ULTRA-SHORT CARBON NANOTUBE QUANTUM DOT TRANSISTORS: ELECTRON-HOLE ASYMMETRY, BENDING VIBRONS, AND THE KONDO EFFECT

Andrew McRae

A thesis

in

The Department

of

Physics

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science (Physics)
Concordia University
Montréal, Québec, Canada

September 2013

Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: **Andrew McRae**

Entitled: **Ultra-Short Carbon Nanotube Quantum Dot Transistors: Electron-Hole Asymmetry, Bending Vibrons, and the Kondo Effect**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Physics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

Dr. Truong Vo-Van	Chair
Dr. Truong Vo-Van	Examiner
Dr. Valter Zazubovich	Examiner
Dr. Alexandre Champagne	Supervisor

Approved _____

Chair of Department or Graduate Program Director

_____ 20 _____   _____

Dean of Faculty

Abstract

Ultra-Short Carbon Nanotube Quantum Dot Transistors:

Electron-Hole Asymmetry, Bending Vibrons, and the Kondo Effect

Andrew McRae

Using an electromigration procedure which we recently developed, we generate 10 nm-scale single-wall carbon nanotube quantum dot (SWCNT-QD) transistors. Because these devices are so short, we can explore fundamental mesoscopic physics, engineer tuneable nanoelectromechanical systems (NEMS) and create ultra-short transistors. These dramatic effects arise from enhanced electron-vibron and QD-lead coupling in short devices.

Contrary to what has been observed in longer SWCNT devices, we observe strong electron-hole asymmetry, due to charge doping from the metallic leads. This asymmetry manifests itself as a striking difference between electron and hole charging energies (up to a factor of 3), and their conductance (0D to 1D transport). The magnitude of this asymmetry depends on the length of the SWCNT.

Suspended SWCNTs can strongly couple to their electrostatic environment through the bending mode, and act as NEMS sensors. Shorter NEMS have higher frequencies and therefore higher sensitivity. By creating very short devices, we observe self-actuated bending mode frequencies up to ≈ 280 GHz, and tune this frequency by electrostatic strain. We clearly resolve the first and second harmonic of the bending resonance and extract their effective coupling $\lambda^* \sim 1$.

In high conductance devices, we observe strong electron-electron interactions, with Kondo temperatures up to $T_K \approx 28$ K, and use these interactions to resolve the energy spectrum of the QD. In devices combining Kondo and bending oscillations, we measure a reduction in charging energy, to the point of complete suppression. This is, to our knowledge, the first time this effect has been observed in molecular transistors.

Acknowledgments

I would first and foremost like to acknowledge the Champagne group. Jimi, Serap, Vahid, Matt, and Colleen, this would not have been possible without your happy faces, motivation, and a little bit of silliness. I would like to thank Joshua, for showing me the ropes and getting me started. I wish you could have stuck around a bit longer. And finally to Alex, for his constant energy, guidance, and great stories. I really appreciate all that you have done for me over the last two years. I would like to acknowledge my pals, Rory, Geoff, Matt, Iain, Chris and Marcus for always keeping in touch and keeping me laughing; I'll see you soon. I would like to thank Dad, for all of his hard work to get me where I am today, and for being there for me unconditionally. And I would like to thank Katie for being Katie. I am so proud of you, and am glad you'll be just a short bus-ride away. I would like to thank Mica because I don't know how I would have survived without her. You are the most important thing in my life; thanks for taking such good care of me. Lastly I would like to thank mum from the deepest of my heart, for being a truly fantastic mother. Thanks for looking out for me. I know you would have been very proud.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

While single-wall carbon nanotubes (SWCNTs) have been well researched down to the 100-nm length scale, the 10-nm scale remains largely unexplored. In this mesoscopic regime, quantum effects begin to play a larger role, providing a platform to study fundamental physics and generate new types of devices. The reason why this rich system has barely been explored is because there exists no reliable method for the production of such devices. We have recently developed a method to generate such devices, and example of which is shown in Fig. 1.1. By understanding the challenges involved in making ultra-short electronics devices, they can be made smaller, faster, and more efficient [1, 2].

In-10 nm scale SWCNT transistors, the channel length is similar to the extent of charge doping by the metallic contacts. Thus the contacts play a larger role and dramatically affect device characteristics differently for positive and negative charge transport. Our goal is to explore the length-dependence of specific interactions in

these devices, such as electron-hole asymmetry, electron-vibron coupling, and the Kondo effect, and to understand how to control them. We find that under the right conditions, these effects can become extremely strong. Ultra-short SWCNT QDs are of fundamental and applicable importance, but to access this physically rich system, we must first produce 10-nm scale devices.
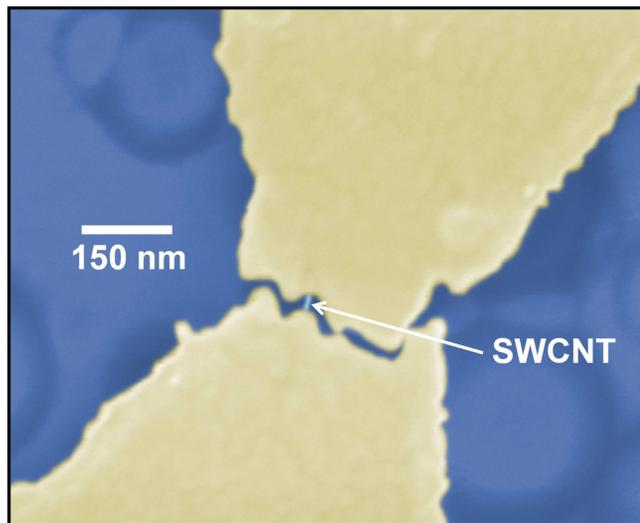
Figure 1.1: Ultra-short SWCNT transistor. One of our ultra-short suspended SWCNT devices ($\approx$ 20 nm).

By using an electromigration method, we create some of the shortest SWCNT devices made to date [3, 4], and can explore the fundamental physics that only becomes observable at this length scale. This technique allows us to produce high quality samples with are suspended, clean, and short, allowing us to probe their properties in a very controlled way. For example, in a sub-10 nm device, we measure a room temperature transistor effect with an on/off ratio of > 4100. This is one

of the shortest high quality SWCNT transistors made to date, and is on par with the work being done using better quality fabrication instrumentation elsewhere [4]. In these devices we can specifically tune between different tunneling regimes, and explore electromechanical coupling in suspended devices.

In our devices, we observe electron-hole (e-h) asymmetry which can allow us to tune the quantum dot (QD) for holes and electrons separately. We find that we can alter the channel length by ≈ 1 nm simply by switching from electron to hole transport because of charge doping from the metallic leads. We observe that the classical energy needed to charge electrons and holes differs by up to a factor of three. In this way, a single device can behave as two; each type of carrier having different transistor properties. This difference in electron vs. hole injection in our devices, could be used for charge pumping by mechanical oscillations [5] or in photovoltaics [6]. This asymmetry can also change the way in which excitons travel through SWCNTs and therefore alter its optical properties [7].

In nearly-metallic SWCNT devices, a dramatic change is observed where holes behave as though they are in a zero dimensional system, while electrons behave as if they are in a one dimensional system. This tuneability allows us to observe the particle-like quantum dot behaviour of a device, while understanding its phase coherence through wave-like interferences. This could be useful for making devices which use both the static QD properties and dynamic coherent electron interactions, such as in spin Q-bits, spintronics devices or in information processing [8, 9].

Because our devices are suspended and very short, they are ideal for exploring

strong electron-vibron (e-v) interactions in carbon nanotubes. We observe very strong electromechanical coupling, allowing us to measure many different vibrational quantum states. In these ultra short devices, the nanotube stretching oscillations reach high frequencies up to 1.2 THz, and have large electron-vibron coupling $\lambda \gtrsim 1$. The bending motion of carbon nanotubes can be harnessed for use in nanoelectromechanical systems (NEMS). These NEMS can be used for many applications such as mechanical switches and Q-bits [10], or ultra-sensitive mass/force sensors [11, 12, 13]. We observe a self-actuated bending mode resonance in our devices with frequencies up to 280 GHz, the highest yet observed in SWCNT NEMS, and quality factors up to $Q \sim 10^6$. The strong e-v coupling in our samples, allows us to observe the second harmonic of the bending mode using DC transport measurements. By applying uniaxial mechanical strain, we show that these devices have large strain-tuneability, allowing us to increase the bending frequency by a factor of 2.

Strong electron-electron (e-e) interactions can occur between the quantum dot and the leads, which are brought closer together in ultra-short devices. We observe a correlation between the electrons on the QD and in the leads which enhances cotunneling, known as the Kondo effect. The strength of the Kondo effect is measured by its onset temperature, and is dependent on the length of the SWCNT. We are interested in pushing the limits of the Kondo temperature in SWCNT QDs by engineering devices with small bandgaps and short lengths. To our knowledge, the previous highest Kondo temperature recorded SWCNTs is 14 K in a several-hundred nm nanotube [14]. We observe a very strong Kondo effect in a short (100 nm)

SWCNT, with gate-tuneable Kondo temperatures from 17 to 28 K. Devices with strong e-e interactions can be used for transistor applications because they reduce shot noise [15], and can allow us to resolve excited states of QD systems [16]. Additionally strong Kondo effect can be used to generate spintronics devices [9, 17].

We observe a theoretically predicted interaction between electrons and vibrons in our devices, causing a polaronic shift in the energy levels of the QD [15, 18]. To our knowledge, this is the first time this effect has been observed in molecular transistors. This interaction can boost the Kondo temperature and reduce the classical electrostatic energy needed to charge the QD. We can to tune the effective charging energy of our quantum dots via temperature, dramatically changing its transistor properties. It is important to understand these effects for molecular devices, as it can play a large role in their transistor properties [15]. Such devices could allow us to explore nano-scale current rectification or spintronics [19].

To present these results, we will structure the thesis as follows. Chapter 2 will give a summary of our ≈ 10 nm device microfabrication methods, and the instrumentation that I developed during my M.Sc. Chapter 3 will give background information on the electronic and mechanical structure of carbon nanotubes, quantum dot theory and Coulomb blockade, before discussing the various manifestations of electron-hole asymmetry we observe in our devices. Chapter 4 will focus on electromechanical coupling between electron tunneling and vibronic states, followed by the Kondo effect. We will finally discuss how e-v interactions cause a change in the observed charging energy of our ultra-short SWCNT quantum dots.

Chapter 2

Ultra-Short SWCNT Devices

The process which forms the groundwork for all of our research is the sequence of microfabrication steps leading to clean and suspended 10-nm scale SWCNT devices. Our microfabrication is guided by the structure of our devices, shown in Fig. 2.1, which allows us to tune the electronic properties of our suspended SWCNT QDs. To properly measure these devices, the appropriate instrumentation and experimental set-up are required. I helped to build, design and repair several instruments during my M.Sc. By optimizing these fabrication and measurement processes, we were able to generate very clean ultra-short SWCNT QDs, and collect high quality data.

This chapter is structured to describe sample preparation and data taking procedures, up to and including electronic measurements. We first describe the full microfabrication process, including wafer preparation, SWCNT growth, electrical contact, suspension, and finally the electromigration process. We will then discuss my work on lab instrumentation, focusing on the ^{3}He cryostat. To conclude, we will
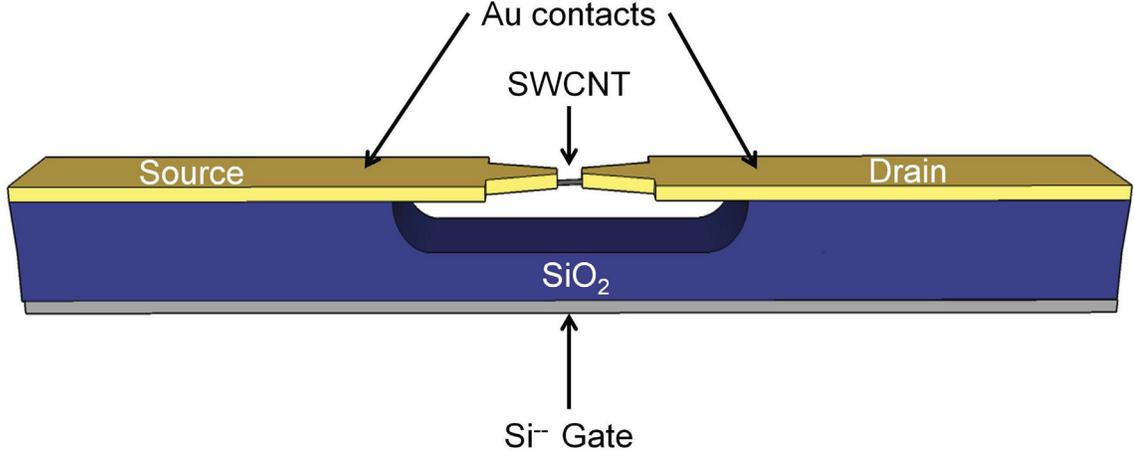
Figure 2.1: Geometric structure of our SWCNT transistors. This diagram depicts our suspended device geometry with source, drain, and gate electrodes.

mention the methods used for data acquisition, and some of the macros coded to simplify data analysis.

2.1 Sample Fabrication

The microfabrication procedures we use are selected to create very short, tuneable SWCNT QDs. It is important that these procedures are performed cautiously, patiently, and diligently; any error made throughout this sequential process nullifies the previous effort. Our specific sample preparation method culminates in the nanoscale etching process known as electromigration, which allows us to make some of the shortest (< 10 nm) SWCNT transistors to date. To control the electronics of our devices, we require source, drain, and gate electrodes. We use two Au contacts, and the degenerately doped Si^{--} substrate respectively, as shown in Fig. 2.1. The

devices must also be suspended, not only to isolate them from any defects which may be introduced by the substrate, but also to insulate the junctions thermally, allowing electromigration. Our full microfabrication process, from nanotube growth to sample measurement, is described in detail here.

2.1.1 Reactive Ion Etching (RIE)

We begin our microfabrication with a degenerately doped 4" $\langle 100 \rangle$ Si^{--} wafer, covered with a 300 nm layer of SiO$_2$. To ensure that we will be able to contact the backside of the wafer electronically, we first perform a step of reactive ion etching (RIE) to remove the oxide from one side of the wafer.

We use a standard O$_2$ etching to clean the RIE chamber with 20 SCCM O$_2$, at a pressure of 200 mTorr, and at 300 mW for 2 min. We then purge and pump the chamber three times with N$_2$ to remove any remaining contaminants before venting the chamber. We wipe the chamber with isopropyl alcohol (IPA) to further clean it before loading the wafer, oxide down. We put the pump the chamber and create a plasma with a mixture of 0.7 SCCM O$_2$ and 6.3 SCCM CHF$_3$ at a pressure of 125 mTorr and at 300 mW for 15 min. During this time, we ensure that the reflected power of the matching unit remains below 50 mW. We again purge the chamber three times with N$_2$ before removing the sample and repeating the oxygen plasma step to clean the chamber for the next use. The bottom of the wafer is now free of oxide and we can ready the top of the wafer for SWCNT growth.

2.1.2 Photolithography

Nanotubes are grown from catalyst islands by a process known as chemical vapour deposition (CVD). The islands consist of a sub-monolayer of sputtered Fe, and are defined by photolithography.

We rinse the wafer with acetone and IPA to ensure its cleanliness before spin coating it with Shipley 1813 photoresist at 4000 rpm for 30 s. This process provides a relatively even photoresist thickness of ≈ 1.4 μm, determined using ellipsometry measurements. The wafer is then soft-baked at 115 °C for 1 minute to solidify the resist. The wafer is kept under yellow light throughout this entire process to prevent UV contamination from ruining the lithography. The wafer is then diced into thirds using a diamond-tipped scribe so that the largest number of patterns can be exposed on its surface.
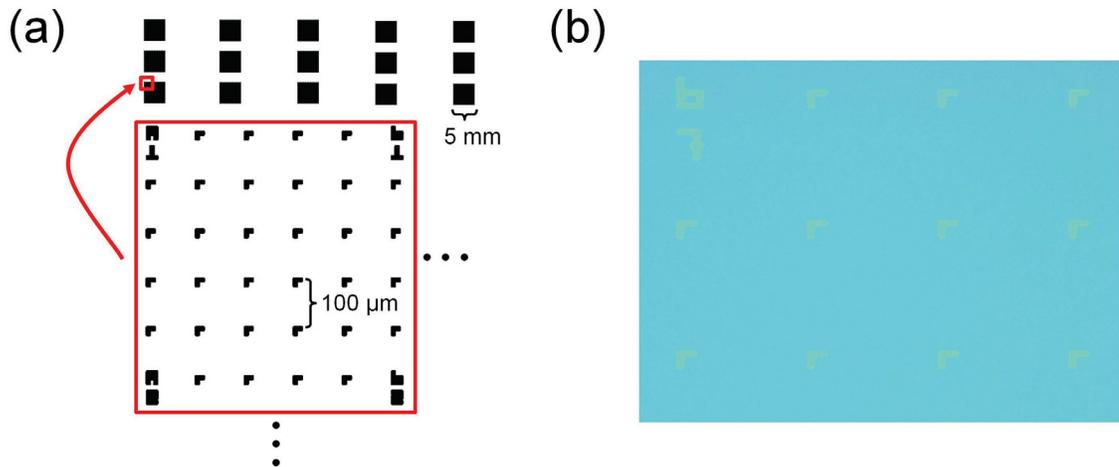
Figure 2.2: Photolithography pattern and sputtering. (a) The photolithography mask is made of fifteen lettered and numbered ⌐ patterns with 100 μm spacing. (b) Catalyst islands made from sputtered iron ⌐ patterns (≈ 7 Å) used for nanotube growth.

The photolithography pattern is shown in Fig. 2.2(a). The fifteen 5×5 mm patterns on the photolithography mask are made up of 500 μm spaced letters and numbers, A-I from left to right and 1-9 from top to bottom. This serves as a way to mark locations on the wafer. These letters and numbers are further divided by $\ulcorner$ shapes every 100 μm. Each third of the wafer is aligned in the SUSS M4A mask aligner and raised into hard contact with the mask to minimize diffraction. The intensity of the UV light is measured in the aligner, and the wafer is exposed for the length of time needed to achieve a dose of 40 mJ/cm^2. The wafer section is then re-exposed in a blank area, approximately 7 mm longitudinally away from the original exposure, giving a total of 30 sets of patterns per section and 90 patterns per wafer. The section of wafer is then submerged in MF-319 photoresist developer for 45 s to dissolve the exposed resist. The wafer is quickly rinsed with deionized water for 20 s to remove the MF-319 and prevent over-developing of the pattern. It is then dried using a nitrogen gun and inspected under optical microscope to ensure that the lithography is of high quality. If the exposure is not acceptable, the resist is fully removed with acetone and the process is repeated. We now deposit a catalyst material for SWCNT growth.

2.1.3 Sputtered Catalyst Islands

Sputtering allows us to deposit a sub-monolayer of Fe catalyst which we use to grow carbon nanotubes (CNTs). An optical image of the sputtered pattern (after liftoff) is shown in Fig. 2.2(b). In a sputtering chamber, we form an Ar plasma at 10 SCCM,

10 mTorr, and 150 mW, with less than 5 mW reflected power on the RF matching unit. From a previous long-term sputter and profilometry test, we determine the rate of deposition of Fe using these settings to be ≈ 3.5 Å/min.

In a fume hood we remove the unwanted Fe and photoresist by placing the sputtered wafers in hot acetone (60 °C) for > 10 min. We then spray them with hot acetone using a syringe to ensure the complete removal of the photoresist. The wafers are inspected under a stereoscope before they are removed from the acetone bath and rinsed with fresh acetone and IPA before being dried with a nitrogen gun. Because they become optically invisible after annealing, we mark the positions of the sputtered patterns on the wafer using a scribe to remember their placement.

Iron thickness is crucial for nanotube growth. If there is not enough Fe, small nanoparticles will form during deposition and nanotubes will not grow. If there is too much Fe, large nanoparticles will form, generating densely packed large diameter carbon nanotubes, which are likely to be multi-walled (made from several concentric CNT shells). We found that an intermediate thickness which forms an acceptable yield of SWCNTs is ≈ 7 Å.

2.1.4   Chemical Vapour Deposition (CVD)

The carbon nanotubes are now grown from the sputtered islands. We use a Thermo Scientific furnace to heat up the growth chamber and a custom built gas-mixing panel. The sample is placed in a vacuum-tight quartz glass tube in the furnace and we flow gases which cause the growth of SWCNTs through chemical vapour deposition.

The exact details of the recipe we use to grow nanotubes come from a mixture of methods used by other groups [20, 21, 22]. We first anneal the catalyst islands at 900 °C under atmosphere for 1.5 hrs to form Fe nanoparticles which seed the growth of carbon nanotubes. The temperature of the furnace is reduced to 500 °C to safely attach the gas lines. The carrying gas, Ar, is flowed at a rate of 0.1 SLM, while the furnace is re-heated to 900 °C. The Ar flow is then stopped and 0.1 SLM of H_2 is added to the chamber for 20 min, which helps to "wet" the iron nanoparticles and catalyze the growth of CNTs [22]. The furnace temperature is raised to 970 °C and allowed to equilibrate for 3 min before the growth gas, CH_3, is introduced at 0.2 SLM for 45 min. During this time, the nanotubes grow on the substrate. Afterwards, the flow of CH_3 and H_2 is stopped, Ar is flowed at 0.1 SLM, and the temperature is reduced down to 400 °C. Once the furnace reaches this temperature the lid is removed to speed the cooling. At 200 °C, the sample is removed, now covered with dispersed carbon nanotubes. We must now give each nanotube their own "address", so that they can be located easily later.

2.1.5 Thermal Evaporation

Before searching for carbon nanotubes using the scanning electron microscope (SEM), we perform a second photolithography to allow us to locate the SWCNTs on a Cartesian plane. We use the same procedure and mask as for making the catalyst islands, but this time evaporate Au instead of sputtering Fe. This process is complicated by the fact that the new patterns must line up with the previous catalyst

$\ulcorner$ shapes, such that they contain the newly grown carbon nanotubes. We carefully align the mask with the help of the scribed marks that were made after the sputter liftoff. The $\ulcorner$ shapes must be sharp and well defined for the alignment of the e-beam lithography pattern, as will be discussed later.

We thermally evaporate metals by heating them up electronically inside of a vacuum chamber. The thickness of the evaporated material is measured using a crystal thickness monitor. We first evaporate 5 nm of Cr to act as a sticking layer for better adhesion between the SiO_2 and the 80 nm of Au, which is evaporated directly afterwards. We then perform the same liftoff procedure as before, using hot acetone and a syringe to remove the unwanted resist. Each 5×5 mm section of wafer is diced, face up, and rinsed with acetone and IPA. These chips are now ready to be imaged to determine the locations of the SWCNTs.

2.1.6   Scanning Electron Microscopy (SEM)

We locate the carbon nanotubes using a scanning electron microscope. SEM images of two of our CNTs are shown in Fig. 2.3 (a)-(b). It was found that for low current (10 μA), low working distance ($\approx$ 6.0 mm), and low electron acceleration voltage (1.0 keV), the nanotubes shone brightly. This is because at low accelerating voltages, the oxide substrate becomes positively charged, while the nanotube accumulates electrons, resulting in a sharper contrast [23]. Because the electron beam focuses contaminants into the imaging area, the nanotubes are observed for the shortest possible time at low magnification (700$\times$).
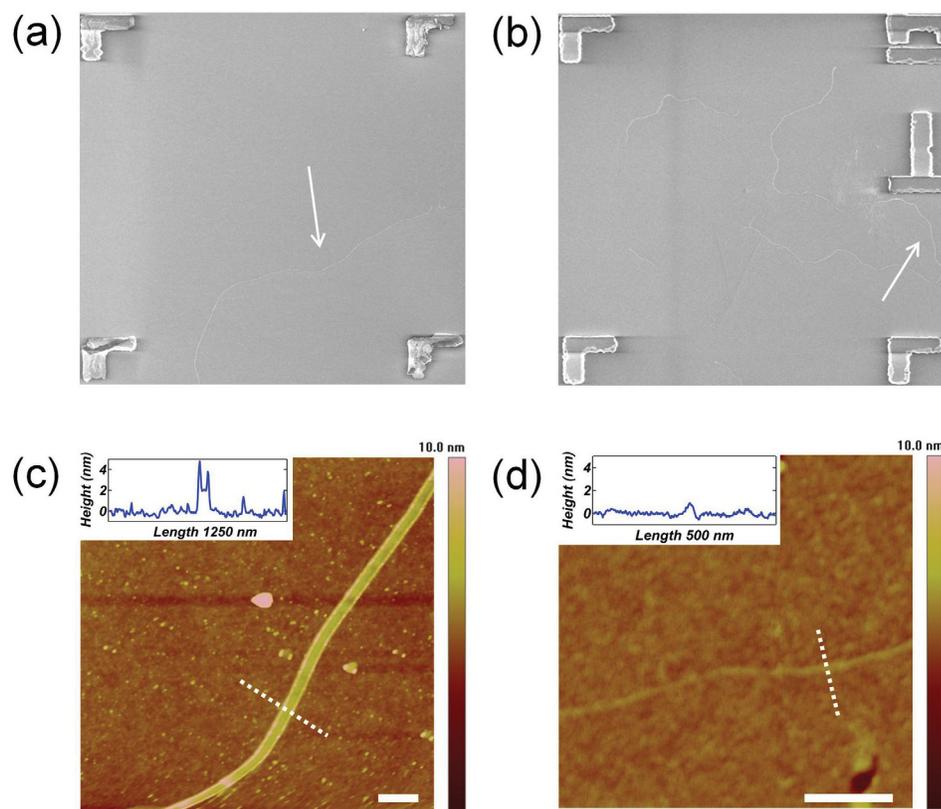
Figure 2.3: SEM and AFM of SWCNTs. (a)-(b) Scanning electron micrograph of "straight" and "meandering" CVD grown nanotubes. For scale, distance between ⌐ markings is 100 nm. (c)-(d) Corresponding AFM scans for the above nanotubes with diameters 4 nm and 0.5 nm respectively with scale bars = 250 nm. Insets show 1D linecuts of AFM height along the dashed lines for each CNT.

Ideal nanotubes are isolated and long enough such that three 25 μm bars will fit along the tube. Usually, three long nanotubes are selected per grid pattern, at least 1.0 mm away from one another, allowing for a maximum of 9 devices per chip. A high resolution SEM image containing the four surrounding ⌐ shapes (later used for alignment) is taken of the candidate nanotubes. We next measure the diameter of the nanotubes using atomic force microscopy (AFM). Through trial and error it was found that, as a rule of thumb, nanotubes which grew with smaller ripples (meandering tubes) were more likely to have a smaller diameter, while nanotubes which curved with a large radius (straight tubes) were more likely to have a larger diameter. Large diameter nanotubes are more likely to be multi-walled. To ensure that the candidate nanotubes are single-walled, we select nanotubes which meander with a smaller curvature, and then measure using AFM.

2.1.7 Atomic Force Microscopy (AFM)

Atomic force microscopy is used to determine the diameter of our carbon nanotubes. We use tapping mode with a tip radius of $\approx$ 10 nm. AFM scans corresponding to the above SEM images are shown in Fig. 2.3 (c)-(d), with dashed lines corresponding to the insets showing nanotube diameter measurements. We locate the nanotube, confirming its position, and slowly scan the tip across the surface. To determine its approximate diameter, nine linecuts are taken within the scan window along the entirety of the nanotube, and the average diameter and deviation are recorded. If the measured nanotube is less than 2.0 nm in diameter, it is likely to be single-walled,

and is ready to be contacted by electrodes. It has been shown that Van der Waals forces between the SWCNT and the AFM tip can cause this diameter reading to be underestimated by ≈ 1 nm [24, 25], and we take this into account for our calculations. Once we determine that these nanotubes are likely to be single-walled, we can now form the device contacts directly on the nanotube.

2.1.8   Computer Aided Design (CAD) of Gold Contacts

We now prepare to deposit micron-sized Au leads directly onto the SWCNTs. Once we have selected the best SWCNT candidates, we adjust the rotation and scaling of our high resolution scanning electron micrograph, such that a 5×5 μm grid lines up with them as shown in Fig. 2.4(a). An origin is placed in the center of the $\ulcorner$ shapes (relative to their top-left corners), and the nanotube is mapped out in the Cartesian plane. We draw sections of the CNT in our CAD software (DesignCAD 16 or Raith e-line plus) and the bowtie break junction patterns used for electromigration are aligned with them, as shown in Fig. 2.4(b). These break junctions are then connected to six rectangular bars around the outside of the pattern, which will later be used to contact large (200 μm) photolithography pads.

If the center of the bowtie pattern is too wide, it becomes difficult to suspend and electromigrate the junction, while if it is not wide enough there is less room for error in the alignment, and we are more likely to miss the SWCNT altogether. We found that a reasonable intermediate was ≈ 350 nm. The full dimensions of the break junction patterns are shown in Fig. 2.4(c). The break junctions must be spaced out
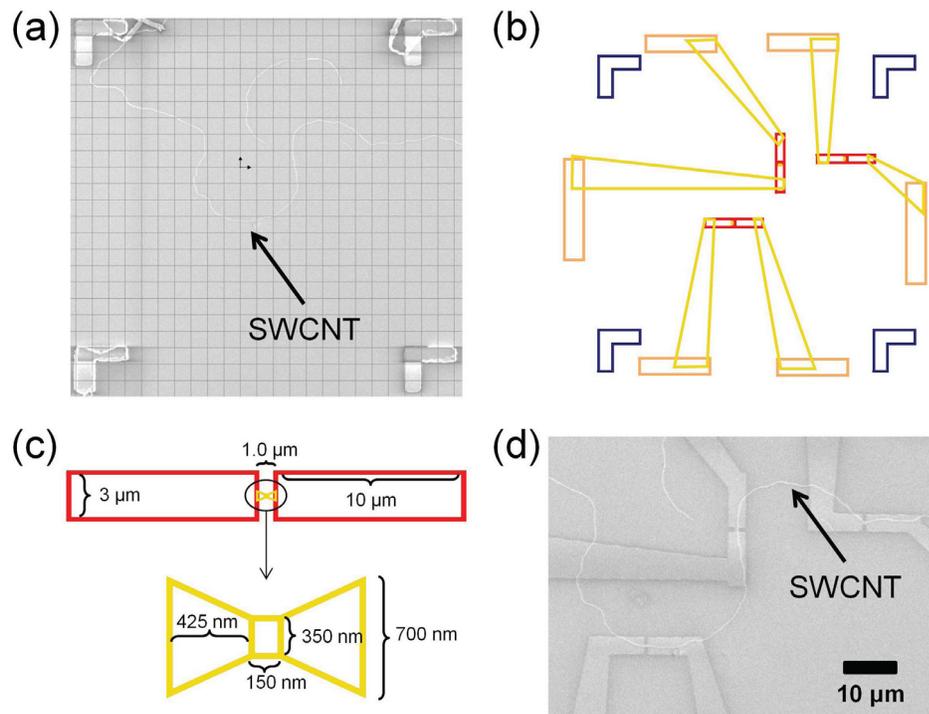
Figure 2.4: Alignment of e-beam lithography patterns. (a) Original scanning electron micrograph of a CVD grown SWCNT. (b) CAD design with break junctions aligned to the carbon nanotube. (c) Close-up of break junction and small contact bars showing their dimensions. (d) Semitransparent overlay of gold break junctions on top of grown nanotubes, showing proper alignment.

($> 3 \ \mu$m) to avoid proximity effects. We next use electron beam (e-beam) lithography to define the gold contacts. An overlay image, showing the nanotube before e-beam lithography and the gold break junctions afterwards, is shown in Fig. 2.4(d).

2.1.9   Electron-Beam Lithography (EBL)

At this stage, we prepare the chip for electron-beam lithography (EBL). This is the process by which the micron-sized contacts are physically defined on the SWCNT. We use a bilayer resist of copolymer EL9 (9% in ethyl lactate) and PMMA A4 resist (polymethyl methacrylate 4% in anisole), which we sequentially spin at 3000 rpm for 1 min, and bake for 15 min each at 170 °C. This gives copolymer and PMMA thicknesses of $\approx$ 300 nm and $\approx$ 200 nm respectively.

We then load the samples into the e-beam writer, where $\approx$ 10 nm contamination spots are made using the electron beam to verify proper stigma and focus. The sample coordinates are correlated to the sample stage using either manual or automatic alignment with the gold $\ulcorner$ shapes. This accounts for rotation, stretch, and position of the pattern. The beam current is measured so that the appropriate dose, nominally 200 μC/cm^2, can be applied to each section. We optimized the doses of the bowties over a wide range and found that to properly expose the bowtie patterns, a dose factor of $1.5 - 2.0$ times that of the bars was necessary. This is due to a reduced proximity effect for the small area bowtie shapes. The chip is then exposed at a magnification of $\approx$ 750$\times$, accelerating voltage of 20 keV, 10 μm aperture giving a current of $\approx$ 40 pA, and a working distance of 8-10 mm. The samples are then carefully removed

from the e-beam writer, developed and prepared for evaporation.

The e-beam resist is developed using MIBK (methyl isobutyl ketone). First, the exposed sample is swirled in a solution of 1:3 MIBK:IPA for 30 s, then quickly transferred to methanol for 15 s to remove excess MIBK, and finally placed in IPA for > 30 s. We inspect the exposure using an optical microscope to ensure that it is acceptable, and the sample is then prepared for evaporation. A Cr sticking layer is not used for the deposition of 40 nm of Au because the metallic bilayer adversely affects the quality of the electromigration. The samples are removed from the evaporator and placed in hot acetone for the bilayer liftoff and inspected with the SEM to ensure that the nanotube is contained within the gold break junction. Once we have contacted our sample on the micron scale, we must then contact it macroscopically.

2.1.10 Large Contact Pads

We form 100 μm contacts using photolithography to contact the EBL pattern, as shown in Fig. 2.5. This six-point pattern will later be used for wire bonding. After spin-coating and soft baking the samples as before, the six-point pattern is carefully aligned in the mask aligner to contact the six edges of the EBL pattern and exposed using a slightly smaller dose of 35 mJ/cm^2. We develop for 60 s in MF-319 and inspect under optical microscope for proper alignment. If the exposure is deemed acceptable, it is evaporated with a sticking layer of 5 nm Cr and 80 nm Au before liftoff in hot acetone. We are now ready to suspend our sample.
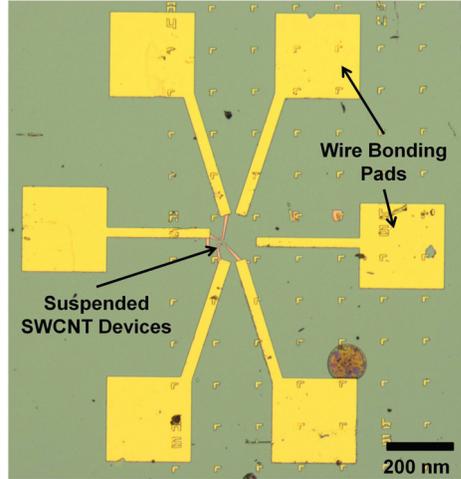
Figure 2.5: Lithographically defined contact pads. To wire bond our devices, we necessitate large sticking pads of gold connected to the EBL pattern. These pads consist of 80 nm/5 nm of evaporated Au/Cr.

2.1.11 Buffered Oxide Etching (BOE)

The oxide layer must now be etched to remove the material underneath the break junction, thereby suspending it. Fig. 2.6 shows two samples after suspension. Inspecting the EBL patterns using SEM deposits carbon residues which can alter the uniformity of the etching rate of the SiO_2 across the surface of the chip. Therefore, for good suspension, we must remove this contamination. Because the carbon nanotube is protected underneath the gold break junctions, we use RIE to etch this residue without damaging the nanotube. This also ensures that no nanotubes will remain on the surface causing short circuits between contacts. We use an oxygen plasma with 20 SCCM O_2 at 200 mTorr and 300 mW for 2 min to clean the chip.

We suspend our devices using a buffered oxide etch, consisting of HF (hydrofluoric acid) in a 1:7 buffer solution of 49%HF:NH_4F. We use an ellipsometer to measure the

20

thickness of oxide and measure the etch rate, found to be $\approx$ 70 nm/min. We found through multiple suspensions on such samples that the inward etch rate underneath the junction is approximately 1.5$\times$ the downward etch rate. Because inward etching occurs on both sides of the junction, the minimum vertical etching distance for device suspension is 1/3 of the junction width, $\approx$ 120 nm for a 350 nm break junction. It is important not to over-etch the sample, as this can cause the contacts to collapse on the surface, while under-etching leaves material under the gold bridge. With this balance in mind, the time needed to fully suspend the junction can be determined from the calibration.
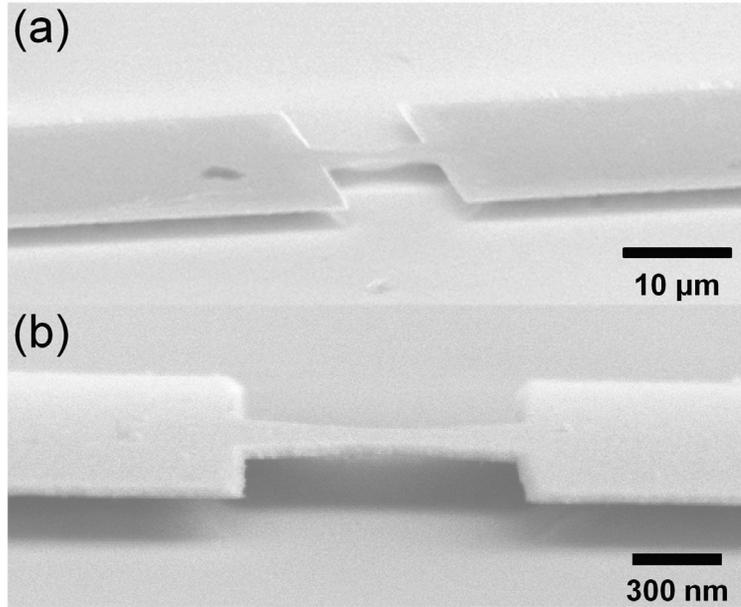
Figure 2.6: Suspension of the gold bridge. (a)-(b) Tilted SEM images taken at 70° and 80° of two break junctions, 300 nm and 375 nm wide, suspended by 105 and 135 nm respectively. The gold bridges are clearly suspended, while the contacts remain well anchored.

The samples are etched by submerging and swirling them in the BOE solution, and then transferring to several deionized water baths. We transfer to a bath of IPA before drying with nitrogen. This makes it is less likely that the gold bridge will collapse from surface tension of the solvent when it evaporates. At this point, the samples are placed on a tilted stage and loaded into the SEM, which can effectively rotate the samples between $60° - 80°$ to view the break junction from the side and confirm suspension. The final step before loading the sample into the cryostat is to contact the chip to the carrier.

2.1.12 Wire Bonding

The suspended devices can now be contacted to the chip carrier giving us direct electrical contact to the sample from the mm-scale. This connection is achieved through wire bonding.
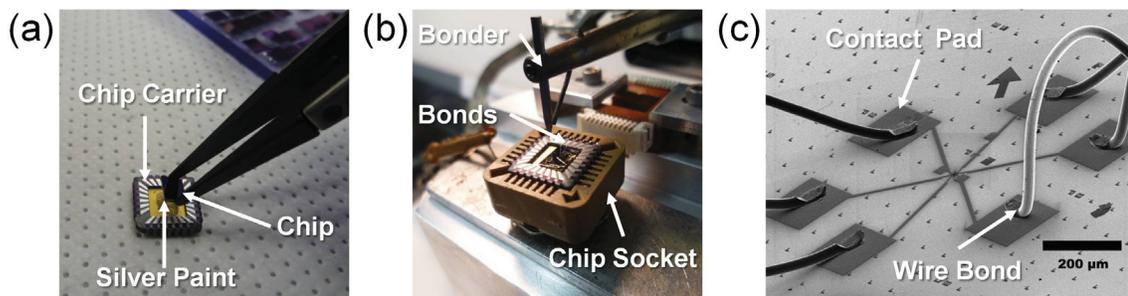
Figure 2.7: Wire bonding. (a) First, the chip is connected mechanically and electronically to the carrier using silver paint. (b) The large gold pads connecting the devices are bonded to the pins of the chip carrier using Al wire. (c) SEM image of wire bond connections to large pads courteousy of Vahid Tayari.

22

The backs of the chips are contacted to the chip carrier using a conductive silver paint, both to hold the chip in place and to make electrical contact to the back gate, as shown in Fig. 2.7(a). We place the chip carrier and sample into a carrier socket with all pins grounded to prevent electrostatic damage. Next, 25 μm diameter Al wires are bonded from the carrier to the large pads. The completed bonds are shown optically and with a scanning electron micrograph in Fig. 2.7(b)-(c). Finally, two wire bonds are made from pins to the conductive "floor" of the chip carrier to control the gate electrode through the silver paint. We have now achieved mm contacts to our nm devices and are ready to connect to our measurement set-up. The device can now be loaded into the probe.

2.1.13   Sample Probe

The chip carrier is loaded very carefully into the fully grounded sample probe socket Fig. 2.8(a). Caution must be taken as the carrier and socket have a unidirectional configuration. A brass "ceiling" is placed over the chip carrier and kept in place by hex nuts to prevent any mechanical damage to the wire bonds in the sample chamber. The contact wires are wrapped around the sample holder to improve thermal contact, and a large amount of vacuum grease is placed around the cone seal at the base of the sample holder to ensure that the sample space is leaktight. The sample holder is then loaded into the sample space of the probe Fig. 2.8(b). A large nut keeps the sample holder in place, while the sensor and heater wires are tied down to the probe and the copper sensor holder is attached. The loaded probe is shown in Fig. 2.8(c).
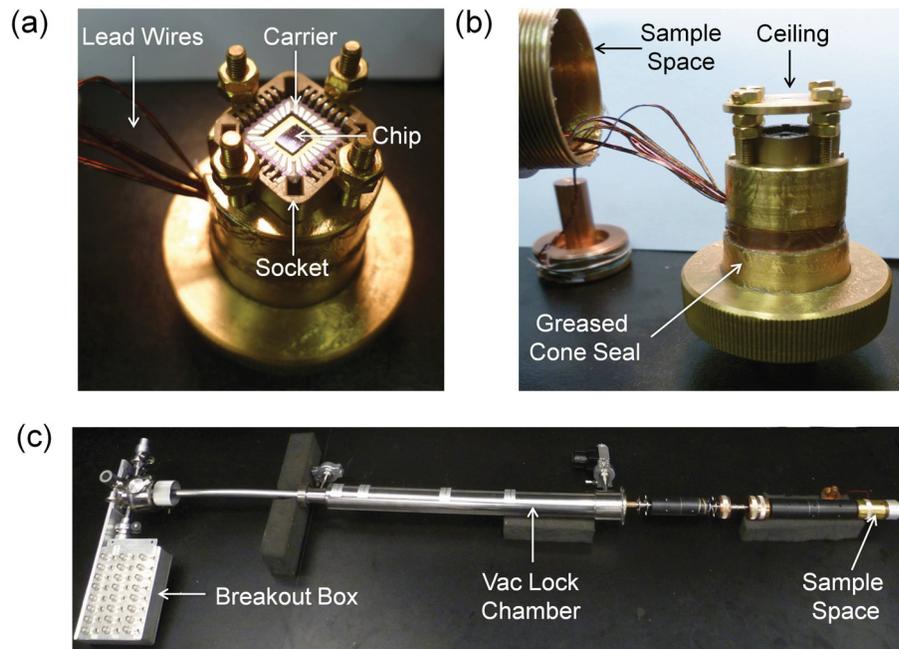
Figure 2.8: Loading the sample. (a) The chip carrier is placed into the socket which is wired to a breakout box at the top of the probe. (b) The cone seal is loaded into the sample space with a brass "ceiling" to prevent damage to the wire bonds. (c) The fully loaded probe is now ready to be inserted into the cryostat.

The sample space is now leak checked and pumped down to $\sim 10^{-6}$ mTorr using a turbo pump and leak detector in series. The probe is then loaded onto the vacuum lock on top of the cryostat. Because this volume is connected to the very delicate and important ^{3}He system, it is well pumped ($\sim 10^{-5}$ mTorr) and leak checked. The gate valve is then opened and the sample probe is slowly lowered into the cryostat to cool the probe over several hours. Once the sample has cooled, we can perform the final step in the fabrication of ultra-short SWCNT devices: electromigration.

2.1.14   Electromigration: Nanoscale Etching

Electromigration is a technique used to form gaps in metallic break junctions. It was developed to make atomic-sized gaps in nanowires, forming contacts for the investigation of molecular devices [26]. In our SWCNT samples, we use electromigration to peel the gold bridge apart, exposing a short section of nanotube. A diagram, depicting the electromigration process is shown in Fig. 2.9.

When a large current is passed through a narrow break junction, the dissipated power causes the atoms therein to heat up. This heating is enhanced by the suspension of the bridge, which prevents heat flow to the substrate. High energy electrons (referred to collectively as the electron wind) can transfer their momentum to the atoms in the break junction, dislodging and moving them. In our devices, the point of highest resistance in the break junction, and therefore the point where most power is dissipated, occurs at the constriction. After migration, the final product is a somewhat irregular geometry with source/drain asymmetry is expected because of the
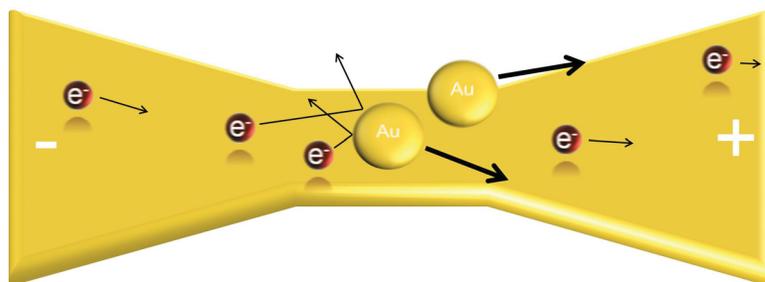
Figure 2.9: Schematic of nanowire electromigration nanowire. High-energy electrons collide with the metal atoms along a nanowire. The transfer of momentum can be so large that it causes metal atoms to become dislodged and move away from the weakest point of the nanowire.

randomness involved in the process, and because the electron wind is unidirectional. As the atoms move away from the constriction during this process, the break junction becomes even narrower, increasing the temperature and therefore the likelihood that atoms will electromigrate. This makes electromigration a runaway process which, if left unchecked, will cause a rapid, uncontrolled breaking of the wire, resulting in a very large gap on the order of 100 nm.

Because our aim is to produce much smaller gaps, we must modify our process. By measuring the resistance of the junction in real time, we can regulate the electromigration process using a feedback loop, as in ref. [27]. A circuit diagram of our set-up is shown in Fig. 2.10. We begin by slowly ramping up the voltage
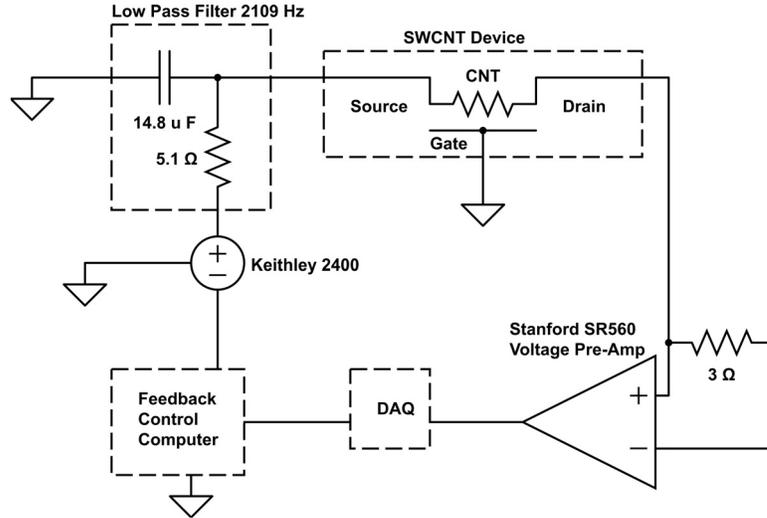
Figure 2.10: Our electromigration circuit. A circuit diagram showing the electromigration system used to open nm-scale gaps in our gold break junctions.

across the sample, while our custom electromigration software measures current and calculates resistance. When a significant ($\approx 5\%$) change in resistance is detected (indicating the start of the runaway process), the voltage is very quickly ramped down to arrest the electromigration. By repeating this process iteratively, we slowly etch away gold atoms and increase the resistance of the junction up to the tunnel barrier resistance. The gaps formed by this method are so small that direct tunneling occurs between the contacts, and we cannot measure the SWCNT. We must therefore once again alter our method to form larger gaps.

We use a two step electromigration process, as shown in Fig. 2.11(a), to form slightly larger gaps (~ 10 nm). A device formed using this method is shown in Fig. 2.11(b). The first step of the process is the iterative feedback etching just described (red). This process is continued down to a predetermined power which will determine
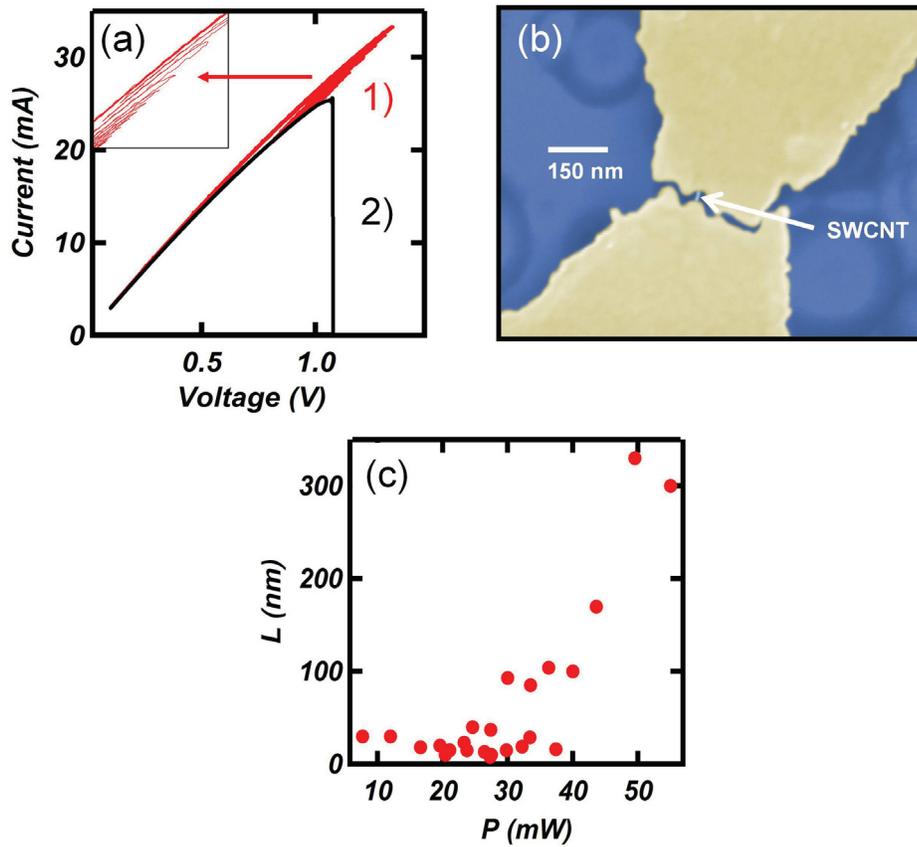
Figure 2.11: Feedback-controlled electromigration technique. (a) Current vs. voltage plot of our two step electromigration process. 1) Feedback-controlled nanoscale etching (red) slowly removes gold from the center of the bowtie, weakening the junction, and 2) set power breaking (black) opens a gap in the gold bridge. (b) False coloured SEM image this device ($\approx$ 20 nm) after electromigration. (c) Dependence of the length of break junction on power of the second electromigration step for all of our samples.

the final length of our devices, and weakens the bridge for the second step: set-power breaking (black). The voltage is ramped up across the weakened junction with the feedback loop disabled. At the predetermined power, the avalanche electromigration process takes over and opens a 10-nm sized gap in the junction, containing an ultra-short exposed section of SWCNT. The voltage across the sample is quickly ramped to zero so as not to damage the SWCNT. It was found that for ≈ 400 nm wide junctions, the optimal breaking power to achieve a nanogap was ≈ 30 mW.

Although this process inherently contains some randomness, there is a strong dependence of break junction length on the set-power of the second step Fig. 2.11(c). This shows that we can roughly tune the size of the gap and therefore the length of our device. By using this feedback controlled two-step electromigration process we can consistently generate ultra-short undamaged SWCNT devices. Usually, this heat generating process removes any chemical microfabrication residues, not only in the SWCNT, but also in the contacts. Sometimes, however, further current annealing is required to fully clean the device.

2.1.15   Current Annealing

After initial data measurements are taken, it is sometimes found that the sample is contaminated by gold, or fabrication residues. This is especially true in devices with very wide (> 400 nm) break junctions. We use current annealing, which locally cleans the suspended nanotubes and contacts through power dissipation. We show a sample annealing curve up to 2.2 μW in Fig. 2.12.
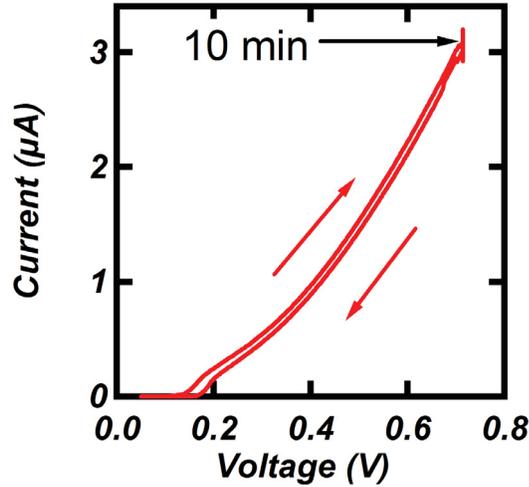
Figure 2.12: Current annealing. Samples were annealed up to several microwatts, to further remove gold and other impurities. This device was annealed at 2.2 μW for 10 min, effectively cleaning the sample.

To anneal, we ramp up the voltage across a device. The power dissipated along the suspended nanotube burns off impurities and moves the gold atoms away from the SWCNT. This effect is again enhanced by the suspension of the device, which thermally isolates the break junction. To anneal, we use the electromigration software, but control it manually, rather than using the feedback loop. Great care must be taken at this point because large electric fields and currents can destroy the nanotube. We sequentially pause the anneal at higher and higher powers for times ranging from $3 - 10$ min. Typical anneal powers range from $1 - 20$ μW. After annealing, we are ready to measure our clean nm-sized devices.

2.2 Instrumentation and Measurement

Over the course of this M.Sc., I spent much time working on measurement set-ups and instrumentation. Without this necessary work, fabrication and measurement would be made much more difficult, if not impossible. I helped to assemble and repair several instruments over this time, mainly the ^{3}He cryostat used to cool our samples to 0.3 K, allowing us to measure the quantized energy spectra of our devices. In this section, we outline the functionality of these instruments, as well as the work done on them, starting with the cryostat and additional instrumentation. Then we describe our measurement set-up and data analysis macros.

2.2.1 ^{3}He Cryostat

The ICE Oxford dryICE cryogen free ^{3}He cryostat allows us to freeze out thermal fluctuations and truly observe the fine electronic structure of our devices down to 0.3 K. The cryostat is outfitted with optical ports for photonic measurements, as well as a 9.0 T superconducting magnet. My main contribution to instrumentation for the lab was assembling and learning to use the ^{3}He cryostat. We will describe the functionality of the cryostat, its assembly, and then the wiring of the sample probe for electromigration and measurement.

Our discussion of the cryostat's functionality will be aided by the diagram in Fig. 2.13. This cryostat does not depend on the addition of cryogenic liquids to cool to base temperature. The outer vacuum chamber (OVC) contains the entire system
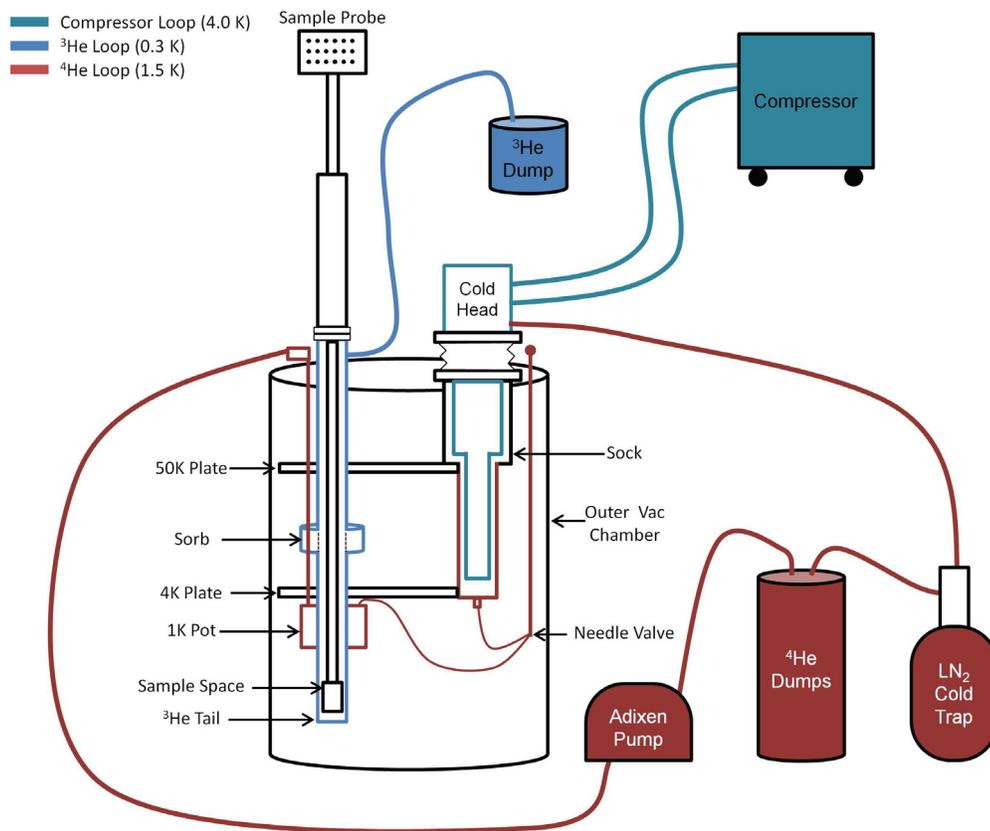
Figure 2.13: Cryostat functionality diagram. This diagram shows the three colour-coded He circuits which sequentially cool the sample to 0.3 K: compressor loop in green, ^{4}He loop in red, and ^{3}He in blue.

preventing heat exchange with the ambient air, and large plates at various stages inside of the fridge block radiation (50 K and 4 K plates). Three isolated helium loops are used concurrently to achieve 0.3 K. The first loop is a closed ^{4}He line (green) which runs from the compressor to the Gifford-McMahon (G-M) cooler. This cools the cold head down to $\approx$ 4 K by gas decompression.

The second circuit is an additional ^{4}He loop (red) which runs from the gas reserves in the dumps to a liquid nitrogen cold trap. This freezes out impurities and prevents contamination of the lines inside the cryostat. The helium gas then enters the sock at the top of the cryostat. Because this is adjacent to the cold head, the gas condenses inside the sock, where it drips down through the needle valve (setting the flow rate), and into the 1K pot. An oil-free Adixen pump reduces the pressure inside the 1K pot, causing the liquid helium to evaporate and cool the surrounding area to $\approx$ 1.3 K. The exhaust from the pump returns the ^{4}He gas back to the dumps.

The final cooling step depends on the much rarer ^{3}He gas (blue). The external ^{3}He dump leads to a heated charcoal sorption pump (sorb) inside of the cryostat, and then to the helium 3 tail, which surrounds the sample space. When the sorb is hot ($>$ 40 K) the ^{3}He is outgassed. Its proximity to the 1K pot causes the gas to condense into the helium 3 tail. When the sorb is cooled, it slowly pumps on the ^{3}He liquid, cooling the sample space to 0.3 K.

With the help of an ICE Oxford technician the system was assembled and prepared for use in the lab. The assembled cryostat is shown in Fig. 2.14(a). For the compressor, we built the electrical connection, provided it with cooling water lines

33
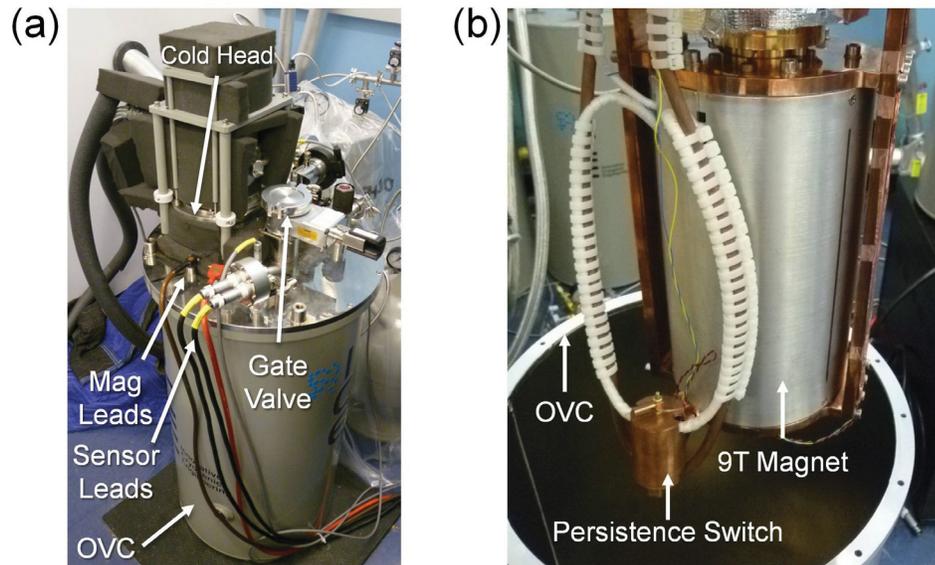
Figure 2.14: Cryogen-free ^{3}He cryostat. (a) ^{3}He cryostat which was assembled and configured with the help of a technician. (b) Installation of new 9 T magnet, replacing the faulty one, which was performed alone.

and a water filter, and charged it with ^{4}He gas. We filled the ^{4}He dumps and leak tested both the ^{4}He and ^{3}He loops before connecting to their respective dumps. After assembly, there were some issues with 9.0 T superconducting magnet, so it was removed and returned to American Magnetics. It was repaired there and returned to the lab. The removal of the magnet is shown in Fig. 2.14(b). I reassembled the system without the technician and confirmed that the magnet was working properly.

Once the cryostat was running, we rewired the sample probe with twenty-four low resistance Manganin wires, necessary for electromigration. I soldered these wires to the connector at the top of the probe, wrapped them down the inner chamber with heat sinking baffles, and painted them with low temperature thermal varnish to improve thermal contact Fig. 2.15(a). A 24 pin BNC breakout box was designed to fit
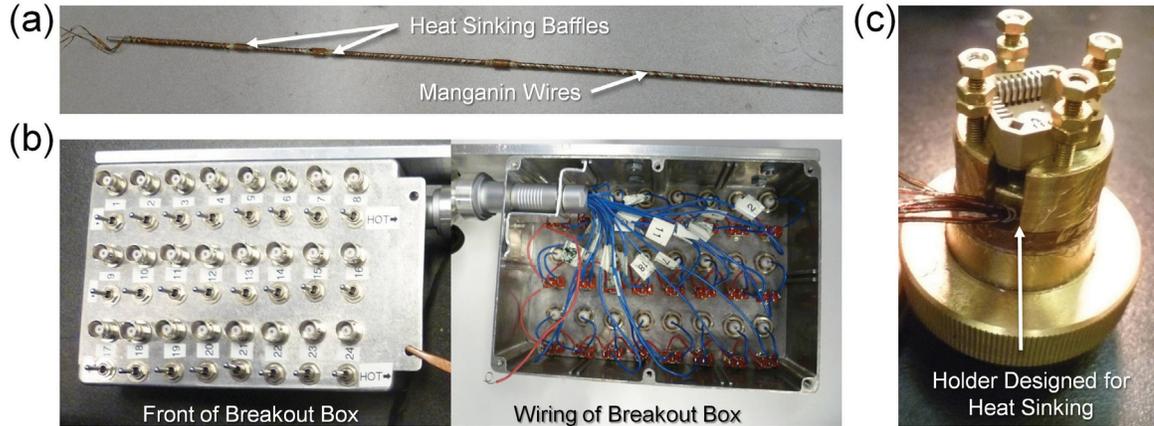
Figure 2.15: Wiring of the probe and sample holder. (a) Wiring of the probe with 24 mangagnin wires. (b) Front and rear of the breakout box which we assembled and wired. (c) Sample holder, which we designed and assembled.

onto the probe for noise reduction. The front and rear of the breakout box are shown in Fig. 2.15(b). We built the breakout box with two-way switches which allow the samples to be either grounded or connected directly to breakout box. Finally a new sample holder, shown in Fig. 2.15(c), was designed to contain the chip-carrier socket, protect the wire bonds, and thermally contacting the wires underneath. Once the probe fully wired, the optimal parameters for cooling to 0.3 K were determined for the new heat load. These settings are described in Appendix A. To prepare samples for use in the cryostat, the further development of other instruments was necessary, which we discuss in the following section.

2.2.2 Additional Instrumentation

Several other instruments in the lab required attention over the course of this degree. The contributions made will be summarized briefly here, starting with the Edwards evaporator (including the repair of a rotary vane pump), and finally mentioning alignment procedures for the Raith e-beam writer.
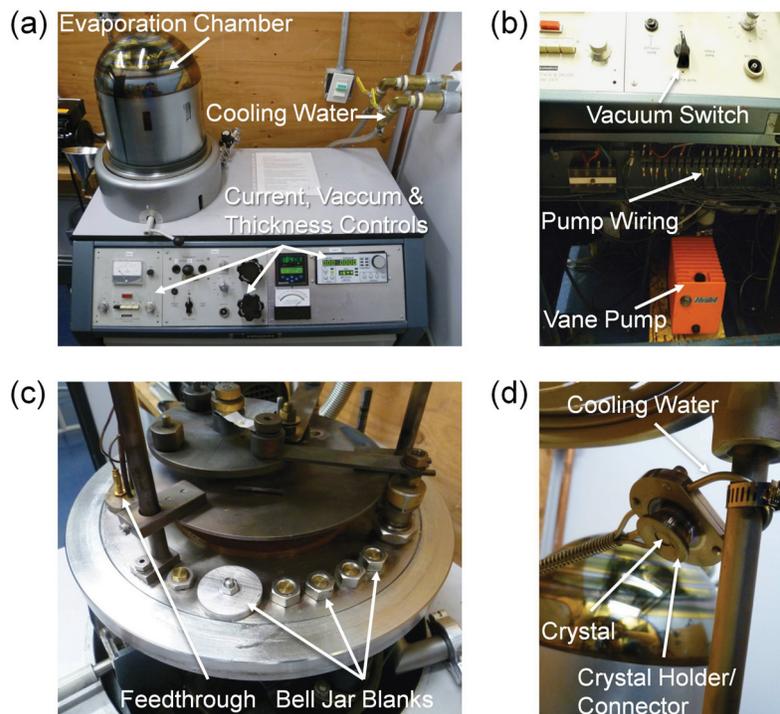
Figure 2.16: Work completed on the Edwards evaporator. (a) Edwards Evaporator. (b) Rotary vane pump, incorporated into the wiring of the evaporator. (c) Vacuum seals for the bell jar with new plugs and O-rings. (d) Crystal thickness monitor sensor head which was designed and soldered in-lab.

The Edwards evaporator was non-functional when I began working on it. The evaporator is shown in Fig. 2.16(a). I replaced the leaking cooling water lines and attached them to our own cooling system, and then refurbished an old Adixen rotary

vane pump. The pump was disassembled and the vanes were removed and replaced. The inside of the pump was cleaned and all O-rings, lip seals, and gaskets were replaced. Electrical connections from the pump were integrated into the evaporator such that the pump could be easily turned on and off with a rotary switch, shown in Fig. 2.16(b). We ensured that the bell jar system and electrical feedthrough were leak-tight by replacing all O-rings and removing unnecessary materials, replacing them with plugs for the baseplate as shown in Fig. 2.16(c). The system was leak tested and found to reach pressures of $\sim 10^{-7}$ with the diffusion pump. Finally, we designed a new crystal thickness monitor system and integrated it into the evaporator. The sensor head is shown in Fig. 2.16(d). I assembled the electronic circuit with the thickness monitor and soldered the sensor head. This allowed us to perform evaporations in our own lab, necessary for photolithography and electron beam lithography.

A new electron beam lithography system became available to the lab and we were in the first generation of trained users. I attended a training course and learned about all aspects of the system. A picture of the new Raith e-beam writer is shown in Fig. 2.17(a). We pioneered two new alignment techniques on this system to create SWCNT break junctions, illustrated in Fig. 2.17(b). The first method is manual alignment. A coordinate system is set up such that corners of ⌐ shapes are equidistant from origin. One corner of each ⌐ is lined up with the coordinate system of the software. If this is done with diligence, it is possible to properly align the break junction on the SWCNT.
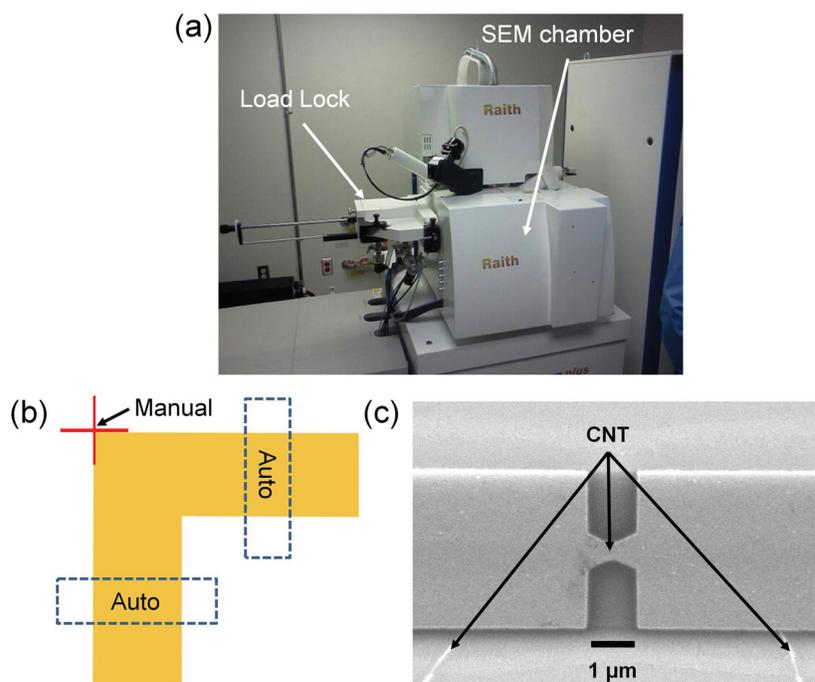
Figure 2.17: Alignment on the Raith electron-beam lithography writer. (a) Raith e-beam writer. (b) Diagram depicting the usage of manual and auto-alignment marks. We pioneered SWCNT break junction alignment using these methods. (c) SEM confirming the resulting alignment, with a large diameter CNT visible inside of the break junction

The second procedure is automatic alignment. For this method, the software locates the center of the ⌐ shapes by using edge detecting algorithms. Thus the coordinate system must be set up so that the widths of these shapes are properly centered in the coordinate spacing. Confirmation that the automatic alignment method works is achieved through SEM and is shown in Fig. 2.17(c). These methods allow us to prepare SWCNT samples, embedded in the gold contacts and ready to be electromigrated.

2.2.3   Measurement Methods

The methods we use to collect data are centered around protecting the sample first, and reducing noise second. We always keep the samples grounded while not being measured, use a personal ground at all times and minimize number of ground loops in our system. To collect data we use a National Instruments data acquisition system (DAQ) with custom measurement software. The circuits we use for both DC and AC measurements are shown in Fig. 2.18(a)-(b). In both cases, the lines from the sample to the Ithaco pre-amp are wrapped in an additional coaxial cable for shielding and kept as short as possible to prevent noise amplification. We optimized the sampling rates, sensitivity, and rise-time averaging on the Ithaco and Lock-in amplifiers, and all pre-amps are zeroed before data taking to prevent offsets. Low pass filters are used on the Keithley to prevent spikes to the gate, and voltage dividers were used when applying bias to gain full DAQ resolution. Data is collected by the DAQ and stored on the controlling computer.
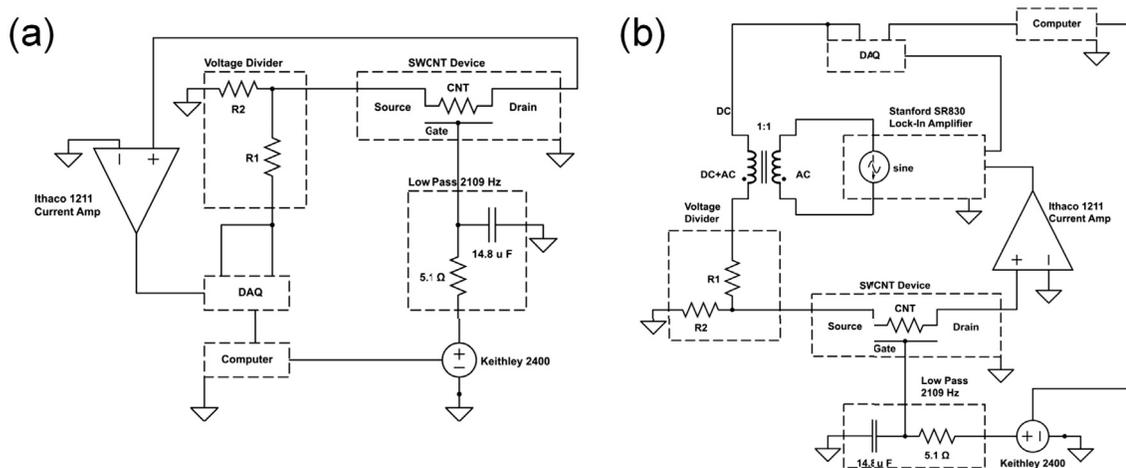
Figure 2.18: DC and AC circuits used for measurements. (a) Electrical circuit for DC measurement. (b) Electrical circuit for AC measurement.

2.2.4 Data Analysis

While collecting data and performing preliminary analysis, we found that much of the graphing and analysis could be readily programmed. I coded several macros with many functions into our graphing software of choice, Igor Pro 6.0. As a result, we can now rebin (average over) both 2D and 1D data, automatically plot collected data, extract all data for a single point in an image plot, extract and plot data along line cuts from colour plots, and plot data on a log scale, amongst other functions. These macros helped to save time and simplify analysis for all lab members. A screenshot of one of these macros, chop_movie, allowing the easy plotting of 1D data along any line cut of a colour plot, is shown in Fig. 2.19. For more information on all of these macros, see Appendix B.
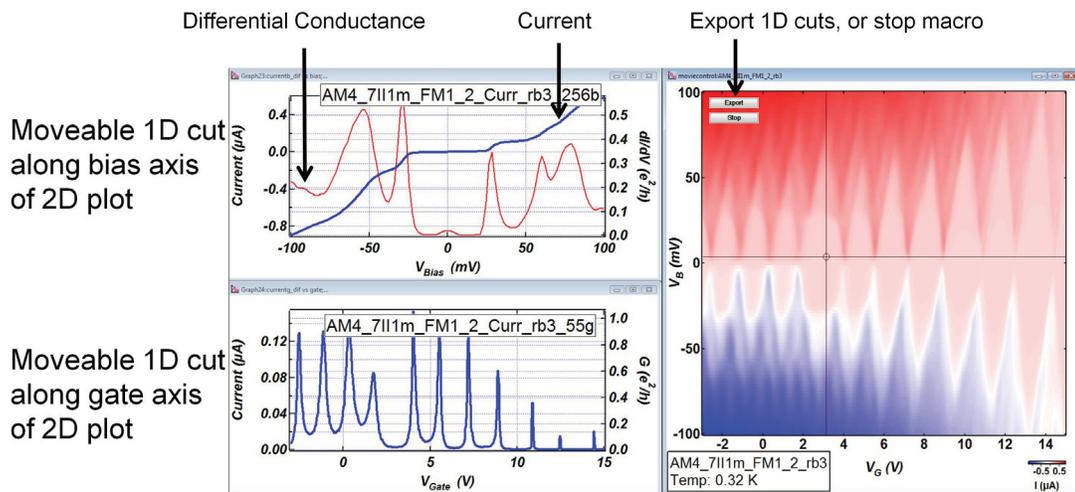
Figure 2.19: Data analysis macros. Screenshot of chop_movie, one of the several analysis macros written using Igor Pro which allows the user to scan through all 1D cuts in a colour plot.

Chapter 3

Electron-Hole Asymmetry in

SWCNT transistors

It has been observed in long (> few hundred nm) semiconducting SWCNT quantum dots that electrons and holes have the same effective mass, energy spectrum and charging energy, as long as the nanotubes are clean and free of impurities [28]. Conversely, in our ultra-short devices, we observe a dramatic and tuneable asymmetry between electron and hole transport. From the regular Coulomb diamond pattern, we determine that our nanotube devices are free of charge impurities and defects, and that the asymmetry is caused by charge doping from the metallic contacts, which alter the tunnel barriers. This asymmetry is reflected in the charging energy, gate capacitance, electronic structure, and conductance of our devices. We measure a difference in electron and hole charging energies up to a factor of 3 in our devices. This charging energy asymmetry could be used to make devices with e-h variable

transistors, photovoltaic devices [7] and improve charge transfer and efficiency in SWCNT transistors [6]. In small bandgap devices, this asymmetry allows us to switch between 0D and 1D transport regimes, which could be used for coherent electronics [29]. By altering the doping of the contacts these asymmetric effects become tuneable, and the difference between electron and hole potentials can be engineered to explore fundamental physics and improve the quality nanoscale electronics devices.

In this chapter, we will first discuss the basic electronic and mechanical structure of carbon nanotubes and the background information necessary to understand quantum dots and Coulomb blockade. We will then describe the different manifestations of *e-h* asymmetry we observe in our devices. For reference, a summary of the devices presented in the following chapters is given in Appendix C.

3.1 Carbon Nanotubes: Background

Carbon nanotubes were discovered as a new isomorph of carbon in 1991 by Sumio Iijima [30]. Since then, carbon nanotubes have been a hot topic of research due to their outstanding mechanical and electronic properties, and their potential for applications as diverse as body armor, solar cells, drug delivery vehicles and much more [31, 32]. In this section, we give a basic overview of CNT structure starting with a 2D planar sheet of atomically thin carbon known as graphene. SWCNTs are made from a rolled up sheet of graphene.
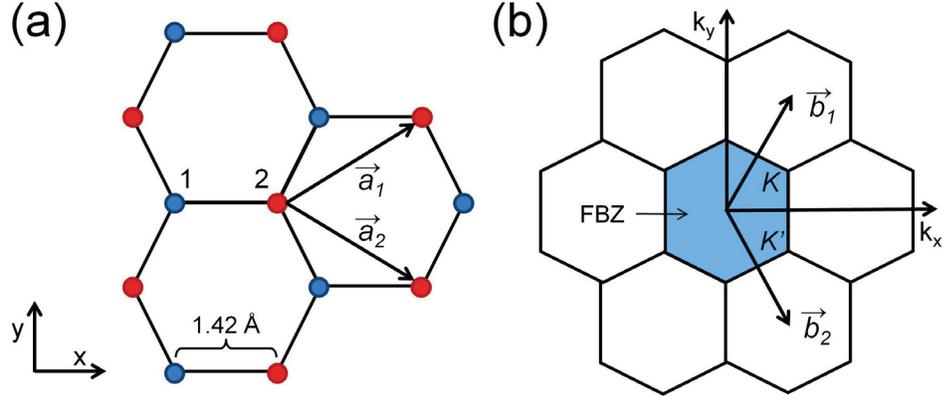
Figure 3.1: Crystal lattice of graphene. (a) Hexagonal crystal lattice of graphene, showing its lattice vectors, two atom basis in red and blue, and C-C bond spacing. (b) First Brillouin zone of graphene (blue), showing lattice vectors in reciprocal space.

3.1.1 Graphene

Graphene was first isolated in 2004 using a scotch tape exfoliation method [33]. The carbon atoms that make up a graphene sheet are sp_2 hybridized, forming three σ-bonds with adjacent carbon atoms. The strength of the σ-bonds is what gives graphene its robust two dimensional structure, forming a planar sheet. What remains is a single free electron in the $2p_z$ orbital for conduction. The trigonal structure of the carbon atoms forms a honeycomb-shaped lattice with C-C spacing $a = 1.42$, and a two atom basis, as shown in Fig. 3.1(a). Translating this lattice into reciprocal space, we recover the hexagonal lattice, rotated by 90°, as shown in Fig. 3.1(b). The lattice vectors of graphene in real and reciprocal space are given respectively by

$$\vec{a}_1 = \frac{a}{2}\left(\sqrt{3}\hat{x} + \hat{y}\right) \qquad \vec{a}_2 = \frac{a}{2}\left(\sqrt{3}\hat{x} - \hat{y}\right) \tag{3.1}$$

and

$$\vec{b}_1 = \frac{2\pi}{\sqrt{3}a}\left(\hat{x} + \sqrt{3}\hat{y}\right) \qquad \vec{b}_2 = \frac{2\pi}{\sqrt{3}a}\left(\hat{x} - \sqrt{3}\hat{y}\right) \tag{3.2}$$

It was Wallace who first derived the electronic structure of graphene in 1947 using the tight binding model [34]. The electrons in graphene are four-fold degenerate owing to the two electron spins and the two equivalent sublattices forming two isospins or equivalently, K-K' valleys. The two atom basis can be treated as a superposition of two hexagonal sublattices which obey Bloch periodicity, giving the electron wavefunction in graphene [34]

$$\psi(\vec{r}) = \sum_1 e^{2\pi i \vec{k}\cdot\vec{r}_1} X(\vec{r} - \vec{r}_1) + \lambda \sum_2 e^{2\pi i \vec{k}\cdot\vec{r}_2} X(\vec{r} - \vec{r}_2) \tag{3.3}$$

where the sums are over each entire sub-lattice, λ accounts for the phase difference between the sub-lattices, and X is the orbital function describing the remaining $2p_z$ electrons. Using the Schrödinger equation, with this wavefunction, we can calculate the band structure of graphene [35]

$$E(k_x, k_y) = \pm\gamma\sqrt{1 + 4\cos^2\left(\frac{\sqrt{3}k_x a}{2}\right) + 4\cos\left(\frac{3k_y a}{2}\right)\cos\left(\frac{\sqrt{3}k_x a}{2}\right)} \tag{3.4}$$

where the $\pm$ sign defines the conduction and valence bands and γ is the hopping integral describing the amount of overlap between nearest-neighbour orbitals. This band structure is shown in Fig. 3.2(a). The excerpt shows the Dirac points, the points

where the valence and conduction bands meet, in detail to the right. Near the Dirac points we calculate a nearly linear dispersion relation, forming cones around these points in momentum space

$$E(\vec{k}) \approx \pm \hbar v_f |\vec{k}| \tag{3.5}$$

where $v_f \approx 10^6$ m/s is the Fermi velocity of Dirac fermions in graphene. We now discuss what happens to the electronic structure of graphene when it is mechanically rolled up on itself, forming a SWCNT.
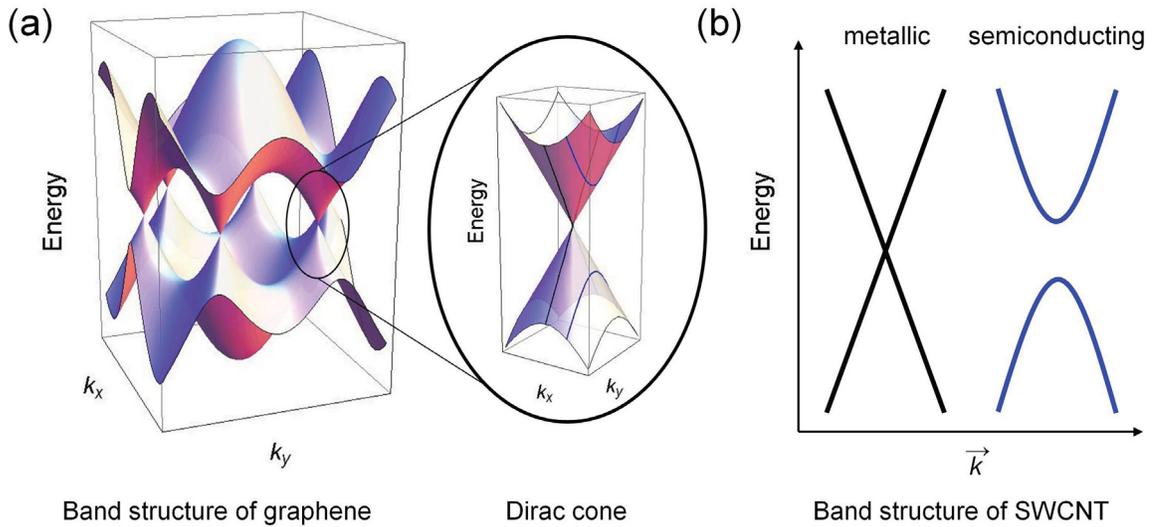
Figure 3.2: Band structure of graphene and carbon nanotubes. (a) Three dimensional band structure of graphene, with an expanded view of the Dirac cone showing planar cuts along the energy axis. (b) Dispersion relation for metallic (black) and semiconducting (blue) carbon nanotubes based on the Dirac cone in panel (a).

3.1.2   Carbon Nanotubes: Electronic Structure

We can understand the origin of the electronic structure of carbon nanotubes by their wrapping [36]. Ultimately, the wrapping determines the position of the planar cuts along the Dirac cone in momentum space, which define the band structure of a carbon nanotube and make it either metallic or semiconducting Fig. 3.2(b). For a given nanotube, two perpendicular vectors describe its mechanical structure: $\vec{T}$, the translational vector, describes the axis around which the nanotube is wrapped while $\vec{C}_h$, the wrapping vector, determines the classification of the nanotube. Fig. 3.3(a)-(b) shows the positions of these vectors on the honeycomb lattice, and a 3D model of a (0,6) zigzag SWCNT respectively.
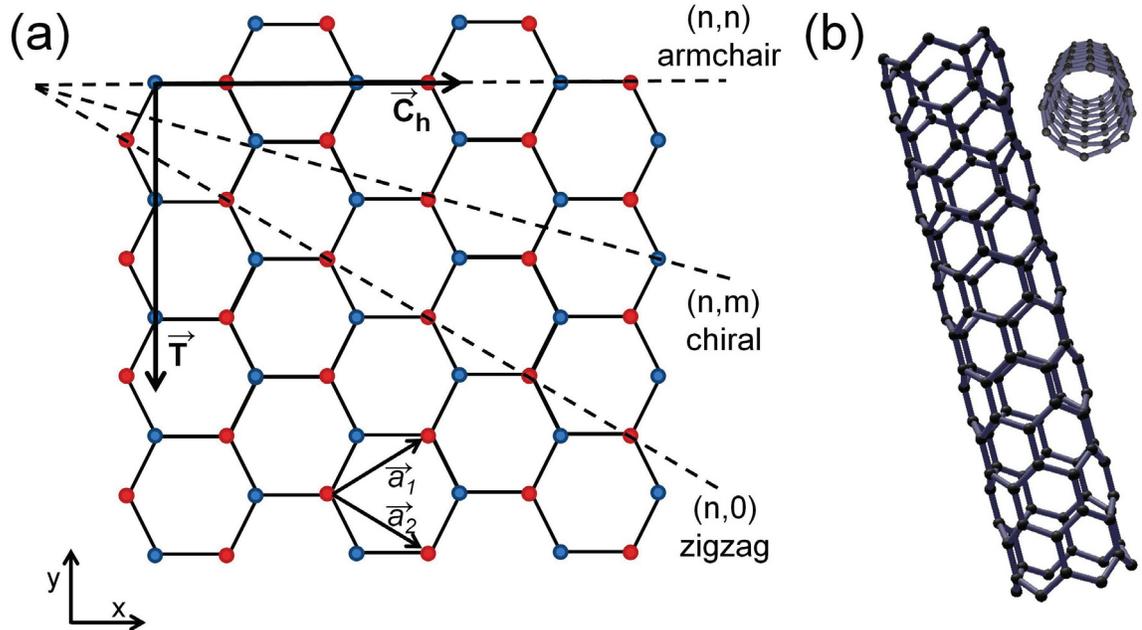
Figure 3.3: Carbon nanotubes from rolled-up graphene.   (a) Real space lattice depicting wrapping vectors for armchair, zigzag, and chiral nanotubes (b) Three dimensional model of a metallic, zigzag carbon nanotube with wrapping vector (6,0).

The chiral vector is defined by

$$\vec{C}_h = n\vec{a}_1 + m\vec{a}_2 \qquad (3.6)$$

If the wrapping vector points along (n,n) the SWCNT is a so-called armchair nanotube, a structure with many C-C bonds parallel to C_h. If the wrapping vector points along (n,0) the SWCNT is known as a zigzag nanotube, its name originating from the zigzag pattern of carbon bonds running across the nanotube. The final class is known as chiral nanotubes with wrapping vectors (n,m).

The quantization of the perpendicular wavevector in the first Brillouin zone upon wrapping the graphene sheet creates the SWCNT band structure from a planar cut of the Dirac cone, causing a tube to be either semiconducting or metallic, see Fig. 3.2(b). Most zigzag and chiral nanotubes are semiconducting, while tubes with wrapping vectors where $2n + m =$ (multiple of 3) have a small bandgap (< 100 meV). Armchair nanotubes are generally metallic, however, curvature induced strain in small diameter nanotubes can introduce small bandgaps [37]. Ignoring strain effects, the bandgap for a given carbon nanotube depends on diameter and chirality, and is given by [38]

$$E_g = \frac{2|q|\gamma a}{d\cos(\theta)} - \frac{q\gamma a^2}{3d^2 \cos^2(\theta)} \qquad (3.7)$$

where $q = 0, \pm 1$ is the remainder of n/3, and $d\cos(\theta) = (2n + m)\sqrt{3}a/2\pi$. If the nanotube has a large diameter, this reduces to an approximate relationship where $E_g \sim 0.7/d$ eV where d is carbon nanotube diameter (in nm) [39].

48

The electronic structure of carbon nanotubes is four-fold degenerate, owing to the two spins of the electrons and two isospin valleys (clockwise and counterclockwise), originating from the two atom basis of graphene. In a one dimensional structure, the maximum conductance that can be achieved is e^2/h due to resistance from the constriction of the dimensions of the system [40]. Because the SWCNT has a four-fold degeneracy, the maximum achievable conductance is $4e^2/h$. This conductance is measurable directly in low contact resistance SWCNTs by measuring current while applying a voltage across it. If we include a capacitively coupled gate electrode, this set-up becomes a SWCNT transistor.

3.1.3 Carbon Nanotube Transistors

It is the bandgap of semiconducting SWCNTs which makes them so exciting in terms of electronics applications such as transistors. A diagram explaining current switching in SWCNT field effect transistors (FETs) is shown in Fig. 3.4(a)-(b). This type of device is both useful as a platform for exploring new physics, as well as an application in and of themselves.

When a semiconducting nanotube is contacted by source and drain electrodes and is capacitively coupled to an additional gate electrode, it forms a nanoscale current switch. When a voltage is applied from source to drain across this device, no current can flow because the density of states in the band gap of the nanotube is zero; the conduction band is empty and the valence band is full. However, by applying a voltage to the gate electrode, we shift the bandgap of the semiconductor.

49
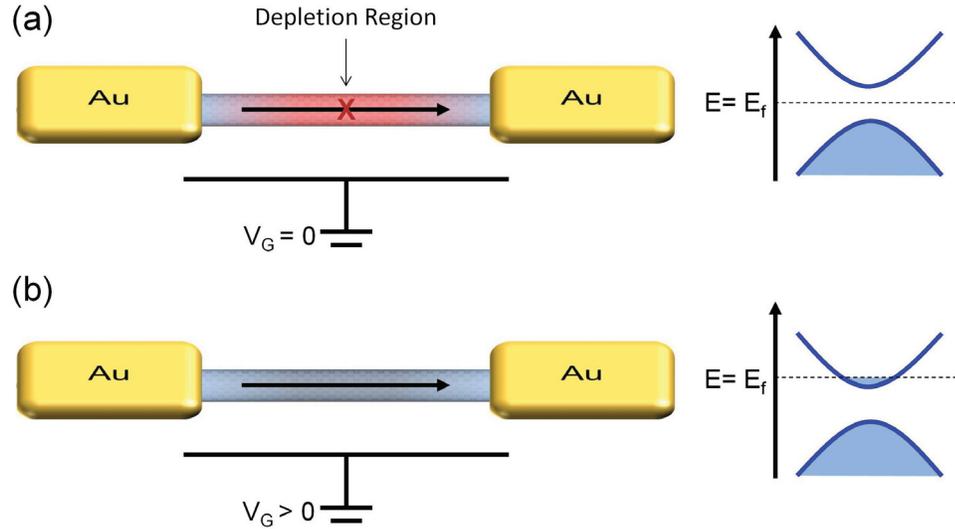
Figure 3.4: SWCNT field effect transistor. (a) In the bandgap of the nanotube there are no charge carriers, creating a depletion region in the nanotube. (b) By applying gate voltage, the bandgap is shifted relative to the Fermi energy of the leads (E_f), adding carriers to the device and allowing current to flow.

This effectively dopes the SWCNT with charge carriers: holes for negative voltage, and electrons for positive. The charge carriers allow the passage of current. We can thus switch the current on and off in a controlled way using the gate electrode. This powerful transistor effect in SWCNTs forms the basis of discussion for the remainder of this thesis. We now discuss the single electron transistor (SET) effect observed in SWCNT-QDs.

3.2   SWCNT Quantum Dots: Background

Before going into detail about our observed *e-h* asymmetry, we will first give some background information on how we determine the electronic structure and effective

50

circuit of our devices. We will first discuss zero-dimensional SWCNT systems and Coulomb blockade, before explaining QD length measurements, electronic energy spectrum measurements, and finally the open quantum dot regime.

3.2.1 SWCNTs in Zero Dimensions

When a one dimensional structure such as a carbon nanotube is reduced to zero dimensions, the system is completely constrained and the density of states (DOS) takes on the form of discrete energy levels. These act as energy levels similar to those in an atom or molecule. This is why these 0D structures, known as quantum dots, are often referred to as artificial atoms. We show this dimensionality reduction in a 1D SWCNT system caused by tunnel barriers in Figure. 3.5(a). The section of the tube which is isolated by the barriers becomes a zero-dimensional system, its energy states become discretized, and a quantum dot is formed.

The effective circuit of a QD is shown in 3.5(b). A quantum dot is structured electronically as three capacitors in parallel from the source (C_S), drain (C_D), and gate (C_G), isolating the electronic island which forms a QD. There is often contact resistance, which acts in series with the source and drain capacitances to reduce the maximum $4e^2/h$ conductance [20].

The energy of a quantum dot is given by [41]

$$U(N) = \sum_{N=1}^{N} E_N + \frac{(Ne)^2}{2C_\Sigma} + NeV_{dot} \tag{3.8}$$
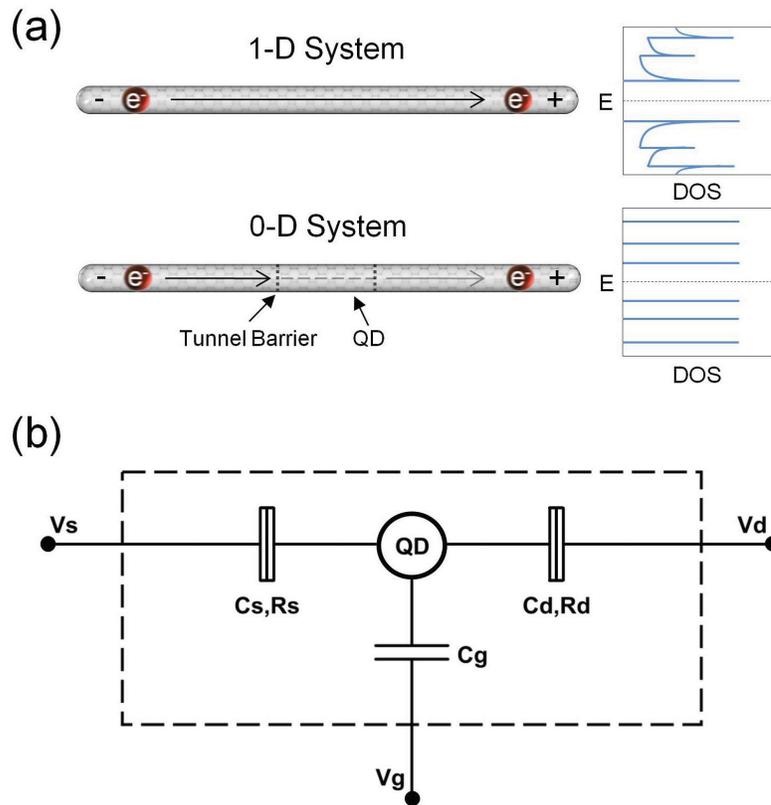
Figure 3.5: From one to zero dimensions in SWCNTs. (a) A carbon nanotube is a 1D system, with a corresponding density of states. When tunnel barriers are added to a SWCNT, a quantum dot is formed with a discretized density of states. (b) Effective circuit of a quantum dot, showing C_S, C_D, C_G, and contact resistance (R_S, R_D).

where E_N is the internal energy of the Nth quantum state of the dot, N is the number of charge states on the dot, $C_\Sigma = C_G + C_S + C_D$ is the total capacitance of system, e is the elementary charge, and $V_{dot} = \frac{C_S V_S + C_G V_G}{C_\Sigma}$ is the total potential of the dot when V_D is grounded, as is the case in all of our devices. Chemical potential, μ_N, is defined as the energy necessary to add the Nth electron to the dot and is easily calculated from eq. 3.8 to be

$$\mu_N = U(N) - U(N-1) = E_N + (N-1/2)\frac{e^2}{C_\Sigma} + eV_{dot} \tag{3.9}$$

and the charge addition energy $\Delta\mu_N$ is given by

$$\Delta\mu_N = \mu_{N+1} - \mu_N = \Delta + E_c \tag{3.10}$$

where Δ is the energy spacing between quantum levels and $E_c = e^2/C_\Sigma$, is the charging energy: the classical energy necessary to charge the QD. These quantities will become very useful in describing the physics of the quantum dot later on.

One of the possible types of tunnel barrier which can form on a SWCNT is known as a Schottky barrier. Schottky barriers forms due to band mismatching between a semiconductor and a metal, as shown in Fig. 3.6, and its height depends on the relative work functions of the metal and semiconductor. Depending on the doping from the contacts, Schottky barriers can form with different heights for electrons and holes. The heights and widths of the Schottky barriers affects the effective size of a quantum dot, as well as its conductance.
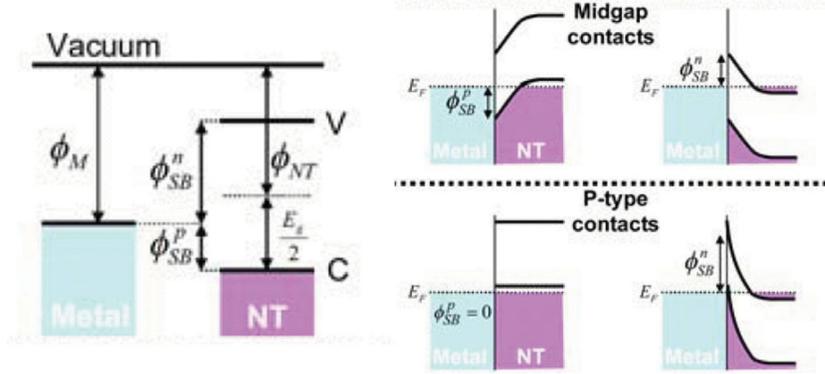
Figure 3.6: Schottky barriers in SWCNTs. Schottky barriers form due to mismatching of the bands in the carbon nanotube and the contacts. Different Schottky barrier heights can form depending on contact doping, applied bias and gate, bandgap, and electron vs. hole transport. Adapted from [39].

Because the carbon nanotube is heavily doped underneath the metallic contacts, charge injection can occur up to tens of nanometers inside the contacts [42]. Therefore, the QD leads are not the metallic contacts themselves, but rather the carbon nanotube underneath the contacts [43]. The equation describing the height of the Schottky barrier is then not dependent on the work function of the metal, but on the metal-modified work function of the carbon nanotube lead, as has been shown in graphene [44]. The barrier height for holes and electrons ($\Phi_{SB}^{P/N}$) is given by [39, 44]

$$\Phi_{SB}^{P} = \Phi_{NT/M} - \Phi_{NT} + \frac{E_g}{2} \qquad \Phi_{SB}^{N} = \Phi_{NT} - \Phi_{NT/M} + \frac{E_g}{2} \qquad (3.11)$$

where $\Phi_{NT/M}$ and Φ_{NT} are the work functions of adsorbed and suspended SWCNT respectively, and E_g is the nanotube bandgap.
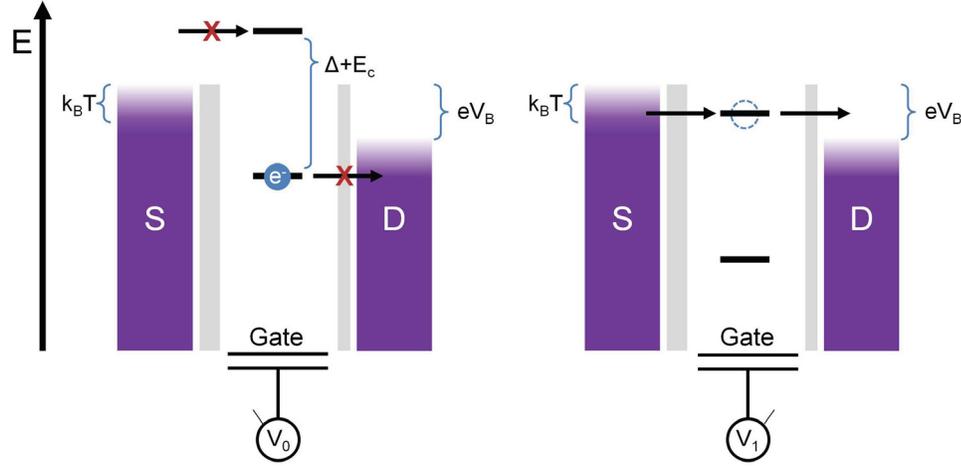
Figure 3.7: A quantum dot single electron transistor. Energy level diagram, illustrating how sequential tunneling is prohibited unless an energy state of the dot resides within the bias window. The applied bias is shown by eV_B, while V_G shifts the energy levels of the dot.

To better understand how current flows through a quantum dot, we use a simple yet illustrative picture known as an energy level diagram, depicted in Fig. 3.7. The Fermi levels of the leads, including applied bias (eV_B) and temperature broadening (k_BT), are shown on the left and right, with asymmetric tunnel barriers blocking the classical passage of electrons. The energy levels of the quantum dot are shown in between. If an energy level of the dot lies below the Fermi energies of the leads, it is always occupied by an electron and there is no room for tunneling. If the energy level of the dot lies above the two Fermi energies, it is impossible for an electron to be excited to that level, and again no current can flow. However, if an energy level of the dot resides between the Fermi energies of the leads, an electron can tunnel from a lead to the dot and then into the opposing lead.

The Fermi levels of the leads are defined by the applied bias voltage (V_B), while

the occupation energy of the dot can be tuned using the gate voltage (V_G). Thus, we can tune the electronic energies of the dot and leads individually, and understand the physical structure of our quantum dots by measuring current.

3.2.2 Coulomb Blockade in Quantum Dots

Coulomb blockade is a phenomenon observed in SETs which appears as a characteristic diamond pattern in a colour plot of current vs. bias vs. gate voltage ($I - V_B - V_G$), as shown in Fig. 3.8. We refer back to the energy level diagram in Fig. 3.7 to understand what happens as we vary bias and gate independently.

In the bias direction, we measure no current when there is no energy state in between the Fermi levels of the leads. However, a sharp increase in current is observed once one of these energy states is encompassed by the leads, as shown in Fig. 3.8(a). This is accomplished by increasing, or decreasing the Fermi level of the source lead (for simplicity, we choose to ground the drain lead and pin its Fermi energy).

As we vary the gate voltage, we sweep the energy states of the dot through the Fermi levels, eventually encompassing an energy state of the leads. Because electron tunneling occurs across each energy state which passes between the leads, we measure pseudo-periodic peaks in current along a gate sweep, as shown in 3.8 (b). If we measure current while simultaneously varying bias and gate voltages, we observe the characteristic Coulomb blockade diamond pattern, shown in 3.8(c).

The stepwise increase in current arising from the opening of new conductance channels can be more easily observed by numerically differentiating current with
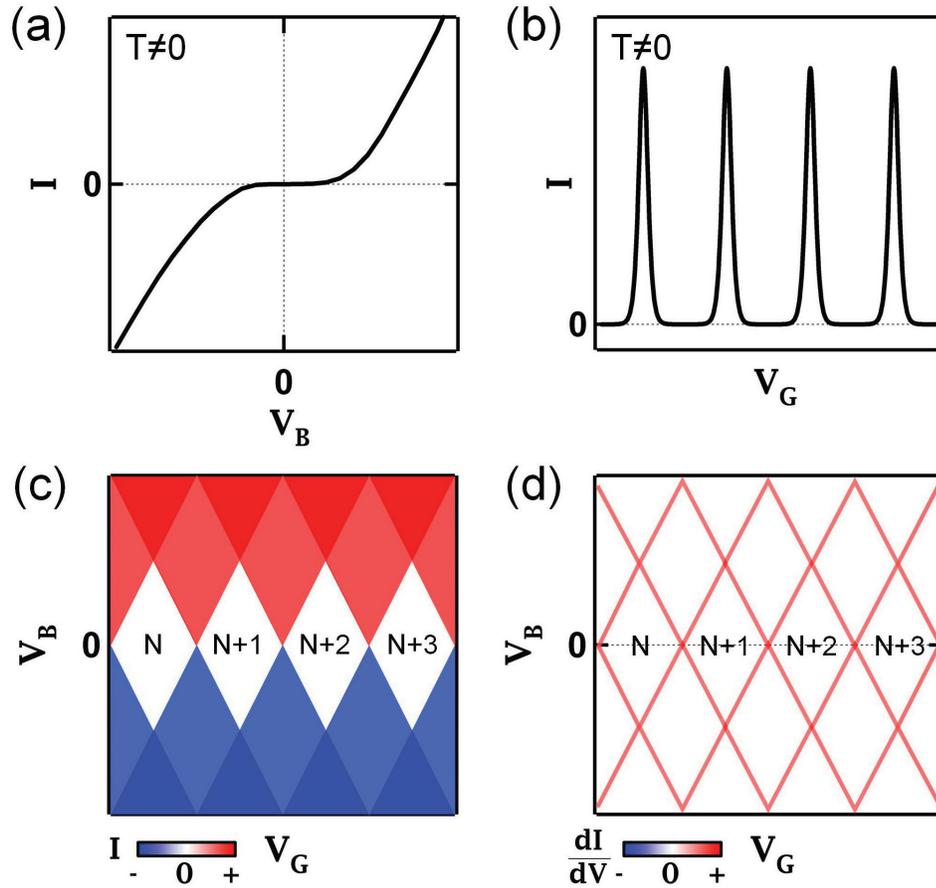
Figure 3.8: Coulomb blockade. (a) Current vs. bias voltage showing Coulomb blockade in a SET. (b) Current vs. gate voltage, showing sweeping of discrete energy states in a SET at finite temperature. (c) Coulomb diamonds result from sweeping V_B, changing the bias window, and V_G, changing the occupation number of the dot, simultaneously while measuring current. (d) Physical features are more visible in differential conductance plots of Coulomb blockade data.

respect to bias voltage, giving sharp peaks of conductance, rather than a smooth increase in current, as shown in 3.8(d). Therefore much of the data in this thesis will be presented as colour plots in $dI/dV - V_B - V_G$ rather than as $I - V_B - V_G$.

The effective circuit of a quantum dot can be extracted from the geometry of the blockade diamonds. In the following discussion, we will assume that all levels are degenerate to give a simple picture of the Coulomb blockade. Further in this chapter we will introduce the quantum confinement energy between electronic energy levels. A more complete discussion of this topic is given in [41]. The way we extract the effective circuit from our data is summarized in Fig. 3.9, with energy level diagrams showing the relative energy levels for different positions in the blockade.
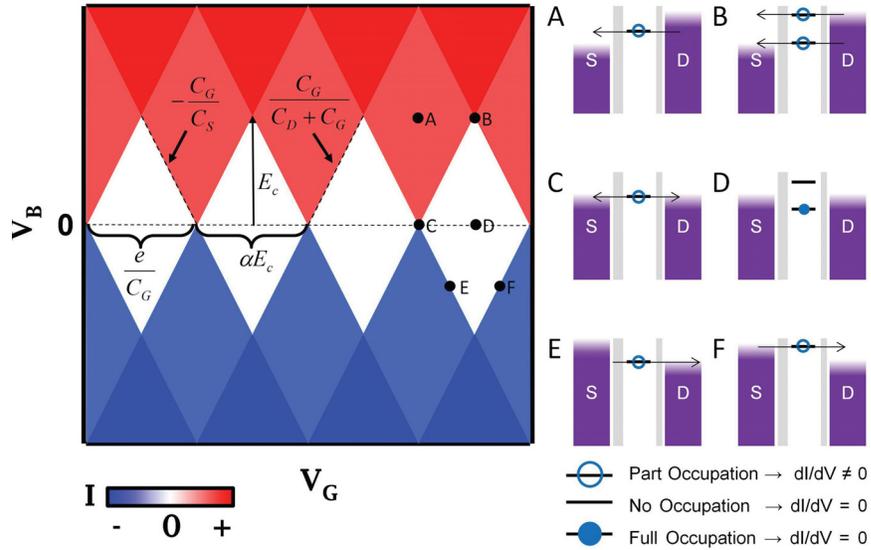
Figure 3.9: Determining the QD circuit from Coulomb blockade. Meaningful physical properties of the quantum dot are extractable from the Coulomb diamonds. Charging energy from diamond height (E_c), and capacitances from diamond slopes (C_S, C_D), and diamond width (C_G). Energy level diagrams for different points in the blockade are shown at points A-F.

The widths of the diamonds are directly dependent on the gate capacitance of the quantum dot, C_G, as given by

$$C_G = \frac{e}{\Delta V_G} \tag{3.12}$$

The charging energy (E_c) is the classical energy needed to add an additional charge onto the quantum dot and is given by

$$E_c = \frac{e}{C_\Sigma} \qquad C_\Sigma = C_G + C_S + C_D \tag{3.13}$$

The charging energy can be read directly from the half height of the Coulomb diamonds. The heights and widths of the diamonds are related by a prefactor (α), which remains constant for all charge states as long as the quantum dot capacitances remain constant

$$\Delta V_G = \alpha \frac{E_c}{e} \qquad \alpha = \frac{C_S + C_D}{C_G} + 1 \tag{3.14}$$

An inequality in source and drain capacitances of the system results in a difference between the positive and negative slopes of the Coulomb diamonds. The source and drain capacitances are given respectively by

$$\Delta V_{SD} = -\frac{C_G}{C_S} \Delta V_G \qquad \Delta V_{SD} = \frac{C_G}{C_D + C_G} \Delta V_G \tag{3.15}$$

The tunneling rate through the quantum dot, Γ, depends on the rate of tunneling through the source and drain leads, Γ_S and Γ_D respectively. These quantities can be determined by measuring the broadening of the source and drain conductance

peaks, corresponding to negative and positive blockade slopes, at full-width-half-max. Together these give the rate of tunneling through the QD

$$\Gamma = \frac{\Gamma_S \Gamma_D}{\Gamma_S + \Gamma_D} \tag{3.16}$$

If the coupling to one lead is much greater than the other, the weak lead acts as a bottleneck and determines the tunneling rate of the system.

It is also possible to determine the bandgap of the nanotube from the Coulomb blockade. If the QD is free of charge impurities, the charge degeneracy point $(N = 0)$ occurs at $V_G = 0$. It is here that we can directly measure the band gap. The energy necessary to add an electron to the $N = 0$ state is equal to the charging energy plus the energy needed to overcome the bandgap: $eV_{add}^{N=0} = E_c + E_g$. Thus, by measuring the height of the central diamond and subtracting the charging energy (measured from the adjacent diamonds), we can determine the band gap of the carbon nanotube.

We observe a falloff of charging energies from the charge degeneracy point. This is because the first few charge states are underscreened, and because the potential of the nanotube is sensitive to a small change in the potential barrier widths for small N [45]. In all of our semiconducting devices, we observe a reasonable agreement between the expected bandgap, based on radius of the nanotubes, and the radii determined from AFM measurements, with a correction of ≈ 1 nm for Van der Waals forces [25]. We can now use the information learned from the blockade diamonds to determine the physical size of the QD.

3.2.3   Length of SWCNT QDs

The majority of the energy scales in our SWCNT devices are length dependent because the size of the QD determines the shape of the electron confinement potential, gate capacitance, and the mechanics of the nanotube. Therefore it is very important that the length of the SWCNT quantum dot be accurately determined. While we can directly measure the length of exposed SWCNT from an SEM image when the tube is visible, screening from the gold contacts and the sheer smallness of the devices, doesn't always make this possible. We find a good agreement between theory and experiment for both longer (> 100 nm) and shorter (< 25 nm) SWCNT devices using a model which treats the nanotube as a wire suspended over a plane [39]. Because our substrate acts as two capacitors in series, the model depends on the thicknesses of the oxide and vacuum layers (t_{ox} and t_{vac}), as well as the dielectric constant of the oxide layer ϵ_{ox}

$$\frac{C_G}{L} = \frac{2\pi\epsilon_{ox}}{\frac{\epsilon_{ox}}{\epsilon_0} \cosh^{-1}\left(\frac{t_{vac}}{r}\right) + \cosh^{-1}\left(\frac{t_{vac}+t_{ox}}{r+t_{vac}}\right)} \tag{3.17}$$

where L is the length of the gateable (exposed) SWCNT, $\epsilon_{ox} = 3.9\epsilon_0$ for SiO$_2$, and $r = d/2$ is the radius of the nanotube. Because gate capacitance depends on length, it becomes apparent that a given device is ultra-short if the diamonds are $\gtrsim 1$ V wide for ≈ 125 nm suspension.

The lengths of our devices can also be confirmed by measuring the energies of the QD excited states. Quantized vibrational mechanical excitations known as vibrons are not directly affected by e-h asymmetry, and will therefore be treated in chapter 4,

but the electronic structure of the quantum dot and its length dependence, will be investigated in the following section.

3.2.4 Electronic Structure of SWCNT QDs

The quantum energy spacing of electrons (Δ), has been ignored up until now for simplicity's sake. This energy arises from the quantum confinement of electrons to a given potential. Several studies have found that the most accurate (and easiest) potential comes from a Dirac particle-in-a-box derivation. The energy spacing for metallic and semiconducting electrons is given respectively by [39]

$$
\Delta = \frac{hv_f}{2L} \qquad \Delta_n = \sqrt{\left(\frac{nhv_f}{2L}\right)^2 + \left(\frac{E_g}{2}\right)^2} - \sqrt{\left(\frac{(n-1)hv_f}{2L}\right)^2 + \left(\frac{E_g}{2}\right)^2} \qquad (3.18)
$$

where $v_f \approx 8.1 \times 10^5$ is the Fermi velocity in carbon nanotubes, n is the electronic energy level number, and E_g is the band gap of the carbon nanotube [39]. Using this model, metallic nanotubes always have the same linear energy spacing, while semiconducting nanotubes have an energy spacing which increases with n. In short devices with a slow varying potential, harmonic or hard-wall models have been employed [28]. To calculate the spacing in these regimes we must know the effective mass of charge carriers in the valence and conduction bands, given by the inverse curvature of the band structure

$$
m_{eff} = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1} = \frac{\left[\left(\frac{hv_f}{2L}\right)^2 + \left(\frac{E_g}{2}\right)^2\right]^{3/2}}{\left(\frac{v_f E_g}{2}\right)^2} \qquad (3.19)
$$

62

For a harmonic potential, the energy spacing is independent of the energy level and is set to be equal to $E_g/2$ at the edges of the QD. The spacing is given by [28]

$$\Delta^{harm} = \hbar\omega_0 = \frac{\hbar}{2}\sqrt{\frac{4E_g}{m_{eff}L^2}} \tag{3.20}$$

Alternatively, the hard wall potential gives an energy spacing which does depend on the energy level number [28]

$$\Delta_n^{hw} = \frac{\hbar^2\pi^2(2n-1)}{2m_{eff}L^2} \tag{3.21}$$

We can measure the electronic energy spacing in our devices directly from the Coulomb blockade, as shown in Fig. 3.10. Because SWCNTs are four-fold degenerate, from the two electron spins and two valley isospins, a new energy level must be opened every $4N^{th}$ charge state. A diagram depicts this phenomenon in Fig. 3.10(a). In panel (b) we show a $I - V_B - V_G$ plot of one of our four-fold degenerate devices, Device A. In the $4N$ charge states, the electron addition energy (eV_{add}) is measured from the heights of the diamonds, and is given by $E_c + \Delta$. For every other charge state, the height of the Coulomb diamonds is simply $V_{add} = E_c$. Therefore, an accurate measure of Δ is attainable by measuring the height of a $4N$ diamond and subtracting the height of an adjacent diamond.

Although this four-fold picture describes the majority of our devices, in others we observe only a two-fold degeneracy, as is illustrated in Fig. 3.11. Mixing of the two degenerate isospin valleys can lift the orbital degeneracy, splitting the two energy
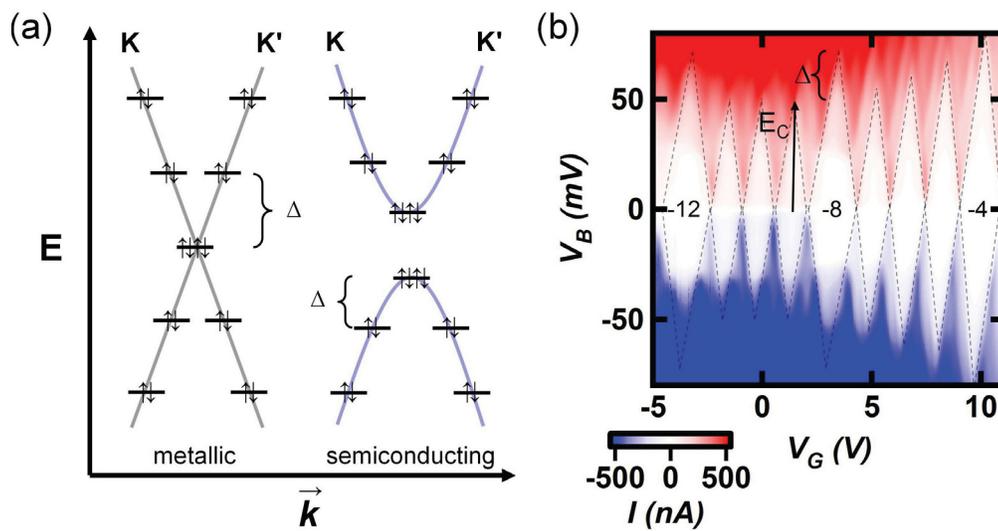
Figure 3.10: Four-fold degeneracy in our SWCNT QDs. (a) Bandgap diagram showing the four-fold degeneracy in metallic and semiconducting carbon nanotubes. (b) In Device A, an increase in charge addition energy from the electronic energy levels (Δ) is visible in every fourth charge state in this devices. The four-fold degeneracy arises from the two spins and two isospin orbitals in a carbon nanotube. Holes are labeled as negative integers.

levels by value $\delta_{KK'}$ [16, 46]. A diagram of the band structure of a two-fold degenerate device with isospin splitting is shown in Fig. 3.11(a). In panel (b) we show a $I-V_B-V_G$ plot of one of our two-fold degenerate devices, Device B. In this plot, we can clearly see the effects of E_c, Δ, and $\delta_{KK'}$ in this device.
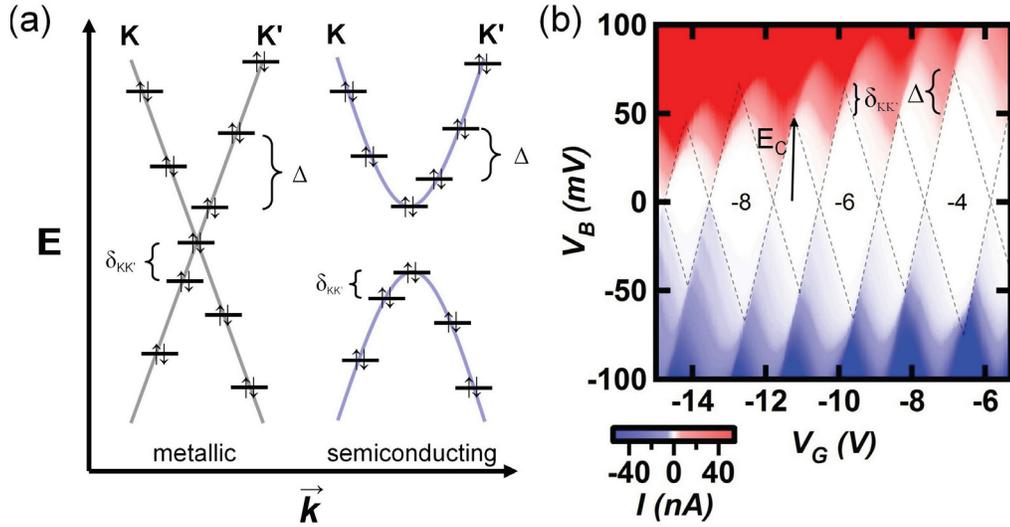
Figure 3.11: Two-fold degeneracy in our SWCNT QDs. (a) Bandgap diagram showing the effect of K-K' scattering, which lifts the isospin degeneracy in metallic and semiconducting carbon nanotubes. (b) In device B we measure a two-fold degeneracy, arising from K-K' scattering. We measure charging energy (E_c) from the odd numbered charge states. The electron spacing (Δ) contributes to the 4N states, but an additional energy to overcome $\delta_{KK'}$ is necessary in the 4N+2 states.

Thus, the effective circuit and energy spacing of a QD can be measured in detail through transport spectroscopy and analysis of the Coulomb blockade. However, if the potential height of the tunneling barriers becomes greatly reduced, a new regime is achieved, from which electronic structure must be determined in a different manner.

3.2.5 Open Quantum Dot Regime

In the case of very small tunneling barriers, the QD becomes what is known as an open dot [47]. Under these high conductance conditions, the tunnel junctions are semitransparent. This causes the nanotube to act as a one dimensional electron waveguide [29], in which the electrons are partially reflected at the contacts and form an interference pattern. This is called the Fabry-Pérot regime, and is only visible in systems where $\Delta > \Gamma > E_c$ [48]. A diagram showing the origin of Fabry-Pérot interferences is shown in Fig. 3.12(a), and the calculated interference pattern from ref. [29] is shown in panel (b). Open dot devices allow us to understand and measure coherent transport in carbon nanotubes by measuring these electron interferences.
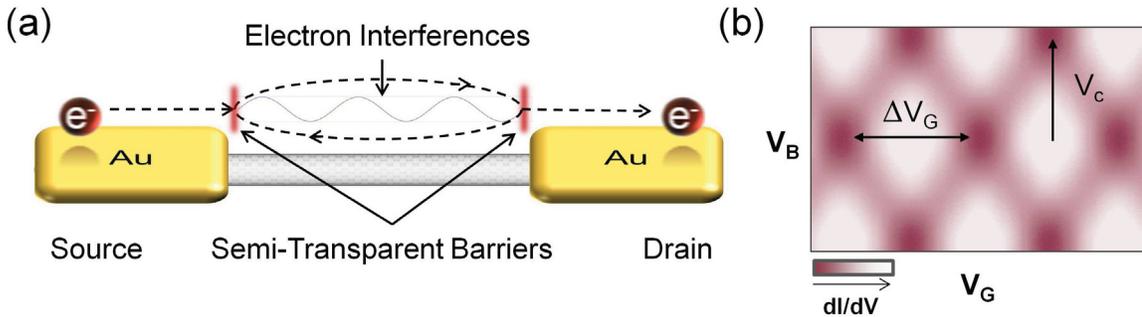
Figure 3.12: Fabry-Pérot interferences. (a) Diagram illustrating Fabry-Pérot interferences. Electrons are reflected at the semi-transparent tunnel barriers and interfere constructively or destructively depending on their energy and phase. (b) $dI/dV - V_B - V_G$ plot, showing interference pattern and relevant voltage spacings. Adapted from [29].

From the interference spacing we can learn about the energy spectrum and size of the quantum dot. The bias at which the first crossing point of positive interference occurs (V_c), corresponds to a single electronic energy spacing, with energy levels

aligned with each of the source and drain leads [48]. This alignment results in constructive interference and a conductance maximum is achieved. This can allow us to measure Δ [29]

$$V_c = \frac{\Delta}{e} \tag{3.22}$$

The interference pattern can also be adjusted by shifting the energy of the QD using the gate electrode. The gate dependence of this interference pattern is described by

$$\Delta V_G = \frac{4e}{C_G} \tag{3.23}$$

This cross-hatching pattern is a distinctive feature and can be used, much like Coulomb blockade, to learn about the energy spectrum of the quantum dot.

We now have the theoretical understanding of electron transport spectroscopy and Coulomb blockade to understand the data collected from our QDs. We now use this understanding to investigate the strong electron-hole asymmetry we observe in our devices.

3.3 Electron-Hole Asymmetry

Symmetry between electron and hole states has been observed in long (> 200 nm) metallic and semiconducting nanotubes [28]. However as semiconducting devices become shorter, the leads play a proportionally larger role, and any *e-h* asymmetry due to charge doping from the metallic contacts becomes magnified. In our devices we

observe several dramatic manifestations of this charge carrier asymmetry, including a factor of 3 difference between electron and hole charging energies, and a tuneability between 0D and 1D behaviour. Since the confinement potential is different for electrons and holes, we can effectively tune between two different transistors in the same device. In addition, because the asymmetry is heavily dependent on the contact region, we can engineer *e-h* asymmetry in future devices by altering the properties of the contacts.

We will first discuss the asymmetries observed in the charging energies and widths of the Coulomb diamonds for electrons and holes, before exploring the ultra-short room temperature SWCNT-FET regime in a highly asymmetric device. We will then investigate conductance asymmetry which allows us to probe different tunneling regimes by switching between positive and negative charge carriers.

3.3.1 Asymmetry in Charging Energy and Blockade Diamond Width

We observe a clear change in charging energy from holes to electrons in our ultra-short devices. This asymmetry is strong, with the ratio between charging energies for holes and electrons, above 300% in Device B.

Because we observe that electrons have a lower charging energy and higher conductance than holes, we can conclude that the electron tunneling barriers are smaller, and therefore the contacts are n-doping the carbon nanotube [45]. In the

diagrams in Fig. 3.13(a)-(b) we show hole and electron transport respectively for n-doping leads. N-doping raises the Fermi levels of the leads in relation to the mid-gap of the nanotube, resulting in a barrier junction which is smaller for electrons than for holes. From eq. 3.11, we can understand that the magnitude of the tunnel barriers depends on the bandgap of the nanotube, but their asymmetry depends on the work function of the carbon nanotube under the metallic leads [39, 44].
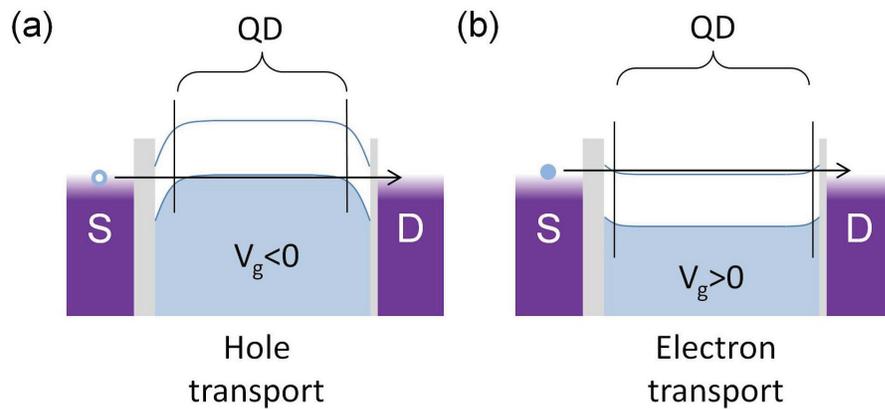
Figure 3.13: Diagram illustrating the origin of e-h asymmetry. Charge doping from the leads shifts their Fermi level relative to the bandgap. In this n-type device, the tunnel barriers are larger for hole transport (a), than electron transport (b).

We show this e-h symmetry about the third charge state in differential conductance colour plots in Fig. 3.14(a)-(b) for Devices B and C respectively. The lengths of these devices are $\approx$ 15 nm and 45 nm. The heights of the diamonds (V_{add}) are clearly different for electrons and holes (e^-/h^+).

One way to quantify this asymmetry is by taking a ratio of diamond heights (V_{add}) or widths (ΔV_G) for holes vs. electrons. We plot the dependence of diamond width
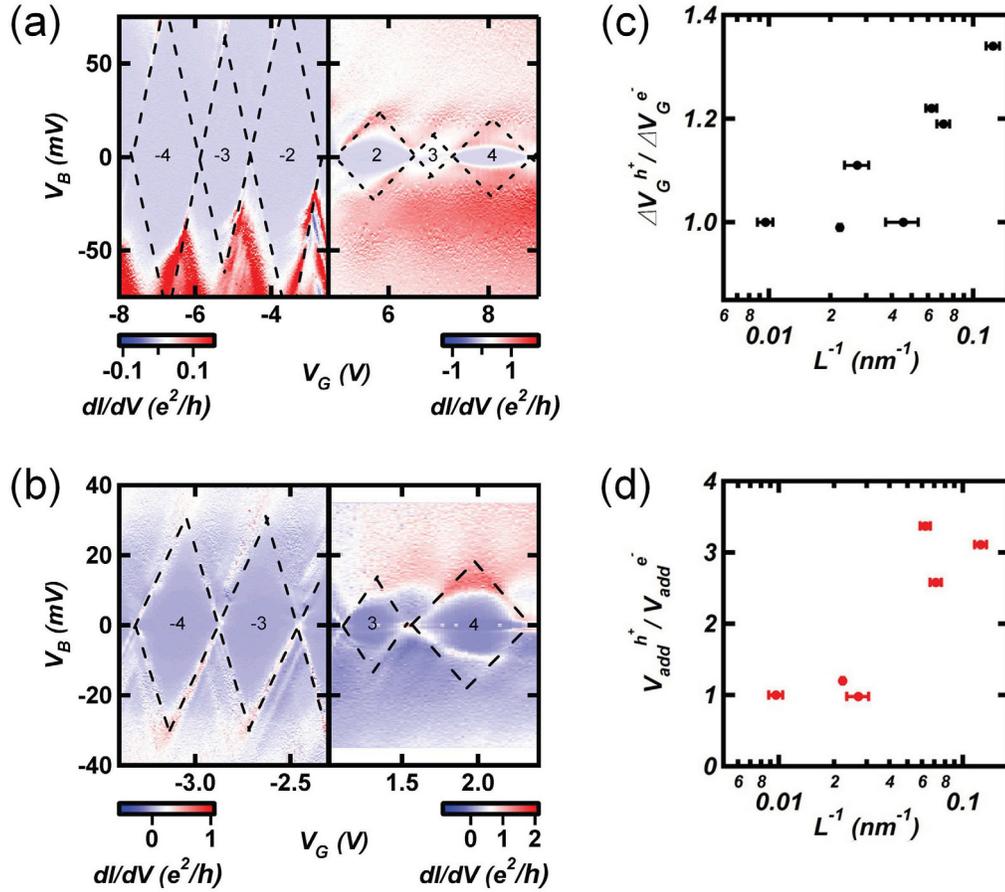
Figure 3.14: Charging energy asymmetry. (a)-(b) Differential conductance plots of Devices B and C with lengths $\approx$ 15 nm and $\approx$ 45 nm respectively, showing clear electron hole asymmetry in diamond heights (V_{add}). (c)-(d) Dependence of average fractional *e-h* asymmetry of V_{add} and ΔV_G vs. length for several of our devices. We note a strong asymmetry ratio in V_{add}, up to a factor of 3.

asymmetry ratio ($\Delta V_G^{h^+}/\Delta V_G^{e^-}$) and addition energy asymmetry ratio ($V_{add}^{h^+}/V_{add}^{e^-}$) vs. L^{-1} in Fig. 3.14(c)-(d) respectively. We observe a weak asymmetry dependence of the diamond widths, corresponding to an overall change in channel length of ≈ 1 nm from electron transport to hole transport. This arises from a difference between the Schottky barrier heights, and therefore tunnel barrier widths, for electron vs. hole transport. In ultra-short (≈ 10 nm) devices, this small change corresponds to a difference in length of $\approx 10\%$, while in longer devices this effect is proportionally much weaker. This explains why $e\text{-}h$ asymmetric channel lengths have not been resolved before in previous SWCNT QD research.

The charging energy $e\text{-}h$ ratio (approximately equal to the V_{add} ratio) has a much stronger length dependence than the diamond width ratio. Recall that $E_c \propto C_\Sigma^{-1}$ and $C_G \propto L$. As length decreases into the ultra-short regime, C_G becomes negligible compared to C_S and C_D, and these latter quantities dominate E_c. Because C_S and C_D depend predominantly on the $e\text{-}h$ asymmetric Schottky barriers at the contacts, the charging energy becomes asymmetric as well. The asymmetry is present even in longer devices if there is any contact doping, but is reduced in magnitude by the gate capacitance to the point where it cannot be resolved. An alternative way to think of the length dependence of the asymmetry is in the relative size of the contact barriers to the full length of the QD. The tunnel barrier sizes do not depend on SWCNT length. Therefore, in ultra-short devices the barriers cover a large portion of the carbon nanotube, magnifying the asymmetry. Note that for none of our samples is the hole channel longer than the electron channel.

The n-type behaviour observed in our devices is contrary to the p-type behaviour normally observed in gold-contacted SWCNTs. However, n-type carbon nanotubes with gold contacts have been achieved in the past through temperature annealing [45, 49]. Annealing gold contacts has been found to remove p-doping impurities such as oxygen, causing a steady n-doping of the contacts by lowering their work function, affecting transport in the SWCNT. In our devices, we likely observe n-doping because in short suspended devices electromigration anneals the contacts as well as the carbon nanotube. This conclusion can be further corroborated when we warm an ultra-short device to room temperature, as we will see in the following section.

3.3.2 Ultra-short Room Temperature SWCNT Transistors

We observe a switching of n-type behaviour to p-type behaviour in Device D which we warm up to room temperature. We determine the length of this device to be ≈ 5 nm from its gate capacitance and observe an n-doping effect at low temperature with higher conductance for electrons than for holes, as shown in Fig. 3.15(a). However, at room temperature we find that the contacts p-dope the device, with higher conductance for holes than electrons as shown in Fig. 3.15(b). This change in doping can be explained in the following way. As we warm our sample to room temperature, the dopants which were annealed out of the device during electromigration, as well as any gases adsorbed to the inner walls of the sample space are desorbed. These impurities can then re-dope the contacts, changing their work function and resulting in a distinctive switch to p-type behaviour.

72
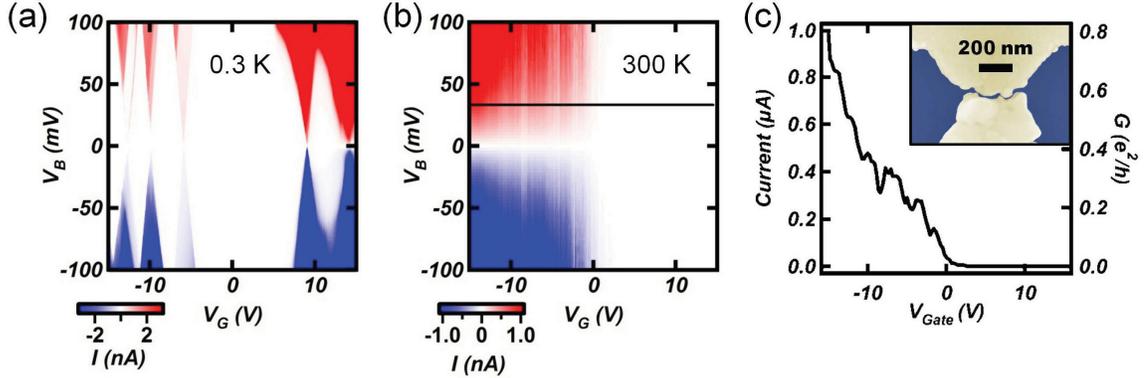
Figure 3.15: Ultra-short SWCNT transistor at 300 K. (a) Low temperature $I-V_B-V_G$ plot, showing n-type behaviour. From diamond widths $L_G \approx 5$ nm. (b) Room temperature $I - V_B - V_G$ plot, showing p-type behaviour. (c) Gate cut along the black line in (b), showing a strong transistor effect with on/off ratio of > 4100. Inset: SEM image of this device with $L_{SEM} \lesssim 10$ nm.

In this sub-10 nm device, we observe a strong transistor effect, with an on/off ratio greater than 4000 at room temperature, as shown in Fig. 3.15(c), taken along the horizontal black line in panel (b). An SEM micrograph of this device is shown in the inset. This device cannot be resolved by SEM measurements, and is therefore $\lesssim 10$ nm in length. It has been predicted that in such short devices, a strong transistor effect is not possible, due to direct source-drain tunneling across the QD [2]. However, our device, much like the 9 nm device from ref. [4], shows otherwise.

The strong conductance asymmetry observed in our room temperature measurements is also reflected in devices with very small bandgaps at low temperature. This can actually result in a difference between electron and hole tunneling regimes, where electrons behave as though they are in a 1D channel, and holes remain in a 0D single electron transistor.

3.3.3   Conductance Asymmetry: Open QD

In metallic carbon nanotubes, there are no Schottky barriers and tunnel junctions arise from defects at the contacts [39]. If the metallic device is shorter than the coherence length of the electrons, it behaves as an electron waveguide and Fabry-Pérot interferences emerge [29]. The electronic energy spacing in metallic Fabry-Pérot devices is the Dirac particle-in-a-box potential, $\Delta = hv_f/2L$. In semiconducting nanotubes, the Schottky barriers are normally too large to be semi-transparent and Fabry-Pérot interferences are not present. However, if the bandgap is small, the barriers are reduced, making the interface behave more like a metallic junction, and allowing one dimensional transport.
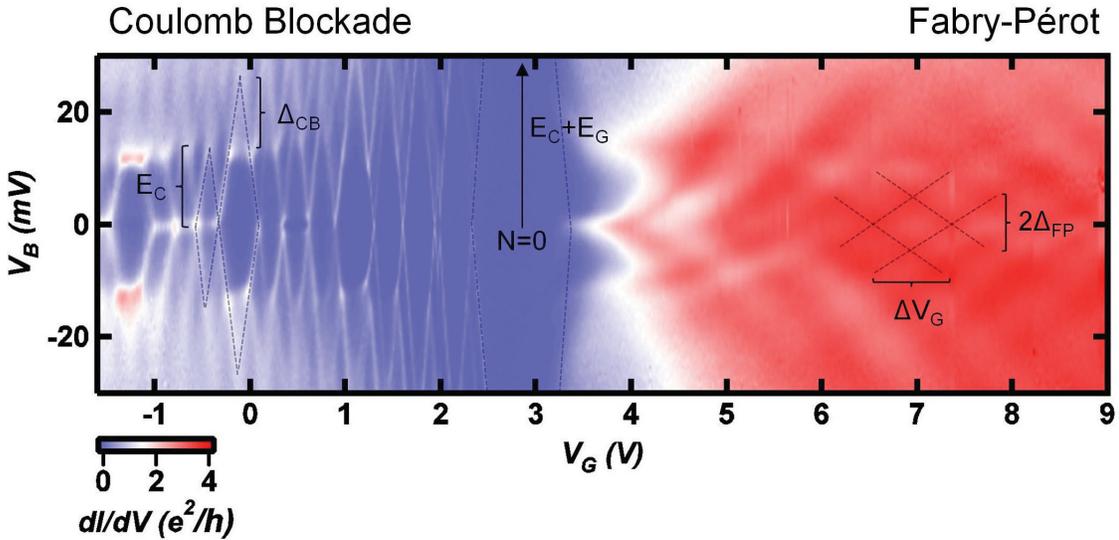
Figure 3.16: Gate tuneable tunnel barriers. (a) One of our high conductance SWCNT devices ($\approx$ 100 nm) showing Fabry-Pérot interferences on the electron side and Coulomb blockade on the hole side. We extract $C_G = 0.78 \pm 0.07$ aF, $\Delta = 13.5 \pm 1$ meV, $E_c = 8 \pm 1$ meV and $E_g = 54 \pm 10$ meV.

We show data from Device E, one of our high conductance devices, in Fig. 3.16(a), with Fabry-Pérot interference on the electron side and Coulomb blockade on the hole side. This behaviour indicates n-doping from the annealed contacts. This conductance asymmetry has been observed previously in longer small-bandgap devices [47], but was not resolved as clearly. Devices do not need to be ultra-short to observe this effect because coherence length is long (> 600 nm) in SWCNTs [29].

From the Coulomb blockade width, we determine the gate capacitance of this sample to be $C_G = 0.78 \pm 0.07$ aF corresponding to a QD of length $L_G = 102 \pm 9$ nm, in good agreement with SEM measurements $L_{SEM} = 104 \pm 9$. From the heights of the diamonds we measure electron spacing $\Delta_{CB} = 13.5 \pm 1$ meV, corresponding to a Dirac particle-in-a-box QD length $L_\Delta = 124 \pm 9$ nm, in reasonable agreement with the other length measurements. We measure the charging energy from the heights of high occupancy odd charge states $E_c = 8 \pm 1$ meV, and extrapolate the bandgap at $N = 0$ to be $E_g = 54 \pm 10$ meV, indicating that this is a small-bandgap nanotube.

We can now compare our device measurements between the Coulomb blockade diamonds and the Fabry-Pérot interferences. From the gate dependence of the oscillations, using C_G from the blockade diamonds on the hole side, we calculate a very similar quantum dot length $L_G = 100 \pm 20$ nm. However the bias crossing point, V_c, gives an electronic energy spacing $\Delta_{FP} = 4.7 \pm 0.6$ meV, quite different from the electron spacing measured for holes. Using the Dirac particle-in-a-box potential, this corresponds to a length of $L = 350 \pm 50$, much longer than any of the other measurements made for this device.

75

This incongruity could be due to the fact that the shape of the potential is different for electrons and holes, resulting in different confinement energies. For example if we use a hard-wall potential to determine the nanotube length from V_c on the Fabry-Pérot side, $L \approx 86$ nm, in much better agreement with the other measurements.

Another possibility is this effect is caused by the extension of the quantum dot under the n-doping leads because the tunnel barriers are so small. If the SWCNT extends far under the leads, it is conceivable that electron injection could occur far under the leads, increasing the length of the waveguide and reducing the electron energy spacing. Screening from the leads prevent this area from being gated, so this effect would not be reflected in the width of the Fabry-Pérot interference pattern.

We have shown in our devices that a strong electron-hole asymmetry arises from doping by the metallic contacts. The magnitude of the asymmetry depends on the Schottky barrier difference between electrons and holes, and the length of the SWCNT device. This can result in up to a factor of 3 difference in charging energy for electrons and holes in our devices. We also observe a strong asymmetry in terms of transport regimes, where electrons behave as though they are in a 1D system and holes behave as though they are in a 0D system. By studying this asymmetry, we can understand electron and hole injection separately and engineer more efficient transistors, gain a better understanding of the scaling limitations of ultra-short nanotubes [2], and develop new types of devices [7]. Now that we have an understanding of contact interactions, we can investigate strong electron-electron and electron-vibron interactions, which are enhanced by the short length of our devices.

Chapter 4

High Kondo Temperature, Self-Actuated Bending Vibrons, and Reduced Charging Energy

We have explored the downscaling of SWCNT transistors to the 10-nm scale. Because these devices are so short, we observe an increased strength of quasi-particle interactions. We now focus on the effects of strong electron vibron (e-v), and electron-electron (e-e) interactions in our devices. We observe three main effects in ultra-short SWCNTs from these interactions.

The first is a strong e-e interaction, resulting in the Kondo effect, which can be used for molecular spintronics applications [50], or as a tool to measure the electronic spectrum of the QD [16]. In our short devices e-e coupling is enhanced and we measure Kondo temperatures up to $T_K = 28$ K, which is to our knowledge a factor

of 2 larger than the highest T_K observed to date in SWCNTs [14].

The second is a strong $e\text{-}v$ coupling to the bending mode of the SWCNT. We have recently shown that ultra-short carbon nanotubes can form self-actuated nanoelectromechanical systems (NEMS) and measured self-actuated bending frequencies up to ≈ 280 GHz using DC transport [11]. The short length and strain tuneability of these devices make them promising candidates for supersensitive mass or force sensors, with down to yoctogram sensitivity [51]. Due to strong $e\text{-}v$ coupling We observe a second harmonic of the bending mode in our devices, and its low-bias onset caused by increased conductance from the Kondo effect.

Finally, we measure a device with both the Kondo effect and bending vibrons. As a result of strong $e\text{-}v$ coupling, a reduction in the classical energy needed to charge the QD is observed. This effect is so strong that we observe a near complete suppression of the charging energy. Although predicted as early as 2004 [18], this is the first time this effect has been observed experimentally to our knowledge. These types of devices could be used as NEMS sensors, current rectifiers, or in spin transport applications such as spin filters [9, 19, 50].

We will first discuss some background information on the Kondo effect before highlighting the very strong Kondo effect we observe in our devices. We will then give background information on vibronic interactions in SWCNT QDs before discussing bending oscillations and the observation of the second bending harmonic that we observe in our devices. We will finally discuss how these strong $e\text{-}e$ and $e\text{-}v$ interactions reduce the electronic charging energy in our devices.

4.1 The Kondo Effect: Background

The Kondo effect has been well documented in several molecular systems, including SWCNTs, and in semiconductor two dimensional electron gas system [15, 52]. This electron-electron (*e-e*) interaction enhances the likelihood of cotunneling through the coupling of electron spins in the QD and the leads. This can serve as a tool to measure excited states of the dot inside of the blockade, and can be used to develop coherent spintronics devices in SWCNTs such as spin filters [9, 53].

In this section we first describe the theory of Kondo effect with SU(2) symmetry, before discussing its length and temperature dependence. We will follow up by discussing the SU(4) Kondo effect and the various Kondo-enhanced inelastic cotunneling processes which occur at finite bias in our devices.

4.1.1 Origin of the SU(2) Kondo Effect

In a given charge state diamond with no bias voltage applied, a quantum dot normally resides within the Coulomb blockade; no energy states of the QD are aligned with the leads and current is blockaded as explained in section 3.2.2. Therefore, in general, we expect to see no conductance at low bias inside of the Coulomb diamonds. When a single electron or hole occupies an energy level of a quantum dot, the charge behaves a spin impurity, which can be screened by the spins in the leads, increasing the likelihood of cotunneling.

Cotunneling is a second order tunneling process which depends on two

simultaneous and coherent tunneling events. An electron from one of the leads tunnels onto the QD, while the electron occupying the dot tunnels onto the opposite lead within the uncertainty time. Because energy is conserved in a cotunneling process, it is not restricted by the Coulomb blockade and appears as a very small zero-bias conductance peak for the odd numbered states of the QD. This second order process is much less probable than first order sequential tunneling across a quantum dot. As a result, these zero-bias conductance peaks cannot normally be resolved, even at temperatures as low as ~ 10 mK.

As shown in Fig. 4.1(a), the spin of the electron on the QD can form an entangled singlet state with a spins in the leads. These electrons are now coupled together, increasing the likelihood of cotunneling and rendering these zero-bias peaks observable. This can be thought of as a cloud of spins on the lead screening the spin impurity on the QD, reducing scattering and allowing the flow of current. This process is known as the SU(2) Kondo effect, whose name comes from group theory nomenclature, and describes the 2D nature of the $S = \frac{1}{2}$ state [54]. Each cotunneling event causes the spin of the dot to change between $S = \frac{1}{2}$ and $S = \text{-}\frac{1}{2}$, as shown in Fig. 4.1(a). For this reason, the Kondo effect is known as a spin-flip process. In differential conductance plots, the SU(2) Kondo effect is manifested as a zero-bias conductance peak with a maximum value of $2e^2/h$ at the oddly occupied states of the QD as shown in Fig. 4.1(b).

The strength of the Kondo effect is characterized in terms of temperature. At higher temperature, the entanglement between dot at lead is disrupted by thermal
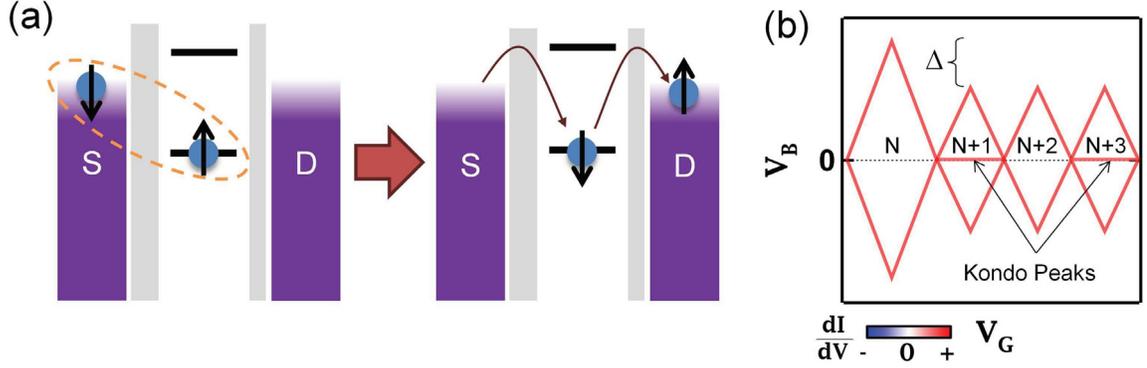
Figure 4.1: Diagram illustrating the Kondo effect. (a) Coupling between an electron on the QD and an electron in the leads can result in an entangled state which increases the likelihood of cotunneling at zero bias. (b) Schematic, showing a Kondo resonance at zero bias for odd occupation.

fluctuations. The Kondo temperature (T_K) is the characteristic temperature at which the Kondo effect becomes observable in a given system. The expected strength of the Kondo effect is derived from Anderson impurity model using perturbation theory with a flat-band assumption [55], giving an analytical expression for the Kondo temperature in the center of the blockade [53]

$$T_K = \sqrt{\Gamma E_c} \exp\left(-\frac{\pi}{4}\frac{E_c}{\Gamma}\right) \tag{4.1}$$

where Γ is electron tunneling rate and E_c is charging energy. Γ is determined by the nature of the tunneling barriers and therefore the bandgap of the SWCNT. The term in the exponent is roughly independent of length, so the strength of the Kondo temperature depends on the product $\sqrt{\Gamma E_c}$. Thus, we expect to observe a stronger Kondo effect in small-bandgap devices, with large Γ. Because E_c increases with

decreasing L through C_G, we expect to observe a stronger Kondo effect in shorter devices, especially when $C_{S/D}$ are small.

The characteristic volume over which Kondo interactions take place in a system is known as the Kondo cloud, whose length scale, (ξ_K), depends on the Kondo temperature: $\xi_K = \hbar v_f / k_B T_K$, where $v_f = 8.1 \times 10^5$ m/s is the Fermi velocity in SWCNTs [56]. In carbon nanotubes with Kondo temperatures on the order of 1.0 K the Kondo cloud is $\xi_K \sim 1$ μm, extending far under the electrodes of our devices. For higher Kondo temperatures the Kondo cloud is reduced in size and therefore requires shorter QDs.

The Kondo temperature can be determined by measuring the height of the differential conductance Kondo peak while varying temperature. To determine the Kondo temperature, this data is fit using the Goldldhaber-Gordon model [57]

$$G(T) = G_0 \left(1 + (2^{1/s} - 1) \left(\frac{T}{T_K} \right)^2 \right)^{-s} \tag{4.2}$$

where G_0 is the conductance of the peak at $T = 0$, and s is a fitting parameter which depends on the spin of the impurity and its screening. For spin-$\frac{1}{2}$ systems, $s = 0.22 \pm 0.01$. It has also been shown that the Kondo temperature can be extracted from the half-width half-max (HWHM) of the Kondo peak at low temperatures [58]. In SWCNTs, the strength of the Kondo effect is greatly increased by the four-fold degeneracy of the electrons, creating a hyperspin interaction and increasing the symmetry of the Kondo resonance to SU(4).

4.1.2 SU(4) Kondo Effect and Zeeman Splitting

In carbon nanotube systems the Kondo effect is not only limited to spin flip-processes: an entanglement between orbital states (the two isospins), occurs as well [17]. The four possible cotunneling pathways become entangled in a four component hyperspin (spin $\otimes$ isospin) configuration which is shown in Fig. 4.2. The SU(4) Kondo effect is characterized by having a conductance of $4e^2/h$, higher than its SU(2) counterpart ($2e^2/h$), as well as being stronger with T_K reaching temperatures up to 14 K [14, 17]. In carbon nanotubes, SU(4) symmetry can only exist in the odd electronic states: every $4N$ state is a full shell and has no levels for zero-bias cotunneling, and every $4N + 2$ state forms a singlet or triplet state on the dot, preventing single electron coupling to the leads. Thus, the SU(4) Kondo effect is manifested as a zero-bias conductance peak in these diamonds, much like the SU(2) configuration.

Under perpendicular magnetic field $B_\perp$, the once degenerate spin states become split by the Zeeman splitting energy

$$E_Z = g\mu_B B_\perp \tag{4.3}$$

where $g = 2.0$ is the g-factor of the electron, and μ_B is the Bohr magneton. This splitting of the Kondo peaks has a slope of ± 0.116 meV/T. Under parallel magnetic field $B_\parallel$, the spin states split in the same way. However, while the isospin states are not magnetically coupled in perpendicular magnetic field, in parallel field these isospin states couple even more strongly that Zeeman splitting. This is a result of the
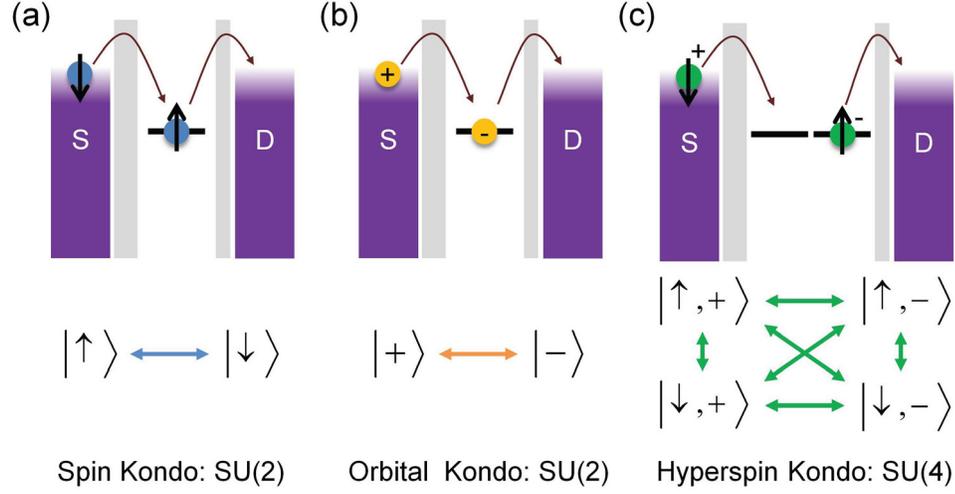
Figure 4.2: Schematic showing SU(2) vs. SU(4) Kondo effect. (a)-(b) Entanglement between two spin states or two orbital states causes SU(2) Kondo effect. (c) When spin and orbital entanglement occur simultaneously in a four-fold degenerate SWCNT, the much more robust SU(4) hyperspin configuration is achieved.

magnetic moments generated parallel to the carbon nanotube by the clockwise and counterclockwise motions of the isospin valleys. This results in an additional splitting of the isospin, with a greater slope [24]. The orbital Zeeman splitting is given by [17]

$$E_Z^{orb} = 2\mu_{orb}B_\parallel \tag{4.4}$$

where μ_{orb} is the orbital magnetic moment of electrons in the two isospin valleys.

In our measurement set-up, we have a perpendicular magnetic field which we can use to probe the Kondo effect, and measure the g-factor of the electrons in our system. Apart from these zero-bias peaks, other Kondo correlations can occur at finite bias. These inelastic cotunneling peaks, can be used as a tool to learn about the excited states of an ultra-short quantum dot with high precision.

4.1.3 Inelastic Kondo Resonances

When two states are non-degenerate, it is not energetically possible for cotunneling to occur, as shown in Fig. 4.3(a). Therefore, no zero-bias conductance peak is observed. However, by applying a bias voltage to the device, the energy needed to make up for the non-degeneracy is provided externally to the QD. This is known as inelastic cotunneling, and is observed as a conductance peak parallel to the gate axis at a finite bias equal to the difference in energy between these two levels. [16].

By measuring inelastic cotunneling it becomes possible to map out the energy states of a quantum dot for a given occupancy. It is only possible to observe Kondo enhanced inelastic cotunneling through an energy state which is smaller than the charging energy of the diamond (so that it appears in the blockade), and if the bias energy is larger than $k_B T_k$, so as not to disrupt the Kondo cloud. Here we will discuss some of the types of inelastic peaks we can observe in our ultra-short QDs, including the singlet-triplet Kondo effect [59], the energy level Kondo effect [16], and the vibron assisted Kondo effect [60], whose energy level diagrams are shown in Fig. 4.3(b)-(d). In these diagrams, the charging energy is larger than the excitation energies and therefore prevents the full occupancy of the SWCNT. The inelastic nature of these effects does not allow them to be characterized as SU(2) or SU(4) [16].

The singlet-triplet Kondo effect, shown in Fig. 4.3(b), is an enhanced cotunneling process that occurs at finite bias between different configurations of the doubly occupied QD [61]. The measured resonances from the singlet-triplet Kondo effect depend on the energy scales of the system, and only occur in states with $4N + 2$
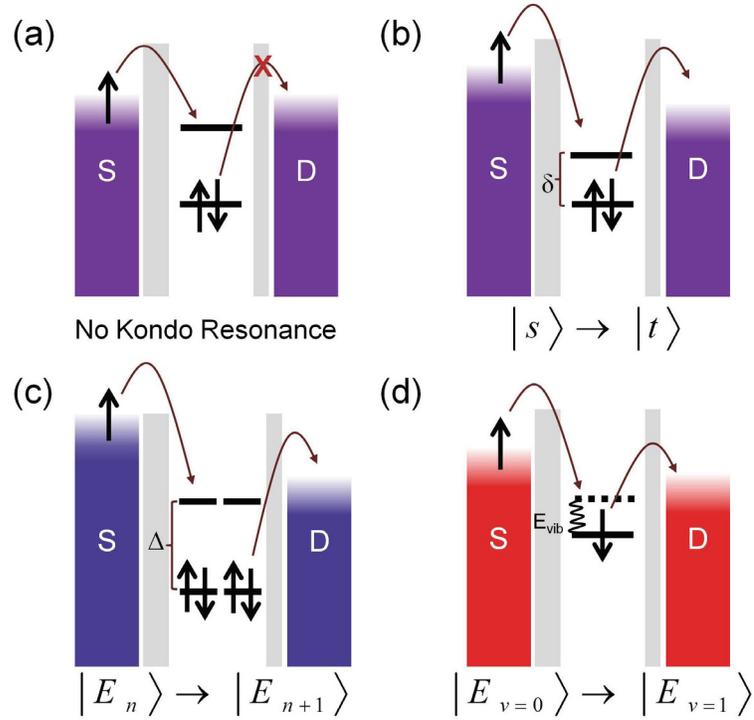
Figure 4.3: Inelastic Kondo effects. (a) When two levels are non degenerate, energy conservation prevents the Kondo effect. However, when V_B reaches the excitation energy, a finite-bias Kondo peak can be observed. (b) The singlet-triplet Kondo effect arises from a triplet excitation state in a doubly occupied QD at $eV_B = \approx \delta_{KK'}$. (c) The energy level Kondo effect can only arise if T_K is large. Inelastic cotunneling can occur through higher electronic excitations and a peak is observed at $eV_B = \Delta$. (d) Vibron assisted cotunneling can also be enhanced by the Kondo effect and is observed as finite bias resonances at multiples of the vibronic energy $eV_B = \Delta E_{vib}$.

charge carriers. The QD singlet and triplet states are shown in Fig. 4.4. A quantum

singlet is made up of two electrons of opposite spin sharing the same energy level. If

the orbital splitting $\delta_{KK'}$ (as defined in section 3.2.4), is smaller than the ferromagnetic

exchange energy J, (a manifestation of Hund's rule), the ground state is the triplet

state, with one electron on each orbital. Since there are unpaired electrons, this allows

zero-bias cotunneling. However, when $\delta_{KK'} > J$ (normally the case), the ground state

is the singlet state which is fully occupied and does not allow cotunneling. If a bias

equal to $\delta_{KK'}$ is applied, the energy difference between the levels is absorbed by the

dot and an inelastic Kondo resonance occurs. This is observed as a cotunneling peak

at $V_B \approx \delta_{KK'}$. Under magnetic field, this peak splits into three, one each for each of

the spins $S = -1, 0, 1$.
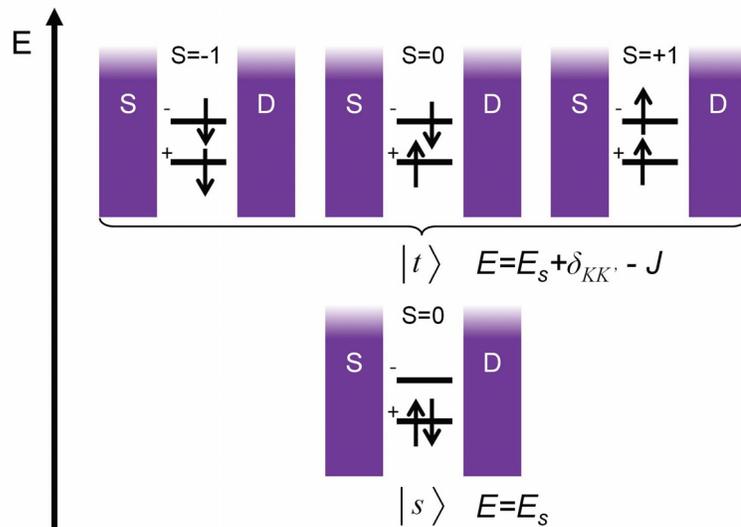
Figure 4.4: Singlet and triplet states of a doubly degenerate QD. The energies of the singlet $|s\rangle$ and triplet $|t\rangle$ states are given for a doubly occupied QD. Positive and negative signs indicate the two isospin orbitals separated by an energy $\delta_{KK'}$, J is the ferromagnetic exchange energy and E_s is the singlet state energy.

The energy level Kondo effect, shown in Fig. 4.3(c), occurs at the electron energy spacing [16]. This is only possible if the Kondo effect is very strong, such that high biases do not disrupt the Kondo cloud ($eV_B < k_B T_K$). The cotunneling peak from this Kondo effect is observable for every blockade diamond, because the subshell spacing is always the same for a given charge state. This state can never cause zero-bias cotunneling because the two energy levels are separated by the electron energy spacing. This results in an enhanced cotunneling peak at $V_B = \Delta$.

Although it has been less commonly observed [15, 60], certain molecular devices have shown Kondo peaks at vibronic energies, matching theoretical predictions [62]. These peaks are caused by a Kondo resonance through the excited electron + vibron(s) state. The details regarding this vibrational energy spectrum will be given in section 4.3, but the general concept is the same for all inelastic Kondo effects. This resonance occurs at $V_B = \Delta E_{vib}$, which depends on the vibronic mode and frequency. In our devices the Kondo effect allows us to observe the vibrational spectrum which would be otherwise unresolvable. We will explore the Kondo effect in our devices, and its interaction with vibronic modes in the following sections.

4.2 Strong Kondo Effect and Cotunneling Spectroscopy

We now turn to our own results to explore the Kondo effect in ultra-short SWCNT devices. One of the limitations of using ultra-short SWCNT devices is that the charging energies are very large. This means that high bias measurements must be made to fully measure the Coulomb blockade. The high-bias features are thermally broadened, $(eV \approx k_B T)$, making them difficult to resolve with high precision. This reduces the accuracy of measurements involving diamond height such as the electron energy spacing (Δ). We observe strong inelastic cotunneling peaks, which we use as a tool to image the electronic energy structure of our QDs at much lower bias, giving high precision measurements of the electronic structure of our ultra-short QDs.

In a short carbon nanotube with a small bandgap, we measure a very strong Kondo effect with Kondo temperatures up to $T_K = 28$ K. We observe that this Kondo temperature is gate-tuneable by nearly a factor of two. To our knowledge, this is the strongest Kondo effect observed in SWCNTs to date.

4.2.1 Inelastic Cotunneling Spectroscopy

In our devices, we observe zero-bias cotunneling peaks as well as several excited states. A differential conductance plot from one such sample, Device A, is shown in Fig. 4.5(a). Using SEM, we measure this device to have $L_{SEM} \approx 15$ nm, matching the length determined from this device's gate capacitance $L_G = 12 \pm 5$ nm.

We observe electron energy level Kondo peaks at $\Delta = 30 \pm 3$ meV across all charge states, matching Δ measured from the diamond heights. Because we cannot observe the charge degeneracy diamond, we determine its bandgap from AFM radius measurements. The expected bandgap, is ≈ 700 meV. This gives an electronic energy spacing of $\Delta^{harm} = 47 \pm 3$ meV. This discrepancy could arise from systematic error regarding the diameter measurement, or perhaps more likely, could occur because this carbon nanotube is of the small bandgap variety.

We first investigate the diamonds with odd hole occupation number to understand the zero-bias Kondo effect observed in this device. We confirm that we are observing the Kondo effect by measuring its Kondo temperature using two methods. We determine the Kondo temperature by measuring the HWHM of the Kondo peak along the dashed line in Fig. 4.5(a) at 0.3 K, shown in panel (b), giving $T_K \approx 4.5$. We then measure the heights of the Kondo peaks vs. temperature as shown in Fig. 4.5(c), and fit them to the Goldhaber-Gordon model, (eq. 4.2), with $s = 0.21$ for a spin-$\frac{1}{2}$ system. We confirm $T_K = 4.4 \pm 0.2$ K in this device.

We now explore the magnetic field dependence of this device. In Fig. 4.5(d), we show magnetic field splitting in differential conductance colour plots from 0 to 9 T. This data corresponds to the boxed area in panel (a) with $4N + 1$ and $4N + 3$ charge states labelled i and ii. We plot the positions of the Kondo peaks with respect to magnetic field for states i and ii in Figs. 4.5(e)-(f) respectively. From linear fits to the splitting, we calculate an average g-factor of $g_{av} = 1.92 \pm 0.08$, in good agreement with the expected value, $g = 2.0$.
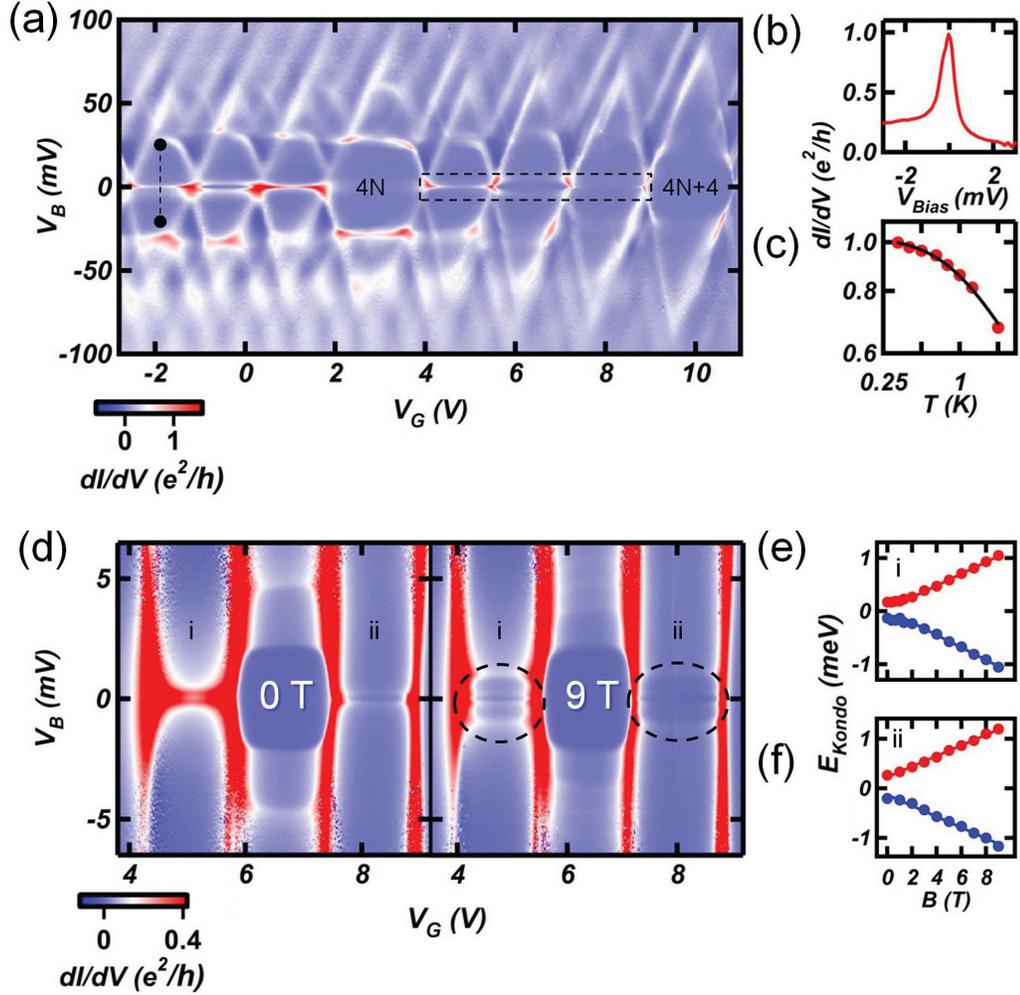
Figure 4.5: The Kondo effect in our devices. (a) Differential conductance plot for Device A showing the Kondo effect. Along the dashed line, we extract a differential bias plot, shown in (b). The HWHM of the Kondo peak gives $T_K \approx 4.5$ K. (c) Temperature dependence of the Kondo peak, fit using the Goldhaber-Gordon model, giving $T_K = 4.4 \pm 0.2$ K. (d) Magnetic field dependence of the Kondo effect at 0 and 9 T, from the boxed section in (a). (e)-(f) We observe a linear Zeeman splitting of the Kondo peaks, noted by i and ii, with respect to magnetic field, giving $g_{av} = 1.92 \pm 0.08$.
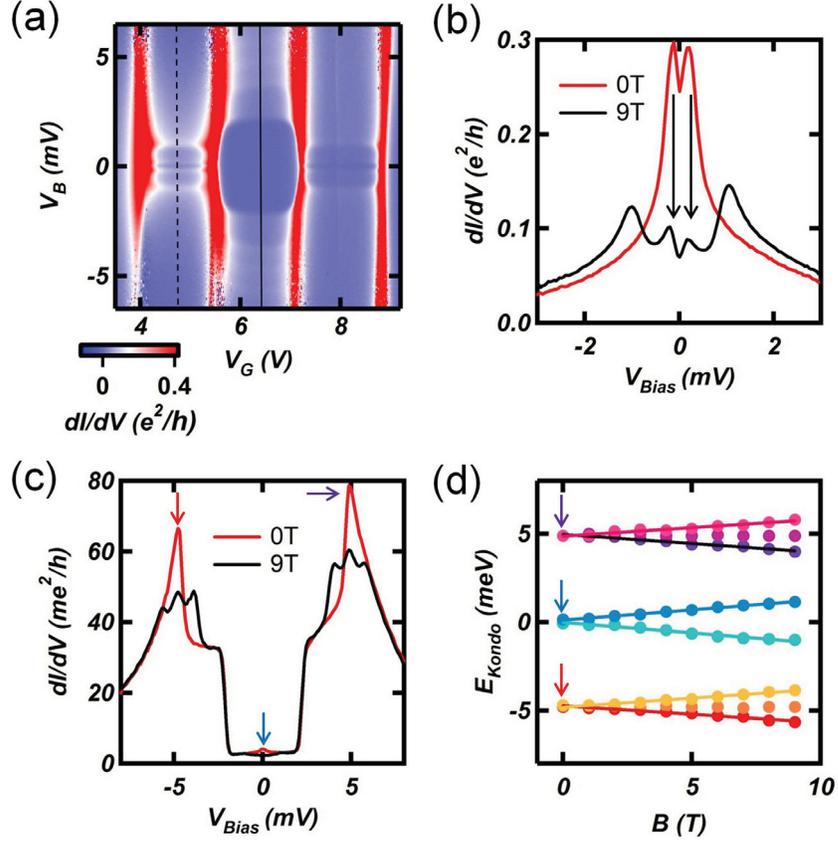
Figure 4.6: Zero field splitting and the singlet-triplet Kondo effect. (a) Device A differential conductance plot at 9T, showing dashed and solid lines corresponding to the locations of the 1D cuts in panels (b) and (c). (b) We observe a split peak which is independent of applied magnetic field, resulting from a magnetic impurity. (c) We observe a 3-way splitting of the triplet Kondo peak with magnetic field and a small zero-bias peak in this device. (d) This 3-way splitting has a linear dependence on magnetic field, giving $g_{av} \approx 1.72$. From the splitting of these peaks we measure $\delta_{KK'} \approx 4.9$ meV. The zero-bias peak, likely from the impurity, splits with $g \approx 2.04$.

Fig. 4.6(a) shows linecuts along the same with dashed and solid lines corresponding to panels (b) and (c) respectively. In panel (b), we show that the central peak is split by ≈ 0.33 meV at $B = 0$, and does not split with $B_\perp$, indicating that this state is isospin based. Using an electron g-factor $g = 2.0$, this is equivalent to a magnetic field of 1.4 T. A similar zero-field splitting of 0.51 T has been previously observed [63], and was attributed to the localized field of a magnetic impurity, most likely a ferromagnetic nanoparticle leftover from SWCNT growth. In our case we observe a zero-field splitting which is three times as large. This indicates that ultra-short carbon nanotubes could be a good platform for developing magnetic impurity sensors.

Now we wish to understand the magnetic dependence of the doubly occupied diamond with charge state $4N + 2$. We show differential conductance data along the solid line in 4.6(a) at 0 and 9 T in panel (c). The large peaks at ± 5 meV split into three as magnetic field is applied, corresponding to overall spins of ± 1 and 0 on the QD. This signifies that we are observing the triplet state, and can therefore measure its energy, approximately equivalent to the magnitude of intervalley scattering: $\delta_{KK'} \approx 4.9$ meV. We show the magnetic field dependence of the triplet state in panel (d), and measure $g_{av} \approx 1.72$ at positive and negative bias (red and purple arrows), which is less than the expected value. A reduction of g has been observed in the past, and was attributed to strong *e-e* interactions or spin orbit coupling [28]. A final feature to point out in this data is the barely visible zero bias peak (blue arrow) which splits with $g \approx 2.04$ in the same diamond. This is likely due to by Zeeman splitting of the electrons on the magnetic nanoparticle itself.

In the same device, we observe a coexistence of the Kondo effect and electromechanical coupling. In Fig. 4.7(a), we show a differential conductance colour plot corresponding to a few-hole state, with solid and dashed lines corresponding to the bias plots in panels (b) and (c). The data in these latter two panels was taken using AC measurement techniques. In panel (a) we observe excitation sidebands which run parallel to the Coulomb blockade (see sec. 4.3.2), some of which may correspond to vibronic excitations. However, it is difficult to determine the energy of the vibronic states in this device because no regularly spaced vibronic sidebands are visible. This is likely due to a somewhat weak electromechanical coupling in this device.
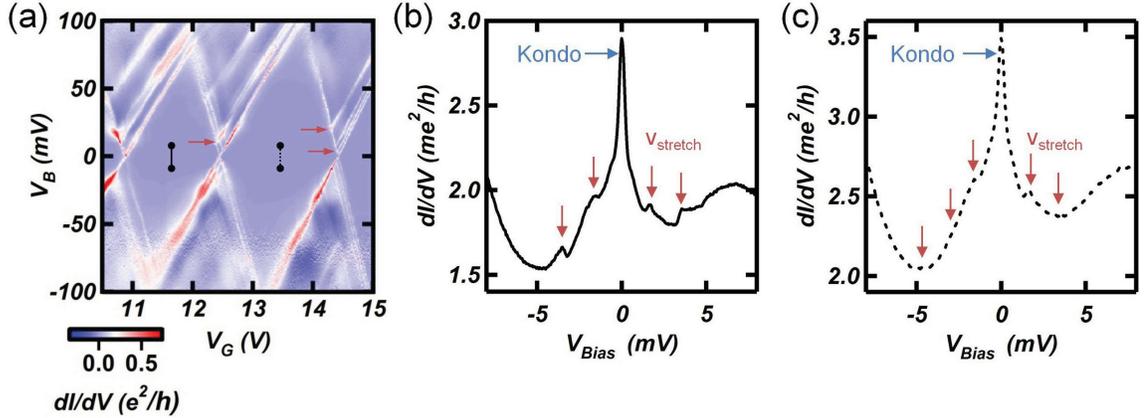
Figure 4.7: Stretching vibron assisted cotunneling. (a) Differential conductance colour plot, showing a partially unresolved sideband excitation spectrum (red arrows). Solid and dashed lines correspond to data in panels (b)-(c). Blue arrows point out weak zero-bias Kondo peaks for few-hole occupation. Red arrows point out the evenly spaced inelastic cotunneling peaks caused by vibronic excitations with energy $\Delta E_{stretch} = 1.65 \pm 0.08$ meV, matching the expected energy for this $L \approx 15$ nm device.

However, we are able to resolve the vibronic spectrum due to Kondo enhanced inelastic cotunneling through vibronic excitations. In panels (b) and (c), in addition to

the zero-bias Kondo peaks (blue arrows), we resolve the first few stretching vibrons (red arrow) with energy spacing $\Delta E_{stretch} = 1.65 \pm 0.08$ meV, corresponding to a vibronic length of $L_{vib} = 22\pm3$ nm. We find that their energies are not gate dependent, as expected for stretching vibrons. The excitation energies are in good agreement with the SEM length of this device $L_{SEM} = 15 \pm 5$ nm, and give a slightly larger length than that of the hole channel of the QD, $L_G^{h^+} = 11 \pm 2$ nm, as is expected for an n-doped nanotube. Because the vibron length is greater than that of the QD, it is no surprise that e-v coupling is weak, as will be explained in section 4.3.2.

We only observe these vibronic peaks for few-hole occupation of the QD (large V_G). This is because at few-hole occupation the tunnel barriers are larger and the zero-bias Kondo resonance which normally masks the faint vibronic peaks becomes weaker. Thus, we find that due to the Kondo resonance, we can measure stretching vibron assisted cotunneling in diamonds where these cannot be resolved in the sequential tunneling sidebands, allowing us to probe the vibronic spectrum and better understand the physics of our devices.

Inelastic cotunneling Kondo resonances allow us to measure the electronic spectrum of ultra-short SWCNTs at low bias, as long as the device is below the Kondo temperature. In other devices, we push the limits of the strength of the Kondo effect in SWCNTs, allowing us to observe these effects at higher temperature.

4.2.2 High Kondo Temperatures

In very short and high conductance devices, Γ and E_c are both large. According to equation 4.1, these factors increase the Kondo temperature and can allow us to make ultra-short devices with stronger electron-electron correlations. Our highest conductance sample is Device E, showing Fabry-Pérot oscillations on the electron side (see section 3.3.3). Hole transport through this ≈ 100 nm long device is in the blockade regime and shows a very strong Kondo effect, as shown in Fig. 4.8(a).
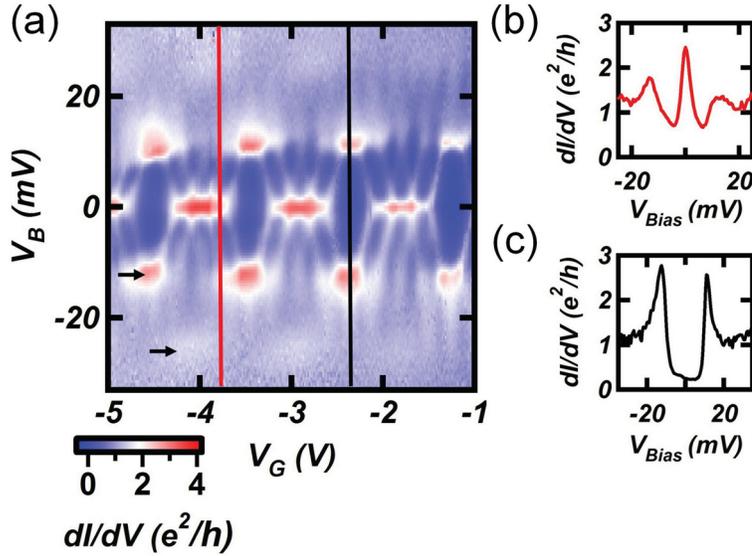
Figure 4.8: High Kondo Temperatures. (a) Differential conductance plot from Device E, showing a strong Kondo effect. We observe inelastic Kondo peaks, corresponding to Δ and 2Δ, pointed out by the black arrows. The red and black lines correspond to the bias cuts in (b) and (c) respectively. The HWHM of the Kondo peak gives $T_K \approx 28$ K. We observe that the inelastic peaks are as strong as the zero-bias peak.

From the HWHM of the zero-bias peaks, we measure T_K ranging from 17 K to 28 K with increasing $|V_G|$. A strong and tuneable Kondo effect with different physical properties has been observed in a 2D sheet of graphene ($T_K = 30 - 90$ K) [64], and

the Kondo temperature has been found to be tuneable by a factor of two in other QD systems ($T_K < 1$ K) [65]. However, to our knowledge, no other system has observed 0D transport with this high Kondo temperature and large degree of tuneability. This high degree of control in Device E arises from the high electron tunneling rate strong tuneability of tunneling barriers as a result of its small bandgap. Along the red line in Fig. 4.8(a) we plot the differential conductance in panel (b). The HWHM of this peak gives $T_K \approx 28$ K. To our knowledge, this is the highest Kondo temperature observed in SWCNT QDs to date.

We observe a strong electron energy Kondo effect at $|V_B| = 13.5 \pm 2$ meV across the entire blockade pattern, and a slightly weaker one at approximately twice this energy $|V_B| = 26 \pm 3$ meV, as pointed out by the black arrows in panel (a). These energies agree with electron spacing measured from diamond heights $\Delta = 11 \pm 2$ meV. This corresponds to a Dirac particle-in-a-box potential of length $L_\Delta \approx 124 \pm 16$ nm, in reasonable agreement with SEM and gate capacitance measurements which give $L = 103 \pm 9$ nm. The strength of these peaks is measured along the black line in Fig. 4.8(a) and shown to be as strong as the zero-bias Kondo peak in panel (c). The second electronic excitation at 2Δ has not been observed before in SWCNT QDs, and shows that the Kondo effect persists to very high energies, confirming its strength. Besides high conductance and short length, the Kondo effect can be further strengthened through electron-vibron coupling [66], allowing us to reach even higher Kondo temperatures. Here we shift focus from strong e-e interactions to e-v coupling in SWCNTs.

4.3 Vibrons in SWCNTs: Background

Quantized mechanical excitations in a finite crystal lattice are known as vibrons. When a tunneling electron interacts with the deformation potential of the lattice structure, vibrons can be excited in the SWCNT. There are several different vibrational degrees of freedom in SWCNTs, as outlined in Fig. 4.9. Each oscillation mode has a different energy (ΔE_{vib}) and e-v coupling (λ_{vib}), which depends on the physical and electronic structure of the nanotube. Electron-vibron coupling describes the strength of the interaction between charge carriers and the mechanical degrees of freedom of the nanotube, and is related to the probability that a tunneling event will excite a vibron [67]. Here we will briefly review the theory of e-v interactions following ref. [68], focusing on the stretching and bending modes. The breathing and twisting modes tend to be very weakly coupled to the longitudinal injection of electrons in SWCNT devices, and are thus not observed in our devices.

4.3.1 System Hamiltonian

Fist we give the Hamiltonian of the system, which comes from the Anderson-Holstein model, which depends on contributions from the electrons on the QD, vibrons, the leads, and electron tunneling. $H = H_{el} + H_{vib} + H_{leads} + H_T$. Here, we only expand the terms which contribute to e-v coupling and the energy of the QD [68]

$$H = \sum_{s=\uparrow,\downarrow} \sum_{v,q} \left[\epsilon + \lambda_q^{(v)} \hbar \omega_q^{(v)} (b_q^{(v)\,\dagger} + b_q^{(v)})^j \right] n_s + U n_\uparrow n_\downarrow + \sum_{v,q} \hbar \omega_q^{(v)} b_q^{(v)\,\dagger} b_q^{(v)} + H_{leads} + H_T$$

$$(4.5)$$

98
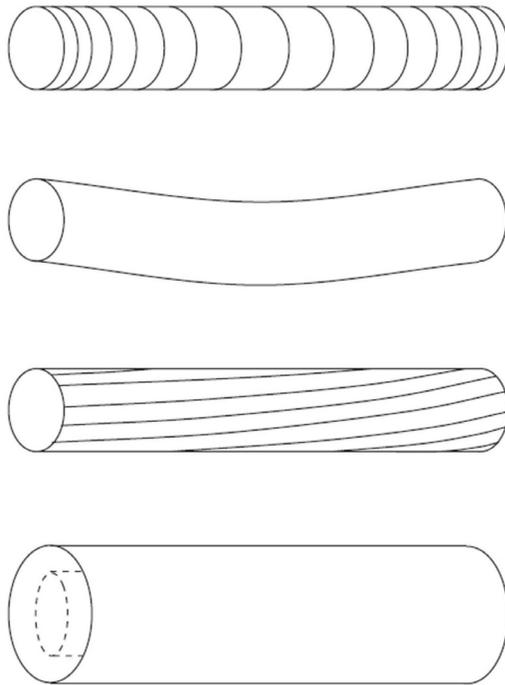
Figure 4.9: Vibrational degrees of freedom in a carbon nanotube. A SWCNT can vibrate in a variety of manners including (top to bottom) stretching, bending, twisting and breathing. Adapted from [68].

where $\uparrow$ and $\downarrow$ denote electron spin, v and q represent the type of vibron and its wavenumber, ϵ is the electronic energy, λ_q is e-v coupling, ω_q is vibron frequency, n_s is the spin-number operator, $j = 1, 2$ describes weather the e-v interaction is linear or quadratic (depending on the symmetry of the vibronic mode) and $b_q^\dagger$ and b_q are the vibron creation and annihilation operators. For all but the bending vibrational mode the coupling is linear and $j = 1$, while $j = 2$ for the quadratically coupled bending mode. This Hamiltonian is the starting point to understand e-v interactions in the QD system. We first discuss the stretching vibrations of the SWCNT which couple most strongly to longitudinal electron tunneling.

4.3.2 Stretching Vibrons

The longitudinal stretching vibrons of a carbon nanotube have been measured in Coulomb blockade data by several groups, and match well with theoretical predictions describing their relative strengths and energies [69]. A tunneling electron can excite a quantum of vibrational energy as long as e-v coupling is strong. These excited states provide additional tunneling pathways for the electron, as shown in Fig. 4.10(a). Therefore, vibrational excitations appear as sidebands to the Coulomb blockade, which onset when the Fermi level of the lead lines up with the excitation energy. A data plot showing these vibronic sidebands from [69] is shown in Fig. 4.10(b). In a differential conductance plot, these sidebands have energy spacings equal to the vibronic energy $\Delta E_{stretch}$.
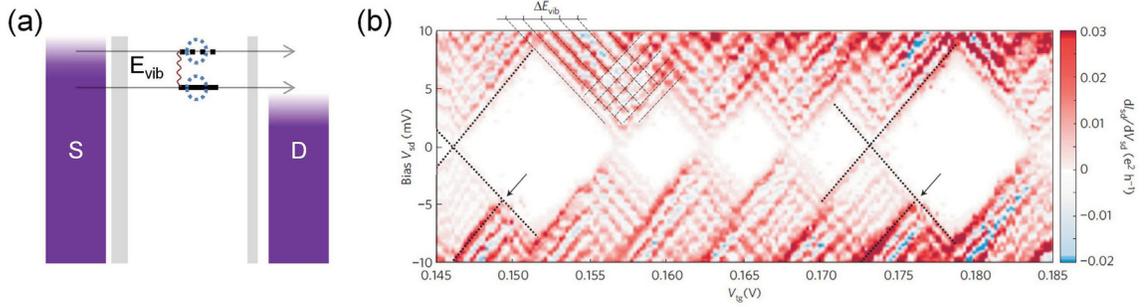
Figure 4.10: Stretching vibron sidebands. (a) Energy level diagram, showing how a vibronic excitation causes an increase in conductance. (b) Adapted from [69]. In this 65 nm device, many sidebands are observed from vibronic excitations. Their spacing is given by $\Delta E_{vib} = 0.8 \pm 0.2$ meV. Electronic excited states ($\Delta = 6.8 \pm 2.5$ meV) are identified by the black arrows. Low bias vibron peaks are weak due to Franck-Condon giving a strong e-v coupling $\lambda_{stretch} \approx 1.8$.

The energy of a stretching vibron is given by [70]

$$E_{stretch} = \frac{h v_{ph}}{2L_v} \qquad v_{ph} = \sqrt{\frac{Y}{\rho}} \tag{4.6}$$

where h is Planck's constant, v_{ph} is the speed of sound in a carbon nanotube, L_v is its length, $Y \approx 1$ TPa is the bulk modulus of CNTs, and $\rho = 3080/d$ kg/m^3 is the carbon nanotube density which depends on d, its diameter in nm. The stretching mode coupling depends on the relative sizes and positions of the dot and vibron. Depending on the mechanical pinning of the CNT, these do not always coincide. For long vibrons which occupy the entire QD ($L_d = L_{vib}$), the electron vibron coupling is only non-zero when the vibronic wavenumber integer (m) is equal to twice the

101

electronic excitation number of the electron n' and is given by [68]

$$\lambda_{stretch} \approx -\delta_{m,2n'}\frac{1.5}{\sqrt{m\pi d}} \tag{4.7}$$

If the quantum dot is much smaller than the length of the vibron $L_d \ll L_{vib}$, the e-v coupling is more dependent on the position of the QD in relation to the vibron. If the dot is localized in a region of intense vibron-induced strain, the coupling reaches a maximum value

$$\lambda_{stretch}^{max} \approx \frac{3.0}{\sqrt{m\pi d}} \tag{4.8}$$

The m and d dependence of these equations ensures that only the only appreciable coupling occurs in long-wavelength modes in small diameter nanotubes.

The relative intensity of the vibron peaks is described by the Franck-Condon model [69]. For strong e-v coupling, this model predicts the suppression of low bias vibronic sidebands, due to a small overlap between the wavefunctions of the electron before and after the excitation of a vibron [69]. The Franck-Condon model allows us to fit the measured differential conductance peaks by

$$\left(\frac{dI}{dV}\right)_m^{max} \propto \frac{e^{-g}g^m}{m!} \tag{4.9}$$

where m is the vibron wavenumber integer, and the Franck-Condon factor (g), is related to the electron-vibron coupling parameter by $g = \lambda^2$. For $g \geq 1$ (strong coupling), the peaks for small m are exponentially suppressed, and only the large m

102

vibronic peaks can be resolved. The e-v coupling can be measured experimentally by fitting the intensity of the vibron peaks with respect to vibron number. Although understanding the nature of stretching vibrons is of fundamental importance, the bending oscillations of a nanotube are more useful for NEMS applications.

4.3.3 Bending Vibrons

Unlike stretching vibrons, bending vibrons are not expected to couple strongly to the longitudinal injection of electrons in to the nanotube. The maximal coupling for the bending mode is given by [68]

$$\lambda_{bend}^{max} \approx \frac{40 L_{vib}}{m^2 \left(\pi d\right)^3} \tag{4.10}$$

for $n_b = 1$, where n_b is the integer harmonic of the bending mode. This relationship defines a strong bending mode coupling only for long wavelength vibrons in devices where the nanotubes are very long, and have small diameter. Experimentally, this quantity can be measured by [71]

$$\lambda_{bend}^{exp} = \frac{F^2}{\hbar m_{tube}^* \left(2\pi f_{bend}\right)^3} \tag{4.11}$$

where F is the electrostatic force on the CNT, $m_{tube}^* = 0.725 \rho L_{tube} \pi r_{tube}^2$ is the effective mass of the oscillator, where ρ , L_{tube}, and r_{tube} are the density, length, and radius, of the nanotube respectively, and f_{bend} is the bending frequency. The electrostatic force

on the SWCNT is given by the QD potential [72]

$$F = \frac{1}{2}C'_G \left(V_G + \frac{C_G V_G - Ne}{C_\Sigma} \right)^2 \tag{4.12}$$

where $C'_G = \partial C_G / \partial t_{vac}$ is the differential gate capacitance with respect to the oscillatory deflection of the nanotube, C_G and V_G are the gate capacitance and gate voltage, and C_Σ is the sum capacitance of the system. For the higher harmonics, $n_b \geq 2$, the coupling is even weaker and decreases with increasing n_b [68]

$$\lambda_{bend}^{max} \approx 10^{-4} \frac{n_b^2 + 1}{n_b^4} \frac{\pi d}{L_{vib}} \tag{4.13}$$
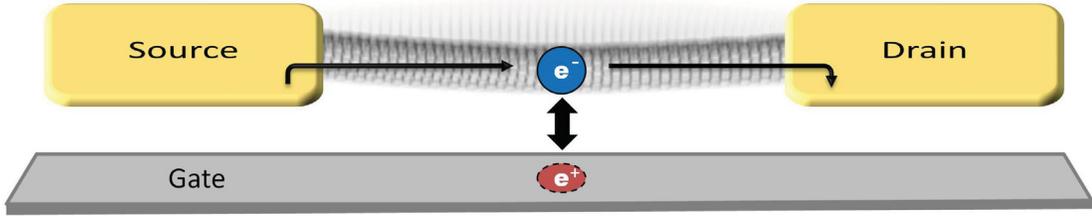
Figure 4.11: Positive feedback mechanism for bending vibrons. For sharp Coulomb diamonds, a deflection of the nanotube can alter the gate capacitance. The causes the electron state to cross the Coulomb threshold, increasing the likelihood of tunneling in phase with the oscillatory motion. The electron on the QD is electrostatically attracted to the gate electrode driving the oscillations through a positive feedback mechanism. This makes it possible to observe the bending mode with DC transport.

Thus, we do not expect to see any sidebands from the bending mode in our data; the coupling is simply too weak. For comparison, strong $e\text{-}v$ coupling to the stretching mode is given by $\lambda_{stretch} > 1$, while typical values for the bending mode are $\lambda_{stretch} \sim 10^{-4}$. However, it does become possible to observe collective resonant oscillations of the nanotube through a positive feedback mechanism [73], which is depicted schematically in Fig. 4.11. An image charge on the gate electrode causes an attractive interaction between the nanotube and the substrate, which is only present when the QD is occupied. This small shift in position modulates gate capacitance. As long as the blockade diamonds are very sharp, this change in capacitance can cause the QD state to cross the Coulomb threshold, increasing the likelihood of tunneling at the nadir of the nanotube's motion. This electromechanical driving further actuates the device. The electrons enhance the bending through an electrostatic attraction in phase with the bending of the nanotube. This forms the positive feedback necessary to observe the bending mode [11]. This feedback causes a deformation of the Coulomb diamond such that we can observe this bending mode resonance deep inside the blockade where current is normally disallowed.

To excite this resonant positive feedback, several conditions are necessary. First, the source-drain bias must be large enough to excite oscillations with energy $\Delta E_{bend} = h f_{bend}$, which for devices under minimal longitudinal strain is given by [74]

$$f_{bend,T=0} = 11.19 \frac{d}{(4\pi L_{vib})^2} \sqrt{\frac{Y}{\rho}} \qquad (4.14)$$

In our 100s of GHz devices we expect $E_{bend} < 1$ meV. Second, the tunneling rate through the device must be at least equal to the bending mode frequency, such that each bending vibron coincides with the tunneling of at least one electron, $\Gamma \geq f_{bend}$. This allows the continuous driving of the oscillatory motion. Third, the Coulomb thresholds must be very sharp so that a small change in C_G results in a large change in conductance.

The final condition necessary for self-actuation of the bending mode is a high quality factor (Q). In large Q devices, the vibrons are long lived, allowing many in-phase vibrons populate the QD, increasing the amplitude of the oscillator [72]. In turn, this increases the likelihood that a tunneling event will coincide with the excitation of a vibron. This gives an effective e-v coupling $\lambda^*_{bend} = \lambda_{bend}\sqrt{Q}$, coming from the probability of electron coupling ($g = \lambda^2$), which depends on the number of vibrons occupying the dot: $g^* = Qg$.

Assuming harmonic motion, the Q-factor in our devices is given by [72]

$$\frac{1}{Q} = \frac{1}{2\pi f_{bend}m^*_{tube}} \left(\frac{2C'_G V_g}{\Gamma C_\Sigma}\right)^2 G \qquad (4.15)$$

where C'_G is again the derivative of the gate capacitance in the direction normal to the substrate, and V_G and G are the gate voltage and conductance at which the resonance occurs respectively. By substituting variables, we find that $Q \propto 1/L^2_{vib}$. Because Q and E_c are both inversely related to length, we only expect to observe this positive feedback in ultra-short devices. This explains why the bending mode was not

until recently observed through DC transport in carbon nanotubes [11]. A schematic diagram showing differential conductance plot including the bending mode resonance and stretching vibrons is shown in Fig. 4.12(a).
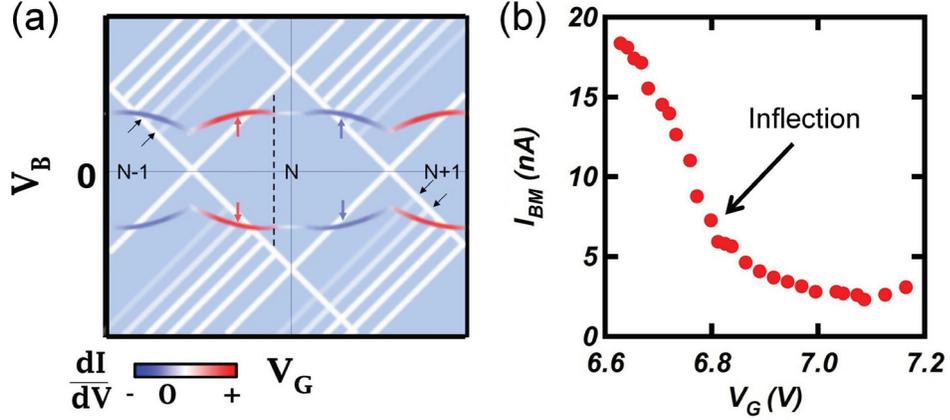
Figure 4.12: Bending frequency measurement from Coulomb blockade. (a) Schematic diagram showing vibronic effects in a differential conductance plot. Black, red and blue arrows point out stretching vibrons, in-phase bending resonance out-of-phase bending resonance respectively. The dashed line shows the ideal position to measure bending frequency, where external contributions are minimal. (b) This ideal point of measurement corresponds to the inflection point, in a $I_{BM} - V_G$ plot from Device B.

Depending on the relative phases of the bending oscillations and electron tunneling, the bending mode can enhance or reduce the flow of current by forward or back scattering of electrons [75]. In our devices the relative phase of bending motion to electron tunneling is determined by the Q-factor, bending frequency, and capacitances of the system [72, 75]. To fully understand this phase behaviour requires complicated modeling and is beyond the scope of this thesis.

Because we observe both current enhancement from the Coulomb peaks and out-of-phase damping further along the bending mode resonance, it is important

to measure the current where these effects are minimal to determine the bending frequency. This occurs at approximately the inflection point of a I_V_G plot along the bending resonance as shown in Fig. 4.12(b). This corresponds to a point approximately one quarter of the diamond width at resonance, along the dotted line shown in panel (a). Here we are deep enough inside the valley to reduce current from sequential tunneling, but to avoid the further reduction from out of phase driving. In this region the only mechanism allowing the tunneling of electrons is the e-v coupling to the bending mode, with approximately one electron tunneled per oscillation [11, 73] Thus, we can measure the bending frequency indirectly though current measurements in DC transport.

$$\Gamma = \frac{I_{BM}}{e} \approx f_{bend} \qquad (4.16)$$

where I_{BM} is the current along the bending resonance deep inside the blockade diamonds. Depending on the magnitude of the bending frequency compared to the background, the error in determining this ideal measurement position can be sizeable, but generally gives less than 30 % error in the frequency measurement.

We can tune the frequency of the bending mode resonance by applying strain to the nanotube. The strain induced frequency is given by [74]

$$f_{bend,T \neq 0} = \frac{1}{2\pi} \left[\frac{3.54}{dL_{vib}} \sqrt{\frac{T}{\rho}} + \frac{\pi d}{2L_{vib}^2} \sqrt{\frac{Y}{\rho}} \right] \qquad (4.17)$$

where T is the tension in the carbon nanotube. Tension in the carbon nanotube can be, built into the geometry of the contacts during suspension, or can be applied

108
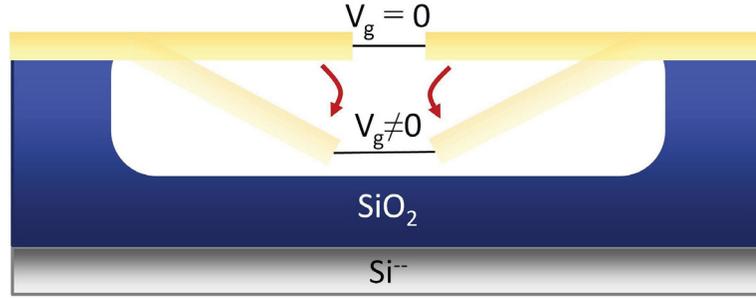
Figure 4.13: Electrostatic strain tuning of a SWCNT NEMS device. A capacitive attraction between the back-gate electrode and the suspended contacts causes them to bend towards the substrate. The suspended contacts act as lever arms which can apply large strains, up to $\approx 10\%$ [11].

electrostatically, by applying a gate voltage which causes the suspended contacts to act as lever arms. A diagram illustrating the nature of electrostatic strain tuning in our SWCNT NEMS is shown in Fig. 4.13. Thus, we can electrostatically alter the tension in the carbon nanotube, and tune its bending frequency.

4.4 Bending Mode Resonances and SWCNT NEMS Devices

We clearly resolve vibronic effects in our devices using DC transport spectroscopy. This is in part due to the fact that our nanotubes are suspended (no damping by the substrate), and in part due to the cleanliness of our devices from the electromigration process. Because our devices are so short, the vibronic energies become large and therefore easy to resolve. Here we describe the vibronic signatures we observe, starting with stretching mode sidebands, and then self-actuated bending mode resonances.

4.4.1 Stretching Vibrons in Ultra-short SWCNTs

We observe stretching mode oscillations with many sidebands in our devices. Differential conductance data from two such devices, are given in Fig. 4.14. The stretching vibrons can be resolved as sidebands, pointed out by the black arrows.
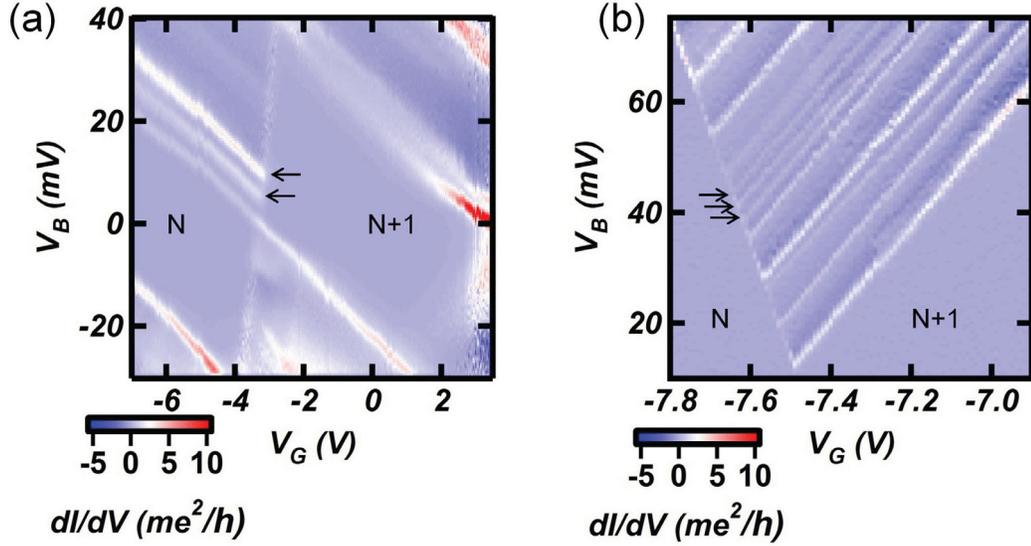
Figure 4.14: Stretching vibrons in our devices. (a)-(b) We measure vibronic energies from two devices as the separation between vibronic sidebands, highlighted by the black arrows. These energies, 4.9 and 2.8 meV respectively, correspond to devices of length 10.8 and 16.3 nm respectively, in good agreement with SEM measurements.

The energies of the vibrons can be read directly from the differential conductance plots at the points where the sidebands cross the blockade thresholds, highlighted by the black arrows. We measure vibronic energies of 4.9 and 2.8 meV respectively, corresponding to stretching frequencies of 1.2 and 0.7 THz. Using equation 4.6, this corresponds to lengths $L_{vib} = 10.8$ and 16.3 nm respectively. Although the nanotubes were not visible in SEM images of these devices due to their short length, we estimate

the lengths of the devices by measuring the gap between the electrodes in the vicinity of the nanotube. We find reasonable agreement with lengths of $L_{SEM} \approx 11.7$ and 18.8 nm respectively. Thus, by measuring the stretching vibron spectrum, we can confirm the length of our ultra-short devices.

We observe the suppression of low bias stretching vibrons in panel (b). This is due to Franck-Condon blockade, resulting of the small overlap of the electronic wavefunctions before and after the emission of a vibron [69]. Because we observe many sidebands of similar amplitudes at high bias, and exponentially suppressed sidebands at low bias in our devices, we deduce from the Franck-Condon model that our devices have a strong e-v coupling with $\lambda_{stretch} \gtrsim 1$. While stretching vibrons are of fundamental interest, bending vibrons are more useful in terms of NEMS applications and will be discussed in detail in the following sections.

4.4.2 Proof of Bending Vibrons in Ultra-short SWCNTs

We will now outline the previous work that has been done on bending vibron resonances in ultra-short carbon nanotubes in our group, showing how we observe the bending mode through a positive feedback mechanism. These results have been published in ref. [11].

We observe a positive feedback mediated bending mode resonance in several of our ultra-short devices. A differential conductance colour plot from one such device with $L \approx 22$ nm shown in Fig. 4.15. As explained in section 4.3.3, we can measure current deep in the blockade along the bending mode resonance peak, giving us

an approximate measure of the bending resonance frequency. In the differential conductance plot in Fig. 4.15(a), we again observe stretching vibronic sidebands running parallel to the blockade lines. However a sharp, non-parallel feature exists around $V_B = 40$ meV, corresponding to the bending mode resonance. We highlight this peak using a dashed line. Current and differential conductance plots along the red and black solid lines are shown in panels (b) and (c) respectively. Panel (c) shows that the bending resonances are much sharper and stronger than the stretching vibron peaks. This is because of the strong positive feedback nature of the bending resonance.
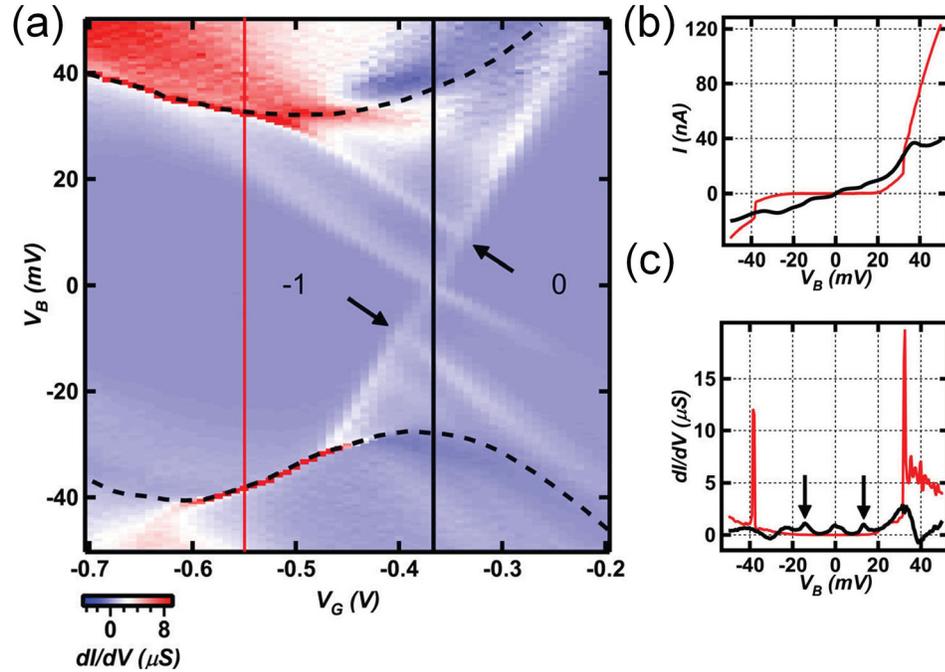
Figure 4.15: Self-actuated bending frequency from DC measurements. (a) Differential conductance plot, showing stretching mode sidebands (black arrows) and bending mode resonance (dashed line). (b) 1D cuts along the red and black lines showing current vs. bias. (c) Corresponding dI/dV plots, highlighting the sharpness of the bending resonance. Extracting frequency from current deep inside the blockade we measure $f_{bend} \approx 75$ GHz, matching the theoretical frequency for a $L \approx 22$ nm nanotube $f_{T=0} \approx 63$ GHz. Adapted from [11].

We measure current deep in the blockade to be $I = 12 \pm 5$ nA at the resonance peak, corresponding to a bending frequency of $f_{bend} \approx \Gamma = 75 \pm 30$ GHz. This is in good agreement with the frequency calculated from equation 4.14, $f_{T=0} \approx 63$ GHz. We observe this agreement in several devices in the 20-30 nm range [11].

We observe that by annealing a given device, the bending resonance becomes sharper, and can even emerge in a sample where bending mode resonances did not at first appear. This is because current annealing increases the quality factor, thereby increasing the effective e-v coupling λ^*_{bend}. In the above device, we calculate the Q-factor to be $Q \sim 10^6$ using equation 4.15, one of the highest Q factors measured in CNTs to date [72]. This enhances the positive feedback and makes DC measurement of this self-actuated system possible.

We find that the frequency of oscillation is tuneable by applying tension to the nanotube. We can actuate this tension using the electrostatic attraction between the suspended contacts and the gate, stretching the carbon nanotube. We model this attraction and determine that tension in the nanotube goes like $T \propto V_G^2$ [11]. The data where this is best exemplified is shown in Fig. 4.16(a). We extract the current along the bending mode (black points) and find that in the Coulomb valleys we have a good agreement with theory for both the high and low tension regimes from eq. 4.14 and 4.17 [11]. From the current at low and high tension we observe that we can tune the bending frequency by a factor of two from ≈ 75 to ≈ 150 GHz.

We have been able to achieve ultra high-frequency devices in a device with built-in mechanical strain. In one of our devices, the contacts became partially collapsed
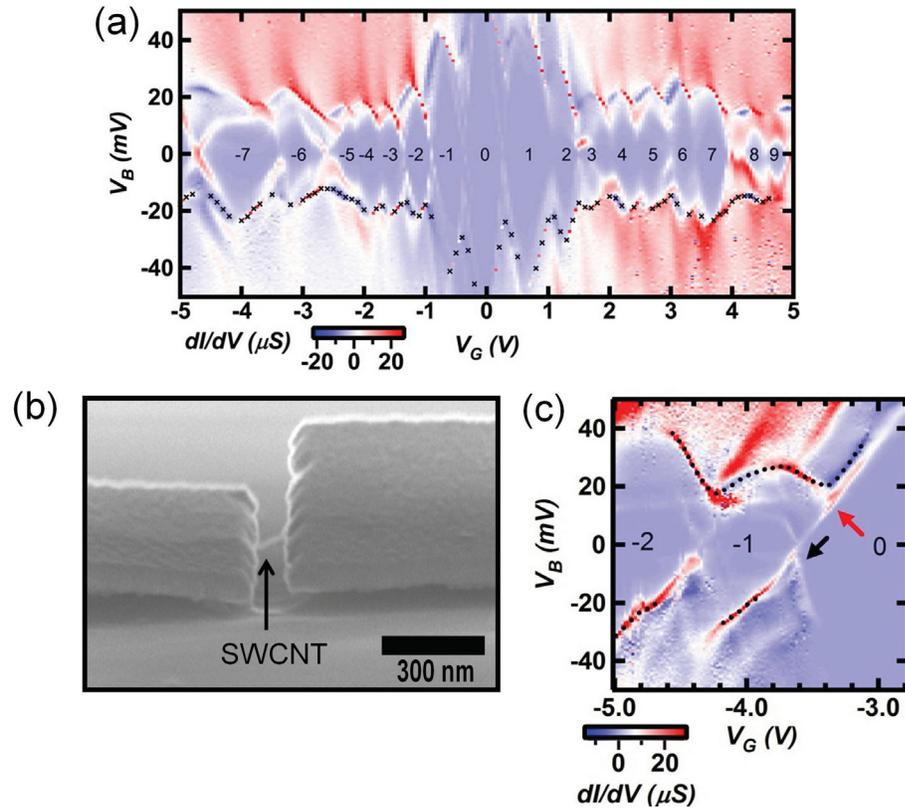
Figure 4.16: Strain tuneable bending mode oscillations. (a) Differential conductance data from an electrostatically strained device. By applying a gate voltage, we can tune the bending resonance frequency from ≈ 80 to ≈ 150 GHz in this sample. (b) SEM image of a mechanically strained junction, with built-in strain from partially suspended contacts. (c) Corresponding differential conductance plot showing a stretching sideband (black arrow) and bending mode resonance (red arrow/dashed line). In this device we measure a bending frequency ≈ 280 GHz. Adapted from [11].

during suspension, applying a large degree of strain as shown in Fig. 4.16(b), with the corresponding differential conductance plot shown in panel (c). By calculating the geometric strain from the positions of the contacts to be $\approx 11\%$, we determine the expected frequency to be $f_{T \neq 0} = 176 \pm 49$ GHz. This matches the frequency extracted from the Coulomb valleys of this device, giving $f_{bend} = 281 \pm 63$ GHz. By carefully engineering very short devices under large strains, we may be able to attain tuneable THz NEMS for ultrasensitive mass/force sensor applications with better than proton mass resolution [51]. This section has summarized the data published in ref. [11], showing that we observe a strong bending resonance in our ultra-short devices. Now we move on to discuss the new, unpublished results we measure from bending mode oscillations.

4.4.3 Bending Vibrons: New Results

We now investigate a sample with both strong e-e interactions, manifested as the Kondo effect, and strong e-v interactions resulting in the bending mode resonance. We show data differential conductance data from Device B, corresponding to electron occupations $N = 2-5$ in Fig. 4.17(a), with panels (b)-(c) showing 1D cuts in current and differential conductance along the red and black vertical lines. The odd and even charge states have very different diamond heights in this device. This will be explored in section 4.5. We highlight the Kondo peaks (which will be analyzed in section 4.5.2), with white arrows, and the bending resonances with black arrows. From the widths of the diamonds we determine the length of the quantum dot to

be $L_{vib} = 16 \pm 2$ nm. Using eq. 4.14, this corresponds to a zero tension resonance frequency of $f_{T=0} = 51 \pm 9$ GHz. The corresponding energy of the bending vibrons is $\Delta E_{bend} \approx 0.2$ meV. We might expect to observe Kondo enhanced vibronic inelastic cotunneling peaks in this device, but this energy is too small to be observed, as it is masked by the large zero-bias Kondo effect.
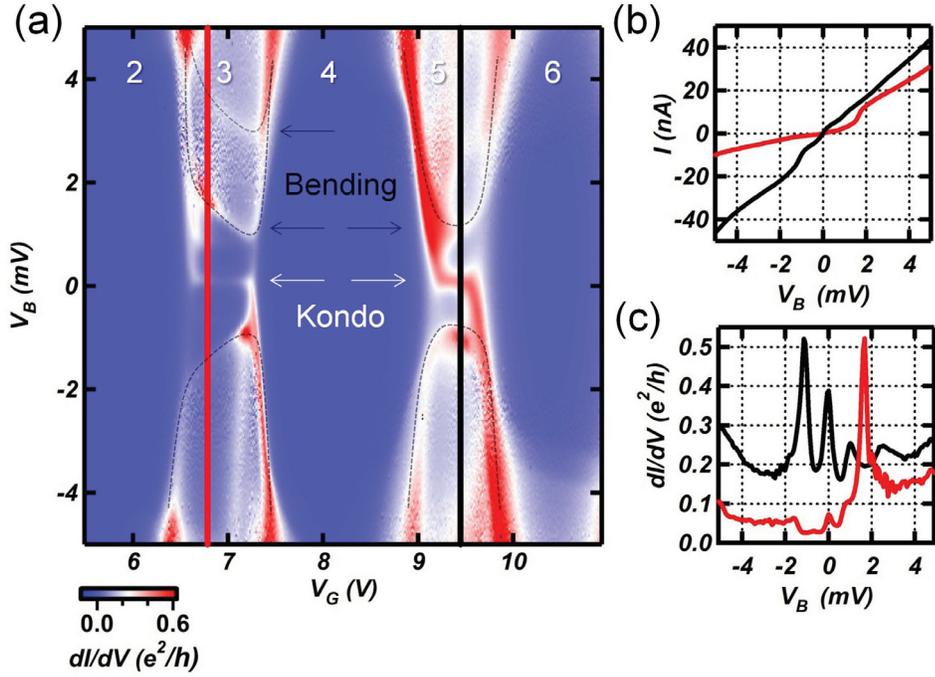
Figure 4.17: Bending vibrons and the Kondo effect. (a) Differential conductance plot from Device B. The measured bending frequency $f_{bend} \approx 45$ is in good agreement with the expected frequency $f_{T=0} \approx 50$ GHz in this ≈ 15 nm device. We observe the second bending harmonic (upper black arrow) with $f_2 \approx 84$ GHz. (b) 1D cuts along the black and red lines, show current vs. bias voltage. (c) Corresponding differential conductance plots illustrate the strength of the Kondo peak at zero bias in relation to the adjacent bending mode peaks at $V_B = 1 - 2$ meV.

We extract the bending frequency deep inside the Coulomb diamond along the red line in Fig. 4.17(a). We note that the onset of the bending mode occurs at

a much lower bias than has been observed in other devices. This is because the Kondo effect enhances the conductance at low bias, allowing the QD to reach the current needed to drive the bending mode at a much lower energy. This emphasizes that the positive feedback truly depends on tunneling rate, and not energy. We extract current of the resonance to be $I = 7 \pm 2$ nA, corresponding to a bending frequency of $f_{bend} = 45 \pm 13$ GHz, in good agreement with theory. We observe an additional, similar looking second peak at 3.1 ± 0.1 meV. This peak occurs at a current of ≈ 13.5 nA, corresponding to a bending mode frequency of 84 ± 10 GHz. Because this is approximately double the frequency of the first bending mode, we conclude that this is the second harmonic. This is the first time that the second harmonic has been observed by DC measurement. Using equations 4.15 and 4.11, we determine the Q-factor in this device to be $Q = 3 \pm 2 \times 10^4$, and e-v coupling $\lambda_{bend} = 5 \pm 2 \times 10^{-2}$. This gives an effective coupling $\lambda_{bend}^* = \lambda_{bend}\sqrt{Q} = 9 \pm 3$. We find that the effective coupling of the second harmonic is somewhat lower $\lambda_{bend}^* = 4 \pm 1$, as expected [68]. This explains why the second harmonic peak is less sharp.

In Fig. 4.17(b)-(c) we show one dimensional bias cuts in current and differential conductance respectively along the red and black lines from panel (a). We observe strong bending mode peaks at finite bias, which are only slightly stronger than the zero-bias Kondo peaks.

We see the effects of current annealing in this device in Fig. 4.18. This device was annealed for 10 min at 8 μA and ≈ 600 meV, corresponding to a power of 4.8 μW. It can be seen that after annealing, the bending modes and Kondo resonance
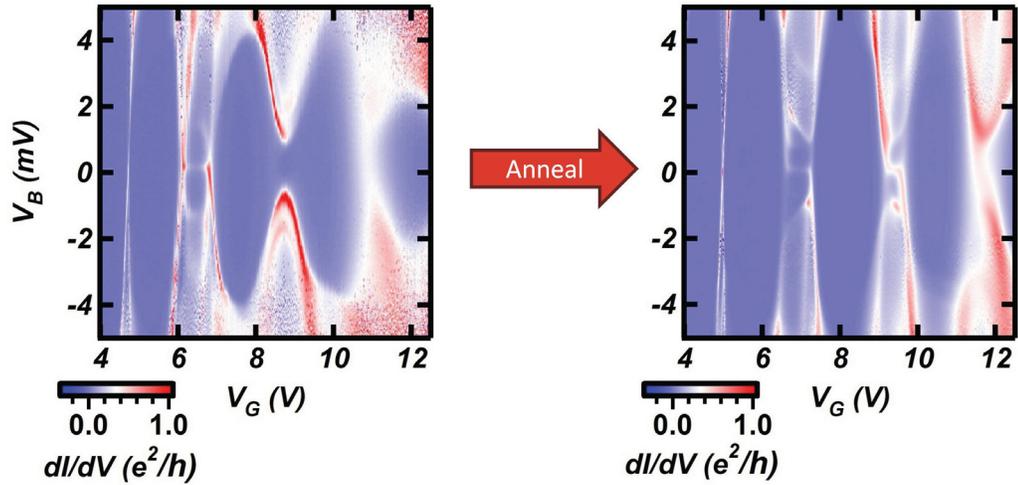
Figure 4.18: The effects of current annealing. Device B was annealed several times, but the best annealing was found to be 10 min at 4.8 μW. The clear result is that the Kondo peaks and bending resonances became stronger and better resolved.

both became sharper, and the second harmonic of the bending mode becomes visible. This is because annealing cleans the carbon nanotube, increasing its quality factor and improving the positive feedback resonance. The Kondo peak becomes sharper though stronger electron-vibron coupling, as explained in ref. [18]. Now we discuss another manifestation of the strong e-v coupling in this device, a renormalization of the classical energy needed to charge the QD.

4.5 Charging Energy Reduction by Bending Vibrons and the Kondo Effect

In devices where the vibronic frequency and coupling are both large, it becomes possible for polarons, quasiparticles made up of a moving electron and its surrounding polarization field in a crystal lattice, to alter the energy states of the system [18]. This can be thought of as a quasiparticle consisting of an electron surrounded by a cloud of vibrons. Depending on the strength of this e-v coupling, two possible regimes emerge. For weaker e-v coupling, the charging energy of odd-numbered diamonds is reduced by a polaronic shift, while for very strong e-v coupling, the effective charging energy becomes negative, causing a charge analogue of the Kondo effect to emerge [19, 76]. Strong electron-vibron coupling can also enhance the Kondo effect by increasing the QD-lead coupling, and allowing us to reach higher Kondo temperatures.

4.5.1 Charging Energy Renormalization: Background

In devices where the vibron frequency and e-v coupling are both large, it becomes possible for polaronic effects to alter the energy states of the system. In our ultra-short devices, both the degree of e-v coupling and frequency can be very high, amplifying the magnitude of this renormalization. The effective charging energy (E_c^*) of a system with strong $e - v$ coupling is given by [18, 66]

$$E_c^* = E_c - 2\lambda_{vib}^{*2}hf_0 \tag{4.18}$$

In our devices, the bending mode is weakly coupled, but the effective coupling can be strong $\lambda_{bend}^* = \lambda_{bend}\sqrt{Q} \sim 1$. In two-fold degenerate QDs, charging energy renormalization is predicted to affect the odd numbered diamonds only. This is because only the odd numbered states are affected by the polaronic shift, as determined from the Hamiltonian of the system [18]. This causes the merger of the blockade peaks in the gate direction, and a reduction of the diamond heights, which can be so strong that it suppresses the charging energy altogether, eliminating odd diamond states [19].

The Kondo coupling (J) is enhanced by e-v coupling in this regime, and therefore, so is the Kondo temperature. The normalized increase in Kondo coupling for weak e-v coupling is given by [18]

$$\frac{J_{vib}}{J_0} \approx 1 + \frac{2\lambda^2 h f_0 / E_c}{1 + E_c/h f_0} + h.c. \qquad T_K \propto \exp\left(\frac{-1}{J_{vib}\rho_0}\right) \qquad (4.19)$$

where ρ_0 is the electronic Green function describing the interaction between the SWCNT and the leads. From this equation, we can understand that as λ and f_0 increase, the Kondo temperature increases. Therefore, electron-vibron coupling enhances e-e interactions in our devices.

For very strong e-v coupling, such that $E_c - 2\lambda_{vib}^{*2} h f_0 < 0$, the negative effective charging energy results in an attractive force between electrons. A charge based analogue of the Kondo effect emerges, which has its own Kondo temperature, and is expected to eliminate Coulomb blockade diamonds [66]. In this case, the ground

state of the system is degenerate between zero and double occupancy, with the charge Kondo effect flipping between these two. This results in pair tunneling across the device [19]. In our devices, we observe the weak coupling regime, with charging energy reduction of the odd charge states, and an absence of the charge Kondo effect.

4.5.2 Observation of Charging Energy Reduction and the Kondo Effect

Because of the short length of our devices, we achieve high bending mode frequencies and strong *e-v* coupling. As a result, we observe a polaronic charging energy renormalization, and a Kondo effect which is enhanced by electron-vibron coupling. In device B, we observe a reduction of the charging energy, and width, of the odd diamonds, as shown in Fig. 4.19.
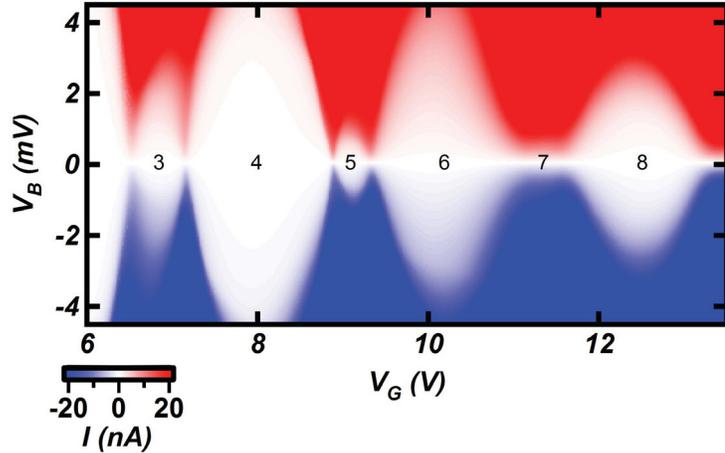
Figure 4.19: Charging energy reduction in Coulomb blockade. We observe for odd electron occupancy both the charging energy (E_c), and the width of the blockade diamonds (ΔV_G) are greatly reduced due to a polaronic shift of the QD energy levels.

First we investigate the Kondo effect in this 2-fold degenerate device. We determine the Kondo temperature, as shown in Fig. 4.20(a)-(b). A differential conductance vs. bias plot shows the HWHM of the Kondo peak to be $T_K \approx 1.6$ K. Using the Goldhaber-Gordon model, we extract a Kondo temperature $T_K \approx 1.5$ K, in good agreement, confirming that this is indeed the Kondo effect. We also observe splitting of the Kondo peak from 0-9 T, shown in Fig. 4.20(c)-(d). However, the g-factor we extract from this splitting is quite small, $g = 1.5 \pm 0.2$. This could be explained by spin-orbit coupling or strong e-e interactions such as the formation of a Wigner crystal [28]. An alternative explanation could be the strong e-v coupling we observe in our devices, interacts with the magnetic field, or that the g-factor of polarons in our devices is smaller than that of electrons.

At high temperature the Q-factor of the oscillator decreases due to thermal dissipation and the blockade peaks become thermally broadened. This disrupts the positive feedback mechanism necessary to excite the bending resonance, and renders the charging energy reduction negligible. We show this effect in Fig. 4.21(a)-(b) with differential conductance data from Device B at 50 K and 0.3 K respectively. In panels (c)-(d) we show the smooth V_{add} and ΔV_G curves at 50 K (black), in contrast to the drastic variations between even and odd occupation observed at 0.3 K (red). This is because as we lower T, the bending mode feedback becomes active, and polaronic shifts reduce the size of the odd diamonds.

By comparing the same charge states at different temperatures we can measure the effective charging energy. For the third electron state, we measure $E_c^* = 3.9 \pm 0.3$ meV,
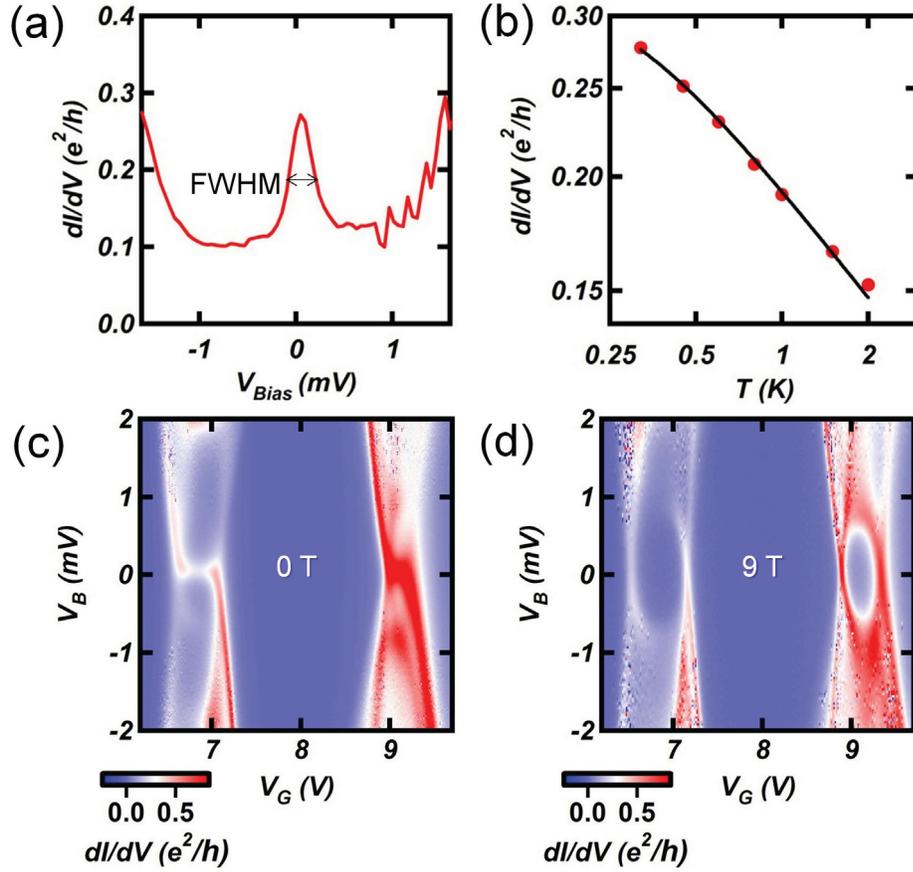
Figure 4.20: Bending mode and Kondo: temperature and magnetic field dependence. (a) Bias sweep along $V_G = 6.91$, showing Kondo peak with HWHM giving a Kondo temperature $T_K \approx 1.6$ K. (b) We observe good agreement with the Goldhaber-Gordon model for the temperature dependence of the differential conductance peak and confirm that $T_K \approx 1.5$ K. (c)-(d) We observe zero bias peak-splitting as we increase from 0 to 9 T, and calculate $g = 1.5 \pm 0.2$.
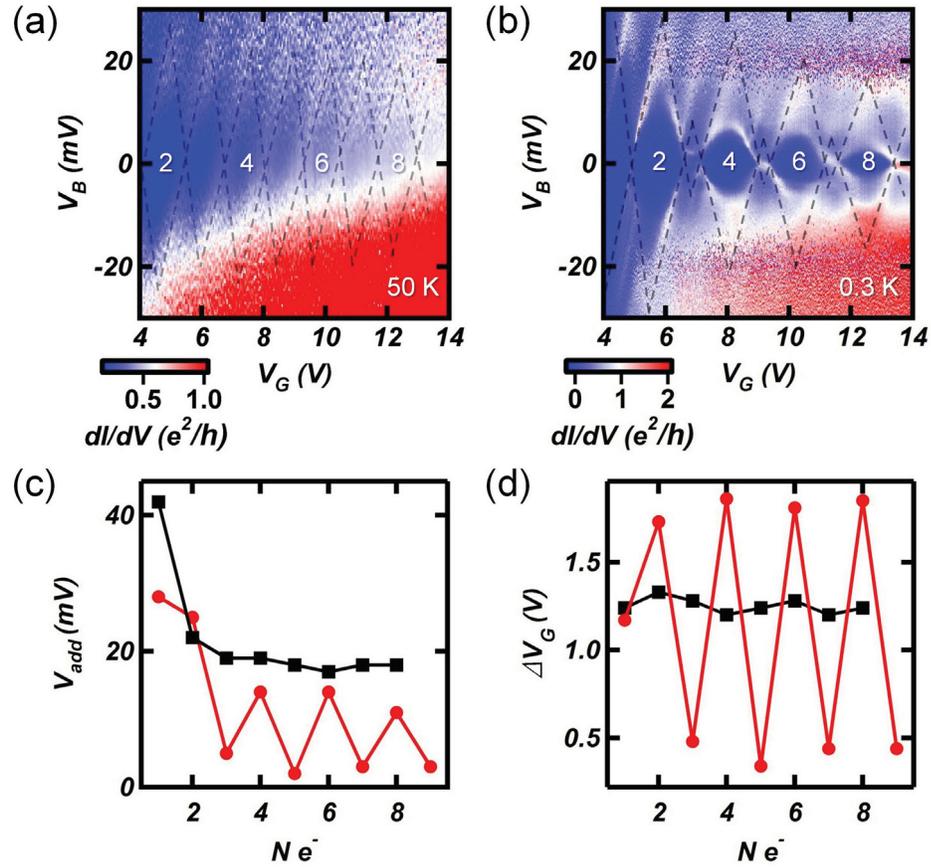
Figure 4.21: Charging energy reduction in device B. (a) Differential conductance plot showing evenly spaced diamonds at 50 K. (b) Similar differential conductance plot at 0.3 K, showing a reduction in charging energy for diamonds with odd occupation from the onset of the bending mode. (c)-(d) Addition energy and blockade diamond width at 50 K (black) and 0.3 K (red). We observe the clear charging energy reduction for odd numbered states at low temperature.

and $E_c = 18.5 \pm 2$ meV. Using the measured $f_{bend} \approx 50$ GHz, we determine the experimental e-v coupling to be $\lambda^*_{bend} = 5 \pm 2$, in agreement which the theoretically determined value of $\lambda^*_{bend} = 9 \pm 3$. Because bending frequency and effective e-v coupling increase with gate in this device, the charging energy of odd diamonds is reduced with increasing charge occupancy to the point of complete suppression.

To illustrate that the diamond spacing truly is regular at high T, we show an $I - V_G$ plot at 32 K in Fig. 4.22(a). We observe an increase in conductance away from the bandgap due to n-doping from the contacts, with a superimposed pattern of periodic blockade peaks. In panel (b), we show the gate sweep behaviour from 0.6 K to 16 K. We can observe the convergence of two phenomena with decreasing T.
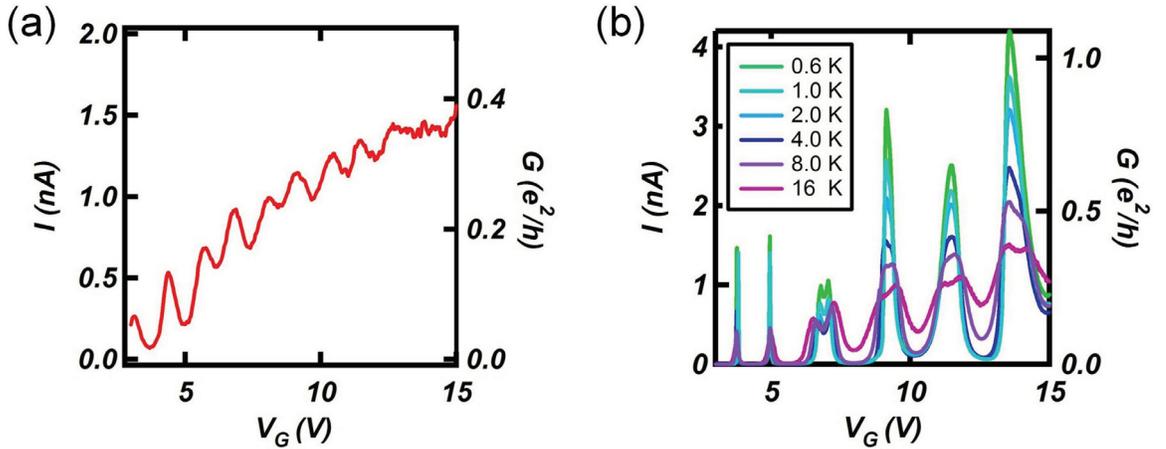
Figure 4.22: Effective charging energy and the Kondo effect. (a) Gate sweep of Device B at 32 K, showing regular Coulomb oscillations which saturate due to n-type doping by the contacts. (b) Gate sweeps as in (a) at temperatures ranging from 0.6 K to 16 K. The Coulomb peaks begin to merge for odd occupation with increasing temperature. This is due to the onset of the bending resonance at low temperature reducing E_c, and an increase in zero-bias conductance from the Kondo effect [18].

First, the effective charging energy becomes smaller. We observe pairs of evenly spaced peaks at high temperature which merge together at lower T. This is due to the thermal smearing of the bending mode resonance, which disrupts the positive feedback and eliminates the polaronic reduction of the charging energy at high T.

Second, the sets of two Coulomb peaks are not just being added linearly as temperature is lowered, and additional conductance peak begins to dominate. This is due to the Kondo effect which is enhanced by strong e-v coupling and high vibron frequency [18]. Above the Kondo temperature, this effect does not play a large role, but as the temperature decreases, the Kondo effect begins to obscure the blockade diamonds completely, as can be seen for high occupation at 0.6 K in panel (b). Thus, in this device, we observe the predicted behaviour of polarons and the Kondo effect [18] in molecular transistors for the first time, to our knowledge.

By attaining this regime with very small or even negative effective charging energy, new applications become available. For example, in devices with asymmetric source-drain tunnel barriers, and a negative effective charging energy, it has been predicted that the shape of the Coulomb peaks becomes gate-asymmetric [19]. Therefore, by tuning the gate to either the positive or negative blockade peak, only positive or negative current could flow respectively. This could allow such a nanometer scale device to rectify an AC signal into a DC signal. This type of device could also be used to explore new applications of electron-vibron coupling, form new types of NEMS devices or, to explore the charge Kondo effect and negative charging energies in ultra-short carbon nanotubes [19, 66].

Chapter 5

Conclusion

We refined an electromigration procedure to fabricate very short (~ 10 nm) and clean SWCNT quantum dots. Although the sub-10 nm QD regime has been touched on at room temperature [4], to our knowledge, no one has created suspended devices below the 10 nm length scale, or investigated ~ 10 nm SWCNTs at low temperature. By creating such short devices, we study contact effects and explore the limits of downscaling molecular transistors, investigate electromechanical interactions which can be used to form high quality NEMS devices [11, 51], study electron coherence effects [9] and investigate quasi-particle interactions [19]. Several of these effects are exemplified in the differential conductance plot from one of our devices shown in Fig. 5.1, which shows conductance asymmetry, Coulomb blockade, Fabry-Pérot interferences and a very strong Kondo effect all in the same device.

We described the fabrication steps and instrumentation necessary to generate and measure these devices, culminating in the electromigration procedure.
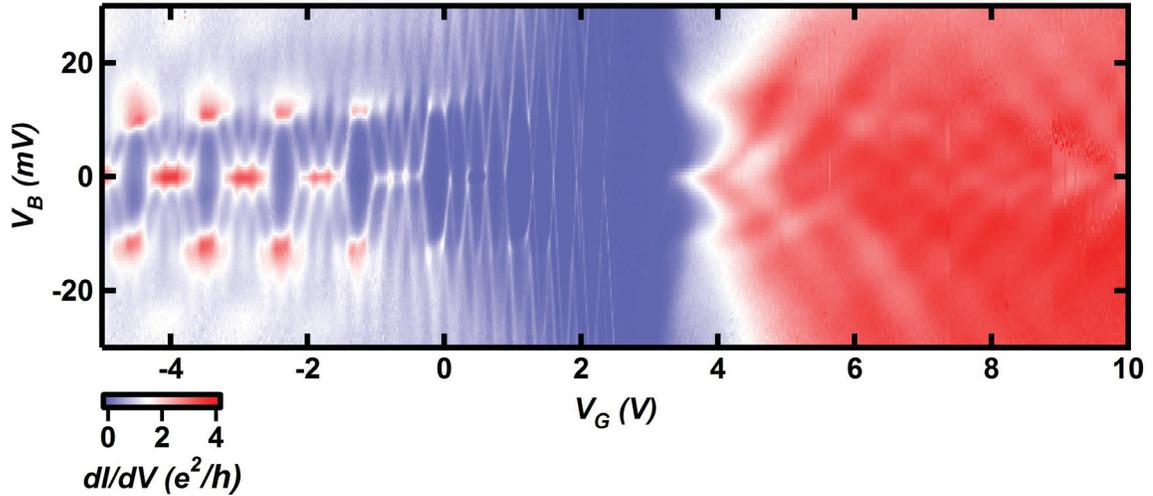
Figure 5.1: Vast tuneability in ultra-short SWCNTs. In our devices, we observe striking differences between the behaviour of electrons and holes and strong electron spin coherence resulting in gate-tuneable Kondo effects with T_K up to 28 K.

Electromigration allows us to etch our gold break junctions on the nanoscale, separating the contacts by just a few nanometers and defining our ultra-short devices, while simultaneously annealing them [3]. We measure a strong transistor effect with an on/off ratio of > 4100 in one of our devices at 300 K. At less than 10 nm, this is one of the smallest SWCNT transistors made to date, and its fabrication only made possible by our electromigration method.

We explored the electron-hole transport asymmetry which arises from doping by the contacts in our short devices. This results in a fundamental difference between the electronic potential, conductance, channel length, and transport regime for electrons and holes. In our devices, we observe an asymmetry in the charging energy of electrons and holes up to a factor of 3. In small bandgap devices, we observe a conductance asymmetry which allows us to tune between the Coulomb blockade regime and the

Fabry-Pérot regime using an electrostatically coupled gate electrode. These gate tuneable barriers could be useful in electronics devices which require coherent electron physics [29], while still preserving QD behaviour.

In our devices, mechanical bending vibrations of the nanotube have been observed with very high frequencies up to 280 GHz [11]. These devices have high Q-factors ($Q \sim 10^6$) which allow us to observe self-actuated bending mode resonances using DC transport. We observe the second harmonic of the bending mode and measure the effective electron-vibron coupling to be $\lambda^*_{bend} > 1$. As NEMS, these devices could be used as sensitive mass sensors [12, 13], or in charge pumping circuits [77].

Due to strong spin interactions between the QD and the leads, we measure some of the highest Kondo temperatures in SWCNTs to date. The Kondo temperature is enhanced by the short length of our devices, giving T_K up to 28 K. In addition, we find that the Kondo temperature is highly gate-tuneable to 17 K, nearly a factor of 2. This coherent electron physics could be exploited in spintronics applications [8, 9], be used as a tool to measure excitation spectra [16] and could further enhance the bending mode [15].

A reduction of the charging in our devices at low temperature indicates polaronic renormalization of the odd occupation energy levels of the quantum dot. This is a result of strongly coupled high frequency bending vibrons. We find that the charging energy is suppressed to the point of elimination for high electron occupation. Although predicted as early as 2004 [18], this is to our knowledge the first time this effect has been observed in molecular transistors. By enhancing this effect, new types

of devices, such as nanoscale current rectifiers or spintronics devices can be made using the charge analogue of the Kondo effect [19, 66].

By pushing scaling limits of SWCNT quantum dot transistors, we have been able to observe new fundamental physics, which helps us to gain understanding of short SWCNT systems and could be used in a variety of nanoscale electronics devices. While some of our research confirms what has been previously predicted, we also find some dramatic differences. These can allow us to engineer new types of devices, and push the scaling limits of SWCNT nanoelectronics into the few-nanometer range, allowing for new types of smaller, faster, and more efficient electronic devices in the future.

Bibliography

[1] A. Javey et al., *High-Field Quasiballistic Transport in Short Carbon Nanotubes*, Physical Review Letters **92** (2004).

[2] J. Guo, S. Datta, and M. Lundstrom, *A numerical study of scaling issues for Schottky-barrier carbon nanotube transistors*, Electron Devices, IEEE Transactions on **51**, 172 (2004).

[3] J. O. Island, V. Tayari, S. Yigen, A. C. McRae, and A. R. Champagne, *Ultra-short suspended single-wall carbon nanotube transistors*, APL **99** (2011).

[4] A. D. Franklin et al., *Sub-10 nm Carbon Nanotube Transistor*, Nano Letters **12**, 758 (2012).

[5] P. J. Leek et al., *Charge Pumping in Carbon Nanotubes*, Phys. Rev. Lett. **95**, 256802 (2005).

[6] X. Cheng et al., *Controlling Electron and Hole Charge Injection in Ambipolar Organic Field-Effect Transistors by Self-Assembled Monolayers*, Advanced Functional Materials **19**, 2407 (2009).

[7] Y. Miyauchi, H. Ajiki, and S. Maruyama, *Electron-hole asymmetry in single-walled carbon nanotubes probed by direct observation of transverse quasidark excitons*, Phys. Rev. B **81**, 121415 (2010).

[8] F. Kuemmeth, S. Ilani, D. C. Ralph, and P. L. McEuen, *Coupling of spin and orbital motion of electrons in carbon nanotubes*, Nature **452**, 448 (2008).

[9] F. Kuemmeth, H. O. H. Churchill, P. K. Herring, and C. M. Marcus, *Carbon nanotubes for coherent spintronics*, Materials Today **13**, 18 (2010).

[10] S. Savel'ev, X. Hu, and F. Nori, *Quantum electromechanics: Qubits from buckling nanobars*, New Journal of Physics **8**, 105 (2006).

[11] J. O. Island, V. Tayari, A. C. McRae, and A. R. Champagne, *Few-Hundred GHz Carbon Nanotube Nanoelectromechanical Systems (NEMS)*, Nano Letters **12**, 4564 (2012).

[12] H.-Y. Chiu, P. Hung, H. W. C. Postma, and M. Bockrath, *Atomic-Scale Mass Sensing Using Carbon Nanotube Resonators*, Nano Letters **8**, 4342 (2008).

[13] V. Sazonova et al., *A tunable carbon nanotube electromechanical oscillator*, Nature **431**, 284 (2004).

[14] F. Wu et al., *Single-walled carbon nanotube weak links in Kondo regime with zero-field splitting*, Physical Review B **79** (2009).

[15] G. D. Scott and D. Natelson, *Kondo Resonances in Molecular Devices*, ACS Nano **4**, 3560 (2010).

[16] J. Paaske et al., *Non-equilibrium singlet-triplet Kondo effect in carbon nanotubes*, Nature Physics **2**, 460 (2006).

[17] P. Jarillo-Herrero et al., *Orbital Kondo effect in carbon nanotubes*, Nature **434**, 484 (2005).

[18] P. S. Cornaglia, H. Ness, and D. R. Grempel, *Many-body effects on the transport properties of single-molecule devices*, Physical Review Letters **93** (2004).

[19] J. Koch, E. Sela, Y. Oreg, and F. von Oppen, *Nonequilibrium charge-Kondo transport through negative-U molecules*, Physical Review B **75**, 195402 (2007).

[20] L. Donev, *Carbon Nanotube Transistors: Capacitance Measurements, Localized Damage, And Use As Gold Scaffolding*, PhD thesis, Cornell, 2009.

[21] Z. Jin et al., *Ultralow Feeding Gas Flow Guiding Growth of Large-Scale Horizontally Aligned Single-Walled Carbon Nanotube Arrays*, Nano Letters **7**, 2073 (2007).

[22] A.-C. Dupuis, *The catalyst in the CCVD of carbon nanotubes–a review*, Progress in Materials Science **50**, 929 (2005).

[23] R. Y. Zhang, Y. Wei, L. A. Nagahara, I. Amlani, and R. K. Tsui, *The contrast mechanism in low voltage scanning electron microscopy of single-walled carbon nanotubes*, Nanotechnology **17**, 272 (2006).

[24] E. D. Minot, Y. Yaish, V. Sazonova, and P. L. McEuen, *Determination of electron orbital magnetic moments in carbon nanotubes*, Nature **428**, 536 (2004).

[25] R. Alizadegan, A. D. Liao, F. Xiong, E. Pop, and K. J. Hsia, *Effects of tip-nanotube interactions on atomic force microscopy imaging of carbon nanotubes*, Nano Research **5**, 235 (2012).

[26] H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, and P. L. McEuen, *Fabrication of metallic electrodes with nanometer separation by electromigration*, Applied Physics Letters **75**, 301 (1999).

[27] D. R. Strachan et al., *Controlled fabrication of nanogaps in ambient environment for molecular electronics*, Applied Physics Letters **86** (2005).

[28] P. Jarillo-Herrero, S. Sapmaz, C. Dekker, L. P. Kouwenhoven, and H. S. J. van der Zant, *Electron-hole symmetry in a semiconducting carbon nanotube quantum dot*, Nature **429**, 389 (2004).

[29] W. Liang et al., *Fabry-Perot interference in a nanotube electron waveguide*, Nature **411**, 665 (2001).

[30] S. Iijima, *Helical Microtubules of Graphitic Carbon*, Nature **354**, 56 (1991).

[31] A. Bianco, K. Kostarelos, and M. Prato, *Applications of carbon nanotubes in drug delivery*, Current opinion in chemical biology **9**, 674 (2005).

[32] M. W. Rowell et al., *Organic solar cells with carbon nanotube network electrodes*, Applied Physics Letters **88**, 233506 (2006).

[33] K. S. Novoselov et al., *Electric field effect in atomically thin carbon films*, Science **306**, 666 (2004).

[34] P. R. Wallace, *The Band Theory of Graphite*, Phys. Rev. **71**, 622 (1947).

[35] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *The electronic properties of graphene*, Rev. Mod. Phys. **81**, 109 (2009).

[36] J. W. Wilder, L. C. Venema, A. G. Rinzler, R. E. Smalley, and C. Dekker, *Electronic structure of atomically resolved carbon nanotubes*, Nature **391**, 59 (1998).

[37] X. Lu and Z. Chen, *Curved Pi-Conjugation, Aromaticity, and the Related Chemistry of Small Fullerenes (<C60) and Single-Walled Carbon Nanotubes*, Chemical Reviews **105**, 3643 (2005).

[38] J. Ding, X. Yan, and J. Cao, *Analytical relation of band gaps to both chirality and diameter of single-wall carbon nanotubes*, Physical Review B **66** (2002).

[39] M. J. Biercuk, S. Ilani, C. M. Marcus, and P. L. McEuen, *Electrical Transport in Single-Wall Carbon Nanotubes*, volume 111 of *Topics in Applied Physics*, chapter 15, pages 455–493, Springer Berlin Heidelberg, 2008.

[40] S. Datta, *Electronic Transport in Mesoscopic Systems (Cambridge Studies in Semiconductor Physics and Microelectronic Engineering)*, Cambridge University Press, 1997.

[41] J. Park, *Electron Transport in Single Molecule Transistors*, PhD thesis, Berkeley, 2003.

[42] N. Nemec, D. Tomanek, and G. Cuniberti, *Contact dependence of carrier injection in carbon nanotubes: An ab initio study*, Physical Review Letters **96** (2006).

[43] J. Knoch et al., An extended model for carbon nanotube field-effect transistors, in *Device Research Conference, 2004. 62nd DRC. Conference Digest [Includes 'Late News Papers' volume]*, pages 135–136 vol.1, 2004.

[44] P. A. Khomyakov, A. A. Starikov, G. Brocks, and P. J. Kelly, *Nonlinear screening of charges induced in graphene by metal contacts*, Physical Review B **82** (2010).

[45] S. Heinze et al., *Carbon Nanotubes as Schottky Barrier Transistors*, Physical Review Letters **89** (2002).

[46] P. Jarillo-Herrero et al., *Electronic transport spectroscopy of carbon nanotubes in a magnetic field*, Physical Review Letters **94**, 156802 (2005).

[47] K. Grove-Rasmussen, H. I. Jørgensen, and P. E. Lindelof, *Fabry-Perot interference, Kondo effect and Coulomb blockade in carbon nanotubes*, Physica E: Low-Dimensional Systems & Nanostructures **40**, 92 (2007).

[48] H. Jrgensen, K. Grove-Rasmussen, K. Flensberg, and P. Lindelof, *Critical and excess current through an open quantum dot: Temperature and magnetic-field dependence*, Physical Review B **79** (2009).

[49] V. Derycke, R. Martel, J. Appenzeller, and P. Avouris, *Controlling doping and carrier injection in carbon nanotube transistors*, Applied Physics Letters **80**, 2773 (2002).

[50] V. Iancu, A. Deshpande, and S.-W. Hla, *Manipulation of the Kondo effect via two-dimensional molecular assembly*, Physical Review Letters **97**, 266603 (2006).

[51] J. Chaste et al., *A nanomechanical mass sensor with yoctogram resolution*, Nature Nanotechnology **7**, 301 (2012).

[52] S. Sasaki et al., *Kondo effect in an integer-spin quantum dot*, Nature **405**, 764 (2000).

[53] J. Nygard, D. H. Cobden, and P. E. Lindelof, *Kondo physics in carbon nanotubes*, Nature **408**, 342 (2000).

[54] A. Makarovski and G. Finkelstein, *SU(4) mixed valence regime in carbon nanotube quantum dots*, Physica B: Condensed Matter **403**, 1555 (2008).

[55] F. Haldane, *Scaling theory of the asymmetric Anderson model*, Physical Review Letters **40**, 416 (1978).

[56] J. Park, S.-S. Lee, Y. Oreg, and H.-S. Sim, *How to Directly Measure a Kondo Cloud's Length*, Physical Review Letters **110**, 246603 (2013).

[57] D. Goldhaber-Gordon et al., *From the Kondo Regime to the Mixed-Valence Regime in a Single-Electron Transistor*, Physical Review Letters **81**, 5225 (1998).

[58] L. Tosi, P. Roura-Bas, and A. A. Aligia, *Transition between SU(4) and SU(2) Kondo effect*, Physica B: Condensed Matter **407**, 3259 (2012).

[59] B. Babić, T. Kontos, and C. Schnenberger, *Kondo effect in carbon nanotubes at half filling*, Physical Review B **70** (2004).

[60] L. H. Yu et al., *Inelastic electron tunneling via molecular vibrations in single-molecule transistors*, Physical Review Letters **93** (2004).

[61] S. Sapmaz et al., *Electronic excitation spectrum of metallic carbon nanotubes*, Physical Review B **71** (2005).

[62] J. Lehmann and D. Loss, *Cotunneling current through quantum dots with phonon-assisted spin-flip processes*, Physical Review B **73** (2006).

[63] J. Nygard, W. Koehl, N. Mason, L. DiCarlo, and C. Marcus, *Zero-field splitting of Kondo resonances in a carbon nanotube quantum dot*, arXiv preprint cond-mat/0410467 (2004).

[64] J. H. Chen, L. Li, W. G. Cullen, E. D. Williams, and M. S. Fuhrer, *Tunable Kondo effect in graphene with defects*, Nature Physics **7**, 535 (2011).

[65] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven, *A tunable Kondo effect in quantum dots*, Science **281**, 540 (1998).

[66] P. S. Cornaglia and D. R. Grempel, *Magnetoconductance through a vibrating molecule in the Kondo regime*, Physical Review B **71** (2005).

[67] M. S. Dresselhaus, G. Dresselhaus, A. Jorio, A. G. Souza, and R. Saito, *Raman spectroscopy on isolated single wall carbon nanotubes*, Carbon **40**, 2043 (2002).

[68] E. Mariani and F. von Oppen, *Electron-vibron coupling in suspended carbon nanotube quantum dots*, Physical Review B **80** (2009).

[69] R. Leturcq et al., *Franck-Condon blockade in suspended carbon nanotube quantum dots*, Nature Physics **5**, 327 (2009).

[70] S. Sapmaz, P. Jarillo-Herrero, Y. M. Blanter, C. Dekker, and H. Van Der Zant, *Tunneling in suspended carbon nanotubes assisted by longitudinal phonons*, Physical Review Letters **96**, 026801 (2006).

[71] Y. M. Blanter, O. Usmani, and Y. V. Nazarov, *Single-electron tunneling with strong mechanical feedback*, Physical Review Letters **93** (2004).

[72] B. Lassagne, Y. Tarakanov, J. Kinaret, D. Garcia-Sanchez, and A. Bachtold, *Coupling mechanics to charge transport in carbon nanotube mechanical resonators*, Science **325**, 1107 (2009).

[73] O. Usmani, Y. M. Blanter, and Y. V. Nazarov, *Strong feedback and current noise in nanoelectromechanical systems*, Physical Review B **75** (2007).

[74] S. Sapmaz, Y. M. Blanter, L. Gurevich, and H. Van der Zant, *Carbon nanotubes as nanoelectromechanical systems*, Physical Review B **67**, 235414 (2003).

[75] E. A. Laird, F. Pei, W. Tang, G. A. Steele, and L. P. Kouwenhoven, *A high quality factor carbon nanotube mechanical resonator at 39 GHz*, Nano Lett **12**, 193 (2012).

[76] P. S. Cornaglia, D. R. Grempel, and H. Ness, *Quantum transport through a deformable molecular transistor*, Physical Review B **71** (2005).

[77] M. R. Buitelaar et al., *Adiabatic Charge Pumping in Carbon Nanotube Quantum Dots*, Physical Review Letters **101**, 126803 (2008).

Appendices

Appendix A

Operation of Cryostat

Putting the system in neutral. When neither condensing nor cooling:

- Sorb Heater: off

- Needle Valve: ≈ 8

- ^{3}He dumps: ≈ 0 mbar

- Pirani: $1 - 5$ mbar

- 1K pot: $1.3 - 1.5$ K

These settings allow a liquid ^{4}He reservoir to build up in the sock and maximal cooling of the 1K pot. The needle valve allows liquid ^{4}He to flow from the sock to the 1K pot. As the needle valve is closed, the 1K pot will drop in temperature; this is because the maximum surface area to volume ratio in the 1K pot will be achieved. Be sure not to completely block the flow through the needle valve.

Condensing Helium 3 into the ^{3}He tail:

- Check Loop 1 settings, 40-45 K on Cryocon

- Sorb Heater: on ("control" button on Cryocon)

- Needle Valve: ≈ 7

- Allow temperature to stabilize at 40-45 K

- Let ^{3}He condense for 1-2 hours

- ^{3}He dump pressure should rise to ≈ 40 mbar, then fall to 0 mbar

- 1K pot 1.3 K

These settings will allow the sorb to heat up, while still maintaining a low temperature on the 1K pot and not boil off much LHe4 in the sock. Heating the sorb outgases the ^{3}He. When this gas passes near (through) the 1K pot, it cools and condenses into the ^{3}He tail.

Running to Base Temperature

- Sorb Heater: off (push "stop" button on Cryocon)

- Needle Valve: ≈ 10.5

- Allow temperatures to drop and stabilize at $\approx$ 0.3 K

- Then close needle valve to ≈ 9

- ^{3}He dump pressure be 0 mbar

- Sorb temperature should drop to $\lesssim 12$ K

Having the needle valve open more at first will cause the 1 K pot to rise in temperature and the LHe4 reservoir in the sock to start depleting (^{4}He dump pressure increases). However, opening the needle valve will cool the sorb more quickly and cause less ^{3}He to be absorbed on cooldown, increasing hold time. Once the sorb has cooled, adjust the needle valve so that the sorb stays below 12 K, but the 1K pot is as cold as possible.

Appendix B

Igor Pro Macros

Introduction:

These macros will make data analysis easier, but they often depend on each other. I will outline the dependencies here, discussing what input is needed and then state the procedure for using each macro, before mentioning why certain errors are called. It is important to know that Igor can only handle wave names with up to 31 characters. Here I limit the char number to 30 just to be safe. In addition, all Igor object names must only contain alphanumeric characters and the underscore. No other characters are allowed. I did not have time to create error detection for this problem.

ChopMacro v 1.2:

Function: This macro is a function which runs off of Matt's Partial Chop and Load macro, making it easier to use. This macro chops a 2D wave into all possible 1D bias or gate waves. This macro requires input with the same formatting as the output from Plot Everything v4.2.

Dependencies: This macro depends on functions within the ChopMovie v7.5 macro and the Partial Chop and Load v3.1 macro.

Usage: To use this macro, select a top graph (2D plot) and click the ChopMacroBGC item from the dropdown list. You will be prompted to select whether you want bias or gate waves. Click ok and the waves will be chopped for you.

Errors: This macro leaves a buffer of 10 characters in the macro name for the output, so the input wave cannot be more than 20 characters. This macro does not work for dI/dV plots.

ChopMovie v 7.5:

Function: This macro chops a given 2D wave into all possible 1D bias or gate waves and then plots them as 1D cuts according to the position of the cursor on the plot. This cursor can be moved by mouse or by the arrow keys to create a movie of the 1D cuts. This macro requires input with the same formatting as the output from Plot Everything v4.2 or Rebin2D v4.0.

Dependencies: This macro depends on functions within the Partial Chop and Load v3.1 function.

Usage: To use this macro, select the top graph (2D plot) and click the chop_movie item from the dropdown list. Now you may drag the cursor around with the mouse or keyboard to see differential and current 1D cuts for bias and gate sweeps. To save a graph as a new plot, click export. It will create the same plot with an "e" extension. When finished, click stop and you will be prompted to either keep the chopped waves, or kill them.

Errors: This macro leaves a buffer of 12 characters in the macro name for the output, so the input wave cannot be more than 18 characters. This macro does not work for dI/dV plots. If an instance of the moviecontrol window is still active, it must be closed. The source waves must be present.

Colourplot Mega v 4.0:

Function: This is a modification of the colourplot macro by Joshua. It is designed to work in sync with the other macros and depict $I - V_B - V_G$ 2D data properly. It is easier to use this function through the Plot Everything macro, but this one still works if that is not possible.

Dependencies: This macro depends on Image Range adjust.

Usage: To use this macro, just choose colourplot_mega from the dropdown list. You will be prompted to fill in data about your sample. "Plot everything" does this for you, but if you need to use this macro: select bias gate and current waves from the dropdown menus of the macro, select output matrix name and choose whether you want image, table or both, Temperature at which the data was taken, prefactor symbol (n,u,m) and corresponding factor (1E-9,1E-6,1E-3). Igor will do the rest.

Errors: None built in.

Extract Cursor (BCG) v 4.0:

Function: This macro creates a table of values from dI/dV and current plots selected by the user, allowing the user to select points highlighting certain features of the plots. Dependencies: None.

Usage: To use this macro bring a 2D image to the foreground, either $I - V_B - V_G$ or $dI/dV - V_B - V_G$. Choose the macro from the dropdown list, name the table where the data will be stored, and choose if your data is dI/dV or current in the z axis. Adjust the cursor to collect the first value. This will be a 3 column $I - V_B - V_G$ table for a current plot. For a dI/dV plot, the macro with store $dI/dV - V_B - V_G$ data and corresponding current data (as long as the current wave is present) in a 4 column table. Click continue after moving the cursor to the first point, select if you would like to continue or stop taking data points. Repeat as necessary.

Errors: Will show error if Chop_Movie macro is running. Both depend on cursors and cannot run at the same time.

Partial chop and Load v 3.1

Function: This function chops specific waves from 2D data. An easier, but more restricted user input is available through ChopMacro_v1.2. For more information, see Matt's PDF on this macro.

PlotEverything_v_4.4:

Function: This macro takes output from motorcurve and electromigrator software by Joshua and plots it in a nice format. This macro works for bias sweeps, gate sweeps and megasweeps (bias ad gate) for plotting differential conductance and current as well as annealing curves and migration curves. All input voltages must be in volts, unless otherwise specified. Input must be made in the following manner: any number of characters and underscores followed by a space, then automatic naming from Joshua's software. Ex. ****_**_****_*_ Mega_Sweep or ****_***** ELECTROM. This is easily achievable in the software by typing in your data set name, followed by a single space. The software will do the rest. Igor will save the waves under the same name with a new prefix ex. _Bias, _Curr, _Gate. This format can then be used in the other macros. For a megasweep, the suffix DIF implies a 2D image wave, while DIF1, is the 1D dI/dV wave.

Dependencies: This macro depends on the colourplot_mega_v4.0 macro for mega sweeps and Image Range Adjust.

Usage: To use this macro select it from the dropdown menu. You will be prompted to input the temperature at which the data was taken and the external resistance of the circuit (input of zero resistance does not take this into account). In addition, this resistance correction will fail if the bias points are too close to zero, causing a round off singularity. The macro will automatically take you to the computer's

data folder where you can select your data file. The software will read the type of data and the current units, incorporating this into its plotting algorithms. If you select a Gate_Sweep data file, you will be prompted to input a bias voltage at which the gate sweep was taken. Igor will save an image of your plot automatically to C:/data/igorimages/(todays date).

Errors: If the file name does not contain a space, Igor will give an error as it cannot find what type of data it is. If the name is too long, Igor will give an error. If waves of the same name already exist, Igor will replot the data, but will not overwrite the old waves. An invalid prefactor in the data file will also produce an error. This would mean that something has been changed in the CVI file.

SaveImage v 1.0:

Function: This macro saves the top graph to C:/data/igorimages/todays date, naming it automatically. This is much faster than going through the "save graphics" menu.

Usage: Bring whichever graph you want to save to the front of the experiment. Select save image from the dropdown menu.

Errors: If no image or trace is detected on the top graph, the macro will abort.

Rebin1D v 1.3:

Function: This macro uses the built-in interpolate function to bin data points together, averaging out noise. Both the input bias/gate wave and current wave will be rebinned.

Usage: To use this macro select it from the dropdown menu. You will be prompted to select the x and y waves for rebinning, as well as how many points you want to average over per new data point: the bin size (does not have to be an integer). Lastly, choose the prefactor of the current wave. Igor will plot and save the new waves.

Errors: If wave name is too long, an error will occur. If x and y waves have different dimensions, an error will occur. If binsize is greater than the number of points or less than one (indicating a greater number of data points), an error will occur. Invalid prefactor will also result in an error.

Rebin2D v 4.0:

Function: This macro uses a function similar to rebin1d after chopping all of the bias (gate) waves. It then parses all of this information back together to be plotted as a 2D image. This macro can handle either one bias rebin, one gate rebin, or one bias and one gate rebin. Due to naming restrictions, further rebinning is not possible Finally, this macro will add a fair bit of text to the wave name, so it is important to keep wave names short, especially if you plan on doing a bias rebin and a gate rebin of the same graph.

Dependencies: This macro depends on colourplot_mega v.4.0 and Image Range adjust.

Usage: Bring the selected I-Vb-Vg plot to the foreground. Select rebin2d from the drop down menu and choose whether you want to rebin the bias or gate waves, and how many points per bin (does not have to be an integer) you would like. The macro will do the rest, plotting the $I - V_B - V_G$ and dI/dV plots.

Errors: This macro will give an error if the plot is not $I - V_B - V_G$, if you try to rebin the same direction on the same plot twice, if the wave name is too long, if the bin size is too big (greater than number of points) or too small (< 1) or if the wave dimensions are off.

Appendix C

Table C.1: Device Summary

Device	Length (nm)	Bandgap (meV)	$V_{add}^{h^+}/V_{add}^{e^-}$	T_K (K)	f_{vib} (GHz)
A[1]	15 ± 5	-	-	4.4 ± 0.2	-
B[2]	15 ± 5	255 ± 35	3.1	1.6 ± 0.2	$f_b = 45 \pm 13$
C	45 ± 10	110 ± 15	1.2	-	-
D[3]	≈ 5	380 ± 80	3.4	-	$f_s \approx 3800$
E[4]	104 ± 9	54 ± 10	-	≈ 28	-

[1]Kondo splitting by magnetic impurity at $B = 0$ and singlet-triplet Kondo effect
[2]Second bending harmonic and polaronic renormalization, $E_c^* \to 0$
[3]Room temperature transistor with $I_{on}/I_{off} > 4100$
[4]Gate-tuneable transition from Coulomb blockade to Fabry-Pérot and very strong Kondo effect