can determine α from which we obtain $x_t \approx$ 0.5 nm for this particular trap, which falls within the range of values measured using the same method by Ralls and coworkers [14].
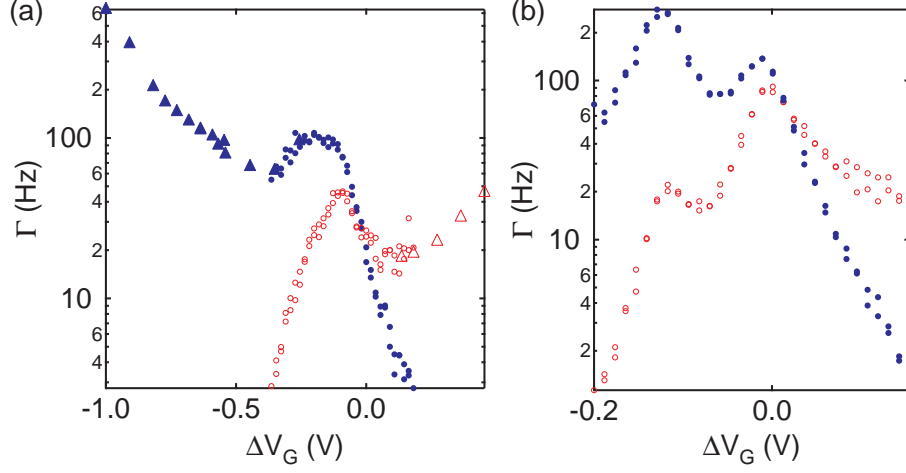
81

Figure 4-11: (a) Tunneling Lineshape for a single MOSFET trap measured at T = 2.3 K, with a MOSFET bias voltage $V_b \approx 5$ mV. Red (blue) circles are Γ_{on} (Γ_{off}) measured from spontaneous tunneling events. Red (blue) triangles are Γ_{on} (Γ_{off}) measured using a pulsed gate technique. (b) The same trap studied in (a) at T = 2.3 K, but using a smaller voltage bias $V_b \approx 1$ mV. Red (blue) circles are again Γ_{on} (Γ_{off}) measured from spontaneous tunneling events.

Given the similarities between the standard model for telegraph noise in MOSFETs and the results discussed in Chapter 2, one might be tempted to apply our model for energy dependent tunneling to telegraph noise in MOSFETs. However, the literature suggests that the tunneling process for MOSFET traps is somewhat more complicated than for GaAs quantum dots. Ralls and collaborators reported that the tunneling process has an activated component [14]. Specifically, if we write $\Gamma_{on}(E) = \Gamma_0 f(E)$, it follows from their results that Γ_0 varies with temperature, with an activation energy ~ 1 meV [14]. One might think that this activated behavior corresponds to hopping over the oxide tunnel barrier, and that the activation energy corresponds to the oxide barrier height. However, for a barrier of height ~ 1 meV and width $x_t \approx 1$ nm, one finds that a direct tunneling process would be much faster than the tunneling rates observed in the experiments, and furthermore this energy scale would be surprisingly small for the energy of the lowest unoccupied molecular orbital in the oxide relative to the Fermi level in the inversion layer. It has therefore been suggested that the observed activation energy is associated with a distortion in the SiO$_2$ bonding configuration near the trap that occurs when the trap is occupied

[14]. Given the more complicated nature of the MOSFET trap tunneling process, one would not expect to be able to describe the tunneling lineshapes measured for MOSFET traps using the same model developed in Chapter 2.

Despite these reservations, we have measured the tunneling lineshape for a MOS-FET trap, and the results are shown in Fig 4-11(a). These results are obtained using the exact same techniques used to obtain tunneling lineshapes for GaAs dots in Chapter 2: When E_t is far from E_F, we use a pulsed gate technique to measure the tunnel rates, and when E_t is close to E_F we measure the rates from spontaneous tunneling events. This tunneling lineshape is clearly more complicated than the ones shown and discussed in Chapter 2 (Fig. 2-7). We see that Γ_{on} and Γ_{off} both have a peak near $\Delta V_G = 0$. Sufficiently far from $\Delta V_G = 0$, Γ_{on} (Γ_{off}) gets faster as ΔV_G is made more positive (negative).

The data shown in Fig. 4-11(a) are taken with a large MOSFET bias, which heats the inversion layer electrons a bit above the cryostat temperature. To resolve finer features in the tunneling lineshape, we apply a smaller bias, allowing the electrons to achieve a colder temperature, and measure the rates from spontaneous tunneling events (Fig. 4-11(b)). We see now two clear peaks in the tunneling lineshape. Cobden and collaborators observe similar peaks in Γ_{on} and Γ_{off} as a function of gate voltage [75]. They ascribe this phenomena to a two-state system located near the silicon-oxide interface: The peak in the tunneling rates occurs when the two states are aligned. Multiple peak structures have also been observed and have been explained by postulating that the two-state system is electrostatically coupled to a separate very fast switch. It is possible that in Fig. 4-11 we are observing the same effect. The fact that the rate begins to grow once ΔV_G is made sufficiently positive or negative may be caused by inelastic processes. However, Cobden and collaborators observed two-state system behavior only for a fraction of traps in electrically stressed devices, where as we observed aspects of the anomalous behavior shown in Fig 4-11 for all of the traps we have studied in detail. It may be that this behavior is instead caused by the localization of electrons in the inversion layer as discussed in Section 4.5, which could lead to peaks in the local density of states in the inversion layer into which an

electron in an oxide trap must tunnel.

4.7 GaAs QPCs and Narrow MOSFETs: A Brief Comparison

It is difficult to determine, quantitatively, whether GaAs or silicon based charge sensors are better. The performance of a charge sensor depends critically upon the nature of background charge noise within the sensor itself. In MOSFETs, as discussed in the previous section, this noise is caused by the charging of traps in the SiO_2 near the silicon-oxide interface. For GaAs/AlGaAs structures, quantitative studies seeking to explain the microscopic origin of switching noise are quite recent, and the origin of the noise is much less well understood [76]. In any case, the frequency spectrum and amplitude of charge noise for either of these two types of sensors can vary substantially with temperature and the voltage applied to the gate(s), and nominally identical devices measured under similar conditions can have quite different noise spectra. We will not provide a thorough comparison between the background charge noise properties of GaAs and MOSFET charge sensors.

Here we will compare the sensitivity of the MOSFET and GaAs charge sensors, as deduced from their current vs. gate voltage characteristics. Our results are shown in Fig. 4-12. There we show measurements of current vs. gate voltage for a narrow MOSFET and a GaAs QPC, with the same fixed bias voltage of 5 mV. For very large bias voltages ~ 1 V, the MOSFET is in the pinch-off regime [40] where the conductance is limited by the region of the inversion layer closest to the $n+$ contact to which the bias is applied. Clearly in this regime the MOSFET would not function as a charge sensor, so one must be somewhat careful about how high a bias is applied to the device. There are oscillations in the current vs gate voltage characteristics (Section 4.5) with $V_b \sim 5$ mV. Because oscillations in the conductance (or current) as a function of gate voltage only occur for narrow MOSFETs, we are assured that for this bias voltage, the conductance of the MOSFET is limited by the inversion
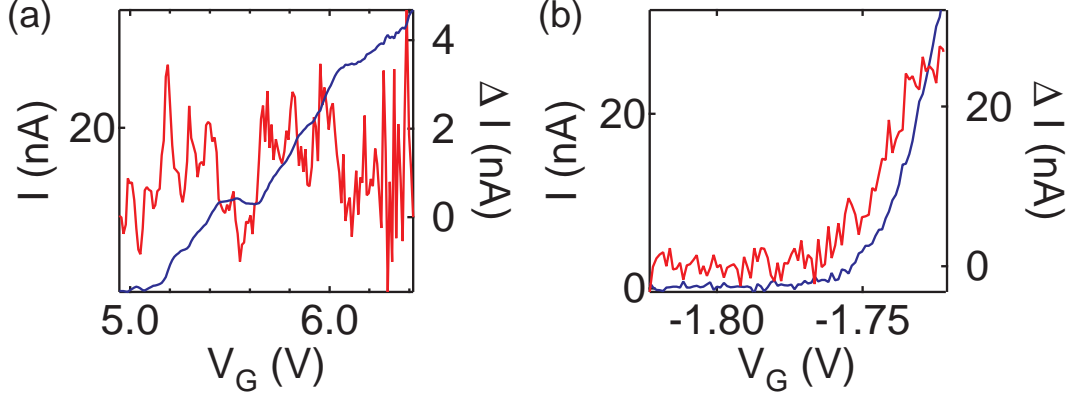
Figure 4-12: I vs V_G characteristics (blue) and current sensitivity ΔI (red) of (a) MOSFET and (b) GaAs QPC, measured with $V_b = 5$ mV, at T = 4.2 K, as discussed in the main text. For the GaAs device, the gate voltage V_G is applied to each of the two gates forming the constriction.

layer under the narrow portion of the gate. We note that there are no oscillations in the I vs V_G characteristics of the GaAs QPC measured over the same range of current: The current monotonically rises with increasing V_G. There is less disorder in the GaAs 2DEG than in the MOSFET inversion layer, as reflected by its higher mobility (Section 4.1). Because oscillations in the MOSFET current as a function of gate voltage are caused by localized states, which result from disorder in the inversion layer, it is not surprising that these oscillations are absent in the QPC current vs. gate voltage trace.

We compare the charge sensitivities of the two devices as follows. We wish to compute the amount by which the current changes for given amount of charge added to the gate. For the amount of charge added, we choose $\sigma_{100} = 1$ electron per 100 nm square. We first compute the change in gate voltage ΔV_G necessary to add this amount of charge to the gate: $\Delta V_G = \sigma_{100}/C$. Here for the MOSFET C is the gate oxide capacitance, and for the GaAs sensor C is the capacitance between the gates and the 2DEG. The change in current for 1 electron added per 100 nm square to the gate is then given by

$$\Delta I = (\partial I/\partial V_G)\Delta V_G = (\partial I/\partial V_G)\sigma_{100}/C \qquad (4.5)$$

85

We calculate ΔI for each device using $\partial I / \partial V_G$ computed from the I vs V_G data, and this quantity is plotted as a function of V_G in Fig. 4-12. We see that, while ΔI smoothly rises for the GaAs QPC as V_G is increased, for the MOSFET ΔI oscillates as a function of V_G. This is caused by the disorder induced localization effects, as the MOSFET becomes insensitive to charge near maxima and minima of the conductance oscillations. The maximum ΔI is approximately 5 times smaller for the MOSFET than for the GaAs QPC measured under these conditions. We therefore see that although the QPC is somewhat more sensitive to gate voltage, and therefore charge, than the MOSFET, the difference between the two is not very dramatic.

Chapter 5

Detecting Charge in an
Amorphous Semiconductor

In this chapter we demonstrate the application of integrated charge sensing to the study of resistive materials, using a nanometer scale MOSFET to study electron transport in hydrogenated amorphous silicon (a-Si:H), a material that is central to a variety of thin film semiconductor technologies. In Section 5.1 we give an introduction to the standard model for electron transport in a-Si:H, and discuss in general the measurement of electron transport in highly resistive materials. In Sections 5.2 and 5.3 we describe our charge detection technique and our theoretical model for extracting conductance from our measurements. In Sections 5.4, 5.5, and 5.6, we use our charge detection technique to measure the a-Si:H conductance as a function of temperature, gate voltage, and time, respectively, and show that our results are consistent with the standard model for electron transport in a-Si:H. In the final two sections we discuss a number aspects of our charge detection technique that are currently only partially understood. Specifically, we discuss variations in the charge sensing signal size, the potential insensitivity of the charge sensing technique to blocking contacts, and the application of charge sensing to the detection of switching noise in the a-Si:H. Parts of this work appear also in K. MacLean *et al.* [77].

5.1 Introduction

5.1.1 Electronic Properties of a-Si:H

We now review the electronic properties of amorphous semiconductors, focusing on a-Si:H. What will be presented here are aspects of the generally accepted model for the electronic structure and transport properties of a-Si:H. Many experimental techniques are used to investigate the electronic properties of a-Si:H, including spin resonance, thermopower measurements, and a variety of electron transport and optical methods: We will not discuss these experiments her; a thorough overview is given in Street's textbook [26].

To understand the electronic properties of amorphous semiconductors, we start by considering a crystalline semiconductor (Figure 5-1, left hand side). For a crystalline semiconductor, the atoms lie on a perfectly ordered lattice. A band gap, where the density of states is zero, separates the valence and conduction bands. The density of states terminates sharply in a Van Hove singularity [53] at the edge of valence and conduction bands, and all of the states within the valence and conduction bands are Bloch wavefunctions that extend throughout the entire crystal. For undoped material, the conduction mechanism is via activation of electrons from the valence to the conduction band, and, because of charge neutrality, which dictates that there are the same number of holes and electrons, the Fermi level lies near the center of the band gap. Because the density of states at the Fermi level is very small, the position of the Fermi level will change by a large amount if a very small amount of charge carriers are added to the crystal. This results in a high sensitivity to substitutional doping, and a large field effect.

The electronic structure of an amorphous semiconductor retains some of the characteristics of that of a crystalline semiconductor, but with some important differences (Figure 5-1, middle). For an amorphous semiconductor, the atoms still prefer to retain the same number of bonds as in the crystalline case (four, in the case of Silicon). There are still a valence and conduction band, which can be thought of as corresponding to bonding and anti-bonding electron wavefunctions. However, for an amorphous

material, the atoms do not lie on a perfect lattice, and there is no long range order. This results in what is called Anderson localization [78]. Rather than a sharp cut-off in the density of states at the conduction band edge, there are states below the conduction band, the density of which falls off exponentially as the energy is moved farther into the band gap as $\rho(E) \sim e^{-E/kT_{BT}}$. The same holds for the valence band. These states are referred to as the band tail, and T_{BT} is the band tail slope, and is typically of order 300 K. The states in the band tail do not extend throughout the crystal, but are instead *localized*. The conduction band energy E_c, sometimes referred to as the mobility edge, denotes the energy above which the states are extended, and below which the states are localized. Likewise, states above the valence band energy E_v are localized, and states below E_v are extended.

In addition to the localized states in the band tail, there are another type of localized state in amorphous semiconductors. Because of the disorder, some atoms cannot form bonds with four other atoms, and instead may be bonded to only three, resulting in what are referred to as dangling bonds. The dangling bonds result in localized states near the center of the band gap. As in the crystalline case, the Fermi level lies close to the center of the band gap. In the amorphous case however, there is a large density of states at the Fermi level resulting from dangling bonds. The electron transport is often dominated by tunneling between these localized states, a process known as hopping conduction [79, 80]. Because of the large density of states, adding charge to such an amorphous material does not change the position of the Fermi level very much. This results in a very small field effect, and a very low sensitivity to substitutional doping. There is in fact another reason why one might expect substitutional doping to be ineffective for amorphous semiconductors. For substitutional doping in a crystalline semiconductor, for instance doping silicon with phosphorous, the phosphorous dopant atom is constrained by the lattice of the crystal to have only four bonds. The fifth electron in the phosphorous atom outer shell is therefore not involved in a making a bond and can reside in the conduction band. However, in the case of an amorphous semiconductor, there is no ordered lattice forcing the phosphorous atom to make four bonds. In fact, one would expect it to be
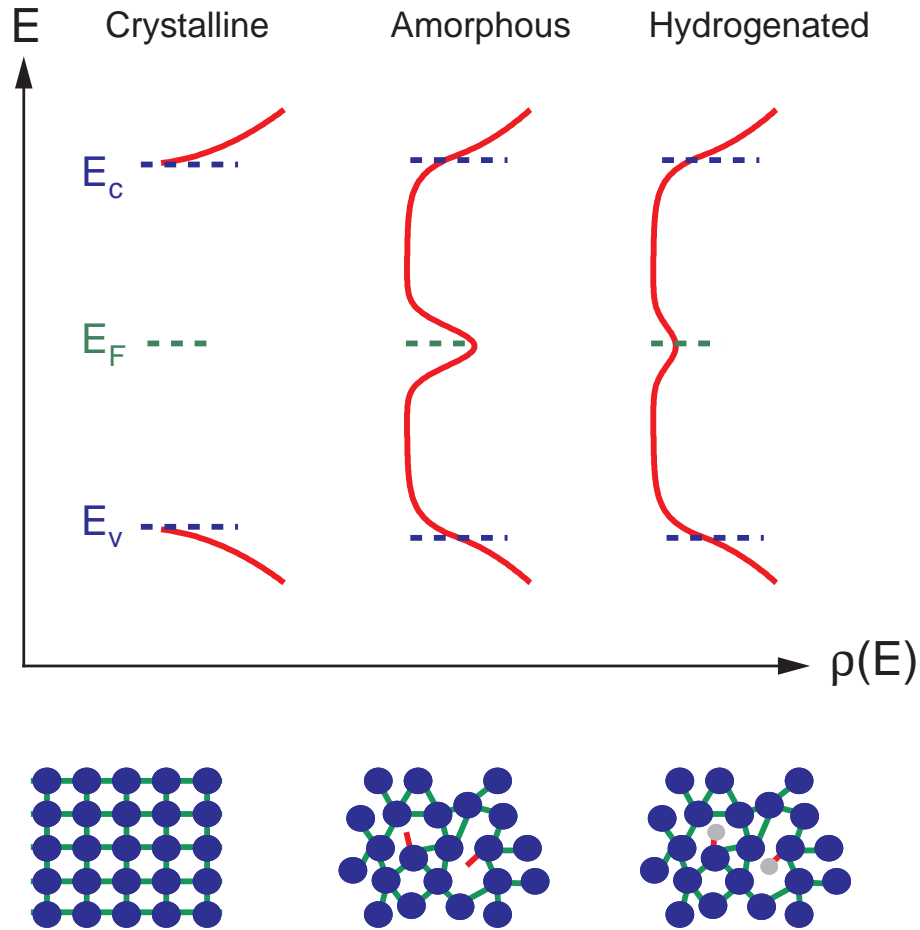
Figure 5-1: Sketch of density of states $\rho(E)$ as a function of energy, and bonding configuration for crystalline, amorphous, and hydrogenated amorphous semiconductors. For each density of states plot, the dashed blue lines indicate the valence and conduction band energies E_v and E_c, respectively, and the dashed green line indicates the position of the Fermi level E_F in undoped material. For the bonding sketches, the red lines indicate dangling bonds, and the gray circles are Hydrogen atoms.

energetically preferable for the phosphorous to make five bonds and not donate any electrons to the conduction band.

For these reasons, before the development of hydrogenated amorphous materials, it was believed by many that amorphous semiconductors would never achieve any of the same functionalities as crystalline materials. Fortunately, amorphous semiconductors alloyed with Hydrogen, in particular a-Si:H, have defied these low expectations (Figure 5-1, right). For these materials, when deposited under optimum conditions [26], many dangling bonds are electrically passivated by bonding with Hydrogen atoms, which lowers the density of states near the center of the band gap by many orders of magnitude. Because of the low density of states in the band gap, the conduction mechanism is via activation of electrons from below the Fermi level to the conduction band and there is a considerable field effect, similar to the crystalline case.

There is also a doping effect in a-Si:H, though its mechanism is more complicated and its effect is much weaker than in the crystalline case. For instance, when phosphorous dopants are introduced into a-Si:H, the vast majority of the dopant atoms simply make five bonds and therefore do not contribute to an increased carrier density. A small number of phosphorous dopants make four bonds, but for the vast majority of these, the random silicon hydrogen bonding network simultaneously forms a dangling bond, which captures the electron donated by the phosphorous. The bonding network does this because it costs less energy than having the electron donated by the phosphorous reside in the conduction band. Because of this effect, known as defect compensation, increasing the number of dopant atoms in a-Si:H has the additional effect of increasing the number of dangling bonds and hence the density of states in the band gap. Only a very small fraction of the phosphorous atoms introduced into a-Si:H, the ones making four bonds and not compensated by a dangling bond, increase the number of charge carriers. However, such atoms do in fact result in a doping effect, as was first demonstrated by Spear and Le Comber [81].

Since the first demonstrations of the field effect and substitutional doping, a-Si:H has found widespread technological use for thin film devices. For instance, thin film field effect transistors, conceptually identical to the MOSFET as described in

Section 4.1, but with the crystalline silicon replaced by a-Si:H and the thermal oxide replaced by deposited a-Si$_3$N$_4$:H, can be fabricated and are used in many commercial applications, such as controlling the voltages applied to liquid crystal displays.

The doping effect in a-Si:H is not strong enough to achieve degenerate doping, where the material becomes metallic. As the number of dopant atoms is increased, the activation energy decreases, typically reaching a minimum $\sim$ 100 meV. However, no one has succeeded in moving the Fermi level into the conduction band or valence for a-Si:H via substitutional doping. As the phosphorous concentration is increased to very high levels, the material effectively becomes a silicon-phosphorous alloy, at which point the activation energy rapidly increases. Because a-Si:H cannot be degenerately doped, even very highly doped samples are very resistive at low temperatures. The sample studied in this chapter is in fact heavily doped, but can be made very resistive by cooling to temperatures $\sim$ 150 K. Before discussing our techniques for probing electron transport in the a-Si:H in this highly resistive state, we turn to a brief review of measurements of electron transport in resistive materials.

5.1.2   Measuring Electron Transport in Resistive Materials

A variety of technologically promising materials and devices have a high electrical resistance. These materials include arrays of semiconducting nanocrystals, which are candidates for solar energy harvesting [82] and other photovoltaic applications [83], and high dielectric constant constant materials [84, 85], which are vital to modern microprocessors. For materials with excessively large resistance, traditional transport measurements fail because the current at reasonable voltages becomes too small to measure. This presents a major problem. The dependence of electrical resistance on temperature, electric and magnetic fields, light exposure, and other variables is the most important probe of novel effects in solid state physics and is essential in providing information about the electronic properties of new materials for specific applications. It is therefore highly desirable to develop an alternative means of measuring electron transport in highly resistive materials.

As we have seen in the previous chapters, in the field of single electron devices,

charge measurement using a sensor integrated with the device has recently been widely utilized to probe quantum mechanical phenomena that would be impossible to observe by measuring current. For the study of highly resistive materials scanning probe techniques have been used to determine the charge distribution and its dynamics [86, 87]. The charge flow transistor (CFT), pioneered by Senturia and coworkers [88], has enabled the measurement of highly resistive materials by making the resistive material under investigation the gate metal for a MOSFET. Since its inception, the CFT has been widely utilized in various sensors, for example, to measure the resistance of a highly resistive organic material that is sensitive to the presence of nitrogen dioxide [89, 90]. Despite its considerable utility for such applications, the CFT is not well suited for detailed studies of electron transport because the resistive material under investigation and the MOSFET sensor cannot be independently gated. This inhibits field effect measurements and other transport techniques necessary for the determination of electronic structure. Prior to the work presented in this chapter, there have been no attempts to use integrated charge sensors, of the type developed for the study of single electron devices, to study resistive materials. As we will see, the integrated charge sensor is in fact ideally suited for detailed transport studies of highly resistive materials.

In the following sections, we illustrate the power of this charge sensing technique by investigating transport in hydrogenated amorphous silicon (a-Si:H). By patterning a strip of a-Si:H thin film adjacent to a nanometer scale silicon MOSFET, we are able to detect charging of the a-Si:H and measure extremely high resistances ($\sim 10^{17}$ Ω) using moderate voltages (~ 1 V). We compare our results with those of current measurements at high temperatures, where the resistance is not too large, and find good agreement. The two methods complement each other in that they probe different ranges of electrical resistance. Our device geometry, in which the MOSFET sensor and a-Si:H can be gated independently, allows us to investigate a variety of transport phenomena, including the field effect [91] and dispersive transport [92, 93], using charge sensing. We use these methods to probe the density of localized states near the Fermi energy, and obtain consistent results.

5.2   Charge Detection Technique

Our charge sensor consists of an n-channel MOSFET that is electrostatically coupled to a strip of a-Si:H. An electron micrograph of the structure is shown in Fig. 5-2(a). The sample fabrication procedure is discussed in detail in Chapter 4; here we quickly review aspects of the device that are vital to an understanding of our charge detection technique. Because of its narrow width, the MOSFET is extremely sensitive to its electrostatic environment [14]. Furthermore, the MOSFET has a relatively thick gate insulator (SiO$_2$, with a thickness of $d_{ox} = 100$ nm), which ensures that the metallic polysilicon gate does not effectively screen the inversion layer from nearby electrostatic fluctuations. Adjacent to the MOSFET, we pattern a strip of phosphorous doped a-Si:H. The a-Si:H is deposited by plasma enhanced chemical vapor deposition [26], with a gas phase doping ratio and hydrogen dilution of [PH$_3$] / [SiH$_4$] $= 2 \times 10^{-2}$ and [H$_2$] / [SiH$_4$ + H$_2$] $= 0.5$, respectively (Section 4.3). Because we use a relatively large doping level, we expect a large defect density $N_D \sim 10^{18}$ cm^{-3} [26]. The a-Si:H strip is connected to two gold contacts (separated by ≈ 2 μm), and the MOSFET inversion layer is contacted through two degenerately doped n^+ regions, none of which are shown in Fig. 5-2(a). After sample preparation, the device is loaded into a cryostat, and kept in Helium exchange gas throughout the course of the measurements discussed here.

The parameters used to deposit our a-Si:H samples are similar to those studied previously, and thicker a-Si:H films than those used for our charge sensing measurements, deposited under the same conditions, have conductivities and activation energies similar to those reported elsewhere [26]. However, the sample studied in this work is nanopatterned (≈ 100 nm wide at its narrowest point) and also is only $\approx$ 50 nm thick. Thus, although the characteristics reported below are similar to what one would expect for thick heavily doped films, we expect that for our sample surface effects may be significant [70], and there may also be differences in morphology and hydrogen content as compared with thicker films.

Our measurement consists of monitoring the MOSFET conductance G_M as a

function of time after changing the voltage applied to one of the a-Si:H contacts (Fig. 5-2(c)). We set the voltage of one of the a-Si:H contacts to 0 V relative to the p-type substrate.[1] We rapidly change the voltage V_{aSi} applied to the other a-Si:H contact from -1.8 V to -2.7 V and back again, as shown in the upper panel of Fig. 5-2(c). The response of G_M to this pulse sequence is shown in Fig. 5-2(c). When V_{aSi} is changed from -1.8 to -2.7, G_M first drops instantaneously by an amount ΔG_{M-Au} because the negative charge added to the gold contact couples to the MOSFET electrostatically. The change in V_{aSi} also causes additional electrons to move onto the a-Si:H strip from the gold contacts. The MOSFET senses this change in charge as well, and G_M slowly decreases by an amount ΔG_{M-aSi}. This decrease in G_M, caused by the charging of the a-Si:H strip capacitor, will henceforth be referred to as a charge transient. When the a-Si:H contact voltage is changed back to its original value of -1.8 V, an instantaneous response to the gold contact followed by a charge transient is again observed, but with the opposite sign. A similar response is observed when the pulse sequence is applied to the other a-Si:H contact, or when the pulse sequence is applied to both a-Si:H contacts at the same time.

To verify that the charge transient is caused by charging of the a-Si:H strip, we measure a separate device for which only one of the gold contacts is connected to a-Si:H close to the MOSFET charge sensor. The results are shown in Fig. 5-2(d). When we change the voltage applied to the contact connected to a-Si:H close to the sensor, we observe both an instantaneous response and a charge transient. However, when we change the voltage applied to the other gold contact, we observe only the instantaneous response. This verifies that the charge transient and instantaneous response are indeed caused by charge added to the a-Si:H strip and gold contact, respectively.

As will be discussed in the following section, the time scale of the charge transient is a measurement of the resistance of the a-Si:H strip. The size of the charge transient ΔG_{M-aSi} depends on screening by the underlying p-type substrate, and is discussed

---

[1]For all of the data shown in this chapter we set the voltage of the substrate, relative to the inversion layer, to $V_{sub} = $ -3 V, unless stated otherwise
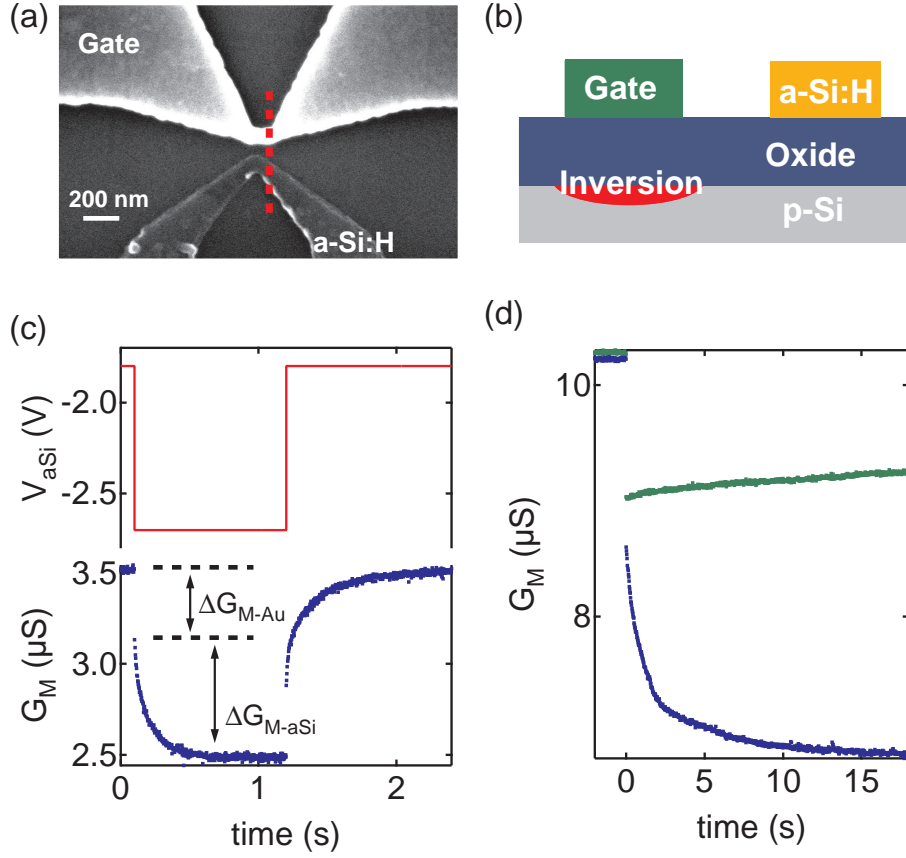
Figure 5-2: (a) Electron micrograph of MOSFET gate and a-Si:H strip. A positive voltage is applied to the MOSFET gate, forming an inversion layer underneath. (b) Vertical sketch of the device geometry along the dashed line in (a). The conductance through the inversion layer formed under the gate is sensitive to the charge on the a-Si:H strip. (c) Measurement technique (T = 125 K). The upper panel shows the voltage V_{aSi} applied to one of the a-Si:H contacts as a function of time. The lower panel shows the resulting time dependence of the transistor conductance G_M, which changes by ΔG_{M-aSi} (ΔG_{M-Au}) when charge is added to the a-Si:H (gold contact), as described in the main text. For this trace multiple charge transients have been averaged to improve signal-to-noise ratio. (d) Result of pulsing the gold contacts for a device where only one of the contacts is connected to a-Si:H close to the MOSFET. Here we pulse the gold contact not connected to a-Si:H (green) and connected to a-Si:H (blue) close to the MOSFET from 0 to -9.9 V at $t = 0$ (T = 79 K), confirming that the charge transient is caused by charging of the a-Si:H strip, as discussed in the main text. For this data $V_{sub} = 0$.

96

in Section 5.7. In order to obtain charge transients with a high signal-to-noise ratio, multiple traces can be averaged together. For some temperatures, the charge transients are obscured by large but infrequent telegraph noise switches (Section 4.5). To alleviate this problem, we implement software in our data acquisition system that checks whether or not there is a telegraph noise switch in each charge transient trace by comparing it to the average of the previous charge transient traces. If the current trace deviates from the average of the previous traces by more than a specified threshold, it is thrown out. If not, the trace is added to the running average. This technique makes it possible to average large numbers of charge transients without having the average be corrupted by the few traces that are obscured by large telegraph switches.

5.3   Diffusion Model

To develop a quantitative model for extracting the resistance of the a-Si:H from charge transients, we consider the strip a-Si:H strip as a distributed RC network [94], as shown in Fig. 5-3. Here the resistance per unit length dx is given by $dR = R_{sq}dx/w$ where w is the width of the strip (measured perpendicular to the plane of Fig. 5-3), and R_{sq} is the resistance per square of the a-Si:H film. C is the capacitance per unit area between the a-Si:H film and the p-type substrate. If $\sigma(x,t)$ is the charge per unit area as a function of position along the strip x and time t, then the voltage of the a-Si:H strip relative to the p-type substrate is given by $V = \sigma(x,t)/C$. Using charge conservation [95], we obtain:

$$w\frac{\partial \sigma(x,t)}{\partial t} = -\frac{\partial I(x,t)}{\partial x} \tag{5.1}$$

Here $I(x,t)$ is the current flowing through the strip. If we set $dV(x,t) = V(x,t) - V(x+dx,t) = -(\partial\sigma(x,t)/\partial x)dx/C$, we have from Ohm's law:

$$I(x,t)R_{sq}dx/w = I(x,t)dR = -dV(x,t) = -(\partial\sigma(x,t)/\partial x)dx/C \tag{5.2}$$
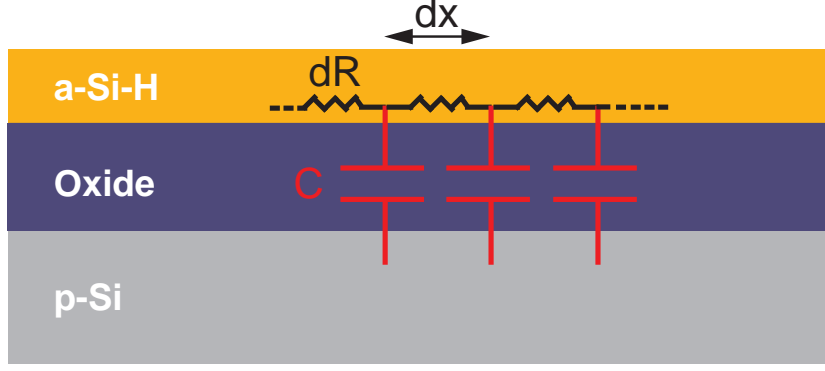
Figure 5-3: Model for charge diffusion along a strip of a-Si:H, as discussed in the main text. The strip has resistance dR and per unit length dx, and capacitance per unit area to the p-type substrate C.

Dividing both sides of this equation by dx, differentiating with respect to x, and using Eqn. 5.1, we obtain a diffusion equation for $\sigma(x,t)$:

$$\frac{\partial \sigma(x,t)}{\partial t} = D\frac{\partial^2 \sigma(x,t)}{\partial x^2} \tag{5.3}$$

where here the diffusion constant is given by $D^{-1} = R_{sq}C$. Because this equation is independent of the width of the strip w, we do not need to include the gradual taper of the a-Si:H strip used in our experiments in our model.

The diffusion equation can be solved by separation of variables [96], and for a strip of length L, the solution is given by:

$$\Delta\sigma(x,t) = \sigma(x,t) - \sigma(x, t \to \infty) = \sum_{n=1}^{\infty} a_n sin(n\pi x/L)e^{-Dn^2\pi^2 t/L^2} \tag{5.4}$$

Here the coefficients a_n are determined by Fourier decomposition of the difference between the initial and final charge profiles of the strip $\Delta\sigma(x,0)$. The slowest exponential term, corresponding to $n = 1$ in the summation, will dominate at sufficiently long times. The next slowest term is $n = 2$, which is 4 times faster because of the factor of n^2 in the exponent. Moreover, at the center of the strip ($x = L/2$), the point to which our charge sensor is presumably most sensitive, the $n = 2$ term vanishes, so that at this point the second slowest term is $n = 3$, which is 9 times faster than

the $n = 1$ term. We can therefore neglect the terms with $n > 1$ except at very small times t, and so we approximate:

$$\Delta\sigma(t) \propto e^{-D\pi^2 t/L^2} \tag{5.5}$$

Now, to lowest order G_M varies linearly with $\Delta\sigma$ so that we can write for the time dependence of the transistor conductance:

$$G_M(t) \approx G_\infty + G_\Delta e^{-\Gamma t} \tag{5.6}$$

Here the sign of G_Δ is opposite to the sign of $\Delta\sigma$, and we define the charging rate $\Gamma = D\pi^2/L^2$. As we will see in the following section, we can fit our charge transient measurements to obtain Γ. From this value, and an estimate of C, we can compute the conductance of the a-Si:H strip.

$$G_{aSi} = w/(R_{sq}L) = wLC\Gamma/\pi^2 \tag{5.7}$$

As will be discussed in Section 5.7, the p-type silicon underneath the a-Si:H can be depleted, depending on the voltages applied to the MOSFET gate, a-Si:H contacts, and p-type substrate. This depletion reduces C below the oxide capacitance $C_{ox} = \epsilon_0\kappa_{ox}/d_{ox}$, where d_{ox} and κ_{ox} are the thickness and dielectric constant of the oxide. As will be discussed in Section 5.7, we estimate that for the range of voltages used in our experiments, $C_{ox}/5 < C < C_{ox}$. For calculations involving C, we use the midpoint of this range.

If we define $\tau_{ch} = \Gamma^{-1}$, $R_{aSi} = G_{aSi}^{-1}$ and $C_{eff} = wLC/\pi^2$, we obtain the relation $\tau_{ch} = R_{aSi}C_{eff}$. We see then that the charging of the a-Si:H strip can be viewed approximately as charging a capacitance C_{eff} through the resistance of the strip R_{aSi}. The power of the charge transient technique is derived in part from the fact that, because of the nanoscale dimensions of our a-Si:H strip, C_{eff} is very small: We estimate $C_{eff} \sim 10$ aF, which is many orders of magnitude smaller than the input capacitance of a voltage amplifier, even when great care is taken to minimize this quantity [54]. Based on the geometry of our device (Figure 5-2(b)), the charge sensor input capac-

itance, between the a-Si:H strip and the MOSFET gate and inversion layer, is even smaller. Because C_{eff} is so small, the charging time τ_{ch} remains measurably short, even for very large resistances R_{aSi}.

Clearly our method cannot be used to probe the resistance of materials that are more insulating than the oxide. In fact, setting aside the obvious practical requirement that τ_{ch} be conveniently short, the finite resistivity of the oxide sets the theoretical limit of how high a resistance we can measure. To include the effect of an oxide conductance in our model, we add to the model shown in Fig. 5-3 a resistor in parallel with the capacitance C. It is straightforward to modify Eqn. 5.3 to include the effect of this leakage resistance:

$$\frac{\partial \sigma(x,t)}{\partial t} = -D\frac{\partial^2 \sigma(x,t)}{\partial x^2} - \frac{\sigma(x,t)}{\tau_d} \tag{5.8}$$

Here the dielectric relaxation time is defined as $\tau_d = 1/(\rho_{ox}\kappa_{ox}\epsilon_0)$, where ρ_{ox} is the resistivity of the oxide. In order for leakage through the oxide to be negligible, we require that the second term on the right hand side of Eqn. 5.8 be very small compared with the other terms which is true provided $\tau_d >> \tau_{ch}$. Our measurement thus requires that the dielectric relaxation time of the oxide is much larger than the charging time.

5.4   Temperature Dependence

To extract the charging rate Γ, we measure a charge transient as discussed in Section 5.2, and fit the data to Eqn. 5.6. The results for two charge transients measured at T = 125 K and T = 140 K are shown in Fig. 5-4(a). Here we show only the charge transient part of the $G_M(t)$ trace immediately after the voltage applied to the a-Si:H contact is made more negative, and do not show the instantaneous response to the gold, or the reverse response that is observed when the a-Si:H contact voltage is changed back to its original value (Fig. 5-2(c)). We see that the data fits well to Eqn. 5.6, and that as the temperature grows, the charging rate Γ gets faster. From the value we extract for Γ we can determine the a-Si:H conductance G_{aSi} using Eqn.
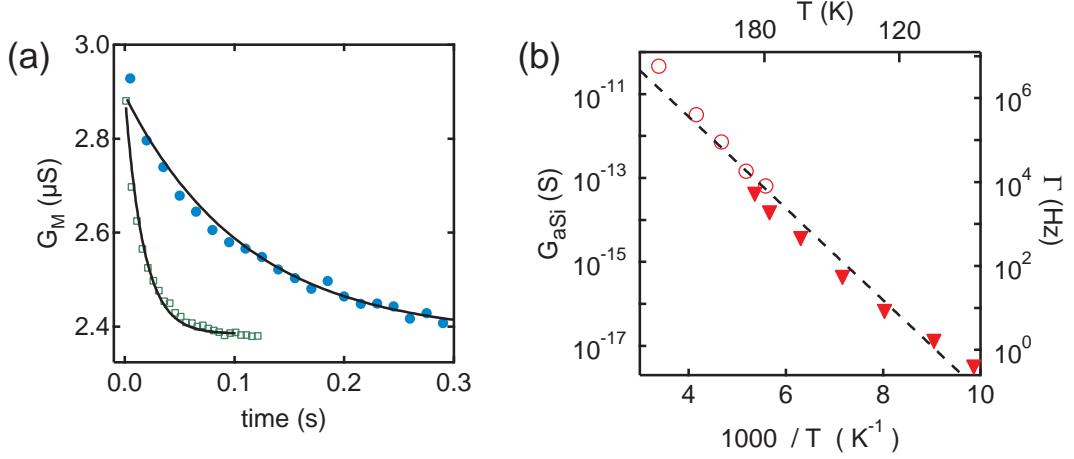
Figure 5-4: (a) G_M as a function of time at T = 125 K (closed circles) and 140 K (open squares), the latter is offset for clarity. At $t = 0$ the voltage applied to one of the a-Si:H contacts is changed from -1.8 V to -2.7 V. For these traces multiple charge transients have been averaged to improve signal-to-noise. The solid lines are theoretical fits described in the main text. (b) Conductance G_{aSi} obtained from charge transients (closed triangles) and direct conductance measurements (open circles), measured with $V_{ds} \approx$ -2.3 V. For the charge transient measurement, Γ is given on the right hand axis. The dashed line is a theoretical fit described in the main text.

5.7.

In Fig. 5-4(b) we plot G_{aSi} and Γ and as functions of temperature. G_M is weakly temperature dependent, and for this data we therefore adjust the MOSFET gate voltage as we vary the temperature to keep G_M approximately constant. At higher temperatures we are able to directly measure $G_{aSi} = dI/dV_{ds}$, where V_{ds} is the voltage between the a-Si:H contacts, and these results are also shown in Fig. 5-4(b). At T $\approx$ 180 K, we can measure G_{aSi} using both techniques, and the results are in good agreement. The measurements are complementary, in that the charge transient technique is easier to implement for smaller conductances G_{aSi} because the charging is slower, while a measurement of current is only possible for larger values G_{aSi}. The dashed line in Fig. 5-4(b) is a fit to $G_{aSi}(T) = G_0 e^{-E_a/kT}$. The data are thus consistent with an activated transport mechanism, with an activation energy $E_a \approx$ 200 meV, as is typically observed for a-Si:H films heavily doped with phosphorous [26]. We note that at the lowest temperatures, we measure resistances as high as $\sim 10^{17}$ Ω.

5.5 Field Effect

We can also measure Γ as a function of gate voltage (Fig. 5-5). For this measurement, we apply the same voltage V_{aSi} to both a-Si:H contacts relative to the p-type substrate. The effective gate voltage is then $V_g = -V_{aSi}$. We then add a small voltage step $\Delta V \approx 0.5V$ to V_{aSi} to produce a charge transient, from which Γ is extracted as in Fig. 5-4(a). Unlike previous reports [88], our geometry allows us to maintain an approximately constant value for the MOSFET conductance, and thus to maintain a high charge sensitivity, as we make large changes in V_{aSi} by applying smaller compensating voltage shifts to the MOSFET gate voltage. This allows us to perform this field effect measurement for a large range of sample conductances.

In Fig. 5-5 we measure Γ as a function of V_g at three different temperatures. We see that Γ increases as V_g is increased, indicating n-type conduction through the a-Si:H, as expected for a phosphorous doped sample. The exponential increase in Γ with V_g is consistent with the activated conduction found in Fig. 5-4(b), provided we assume an approximately constant density of localized states. We have:

$$\Gamma = \omega_0 e^{-E_A/kT} \tag{5.9}$$

Here ω_0 is a prefactor that depends only weakly on temperature, and the activation energy E_A is reduced as the gate voltage moves the Fermi level closer to the mobility edge. The logarithmic slope $\alpha = \partial ln(\Gamma)/\partial V_g$ is then given by $\alpha = \frac{1}{kT}\partial E_A/\partial V_g = C/(ekT\rho(E_F)s_{tf})$, where $\rho(E_F)$ and s_{tf} are the density of states at the Fermi energy and Thomas-Fermi Screening length, respectively [97].[2] Thus we expect an exponential increase in Γ with V_g as long as the product $\rho(E_F)s_{tf}$ is constant.

At each temperature, we fit the data to obtain α (solid lines in Fig. 5-5), and, in the inset to Fig. 5-5, we plot α as a function of inverse temperature. The dashed line is a linear fit (constrained to pass through zero) and is consistent with the data. From the slope of this fit we obtain $s_{tf}\rho(E_F) \approx 5 \times 10^{13}$ eV^{-1}cm^{-2}, and, expressing

---

[2]For these expressions we have used the T = 0 limit, which is justified provided $\rho(E)$ does not vary substantially over a range of energy kT about E_F [91].
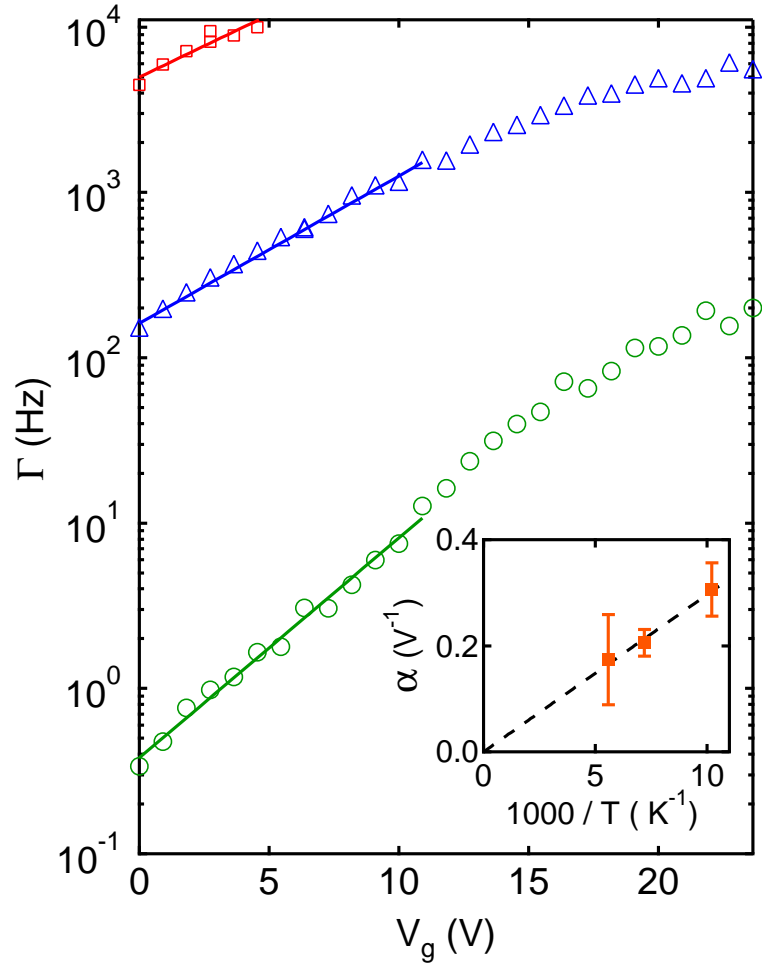
Figure 5-5: Γ as a function of V_g at T = 98 K (circles), 139 K (triangles), and 179 K (squares). The solid lines are theoretical fits described in the main text. (Inset) α as a function of inverse temperature. The dashed line is a theoretical fit described in the main text.

s_{tf} in terms of $\rho(E_F)$ and the a-Si:H dielectric constant [97], we solve for $\rho(E_F) \sim$ 10^{20} eV^{-1}cm^{-3}. The density of states at the Fermi level for phosphorous doped amorphous hydrogenated silicon obtained from more commonly used transport techniques is typically $\sim 10^{19}$ eV^{-1}cm^{-3} [91]. The fact that our number for $\rho(E_F)$ is somewhat high is not surprising, given the large gas phase doping level used in our a-Si:H film deposition. For the range of voltages used here E_F moves by an amount comparable to values of the band tail width commonly found for a-Si:H films [26], so we expect that $\rho(E_F)$ should increase somewhat as V_g is made more positive and the Fermi level is moved into the band tail. This may cause the observed decrease in logarithmic slope α for $V_g >$15 V in Fig. 5-5.

5.6   Dispersive Transport

At lower temperatures, where the time scale for charging is longer, we observe hysteretic behavior and dispersive transport. Dispersive transport, where the mobility of the sample has a power law dependence on time, has been studied extensively in a-Si:H and other amorphous semiconductors using various optical techniques [92, 93], and results from localized states with a broad distribution of energies. As we will see below, here dispersive transport is manifested as a power law time dependence of the rate at which the charge on the a-Si:H changes following a large change in V_{aSi}.

An example of the hysteretic behavior we observe is shown in Fig. 5-6(a). First, we sweep the voltage applied to both a-Si:H contacts from 0 to -25 volts. We see that G_M falls as V_{aSi} is made more negative. This is expected because making V_{aSi} more negative adds additional electrons to the a-Si:H strip. However, when V_{aSi} is swept back to 0 V, we see that G_M does not regain its original value. Evidently the electrons that were added to the strip are now trapped. The size of this hysteresis loop shrinks with increasing temperature. After observing hysteretic behavior, we bring the device back to its original state by warming up the cryostat.

The mechanism underlying the hysteretic behavior shown in Fig. 5-6(a) is illustrated in Fig. 5-6(b), which shows a diagram of the charge state of the a-Si:H at

different points along the hysteresis curve. As V_{aSi} is made more negative, the a-Si:H conduction band energy E_c is brought closer to the Fermi level in the gold contacts. This allows electrons to charge the a-Si:H at an appreciable rate. However, as time elapses, these electrons are retrapped in localized states farther below E_c, from which they cannot escape on the time scale of the experiment. When V_{aSi} is swept back to 0, these electrons therefore remain in the a-Si:H, and G_M does not regain its original value.

We can make use of this hysteretic behavior to observe dispersive transport, from which we can extract the density of states at the Fermi level of the a-Si:H. To accomplish this, we use the pulse sequence shown in Fig. 5-7. When we quickly step V_{aSi} from 0 V to -24 V, G_M quickly drops. However, when V_{aSi} is stepped back to 0 V, G_M rises at slower rate, and does not regain its original value. This behavior can be understood in the same way as the hysteretic behavior shown in Fig. 5-6: When V_{aSi} is stepped more negative, the a-Si:H quickly charges, as electrons can enter the a-Si:H at energies close to the conduction band. However, as time progresses, these electrons get trapped in localized states deeper in the a-Si:H band gap. When the voltage is returned to its original value, the a-Si:H therefore takes a much longer time to discharge: From Eq. (5.9), the time necessary to release electrons from states at an energy E_A below the transport energy is $t \sim \Gamma^{-1} = \omega_0^{-1} e^{-E_A/kT}$. As electrons deeper and deeper in the gap are released, t grows, and thus the transport process becomes dispersive [92, 93].

We can understand the time dependence following the voltage step quantitatively. At a time t after the negative voltage step, only electrons in localized states with energies $E_A < E_{max} = kT ln(\omega_0 t)$ [92, 93] are able to escape from the a-Si:H. The charge on the a-Si:H is then given by $\sigma(t) = e \int^{E_{max}} s_{tf} \rho(E_A) dE_A$ (up to an additive constant). Assuming a constant density of states and differentiating with respect to time we obtain:

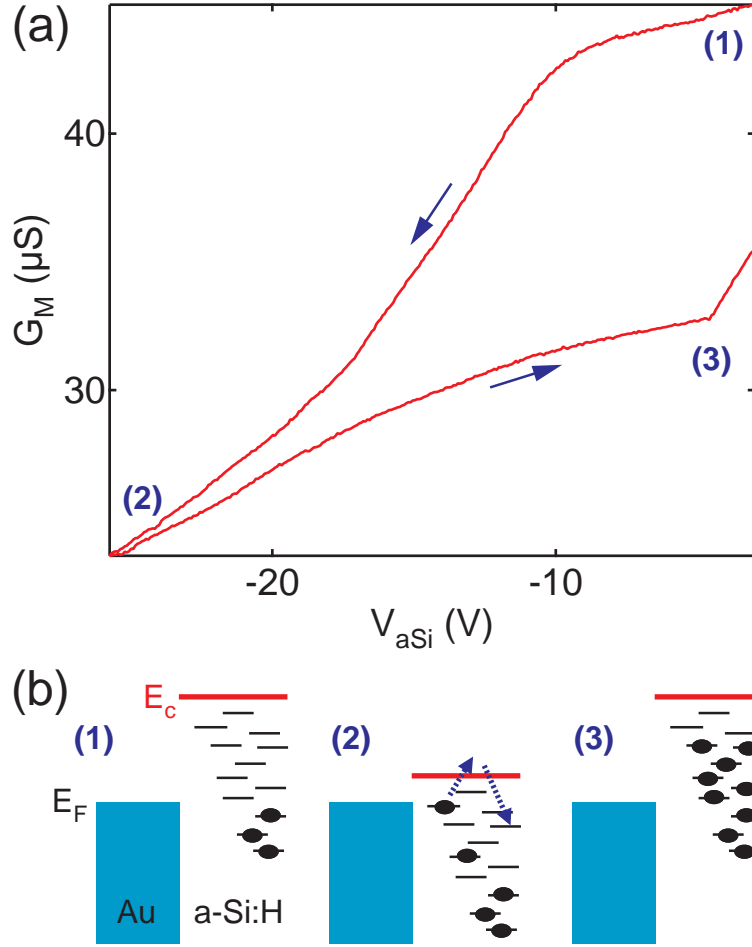$$d\sigma/dt = e s_{tf} \rho(E_F) kT/t \qquad (5.10)$$

105

Figure 5-6: (a) Hysteresis in transistor conductance G_M as a function of the voltage applied to both a-Si:H contacts V_{aSi}, measured at T = 57 K. (b) Model for hysteresis, as discussed in the main text. We start (1) with a large number of localized states in the a-Si:H above the Fermi energy in the gold (E_F). As the voltage applied to the contacts V_{aSi} is made increasingly negative, these localized states are brought below E_F. (2) Electrons enter the a-Si:H at energies close to the conduction band energy E_c, reducing G_M, but as time progresses, these electrons are trapped farther and farther below E_c. (3) When V_{aSi} is returned to its original value, electrons sufficiently far below E_c are trapped. These electrons do not leave the a-Si:H, so that G_M remains below its original value.
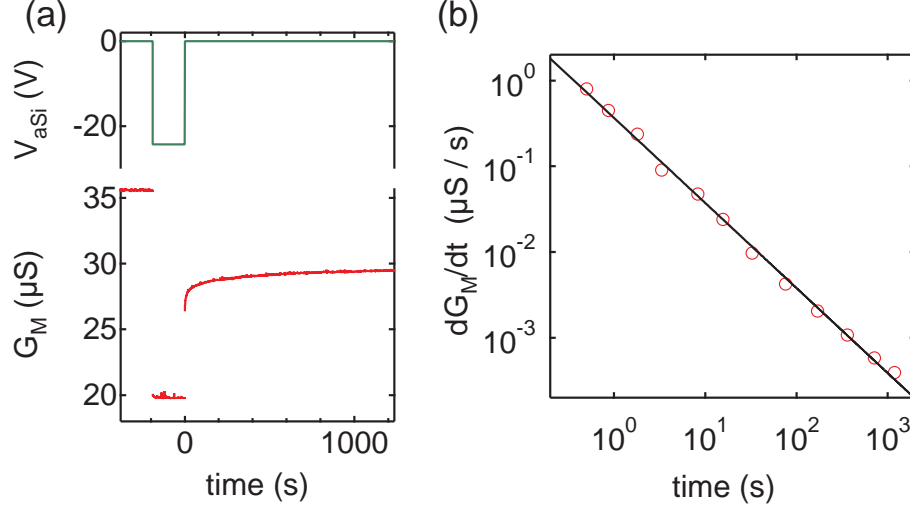
Figure 5-7: (a) Pulse sequence (top panel) and transistor conductance (bottom panel) used to measure dispersive transport at T = 89 K, as discussed in the main text (b) $dG_M(t)/dt$ extracted from the data shown in (a). The solid line is a theoretical fit described in the main text.

In Fig. 5-7(b) we plot the derivative of G_M with respect to time on a log-log plot: A fit to a power law dependence (solid line) yields a power of -1 ± 0.1 as expected for a constant density of states. Moreover, the prefactor of this power law is $es_{tf}\rho(E_F)\beta$, where β is the conversion between σ and G_M that can be estimated from the decrease in G_M after pulsing V_{aSi} to -24 volts. We obtain $\rho(E_F) \sim 10^{20}$ eV^{-1}cm^{-3}, consistent with the value extracted from the data in Fig. 5-5.

From the dependence of the a-Si:H charging rate on temperature (Fig. 5-4), gate voltage (Fig. 5-5), and time (Fig. 5-7), we have established a picture of the transport in the a-Si:H strip that is consistent with the standard model for transport in heavily doped hydrogenated amorphous silicon [26]. The Fermi level lies $\sim$ 200 meV below the conduction band, where the density of states is $\sim 10^{20}$ eV^{-1}cm^{-3}, and conduction occurs via activation of electrons from the Fermi level to the conduction band. In the following section, we discuss the magnitude of the MOSFET response to changes in the a-Si:H charge.

107

5.7 Screening Effects

The sensitivity of the MOSFET detector to its electrostatic environment depends on screening by the underlying p-type silicon substrate. To demonstrate this, we examine the response of the MOSFET to changes in the charge on the gold contacts. As we saw in Section 5.2 (Figure 5-2(c)), the MOSFET conductance responds to changes in the charge on the gold contacts, and, at temperatures of ~ 100 K, this response is instantaneous. In fact, the response of the MOSFET to the gold contact would be much larger were it not screened by the silicon substrate. To see this, we cool down to lower temperatures ~ 10 K, where the resistance of the silicon substrate is much larger. When we then change the voltage applied to the gold contacts, we see a large jump in the MOSFET conductance which gradually dies away as time progresses (Fig 5-8(a)). This transient response has nothing to do with the amorphous hydrogenated silicon: At T = 10 K the a-Si:H is too resistive to charge on any reasonable time scale [3]. This response can be understood in terms of screening. When we add charge to the gold contact, an opposing charge in the p-type substrate is induced, reducing the over all effect of the charge on the MOSFET conductance. At low temperatures, where the resistance of the substrate is high, this charge is induced at a measurably slow rate r. To obtain a value for r, we fit the $G_M(t)$ trace to an exponential, as shown in Fig 5-8(a).

The screening rate r is proportional to the conductance of the p-type substrate. To see this, we measure r as a function of inverse temperature. The results are shown in Fig 5-8(b). As the temperature is reduced, r drops, saturating at a minimum value $r_{min} \approx 8$ Hz. In Fig 5-8(b), we plot $\Delta r = r$ - r_{min} as function of inverse temperature, and fit to an activated temperature dependence $\Delta r \propto e^{-E_{act}/kT}$. We obtain $E_{act} =$ 45 meV, which is exactly the Boron acceptor binding energy [40], demonstrating that conduction through the Boron doped substrate is what determines r.[4] Presumably r

---

[3]This response is also observed for a device where the gold contact to which the voltage change is applied is not connected to a-Si:H close to the MOSFET

[4]In principle, for a p-type silicon substrate with no defects at low temperatures, the Fermi level lies between the acceptor and valence bands, so that the expected activation energy is closer to half of the binding energy. However, even a very small number of compensating defects $N_D \sim 10^{10}$ cm^{-3} will move the Fermi level into the acceptor band, so that the activation energy is equal to the
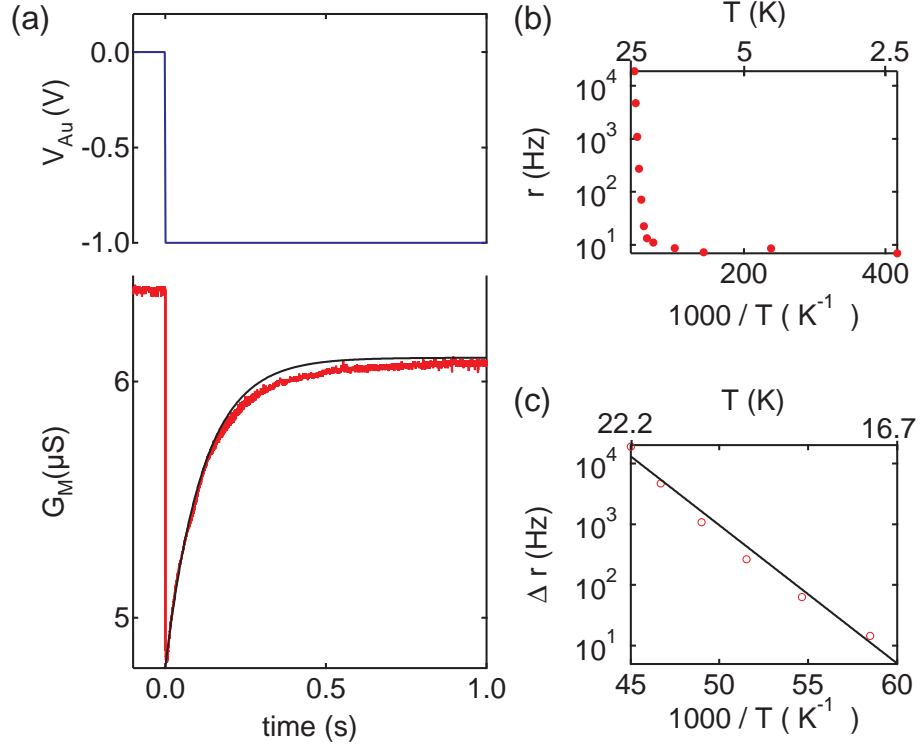
Figure 5-8: (a) Observation of screening effect at T = 9.8 K. Here we pulse the voltage applied to both gold contacts from 0 to -1 V at $t = 0$ (top panel). G_M drops by a large amount because of the negative charge added to the gold, but then rises as this charge is screened by charge in the p-type substrate. The solid black curve is a fit to an exponential, as discussed in the main text. (b) Screening rate r as a function of inverse temperature. (c) Change in screening rate Δr from low temperature value as a function of inverse temperature, as described in the main text. The solid line is a theoretical fit from which the Boron acceptor activation energy $E_{act} \approx 45$ meV is extracted, as described in the main text. For all of this data, $V_{sub} = 0$.

saturates at a minimum value r_{min} because some conduction path other than through the p-type substrate by activation of holes in the valence band dominates at low temperature, though it is not clear what this conduction path could be. Tunneling of electrons between neighboring acceptor atoms seems unlikely, as the mean spacing between Boron atoms ≈ 60 nm is much larger than the size of the Boron acceptor wavefunction ~ 1 nm [53]. In any case, from this data it is clear that screening by the substrate significantly reduces the response of the MOSFET to its electrostatic environment.

At higher temperatures $T > 25$ K, r becomes very fast, so that the screening takes place much more rapidly than the charge diffusion in the a-Si:H. In this regime, we quantify the effect of the screening on the charge sensing signal by measuring the response of the charge detector to changes in the charge on either the gold contacts or the a-Si:H. The changes in G_M for a given change in charge on the gold or a-Si:H, denoted ΔG_{M-Au} and ΔG_{M-aSi} respectively, are extracted from charge transient traces as shown in Fig. 5-2(c). We plot these quantities as a function of $V_g = -V_{aSi}$ in Fig. 5-9(a), at three different temperatures. We see that at each temperature, both ΔG_{M-Au} and ΔG_{aSi} decrease with increasing V_g.

To understand this effect qualitatively, we refer to the sketch in Fig. 5-9(b). The negative bias applied to the p-type substrate relative to the inversion layer prevents inversion anywhere at the Si/SiO₂ interface [40] except underneath the gate (to which, of course, a large positive voltage is applied). Therefore, the p-type silicon underneath the a-Si:H and gold contacts is in a state of either depletion or accumulation. For the depleted state, as shown in Fig. 5-9(b), the screening of the charge in the a-Si:H (or gold) is done by charge in the substrate that is a distance L_D away, where L_D is the depletion length. The maximum extent of the depletion region, for the voltages used in our experiment, is quite large for our lightly doped substrate: We estimate $L_D \approx 1.4$ μm as the maximum depletion length [40]. The depletion reduces the capacitance C between the a-Si:H and the silicon substrate by adding a capacitance $C_{Si} = \epsilon_0 \kappa_{Si}/L_D$, where κ_{Si} is the dielectric constant of Silicon, in series with the
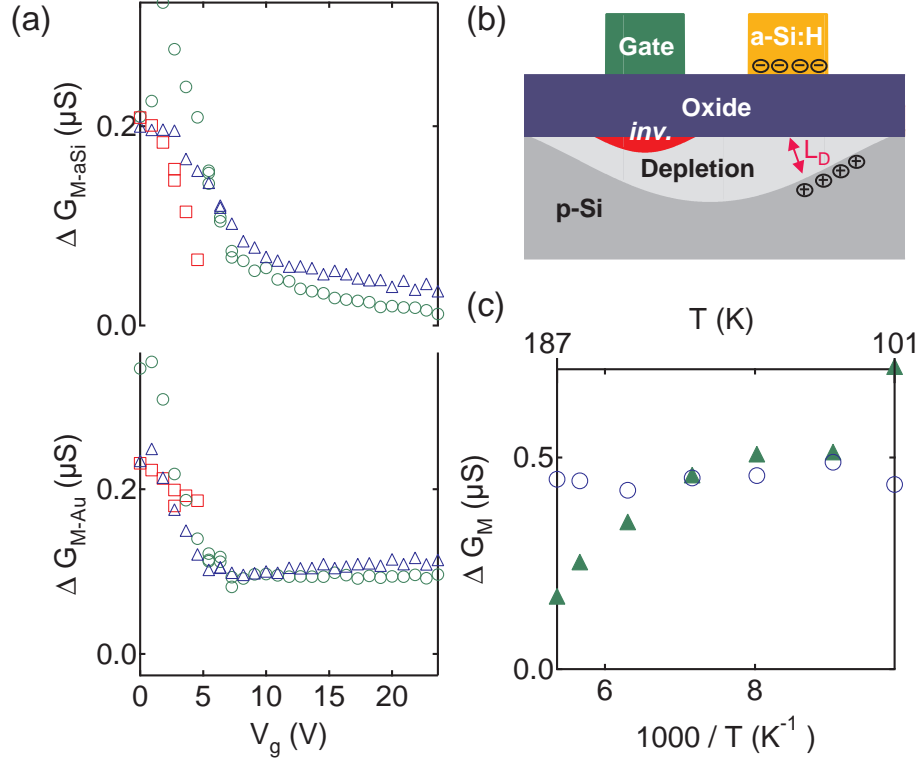
binding energy.

110

Figure 5-9: Transistor response ΔG_{M-aSi} (ΔG_{M-Au}) to charge added to the a-Si:H and gold contacts respectively. (a) Variations in ΔG_{M-aSi} (top panel) and (ΔG_{M-Au}) (bottom panel) with effective gate voltage V_g at 98 K (circles), 139 K (triangles) and 179 K (squares). These data are extracted from the same charge transient traces from which we extract the data shown in Fig. 5-5. (b) Model for dependence of transistor response on gate voltage, as discussed in the main text. (c) Transistor response ΔG_{M-aSi} (closed triangles) and ΔG_{M-Au} (open circles) as a function of inverse temperature. These data are extracted the same from the same charge transient traces from which we extract the data shown in Fig. 5-4.

111

oxide capacitance C_{ox}. For the maximum depletion length we calculate $C \approx C_{ox}/5$. The reduction of C caused by depletion would be more severe, were it not for the fact that κ_{Si} is 3 times larger than the oxide dielectric constant κ_{ox}. When a negative voltage is applied to the a-Si:H relative to the substrate, corresponding to a positive V_g, the depletion layer shrinks, and the capacitance increases toward C_{ox}. In addition to these changes in capacitance, we expect large changes in the signal size. As V_g is made more positive, L_D shrinks. This makes the screening more effective, and both signal sizes ΔG_{M-Au} and ΔG_{M-aSi} shrink. Depletion therefore constitutes a qualitative explanation for the data shown in Fig. 5-9(a).

For the response of the gold to V_g, we see that the signal is large for V_g close to zero, and decreases with increasing V_g, saturating at a minimum value at $V_g \approx 7$ V. For $V_g > 7$ V, the response is roughly independent of V_g. This is to be expected: Presumably the silicon under the gold reaches accumulation at $V_g \approx 7$ V. The distance between the charge on the gold and the screening charge in the substrate remains fixed at the oxide thickness, and further increases in V_g do not change the signal size. The shape of the response does not seem to depend strongly on temperature, which is also to be expected, as the depletion length does not depend strongly on temperature [40].

For the response of the a-Si:H to V_g, the characteristics are qualitatively similar, but there are some unexplained differences. For instance, ΔG_{M-aSi} drops more rapidly with V_g at the highest temperature. Also, at the two lower temperatures, ΔG_{M-aSi} does not seem to saturate at a minimum value. This may be a result of the fact that the center of the a-Si:H strip is very close to the inversion layer. This may inhibit accumulation under this region because a depletion layer must separate the inversion layer and the p-type substrate. Detailed simulations would be required to ascertain whether or not this could be the reason that ΔG_{M-aSi} does not saturate at a minimum value . The temperature dependence of the response to the a-Si:H is also not understood (Fig. 5-9(c)). The gold response ΔG_{M-Au} is roughly temperature independent, as one would expect. However, the response to the a-Si:H $\Delta G_{a-Si:H}$ decreases with increasing temperature. This effect, which along with the increased charging rate Γ ultimately limited our charge transient measurements to temperatures

T< 180 K, is currently not understood.

5.8   a-Si:H IV Characteristics and Noise Correlations

There are other aspects of our charge transient results that are not fully understood. While the measurements of Γ shown in Fig. 5-4 and Fig. 5-5 do not depend strongly on the voltage applied between the a-Si:H contacts V_{ds} for $V_{ds} < 1$, we observe a large nonlinearity in G_{aSi} at room temperature when we measure current I_{aSi} as a function of voltage at $V_{ds} \approx 500$ mV (Fig. 5-10). While the source of this disagreement is unclear, it is possible that at zero bias, either because of surface effects or because of the restricted geometry of the a-Si:H strip, transport through the a-Si:H is limited by the narrowest segment: The charge detection method is probably not sensitive to such effects because it only requires that charge diffuse into the a-Si:H, and not that the charge traverse all of the way from one contact to the other. More interestingly, it is possible that the observed nonlinearity is caused by contact resistance: Our charge sensing method is effective even in the presence of blocking contacts; it can detect charge diffusing toward the contact even with infinite contact resistance as long as there is significant contact capacitance. Future work will seek to demonstrate this explicitly by using samples where there is an insulator between the a-Si:H and the gold contacts.

Also intriguing but not fully understood is the sensitivity of the MOSFET to telegraph noise switches in the a-Si:H. $1/f$ noise and discrete telegraph switches have been observed previously in macroscopic a-Si:H samples [98]. The discrete switching that is sometimes observed presumably occurs for samples where the conductance is dominated by filaments small enough to be affected by a single switch. While the microscopic origin of $1/f$ noise in a-Si:H is unclear, its phenomenology is quite rich. Kakalios and coworkers demonstrated that this noise can be statistically non-Gaussian, having a high degree of correlation between the noise power at different
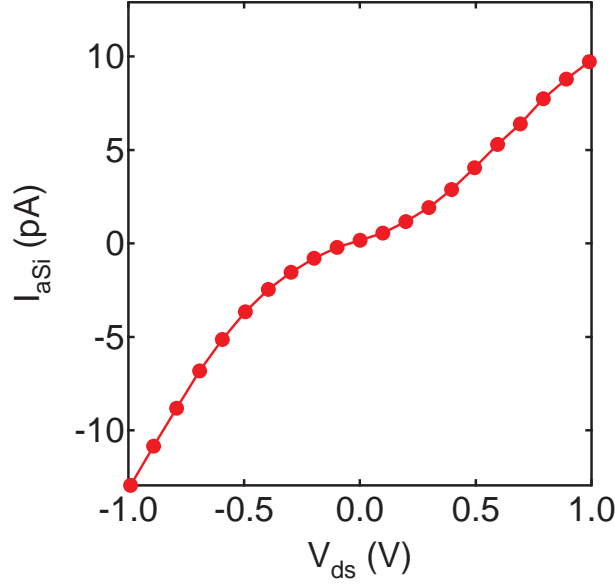
Figure 5-10: Current vs. voltage characteristics of a-Si:H strip, measured at room temperature.

frequencies [99]. The noise becomes Gaussian when the sample is exposed to light, and can be reversibly brought back to the the non-Gaussian state by annealing at 150 C [100]. Reversible changes in the electronic properties of a-Si:H with light exposure and annealing, known in general as the Staebler-Wronski effect [26], are one of the great unsolved problems in the physics of amorphous semiconductors, and lead to limitations on performance thin film solar cells made with amorphous materials. The phenomenology of $1/f$ noise in a-Si:H shows its close connection with the Staebler-Wronski effect.

In Fig. 5-11(b), we show measurements of the current vs. voltage characteristics of the a-Si:H strip at room temperature, in which we observed telegraph noise in the current when $V_{ds} \approx 2$ V. This switching appeared and disappeared apparently randomly, lasting ~ 1 day. Because of the potential sensitivity of our sample to surface effects, it is not clear whether the origin of the telegraph noise we observe is the same as the origin of the noise found in bulk a-Si:H samples. However, because the gate effect for our a-Si:H strip is fairly week, the narrow a-Si:H strip is not very sensitive to nearby electrostatic fluctuations, and it therefore seems likely that the
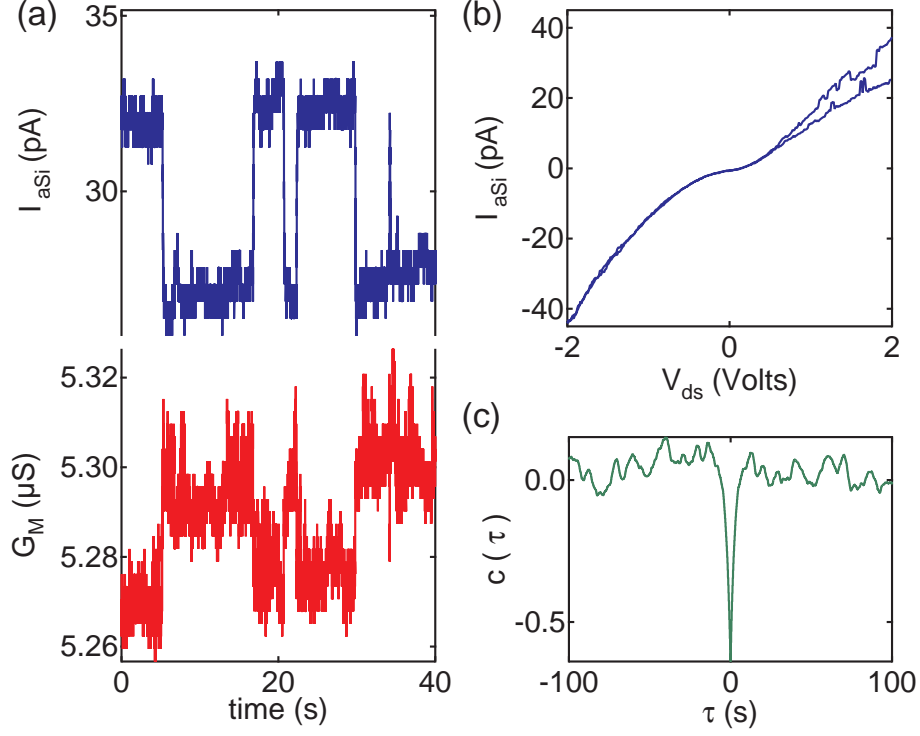
114

Figure 5-11: Noise correlations measured at room temperature. (a) Current through a-Si:H strip I_{aSi} (top panel) and transistor conductance G_M (bottom panel) as a function of time. Here we apply a voltage bias $V_{ds} = 2$ V across the a-Si:H strip. (b) Current vs. voltage characteristics of a-Si:H strip, showing telegraph noise that appears around $V_{ds} = 2$ V. (c) Correlation between I_{aSi} and G_M calculated from data such as that shown in (a), as discussed in the main text.

source of this noise resides inside the a-Si:H strip or on its surface.

In the top panel of Fig. 5-11(a), we plot I_{aSi} measured as a function of time with $V_{ds} \approx 2$ V. We measure G_M as a function of time simultaneously, and this is plotted in the bottom panel of Fig. 5-11(a). We see that the two are *correlated*: When I_{aSi} jumps up, G_M jumps down, and vice versa. This correlation is demonstrated explicitly in Fig. 5-11(c). Here we measure I_{aSi} and G_M simultaneously for a long time T, and compute the correlation function:

$$c(\tau) = \frac{1}{T} \int_0^T \tilde{I}_{aSi}(t + \tau)\tilde{G}_M(t)dt \qquad (5.11)$$

Here we define $\tilde{G}_M(t)$ by subtracting the mean and dividing by the standard deviation $\tilde{G}_M(t) = G(t) - <G(t)>/\sqrt{<G(t)^2> - <G(t)>^2}$, where here the averages

115

are over time. $\tilde{I}_{aSi}$ is defined in the same way. For two completely uncorrelated quantities $f(t)$ and $g(t)$, $c(\tau) = 0$. For two perfectly correlated quantities, where one quantity is a linear function of the other $f(t) = a + bg(t)$, $c(\tau)$ has a peak at $\tau = 0$. One obtains $c(0) = 1$ or -1, depending on whether b, the slope of the linear function, is positive or negative, respectively. We see that for our data $c(\tau)$ has a clear peak at $\tau = 0$ with a value ≈ -0.6, indicating a high degree of correlation between the two signals.

From this, data it is clear that the MOSFET can detect the $1/f$ noise in a material adjacent to it. The mechanism of this detection is not clear. It may be that electrostatic fluctuations that give rise to the switching noise in the a-Si:H current are detected by the MOSFET directly, or that these fluctuations change the charge profile along the a-Si:H strip to which, as we saw in the previous sections, the MOSFET is extremely sensitive. We have observed telegraph noise in the current through strips of a-Si:H other than the one studied here, but these samples were not fabricated adjacent to a MOSFET charge sensor. The intermittency of the switch investigated here made it difficult to study in detail, and more work is required to determine mechanism by which the MOSFET senses these switches.

Chapter 6

Conclusion

To conclude this thesis, here we briefly summarize the results presented in the previous chapters, and then present ideas for possible future experiments. For all of the work presented in this thesis, we used integrated charge sensors to study electron transport in solid state systems. In the first part (Chapters 2 and 3), we used a GaAs quantum point contact as a charge sensor, and studied single-electron tunneling in and out of a GaAs quantum dot. We found that the tunneling was purely elastic, and that the tunnel rate depended exponentially on the energy of the electron relative to the tunnel barrier potential, as expected from a simple quantum mechanical model for the tunneling process. In a large magnetic field, the rate of tunneling into the excited spin state was suppressed relative to the rate of tunneling into the ground spin state. This spin dependent tunneling effect depended on the shape of the electrostatic potential defining the quantum dot, and remains unexplained.

The impact of integrated charge sensors on the study of GaAs quantum dots has been dramatic, enabling a wide range of experiments that would not otherwise be possible. We decided to apply this technique to the study of other solid state systems, to see if it might have a similar impact. For a charge sensor, we fabricated a narrow MOSFET. We used this sensor to study electron transport in a nanometer scale strip of hydrogenated amorphous silicon (a-Si:H). We found that we could use the sensor to measure the resistance of the a-Si:H, even at low temperatures ~ 100 K where the a-Si:H resistance is extremely large ($\sim 10^{17}$ Ω), using fairly small voltages

(~ 1 V). We showed that at higher temperatures, where the a-Si:H resistance is not too large, the resistance obtained from the charge sensing technique was consistent with the resistance obtained from a measurement of current. We used the field effect and dispersive transport to probe the density of localized states in the a-Si:H with the charge sensor, and obtained consistent results.

None of our results for the transport properties of a-Si:H are particularly surprising: They are all consistent with the generally accepted model for electron transport in a-Si:H. What is significant about these results is the manner in which they are obtained: Our integrated charge sensing technique makes it possible to characterize electron transport in highly resistive materials. The results presented in this thesis should therefore enable studies of electron transport in highly resistive materials that would otherwise be impossible. As mentioned in Chapter 5, it is likely that our technique is not sensitive to contact resistance. In order to demonstrate this explicitly, we plan to repeat the experiments discussed in Chapter 5, but with an insulator (for instance, a-Si_3N_4:H) placed between the gold contacts and the a-Si:H. Following such a demonstration, the next next thing to do is to apply our technique to materials with a high electrical resistance or contact resistance in order to elucidate their electronic properties. As mentioned in Chapter 5, there are in fact a variety of highly resistive materials for which a study of electron transport would be valuable, and we now review a few of these materials in detail.

Electron transport through arrays of semiconducting nanocrystals has been studied intensively by our group [97, 101] and by others [102, 103, 104]. The nanocrystals consist of a semiconducting (PbSe) sphere with a diameter ~ 5 nm, capped by an organic material (oleic acid) of thickness ~ 1 nm, which serves to passivate the surface of the semiconducting sphere, and to insulate the sphere from its surroundings. These nanocrystals can be deposited from solution onto the surface of a chip, and if this is done properly, they self-assemble into an ordered array [105]. It is possible that electron transport through such ordered arrays exhibits collective effects caused by Coulomb interactions [106, 107]. However, transport studies of these arrays are inhibited by their extremely large electrical resistance. There have been a number

of studies in which the high electrical resistance is ameliorated by either chemical treatments [102] or by annealing [97]. However, in addition to lowering the electrical resistance, these procedures have other undesired effects. For instance, in Mentzel *et al.* [97], we found that annealing arrays of nanocrystals increases their conductivity, as the average distance between neighboring nanocrystals is reduced. However, the films also become highly disordered following the annealing procedure, as is clearly observable from transmission electron microscope images and grazing incidence small angle X-ray scattering data. We found that transport through such disordered arrays can be described in terms of a simple model where the conduction occurs via activation of holes from above the Fermi level to the valence band, and electron-electron interactions are ignored.

With our charge sensing technique, we may be able to probe nanocrystal arrays without subjecting them to annealing or chemical treatments, and thus may be able to observe more interesting forms of electron transport. Our technique would also be interesting to apply to annealed or chemically treated arrays of nanocrystals. Hu and collaborators [108] demonstrated, using a combination of electrostatic force microscopy and transmission electron microscopy, that for annealed arrays of PbSe nanocrystals, the transport can be limited by cracks in the film that form during the annealing process. Therefore, traditional transport measurements do not necessarily probe the bulk properties of PbSe nanocrystal arrays, but rather measure the resistance of thin ($\sim$ 20 nm) wide channels of NC's that extend across the cracks [108]. Our technique is not sensitive to cracks in the film, as it does not require that a continuous path of nanocrystals extend from one contact to the other, but rather simply that a section of nanocrystals is charged by the contact.

Another class of materials of technological relevance to which our technique could be applied with great effect are high dielectric constant materials. As integrated circuits (IC's) have gotten smaller, the required oxide thickness for MOSFETs in state of the art IC's has shrunk to only a few nanometers. Quantum mechanical tunneling through these thin gate oxides results in a large leakage from the gate to the inversion layer. This leakage increases power consumption, and over the last decade

has been one of the main roadblocks to reducing the size of IC's. The solution that has been embraced by the engineering community (and has recently been implemented by the large IC manufactures [109]) is to replace SiO_2 with a higher dielectric constant material. With a higher dielectric constant, one can use a thicker insulator while maintaining the same gate capacitance, resulting in an exponential decrease in the tunneling current.

There are a number of issues with high dielectric constant materials, but one of the main problems is a high bulk defect density as compared to SiO_2 [109]. Consider for example HfO_2. This material, along with a number of other transition metal oxides, has a high dielectric constant owing to optical phonons [109], and is therefore a good candidate to replace SiO_2. However, it is of inferior electrical quality, for the following reasons: First, it cannot be grown thermally on silicon like SiO_2, and therefore must be deposited, using a method such as atomic layer deposition (ALD). In general, deposited oxides have higher defect densities than oxides grown thermally. The second reason that HfO_2 has a high defect density is that is has a high number of bonds per atom, or coordination number. SiO_2 has a low coordination number, and this allows it to easily relieve structural defects. HfO_2 cannot relax so easily, and therefore remains in a nonequilibrium, high defect density state [109]. The defects in HfO_2 result in a number of problems, including gate leakage that is higher than the expected tunneling contribution, and hysteresis in the transistor threshold voltage caused by the filling and emptying of traps in the HfO_2. The methods developed in this thesis are well suited to studying electron transport in these highly resistive dielectrics, and may thereby be able to shed light on the nature of the defects in these materials. A preliminary experiment would be straightforward to implement, as high dielectric constant materials are typically deposited with ALD at $T \sim 200$ C, and can probably be patterned using the same process we used to pattern a-Si:H.

Other than for measuring charge transport in resistive materials, there may be a number of additional ways that integrated charge sensors can be used to elucidate the electronic properties of solids. As a specific example, consider resistive memory devices, for which the resistance of a transition metal oxide (for instance $SrTiO_3$) is

switched between an insulating and metallic state by the application of an electric field, providing a new form of nonvolatile electronic memory [110]. The mechanism for this switching phenomena is not fully understood, but it is observed in a wide range of materials [110]. One common feature of all materials exhibiting this effect is the presence of mobile oxygen vacancies, and a particularly promising theory involves the motion of oxygen vacancies, which are positively charged, in response to the electric field [111]. It is possible that using integrated charge sensing we could observe the motion of the charged oxygen vacancies directly.

It is thus possible that charge sensing will make contributions to a wide variety of disciplines. Our group has already begun the steps necessary to apply this technique to the study of electron transport in nanocrystal arrays, where hopefully it will have a large impact.

Appendix A

Fabrication

In this appendix, we give the details for our fabrication sequences. We give step by step fabrication sequences in A.1 and A.2 for our MOSFET and a-Si:H strip respectively. In A.3 we outline our electron beam lithography alignment procedure. All of the MOSFET fabrication steps were performed in either in the Microsystems Technology Laboratory (MTL) (in either the Integrated Circuits Laboratory (ICL) or Technology Research Laboratory (TRL)) or in the Scanning Electron Beams Laboratory (SEBL), at MIT. For the a-Si:H patterning, we also used the Raith electron beam lithography system at Harvard. This system is essentially identical to the Raith 150 in the SEBL facility, and we used the Harvard facility simply because the SEBL facility is very heavily used. A few minor processing steps (cleaning in solvents, lift-off, ect.) were performed in our laboratory or in Moungi Bawendi's laboratory.

A.1 MOSFET Fabrication

The starting material is a "prime" electrical quality p-type silicon 6 inch wafer, doped with Boron $N_B \approx 3 \times 10^{15}$ cm^{-3}, which we purchased from the MTL. The doping level can be verified by measuring the resistivity of the wafers $\rho \approx 5$ Ω·cm with the ICL's four point probe.

1. **RCA Clean:** This is an industry standard pre-diffussion cleaning procedure, and was done at the ICL's RCA cleaning station. The wafers should be

transferred to the furnace immediately after this step. The three cleaning steps are supposed to remove organic contamination, native oxide, and ionic contamination, respectively.

SC1 5:1:1 H_2O:H_2O_2:NH_4OH, T = 80 C, time = 10 minutes. Rinse.

HF Dip 50:1 H_2O:HF, 1 minute. Rinse.

SC2 5:1:1 H_2O:H_2O_2:HCL, T = 80 C, time = 15 minutes. Rinse. Spin Dry.

2. **Grow Field Oxide:** We grow a wet thermal oxide. We use the furnace 5D-ThickOx in the ICL's diffusion system (Thermco 10K 4 Furnace Systems). The recipe (4W1000, time = 127 minutes) grows $\approx$ 650 nm of wet thermal oxide at T = 1000 C. The oxide thickness can be verified using ellipsometry, using the ICL's UV1280.

3. **Photolithography:** Pattern the field oxide for etching. This step was accomplished using the ICL's coater/developer track (SSI 150) and optical stepper (Nikon NSR-2005i9, an i-line optical stepper). The masks are chrome on quartz. Some of our masks were patterned for us in the MTL by Dennis Ward, and some we purchased from external vendors. This first photolithography step along with the following etch, in addition to patterning the active region of the device, patterns alignment marks in the oxide to which all subsequent photolithography steps, as well as the e-beam lithography step for patterning the gate, are aligned.

   Coat Use recipe T1HMDS, which, after application of the adhesion promoter HMDS, coats the wafers in $\approx$ 1 μm of SPR 700-1.0 (from Shipley), a positive photoresist.

   Expose Exposure time $\approx$ 165 ms (the lamp intensity is $\approx$ 475 mW/cm^2).

   Develop Use recipe Dev6. This develops the wafers in LDD-26W, a TMAH based developer.

4. **Etch Oxide:** Etch the field oxide in BOE (Buffered oxide etchant, a mixture of HF and NH_4OH) at the ICL's oxide etch wet bench. The etch rate should be ≈ 1.5 nm/s, and we overetch by about 20 percent. The etch rate can be measured using ellipsometry, and we can verify that all of the oxide in the active regions has been removed using ellipsometry as well.

5. **Strip Photoresist:** Using the ICL's O_2 plasma asher (Matrix 106 Plasma Asher). This takes about 3 minutes, depending on the power.

6. **Photolithography:** As above. This step defines the ion implanted regions.

7. **Ion Implantation:** We send the wafers out to Innovian for ion implantation. They should be doped with phosphorous. The dose is 3.5×10^{15} cm^{-2}, at an energy of 180 keV. We asked for a small, 7 degree tilt, which is supposed to prevent "channeling" of the phosphorous along the crystalline axes of the wafer. For a few of our wafers, we had trouble contacting the implanted regions. This may have been caused by damage to the silicon surface during the ion implantation. We never fully resolved this problem, but for future processes, it is probably advisable to have the phosphorous implanted only after growing a thin (~ 10 nm) protective oxide on the wafer.

8. **Strip Photoresist/Clean:** Following ion implantation, the photoresist is much harder to remove than it normally is. It can be removed using a combination of Piranha cleans (3:1 H_2SO_4:H_2O_2) and O_2 plasma ashing such as:

   Piranha resist strip Blue pre-metal wet station in the ICL, 10 minutes. Rinse. Dry.

   O_2 Plasma clean In the ICL's O_2 plasma asher. Repeat until photoresist is gone.

   Piranha clean Green pre-metal wet station in the ICL, 10 minutes. Rinse. Dry.

9. **RCA Clean:** As above.

10. **Grow gate oxide:** We use furnace 5A-GateOx in the ICL's diffusion system. The recipe (1D1000, time = 160 minutes) grows 100 nm of oxide at T = 1000 C. It also serves as to "drive in" the donor atoms: The phosphorous atoms diffuse ~ 0.5 μm into the substrate during this step. In order that the donor atoms do not diffuse out of the silicon as the furnace ramps up, the recipe can be changed to introduce O_2 into the furnace as it ramps up, thus growing a thin layer of oxide during the ramp up that prevents this so called ex-diffusion of phosphorous out of the silicon substrate. To doping level can be verified by etching off the oxide and measuring the resistivity of the ion implanted silicon with the ICL's four point probe. From this measurement we obtain $N_P \sim 10^{20}$ cm^{-3}, so the implanted regions are degenerately doped. The thickness of the oxide can be verified with ellipsometry, and it is generally thicker over the ion implanted regions. This is to be expected, as oxidation of silicon is known to speed up with doping. After the oxide has been grown the wafers should be transferred directly to the polysilicon deposition furnace.

11. $n+$ **Polysilicon Deposition:** We perform LPCVD (Low pressure chemical vapor deposition) in the ICL's diffusion system. The recipe (560Doped PH3 Flat, time = 77 minutes, in tube 6A-npoly) deposits 80 - 100 nm of phosphorous doped polysilicon at T = 560 C. The wafers should be stored in a clean box until proceeding to the next step, which, because of time constraints, must usually be performed the following morning.

12. **Polysilicon Anneal:** This step is intended to reduce the grain boundaries in the polysilicon and thereby increase its conductivity. The recipe anneals at 900 C (We used the furnace 5B-Anneal in the ICL, 2A900, time = 30 minutes). This step changes the polysilicon color from brown to green. The resistivity of the polysilicon can be measured with the ICL's four point probe after this step, and for our wafers it is ≈ 1.5 mΩ·cm.

13. **Polysilicon Oxide Strip:** We etch the wafers in BOE as above for ≈ 1 minute. This removes a somewhat thick ~ 5 nm layer of oxide that is on top the polysilicon after the anneal. This step is actually very important because the reactive ion etching used for etching the polysilicon gate etches oxide very slowly, and with such a thick native oxide the time required for the gate etch becomes longer than it should be, and varies from wafer to wafer. We attributed this rather thick native oxide to the fact that the wafers are still quite hot as they are unloaded into atmosphere from the annealing furnace. This thick native oxide did not grow back on the polysilicon at room temperature, even when the wafers were left sitting in atmosphere for days.

14. **Backside Etch:** We remove the polysilicon and oxide from the back of the wafer so that the p-type substrate can be grounded during the e-beam lithography. First the wafers are coated in photoresist on the coater track as described above (to protect the front side of the wafer), and then etched upside down using two magnetically enhanced reactive ion-etching steps in the AME Model P5000 to etch polysilicon (recipe Keith CP, with the main step time $t_m \approx 60$ s as described in the gate etching step below) and oxide etch (such as Baseline Ox New for 90 seconds). These etch steps are not very sensitive, so we overetched quite a bit. The resist is then be stripped in O_2 plasma as described above.

15. **Spin E-beam Resist (HSQ):** We use the negative e-beam resist XR-1541 4 HSQ from Dow Corning. This should be kept refrigerated, or else it degrades, and generally you must spin it on the wafer not too long before the e-beam write (the night before seemed to be all right). We deposit the HSQ on the wafer with a pipet, spread at 500 rpm for 5 seconds, and then spin at 2000 rpm for 45 seconds (We use the PMMA spinner in the TRL Heidelberg room. We did not precisely measure the HSQ film thickness, but it is between 55 and 115 nm). After spinning the wafer is baked on a hot plate for 2 minutes at T = 110 C.

16. **Electron Beam Lithography:** We use SEBL's Raith 150. See Appendix A.3 for details. We wrote the pattern at a voltage of 30 kV, with a step size = 10 nm, field size = 100 μm. The required dose is discussed in A.3.

17. **Develop HSQ:** 5 minutes in LDD 26W, followed by a rinse in water (this is done in the TRL photo area). The patterns were imaged using the ICL's Zeiss electron microscope after this step or after the next.

18. **Post E-beam Clean:** We were required by standard ICL procedure to clean the wafers after electron beam lithography before they could be processed further in the ICL. We performed a Piranha clean (as above) followed by an SC1 clean (as in the RCA clean description above, but at room temperature) at the TRL acid wet bench.

19. **Photolithography:** As described above. This step patterns the larger parts of the polysilicon gate.

20. **Polysilicon Etch:** This step uses magnetically enhanced reactive ion etching in a chlorine/hydrogen bromide chemistry to create the polysilicon gate of the transistor. The time of the main step must be long enough to etch through the polysilicon but not too long or it will etch through the HSQ. We use the recipe Keith CP with the following parameters:

    Gas Stabilization for Main Etch Cl_2:HBr 20:20 Scc, $t_{s,m} = 10$ s, P = 200 mtorr, B = 50 Gauss, RF = 0.

    Main Etch Cl_2:HBr 20:20 Scc, $t_m = 24$ sec, P = 200 mtorr, B = 50 Gauss, RF = 350 W.

    Gas Stabilization for Overetch HBr 40 Scc, $t_{s,o} = 10$ s, P = 100 mtorr, B = 50 Gauss, RF = 0

    Short Overetch Step HBr 40 Scc, $t_o = 40$ s, P = 100 mtorr, B = 50 Gauss, RF = 50 W.

21. **Strip Photoresist:** As described above.

22. **Photolithography:** As described above. This step covers the polysilicon gate structure in photoresist, for the "stringer" etch.

23. **"Stringer" Etch:** This step is designed to remove any remaining polysilicon from the edge of the active region, as described in Section 4.2. We use the same recipe used for the gate etch step, except that the times for the first two steps are set to zero, $(t_{s,m} = t_m = 0)$ and the overetch time is made very long $(t_o = 900 \text{ s})$. The overetch step does not etch the oxide at a measurable rate.

24. **Strip Photoresist:** As described above.

25. **HF Dip:** 50:1 H_2O:HF, in the TRL acid hood. This is intended to remove the HSQ before the next oxide growth step. This step will only remove a little bit of the oxide (6 ± 3 nm). The mixture must be stirred very well in order to achieve a uniform etch.

26. **Long Piranha Clean:** 20 minute piranha clean (Done in the green pre-metal ICL station). We did this instead of an RCA clean because the complete RCA process would etch the exposed gate oxide.

27. **Grow Thin Oxide:** We use furnace 5C-FieldOx in the ICL diffusion system. The recipe (3D900, time = 25 minutes) grows a thin (~ 10 nm) dry thermal oxide at 900 C. This was intended to insulate the polysilicon from anything it might touch, but ended up being unnecessary for the experiments discussed in this thesis.

28. **Photolithography:** As described above. This step patterns holes to be etched in the oxide over the polysilicon gate and ion implanted regions through which electrical contact will be made.

29. **Etch Oxide:** As described above, with an etch time = 4 minutes (this should overetch quite a bit, which is all right for this step). Ellipsometry can be used to verify that no oxide remains above the implanted region where the holes have been etched. One should proceed immediately (e.g. within an hour) to

the following three steps in order to minimize the native oxide that will grow where the holes have been etched. For a wafer that we had trouble making contact to, we noticed that the ellipsometer registered a small ($\sim$ 3 nm) thickness of oxide that did not go away even for very long BOE etches. This may have been caused by damage of the silicon surface by the ion implantation process.

30. **Strip Photoresist:** As described above.

31. **Long Pirhana Clean:** As described above. This is a standard pre-metal deposition clean.

32. **Deposit Aluminum:** We sputter 250 nm of pure aluminum (using the Applied Materials Endura system). Originally, we were advised to deposit an Al-Si alloy, which is a standard procedure used to minimize "spiking" of Al into the substrate. Spiking occurs because silicon is soluble in aluminum, so that pits form in the silicon as it is dissolved into the Al during the sintering step. This can cause electrical shorts between the ion-implanted regions and the substrate. This is alleviated by using aluminum that is already saturated with silicon. For our process however, this caused problems. The standard PAN etch of aluminum (see below) will not remove silicon, so after this etch there are small so called "freckles", small bits of silicon visible with an electron microscope. Using any silicon etch to remove these would risk damaging the gate. We therefore switched to pure aluminum. We did not encounter any serious problems caused by spiking, probably because our phosphorous implanted regions were fairly deep ($\sim$ 0.5 μm), so only very deep spikes could cause problems.

33. **Photolithography:** As described above. This step patterns the Aluminum for the etch in the following step.

34. **PAN Etch** We etch the aluminum in standard PAN aluminum etchant (H_3PO_4:CH_3COOH:HNO_3:H_2O, 16:1:1:2). Because the of the native oxide on

the aluminum, the time required for the PAN etch can vary a bit. We typically etched for $\sim$ 5 minutes, and examined the wafers afterward to verify there was no aluminum remaining.

35. **Strip Photoresist:** As described above.

36. **Sinter contacts:** Standard thermal process for improving the contact between aluminum and doped silicon. The wafer is sintered at 425 C for about 30 minutes in forming gas (This is done in the TRL's sintering tube).

37. **Photolithography with Negative Resist:** This step is to pattern the gold a-Si:H contacts using a lift off process. We used the negative resist NR9-1000PY from Futurex:

    Bake 20 minutes at 120 C to remove moisture (TRL photo area oven).

    Coat Deposit two pipets of the resist on the wafers, spread at 750 rpm for 6s, spin at 5000 rpm for 40 s. The resist will be 1 μm thick (TRL coater).

    Soft Bake The recipe (NR9soft, on the ICL coater track) bakes the wafer at 150 C for 1 minute.

    Expose Exposure time $\approx$ 650 ms.

    Develop Use LDD-26W (using the recipe NR9dev on the ICL's developer track).

38. **Electron Beam Evaporation:** Deposit Titanium/Gold contacts (30 nm Ti deposited at 0.1 nm/s, 300 nm Au deposited at 0.3 nm/s. You can use either of the TRL's two electron-beam evaporators).

39. **Lift-Off:** Put the wafers in acetone and leave for a few hours. One can also spray with acetone to speed the process up a bit. Clean in methanol and water and spin dry.

40. **Spin Photoresist:** We spin on the photoresist AZ514 at 3000 rpm and bake it at 90 C (in the TRL photo area), which serves to protect the wafer surface

during diesaw step. Any type of resist of sufficient thickness should be appropriate for this step.

41. **Diesaw:** Wafers are diced up with a 220 μm silicon blade.

A.2   a-Si:H Patterning

1. **Scratch Back of Chip:** Use a diamond scribe. This is so that the chip is well grounded during the electron beam lithography.

2. **Clean Chip:** The chips have been coated with a protective layer photoresist for the diesaw procedure. Before spinning e-beam resist onto the chips, we need to clean this photoresist off.

   Acetone Soak chip acetone $\sim$ 7 minutes.

   Methanol Sonicate chip in methanol $\sim$ 3 minutes. Blow dry in Helium or Nitrogen.

3. **Spin on PMGI:** We use PMGI SF3, from Microchem. We spin fairly fast in this step: For small chips, it is important to do this, as it reduces the edge bead. For the steps used here, we spin two layers of PMGI, and the total thickness is $\approx$ 120 nm (measured by cleaving a chip and imaging it from the side with an SEM). It is good to examine the chips carefully after spinning, to check that the resist looks to be of a uniform thickness.

   Bake 1 180 C, $\sim$ 1 minute, to drive off any moisture.

   Spin 1 Deposit PMGI on chip with a pipet. Put the chip on top of a fab wipe, and allow the PMGI it to run down over the sides of the chip. Spin at 6000 rpm for 1 minute.

   Bake 2 180 C, 5 minutes.

   Spin 2 Repeat the step Spin 1 above.

   Bake 3 250 C, 15 minutes.

4. **Electron Beam Lithography:** For this step we use SEBL's Raith 150, or the Raith at Harvard. See Appendix A.3 for details. For PMGI spun and developed using the conditions described here, we wrote at 10 kV with a dose of 520 μC/cm^2, with a step size = 10 nm, and a field size = 100 μm.

5. **Develop Pattern:**

   Develop MIBK:IPA 1:3 (premixed), 5 minutes while sonicating.

   Clean IPA, for a minute or so.

6. **O$_2$ Plasma Clean:** The exact parameters one should use depend on the plasma cleaner. Using a small table top plasma cleaner, which typically remove organic material at a rate $\sim$ 10 nm /minute, we plasma cleaned for 50 seconds: Not so long that it changes the PMGI thickness appreciably, but long enough to remove any small amount of organic material left where the developer has removed the PMGI.

7. **a-Si:H Deposition:** Using the STS CVD in the TRL: Time = 5 minutes, Temperature = 200 C, Pressure = 300 mtorr, SiH$_4$ flow rate = 80 Sccm, PH$_3$ (2 % in H$_2$) flow rate = 80 Sccm, RF Power = 20 W, RF frequency = 13.56 MHz. This will deposit $\approx$ 50 nm of a-Si:H heavily doped with phosphorous. After the chips are unloaded from the deposition chamber, wait $\approx$ 3 minutes to let them cool a little before opening the load lock (to reduce oxidation).

8. **Lift Off:**

   Scratch Scratch the a-Si:H at the four corners of the chip with a diamond scribe, so that the remover can get under the a-Si:H.

   Soak Soak in PG remover from Microchem (PG remover is just NMP with a proprietary surfactant added to it) at 70 - 80 C for > 12 hours.

   Lift off Sonicate in the hot PG remover for $\sim$ 2 minutes.

   Clean Acetone, methanol, and IPA. Blow dry

A.3   Electron Beam Lithography

For our process, the most difficult parts of the electron beam lithography procedure were properly processing HSQ in order to pattern the polysilicon gate, and aligning the a-Si:H strip to the polysilicon gate with high precision. Here we briefly discuss the HSQ procedure, and then give a detailed description of our alignment procedure. Our description of the alignment procedure will be in general terms, details on how to manipulate the Raith software can be found elsewhere [112]. For our device (Fig. 4-6), our polysilicon gate tapers down to a width of 60 nm. While patterning a 60 nm line in HSQ is fairly straightforward, patterning wide sections that rapidly taper down to a 60 nm width is more difficult, a result of the proximity effect. This effect can be understood as follows. Suppose we wish to write a small square box (Box S), located just to the side of a larger box (Box L). We first expose Box S to a give electron dose. When we subsequently expose Box L, electrons in the substrate excited by the incoming beam, called secondary electrons, as well as x-ray radiation caused by the relaxation of high energy electrons [113], will impinge upon Box S. The amount of this excess exposure depends on the area of Box L, and how far away it is from Box S. Points within Box L will receive even more extra exposure than points within Box S, and, in general, the result of the proximity effect is that larger features are overexposed. A commonly used technique for ameliorating this effect is to expose the smaller regions to a relatively higher dose [113]. We exposed the narrowest portion of our gate to a dose of 430 μC/cm^2, the regions adjacent to the narrowest portion to a dose of 295 μC/cm^2, and the larger regions to a dose of 245 μC/cm^2. Our device geometry required precise alignment of the a-Si:H pattern to the MOSFET gate. The Raith is perfectly capable of accomplishing this, but the procedure for doing so is somewhat complicated. Here we give a step by step process for the electron beam lithography procedure we used for patterning the a-Si:H.

1. **NC Deposition:** Put gold 100 nm nanocrystals on the chip with a micro pipet tool. We use these nanocrystals to focus and stigmate the electron

beam. Obviously, you should not put the nanocrystals exactly where you want to write. But you should put them fairly close (a few mm away) because the focus can change when you move large distances. Also, you should not put the nanocrystals on top of the chip where the resist is thick (i.e. the edge bead), because if the resist under the nanocrystals is too thick, it will charge up and make it impossible to focus the beam properly.

2. **Load Sample, Etc:** The procedure for loading is different at Harvard and MIT. We usually made sure to to align the axes of our sample pattern (to which we align the write) with the axes of the Raith stage (by eye), as this generally makes navigating the wafer a lot easier. We followed the appropriate loading procedures at Harvard and MIT, and usually measured the beam current at the Faraday cup immediately after loading.

3. **Focus, Stigmate, and Aperature Align Using Gold NC's:** The times when this was particularly difficult, the problem was usually that the sample was not well grounded, or that the resist under the NC's is too thick and is charging. If a grounding problem is suspected, it is useful to diagnose whether the wafer chuck is not contacting the stage or the chip is not contacting the wafer chuck. If the current measured at the Faraday cup is zero or unstable, then wafer chuck is not grounded to the stage (this frequently happened at MIT), and the best thing to do is run a routine that shakes the stage, which hopefully moves the chuck into a position where it is better grounded. If the current is a reasonable, stable value, then the best thing to do is unload the sample and scratch the back of it. Immediately after focusing etc., it is good to check that the image does not drift too much (it should drift less than ~ 50 nm in 5 minutes), a problem that is also usually caused by a build up of charge on the sample, which deflects the beam.

4. **Field Alignment:** Use gold NC's. We used the standard field alignment procedure at MIT, and at Harvard we used the standard field alignment procedure three times with the step size set to 20,5, and 1 μm. This step

aligns the electron beam to the stage axes.

5. **UV Alignment:** Perform either a UV coordinate transformation, or an angle correction followed by an origin correction. It is important to use the Raith side for imaging during this step. The SEM view is not necessarily perfectly aligned with the Raith side view, so if you align using the SEM view, the Raith will not write in the proper location.

6. **Field Alignment:** As above. Following the UV alignment, the axes of the electron beam are not aligned to your UV axes, and this step is needed to perform the alignment.

7. **Origin Correction:** Following this step, the alignment is complete. To check how good the alignment is, drive to the UV coordinates of any feature of the design to which you are aligned and image it on the Raith side.

Bibliography

[1] W. Shockley, J. Bardeen, and W. Brattain, Science **108**, 678 (1948).

[2] B. G. Streetman and S. Banerjee, *Sold State Electronic Devices* (Prentice Hall, 2000).

[3] R. C. Ashoori, Nature **379**, 413 (1996).

[4] R. C. Ashoori, H. L. Stormer, J. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Phys. Rev. Lett. **71**, 613 (1993).

[5] L. P. Kouwenhoven, T. H. Oosterkamp, M. W. Danoesastro, M. Eto, D. Austing, T. Honda, and S. Tarucha, Science **278**, 1788 (1997).

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

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

[8] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R. Huang, J. Majer, S. Kumar, S. M. Girvin, and R. J. Schoelkopf, Nature **431**, 162 (2004).

[9] M. A. Kastner, Rev. Mod. Phys. **64**, 849 (1992).

[10] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. Kastner, Nature **391**, 156 (1998).

[11] S. M. Cronenwett, T. H. OosterKamp, and L. P. Kouwenhoven, Science **281**, 540 (1998).

[12] M. A. Kastner, R. F. Kwasnick, J. C. Licini, and D. J. Bishop, Phys. Rev. B **36**, 8015 (1987).

[13] K. J. Thomas, J. T. Nicholls, M. Y. Simmons, M. Pepper, D. R. Mace, and D. A. Ritchie, Phys. Rev. Lett. **77**, 135 (1996).

[14] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, Phys. Rev. Lett. **52**, 228 (1984).

[15] M. J. Kirton and M. J. Uren, Advances in Physics **38**, 367 (1989).

[16] M. Pioro-Ladriére, J. H. Davies, A. R. Long, A. S. Sachrajda, L. Gaudreau, P. Zawadzki, J. Lapointe, J. Gupta, Z. Wasilewski, and S. Studenikin, Phys. Rev. B **72**, 115331 (2005).

[17] D. H. Cobden and B. A. Muzykantskii, Phys. Rev. Lett. **75**, 4274 (1995).

[18] M. Xiao, L. Martin, E. Yablonovitch, and H. W. Jiang, Nature **430**, 435 (2004).

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

[20] L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M. Westervelt, and N. S. Wingreen, Electron transport in quantum dots, in *Mesoscopic Electron Transport*, edited by L. Sohn, L. Kouwenhoven, and G. Schon, NATO ASI Series E 345, pp. 105–214, Kluwer Acadenic Publishers, 1997.

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

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

[23] S. Gustavsson, R. Leturcq, B. Simovič, R. Schleser, T. Ihn, P. Studerus, and K. Ensslin, Phys. Rev. Lett. **96**, 076605 (2006).

[24] T. Fujisawa, T. Hayashi, R. Tomita, and Y. Hirayama, Science **312**, 1634 (2006).

[25] S. Amasha, K. MacLean, I. P. Radu, D. M. D. M. Zumbühl, M. A. Kastner, M. Hanson, and A. C. Gossard, Phys. Rev. Lett. **100**, 046803 (2008).

[26] R. A. Street, *Hydrogenated Amorphous Silicon* (Cambridge University Press, 1991).

[27] S. Amasha, *Electron Tunneling and Spin Relaxation in a Lateral Quantum Dot*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 2008.

[28] T. Ihn, S. Gustavsson, U. Gasser, B. Küng, T. Müller, R. Schleser, M. Sigrist, I. Shorubalko, R. Leturcq, and K. Ensslin, Solid State Communications **149**, 1419 (2009).

[29] I. P. Radu, *Electron Transport in Confined Structures in Very High Mobility GaAs in Perpendicular Magnetic Fields*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 2009.

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

[31] H. Churchill, F. Kuemmeth, J. W. Harlow, A. J. Bestwick, E. I. Rashba, K. Flensberg, C. H. Stwertka, T. Taychatanapat, S. Watson, and C. M. Marcus, Phys. Rev. Lett. **102**, 166802 (2009).

[32] G. Gotz, G. A. Steele, W. Vos, and L. P. Kouwenhoven, Nanoletters **8**, 4039 (2008).

[33] C. B. Murray, D. J. Norris, and M. G. Bawendi, J. Am. Chem. Soc. **115**, 8706 (1993).

[34] S. Kim, J. P. Zimmer, S. Ohnishi, J. B. Tracy, J. V. Frangioni, and M. G. Bawendi, J. Am. Chem. Soc. **127**, 10526 (2005).

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

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

[37] T. Hayashi, T. Fujisawa, H. D. Cheong, Y. H. Jeong, and Y. Hirayama, Phys. Rev. Lett. **91**, 226804 (2003).

[38] G. Granger, *Spin Effects in Single-Electron Transistors*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 2005.

[39] J. H. Davies, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, 1998).

[40] S. M. Sze, *Physics of Semiconductor Devices* (Wiley-Interscience, 2007).

[41] R. E. Collin, *Foundations for Microwave Engineering* (Wiley-Interscience, 2001).

[42] R. Schleser, E. Ruh, T. Ihn, K. Ensslin, D. C. Driscoll, and A. C. Gossard, Appl. Phys. Lett **85**, 2005 (2004).

[43] O. Naaman and J. Aumentado, Physical Review Letters **96**, 100201 (2006).

[44] L. D. Landau and E. Lifshitz, *Quantum Mechanics* (Butterworth Heinemann, 1991).

[45] D. H. Cobden, M. Bockrath, P. L. McEuen, A. G. Rinzler, and R. E. Smalley, Phys. Rev. Lett. **81**, 681 (1998).

[46] J. C. Chen, Z. An, T. Ueda, S. Komiyama, K. Hirakawa, and V. Antonov, Phys. Rev. B **74**, 045321 (2006).

[47] S. Gustavsson, R. Leturcq, B. Simovič, R. Schleser, P. Studerus, T. Ihn, K. Ensslin, D. C. Driscoll, and A. C. Gossard, Phys. Rev. B **74**, 105305 (2006).

[48] U. Bockelmann, Phys. Rev. B **50**, 17271 (1994).

[49] S. Gustavsson, R. Leturcq, B. Simovic, R. Schleser, P. Studerus, T. Ihn, and K. Ensslin, cond-mat/0605365, 2006.

[50] J. I. Climente, A. Bertoni, G. Goldoni, and E. Molinari, Phys. Rev. B **74**, 035313 (2006).

[51] S. Amasha, K. MacLean, I. P. Radu, D. M. D. M. Zumbühl, M. A. Kastner, M. Hanson, and A. C. Gossard, Phys. Rev. B **78**, 041306 (2008).

[52] R. Hanson, B. Witkamp, L. M. K. Vandersypen, L. H. W. van Beveren, J. M. Elzerman, and L. P. Kouwenhoven, Phys. Rev. Lett. **91**, 196802 (2003).

[53] C. Kittel, *Introduction to Solid State Physics* (John Wiley and Sons, Inc., 2005).

[54] R. C. Ashoori, H. L. Stormer, J. S. Weiner, L. N. Pfeiffer, S. J. Pearton, K. W. Baldwin, and K. W. West, Phys. Rev. Lett. **68**, 3088 (1992).

[55] M. S. Rudner and B. I. Halperin, private communication.

[56] E. B. Foxman, P. L. McEuen, U. Meirav, N. S. Wingreen, Y. Meir, P. A. Belk, N. R. Belk, and M. A. Kastner, Phys. Rev. B **47**, 10020 (1993).

[57] S. D. Franceschi, S. Sasaki, J. M. Elzerman, W. G. van der Wiel, S. Tarucha, and L. P. Kouwenhoven, Phys. Rev. Lett. **86**, 878 (2001).

[58] R. Leturcq, D. Graf, T. Ihn, K. Ensslin, D. D. Driscoll, and A. C. Gossard, Europhys. Lett. **67**, 439 (2004).

[59] M. P. Marder, *Condensed Matter Physics* (Wiley-Interscience, 2000).

[60] K. Ensslin, private communication.

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

[62] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[63] R. Williams, *Modern GaAs Processing Methods* (Artech House, 1990).

[64] S. A. Campbell, *The Science and Engineering of Microelectronic Fabrication* (Oxford University Press, 2001).

[65] A. E. Grigorescu and C. W. Hagen, Nanotechnology **20**, 292001 (2009).

[66] A. Madan, P. L. Comber, and W. E. Spear, J. Non-Cryst. Solids **20**, 239 (1976).

[67] W. Beyer and H. Overhof, Semiconductors and Semimetals **21**, 257 (1984).

[68] B. Cui and T. Veres, Micro. Eng. **85**, 810 (2008).

[69] B. Cord, C. Dames, and K. Berggren, J. Vac. Sci. Technol. B **24**, 3139 (2006).

[70] H. Fritzsche, Semiconductors and Semimetals **21**, 309 (1984).

[71] P. Horowitz and W. Hill, *The Art of Electronics* (Cambridge University Press, 1985).

[72] R. F. Kwasnick, *Electronic Conduction in Extremely Narrow Silicon Inversion Layers*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 1984.

[73] J. C. Licini, *Quantum Conductance Fluctuations in Extremely Narrow Inversion Layers in Silicon*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 1987.

[74] E. Kamieniecki, J. Appl. Phys. **54**, 6481 (1983).

[75] D. H. Cobden, M. J. Uren, and M. Pepper, Phys. Rev. Lett. **71**, 4230 (1993).

[76] M. Pioro-Ladrière, J. H. Davies, A. R. Long, A. S. Sachrajda, L. Gaudreau, P. Zawadzki, J. Lapointe, J. Gupta, Z. Wasilewski, and S. Studenikin, Phys. Rev. B **72**, 115331 (2005).

[77] K. MacLean, T. S. Mentzel, and M. A. Kastner (2009), cond-mat/0908.3181.

[78] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[79] H. Böttger and V. Bryksin, *Hopping Conduction in Solids* (VCH, 1985).

[80] M. L. Knotek, M. Pollak, T. M. Donovan, and H. Kurtzman, Phys. Rev. Lett. **30**, 855 (1973).

[81] W. E. Spear and P. G. L. Comber, Solid State Communications **17**, 1193 (1975).

[82] R. D. Schaller and V. I. Klimov, Phys. Rev. Lett. **92**, 186601 (2004).

[83] V. L. Colvin, M. C. Schlamp, and A. P. Alivisatos, Nature **370**, 354 (1994).

[84] A. I. Kingon, J. Maria, and S. K. Streiffer, Nature **406**, 1032 (2000).

[85] G. D. Wilk, R. M. Wallace, and J. M. Anthony, J. Appl. Phys. **89**, 5243 (2001).

[86] N. B. Zhitenev, T. A. Fulton, A. Yacoby, H. F. Hess, L. N. Pfeiffer, and K. W. West, Nature **404**, 473 (2000).

[87] M. Drndic, R. Markov, M. V. Jarosz, M. G. Bawendi, M. A. Kastner, N. Markovic, and M. Tinkham, Appl. Phys. Lett. **83**, 4008 (2003).

[88] S. D. Senturia, C. M. Sechen, and J. A. Wishneusky, Appl. Phys. Lett. **30**, 106 (1977).

[89] D. Xie, Y. Jiang, W. Pan, and Y. Li, Thin Solid Films **424**, 247 (2003).

[90] P. S. Parker, A. P. Monkman, M. C. Petty, and R. Pride, IEEE Proc. **144**, 111 (1997).

[91] J. D. Cohen, Semiconductors and Semimetals **21**, 9 (1984).

[92] J. Orenstein and M. A. Kastner, Phys. Rev. Lett. **46**, 1421 (1981).

[93] T. Tiedje and A. Rose, Solid State Communications **37**, 49 (1980).

[94] S. C. Agarwal, Phys. Rev. B **10**, 4340 (1974).

[95] J. Schwinger, L. L. DeRaad, K. A. Milton, and W. Tsai, *Classical Electrodynamics* (Westview Press, 1998).

[96] W. E. Boyce and R. C. Diprima, *Elementary Differential Equations and Boundary Value Problems* (John Wiley and Sons, Inc., 1992).

[97] T. S. Mentzel, V. J. Porter, S. Geyer, K. MacLean, M. G. Bawendi, and M. A. Kastner, Phys. Rev. B **77**, 075316 (2008).

[98] R. E. Johanson, M. G. Günes, and S. O. Kasap, J. Non-Cryst. Solids **266 - 269**, 242 (2000).

[99] C. E. Parman, N. E. Israeloff, and J. Kakalios, Phys. Rev. Lett. **69**, 1097 (1992).

[100] J. Fan and J. Kakalios, Phys. Rev. B **47**, 10903 (1993).

[101] N. Y. Morgan, *Electronic Transport in CdSe Quantum Dot Arrays*, PhD dissertation, Massachusetts Institute of Technology, Department of Physics, 2001.

[102] D. V. Talapin and C. B. Murray, Science **310**, 86 (2005).

[103] D. Yu, C. Wang, and P. Guyot-Sionnest, Science **300**, 1277 (2005).

[104] H. E. Romero and M. Drndic, Phys. Rev. Lett. **95**, 156801 (2005).

[105] C. B. Murray, C. R. Kagan, and M. G. Bawendi, Science **270**, 1335 (1995).

[106] C. B. Whan, J. White, and T. P. Orlando, Appl. Phys. Lett. **68**, 2996 (1996).

[107] L. S. Levitov and B. Kozinsky, cond-mat/9912484, 1999.

[108] Z. Hu, M. D. Fischbein, and M. Drndić, Nano Letters **5**, 1463 (2005).

[109] J. Robertson, Rep. Prog. Phys. **69**, 327 (2006).

[110] G. I. Meijer, Science **319**, 1625 (2008).

[111] M. Janousch, G. I. Meijer, U. Staub, B. Delley, S. E. Karg, and B. P. Andreasson, Adv. Mat. **19**, 2232 (2007).

[112] J. B. Miller, *Electron Transport in GaAs Heterostructures at Various Field Strengths*, PhD dissertation, Harvard University, Division of Engineering and Applied Science, 2007.

[113] M. Köhler and W. Fritzsche, *Nanotechnology* (Wiley-VCH, 2007).