EXPERIMENTAL STUDY OF INTERFACES IN VAN DER WAALS HETEROSTRUCTURES

By

MICHAEL ALTVATER

A dissertation submitted to the
School of Graduate Studies
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Physics and Astronomy
written under the direction of
Eva Y. Andrei
and approved by

____________________________

____________________________

____________________________

____________________________

New Brunswick, New Jersey
October, 2020
ABSTRACT OF THE DISSERTATION

Experimental Study of Interfaces in van der Waals Heterostructures

By MICHAEL ALTVATER

Dissertation Director:
Eva Y. Andrei

The advent of 2D materials and heterostructures built from them has unleashed a huge potential to create and control novel electron systems. This thesis experimentally explores physical effects on the electronic states at the interfaces of 2D materials using electrostatic atomic force microscopy (AFM) and scanning tunneling microscopy (STM). The findings described herein include the observation of atomically thin graphene and its local charge distributions through an insulating encapsulant using electrostatic force microscopy (EFM) and Kelvin probe force microscopy (KPFM). Similar to encapsulation with hexagonal boron nitride (hBN), it is shown that covering a 2D material surface using monolayer graphene protects from damaging interactions with the environment, preventing detrimental oxidation of the sample surface in ambient conditions, with the added benefit of being able to image local charge distributions at atomic scales through the cover layer. Using this method, we image the room temperature charge density wave (CDW) phase of $1\text{T-TaS}_2$, a
highly correlated 2D material, and identify the ordering of topological defects of the 2D CDW state to be related to that of Abrikosov vortex lattices in type II superconductor films. The interaction between the graphene layer and the 1T–TaS$_2$ surface is further probed at 77K where 1T–TaS$_2$ exhibits a commensurate CDW coupled to a Mott insulating electronic phase. Itinerant carriers within the graphene layer are found to screen the electron-electron interactions and reduce the Mott gap size at the 1T–TaS$_2$ surface. Simultaneously, a charge density wave is observed to be induced in the graphene layer that is analogous to the superconducting proximity effect. The novel CDW proximity effect is found to be well captured within density functional theory (DFT) as well as with a simplified mean field Hamiltonian which allows the effect to be generalized to other materials.
Acknowledgments

Foremost, I would like to thank my adviser Eva Andrei for all of her help and guidance throughout my years in graduate school. Her enthusiasm and love of physics has been an inspiration. Additionally, I am thankful to Guohong Li who has been an invaluable source of expertise and wisdom during my time at Rutgers. I am grateful to the previous graduate students and postdocs, Adina Luican-Mayer, Chih-Pin Lu, and Ivan Skatchko, who acquainted me with the lab and started me off in the right direction. Their work has greatly helped shape my own research and my goals as a scientist as well as an educator. I would like to thank the postdocs, Junxi Duan, Jin-hai Mao, and Yuhang Jiang, for many helpful discussions as well as fellow graduate students, Nikhil Tilak, Angela Coe, Xinyuan Lai, and Zhenyuan Zhang, for stimulating lunches and patiently lending an ear. Finally, I am indebted to my family and my wife, Brianna, for always helping me push forward toward my goals; I could not have completed this work without their support.
Dedication

For the next student
# Table of Contents

Abstract ....................................................... ii

Acknowledgments ........................................ iv

Dedication .................................................. v

List of Figures .............................................. ix

1. Introduction ........................................ 1
   1.1. Introduction to 2D Materials and Heterostructures .... 1
   1.2. Overview of Thesis ................................ 4

2. Theory and Background .............................. 6
   2.1. Tight Binding Model .............................. 6
       2.1.1. Graphene .................................... 7
       2.1.2. Strain and Pseudomagnetic Fields in Graphene .... 13
       2.1.3. Moiré Patterns and Superpotentials ................. 20
   2.2. Charge Density Waves ............................ 23
       2.2.1. CDWs in Two Dimensions ..................... 31
       2.2.2. Chiral CDWs ................................ 36
   2.3. Electronic Phases of Tantalum Disulfide ........... 37
       2.3.1. Undistorted Phase ........................... 39
       2.3.2. Incommensurate Charge Density Wave Phase ....... 40
       2.3.3. Commensurate/Mott-Insulating Phase ............... 43
2.3.4. Nearly-Commensurate Phase .............................................. 45
2.3.5. T Phase ........................................................................ 48
2.3.6. Manipulation and Metastability of CDW Phases .......... 50
2.4. 1T-TaS$_2$ as a Quantum Spin Liquid ............................. 54
  2.4.1. What is a Quantum Spin Liquid? ................................. 54
  2.4.2. Is 1T-TaS$_2$ a QSL? .................................................. 56
2.5. Intrinsic Stacking Order in the CCDW Phase of 1T-TaS$_2$ 57

3. Experimental Methods ............................................................... 61
  3.1. Scanning Probe Microscopy ............................................. 61
    3.1.1. STM and STS .......................................................... 62
  3.2. Atomic Force Microscopy ................................................. 66
    3.2.1. Atomic Force Microscopy ........................................ 66
    3.2.2. Electrostatic AFM Measurement Modes .................... 68
  3.3. STM Design and Construction ......................................... 71
    3.3.1. STM Tip ............................................................... 71
    3.3.2. Scanning ............................................................. 76
    3.3.3. Stepper Motors ..................................................... 78
    3.3.4. Housing Design ..................................................... 82
  3.4. Noise Testing ................................................................. 85
    3.4.1. Signal Control Theory ............................................ 85
    3.4.2. Control System Outputs ......................................... 91
    3.4.3. Grounding and Shielding ....................................... 92
    3.4.4. Vibration Isolation ............................................... 94
  3.5. Capacitive Alignment ....................................................... 97
    3.5.1. Capacitive Sensing ............................................... 97
    3.5.2. Alignment Algorithm ............................................. 101
List of Figures

2.1. Graphene Lattice ............................................. 8
2.2. Graphene Band Structure .......................................... 9
2.4. Uniaxial Strain in Graphene ........................................ 15
2.5. Pseudo-Landau Levels in Triaxially Strained Graphene ....... 18
2.6. Moiré Patterns in Twisted Atomic Layers ......................... 20
2.7. 1D Charge Density Wave ........................................... 25
2.8. 2D CDW Components .............................................. 32
2.9. 2D CDW Harmonic Expansion ...................................... 34
2.10. 2D CDW Registrations ............................................. 36
2.11. Undistorted 1 T–TaS$_2$ ........................................... 39
2.12. Incommensurate Charge Density Wave Phase ................. 41
2.13. Commensurate Charge Density Wave Phase ..................... 43
2.15. Introduction to Quantum Spin Liquids ......................... 55
2.16. Stacking Orders .................................................. 58
3.1. STM Principles of Operation ....................................... 62
3.2. AFM Principles of Operation ....................................... 66
3.3. Principles of KPFM and EFM ........................................ 68
3.4. Tungsten Tip Etching Procedure ..................................... 73
3.5. Coarse Etching Supplies ........................................... 74
3.6. Fine Etching Supplies .............................................. 75
3.7. Scanner Piezo .................................................... 77
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3. Work Function vs Thickness of Exfoliated 1T–TaS₂</td>
<td>144</td>
</tr>
<tr>
<td>5.4. Protection from Oxidation By Graphene Encapsulation</td>
<td>146</td>
</tr>
<tr>
<td>5.5. Room Temperature STM of G/TaS₂</td>
<td>149</td>
</tr>
<tr>
<td>5.6. CDW Domains in the NCCDW Phase</td>
<td>151</td>
</tr>
<tr>
<td>5.7. Room Temperature Spectroscopy of G/TaS₂</td>
<td>155</td>
</tr>
<tr>
<td>5.8. Comparing Bare and Graphene Covered TaS₂ at 77K</td>
<td>157</td>
</tr>
<tr>
<td>5.9. Moiré Pattern between Graphene and 1T–TaS₂</td>
<td>160</td>
</tr>
<tr>
<td>5.10. Band Structures of 1T–TaS₂ and Graphene on 1T–TaS₂</td>
<td>162</td>
</tr>
<tr>
<td>5.11. Density of States of Graphene on Commensurate 1T–TaS₂</td>
<td>167</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Introduction to 2D Materials and Heterostructures

Freestanding, two-dimensional (2D) systems do not exist spontaneously in nature as low energy thermal fluctuations are enough to destabilize long-range ordering [1, 2, 3, 4]. Never-the-less, 2D electron systems can be found weakly coupled to the boundaries of finite systems in 3D. Several examples of these have been experimentally fabricated and probed such as electrons on the surface of superfluid helium [5, 6, 7, 8] or confined to the surfaces and interfaces of metals and semiconductors [9, 10]. These 2D electron systems are highly sensitive to perturbations making fabrication and probing their properties a hard-fought battle. Following the Nobel Prize-winning isolation of atomically thin flakes of graphene from the layered 3D material, graphite [11], the availability and interest of 2D electron systems has grown rapidly.

Carbon sheets in graphite are covalently bonded within the atomic plane but weakly bound to neighboring layers by van der Waals (vdW) forces. Mechanically cleaving graphite results in easy separation of vdW-bonded layers without damaging the strong in-plane carbon-carbon bonds. In 2004, the anisotropic bonding nature of graphite allowed the separation of a single layer of graphite, called graphene [11]. The isolation and measurement of single layer graphene introduced a new class of materials: 2D materials. These are, like graphite, characterized by strong in-plane covalent bonding and weak van der Waals bonds between atomic planes. The anisotropic nature of their bonds is reflected in their electronic properties [12, 13, 14, 15]. Being atomically thin, electrons in these materials are confined
to a 2D plane leading to observations of features expected of two-dimensional electronic systems such as the generation of Landau levels and quantum Hall effect in magnetic field [16, 17, 18, 19, 20, 21], frustrated spin states [22, 23], and a high-degree of electronic and lattice interactions [24].

Since 2004, 2D materials with a variety of electronic properties have been investigated including 2D metals, semimetals, half-metals, insulators, semiconductors, and superconductors although there are well over 1000 predicted 2D materials which have yet to be synthesized and measured [25, 26, 27, 28]. With so many materials and properties to choose from, the potential for novel scientific discoveries and emerging new technologies is high. Additionally, two dimensional crystals offer several benefits over traditional 2D electron gas systems. Foremost, making electrical contact to probe these systems is much easier than those where the electron gas lies beneath the material surface and/or is highly sensitive to perturbations. The electron gasses provided by 2D crystals often exhibit high carrier mobilities which has lead to the observation of many interesting properties of ballistic electrons including micron-scale, room temperature ballistic transport [29], electron optics [30, 31], and magnetic focusing [32, 33, 34]. Changing atomic content and lattice structures gives different electronic dispersions, carrier densities, and interactions further extending the phase space of properties to explore. Additionally, the existence of ordered states such as ferromagnetism, charge density waves, and superconductivity in two dimensions leads to interesting questions regarding how these systems order (and disorder), how their properties are affected by confined degrees of freedom, and how they might be manipulated to develop new technologies.

The properties of 2D materials can be controllably manipulated. Further, isolated, atomically-thin layers often exhibit properties that are different than those of their bulk counterparts [35]. The electronic properties are highly sensitive to strain. Are able to withstand much larger strain than their 3D counterparts. Using atomically thin layers allows one to produce large strain through poking, bending, and folding
the material; inducing controllable changes in electronic and structural properties [36, 37, 38, 39]. can be cut and folded into complex structures. They can be doped chemically [40, 41, 42, 43] and electrostatically. Already, 2D materials have been used to create new devices for solar energy conversion and storage [44], in sensors and radio transmitters (to be used in internet of things (IoT) applications), as well as gas sensors and bio-sensors [45, 46]. They are highly sensitive to local potentials from their environment and can be lifted and stacked onto arbitrary substrates. Therefore, we can combine layers together to create novel systems with tailored functionalities beyond the capabilities of their constituent parts [47, 48, 49, 50, 51, 52, 53, 54].

Heterostructures made from 2D materials then hold a great potential to generate novel technologies which overcome the limitations of today’s electronic devices. With smaller transistors and interconnects, integrated circuits are able to be produced much cheaper and faster with many benefits to circuit speed and power consumption. However, due to their small sizes, today's switches are becoming more affected by thermal fluctuations, parasitic currents, tunneling, and other quantum effects. These reduce the stability of the device during operation and contribute to generation of waste heat. In addition to these performance issues, fabrication of modern switches has become increasingly difficult as interconnects, oxide thickness, channel length, gate length, etc. become smaller [55].

Switches utilizing electronic phase transitions in strongly correlated materials use the sudden changes in a system's band structure near the Fermi energy during a phase transition to produce drastic changes in the device conductivity. Compared to today's FETs, which modulate the Fermi energy relative to a material's band structure, these devices can achieve sub-threshold swings of $S \propto I \cdot \frac{dV}{dI} = \sigma \cdot \frac{dV}{d\sigma} \approx 0$ making them highly attractive for electronic switches and logic circuits [56, 57, 58, 59, 60, 61, 62].

However, the microscopic properties of correlated materials, how they relate to global transport properties, and how correlated electronic states interact with other
materials in contact are not yet fully understood. In this thesis, we provide methods, instrumentation, and novel surface probe measurements revealing the richness and potential for new physical effects, device architectures, and technological applications of 2D heterostructures. In particular, we investigate the charge distributions at the interfaces of prototypical van der Waals heterostructures using several widely available scanning probe microscopy methods. Sub-microscopic details of the charge distributions in heterostructures of graphene encapsulated in hexagonal boron nitride are explored using electrostatic AFM modes. The charge density wave phases of $1T$–TaS$_2$ are imaged through a graphene passivation layer in ambient conditions, revealing the long-ranged ordering of topological defects of the 2D charge density wave (CDW). Finally, through scanning tunneling microscopy and spectroscopy, the interaction between Dirac electrons in graphene and the highly correlated Mott insulating CDW phase of $1T$–TaS$_2$ when in van der Waals contact. Exploring new instrumentation and scanning probe methods which can see into van der

1.2 Overview of Thesis

In Chapter 2, theoretical framework for describing electronic properties of crystalline materials is given. The basics of the tight binding (TB) model are presented using graphene as an example including notable effects of strain and periodic superpotentials on its electronic properties. The common description of charge density waves is given as well as an extension to describe CDWs in two dimensions. Finally, the various electronic phases of $1T$–TaS$_2$ are laid out and described within the theoretical framework presented.

Chapter 3 describes the experimental methods used in this thesis including the basic operating principles of scanning tunneling microscopy (STM) and atomic force microscopy (AFM) as well as their associated electronic measurement modes. Key elements of the design, construction, and noise testing of the STM instrumentation
developed for this thesis work are outlined. Finally, fabrication techniques to construct van der Waals heterostructures from exfoliated 2D materials are given.

Chapter 4 demonstrates non-invasive, electrostatic scanning probe techniques for imaging conductors and charge distributions encapsulated in an insulator. Heterostructures built from graphene encapsulated in hBN are used to show that the location of the buried graphene edge, its defects, and charge variations are observable without requiring electrical contact to the graphene layer. This value of this method is demonstrated as it is able to identify and locate defects and contamination on sub-micron length scales.

Chapter 5 discusses measurements of graphene covered 1T–TaS$_2$. It is shown that the graphene layer acts as an inert gas barrier which prevents oxidation of the 1T–TaS$_2$ surface in ambient conditions, a common problem in the material science and condensed matter physics communities. Additionally, it is shown that nanometer scale charge density wave modulations, native to 1T–TaS$_2$, are observable by STM through the graphene layer. The ordering of a lattice of topological defects within the 2D CDW state is investigated and, for the first time, is experimentally related to the lattices of topological vortices observed in type II superconducting films. STM and STS measurements of the bare and graphene covered 1T–TaS$_2$ surfaces are compared with each other and with calculations based on theory to identify the effect of proximity between the two materials. The modification of the Mott insulating state in 1T–TaS$_2$ by carriers in graphene and a proximity induced CDW in graphene are identified.

In Chapter 6, the work is summarized and an outlook on future experiments and directions that might follow this work is given.
Chapter 2
Theory and Background

2.1 Tight Binding Model

The tight binding model (TB) is a relatively simple model which can be used to capture the basic physics of electrons in crystalline materials. In the tight binding model, electrons are assumed to be localized within the strong potential of the (non-vibrating) ionic lattice and are thus separated from adjacent lattice sites by a potential energy barrier. We can then assume that the wavefunction of an electron at one lattice site has little weight at any lattice site far away but has some probability of tunneling into neighboring lattice sites.

Taking a Bloch wavefunction basis,

$$|\psi_k(r)\rangle = \frac{1}{\sqrt{N}} \sum_R e^{i\mathbf{k} \cdot \mathbf{R}} \phi(r - \mathbf{R})$$ (2.1)

we can write the resulting dispersion relationship as

$$E(k) = \langle \psi_k | \hat{H} | \psi_k \rangle = \epsilon_0 + \sum_\delta e^{ik\delta} \int \phi^*(\mathbf{r}) \hat{H} \phi(\mathbf{r} - \delta) d\mathbf{r}$$ (2.2)

where $\epsilon_0 = \int \phi^*(\mathbf{r}) \hat{H} \phi(\mathbf{r}) d\mathbf{r}$ is the onsite energy and the sum is over nearest neighbor vectors $\delta$. Defining the overlap integral:

$$t \equiv - \int \phi^*(\mathbf{r}) \hat{H} \phi(\mathbf{r} - \delta) d\mathbf{r},$$ (2.3)
we have

\[ E(k) = \epsilon_0 - t \sum_\delta e^{ik\cdot\delta}. \]  

(2.4)

The resulting dispersion describes the single-electron properties of a material given lattice parameters and using the value of \( t \) as a fitting parameter [63]. This highly simplified model does a surprisingly good job of approximating the electronic properties of materials. Further, the TB model provides a framework to describe the effects of strain and doping and is often a starting point for more detailed theories.

For those interested in performing TB calculations themselves, there are several available tight binding model solvers with detailed descriptions and tutorials (see PyBinding[64] or PythTB[65] for python based examples).

2.1.1 Graphene

The simple chemical composition of graphene makes it a perfect starting point for understanding the nature of electrons in condensed matter systems. The carbon atoms in graphene are arranged in a honeycomb pattern which can be written as a triangular lattice with spacing \( \alpha = 2.46\text{Å} \) composed of a basis of two atoms labeled "A" and "B" separated by a carbon-carbon bond length \( a_{cc} = 1.24\text{Å} \). Each unit cell can be related to any other by an integer number of translations along one or both of two fundamental lattice vectors, \( \vec{a}_1 \) and \( \vec{a}_2 \).

\[ \vec{a}_1 = \frac{a}{2}(1, \sqrt{3}), \vec{a}_2 = \frac{a}{2}(1, -\sqrt{3}) \]  

(2.5)

In the TB model, we assume that electrons are "tightly bound" to the atomic positions but are able to "hop" (tunnel) from one site with kinetic hopping energy, \( t \) (\( t \approx 2.7\text{eV} \) for graphene).
The form factor which describes the scattering amplitude due to the atomic centers in momentum space, including only the three nearest neighboring atoms of one sublattice site, is given by

\[
    f(\vec{k}) = \sum_{j=1}^{3} \exp i\vec{k} \cdot \delta_j = \exp i\sqrt{3}a_1 \cdot k_y + \exp i\sqrt{3}a_2 \cdot k_y \cdot 2 \cos \left( \frac{ak_x}{2} \right).
\]

(2.6)

Assuming the A and B sublattice states are orthogonal (no overlap), the electronic Hamiltonian can be constructed as a matrix with diagonal components corresponding to onsite hopping potentials and off-diagonal components corresponding to hopping from one sublattice to the other.

\[
    H = \begin{pmatrix}
        H_{AA} & H_{AB} \\
        H_{BA} & H_{BB}
    \end{pmatrix} = \begin{pmatrix}
        \epsilon_0 & t \cdot f(\vec{k}) \\
        t \cdot f^*(\vec{k}) & \epsilon_0
    \end{pmatrix}
\]

(2.7)

Without loss of generality, we take the on-site potential to be zero, \(\epsilon_0 = 0\), and solve the characteristic equation for eigenenergies, \(E\),

\[
    \det |H - E| = E^2 - t^2 \cdot |f(\vec{k})|^2 = 0
\]

(2.8)
We find,

\[ E = \pm t \cdot |f(k)| = \pm t \sqrt{1 + 4 \cos \left( \frac{\sqrt{3}}{2} a k_y \right) \cos \left( \frac{1}{2} a k_x \right) + 4 \cos^2 \left( \frac{1}{2} a k_x \right)}. \]  

(2.9)

Figure 2.2: **Graphene Band Structure** a) Tight binding band structure of graphene considering nearest and next-nearest neighboring atoms. Positive energy states are shown in red while negative energy states are displayed in blue. The band structure crosses zero energy at six points within the first Brillouin zone. b) Low energy band structure of graphene near the K and K' points showing a linear "light-like" dispersion c) A contour map of electron energy states of graphene within the first Brillouin zone (highlighted with a dashed line). Arrows mark the path of the linecut displayed in (d) d) A linecut of the energy dispersion (a) along high symmetry directions within the first Brillouin zone e) Energy dependent density of states obtained from the energy dispersion given in (a) shows a V-shaped low energy density of states with diverging peaks at the energies of van Hove singularities.

The electronic energy spectrum is plotted along high symmetry directions in the Brillouin zone in Figure 2.2d. We can see that the low energy electronic states are
centered around the non-equivalent corners of the Brillouin zone, denoted K and K'.

Then, in the limit that $ka \ll 1$ (low energy/long wavelength limit), we can expand the energy dispersion around these points, $\vec{K} = (\frac{4\pi}{3a}, 0)$ and $\vec{K}' = (-\frac{4\pi}{3a}, 0)$. Near these points, the structure factor becomes

$$f(\vec{k}') \approx -a \frac{\sqrt{3}}{2} (p_x - ip_y)$$

(2.10)

where $\vec{p} = \begin{pmatrix} p_x \\ p_y \end{pmatrix}$ is the momentum vector relative to the point K. Near this point (and also near K'), the energy relation is then

$$E \approx \pm ta \frac{\sqrt{3}}{2} \cdot |p| = \pm \hbar v_F |p|$$

(2.11)

where $v_F \equiv \frac{\sqrt{3} ta}{\hbar} \approx 1.1 \times 10^6 m/s$ defines the Fermi velocity of the low energy particles. The resulting linear dispersion is plotted in Figure 2.2b.

For a given energy dispersion in two dimensions, the density of electron states is $g(E) = \frac{dN}{dE}$ where $N = \frac{4k^2}{(2\pi)^2}$ is the total number of states within a circle in k-space of radius $k$. Or $g(E) = \int_{BZ} \frac{d^2 k}{(2\pi)^2} \delta \left( E - E(\vec{k}) \right) = \frac{L^2}{(2\pi)^2} \int_{S(E)} \frac{dS}{|\nabla_k E(\vec{k})|} \int_{BZ} \frac{d^2 k}{(2\pi)^2}$ where $S$ is a constant energy circle in k-space and $L$ is the sample size. The density of states for graphene is plotted in figure 2.2e. The linear low energy dispersion results in a linear density of states, however, at points where the energy dispersion becomes flat (i.e. $|\nabla_k E(\vec{k})| = 0$), the density of states diverges logarithmically [66].

At low energies, the independence of the Fermi velocity and energy is analogous to that of the dispersion of light. Taking these particles as analogs to photons, we can see that their motion is governed by the relativistic wave equation known as the Dirac equation (rather than the typical Schrodinger’s equation) [67]

$$-i\hbar v_F \begin{pmatrix} 0 & \partial_x - i\partial_y \\ \partial_x + i\partial_y & 0 \end{pmatrix} \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix} = E \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix}$$

(2.12)
Or, more briefly:

\[ v_F \mathbf{\sigma} \cdot \mathbf{p} \Psi = E \Psi \] (2.13)

Where \( \mathbf{p} = -i\hbar(\partial_x, \partial_y) \) is the momentum operator, \( \mathbf{\sigma} = (\sigma_x, \sigma_y) \) is the vector of Pauli matrices, \( \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \) and \( \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \). The components \( \Psi_A \) and \( \Psi_B \) are the sublattice components of the wavefunction \( \Psi_K = \Psi_A(\vec{r})e^{i\mathbf{K} \cdot \vec{r}} + \Psi_B(\vec{r})e^{i\mathbf{K'} \cdot \vec{r}} \) near the K-point. The valley at the opposite corner, where \( p_x \rightarrow -p_x \), also contains an independent set of states with amplitudes \( \Psi'_A \) and \( \Psi'_B \) so that the complete wavefunction can be represented as a bispinor

\[
\Psi = \begin{pmatrix} \Psi_A \\ \Psi_B \\ \Psi'_A \\ \Psi'_B \end{pmatrix}.
\] (2.14)

From equation 2.13, we can define a chirality operator, \( \hat{\sigma} = \frac{\mathbf{\sigma} \cdot \mathbf{p}}{|\mathbf{p}|} \), which commutes with the Hamiltonian and is thus, a conserved quantity. This conservation of chirality has profound consequences on the electronic properties of this material.

Finally, we identify the symmetries of this Hamiltonian. One simple symmetry which is obvious from the lattice structure (considering nearest-neighbors) is rotations by \( 2\pi/3 \). The \( C_3 \) symmetry operation should then preserve the electronic wavefunction with

\[
C_3 = \begin{pmatrix}
\omega^* & 0 & 0 & 0 \\
0 & \omega & 0 & 0 \\
0 & 0 & \omega^* & 0 \\
0 & 0 & 0 & \omega
\end{pmatrix}
\] (2.15)

where \( \omega = \exp(2\pi i/3) \). When acted on the Hamiltonian, such rotation simply moves the K and K’ points to another, equivalent corner, unchanging the form of the energy
Another symmetry of the underlying lattice is spatial inversion symmetry, \( \{x, y\} \rightarrow \{-x, -y\} \). Written as an operator:

\[
P = \begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\] (2.16)

Acting on the wavefunction, spatial inversion reverses the sign of momentum (\( \Psi_i \rightarrow \Psi'_i \)) and also exchanges the graphene sublattices (\( A \rightarrow B \)).

Next, we consider the operation of time reversal: \( t \rightarrow -t \). The time reversed wavefunction is

\[
T \Psi = T \left[ \Psi_A(\vec{r}) e^{i \vec{K} \cdot \vec{r}} + \Psi_B(\vec{r}) e^{i \vec{K}' \cdot \vec{r}} + \Psi'_A(\vec{r}) e^{-i \vec{K} \cdot \vec{r}} + \Psi'_B(\vec{r}) e^{-i \vec{K}' \cdot \vec{r}} \right] \\
= \Psi^*_A(\vec{r}) e^{-i \vec{K} \cdot \vec{r}} + \Psi^*_B(\vec{r}) e^{-i \vec{K}' \cdot \vec{r}} + \Psi'^*_A(\vec{r}) e^{i \vec{K} \cdot \vec{r}} + \Psi'^*_B(\vec{r}) e^{i \vec{K}' \cdot \vec{r}}.
\] (2.17)

We can see that time reversal has the effect of interchanging the valleys (\( K \) and \( K' \)). Then, the time reversal operator can be written as

\[
T = \begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix} C
\] (2.18)

where \( C \) is the complex conjugation operator. It is then easy to see that the derived Hamiltonian is unchanged by time reversal: \( THT^{-1} = H \). As we should expect from conservation of energy.

Note that both spatial inversion and time reversal change the sign of the wavevector. Thus, the combined symmetry operation \( PT \) preserves the Fermi points. This
subtle result has strong implications on the topological stability of these points in graphene: perturbations which preserve this symmetry will not open a gap in graphene \[68, 69\].

### 2.1.2 Strain and Pseudomagnetic Fields in Graphene

Aside from having extremely interesting intrinsic electronic properties, graphene boasts to be the stretchiest crystal in existence, capable of withstanding strain up to 25 percent. Additionally, it has been theoretically proposed that at very large strain values, graphene may undergo a metal to insulator phase transition, obtaining a band gap at the Fermi level \[70\]. When placed in contact with an atomically flat substrate, as in a 2D heterostructure configuration, lattice mismatch between graphene and neighboring atoms typically generate strain in the graphene sheet. The magnitude and direction of strain generally depends on the spacing mismatch and relative directions of the two lattices. To deduce the effects of strain on the electronic properties of graphene, we extend the tight binding model derived in the previous section.

Consider a perturbed graphene lattice (indexed by position \(\vec{r}_i\)) whose difference from the unperturbed lattice \((\vec{r}_i^0)\) is given by the displacement field \(\vec{u}(\vec{r}) = (u_x, u_y, z)\) such that \(\vec{r}_i = \vec{r}_i^0 + \vec{u}\). The strain field can be related to the classical strain tensor as

\[
\varepsilon_{lm} = \frac{1}{2}[\partial_l u_m + \partial_m u_l + (\partial_l z)(\partial_m z)], l, m \in \{x, y\}. \tag{2.19}
\]

As strain is applied, the equilibrium nearest-neighbor bond length changes to \(d_{ij} = |\vec{r}_i - \vec{r}_j|\). Then, the energy required for an electron to hop from an atom at location \(\vec{r}_i\) to one at \(\vec{r}_j\), the hopping parameters, \(t_{ij} = t(\vec{r}_i, \vec{r}_j)\), become modified as

\[
t_{ij} = t_0 e^{-\beta(d_{ij}/a_0 - 1)}, \tag{2.20}
\]

where \(a_0\) and \(t_0\) are the unperturbed bond length and hopping energy, respectively,
and $\beta = \frac{\partial \log(t)}{\partial \log(a)}|_{a=a_0} \approx 3.37$ [70]. Thus, for arbitrary strain, the hopping integral will take on different values along the three nearest neighbor bonds; however, for uniform strain, each atom will see the same three hopping parameter values. We label them as $t_1, t_2,$ and $t_3$. Now, we can write the dispersion relation of graphene under uniform strain as

$$E = \pm \sqrt{(t_2 + t_3 \cos(\vec{k} \cdot \vec{a}_1) + t_1 \cos(\vec{k} \cdot \vec{a}_2))^2 + (t_3 \sin(\vec{k} \cdot \vec{a}_1) + t_1 \sin(\vec{k} \cdot \vec{a}_2))^2}$$

(2.21)

(Compare to equation 2.9, the un-strained case) Notice that not only do the $t_{ij}$ change with strain, but also the primitive lattice vectors, $\vec{a}_1$ and $\vec{a}_2$. The new lattice vectors are related to the unperturbed lattice vectors through the strain tensor.

For simplicity, we start by considering uniaxial strain along some angle $\theta$ with respect to the graphene crystal lattice. The strain tensor can then be written as

$$\varepsilon = \varepsilon \begin{pmatrix} \cos^2 \theta - \sigma \sin^2 \theta & (1 + \sigma) \cos \theta \sin \theta \\ (1 + \sigma) \cos \theta \sin \theta & \sin^2 \theta - \sigma \cos^2 \theta \end{pmatrix}$$

(2.22)

where $\sigma = 0.165$ is the measured Poisson’s ratio for graphite. Then the unperturbed primitive lattice vectors, $\vec{a}_{1,2}^0$, move (to leading order in strain) to

$$\vec{a}_{1,2} = (1 + \varepsilon) \cdot \vec{a}_{1,2}^0.$$

(2.23)

As the atomic lattice sites distort with strain, so does the reciprocal lattice. The Dirac point at $\vec{K} = \left( \frac{4\pi}{3\sqrt{3}}, 0 \right)$ shifts to the new position in reciprocal space

$$\vec{K} \approx \frac{4\pi}{3\sqrt{3}} \left( 1 - \frac{\varepsilon(1 - \sigma)}{2}, -\varepsilon(1 + \sigma) \sin(2\theta) \right).$$

(2.24)

We can visualize this effect by plotting the energy dispersion for several strain values along different directions. Figure 2.4 displays the energy dispersion for the cases where there is no strain ($\varepsilon = 0$), strain along the armchair direction (with $\varepsilon = 0.2$), and strain
along the zigzag direction ($\varepsilon = 0.2$). When strain is applied along the zigzag direction, the Dirac points move closer together (see fig. 2.4d), eventually merging together and opening a gap at a critical value, $\varepsilon \approx 0.23$ (on the order of 20%). Thus, it has been proposed that highly strained graphene might be used in technological applications such as strain sensors and switches. Such large strain values are difficult to obtain (and maintain) across large areas of graphene; however, large strain values may be obtained locally, leading to a suppression of conductivity through these regions.

Uniaxial strain in graphene breaks the intrinsic $C_3$ symmetry by weakening bonds
in one direction while strengthening them in another. Tension on the graphene lattice can then lead to anisotropic transport properties. From figure 2.4d, we can see that the slope of the low energy, linear dispersion (related to the Fermi velocity) varies with strain. Additionally, the shape of the Fermi surface becomes elliptical, having different Fermi velocities along different directions. As a result, the transport properties of strained graphene have been shown to be different along directions parallel and perpendicular to the applied strain, consistent with the model presented here [71].

Early considerations of electron-phonon coupling in graphene and carbon nanotubes introduced an effective electronic potential induced by the motion of the atomic lattice. It can be shown that the effect of strain on the electronic Hamiltonian is captured by introducing both a scalar and a vector potential. The cases of in-plane and out-of-plane deformations are typically treated separately in the literature. In the following discussion, we will consider only in-plane strain and leave it to the reader to repeat the analysis using the scalar and vector potentials due to deformations in the z-direction given in the literature [72, 73, 74]. For in-plane strain, the scalar potential is given by

\[ \Phi = g_s (\varepsilon_{xx} + \varepsilon_{yy}) \]  

(2.25)

which describes the change in local charge density due to the strain variation \(\varepsilon_{xx} + \varepsilon_{yy}\). The value of \(g_s\) is not yet known with certainty although calculations typically adopt a value near \(3eV\) [72, 73, 69, 75, 74]. The vector potential component is written as

\[ \vec{A} = -\frac{\hbar}{2e a_0} \begin{pmatrix} \varepsilon_{xx} - \varepsilon_{yy} \\ -2\varepsilon_{xy} \end{pmatrix} \]  

(2.26)

where \(a_0 = |\vec{a}_{1,2}|\) is the unperturbed bond length. The notation has been chosen suggestively, as the potentials enter the tight-binding Hamiltonian similar to the
electromagnetic scalar and vector potentials

\[ H_K = v_F \vec{\sigma} \cdot (\vec{p} + e\vec{A}) + \Phi \]  

(2.27)

with one major difference: The vector potential changes sign in the opposite valley, such that

\[ H_{K'} = -v_F \vec{\sigma} \cdot (\vec{p} - e\vec{A}) + \Phi. \]  

(2.28)

Because of this sign difference, the strain field does not break time reversal symmetry whereas a real magnetic field does.

This analogy can be taken surprisingly far and has some peculiar consequences on electronic motion in the strained graphene lattice. Continuing the analogy, we define a pseudo-magnetic field in the same manner as a real, external magnetic field

\[ B_s = \nabla \times \vec{A} = \partial_x A_y - \partial_y A_x. \]  

(2.29)

Thus, for regions where the strain field has a curl, electrons experience an effective magnetic field but with opposite directions at the K and K' points in the Brillouin zone.

It has been proposed that through strain engineering, one can utilize this sign difference to produce a device which can filter and detect electrons in different valleys. Such "valley-tronic" devices could then be used for logic circuits in a similar manner to spin-tronics [76].

A profound consequence of the vector field induced in strained graphene is that within regions of constant pseudo-magnetic fields, experimental probes observe pseudo-Landau levels; analogs of true Landau levels, the quantized energy states observed in two-dimensional systems in an external magnetic field. These discrete
energy levels follow the same sequence as real Landau levels in graphene [77, 21, 78].

\[ E_n = \text{sgn}(n)\sqrt{2e\hbar v_F^2 B_s |n|} \]  

(2.30)

where \( E_n \) is the energy of the Landau level with index \( n \). These can be observed as sharp peaks in the electronic density of states spaced by unoccupied, incompressible regions so long as the pseudomagnetic field is uniform across an area which is larger than the pseudomagnetic field length \( l_B = \sqrt{\hbar/eB_s} \).

Figure 2.5: Pseudo-Landau Levels in Triaxially Strained Graphene

a) Diagram of a triaxially strained graphene. Arrows designate directions of the applied force  
b) Local density of states at the center of the strained graphene (a) shows discrete peaks associated with the pseudo-Landau levels  
c) The energy of the peaks in (b) plotted against \( \text{sgn}(n)\sqrt{|n|} \) gives a linear relationship, just as in Landau levels for 2D Dirac electrons in a magnetic field (Data generated using [64])

Finally, we demonstrate that this can be captured in the tight binding model by considering a simple strain profile for which \( \nabla \times \vec{A} \neq 0 \); triaxial strain. Consider a
non-uniform strain field of the form \( \vec{u}(\vec{r}) = (2cxy, c(x^2-y^2), 0) \) where \( x \) and \( y \) are the Cartesian positions of lattice points with respect to the center of the graphene flake and \( c \) is a constant. This strain profile is that obtained from stretching a graphene flake by pulling equally along the nearest neighbor bond directions; as shown in figure 2.5a. Using equations 2.19, 2.26, and 2.29, it is easy to see that such a strain profile leads to a constant and uniform pseudo-magnetic field. The electronic density of states is calculated using the tight binding model and displayed in 2.5b. We can see sharp peaks corresponding to the energy positions of pseudo-Landau levels. Fitting the sequence using equation 2.30 gives a measurement of the pseudo-magnetic field, \( B_s = 129.7T \) for \( c = 0.0025 \) (figure 2.5c).
2.1.3 Moiré Patterns and Superpotentials

Figure 2.6: Moiré Patterns in Twisted Atomic Layers a) Two overlapping square lattices (top) and honeycomb lattices (bottom) with different inter-layer twist angles b) A plot of the wavelength of the Moiré pattern versus inter-layer rotation angle for three different values of $\delta$. The chosen values of $\delta$ correspond to the lattice mismatch between graphene and graphene ($\delta = 0$), hBN ($\delta = 0.032$), 1T-TaS$_2$ ($\delta = 0.280$)

When two lattices are placed on top of one another with arbitrary angle between them, a larger periodic structure is generated due to interference of the two underlying (smaller-period) lattices. This can be simply visualized in real space by placing two misaligned dot arrays on top of each other. Examples of Moiré interference between
two identical square lattices and honeycomb lattices are given in figure 2.6a with twist angles of 5 and 12 degrees.

We can see that there becomes a long-wavelength periodic modulation of lattice point density. This superstructure oscillates in space between overlapping and misaligned dot regions. The observed Moire super-period is related to the mismatch of the lattice spacing, $\delta$, and the angle between the two lattices, $\theta$. From figure 2.6a, the Moiré superperiod decreases in size as the angle between the two lattices increases. The Moiré wavelength can be written as [50]:

$$\lambda = \frac{(1 + \delta)a}{\sqrt{2(1 + \delta)(1 - \cos \theta) + \delta^2}}. \quad (2.31)$$

Moire patterns were observed by STM shortly after its invention in decoupled layers at the surface of graphite [79]. They were initially recognized as a simple mathematical oddity as a result of two overlapping lattices. It was realized after the isolation and measurement of graphene that Moire patterns can have a strong impact on the electronic and vibrational properties of two dimensional materials.

As an example, we first consider two misaligned graphene layers. Since the lattice spacing of both layers is the same, $\delta = 0$. Lattice sites are simply related to each other through a rotation about some point in space by the angle $\theta$. Similarly, reciprocal lattice points associated with each layer can be related to the other layer. The Dirac cones located at the K-points of each layer is separated from the other in reciprocal space as

$$\Delta K = 2K \sin(\theta/2) \quad (2.32)$$

with $K = 4\pi/3a$ [80]. As the relative angle between the two lattices decreases, the Dirac cones intersect with each other. Due to interlayer coupling, the cones hybridize with each other, reducing the electronic energy near their intersection in k-space. Hybridization leaves a saddle point in the band structure between the two K-points (and also the K'-points), leading to a van Hove singularity (vHS) in the electronic
density of states. The energy separation between two adjacent vHSs is simply related to the Dirac cone separation and the interlayer hopping energy, $t_{\perp}$, by

$$\Delta E_{vHS} = \hbar v_F \Delta K - 2t_{\perp}. \quad (2.33)$$

From this example, we can see that the twist angle between two 2D crystals becomes another knob which allows us to tune the electronic properties of the crystals. At large angles, the electrons behave as in two decoupled graphene monolayers with linear dispersion at low energy and the K-points are separated in k-space [81].

As the van Hove singularities move towards each other in energy (as twist angle decreases) the Dirac cone becomes compressed. Thus, the hybridization reduces the Fermi velocity near the K-points. To second order in a continuum model, the renormalization of the Fermi velocity is given by [80]

$$\frac{v_F(\theta)}{v_F^0} = 1 - 9 \left( \frac{t_{\perp}}{\hbar v_F^0 \Delta K} \right)^2. \quad (2.34)$$

Equation 2.34 depends on the superpotential period (through $\Delta K$) and the coupling strength between adjacent layers. It can be shown that any general periodic potential coupled to graphene has a similar effect on the Fermi velocity of the graphene quasiparticles [82, 83, 84].

There has been a huge surge in research on the effects of Moiré patterns and superpotentials in graphene (especially between two layers of graphene) which has led to intriguing observations including correlated electron states and has rapidly changed the way we think about and model interactions in materials. The most basic principles have been laid out here, however, this simple topic has proven to be very rich with physics and technological promise.
2.2 Charge Density Waves

A charge density wave (CDW) is a periodic distortion in charge distribution accompanying a periodic distortion of the underlying lattice which arises from coupling between electrons and the motion of the atomic lattice. Together, these open a gap in the electronic spectrum, reducing the total energy of the system to generate a new ground state. In order to discuss the CDW in one dimension, we must first model the unperturbed 1D atomic chain using a simple tight binding model.

Consider a periodic string of atoms with interatomic spacing, a. An arbitrary lattice site will then have two nearest neighbors related by the vectors $\delta_{\pm} = \pm a \hat{\imath}$. From equation 2.4, the electronic dispersion is given by

$$E(k) = \epsilon_0 - t \sum_{\delta} e^{i k \cdot \delta} = \epsilon_0 - t (e^{ika} + e^{-ika}) = \epsilon_0 - 2t \cos(ka) \quad (2.35)$$

The energy dispersion is plotted in figure 2.7b for $\epsilon_0 = 0$ and $t = 1eV$. We can see that despite assuming the electrons in the TB lattice to be highly localized in space, any arbitrary amount of overlap, t, causes the electron energy states to spread into a continuous band. We will see how this tight binding model for a 1D metallic chain and its resulting dispersion is modified by correlation effects.

In the CDW state, the charge density is modulated such that:

$$\rho(x) = \rho_0(x)[1 + A \cos(q_0x + \phi)], \quad (2.36)$$

where $\rho_0(x)$ is the unperturbed electron density, A, $q_0$, and $\phi$ are the modulation amplitude, wavenumber, and phase, respectively. With electron-phonon coupling, the charge modulation causes the ions in the lattice to see different potentials, inducing a periodic lattice distortion (PLD) of the form:

$$u_n = u_0 \sin(nq_0a + \phi), \quad (2.37)$$
where \( n \) denotes the lattice site and \( u_0 \) is the distortion amplitude. It is easy to see that a periodic lattice distortion as described by 2.37 will also lead to a charge density modulation as described by 2.36. The coupled emergence of the charge density wave and periodic lattice distortion makes the driving mechanism for CDW formation in real systems ambiguous. Additionally, there are many proposed mechanisms and factors which might drive a system into such a phase.
Figure 2.7: **1D Charge Density Wave** a) A cartoon depiction of a uniform 1D metallic chain (left panel) and a 1D charge density wave state (right panel) which has a modulated charge density along with a periodic lattice displacement of 1D atomic chain (figure modelled after [85]). b) The electronic energy dispersion for the 1D metallic chain c) Single-particle electron energy dispersion for the 1D CDW system exhibiting a gap induced at the Fermi level d) The real part of the static Lindhard susceptibility function in 1D, 2D, and 3D shows that the 1D metal is highly susceptible to the CDW modulation given by a strong divergence for the 1D case.

Following the work of Peierls, I show that under certain conditions, the charge density wave phase is stable and energetically favorable over the undistorted phase. For simplicity, we restrict the model to a one-dimensional atomic chain with electronic energy dispersion, $E(k)$. The Hamiltonian describing the atomic lattice, the
electronic energy spectrum, and the coupling between the two is given by the 1D Fröhlich Hamiltonian (in second quantized notation):

\[ H = \sum_k \epsilon_k c_k^\dagger c_k + \sum_q \hbar \omega_q b_q^\dagger b_q + \sum_{k,q} g(q) c_{k+q}^\dagger c_k (b_q + b_{-q}^\dagger), \]  

(2.38)

where \( \epsilon_k \) is the electron kinetic energy with wavevector \( k \), \( \hbar \omega_q \) is the energy of phonon with wavevector \( q \), and \( c_k^\dagger \) (\( c_k \)) and \( b_q^\dagger \) (\( b_q \)) are the electron and phonon creation (annihilation) operators. Then, a static ion displacement \( u_q \),

\[ \sqrt{\frac{\hbar}{2M\omega_q}} \left\langle b_q + b_{-q}^\dagger \right\rangle = u_q, \]  

(2.39)

produces a potential

\[ v_q = g(q) u_q \sqrt{\frac{2M\omega_q}{\hbar}}. \]  

(2.40)

This is the electronic potential produced by the lattice distortion. If the new lattice potential and the modified electronic energy reduce the total energy of the system, then the CDW phase is energetically stable.

The change in lattice energy is then

\[ \delta E_{\text{lattice}} = \frac{1}{2} M \omega_q^2 u_q^2. \]  

(2.41)

Treating the potential \( v_q \) as a perturbation to the electronic spectrum we can calculate the second order shift of the band energy

\[ \delta E_{\text{band}} = -|v_q|^2 \chi_0(q), \]  

(2.42)

where \( \chi_0(q) \) is known as the static Lindhard susceptibility function

\[ \chi_0(q) = \frac{1}{L} \sum_k \frac{f_{k+q} - f_k}{\epsilon_k - \epsilon_{k+q}}. \]  

(2.43)
Here, \( f(\epsilon) \) is the Fermi function and \( L \) is the length of our 1D chain. We can now see that the CDW is energetically favorable for

\[
\delta E_{\text{lattice}} + \delta E_{\text{band}} < 0
\]  

or

\[
\chi_0(q) > \frac{\hbar \omega_q}{4g_q^2}
\]  

Thus, for large electron-phonon coupling and for large electronic susceptibility, the distorted CDW phase is energetically favorable. Note that in 1973, Chan and Heine modify this criteria to account for the effects of Coulomb repulsion and exchange interactions, with energies \( U_q \) and \( V_q \), respectively. The authors find that the CDW phase is stable for

\[
\frac{4g_q^2}{\hbar \omega_q} - 2U_q + V_q \geq \frac{1}{\chi_0(q)}
\]  

with

\[
\frac{4g_q^2}{\hbar \omega_q} > 2U_q > V_q > 0.
\]

The Lindhard susceptibility function, \( \chi_0 \), has two important features to note. First, is the dependence of electronic susceptibility on wave vector, \( q \). In the particular case of our previously described one dimensional metallic chain,

\[
\chi_0(q) \propto \frac{1}{q^2} \left[ 1 + \frac{k_F}{q} \left( 1 - \frac{q^2}{4k_F^2} \right) \ln \left| \frac{q + 2k_F}{q - 2k_F} \right| \right].
\]  

The electronic susceptibility diverges at \( q = 2k_F \):

\[
\chi_0(q) \approx \frac{1}{2k_F} \left[ 1 + \frac{1}{2k_F} (q - 2k_F) \ln |q - 2k_F| \right] \approx \infty.
\]  

Thus the 1D CDW state is stable for arbitrary values of the e-ph coupling constant \( g(q \approx 2k_F) \) (fig. 2.7d) [86, 87, 88]. In real systems, however, small deviations from the ideal case, such as temperature fluctuations, momentum fluctuations, and coupling
into other dimensions (say, to adjacent 1D metallic chains) kill the divergence of this function and reduce the infinite peak to a finite value. However, if there exists large parallel portions of the Fermi surface, i.e. if there are many k such that \( \epsilon_k = \epsilon_{k+2k_F} \), then the Lindhard susceptibility function as given by Eqn. 2.43 is large and the CDW formation is stable again. This concept is known as Fermi Surface nesting [89, 90, 91].

The second feature of \( \chi_0 \) to note is its temperature dependence. Under a mean-field approximation (still assuming a 1D tight-binding metallic chain), the electronic susceptibility diverges with decreasing temperature

\[
\chi_0(2k_F, T) = \frac{1}{2} N(0) \ln \left( \frac{2.28 E_F}{k_B T} \right) \tag{2.50}
\]

Thus, the screening ability of electrons at the wavevector \( 2k_F \) becomes very large.

It was noticed by Kohn that for large electronic susceptibility and electron-phonon coupling, the electrons are able to screen the motion of the atomic lattice, reducing the frequency of vibrational modes:

\[
\tilde{\omega}_q^2 = \omega_q^2 (1 - \frac{4g_q^2}{\hbar \omega_q \chi_0(q)}) \tag{2.51}
\]

Where \( \omega_q \) and \( \tilde{\omega}_q \) are the unscreened and renormalized phonon frequencies, respectively. Combined with the divergence of \( \chi_0(q) \) with decreasing temperature, we can expect a sharp reduction of the phonon frequency near \( 2k_F \) and for \( \frac{4g_q^2}{\hbar \omega_q \chi_0(q)} \geq 1 \), the frequency can reach zero or imaginary values, becoming a static PLD. With 2.50, we get a mean-field transition temperature of

\[
k_B T_0 = 2.28 \exp(-1/\lambda) \tag{2.52}
\]

where \( \lambda \) is a dimensionless coupling constant \( \lambda = \frac{2g_0^2 N(0)}{\hbar \omega_q} \).

Now we turn to the effect of CDW formation on the electronic spectrum of the material. When the electronic energy discontinuity is at \( k = \pm k_F \), the Fermi level, it
becomes favorable for the electron energy band to open a gap at the Fermi level with energy $2\Delta$ and split into two bands (see fig. 2.7c):

$$E_{\pm}(k) = \frac{\epsilon_k + \epsilon_{k+q_0}}{2} \pm \sqrt{\left(\frac{\epsilon_k - \epsilon_{k+q_0}}{2}\right)^2 + \Delta^2}$$  \hspace{1cm} (2.53)

with the electronic energy gap related to the electron-phonon coupling constant and the average static ion displacement:

$$\Delta = u_{q_0} g_{q_0} \sqrt{\frac{2M\omega_{q_0}}{\hbar}}.$$  \hspace{1cm} (2.54)

Although the single particle spectrum obtains a gap, the energy of the collective charge density wave is independent of its phase and is thus capable of conducting. We can take the CDW phase as a classical field and in the presence of an applied electric field $E(\omega) = E_0 e^{i\omega t}$, we expect the equation of motion

$$\frac{d^2\phi}{dt^2} + \frac{m}{m^*} \hbar v_F^2 \frac{d^2\phi}{dx^2} = \frac{2k_F e E(\omega)}{m^*}$$  \hspace{1cm} (2.55)

and the frequency dependent conductivity

$$\sigma(\omega) = \frac{j(\omega)}{E(\omega)} = \frac{m}{m^* 4\pi (\omega + i\delta(\omega))},$$  \hspace{1cm} (2.56)

where $\omega_p = 8v_F e^2$ is the plasma frequency. However, it has been shown by [89, 90] that unlike the BCS superconducting wavefunction, the CDW can be pinned to impurities. The effect of pinning is to shift the collective mode conductivity to non-zero frequencies. We can write down a phenomenological equation of motion including the effects of pinning by introducing an average pinning frequency $\omega_0$ and a damping time (due to internal effects of the collective mode or inhomogeneity of
pinning centers in space), $\tau$. We have,

$$\frac{d^2\phi}{dt^2} + \frac{1}{\tau} \frac{d\phi}{dt} + \omega_0^2 \phi = \frac{2k_F e E(\omega)}{m^*}$$

(2.57)

and

$$\sigma(\omega) = \frac{n_e e^2}{\omega m^*} \frac{\omega^2}{\omega_0^2 - \omega^2 - i\omega/\tau}.$$  

(2.58)

In this case, an applied electric field needs to be large enough to first overcome the depinning of the CDW from defects before it is able to slide.

Near the Peierls transition temperature, the CDW properties can be described by a simple Ginzburg-Landau theory. We perform a low order approximation of the free energy of the CDW system by expanding in even powers and neglecting fluctuations of the order parameter

$$F = F_0 + f \int dx \left( \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 \right),$$

(2.59)

where $\psi = \rho(x) - \rho_0(x) = |\psi|e^{i\phi}$ is the order parameter of this model, $\alpha$ and $\beta$ are positive constants, and near the mean field transition temperature, $T_0$ (from equation 2.52), $t \approx \frac{T-T_0}{T_0}$. Minimizing the free energy density with respect to the order parameter, we find that the order parameter has the temperature dependence given by

$$\psi = \begin{cases} 
0 & t > 0 \\
\pm \sqrt{\frac{-\alpha t}{\beta}} & t \leq 0 
\end{cases}.$$  

(2.60)

We can see that the CDW amplitude decreases continuously before disappearing at the CDW-normal state transition. Thus, this transition should be of second order in temperature.

The arguments presented here introduce the idea of a charge density wave and
suggest how one might be a natural result of arbitrarily small electron-phonon coupling in one dimension. Next we see that both dimensionality and strong momentum-dependent electron-phonon coupling lead to deviations from the 1D Peierls-like CDW behavior. The next section is devoted to extending the arguments presented in this section into two dimensions in order to describe the CDWs observed in 2D materials.

2.2.1 CDWs in Two Dimensions

In the 1970s, the topic of charge density waves promoted great interest in the transition metal dichalcogenides [92]. Transition metal dichalcogenides are van der Waals stacked materials, like graphene, and can similarly be mechanically exfoliated and stacked into various heterostructures. This provides a new platform for studying these materials and their interaction with a substrate. Being van der Waals bonded out of their transition metal planes, these materials often exhibit highly anisotropic transport with nearly two dimensional electronic behavior. Many TMDs exhibit strongly correlated behavior and nearly all of these materials have charge density waves, however, their properties tend to deviate from Peierls’ 1D CDW picture.

The Peierls transition described in section 4.1 should be thought of as a perturbation of the undistorted lattice and charge density. As we "turn on" electron-phonon coupling, the CDW and PLD begin to emerge. The typical Peierls picture, then, describes the "weak coupling regime," where we can assume a weakly perturbed lattice and charge density. This description captures many of the main features of CDWs including the low-energy excitations and the normal-incommensurate and incommensurate-commensurate phase transitions. However, the 2D TMDs tend to deviate from the Peierls picture due to intrinsically strong electron-phonon coupling as well as some degree of Fermi surface nesting.

Compared to the weakly coupled case, a strongly coupled CDW (also called a Type II CDW [93]) exhibits a larger lattice distortion amplitude, a larger charge density modulation, larger energy gaps, and smaller coherence lengths. Above the CDW
transition temperature, long-range coherence is lost but short-range fluctuations remain, causing strong carrier scattering. As a result, CDW-like charge modulations are often observed near defects in CDW materials even in the metallic state. For the strongly coupled CDW, the incommensurate-commensurate transition is driven by lattice entropy rather than energy competition and the transition temperature is given (without derivation) by

\[ k_B T_0 = 0.296 C \frac{\left( \pi \xi_0 u_0 \right)^2}{2 \Omega} \] (2.61)

where \( \xi_0 \) is the CDW coherence length, \( u_0 \) is the periodic lattice distortion amplitude, \( \Omega \) is the normal state unit cell area, and \( C \) is an electronic energy scale (in units of energy per area). While weakly coupled CDW transitions were found to be of second order, the strongly coupled case gives a first order commensurate-incommensurate transition and a second order incommensurate-normal metal transition [94].

Figure 2.8: 2D CDW Components Top row: The 2D CDW can be broken up into three one-dimensional components and the atomic lattice. Bottom: The sum of the charge density modulation due to the lattice and the three 2D CDW components exemplifies the construction of a 3 by 3 CDW in two dimensions from 1D components.
Similar to the one dimensional case, the important features of charge density waves in two dimensions and their phase transitions are well-captured with a phenomenological Landau theory. In reference [95], McMillan shows that the 1D CDW picture can be extended to two dimensions by considering three coexisting 1D CDWs with reciprocal vectors that are simply related to each other by a symmetry of the fundamental lattice. To describe a two dimensional CDW, we extend 2.36 to account for all three CDW components:

\[
\rho(x) = \rho_0(x)[1 + \alpha(\vec{r})],
\]

(2.62)

where \(\alpha(\vec{r})\) is the real order parameter composed from the three complex CDW order parameters \(\psi_i(\vec{r}) = A_i \exp(iQ^{(i)} \cdot r + \phi_i)\):

\[
\alpha(\vec{r}) = \text{Re}[\psi_1(\vec{r}) + \psi_2(\vec{r}) + \psi_3(\vec{r})].
\]

(2.63)

where \(A_i, \phi_i, \) and \(Q^{(i)}\) are the amplitude, phase, and wavevector of the ith CDW component, respectively. Figure 2.8 shows the sum of three 1D CDW components, \(\text{Re}[\psi_i]\). Together, the three components make a \(C_3\) symmetric, 2D charge density wave. A periodic modulation due to the lattice is added for demonstration assuming a 3 by 3, commensurate CDW modulation.

To model the 2D CDW phase transitions, the Landau free energy is written as an expansion in powers of the order parameter and its derivatives:

\[
F = \int d^2 r [a(\vec{r})\alpha^2 - b(\vec{r})\alpha^3 + c(\vec{r})\alpha^4
\]

\[
+ d(\vec{r})(|\psi_1(\vec{r})\psi_2(\vec{r})|^2 + |\psi_2(\vec{r})\psi_3(\vec{r})|^2 + |\psi_3\psi_1|^2)]
\]

\[
+ e(\vec{r}) \sum_i |(\vec{Q}^{(i)} \cdot \vec{\nabla} - i\vec{Q}^{(i)2})\psi_i|^2 + f(\vec{r}) \sum_i |\vec{Q}^{(i)} \times \vec{\nabla} \psi_i|^2]
\]

(2.64)

where the coefficients \(a(\vec{r}), b(\vec{r}), c(\vec{r}), \) and \(d(\vec{r})\) exhibit the periodicity of the lattice and can be written as \(b_0 + b_1 \sum_{i=1}^{6} e^{iG_i \cdot \vec{r}}, \) etc., with \(G_i\) are the six shortest reciprocal
lattice vectors. The vectors $\mathbf{Q}^{(1)}$, $\mathbf{Q}^{(2)}$, and $\mathbf{Q}^{(3)}$ are the three reciprocal lattice vectors of the incommensurate CDW. Their directions differ by 120° and the magnitudes are equal to $2\pi/\lambda_{CDW}$, with $\lambda_{CDW}$ being the wavelength of the ICCDW.

Following the usual approach, the coefficient of the first term in 2.64 much change sign at the transition temperature $T_0$, so we take $a_0 = a(T - T_0)$. The second term integrates to zero when considering the 1D CDW, however in the 2D CDW, the cubic term is minimized when the phases of the three CDW components align relative to each other and be considered a phasing energy. The fourth term (proportional to $d(\mathbf{r})$) describes another interaction between the three CDW components in 2D, stabilizing the triple CDW over the single CDW case (for a particular range of values of the ratio $\frac{d_0}{c_0}$). The final two terms (the gradient terms) describe the elastic energy cost of stretching and rotating the CDW towards a commensurate wavevector, $\mathbf{Q}^{(i)}_C$, defined by the conditions: $\mu\mathbf{Q}^{(i)}_C + \nu\mathbf{Q}^{(i+1)}_C = \mathbf{G}_i$ and $|\mathbf{Q}^{(i)}_C| = |\mathbf{G}_i|/\sqrt{\mu^2 + \mu\nu + \nu^2}$ ($\mu$ and $\nu$ are integers) [95, 96].

![Figure 2.9: 2D CDW Harmonic Expansion](image)

Figure 2.9: 2D CDW Harmonic Expansion a) Comparing an arbitrary set of CDW wavevectors $\mathbf{Q}^{(i)}$ with those of the commensurate CDW $\mathbf{Q}^{(i)}_C$ defines the discommensuration vectors $\mathbf{q}^{(i)}$. b) The wavevectors corresponding to the domain walls $\mathbf{k}^{(i)}$ are built from multiples of the vectors $\mathbf{q}^{(i)}$ (the case of $(\mu, \nu) = (3, 1)$ is shown) c) A general CDW state can be described as a superposition of components with wavevectors $\mathbf{Q}^{(i)}_{lmn}$. Figure based on [97]

In a strongly coupled CDW system, atoms tend to form pairs or clusters with shortened bonds. Due to larger atomic displacements, non-linear and higher harmonic
effects of the electron-phonon coupling interaction become increasingly important [85]. From neutron scattering experiments on several TMDs, Moncton et. al noticed diffuse scattering at wavevectors which are higher harmonics of the fundamental CDW mode. These authors then recognized that higher harmonics than those considered in McMillan’s theory may play a strong role near the commensuration transition by creating phase shifts or commensurate domains within the ICCDW state [98]. Following the suggestion of Moncton et. al, Nakanishi and Shiba have extended the analysis of McMillan and coworkers by including the effect of higher harmonics on the CDW order parameter. To consider the harmonics, it becomes beneficial to expand the order parameter near the commensurate wavevector such that $\psi_i(\vec{r})$ can be written as

$$\psi_i(\vec{r}) = \phi_i(\vec{r}) e^{i\vec{Q}_C \cdot \vec{r}}$$

(2.65)

where $\phi_i(\vec{r}) = \sum_{l,m,n \geq 0}^{l,m,n} \Delta_{lmn}^{(i)} e^{i\vec{Q}_nm \cdot \vec{r}}$, l, m, and n are integers and $\vec{Q}_{nm}^{(i)} = l\vec{k}^{(i)} + m\vec{k}^{(i+1)} + n\vec{k}^{(i+2)} + \vec{Q}^{(i)}$, with $k^{(i)} = \mu \delta q^{(i)} - \nu \delta q^{(i+1)}$, $\delta q^{(i)} = Q^{(i)} - Q^{(i)}_C$ and $i = 1, 2, 3$ are cyclically permuted. These vectors and how they relate to the commensurate CDW wavevector is displayed in figure 2.9. Note that in the figure, the 120° symmetric constraint is relaxed for generality. The harmonic amplitudes, $\Delta_{lmn}^{(i)}$, are found by solving the coupled differential equations $\frac{\partial F}{\partial \Delta_{lmn}^{(i)}} = 0$ for a CDW given by the integers $\mu$ and $\nu$. We will see in later sections that this expansion is necessary for understanding ordered CDW phases beyond the typical commensurate and incommensurate phases [95, 96, 99, 100, 101, 97].
2.2.2 Chiral CDWs

Figure 2.10: **2D CDW Registrations**

(a) Symmetric

\[ \psi_1 + \psi_2 + \psi_3 \]

The left-most image displays an atomic lattice as well as a 3x3 charge density modulation associated with the 2D CDW state. The middle image reproduces the charge density modulation without including the atomic lattice. The right-most image displays the Fast-Fourier Transform (FFT) of the left-most image. The six outermost peaks of the FFT correspond to the atomic lattice and the inner six peaks reflect the CDW periodicity.

(b) Clockwise

\[ \psi_1 + 0.6\psi_2 + 0.3\psi_3 \]

A symmetry broken, “chiral” charge density wave with non-equivalent CDW amplitudes. The three different CDW amplitudes are reflected in the amplitudes of the corresponding peaks in the FFT.

(c) Counter-Clockwise

\[ 0.3\psi_1 + 0.6\psi_2 + \psi_3 \]

The same as in (b) with opposite chirality.
In general, the three CDW components can have independent phases and amplitudes. For a crystal with $C_3$ symmetry, we expect the three CDW vectors to be related by this symmetry (rotation by 120°) and the three CDW amplitudes should be equivalent, $|\psi_1(\vec{r})| = |\psi_2(\vec{r})| = |\psi_3(\vec{r})|$. An example of such a case is given in figure 2.10a displaying a 2D map of the charge density by modeling the CDW and lattice components as a sum of plane waves (see equation 2.63). From the left, figure 2.10a shows the model charge density wave with the lattice included, the CDW amplitude without the atomic lattice plotted, and the fast Fourier transform of the leftmost image. In the FFT image, the peaks at the CDW wavevectors are equal in amplitude, reflecting the symmetry of the CDW order parameters. Mechanisms which break this symmetry (ie. anisotropic strain or orbital hybridization) may lead to changes in the local CDW registration; that is, the relative amplitudes and phases of the 3 CDW components. In figures 2.10b and c, we show examples of asymmetric charge density waves where the amplitudes of the three CDW components are no longer equal. We can see that the individual peaks of the CDW become distorted and the amplitudes of the CDW spots in the FFT spectra are no longer equal. Such is the case of the charge density wave in 2H–TiSe$_2$ where hybridization between Ti and Se orbitals breaks the underlying $C_3$ symmetry of the lattice leading to local ”chiral” charge density waves which may be considered clockwise or counter-clockwise depending on the relative amplitudes of the three CDW components.

### 2.3 Electronic Phases of Tantalum Disulfide

In general, 2D TMDs have already gained a lot of attention in the research community, however, most of these studies have been limited to the group VI TMDs (MoX$_2$ and WX$_2$). Unlike their semiconducting cousins, the group-V TMDs (NbX$_2$ and TaX$_s$) have an additional non-bonding d-electron making these materials metallic. In
this section we begin to describe the electronic properties of 1T–TaS$_2$, a fascinating material that exhibits a large number of distinct electronic phases due to high degree of both electron-phonon and electron-electron interactions. These give rise to several correlated states including several charge density wave and Mott insulating states. Each phase comes with its own ordering physics and different electronic properties with sharp transitions between them. Particularly attractive are two first-order phase transitions accompanied by sudden changes in material resistivity. Additionally, it has been found that the properties of these phases and their transitions can be manipulated using external control knobs such as electric fields, carrier doping, and pressure. The large changes in material properties, the ability to manipulate transitions between phases, and the ability to separate this material into very thin films has made it a promising candidate for use in next-generation memristive and switching devices.

Low temperature optical and STM studies of 1T–TaS$_2$ have additionally revealed metastable, ordered states that arise after electrical or optical pulses which seem distinctly different from the equilibrium phases. Perhaps even more exotic than the charge ordered states is the magnetic properties of 1T–TaS$_2$. Peculiarly, 1T–TaS$_2$ shows a suppression of magnetic ordering down to the lowest measurable temperatures, identifying it as a proposed quantum spin liquid. The strong interactions in this material present an opportunity to test and study a variety of complex and novel phenomena.
2.3.1 Undistorted Phase

Figure 2.11: Undistorted 1T-TaS$_2$ a) Ball and stick model of the 1T-TaS$_2$ unit cell from a top view (left) and side view (right) blue and yellow spheres depict Ta and S atoms, respectively [102] b) Band structure of undistorted 1T-TaS$_2$ determined from density functional theory c) Phonon band structure for undistorted 1T-TaS$_2$ shows instabilities (negative frequency solutions) at several locations within the Brillouin zone (marked with a blue circle and yellow diamond).

Above its charge density wave transition temperature (550K), in its undistorted phase, layers of 1T-TaS$_2$ are composed of planes of tantalum atoms sandwiched between octahedrally coordinated sulfur atoms. Similar to graphene, atomic planes are weakly bound together; separated by a van der Waals gap, allowing this material to be mechanically cleaved into thin layers.

Figure 2.11 displays the atomic structure of one unit cell of 1T-TaS$_2$. Unit cells are linked by multiples of the in-plane lattice vectors, $\vec{a}$ and $\vec{b}$, with lengths 3.365Å and rotated 120° with respect to each other and adjacent layers are separated by the out-of-plane lattice constant $c = 5.897\text{Å}$ [103]. The reciprocal lattice vectors, $\vec{a}^*$ and $\vec{b}^*$, are defined by the relationship $\vec{a}^* \cdot \vec{a} = 2\pi$ and $\vec{b}^* \cdot \vec{b} = 2\pi$, respectively. To connect with the Landau theory of CDWs, the six shortest reciprocal lattice vectors are given as $\vec{G}_1 = \vec{a}^*$, $\vec{G}_2 = \vec{b}^*$, $\vec{G}_3 = -\vec{a}^* - \vec{b}^*$, $\vec{G}_4 = -\vec{G}_1$, $\vec{G}_5 = -\vec{G}_2$, and $\vec{G}_6 = -\vec{G}_3$. 
We can see from the band structure in figure 2.11b that undistorted 1T–TaS$_2$ is metallic due to the bands crossing the Fermi level between the Γ–M and M–K directions. As suggested by the observation of CDW formation at low temperature, the calculated phonon spectrum of 1T–TaS$_2$ shows negative phonon frequencies at certain points in the Brillouin zone indicating structural instability. Figure 2.11c shows the calculated phonon spectrum for a monolayer 1T–TaS$_2$. The three lowest energy bands are attributed to acoustic vibrational modes corresponding to in-plane (LA and TA) and out-of-plane (ZA) oscillations. We can see that phonon anomalies occur in one of the acoustic branches with the most pronounced anomaly near the known CDW wave vector, $Q_{CDW}$ (blue dot in fig. 2.11c). Comparing to the k-dependent electronic susceptibility, it has been argued that Fermi-surface nesting plays a weak role in the CDW formation in this material and that strong electron-phonon coupling plays the primary role [104].

2.3.2 Incommensurate Charge Density Wave Phase

The normal to incommensurate CDW transition behaves differently in 2D than in the Peierls-like 1D case. To see this, we start with McMillan’s free energy for the 2D CDW (2.64) and assuming the three CDW order parameters to have the same magnitude $\psi_i = |\psi_0|e^{i\vec{Q}_i \cdot \vec{r}}$, we get

$$F = \frac{3}{2}a_0|\psi_0|^2 - \frac{3}{2}b_0|\psi_0|^3 + \frac{3}{8}(15c_0 - 8d_0)|\psi_0|^4.$$

(2.66)

The transition temperature is determined from the point where the free energy goes to zero,

$$T_{IN} = T_0 + \frac{b_0^2}{a(15c_0 - 8d_0)}.$$

(2.67)

The normal-to-ICCDW transition temperature of the 2D CDW differs from the 1D case, $T_0$, due to the cubic, ”phasing energy,” term and the fourth order interaction term between the three CDW components. To evaluate the change in free energy
across the transition, we must first solve for the order parameter in the incommensurate state by minimizing the free energy (find $\psi_0$ at $\frac{F(T<T_{IN})}{\psi_0} = 0$ s.t. $F(\psi_0)$ is a minimum). We find,

$$\psi_0 = \frac{3b_0}{2(15c_0 - 8d_0)} + \left[\frac{3b_0}{2(15c_0 - 8d_0)}\right]^2 - \frac{2a_0}{15c_0 - 8d_0}\right]^{1/2}.$$  

(2.68)

Now we can determine the change in entropy across the transition,

$$\Delta S = \frac{\Delta F}{T_{IN}} = \frac{3ab_0}{(15c_0 - 8d_0)},$$  

(2.69)

and see that the phase transition is of first order with transition entropy proportional to $b_0$. Thus, the interaction between the three CDW components results in a first order transition where the 1D CDW exhibits a second order transition.

Figure 2.12: **Incommensurate Charge Density Wave Phase** a) Model of the real space charge density distribution in the incommensurate CDW phase. Two 2D modulations can be seen corresponding to the Ta lattice and the CDW modulation b) FFT of the incommensurate CDW phase shows peaks corresponding to the two periodic modulations seen in (a). The lattice modulation is characterized by the reciprocal lattice vectors $\vec{G}_1$ and $\vec{G}_2$ while the ICCDW modulation is characterized by the three wavevectors $\vec{Q}_{IC}^{(i)}$.

At last, we can begin to connect McMillan’s Landau theory to the actual material of interest, 1T–TaS$_2$, by relating to experimentally determined details of the CDW
phases. Below approximately 543K, the lattice shifts in a weakly first-order phase transition into its incommensurate CDW state with wavevectors $\vec{Q}_{IC}^{(i)} = 0.283\vec{G}_i$ ($i = 1, 2, 3$) [105, 106, 107, 92, 108, 109, 103]. Figure 2.12 shows an illustrative example of the CDW in the incommensurate state in real space and Fourier space. In this example, a modulation due to the Ta lattice is added to help illustrate the features of the ICCDW phase. The charge density modulation is aligned in angle with the atomic lattice, but the wavelength of the CDW does not commensurate with the lattice. The fast Fourier transform (FFT) of the real space image reflects the periodicity of both the CDW and the atomic lattice by showing discrete peaks at the ends of the reciprocal lattice vectors for the lattice, $\vec{G}_i$, and for the CDW, $\vec{Q}_{IC}^{(i)}$. The lattice spacing and CDW wavelength can be determined from the peak positions through the relationships $|\vec{G}_i| = \frac{4\pi}{\sqrt{3}|\vec{a}|}$ and $|\vec{Q}_{IC}^{(i)}| = \frac{4\pi}{\sqrt{3}\lambda_{CDW}}$, respectively.

Recall that in the 1D case, the free energy of the CDW state is independent of its phase and is thus, able to conduct through sliding of the CDW [110, 87, 88]. In the 2D case, however, there is a phasing energy which relates the three CDW phase in space. The phasing energy prevents the translation of one CDW component with respect to the other two. Possibly due to this phase interaction, sliding charge density waves, identified by non-Ohmic resistivity, are not observed in the 2D, three component CDWs at small fields [111, 112]. However, there is evidence, from noise spectroscopy measurements, of a depinning threshold similar to that observed in 1D CDWs with strong pinning defects [113, 114]. Above the threshold, the CDW might conduct by sliding as discussed for the 1D case [110, 87, 88]. The strength of the depinning potential is larger than what is observed in the doped 1D CDWs leaving the question of whether the threshold observed in 2D is due to pinning impurities or interaction of the 3 CDW components in space still up for debate.
2.3.3 Commensurate/Mott-Insulating Phase

Figure 2.13: **Commensurate Charge Density Wave Phase** a) Lattice structure of the commensurate CDW superstructure. Ta atoms arrange into a Star of David (SD) pattern involving 13 Ta atoms each. The SDs form a 2D triangular lattice with lattice spacing $\lambda_{CDW} = 1.21\text{nm}$. Ta atoms are shown in brown and S atoms are shown in yellow b) Band structure of CDW reconstructed $1\text{T-TaS}_2$ with a narrow, flat band isolated at the Fermi level. c) Band structure of CDW reconstructed $1\text{T-TaS}_2$ including electron-electron interactions (with the Mott-Hubbard $U=2.27\text{eV}$). The flat band has been split into a lower (near the Fermi level) and upper (touching the conduction band) Hubbard bands with a gap between them. d) Model of the real space charge density distribution in the commensurate CDW phase. Two 2D modulations can be seen corresponding to the Ta lattice and the CCDW modulation d) FFT of the commensurate CDW phase shows peaks corresponding to the two periodic modulations seen in (d). The lattice modulation is characterized by the reciprocal lattice vectors $\mathbf{G}_1$ and $\mathbf{G}_2$ while the ICCDW modulation is characterized by the three wavevectors $\mathbf{Q}_C^{(i)}$ which are rotated with respect to the lattice by $13.9^\circ$.  

From the Landau model, we see that commensuration costs energy to both change the wavelength and rotate the CDW toward commensuration, however, energy is gained
from the Umklapp terms ($\propto \alpha c_1$). Unlike many of the triple-q CDWs observed in the 2D TMDs, the incommensurate CDW wavelength in TaS$_2$ is too far from $|\vec{G}|/3$ and instead of locking in at that wavevector, the CDW lattice rotates with respect to the atomic lattice to interlock into commensurate super structure with wavevector $\vec{Q}_C = 3\vec{a} \pm \vec{b}$ (see figure 2.13a). The hexagonal CDW lattice is, thus, rotated by 13.54° with respect to the atomic lattice either clockwise or counterclockwise (defining a chirality). The periodic lattice distortion generates Ta atom clusters in the shape of a 6-pointed ”Star of David” (SD). Within each SD, twelve Ta atoms converge toward a central, thirteenth Ta atom (with a slight rotation). The atomic displacements in TaS$_2$ are the largest of the known triple-q CDW materials, up to 0.24Å. While the largest displacements are in-plane, there is a slight bulging of the sulfur atoms out-of-plane (c-axis) [103]. Bonding in the c-direction may then be phase dependent. Further, the combination of the three dimensional distortion and several metastable CDW stacking orders can lead to very different band structures [115]. Whether the out-of-plane stacking order of the CDW is ordered or not is still debated. Early diffraction measurements have suggested the possibilities of interlayer stacking orders with periods of 2 layers, 3 layers, and 13 layers. However, in all cases, the interlayer coupling strength is expected to be small leading to a large degree of stacking disorder in all measured samples.

Recall that a typical CDW is metallic, however, this one is semiconducting. To investigate, we can look to the details of the band structure in the commensurate phase. During commensuration, 12 bands associated with 12 Ta atoms per CDW unit cell are pushed away from the Fermi level, generating a CDW gap below it. A 13th band, belonging to the central Ta atom, is then isolated at the Fermi level by spin-orbit interaction (see figure 2.13b) [116, 117]. The isolated, narrow ($\approx$80meV), half-filled band is then split by a Mott-Hubbard interaction. That is, electron-electron (Coulomb) interactions between Ta atoms at the centers of the SDs become large enough to localize the conduction electrons and open a gap in this narrow band
(figure 2.13c). Then, the CCDW phase should be considered a Mott insulator with a gap $\Delta_{\text{Mott}} \approx 0.32\text{eV}$ [118].

Experiment has shown a corresponding, abrupt increase in resistivity by nearly one order of magnitude [119]. Since the Mott state is brought on by the spontaneous CDW transition, the Mott transition might also be controlled by the CDW order. The first order phase transition while cooling into the CCDW phase contrasts the second-order transition that arises from McMillan’s Landau theory [95, 96]. However, using an approximate model considering the small coherence length limit, a first-order ICCDW to CCDW phase transition can be achieved [94]. We will show in the next subsection that the ICCDW to CCDW phase transition in 1T-TaS$_2$ is far more complex and requires consideration of higher order components of the CDW.

2.3.4 Nearly-Commensurate Phase

Experimentally, 1T-TaS$_2$ is observed to exhibit peculiar region of its phase diagram below its incommensurate transition temperature and before the CDW is able to commensurate with the lattice. As the temperature of TaS$_2$ decreases below $T_{IN}$, electron-lattice interactions become stronger (relative to the average kinetic energy of the system). Between 180K and 350K, the CDW forms regions of commensuration (where $\lambda_{\text{CDW}} = \frac{p}{q}a$, ie. the charge fluctuation aligns with the underlying Ta lattice) bounded by domain walls of phase dislocations (See Fig. 2.14) [120]. This phase is known as the nearly commensurate charge density wave (NCCDW) phase [121, 109, 122, 123, 124, 125, 106].

In terms of our working Landau theory, the effect of Umklapp terms (proportional to $b_1$, $c_1$, etc.) and the effects of CDW harmonics play a stronger role and can no longer be ignored (as was done treating the ICCDW). By including the harmonic components (see equation 2.65), Nakanishi and Shiba showed that a new energy minimum appears in the free energy landscape as a function of wavevector (for some regions of the parameter space). Using parameters consistent with experimental observations of
TaS$_2$ in the ICCDW phase, the wavevector of the new free energy minimum agrees well with the experimentally determined value in the NC phase. Thus, it is the effect of higher harmonics which leads to this interesting CDW phase. Solving for the harmonic amplitudes as a function of wavevector, it is found that the higher harmonic components become substantially large near the commensurate value, $Q_C^{(1)}$. Considering the strongest two harmonic components, $\Delta_{000}$ and $\Delta_{011}$, is sufficient to illustrate the first order phase transition between the ICCDW and NCCDW phases (see [99] for details).
Figure 2.14: Nearly Commensurate Charge Density Wave Phase a) Model of the real space charge density distribution in the nearly commensurate CDW phase. Phase and amplitude modulations show the hexagonal domain structure b) FFT of the NCCDW phase shows strong peaks corresponding to the NCCDW wavevectors $\mathbf{Q}_{NC}^{(i)}$. Weaker spots can be seen in the vicinity of the fundamental peaks corresponding to the strongest harmonic components ($\mathbf{Q}_{011}$ is labeled for example). The difference between the fundamental and satellite wavevectors give the inter-domain wavevector, $\mathbf{k}_d$. c) A closer look at the NCCDW model shows both the CDW modulation as well as the Ta lattice (added for illustration) shows the local commensuration of the CDW and atomic lattices within the domains and distorted CDW maxima between domains d) FFT of the NCCDW phase shows peaks corresponding to the modulations seen in (c). The lattice modulation is characterized by the reciprocal lattice vectors $\mathbf{G}_1$ and $\mathbf{G}_2$ while the ICCDW modulation is characterized by the three wavevectors $\mathbf{Q}_{NC}^{(i)}$ (related by $120^\circ$ rotations)

We can use Nakanishi and Shiba’s result to plot the full CDW order parameter in space. The result reproduces the domain-like CDW structure observed experimentally. The CDW order parameters take on a more complex form than in the ICCDW
case, however, we can similarly use the theoretical result to demonstrate how the CDW looks in real space. The charge density of the nearly commensurate phase is modeled in figure 2.14a. Upon close inspection, we can see that the CDW amplitude modulates in space leaving roughly hexagonal regions where the CDW is close to commensuration and bounded by domain walls with a reduced CDW amplitude. The FFT of the charge density image shows peaks (black spots) at the reciprocal lattice vectors and the fundamental CDW components, $\vec{Q}_{NC}^{(i)}$. In addition to these fundamental peaks, there are smaller intensity peaks near the fundamental CDW wavevector corresponding to the strongest harmonic components in the NCCDW phase. The difference between the fundamental and satellite wavevectors gives the wavevector associated with adjacent domains, $\vec{k}_d$, related to the domain spacing by $|\vec{k}_d| = \frac{2\pi}{\lambda_d}$.

It is well known from experiment that this phase is conducting; about 10 times as much as the CCDW phase, but half as much as the ICCDW phase. Despite the lack of microscopic information, several mechanisms of conduction in this phase have been proposed. It may be the result of the domain walls being metallic [126], screening from within the domain walls contribute to make the Mott-CDW domains metallic [127], or by shifts in the stacking order of the commensurate domains [115]. Most experimental work, up to now, has focused on the transition into and out of this phase rather than the dynamics of the electrons within the phase. Thus, more experimental work is needed to pinpoint the mechanism of conduction in this unique CDW state.

### 2.3.5 T Phase

Upon warming from the commensurate phase, the long-ranged CCDW/Mott state breaks into elongated hexagonal domains unlike the nearly regular hexagonal domains of the NCCDW phase. The existence of this phase has been revealed through measurements of thermal expansion [128], Seebeck effect, resistivity [129], STM [106], and x-ray diffraction [130] and has been termed the T-phase.
Notice that the T-phase did not appear in Nakanishi and Shiba’s original extension of McMillan’s theory of 2D CDWs. In all of the previous sections, we have been assuming that the CDW structure maintains the underlying $C_3$ symmetry of the lattice. Clearly, the T-phase does not satisfy this condition. It has been reasoned that interlayer interactions are responsible for this unusual symmetry breaking, however, interlayer coupling might not be a requirement for the observed T-phase. Recent work has shown that relaxing the $C_3$ symmetry requirement in Nakanishi and Shiba’s model allows for the appearance of additional free energy minima (vs. wavevector) corresponding to the T-phase as well as some other possible, metastable CDW structures.

As was noted by McMillan, the low energy excitations of the 2D CDW are 1D line defects. The generation of 1D line defects in the 2D isotropic background breaks the $C_3$ symmetry. This spontaneous reduction of symmetry increases the entropy of the system, minimizing the free energy. Thus, the symmetry breaking of the T-phase might be a natural aspect of melting of the 2D CDWs.

This proposed BKT-like transition between the CCDW and T-phase is not supported by experiment, which shows strong, hysteretic jumps of many measurable values (resistance, interlayer spacing, etc.) indicative of a first order transition. To add more mystery to the problem, ARPES measurements have shown that the Mott-Hubbard bands remain unaffected through the CCDW to T-phase transition and the Mott state does not collapse until the NCCDW phase [131]. The first-order behavior of the CCDW to T-phase transition is likely not due to a collapse of the Mott state. The effect of the addition of domain walls on stacking order and in-plane Coulomb interactions may need to be further investigated in the T-phase to elucidate the nature of the CCDW to T-phase transition.

As temperature increases further, the T-phase undergoes a phase transition (either second-order or weakly first order) into the NCCDW phase when the T-phase domain size becomes comparable to that of the NCCDW phase [130, 97]. The fact
that the T-phase is only clearly observed upon warming the sample might indicate that the T-phase is a metastable phase (not a global minimum in free energy phase space), however, x-ray diffraction measurements have shown that both the T-phase and NCCDW phase can coexist within the same sample during cooling [130]. The existence of the T-phase during both cooling and heating does not rule out the T-phase as the ground state within this temperature range. Temperature-controlled, local STM measurements would be useful to monitor the transitions into and out of this phase to truly identify if the line defects of the T-phase evolve as BKT-type defects (with a continuous transition) and to directly observe if the T-phase and NCCDW phase domains coexist within the sample upon both heating and cooling.

2.3.6 Manipulation and Metastability of CDW Phases

To produce a novel switch based on correlated materials, one needs to be able to control the electronic phases with an external stimulus. For a traditional, semiconductor FET, the external stimulus is a gate voltage which reduces the barrier for electrons to be excited into a conducting channel. Due to its technological promise, many methods of manipulating the electronic phases of TaS$_2$ have recently been explored including doping, alloying, intercalation, pressure, strain, electric fields, and optical pulses. The effects observed in these experiments deserve a comprehensive review of their own; however, as these methods are not employed in this thesis work, only a brief description of the main observations is provided here.

Unlike semiconductors (also semimetals and graphene), metallic 1 T–TaS$_2$ has a large carrier concentration ($n_{2D} \approx 10^{15} cm^{-2}$) making it difficult to move the Fermi level using electrostatic gating. Traditional methods of electronic doping such as chemical substitution has been used to demonstrate the suppression of the commensurate CDW/Mott state and leads to the emergence of superconductivity at low temperatures [132], however, non-uniform chemical substitution leads to both electronic and lattice disorder and, at large doping, strongly modify the material’s intrinsic band
structure [133].

Intercalation with ions is another popular method of achieving different carrier concentrations, modifying interlayer interactions, and altering the CDW symmetry. Intercalation with certain ion species has the effect of distorting the commensurate CDW and weakening Ta-Ta bonds, hindering commensuration [134, 135, 136]. While other ions strain the lattice such as to squeeze the David stars, enhancing the commensurate CDW phase above room temperature [137]. Electronic control of carrier density in thin flakes of TaS$_2$ can be achieved using gate-controlled intercalation of lithium ions. Similar to the effect of selenium doping, doping to high carrier densities with electrolyte gates suppresses the commensurate phase and leads to superconductivity at low temperatures [138]. In accordance with these observations, it has been shown through theoretical modeling that electronic disorder and large doping can disturb and even completely suppress CDW commensuration [139]. In general, the coincident suppression of the commensurate CDW phase and the emergence of superconductivity points to a competition between superconductivity and charge ordered electronic states. However, both the commensuration of the CDW and formation of a Mott gap reduce the carrier density at the Fermi level, naturally suppressing superconductivity. Thus, the lack of superconductivity in the CCDW state is not necessarily due to a competition between CDW and superconducting ordering, but might be a simple consequence of having no carriers at the Fermi level due to Mott localization.

With the application of an in-plane electric field, Joule heating can cause a sudden transition between CDW phases. When a field is applied at room temperature, the NCCDW state switches abruptly into the ICCDW phase (rather, a state which has the same resistivity as that of the equilibrium ICCDW) [140]. When applied to the CCDW phase at 77K, both transitions into the NCCDW and ICCDW phases have been observed [119]. The resulting non-linear conduction properties can be exploited for novel electronic devices as was demonstrated by constructing a nanoscale oscillator.
circuit utilizing TaS$_2$ and graphene. The ability to alter the graphene resistance with a backgate allowed the authors to tune the frequency of the current oscillations of the circuit [141]. It was later shown that the critical field at which the CDW phase of the 1T–TaS$_2$ transitions is sensitive to light, allowing the oscillator to be used as a photodetector [142].

Interaction between different CDW components, weak interlayer interactions, and the interplay between the effects of electron-phonon and electron-electron interactions leads to very high degree of metastability in this material. This is evident in the sensitivity of the CDW phases to the rate at which the sample is cooled. Quickly cooling 1T–TaS$_2$ films can quench the system into a metastable nearly commensurate phase without Mott localization [143, 144, 145]. The critical cooling rate depends on the thickness of the sample and the type of substrate used indicating a possible role of disorder on the suppression of the CCDW state. As the resistivity continuously increases from the NCCDW phase into the quenched state, authors often consider the low temperature state to be a super-cooled version of the nearly commensurate phase, however, without local measurements it is not clear if the quenched state is ordered the same as the NCCDW state.

Perhaps similar to the states induced by quenching, one can access metastable, charge-ordered ground states with very long lifetimes upon application of an optical or local electronic pulse. Local voltage pulses from an STM tip can inject charge into the system by redistributing the CDW polarons and forming non-equilibrium domain walls which are frozen in place at low temperatures [146, 120, 147]. Interestingly, the islands of commensurate CDW bounded by the induced domain walls is found to be metallic and is named a "melted Mott" state. Additionally, the domain walls themselves can host 1D electronic states within and alongside them [148, 149]. Similarly, one can induce metastable charge configurations within the CCDW phase using laser pulses [150, 151, 101]. However, unlike STM pulses, optical pulses are charge neutral and cover a much larger area of the sample surface. These have led to interesting
observations of how charge conservation dictates the configurations of domain walls induced by laser pulses [147] and leads to a non-trivial long-ranged ordering of domain wall density and orientation [101]. Further application of laser pulses at low temperatures can lead to a redistribution of the CDW polarons which is even more tightly packed than the natural CCDW ground state of the system [152].

As TaS$_2$ is a layered material, it is expected that one would ask: What happens to the electronic properties of this material when thinned to the 2D limit? First, it was noticed by many groups that when thinned below 20nm (33 layers), the commensurate charge density wave ground state is no longer observed. Instead, only the nearly commensurate phase transition is seen in transport measurements. As the temperature is reduced, thin flakes between 5-12nm (8-20 layers) remain metallic whereas flakes below 5nm (8 layers) are insulating at low temperature [143, 138]. The exact thickness at which the commensurate phase is lost varies from sample to sample. Then, the loss of the commensurate phase may not be an intrinsic material property, instead it is evidenced that lattice disorder, which occurs during cleaving and depositing the thin flakes onto relatively rough substrates, may be the leading cause for the lack of commensuration and formation of the associated Mott state [153]. This question has led to several recent studies of monolayer 1 T–TaS$_2$ crystals which are grown (by molecular beam epitaxy) on atomically smooth substrates [154] or exfoliated and suspended without a substrate [155]. However, STM measurements of monolayer TaS$_2$ on a substrate show a commensurate CDW populated with many defects [154], while TEM measurements of a suspended monolayer region show a distorted CDW akin to that of the T-phase [155]. Thus, the natural ordering of the ground state of monolayer TaS$_2$ is still unclear at this time and requires further measurements of high-quality monolayer samples to address.
2.4 1 T–TaS$_2$ as a Quantum Spin Liquid

During the course of this thesis, several breakthrough measurements of 1T-TaS$_2$ revealed a theoretically proposed, magnetically disordered ground state known as a quantum spin liquid (QSL). QSLs are a topic of much interest because such a ground state is not characterized by any order parameter or spontaneous symmetry breaking and must be described by physics beyond the typical Landau picture. Modern theory currently describes this phase with the language of long-range entangled phases and topological order. These topics are extremely rich with exciting new, fundamental physics, however, the details of this phase and its viability for technological applications are not yet well-understood and fall beyond the scope of this thesis work. As a starting point for readers interested in this topic, I provide a brief introduction to quantum spin liquids, the measurements made which place 1T-TaS$_2$ as a candidate for the QSL ground state, and briefly discuss measurements which might better probe the details of this state.

2.4.1 What is a Quantum Spin Liquid?

There are several reviews available on the topic of quantum spin liquids (QSLs) which provide a useful overview of the topic and relevant resources [156, 157]. I will only briefly reiterate select points for a beginner’s introduction.

A magnetic material tends to order itself due to the exchange interaction between electrons. Typically, a material will order into a spin-aligned ferromagnetic state or a spin-misaligned antiferromagnetic state depending on the sign of the superexchange coupling constant, $J$, as displayed in fig 2.15a,b for the one-dimensional Heisenberg model. However, for $J > 0$ on a two-dimensional triangular lattice, spins between two neighboring atoms may anti-align leaving the third spin undetermined (see figure 2.15c). There is a solution (realized in nature) which resolves the geometrical frustration in which the three spins per triangle align at 120° with respect to each other.
so that the net spin per triangle is zero in the lattice plane (called the Néel state) (see fig 2.15d).

P.W. Anderson proposed a less trivial ground state where electron pairs on neighboring lattice sites form spin singlet pairs. Without asymmetric interactions between neighboring sites (like in the Kitaev model), the ground state of the system is described by a superposition of possible singlet pairs (fig 2.15e). In this model, the spins do not develop a long-range magnetic order (like the Néel state) and behave as a liquid, even at zero temperature. Thus, this state is termed a quantum spin liquid.

Recall that in the low temperature phase of 1T-TaS$_2$, a Mott-Hubbard interaction localizes a single electron state on the center of each star of David in the triangular commensurate CDW lattice. As a Mott insulator on a two-dimensional triangular lattice, it is interesting to ask what is the magnetic ground state of this system [158, 159, 160, 161, 162].

Figure 2.15: *Introduction to Quantum Spin Liquids* a) For a positive value of J the 1D metallic chain becomes ferromagnetic according to the Heisenberg model b) For negative J, the Heisenberg model gives an antiferromagnetic atomic chain. c) For negative J on a 2D triangular lattice, the spin ground state is not obvious. d) The Néel state is a geometric solution to the antiferromagnetic state on a 2D triangular lattice observed experimentally. e) The proposed quantum spin liquid state is made from a superposition of all possible spin singlet pairs. Figure adapted from [22]
2.4.2 Is 1T-TaS2 a QSL?

The first step to proving 1T-TaS2 as a quantum spin liquid is to show that it does not exhibit any magnetically ordered phases down to the smallest measurable temperatures. Various methods can be used to rule out magnetically ordered phases at low temperature. Muon spin relaxation ($\mu$SR) is a powerful probe for local magnetic moments as the polarized muons intercalate into the material and are sensitive to local and long-range magnetic fields. $\mu$SR measurements of 1T-TaS2 at low temperatures reveal a lack of fixed local magnetic moments indicating that this material exhibits no magnetically ordered phase.

Several authors report a temperature independent component to magnetic susceptibility at low temperatures further supporting the QSL state. Fitting the high-temperature dependence of magnetic susceptibility to a Curie-Weiss law provides a measure of the interaction coupling strength, $J$. There is currently a discrepancy in the estimation of exchange interaction constant. Authors [22] et. al fit magnetic susceptibility and find $J \approx 0.13eV$ while [159] et. al report $J \approx 0.1meV$ from a similar fit to SQUID measurements. The two extreme values have different implications on the degree of frustration in this system, the size of the excitation gap (if any), and the temperature range in which the material exhibits QSL behavior. Further measurements should resolve this issue.

At low temperatures, the charge and lattice degrees of freedom become frozen. Then, any residual heat capacity or thermal conductivity comes from ordering of the spin system. Indeed, measurements of the heat capacity show some anomalous magnetic field-dependent behavior although the mechanism is currently unknown [159].

These measurements strongly support TaS2 as a QSL by revealing anomalous low temperature spin excitations. These works, however, do not describe the nature of the spin excitations in this state. QSLs may be either gapped to excitations and exhibit edge states or it may be gapless. To uncover the details of the spin excitations at low
temperatures, nuclear quadrupole resonance gives a statistical scaling relationship between spin relaxation times at temperature. Describe the figure Recently, authors measure thermal conductivity and claim that there is no residual thermal conductivity as T goes to 0, suggesting a gapped QSL state [162].

The nature of the low temperature spin excitations is still an open question. There appears to be two possibilities: a gapless, nodal Fermi surface with a non-trivial density of states at the Fermi level or a very large gapped Fermi surface. The former case cannot explain the lack of thermal transport at low temperatures and the latter case does not explain the strange behavior of spin relaxation with temperature. It is possible that TaS2 is a gapped QSL where the strange temperature scaling of spin relation rates as well as the magnetic field-dependent heat capacity might be attributed to stacking disorder of the CDW and/or out of plane conduction of spins [162].

2.5 Intrinsic Stacking Order in the CCDW Phase of 1 T-TaS2

As TaS2 in the commensurate CDW phase has an odd number of conduction electrons (13 electrons per CDW unit cell), simple band theory predicts that the CCDW phase should be metallic. Indeed, calculations starting from a tight binding or DFT model have shown that TaS2 has a band which crosses the Fermi level along the $\Gamma \rightarrow A$ cut through its first Brillouin zone [116, 117]. This implies 1D metallic conduction along the out-of-plane direction. As noted in previous sections, experiment observes insulating behavior consistent with that of a Mott insulator. For example, early STS measurements show Mott-Hubbard bands aside a wide band gap [163], consistent with the predicted Mott state [164], and transport measurements show resistivity that diverges to infinity as temperature is reduced [165]. However, careful inspection shows that a small, but finite density of states persists into much of the gap in STS measurements and fits to an Arrhenius law reveal an abnormally small gap size
compared to the expected Mott gap (∼10^4 times smaller). Are the "mid-gap states" due to the expected out-of-plane hopping? Why then, do transport experiments not observe metallic behavior?

Several mechanisms have been proposed to explain the observed insulating behavior. The half-filled, 1D metallic channels could be subject to a Peierls distortion in the out-of-plane direction, opening a single particle gap. However, a 1D CDW should still conduct [167]. Random stacking disorder might induce Anderson localization in these 1D metallic chains, however, introducing more disorder by doping is seen to decrease resistivity [164]. Finally, alternative stacking orders of adjacent planes of CDW orbitals are shown to be able to produce different band structures, however, the resulting density of states do not match those observed in STS experiments and do not lead to out-of-plane insulating behavior for any coherent stacking order tested [115, 168]. Following the notation of [166], the 13 Ta atom sites within the CCDW unit cell (A through M), the 5 distinct stacking orders can be described with respect to the 'A' site of a reference layer: A → A, A → B, A → M, A → C, and A → L, or equivalently, by the stacking vectors \( \vec{T}_s = \vec{c}, \pm \vec{a} + \vec{c}, \) and \( \pm 2\vec{a} + \vec{c} \) (see figure 2.16).

Recently, there has been theoretical and experimental evidence of bilayer Peierls-like dimerization, the formation of inter-molecular bonds between adjacent CDW cells in CCDW TaS_2 [166, 169, 170]. A doubling of the CDW unit cell then implies an even number of electrons per unit cell, opening the possibility that the CCDW
state is a band insulator in the out-of-plane direction. DFT calculations of different stacking structures have revealed that the AL stacking sequence is the lowest energy, however, the L sequence is also very stable with a small energy difference of 1.1 meV/star (compared to AL stacking). These DFT results are consistent with the predictions of Nakanishi and Shiba, who extended the Landau-Ginzburg-McMillan CDW theory [171]. The calculated band structure of the AL stacked ground state shows a band gap of approximately 150 meV which explains the insulating behavior. The L configuration, however, is found to be metallic. Additionally, the interlayer spacing in the L configuration is reduced and weak \( \pi \) bonds form between sulfur atoms in adjacent clusters, mediating interlayer transport [166].

Recent x-ray diffraction measurements have provided some experimental evidence of bilayer dimerization in the out-of-plane direction. Diffuse diffraction spots in XRD measurements suggest a layer doubling with some large degree of disorder which causes the peaks to smear out [169]. Previous low energy electron diffraction has provided one example of a surface with an AL stacking termination [172]. Additionally, STM measurements of TaS\(_2\) crystals cleaved at low temperatures have provided local evidence of the AL ground state stacking sequence. Out of 24 surfaces measured, 75\% displayed characteristics of a A-stacked surface while 25\% showed a L-stacked surface and, in several instances where unit-layer steps were observed, the stacking sequence was found be the AL stacking order. Interestingly, the unpaired, L termination is found to be insulating, contrary to conventional wisdom. Even in the unpaired layer, the strong Coulomb interaction causes Mott physics to dominate over conventional band theory [170].

These low temperature STM measurements also reveal a collapsed Mott state when the stacking order assumes a the "B" configuration, similar to the collapsed Mott spectra observed in STM-pulsed and photo-excited surfaces. This might suggest that the metastable, collapsed-Mott and "hidden" phases might be realizations of quenched, non-equilibrium stacking orders [170]. DFT calculations show that while
there is a small total energy difference between AL and L stacking configurations, there is a large van der Waals energy difference ($\sim$170 meV/star). Perhaps resulting in a strong interplay between stacking configurations and doping (governing electronic energy gain from gap opening) or pressure (governing van der Waals interaction energy) might explain why stacking order is easily perturbed by electronic pulses and pressure as well as why cooling rate and field cooling govern the ground state resistivity [166].

Notice that for the monolayer triangular lattice, there is an intrinsic spin frustration which is the basis for understanding the observed quantum spin liquid (QSL) behavior. The doubled unit cell of the AL-stacked ground state, however, has a more obvious, antiferromagnetic solution to its ground state due to coupling within the bilayer. It seems, then, that the observed spin disorder at very low temperatures (see section 2.4) requires some degree of stacking disorder in order to leave orphan spins within the unpaired layers. Locally manipulating the stacking order (say, through STM pulses) might be one method of controllably (and reversibly) generating a surface-exposed QSL system within an antiferromagnetic background (ie. a "quantum spin puddle").

Ultimately, the bilayer picture with alternating stacking sequence is consistent with previous experimental results and likely represents the true, 3D ground state of the CCDW phase. However, it should be noted that stacking disorder seems to be intrinsic to this system due to the low energy difference and high energy barriers between stacking configurations as well as degenerate stacking orders. It seems that an accurate description of the 3D magnetic and conduction properties of TaS$_2$ requires elements of both bilayer dimerization as well as stacking disorder.
Chapter 3
Experimental Methods

3.1 Scanning Probe Microscopy

Scanning probe microscopy is a family of techniques used to map the surface topography of materials on nanometer to Angstrom scales by bringing a sharp probe close to a surface while monitoring the interaction between them. The earliest and simplest of the scanning probe family is the scanning tunneling microscope (STM), developed in 1981 by Binnig, Rohrer, and Gerber [173, 174]. The authors show that the STM is capable of imaging individual atomic positions in conductors and semi-conductors. Due to its relatively simple construction and superior resolution, STMs are commonly used in many fields including physics, material science, surface science, organic chemistry, and biology. Due to the basic principle by which they operate, STMs are unable to image insulating surfaces. For this purpose, Binnig, Quate, and Gerber went on to develop a related scanning probe microscope, the atomic force microscope (AFM) [175]. The AFM has the added advantage of being able to image insulating surfaces, however, AFMs typically have lower spatial resolution and their measurement signal is more complicated to model. Since the invention of the STM and AFM, several related microscopy modes have been developed to probe different features of the sample surface. For example, the STM can be used to perform spectroscopic measurements (STS) of the electronic energy states of the surface or sliding an AFM tip across a surface laterally may provide information about the surface adhesion. The following sections provide a brief introduction to the scanning probe methods used in this thesis starting with STM, STS, and AFM before moving onto conductive tip AFM modes,
Kelvin probe force microscopy (KPFM) and electric force microscopy (EFM).

### 3.1.1 STM and STS

A standard STM operates by bringing a sharp metallic tip within close proximity of a metallic (or semiconducting) surface so that they are able to interact at the atomic-scale. At close enough distances ($\approx 1\text{nm}$), the electronic wavefunctions of the surface and tip apex begin to overlap, allowing quantum tunneling to occur, passing electrons back and forth between them, generating a current.

Figure 3.1: **STM Principles of Operation** a) Schematic diagram of the STM measurement. A metallic tip is brought close to a metallic sample held at a bias voltage $V_b$. The resulting tunneling current is passed through a high gain amplifier and recorded digitally. b) The tunneling current measures the electronic transfer from the sample (left) for positive bias voltages or into the sample for negative bias voltages (right).
Considering many impeding electrons and electron states, the rate of transition can be quantitatively analyzed by relating to Fermi’s Golden Rule. The rate of transition from an initial eigenstate of a system described by a Hamiltonian $H_0$ into a continuum of final states described by $H = H_0 + H'$ is given by:

$$
\Gamma = \sum_f \frac{2\pi}{\hbar}|\langle f|H'|i\rangle|^2 \delta(E_f - E_i) \quad (3.1)
$$

where $|i\rangle$ and $|f\rangle$ are the initial and final state vectors and $E_i$ and $E_f$ are the initial and final energies of each respective state.

Now, to connect with the experiment, we consider the transition of electrons (at finite temperature, $T$) from sample states, $|s\rangle$, to the STM tip states, $|t\rangle$:

$$
\Gamma = \sum_t \sum_s \frac{2\pi}{\hbar}|M_{ts}|^2 \delta(E_t - E_s)(1 - f(E_t))f(E_s) \quad (3.2)
$$

where $f(E)$ is the Fermi-Dirac distribution function and $M_{ts} = \langle t|H'|s\rangle$ is the transfer matrix element, which comes from the overlap of the tip and sample wavefunction.

With an applied sample bias, $V_b$, the metallic tip hosts a continuum of electronic states at an energy $-eV_b$ with respect to the sample, so we integrate over the tip density of states, $\rho_t(E_s - eV_b)$:

$$
\Gamma = \sum_s \frac{2\pi}{\hbar}|M_{ts}|^2 \delta(E_t - E_s)(1 - f(E_s - eV_b))f(E_s) \times \rho_t(E_s - eV_b) \quad (3.3)
$$

Similarly, we take the sample states to make up a continuous density of states, $\rho_s(E_F)$, between the Fermi level, $E_F$, and the applied bias energy, $eV_b$:

$$
\Gamma = \int_{-\infty}^{\infty} \frac{2\pi}{\hbar}|M_{ts}|^2 (1 - f(E_F - eV_b + \varepsilon))f(E_F + \varepsilon) \times \rho_t(E_F - eV_b + \varepsilon)\rho_s(E_F + \varepsilon)d\varepsilon \quad (3.4)
$$

Then, the tunneling current should be proportional to $e$ times the probability of an electron transitioning from sample to tip minus the probability of an electron
transitioning from tip to sample:

$$I(V_b) \propto \frac{2\pi e}{\hbar} \int_{-\infty}^{\infty} \{f(E_F-eV_b+\varepsilon)-f(E_F+\varepsilon)\} |M|^2 \rho_t(E_F-eV_b+\varepsilon)\rho_s(E_F+\varepsilon)d\varepsilon$$  \hspace{1cm} (3.5)

Note that we have assumed the tunneling matrix to be symmetric such that $M_{ts} = M_{st} \equiv M$.

Next, we make some simplifying assumptions based on the experimental setup. In practice, the metallic tip is made from a material with a nearly constant density of states near the Fermi level ($\rho_t(E) \propto const$). Further, for low enough temperatures, the Fermi distribution function can be approximated by a step function. The tunneling current is then

$$I(V_b) \propto \int_{0}^{eV_b} |M|^2 \rho_s(E_F + \varepsilon)d\varepsilon$$  \hspace{1cm} (3.6)

Taking the simple model of a square potential barrier made by the vacuum gap between the tip and sample, the wavefunctions decay exponentially into the gap. Thus,

$$|M|^2 \propto e^{-\frac{2z}{\hbar}\sqrt{2m\phi}}$$  \hspace{1cm} (3.7)

where $z$ is the tip-sample separation distance, $m$ is the mass of the electron in vacuum, and $\phi$ is the vacuum barrier height. We have,

$$I(V_b) \propto e^{-\frac{2z}{\hbar}\sqrt{2m\phi}} \int_{0}^{eV_b} \rho_s(E_F + \varepsilon)d\varepsilon.$$  \hspace{1cm} (3.8)

There are two key features of equation 3.8 used in STM operation. Firstly, the tunneling current is exponentially sensitive to the tip-sample distance, $z$. This allows the tunneling current to be used as a sensitive probe of changes in the surface height. There are two methods used to map the surface morphology with STM: constant height mode and constant current mode. The most commonly employed method is constant current mode. In this mode, the tunneling current is fed through a feedback system which controls the tip-sample separation. The feedback moves the position
of the tip in order to maintain a constant current as the tip is raster scanned across the sample surface. By mapping the extension of the scanner, we obtain a map of the sample topography. Due to the use of a fast feedback system, this method prevents the tip from accidentally crashing into the sample, ruining both the tip and sample surface. However, the feedback system may introduce artifacts into the image for certain feedback parameters and cannot resolve changes in height which happen faster than the feedback loop rate.

In constant height mode, the scanner extension is fixed above the sample surface while the tip is quickly raster scanned, mapping the changes in tunneling current due to changes in the surface topography (as related through eqn. 3.8). Constant current mode has the advantage of operating at faster scan speeds without introducing processing artifacts, however, without monitoring the tip height during the scan, this mode is more prone to crashing the tip.

The second key feature of equation 3.8 is that the derivative of the tunneling current with respect to the bias is proportional to the sample density of states at the bias voltage:

$$\frac{dI}{dV_b} \propto \rho_s(E_F + eV_b).$$

(3.9)

Furthermore, the tunneling spectrum can be measured as a function of tip position to resolve the local density of states in space as well as energy ($\frac{dI}{dV_b}(r, E_F + eV_b)$). In this measurement, the tip is initially held fixed above the sample surface by a distance determined by an initial bias and current set point. The tunneling current is then measured as the bias voltage is varied. Alternatively, an additional, small AC bias can be applied to the sample and the differential conductance is measured by lock-in technique [173, 174, 176, 177, 178, 179].
3.2 Atomic Force Microscopy

3.2.1 Atomic Force Microscopy

Figure 3.2: **AFM Principles of Operation** a) Schematic diagram of an atomic force microscopy measurement. A flexible cantilever, with an attached sharp tip, interacts with a sample held on a piezoelectric scanner. The displacement of the cantilever due to the tip-sample interaction is measured using a laser and split photodiode b) Interaction force between the AFM tip and the sample surface. Far from the sample surface, the tip experiences a small, negative (attractive) force. This regime is where semi-contact AFM is employed. As the distance decreases, the force becomes repulsive. This is considered the contact regime

It is often beneficial to perform preliminary characterization of a device prior to lead fabrication or low-temperature STM and transport measurements. Atomic force microscopy (AFM) is a quick method useful for determining sample thickness, identifying defective regions, and probing surface quality. In AFM measurements, a sharp probe, attached to a cantilever, is brought into close contact with the sample surface. As the interaction forces between the tip and sample increase, the cantilever will bend. This displacement is then monitored using a laser beam reflected off the back of the cantilever shining onto an array of photodiodes (see Fig. 3.2). By moving the sample stage beneath the cantilever, one can map both the topographical and force interaction properties of the sample surface.
To avoid damage to the AFM tip and sample surface, a form of AFM called non-contact (also semi-contact or tapping) mode oscillates the AFM cantilever at or near its resonant frequency such that it taps the sample surface repeatedly as it is brought into contact. The amplitude and frequency of oscillation is modified by interaction with the surface and can be monitored by the laser and photodiode system. This method avoids the build up of frictional forces between the tip and sample and is thus less damaging to both surfaces. The topographical AFM data in this work is taken in non-contact mode to avoid damaging the sample surfaces.
3.2.2 Electrostatic AFM Measurement Modes

Figure 3.3: Principles of KPFM and EFM

a) During the two-pass measurement scheme, the first pass measures the sample topography and the second pass traces the topography but at a fixed height away to probe the long-range Coulomb forces.

b) The change in amplitude (left) and phase (right) of the AFM cantilever oscillations for a freely oscillating cantilever (black), a cantilever experiencing an attractive force (red), experiencing a repulsive force (blue) c) The principle of the Kelvin probe measurement. (Left) When the metallic tip and sample are electrically isolated, their Fermi energies are not aligned due to the differences in their chemical work functions. (Right) When electrical contact is made, a current passes between the tip and sample so that their Fermi energies align, charging their surfaces and generating a Coulomb attraction between them.
Standard AFM topography measurements interact with the sample surface via short ranged van der Waals forces. In the electrostatic measurement modes, the AFM tip is coated with a metallic film and retracted above the surface. Outside of the van der Waals regime, the long-ranged Coulomb forces dominate (figure 3.3a) and can be probed by applying a bias between the tip and sample. In this range, measurement of these forces probes the electrostatic properties of the sample surface, such as the locations of trapped charges and dielectric properties. In this work, electrostatic measurements are performed in a two-pass scheme, where the sample topography is recorded by non-contact AFM while on the second pass, the tip traces the topography but at a fixed distance above the surface. With this method, we employ two different measurement schemes known as Kelvin probe force microscopy (KPFM) and electrostatic force microscopy (EFM) described below.

In the second pass of KPFM, the gold-coated AFM cantilever is oscillated near resonance by applying a DC + AC voltage signal. This excitation generates an electrostatic force between the tip and the sample surface:

\[ F_{es}(z) = -\frac{1}{2} \frac{\partial C}{\partial z} (V_b - V_s)^2 \]  \hspace{1cm} (3.10)

where \( V_b = V_{DC} + V_{AC} \cos \omega t \) is the applied tip bias, \( \omega \) is the AC excitation frequency, \( V_s \) is the potential due to the work function difference between the tip and the sample surface, \( C \) is the local capacitance between the tip and the grounded backgate, and \( z \) is the tip-surface separation distance. Choosing the AC signal frequency to be near the resonant frequency of the cantilever leads to a large cantilever oscillation amplitude due to the electrostatic force component that oscillates at \( \omega \):

\[ F_\omega(z) = -\frac{\partial C}{\partial z} (V_{DC} - V_s) \cdot V_{AC} \cos \omega t, \]  \hspace{1cm} (3.11)

Thus, by means of a feedback circuit, we find the value of \( V_{DC} \) which nullifies the cantilever oscillation amplitude (\( F_\omega(z) = 0 \)). This value determines the KPFM signal,
It should be noted that the role of the bias in this measurement is to identify the local surface potential beneath the tip. The signal of this measurement mode is independent of the applied bias.

During typical KPFM measurements of metallic samples, electrical contact is made between the sample and the gold-coated tip, in this case, the work function difference between the two materials generates a static charge build-up at the sample surface. The measured surface potential is then simply related to the workfunction difference between the two materials, \( V_s = (-\Delta \Phi)/e \), where \( \Delta \Phi = \Phi_t - \Phi_s \) is the workfunction difference between the gold tip and sample and \( e \) is the charge of an electron.

Typically, HOPG is used as a calibration sample to determine the workfunction of the tip as the workfunction of HOPG is well-cited in the literature to be \( \approx 4.6eV \). After the workfunction of the tip is known, KPFM can be used to determine the workfunction of the sample.

We now turn to EFM. In this measurement mode the tip is mechanically resonated using a piezoelectric block with an excitation signal near the fundamental resonant frequency of the cantilever while a DC bias, \( V_b \), is applied to the tip. As a result of the electrostatic force between the tip and the sample, a phase difference is established between the driving signal of the cantilever and the measured oscillation (see 3.3). The phase shift is related to the local electrostatic force gradient by

\[
\Delta \phi = -\arcsin \left( \frac{Q}{k} \frac{\partial F_{es}(V_b)}{\partial z} \right)
\]

(3.12)

Where \( Q \) and \( k \) are the AFM cantilever quality factor and stiffness, respectively, and \( F_{es}(V_b) \) is the electrostatic force between the tip and sample surface 3.10.

As an imaging technique, EFM has the advantage of a higher intrinsic lateral resolution than KPFM. Because EFM measures the force gradient rather than the force itself, it is less susceptible to stray signals than KPFM and therefore can achieve
better spatial resolution. This is a direct consequence of the fact that $\frac{\partial^2 C}{\partial z^2}$ is focused more at the tip apex than $\frac{\partial C}{\partial z}$. The fact that the EFM signal is nearly quadratic in the applied bias makes it possible to easily tune the measurement sensitivity and image contrast [180, 181].

3.3 STM Design and Construction

While the design and construction of a homemade STM head is not itself an impactful contribution to modern physics, a considerable amount of time and thought went into the construction of this measurement system. The purpose of this section is to summarize the construction and design considerations for the homemade STM head used in this thesis work.

3.3.1 STM Tip

Scanning tunneling microscopy relies on the electronic transfer between a sample surface and a sharp wire tip. The images and structures observed by STM can be highly dependent on the tip material and the precise shape of the tip apex. Therefore, the tip is an extremely sensitive and important part of the microscope.

The STM tip material is chosen to have a nearly constant density of states at low energy so that tunneling spectroscopy (STS) reflects the density of electron states in the sample materials only. Tungsten and Platinum-Iridium alloy are commonly chosen metals for tunneling measurements because they are fairly inert and thus won’t naturally become contaminated with time and exposure to air. Both materials can be mechanically pulled and cut or electrochemically etched into atomically sharp points. The particular choice of material is up to the user. Tungsten wire is more stiff, resistant to chemical contamination during etching, and can be heated to high temperatures to clean. However, tungsten slowly oxidizes in air and is primarily chosen for use in high-vacuum and ultra-high-vacuum systems. Platinum-iridium
alloys are less prone to oxidation and can be cut or etched and left in air until used. PtIr wire is softer than tungsten, making it a bit harder to manipulate into the STM system and making it easier to melt and deform the tip apex with heating or contact with a sample.

No matter the choice of material, tips are fabricated from wire in similar ways. The "traditional" method of fabricating tips involves mechanically cutting the wire to produce a sharp point. Each STM user develops their own technique for reliably cutting atomically sharp tips but the general method is as follows: The wire is cut at approximately a 45deg angle using sharp, flat-edged hard wire cutters while simultaneously pulling the wire along its length. This combination produces a weak spot as you begin to cut the wire while the pulling motion teases the soft metal tip to a point. Applying too much force to the wire cutters results in a pinched tip with a large point while pulling too hard breaks the wire and produces a flat tip. One can learn to feel when they have applied the correct amount of pressure and cut a good tip.
Figure 3.4: Tungsten Tip Etching Procedure The supplies and setup required to electrochemically etch tungsten STM tips. A DC power supply gives 10-12V between a stainless steel ring (cathode) and the tungsten wire to be etched (anode) through a film of 2M NaOH solution.

A. DC power supply (10-12V)
B. 2M NaOH solution (20mL)
C. Stainless steel ring (cathode)
D. Tungsten wire (anode)

The electrochemical etching process depends on the metal chosen for your STM tip. Tungsten wire can be etched in sodium hydroxide (NaOH) (2 molar) under an applied DC potential of 10-12V. To produce an atomically sharp tungsten tip, we suspend a film of NaOH etching solution in a loop of stainless steel wire while passing a tungsten wire approximately 1cm through the loop (see figure 3.4). A 10-12V DC voltage is applied between the tungsten tip wire (+) and the stainless steel ring (-). The dynamics of the etching process are such that the tungsten wire gets etched in a conical fashion, forming a sharp apex at the point where the wire breaks. After a few (5) minutes, the 1cm segment of tungsten wire falls into a beaker of etching solution. The etched tip is then carefully removed from the etchant and rinsed thoroughly with DI water and dried with a nitrogen gas flow [182].

Etching platinum-iridium tips is performed in two steps using calcium chloride
(CaCl) in DI water as an etchant. In the first step, coarse etching of the tip, the STM tip wire is inserted approximately 1mm into the etching solution while an AC voltage (35V at 60Hz) is applied between the tip wire and a graphite counter-electrode. The current is measured through the tip and solution during the etching process and is stopped when the current decreases by roughly a factor of four (an experimentally determined value). By stopping the etching process before the tip wire is etched completely through results in a narrow neck (approximately 10um) formed near the surface of the etching fluid.

Figure 3.5: **Coarse Etching Supplies** Supplies required for the coarse etching step for PtIr tips includes a 1.M solution of CaCl₂, a small clamp to hold the tip wire, a graphite counter-electrode, an AC power supply, and AC ammeter to monitor the current. (Bottom) Optical microscope images of the PtIr tip wire after completing the coarse etching step. The goal of the coarse etching step is to create a narrow neck in the PtIr wire.
In a second step of the etching process, the narrow neck is carefully thinned by locally etching it using a wire loop beneath a stereo microscope. The tip is inserted into a capillary mounted on a micro-manipulator stage and a wire loop made of tungsten is wet with a droplet of calcium chloride etching solution. The wire loop and capillary are connected to a power supply (3V AC at 60Hz) and the micro-manipulator is used to advance the neck of the coarse-etched tip wire into the etchant droplet. The etchant droplet is replaced every few seconds during the etching process which continues until the etching makes it through the neck, forming an atomically sharp tip. The tip is then dipped in hot DI water for one minute to clean.

Figure 3.6: **Fine Etching Supplies** The fine etching step requires the coarse etched tips to be held by a micromanipulator under an optical microscope. The coarse etched tip is passed through a small, tungsten loop coated which is coated with a droplet of etching solution while a small AC voltage (3V) is applied across them. The micromanipulator allows one to further etch the neck portion of the coarse etched tip until a sharp point is formed.

The type of tip used does not strongly impact the quality of tunneling data and should be chosen based on the experiment as each has its own pros and cons. When
UHV and high vacuum can be used, an etched tungsten tip will be the best choice because of its stability, hardness, and reliably sharp apex. However, tungsten tips oxidize in air and need to be annealed at high temperatures. For measurements in air, cut PtIr tips can be produced quickly by a skilled hand, but have unreliable, irregular tip shapes. Etched PtIr tips have a reliably conical apex but are fragile and typically more dull than etched tungsten (though this can be improved using fancier fabrication methods).

3.3.2 Scanning

To achieve atomic-scale resolution, scanning probe measurement systems utilize the fine motion achievable by the piezoelectric effect. The piezoelectric effect is the deformation of a crystal under an applied electric field (or vice-versa). A piezoelectric effect is common in materials which exhibit a dipolar unit cell. Similar to magnetic dipoles in ferromagnets, nearby dipoles within the material align with each other forming ordered Weiss domains. Naturally occurring Weiss domains are not necessarily aligned with each other, however, by applying a large electric field across the material significant fractions of the material become electrically aligned with the applied field (or "poled"). As the poled crystals are strained, the crystal structure gets deformed causing a shift in the polarization of the embedded dipoles within the crystal, generating an electric field. Similarly in the converse piezoelectric effect, applying an external electric field changes the orientation of the dipoles, physically changing the shape of the crystal.

Piezoelectric materials occur naturally (like quartz) and can be man-made (usually ceramic). Man-made piezoelectric ceramics tend to be more efficient than their natural counterparts, but their crystal structure is complex and difficult to produce. The most commonly used piezoelectric ceramic is Lead Zirconate Titanate (PZT) due to its efficiency and its low cost).
Figure 3.7: **Scanner Piezo** a) Simulated model of the scanner tube assembly; combining two piezoelectric tubes to decouple the x,y, and z directions b) An optical image of the piezoelectric tubes and the capillary which holds the STM tip within a Macor sleeve c) A top-down drawing of the piezoelectric tube showing the radial polarization with black arrows d) The split scanner tube can be bent along different directions by applying different combinations of voltages to the four quadrants

The piezoelectric scanner in our homemade STM is compound in design, using two concentric piezoelectric tubes to extend its measurement range and to easily separate lateral and vertical degrees of freedom. The tubes are radially poled so that applying a voltage between the inside and outside of the tubes causes the entire tube to extend or retract. The extension of the tube as a function of applied voltage can be determined from the geometry of the tube (diameter, $D$, wall thickness, $t$, and length, $L$) and the piezoelectric coefficient of the tube material, $d_{31}$:

$$
\Delta L = V \cdot \frac{d_{31} L}{t}.
$$

(3.13)

Using the geometry material properties of the tubes in this system, we have $\frac{\Delta L}{V} \approx 2.85nm/V$.

By segmenting the outer electrode of the inner-most piezo-electric tube (scanner tube) into four quadrants, we are able to separately apply voltages to different sides
of the tube, causing the tube to deflect in the x-y plane (see figure). The amount of lateral deflection of the end of the scanner tube is given by:

\[
\Delta r = V \cdot \frac{\sqrt{2}d_{31}L^2}{\pi Dt}.
\]  

(3.14)

Given the geometry of the tube scanner used in our system, we get a lateral displacement per volt: \(\frac{\Delta r}{V} \approx 14.6 \text{nm/V}\). Note that the piezoelectric coefficient is a temperature-dependent material property. The values given here are the room temperature reported values. The piezo coefficient decreases monotonically, but non-linearly as temperature decreases such that it reaches approximately half of its room temperature value at liquid nitrogen temperature (77K) and approximately one fifth at liquid helium (4K). Atomic-scale measurements should be made at these temperatures using a material with known lattice spacing (like HOPG) to determine the precise piezoelectric coefficient.

### 3.3.3 Stepper Motors

The lateral position of the sample (with respect to the tip) and the vertical position of the tip is controlled using homemade piezoelectric inertia motors. This type of motor is chosen due to their simple operation principle, simple electronic control, their ability to be miniaturized, and their small intrinsic spatial resolution. In general terms, our motors are made from a fixed piezoelectric actuator with a movable slider which holds the STM sample. The actuator includes three stacks of piezoelectric slabs (PZT 8) capped with an alumina wearplate, of which two are fixed and the third is held against the slider using a beryllium-copper (BeCu) spring and a ruby ball (for point contact). The faces of the slider are protected with sapphire which presses against the alumina wearplates of the actuator.

The motors move via the stick-slip method in which a voltage is used to first slowly deform the piezo-stacks, pushing the slider forward by friction. Then, the
voltage polarity is rapidly reversed, causing the alumina wearplates to slide against the sapphire on the slider. This repeated motion causes the motor to step along one direction.

Figure 3.8: **Stepper Motors** a) A design for the miniature stick-slip motors used to control the sample stage position b) A side view of (a) c) A diagram of one of the piezoelectric stacks used in the stick-slip motors shows the polarization and electrode placement of each layer d) The piezoelectric stack shown in (c) after a voltage is applied between electrodes. The induced displacement of the free end of the stack is labeled ‘s’ d) A cartoon depiction of the voltage signal applied to the piezo stacks during stick-slip operation. A slow voltage increase deforms the stack while friction moves the slider (stick action). Suddenly, the voltage sign is reversed, causing a rapid deformation of the piezo-stack in the opposite direction. The rapid deformation should overcome static friction, allowing the piezo-stack to slide along the surface of the slider (slip action). This process is repeated rapidly to cause the motors to move (several mm)

The stick-slip motors are first performance tested and calibrated on a bench top. The motor step size should be remeasured after changing the spring tightness or
cleaning the motor surfaces. First, test the motor condition by measuring their capacitances. Open or shorted circuits are likely due to wiring issues such as broken down insulation between two wires or broken wire contacts. Sudden changes in capacitance values indicate a more serious problem with the piezo motor stacks. Make sure that adjacent electrodes within the stack are not shorted to each other and check that the piezo-ceramic crystals are not physically broken.

Next, with the motor spring at its loosest position (while still able to move), walk the motor back and forth across its full range of motion to make sure that the slider does not get stuck, slow down, or change frequency at any position. Repeat this test while gradually tightening the springs (try not to deform the beryllium copper springs). Issues with the motor motion are typically due to a) a contaminated junction between the alumina and sapphire plates, b) the motors are physically blocked by wires, c) springs are too tight or too loose, or d) the spring clamps are too deformed to apply enough pressure.
Figure 3.9: **Stepper Motor Calibration**

a) Stepper motor step size versus spring tightness for the X motor both in the forward and backward directions. Motors stop moving when the spring is too loose or too tight with the step size maximized at some point between. b) X and Y motor step size versus peak driving voltage in the forward and backward directions. Hysteresis can be minimized by applying different driving voltages to move the motors in different directions.

By measuring the step size as a function of spring tightness, one can (sometimes) find an optimal point where the step size is maximized (figure 3.9a). This spring tightness is chosen as it will minimize the chances of the motors getting stuck at reduced temperatures.

After the spring tightness is optimized, measure the step size(s) as a function of driving voltage. The step size in opposite directions are typically not the same due to subtle differences in the contact junction between the motors and the slider. By controlling the driving voltages in different directions separately, we can reduce the asymmetry between step sizes in different directions (figure 3.9b).
3.3.4 Housing Design

Our homemade STM head is designed with a segmented body so that parts can be easily interchanged and upgraded to accommodate different experiments. The main segments include the Z-motor housing which holds the tip and its positioners, and the sample housing, which holds the sample stage and its positioners.

Figure 3.10: STM Head Overview Computer design of the assembled STM head (left) next to an image of the real, assembled head (right). The upper chamber houses the Z-motor and the STM tip while the lower chamber houses the sample and XY-motors. A zoomed in image shows the tip near the sample inside the STM head.

Inside the Z-motor housing, we have a sapphire prism which holds the piezoelectric tubes that control the fine motion of the tip during scanning as well as the STM tip holder and tip. The sapphire prism itself acts as the slider for the coarse z-direction stick-slip motors which bring the tip into close proximity with the sample surface. The sapphire prism is held by friction with a BeCu spring at two ruby ball point contacts. It is essentially an elongated version of the X and Y-direction stick-slip motors described in the previous section.
This body segment is designed so that the tip control motors (and their electrical attachments) are completely contained within the sapphire prism and the coarse positioners are attached to removable titanium walls. This way, we are able to replace and repair the coarse positioner motors as they may wear over time. More importantly, we are able to easily remove and replace the scanner for different experiments. For example, we have constructed a low-cost, 3D-printed prism out of sapphire-plated titanium with half-length scanner tubes for use as a quick, inexpensive, room-temperature scanner. Additionally, we have proposed the future implementation of a tuning fork.
based SPM attachment for frequency-sensitive experiments and measurement of non-conducting surfaces which utilizes the versatility of this head design.

The sample housing contains two linear stick-slip positioners (described in the previous section) which are stacked perpendicular to each other to move the sample in two Cartesian directions (labeled X and Y). Together, we called these motors the "XY-motors." On top of the XY-motors sits the sample stage.

Figure 3.12: Sample Housing Computer design of one of the assembled stick-slip motors (left, top), the sample stage (left, bottom), and the fully assembled sample housing (right)

The sample stage is a gold-coated copper block which contains a constantan wire heater and platinum film thermometer to control and monitor the sample temperature. A typical sample is glued to the surface of a sapphire plate which sockets into the sample stage and is held fixed with two BeCu spring clamps. The sample is then electrically connected to pads on the sample stage designated for bias and backgate control.

Control and organization of the wiring of the STM head is not only beneficial for testing and assembly. It is necessary for maintaining low electronic and mechanical noise levels. Wire anchor points are designed at the top and bottom of the heads to pin the Z-housing and sample housing wires (respectively). The scanner tube connections and the shielded tunneling line are anchored to the top of the sapphire
prism and the coarse Z-motor wires are attached to the top of the Z-motor housing. It is important to prevent the shielded tunneling line from making contact with any other wires to prevent excess noise generation. The XY-motor connections, stage heater, thermometer, backgate, and bias connections are attached to a copper-clad Kapton film on a macor plate at the bottom of the head. The wiring from within the head passes through small holes on the bottom plate of the sample housing and are rigidly anchored to the plate before attaching to pin connectors. The use of these pin connectors makes replacing the sample and repairing the sample stage much easier than previous designs. In this design, the bottom plate connector can be detached from the sample housing and the sample, sample stage, and XY-motors can be removed from the head along with it. The sample can then be adjusted elsewhere before being replaced into the head.

3.4 Noise Testing

3.4.1 Signal Control Theory

During a constant-current STM measurement, the tunneling junction current is maintained by a computer-controlled feedback loop. A simple model is illustrated in figure ???. A constant current set point is compared to the measured tunneling current (which depends on the tunneling junction stability) and the difference is recorded as an error signal. The error signal is fed to a digital proportional-integral (PI) feedback loop which then adjusts the voltage applied to the piezoelectric scanner of the tunneling microscope in order to reduce the value of the error signal.

To analyze an STM signal, we consider the Laplace transform of a time-dependent input function, $f(t)$, defined as

$$F(s) = \int_0^\infty f(t)e^{-st} dt$$  \hspace{1cm} (3.15)
where $t$ is time and $s$ is a complex-valued frequency. As a side note: notice that the Fourier transform is simply related to the Laplace transform for the cases where $s \rightarrow i\omega$ with $\omega \in \mathbb{R}$. Where the Fourier transform decomposes a function into a basis of sinusoidal functions, the Laplace transform decomposes a function into a basis of \textit{exponentially decaying} sinusoidal functions where the real part of $s$ determines the decay (or growth) time.

Thanks to the convolution properties of the Laplace transform, the frequency-dependent response of the tunneling junction, the digital feedback control, and electronic filtering can be treated separately (if we assume them to be components of a linear system). The response of a component of such a feedback loop can be characterized by its transfer function, the output due to an impulse excited at time, $t = 0$. 
Figure 3.13: **Control Flow Diagram** a) Propagation of the PID control signal through different elements of the circuit. The current signal is compared with the current set point to create an error signal which is fed into the digital PI control circuit. The PI control signal is fed through a high voltage amplifier which controls the piezo tube extension. At a fixed bias, the piezo tube extension is converted into a tunneling current which is then amplified before returning to the control system b) The response of a PI feedback loop to a step function input at time $t = 0$ for different integral gain values. Higher gain values cause the set point to be reached quicker but overshoots the set point more, initially.

The input of the feedback loop is the time-dependent error signal, $E(t) = I_{SP} - I_e(t)$, which is first fed through a first-order low pass filter. For a linear system, the only frequency-dependent part of our control system is the low pass filter. The transfer function for a simple low pass filter with cut-off frequency, $\omega_c$, is $G_{LP}(s) = \frac{\omega_c}{s + \omega_c}$. By
the convolution property of the Laplace transform, the filtered error signal can be written as a product of the Laplace transform of the original error signal and the transfer function of the low pass filter:

\[ E'(s) = E(s)G_{LP}(s). \] (3.16)

The filtered error signal is then fed into a feedback system using a proportional and integral control function. That is, given a time-dependent, filtered, error signal as an input, \( E'(t) \), the feedback loop, \( V(t) \), output has the form

\[ V(t) = G_P E'(t) + G_I \int E'(\tau) d\tau, \] (3.17)

where \( G_P \) and \( G_I \) are constants which determine the gain of each term in the control function. The Laplace transform of this equation gives an output that looks like

\[ V(s) = \left(G_P + \frac{G_I}{s}\right) \cdot \frac{\omega_c}{s + \omega_c}. \] (3.18)

The feedback parameters can then be used to tune the response to signal disturbances [177]. For illustration, the step response of this model system is plotted in figure 3.13b taking a cutoff frequency of \( \omega_c = 250Hz \) and changing the integral gain value to a percentage of its optimized value.
Figure 3.14: **Control Flow with Noise** a) The control flow diagram given in figure 3.13a with the mechanical $H_n(s)$ and electronic $e_n(s)$ noise signals added b) Power spectrum of the open-circuit current signal with strong electronic and mechanical noise sources nearby

For a higher detailed model, the control system should include multiple inputs to account for noise which gets coupled into the system from outside of the control electronics (see figure 3.14). Modeling such a system can be done using MATLAB, SPICE, or other numerical methods but will not be performed here. Instead, we simply aim to minimize the noise coupling into the system rather than characterize it. There are two primary ways that unwanted disturbances enter the tunneling signal: mechanically and electrically. Mechanically coupled noise ($H_n(s)$) arises from ground and acoustic vibrations which travel from outside of the STM system to the tunneling junction. Due to the exponential dependence of the tunneling current on tip position, even small mechanical vibrations can lead to large error signals in
the tunneling current signal. Electronically coupled noise \((e_n(s))\) primarily arises from both radio-frequency AC electric fields emitted by nearby electronics and power sockets which capacitively couple to the STM tunneling line and fluctuations of the reference ground potential.

Noise in the tunneling line can be separated into sources and coupling mechanisms and the effects of each should be mitigated as much as possible. The noise functions \(H_n(s)\) and \(e_n(s)\) can be decomposed:

\[
H_n(s) = \sum_i T_i(s)W_i(s)
\]

\[
e_n(s) = \sum_j c_j(s)w_j(s)
\]

where \(i\) and \(j\) index mechanical \((W_i(s))\) and electronic \((w_j(s))\) noise sources and \(T_i(s)\) and \(c_j(s)\) characterize the degree of noise coupling into the system i.e., the mechanical transfer function and the source-tunneling line capacitance. Experimentally, we do not measure the transfer function of these noise components as we do not know the form of the noise source. Instead, we can monitor the effect of noise on the open circuit current, \(I(s)\), without the PI feedback engaged. Figure 3.14b shows an example measurement of the discrete, finite-time Fourier transform of the open circuit current noise. During this measurement, all vibration isolation has been shorted and radiation shielding has been removed. Additionally, a RF noise source (a fluorescent lamp) is placed near the STM head to show the effect of radio noise. The main features to note are a broad mechanical resonance around 6776Hz (this value depends on the stiffness of the XY motor springs) as well as many sharp peaks at all integer multiples of 60Hz. In the following sections, I outline methods to reduce the effects of mechanical and electronic noise on the STM measurement system by removing nearby noise sources, reducing the coupling into the STM system, and by damping any resonators within the system.
3.4.2 Control System Outputs

Electrical noise is the easiest to monitor and identify as the different possible contributions can be separated and measured independently. Here, we aim to measure and reduce the effect of transmitted and intrinsic noise in the system. The largest intrinsic electronic noise contribution comes from the preamplifier circuit due to Johnson noise. Additionally, this circuit is designed to amplify small variations in the tunneling line current and is thus highly sensitive to electronic noise which is transmitted through the tunneling circuit. The first step to measuring the system’s electronic noise floor is to measure the DC bias output. The output should be very stable and accurate as the bias determines the energy resolution of the STM’s spectroscopic measurements. The bias output is measured using an oscilloscope to verify the peak-to-peak noise meets our expectations. Next, we determine the response which comes from the preamplifier itself. To do this, we utilize a ”tunnel gap simulator” (a large resistor). We apply a bias voltage across this resistor and use the preamplifier to measure the current. Given that the applied bias signal is stable, any noise peaks that appear in the current or its Fourier transform must be induced by or coupled in through the preamplifier. It is crucial that any noise coming through the preamp is minimized through proper shielding and grounding of the preamp circuit (see the next subsection).

While voltage noise coupled through the tunneling line and preamplifier primarily limits spectral resolution, noise in the high voltage (HV) amplifiers, which control the scanner motion, primarily limits spatial resolution. Using equation 3.14, we can establish a voltage noise tolerance for atomic resolution. Assuming that we require at least 5 data points per atom to achieve decent resolution, and taking the atomic spacing of the desired sample, $a$: $\Delta r \leq \frac{a}{3}$ gives $|V_{\text{noise}}| \leq \frac{\alpha \pi D t}{5 \sqrt{2d_1 L^2}} \approx 3mV$ (assuming the sample is HOPG, $a=0.246\text{nm}$). This level is achievable with commercial control electronics, thus, we assume that our resolution is not limited by the geometry and electronic control of our chosen piezo tube scanner.

Even with optimized control electronics, electronic noise may sneak in through
the reference ground. If the ground of the system is not clean or large in capacity, all signals will be susceptible to noise and sparking. We have seen samples burn when an STM user accidentally shocks their desk on a dry winter day. Ideally, the ground source should be large enough such that no effect of shocking the desk is seen at the tunneling junction. For this, the metallic machine body is grounded through a thick braid of copper to the copper pipes of the building. In any case, if higher resolution is desired, either in space or in energy, voltage attenuators can be employed at the outputs of the high voltage boards or the bias. Attenuators can reduce the noise which enters the system but have the drawback of reducing the maximum applied voltage.

When the outputs of the tunneling line pre-amplifier and high voltage amplifiers have been verified to exhibit no large, frequency-dependent noise, we can approximate their contributions to the feedback circuit simply as their DC gain values. For example, the high voltage amplifiers amplify their input from the feedback system (in Volts) by a factor of 15, therefore, we take \( G_{HV} = 15 \). Similarly, we take \( G_{PA} = 1V/nA \) from the preamplifier circuit. For feedback frequencies below any resonances of the piezoelectric scanner tube, we also take \( P_z = \Delta L/V \approx 2.85nm/V \) from equation 3.13 and evaluating the parameters for the STM head used.

### 3.4.3 Grounding and Shielding

After verifying the stability of the measurement circuit, we measure the noise coupling in through the STM wiring: interference noise. For obvious reasons, the most sensitive wires are the tunneling line and shield followed by the bias line and shield. To reduce the effect of electronic noise pickup through the tunneling line, it should be continuously shielded with the capacitance between the shield and the tunneling line to be as small as possible. To reduce any mechanical noise, the tunneling line needs to be firmly anchored to the machine body. The tunneling signal noise can be monitored with the circuit open by simply monitoring the FT spectrum. With the circuit open,
periodic noise observed is due to both electronic pickup from AC power lines and
nearby electronic equipment as well as mechanical resonances from poorly anchored
tunneling lines. The same procedure is performed for the bias signal line.

Electronic noise sources which are picked up by the tunneling line include: RF
fields from AC power lines, nearby electronic equipment (including lamps and cell
phone chargers), and ground loops within the circuit. As mentioned in a previous
section, interference noise can be broken into three elements: the noise source, cou-
pling mechanism, and the receiver [183, 184]. For example, the rapidly changing, large
voltages used to driven the piezoelectric scanner may source noise which is capaci-
tively coupled to the tunneling current line. We can then work to reduce the effect
of both the source and the coupling, i.e. reduce the voltage range of the piezoelectric
scanner and adding a dielectric medium between the scanner and the tunneling line.

Due to the sensitivity of the tunneling line, it is very important to properly shield
this line as well as any high voltage signals entering the system. The key to shielding
a wire without generating more noise is to keep the shield as continuous as possible
and be sure to ground the shield at one location only. If a break in the shield needs
to be made, the shield segments should be connected together, not each connected to
ground. A shield which connects to ground in multiple locations can generate currents
within the shield itself and can act as an antenna, picking up RF noise. These often
show up as "ground loops" and generate noise at 60Hz and its harmonics. Thus, it is
not always beneficial to ground extra wires or wrap electronic components with metal
foil.
3.4.4 Vibration Isolation

The largest and most detrimental noise components often come from mechanical vibrations. In order to image atoms at a material surface, the relative motion between the tip and sample surface must be minimized. To this effect, STM heads are often
designed to be very rigid, however, some residual motion always persists. We can model the tip-sample junction as a damped oscillator with resonant frequency, $\omega_t$. The STM head should be constructed rigid enough so that $\omega_t$ is much larger (several orders of magnitude) than the frequency of vibrations entering the system. Typical floor vibrations are below 2Hz, therefore, having $\omega_t > 2$kHz is required. A schematic cartoon of the tip-sample stage junction is given in figure 3.15a. The junction is modeled as a spring with stiffness $k_t$ and damping coefficient $b_t$. Given the effective mass of the tip, $m_t$, the motion of the junction separation ($x_t - x_h$) is governed by the differential equation:

$$m_t \ddot{x}_t + b_t (\dot{x}_t - \dot{x}_h) + k_t (x_t - x_h) = 0$$

(3.20)

which defines the characteristic resonant frequency $\omega_t = \sqrt{\frac{k_t}{m_t}}$ and damping ratio $\gamma_t = \frac{b_t}{2m_t}$. If we assume the STM head is driven such that $x_h(t) = Ae^{i\omega t}$. Then the transmission function, describing the motion of the tip-stage junction as a function of frequency, is given by

$$|T_1(\omega)| = \left| \frac{x_t - x_h}{A} \right| = \left| \frac{\omega^2}{\omega_t^2 - \omega^2 + 2i\gamma_t \omega} \right|$$

(3.21)

Equation 3.21 is plotted in figure 3.15b for different values of the damping ratio $\gamma_t$. Vibrations entering the system which have a lower frequency than $\omega_t$ are damped while higher frequencies are transmitted to the junction. When the driving frequency is near the resonant frequency, vibrations are amplified unless damping is implemented.

Further reducing the vibrational noise requires reducing the amplitude of the motion of the head. This is done by removing the STM head from all noise sources with a vibration isolation stage. This stage of vibration isolation has its own resonant frequency determined by the mass of the STM head, $m_h$, and the stiffness of the springs used. Vibration isolation is performed by suspending the STM head from a cryogen refrigeration system ("fridge") using spring bellows. Similar to the tip-stage
junction, the fridge-head junction is modeled as a damped harmonic oscillator with
stiffness $k_F$ and damping coefficient $b_f$ (figure 3.15c) following the equation of motion

$$m_h \ddot{x}_h + b_f (\dot{x}_h - \dot{x}_f) + k_F (x_h - x_f) = 0$$

(3.22)

defining $\omega_h = \sqrt{\frac{k_h}{m_h}}$ and $\gamma_h = \frac{b_f}{2m_h}$.

Now, assume the cryostat (fridge) vibrates with an amplitude $A'$ and frequency
$\omega$. Then, the transmission from the fridge (amplitude $A'$) to the head (amplitude $A$)
is given by:

$$|T_2(\omega)| = \frac{|A|}{|A'|} = \left| \frac{\omega_h^2 + 2i\gamma_h \omega}{\omega_h^2 - \omega^2 + 2i\gamma_h \omega} \right|.$$ 

(3.23)

The transmission from the ground to the tunneling junction will be the product of
the transmission function for the head and for the isolation stage $T_1 \cdot T_2$ (figure 3.15d).
Here we choose the resonant frequency of the isolation stage arbitrarily to be $\omega_h = \omega_t / 1000$ and choose a range of values of the head damping ratio $\gamma_h$ (plotted in figure 3.15d). Adding the isolation stage drastically reduces the transmission of ground
vibrations to the tunneling junction at frequencies above $\omega_h$ but adds a resonance
peak there. Damping the isolation stage reduces the resonance peak at the cost of
increasing transmission above $\omega_h$ (compared to the no damping case) [177].

Adding more stages of isolation continues the trend. Further isolation reduces
transmission at frequencies above the resonant frequency of the isolation stage while
adding peaks there. In addition to having spring bellows between the fridge and STM
head, the fridge is mounted on a one ton steel plate on a cushion of compressed gas.
This heavy, soft vibration system has a low resonant frequency (< 0.2Hz), filtering
out most ground vibrations. The bellows then act to further reduce transmission
of acoustic vibrations which might couple into the system by resonating the fridge.
Similar to electronic noise testing, the mechanical noise can be monitored using the
power spectrum of the open-circuit current as well as the tunneling current with
feedback in place. Performing such tests with different isolation stages shorted can
help identify noise sources and coupling paths within the system.

3.5 Capacitive Alignment

3.5.1 Capacitive Sensing

Following graphene device fabrication methods, other 2D materials are commonly exfoliated from bulk crystals prior to and transferred onto insulating substrates. These flakes are typically only a few microns across and a few nanometers in thickness creating a challenge in aligning the targeted sample and STM tip. Further, the insulating substrate is invisible to the microscope and approaching its surface may damage or contaminate the tunneling tip. To align the STM tip with a conducting sample on an insulating substrate, we use the STM as a local capacitance sensor to detect the location of conductors on the substrate.

To use the STM system as a capacitance sensor, we apply an AC bias voltage between the sample bias and the STM tip at a frequency, $f$. Due to the capacitive coupling between the STM tip and the conducting sample, a current is induced at the applied frequency which is out-of-phase with the bias.

\[ I = i2\pi fCV \]  

The current measured at the STM tip is fed through a lock-in amplifier tuned to the driving frequency. The out-of-phase component of the lock-in output is monitored while the sample moves beneath the STM tip using stepper motors. Then, the measured capacitance as a function of space can be used to determine the positions of conductors beneath the tip. Figure 3.16a shows a cartoon depiction of the measurement scheme (left) and an optical image of the tip-sample junction.

Unfortunately, the name "scanning capacitance microscope" (SCM) has been taken. It describes a different (AFM-based) technique that also measures capacitance
as a function of space through similar principles. To distinguish between normal STM operation and capacitive sensing mode, I will label this operation mode "capacitance sensing microscopy" (CSM).

Typical frequencies used are in the 1-5kHz range and voltages are 100mV-10V. The value of capacitance is typically in the pF-nF range with changes in the vicinity of the electrode on the order of fF-pF. Thus, typical AC currents sit at the pA to nA scale; where the STM amplification system is already tuned for tunneling measurements. This makes the STM measurement system easy to convert into a capacitance sensor.
To gain a qualitative understanding of the principle behind the method, we adopt a simplified geometric model of the measurement. During the capacitive approach procedure, the STM tip is kept several microns above an evaporated gold pad roughly 50nm thick. We thus take the gold pad to be of negligible thickness. Additionally, we assume that the gold pad is long enough such that the length of the pad has far less effect on the change in capacitance than the pad width. Since the electric field and charge density is largest at the tip apex, we consider the tip to be a point probe.
Finally, we assume that the backgate present in many devices is spatially uniform and will contribute only a uniform offset to the measured capacitance. Under these assumptions, we can model the measurement as a point probe above a line of gold with width, L (See Fig 3.16c). When a bias, V, is applied between the gold pad and the tip, a charge, Q, gathers in the line of gold such that a uniform charge per unit length, \( \lambda \), is established. The capacitance between the point probe and the line of gold is defined as \( C = \frac{Q}{V} \), where:

\[
V = \int_{-L/2}^{L/2} \frac{1}{4 \pi \epsilon_0 \sqrt{x^2 + z^2}} \lambda \, dx.
\]

We find that:

\[
C = 4 \pi \epsilon_0 \ln \left[ \frac{x + \frac{L}{2} + \sqrt{(x + \frac{L}{2})^2 + z^2}}{x - \frac{L}{2} + \sqrt{(x - \frac{L}{2})^2 + z^2}} \right].
\] (3.25)

The modeled capacitance and spatial derivative of the capacitance are plotted vs position in units of pad width, L, at a distance above the pad \( z = 0.3L \) in Figure 3.16c. Figure 3.16d displays the measured data, showing remarkable agreement with the model. The key features are the peaks in the capacitance curve and its derivative. For a constant background, the peak in capacitance should occur precisely at the center of the gold electrode and the inflection points (maxima and minima in the derivative) occur near the edges of the gold pad. Thus, we can use these features to measure the size and locate the position of the conductors on an insulating substrate.
3.5.2 Alignment Algorithm

By designing electrodes to include a staircase of decreasing sized gold pads, we can guide the STM tip toward a micron-sized sample with a reduced risk of damaging the tip or sample. Next, we outline the approach procedure and provide some tips for focusing the capacitance microscope.

STM guiding electrodes are designed with several different sized gold pads linked to each other with at least a factor of 3 difference between adjacent pad widths. This ensures that we can tell the difference between the nearest gold pad and the next pad. The electrodes are usually written using electron beam lithography and gold is deposited using thermal evaporation; however, such high precision is not necessary. Photolithography or a shadow mask could also be used to write the guiding electrodes with sufficient precision. The capacitance sensing microscope (CSM) is
highly sensitive to conductors which protrude out of the sample surface. Thus, care has to be made when making electrical contact to the guiding electrodes. Typically, a thin, aluminum wire is wire-bonded to the largest (500µm) gold pad. Using silver paint and a thick wire generates a large background capacitance and may obscure the locations and distort the observed shape of the gold pads.

STEP 1

After electrically connecting the guiding electrode and loading into the STM head, a long working distance microscope is used to visually align the STM tip with the large gold pad (∼200µm) and approach the tip to within ∼500µm of the surface. Using the CSM-mode, move the sample stage so that the tip passes across the gold pad and locate the peak in capacitance. With the tip still far from the gold pad, the change in capacitance across the gold pad may be small (the bump will be broad) and it may be difficult to locate the pad edges precisely. This is ok. We will need to assume that the local maximum in capacitance coincides with the center of the largest gold pad. If possible, this should first be performed with visual access to the tip-sample junction to verify the position of the tip with respect to the gold electrode. Now, we approach the STM tip to the pad center with the STM feedback engaged. We will land on the gold pad to be able to calibrate the coarse motor in the Z-direction as well as maintain a known distance from the gold pad during capacitance measurements.

The user might find it beneficial to monitor the capacitance as the tip approaches the sample surface. As the tip-sample distance decreases to become comparable to the STM tip size (the conical part), magnitude of the capacitance (and its derivative) grows rapidly. The first and second derivatives of the capacitance can be used to speed up the approach process by setting a critical value above which the feedback is engaged [185, 186].

At the surface of the large gold pad, the coarse motor step size can be calibrated. Assuming the extension of the piezoelectric scanner tube is accurate, take coarse steps toward and away from the surface (one at a time). The change in scanner extension
during one step measures the step size of the coarse motor. If this is the first time you are testing the coarse motor, start by using small stepping voltages. The step size of the coarse motor should be approximately less than one quarter of the total scanner extension range to avoid accidentally crashing the tip into the sample. For reference, the motors used in our setup show $\approx 150\text{nm}$ per step at a driving voltage of 100V with a sweep time of 5ms at room temperature. Keep in mind that the step size not only depends on driving voltage and sweep time, but is highly dependent on the loading force applied to the piezo motor, the cleanliness of the stick-slip interface, temperature, and vacuum level. Additionally, the step size is not necessarily linear in voltage, so calibration should be performed at each different driving voltage used.
Figure 3.18: Capacitive Approach Procedure Three successive steps of the capacitive approach procedure. Design of the gold pad is shown on the left with the path of the STM tip above the sample surface during each measurement. The right shows the capacitance and its derivative as a function of position of the tip with respect to the sample.

STEP 2

The next step is to re-focus the capacitance microscope away from the sample surface. The tip needs to be retracted far enough away from the surface to avoid accidental crashing while it moves but not so far as to lose the features of the gold pad. From figure 3.16, we can see that at a height of 0.3 times the pad width, the edges of the pad are clear and the error in the measured pad width is still small (∼6%). Thus, we choose to retract the tip approximately 0.3 times the pad width.
STEP 3

Now, we measure the width of the pad along the X-direction in steps. The capacitance and its derivative is measured as a function of number of steps. As presented in figure 3.16 and equation 3.25, the distance between inflection points (in number of steps) can be used to calibrate the coarse motor step size. Record this value and move the tip back toward the center of the pad, stopping just inside of the gold pad edge (see figure 3.18). The width of the gold pad is known from the design and the error in the width measurement is estimated from the line-charge model presented in the previous section. The step size calibration is given by

\[
\text{step size [nm/step]} = 1.06 \cdot \frac{\text{pad width}}{\text{steps between inflections}}.
\]  

(3.26)

Repeat this step along the Y-direction for calibration. Be careful: there is often a background slope in capacitance along the Y-direction due to the presence of the other gold pads and bonding wires. This background may obscure any sample tilt and may lead to a tip crash. The slope in capacitance also makes it difficult to precisely determine the center of the gold pad along this direction. However, a slope in capacitance translates to a simple offset in the measurement of the derivative of capacitance, \( \frac{dC}{dy} \), so that the inflection points can still be used reliably to determine the pad edges.

STEP 4

After calibrating the step sizes in the X and Y-directions, we need to re-focus the microscope to measure the next smaller pad (80\( \mu m \) in our example). This is performed similarly to the first re-focusing. The tip is approached to the sample surface and subsequently retracted a height of 0.3 times the next pad's width (ie. 24\( \mu m \) for a 80\( \mu m \)-wide pad). Then, walk the stage along the Y-direction until the tip passes the lower edge of the pad.
Continue by repeating steps 3 and 4 to move on to smaller and smaller gold pads until you are finally able to land on your micron-sized sample contacted by the smallest gold pad. As you move to smaller pads (requiring fewer steps), the effect of motor hysteresis becomes appreciable. When dealing with backlash in any machine, it is important to always approach your sample from the same direction as switching directions may cause the stage position to shift suddenly and calculated distances may be inaccurate. As mentioned previously, the amount of hysteresis depends primarily on the cleanliness of the motor interfaces but also on the direction and magnitude of the loading force. Keeping the motors clean and adjusting the spring tightness can help reduce the amount hysteresis present.

### 3.6 Sample Fabrication

Since the first experimental realization of isolated graphene flakes from graphite in 2004 [11], methods and techniques for producing and identifying layers of 2D materials have evolved rapidly. Details of 2D device fabrication are still changing, although, several key aspects and techniques have emerged as standard. For example, the "Scotch Tape Method" of material exfoliation remains a quick and easy way of isolating atomically thin layers of van der Waals stacked materials, oxidized silicon wafers are often used as a substrate for imaging and electrostatic gating 2D materials, and heterostructure fabrication is facilitated using thin polymer films to lift and manipulate thin flakes of 2D materials. This section describes these basic methods for producing devices of 2D materials and their heterostructures as well as the physics governing the ability to visualize thin flakes exfoliated onto SiO$_2$ substrates.
3.6.1 Mechanical Exfoliation

![Exfoliation of 2D Materials](image)

Figure 3.19: **Exfoliation of 2D Materials** a) Thin flakes of a van der Waals-stacked material are removed from a bulk crystal using tape b) The removed flakes are thinned out using more tape c) Atomically thin flakes are transferred from the tape onto a Si/SiO$_2$ substrate

The traditional method of mechanically isolating 2D materials, the ”Scotch Tape” method, was the first to produce and measure graphene on SiO2 using Scotch Tape brand tape [11]. In principle, the method uses a piece of tape to cleave away thin layers of a bulk, vdW-stacked material. The layers adhered to the tape are then pressed onto a substrate (usually SiO2). The weak vdW bonds between layers (compared to the in-plane bonds) gives a small probability of leaving behind a thin (sometimes monolayer) flake of the 2D material when the tape is removed from the substrate.

To exfoliate a 2D material, one should start with a clean bulk crystal. In this work, we use Nitto brand ”blue tape” rather than Scotch brand tape because its adhesive is weaker and easier to remove with solvents. The surface of the bulk crystal should be cleaned with tape to remove loosely bound particles. Then, a fresh piece of tape is pressed onto a clean, bulk crystal surface (gently) to remove a small portion
of the crystal onto the tape.

In the next step, the cleaved crystal is thinned. The removed flake is then pressed onto another piece (or section) of blue tape to further thin it down. This process is repeated until the desired thickness and coverage is acquired. Describe the figure

There are many subtleties here which might have an impact on the quality of the crystal deposited onto the substrate. Having thinner flakes on the tape before deposition yields thinner deposited flakes. Similarly, starting with larger area crystals on the tape will deposit larger area thin flakes onto the substrate. If the tape surface is completely covered by the cleaved and thinned 2D material, many thin flakes will transfer to the substrate. However, aside from graphene, 2D materials are usually rare, grown in small batches, and are prone to chemical reactions with the ambient environment. Therefore, one must decide carefully on the size of the tape used and how much of the bulk crystal to use up during sample fabrication.

After the cleaved crystal has been suitably thinned, the tape is pressed, crystal-side down, onto the substrate of choice. As the tape is removed, weakly bound 2D flakes are deposited onto the substrate surface. Here there are some more subtle effects which affect the statistical probability of acquiring a thin, large-area 2D flake. Factors such as temperature, pressure, and humidity play an important role during sample exfoliation, however, these effects are rarely controlled in practice. Different publications and groups cite different recipes for the "optimal" temperatures and pressures to use, however, these results are often not reproducible when carried out in other labs. From experience: higher temperature yields larger area flakes but they are generally thicker much more abundant (making isolating a single flake difficult), higher pressure yields larger area flakes but more tape residue, and lower humidity yields larger area flakes but is rarely controlled reliably.

Typically, for a Si/SiO$_2$ substrate, many flakes (typically tens of microns in diameter) of the 2D material will be left after the tape is removed with varying thicknesses. We then use an optical microscope to search the substrate surface to identify flakes of
desired thickness using the color and optical contrast (described in the next section).

Figure 3.20: **Electrode Fabrication** a) A thin piece of 2D material on a substrate is spin-coated in PMMA b) The metallic lead locations are defined using electron beam lithography. The electron beam locally damages the PMMA film, allowing those regions to be selectively washed away c) A thin film of metal is evaporated onto the entire surface d) The sacrificial PMMA is washed away, removing the excess gold with it. The result is an electrically contacted 2D flake

After a desired 2D flake is found, the next step toward fabricating an electronic device is to deposit electrical contacts. The flake’s position on the substrate is recorded and the surface is spin-coated with poly(methyl methacrylate) (PMMA) to serve as an electron beam resist. E-beam lithography is performed using an FEI Sirion electron microscope with an accelerating voltage of 30kV. The electron exposure dose is chosen according to the molecular density and thickness of the PMMA used. The exposed regions of PMMA are then developed away using a 1:3 mixture of H$_2$O:IPA kept at 5°C followed by a rinse in pure IPA and blown dry with nitrogen gas. Metallic electrodes (typically $\approx$5nm:45nm of Ti:Au) are deposited by e-beam evaporation (Thermionics) under ultra-high vacuum (UHV) ($<10^{-9}$ torr). The metal film thickness is monitored using a quartz microbalance (Q-pod). The excess gold is removed with the remaining PMMA in a final "lift-off" step. The sample is submerged in warm (80°C) acetone and the liquid is gently agitated using a pipette. After a few minutes, the excess gold film floats away from the sample surface. The sample is then
transferred into a beaker of fresh acetone followed by a wash in IPA and blown dry with nitrogen gas. Finally, before measurement, the sample should be annealed for several hours (≥4 hours) in forming gas (H₂:Ar 10:90) below 250°C to avoid melting or displacing the gold electrodes.

3.6.2 Optically Imaging 2D Flakes

Figure 3.21: **Contrast vs. Oxide Thickness** a) Diagram of thin film interference in 2D materials on a Si/SiO₂ substrate. The difference in indices of refraction between the materials in the path of the light cause an interference effect. Changing the thickness of the oxide changes the contrast between the 2D flake on a substrate and the substrate itself b,c,d) The calculated contrast as a function of incident light wavelength and SiO₂ thickness for monolayers of (b) graphene, (c) hBN, (d) 2H–MoS₂, and (e) 1T–TaS₂
The ability to see atomically thin materials is often taken advantage of but is not a trivial task. Although it is only one atom thick, free-standing graphene absorbs a considerably large fraction of light due to the finite and constant high frequency conductivity of Dirac fermions which can be calculated as \( \pi \alpha = \pi \frac{e^2}{h v_F} \approx 2.3\% \) absorbed. It is not only interesting that this quantity depends only on physical constants, but is also useful in device fabrication as it allows us to see individual flakes optically \[187\]. To visualize arbitrary 2D materials with low reflectivity, we cannot rely on the opacity of the material itself, instead, we make use of thin film diffraction.

Fabrication of isolated graphene devices (including its first observation) has relied on the fact that graphene is visible when placed on top of silicon crystals with certain thicknesses of silicon oxide grown on top. During the first observation of graphene, the thin film was placed circumstantially on a silicon oxide film roughly 300nm thick. This thickness happens to be one of the most ideal oxide thicknesses for generating contrast between bare \( \text{SiO}_2 \) and \( \text{SiO}_2 \) with graphene placed on top. We can see that this is the case by solving Fresnel’s equations for incident light on an interface of graphene (or any arbitrary 2D material) on \( \text{SiO}_2 \) on Si (See Figure 3.21a).

Each interface in this geometry will generate a phase shift between incident and reflected light rays due to differences in the material specific indices of refraction. Following \[188\], we find that the reflected light intensity can be written as:

\[
I(n_1) = \left| \frac{r_1e^{i(\Phi_1+\Phi_2)} + r_2e^{-i(\Phi_1-\Phi_2)} + r_3e^{-i(\Phi_1+\Phi_2)} + r_1r_2r_3e^{i(\Phi_1-\Phi_2)}}{e^{i(\Phi_1+\Phi_2)} + r_1r_2e^{-i(\Phi_1-\Phi_2)} + r_1r_3e^{-i(\Phi_1+\Phi_2)} + r_2r_3e^{i(\Phi_1-\Phi_2)}} \right|^2
\]  

(3.27)

where \( r_1 = \frac{n_3-n_1}{n_0+n_1}, r_2 = \frac{n_1-n_2}{n_1+n_2} \), and \( r_3 = \frac{n_2-n_3}{n_2+n_3} \) are relative indices of refraction and \( \Phi_1 = 2\pi n_1 t_{2D}/\lambda \) and \( \Phi_2 = 2\pi n_1 t_{\text{ox}}/\lambda \) are the phase shifts due to the 2D material and the oxide layer, respectively. The complex indices of refraction, \( n_i \), are taken from literature (graphene: \[188\], hBN: \[189\], MoS2: \[190\], TaS2: \[191\]). Error in the interpolation of literature data might introduce error in the calculations shown (error bars not shown).
Using the calculated reflected light intensity, we define the material contrast as the difference between the reflected light intensities with and without the 2D material present

\[ C = \frac{I_0 - I}{I_0} \]  

(3.28)

where \( I_0 \) is the reflected light intensity of the bare SiO\(_2\) surface without the 2D flake present. A positive contrast value then corresponds to absorption of more light by the 2D material than bare SiO\(_2\) and a negative value means light is more strongly reflected by the 2D material.

We can then plot the contrast provided by a monolayer as a function of oxide thickness (see Figure 3.21b-e). From the figure, we can see that the contrast is large only for certain combinations of wavelength and oxide thicknesses. Ideally, we would like to observe high contrast near the middle of the optical light spectrum (say \( \approx 600\text{nm} \)). Then, using graphene as an example, we find the optimal thicknesses of SiO\(_2\) to be near 90nm and 300nm, precisely the two thicknesses commonly used in graphene device fabrication. Notice that the maximum contrast value obtained for graphene is approximately 14\% where the values for hBN are less than 3\%. This shows that it is extremely difficult to see monolayer hBN on SiO\(_2\) at all thicknesses. To maximize the ability to see thin hBN flakes, one should use 90nm of SiO\(_2\). The contrast of the two TMDs considered, 2H–MoS\(_2\) and 1T–TaS\(_2\), maximize due to reflection rather than absorption and provide good contrast on either 90nm or 300nm SiO\(_2\).
Figure 3.22: **Graphene and hBN Contrast on SiO₂ vs. Layer Number**  

a,b) Wavelength-dependent contrast of (a) graphene and (b) hBN vs. layer number on 285nm of SiO₂ on Si c,d) Linecuts of contrast vs. wavelength for 1, 5, 30, and 100 layers of (c) graphene and (d) hBN. Insets depict the approximate color of the 2D flake on substrate.
Next, we investigate the contrast dependence on thickness of the 2D material layer. From figures 3.22 and 3.23, we see that for a constant SiO₂ thickness, the contrast of a 2D material flake varies dramatically in the low-thickness range. To make this calculation into a useful experimental tool, we can map the optical wavelength-dependent intensities to R, G, and B channels of our optical microscope camera to estimate the actual color of the thin flakes that we expect to see. The precise color of the observed flakes depends on details including the type of camera and light source used, however, the observed trends in color are reproduced (see insets in plots of figures 3.22 and 3.23 for calculated color estimation). Using this simple analysis, we are able to estimate the thicknesses of exfoliated thin flakes on SiO₂ which has been
tuned to a thickness which maximizes the optical contrast of thin flakes.

3.6.3 Heterostructure Fabrication Techniques

When measured on an insulating substrate, 2D materials can exhibit remarkable intrinsic properties. However, with all of their atoms at the surface, these properties are easily obscured by interactions with the environment such as substrate-induced random potentials and adsorbed species. It is then desirable to protect these layers with flat, inert substrates and capping materials.

Encapsulated graphene is a prime example of how protecting 2D materials from the environment might improve electronic quality and uniformity. Encapsulated graphene devices have been shown to exhibit significantly less electronic scattering than graphene on silicon dioxide. The high mobility and reduced scattering centers has lead to interesting experiments on electrons including micron-scale ballistic transport [29], electron optics [30, 31], and magnetic focusing [32].
Figure 3.24: **PMMA Transfer Film Method** a) PMMA transfer film is made from a spin-coated stack of PMMA on PVA on a Si/SiO$_2$ substrate. A piece of blue tape with a hole cut in it is adhered to the PMMA surface b) A 2D material is exfoliated onto the PMMA within the center of the tape window c) The PMMA film is peeled from the PVA film using the blue tape window and placed, 2D flake-side down, onto a substrate of choice under an optical microscope d) The final structure is a 2D flake, placed onto any arbitrary substrate with a PMMA overlayer; ready for lithography

One of the earliest techniques developed for heterostructure fabrication, using a poly(methyl methacrylate) (PMMA) supporting film, has remained one of the most reliable methods of transferring 2D materials. In this method a thin layer of the water-soluble polymer poly(vinyl alcohol) (PVA) is spin-coated onto a silicon wafer (with or without an oxide layer) ($\sim$1200 RPM) and baked at 80$^\circ$C for about 5 minutes. The PVA film is then spin-coated with a layer of PMMA and baked at 80$^\circ$C for at least 30 minutes. In this step, the PMMA curing temperature is lower than the suggested recipe in order to avoid over-heating the PVA layer and the curing time is long to produce a film which is more easily removed from the PVA in a later step.

After curing, the edges of the resulting stack is gently scratched to release the polymer films from grabbing onto the Si wafer and a tape with a hole (smaller than the Si wafer) is placed onto the polymer films (see top left of figure 3.24).
At this stage, the 2D material can be exfoliated onto the polymer film, taking care not to rip the sensitive film with the tape. With the 2D material exfoliated onto the PMMA surface, the polymer stack is removed from the backing silicon wafer using the tape window. After being removed from the Si wafer, the sacrificial PVA layer can be peeled from the PMMA layer using tweezers or dissolved with DI water. Finally, the tape window, which now holds only the PMMA film with 2D flakes on it, is placed onto a custom-made holder which can be placed under an optical microscope (flake-side down for highest contrast). Under the microscope, thin flakes are identified by their contrast and can be placed onto any substrate of choice. The PMMA film is subsequently used as an e-beam resist or is removed with acetone.

Another method utilizes a film made from the viscoelastic polymer poly(dimethyl siloxane) (PDMS). The PDMS film is prepared according to specifications from the manufacturer (including degassing in vacuum) after being poured onto a clean, flat surface (a fresh glass slide or blank CD works well). After curing, materials can be exfoliated onto the PDMS surface similarly to mechanical exfoliation on SiO$_2$, however, the tape should be removed from the PDMS surface quickly to maximize the size of the flakes transferred. Note that exfoliation onto PDMS shows a much lower transfer yield compared to SiO$_2$ due to a lower surface energy. To combat this effect, we use large area PDMS films to increase the chance of yielding a useable 2D flake. The PDMS stamp is then placed under an optical microscope, flake-side down, and the 2D flakes can be located. The flakes are transferred onto the desired substrate by placing the flake and substrate in contact followed by slowly removing the PDMS film. One of the benefits of this method is that the 2D flake can be deposited without contacting any liquid solvents.
For heterostructures requiring high mobility graphene, the most reliable method is called the "vdW Pick-Up Method." This technique takes advantage of the small lattice spacing mismatch between graphene and hBN, which leads to a strong vdW bond between flakes of the two materials when placed in contact [192, 193]. Because this interaction is stronger than the graphene-substrate interaction, hBN can be used to selectively pick up graphene flakes from a variety of substrates [194]. This method requires a homemade "handle" tool which is able to choose and grab a hBN flake.
from the surface of SiO\textsubscript{2}. The handle is made from a glass slide which has a PDMS (sometimes clear epoxy) cushion which is shaped and protected by a piece of clear tape. The tape should be placed as to compress the PDMS cushion at the edges, leaving the surface slightly rounded. It has been shown that by rounding the cushion, one obtains a well-defined apex which contacts the substrate first, improving the level of control [195].

For each sample, the handle is spin-coated (or blanketed) with a thin layer of poly(propylene carbonate) (PPC), a clear polymer which exhibits much larger surface adhesion than both PDMS and PMMA, at a low spin rate (\(\leq 1000\) RPM) and cured at 90\(^\circ\)C. Next, hBN and graphene are exfoliated onto separate Si/SiO\textsubscript{2} chips and located under an optical microscope in the normal fashion, using thin film interference to determine thickness. After a piece of graphene and a thin flake of hBN are chosen, the handle is loaded onto a micromanipulator under the optical microscope. The apex of the handle is then aligned with the chosen hBN flake and the handle is lowered to make contact between the PPC film and hBN flake. As the handle is lifted from the surface, it lifts the hBN flake with it. The probability of removing the hBN from the SiO\textsubscript{2} surface depends on several factors. Primarily, the hBN should be freshly exfoliated and the PPC film freshly cured. Leaving the hBN on SiO\textsubscript{2} for long periods of time or annealing the hBN flake will make the bond between hBN and SiO\textsubscript{2} too strong.

Next, the hBN on the handle is aligned with a graphene flake on SiO\textsubscript{2} and lowered to make contact between the hBN and graphene layer. Again, as the the handle is lifted from the surface, the graphene is removed from the SiO\textsubscript{2}, adhering strongly to the hBN. The probability of removing graphene from the SiO\textsubscript{2} surface is usually very high (given that the surfaces are clean). The bond between hBN and graphene is even large enough to tear micron-sized graphene layers at the edge of the hBN flake [195].

Finally, the hBN/graphene stack can be placed onto an arbitrary substrate by
placing the stack in contact and heating the sample above 90°C to partially melt the PPC film, leaving it on the substrate when the handle is lifted. Perhaps the only drawback to this method is that the resulting graphene layer must be contacted on (at least) one side with hBN. To perform tunneling measurements or to place electrical contact, the stack must either be inverted or the hBN layer partially removed by reactive ion etching, respectively.

3.6.4 In Air vs. Glovebox

Samples made from graphene and hBN are highly inert and can be easily annealed clean prior to electrode fabrication and measurement. This allows these materials to be exfoliated, manipulated, stacked, and measured in ambient conditions with little risk of permanent degradation of device qualities. Many materials, however, are sensitive to degradation due to chemical interactions with ambient molecules. For the cases where material degradation is a risk, or if a device requires an exceptionally high degree of cleanliness, sample fabrication should be performed in a glovebox with an inert atmosphere.

Figure 3.26: Glovebox Fabrication Station
Heterostructures containing air-sensitive materials are constructed inside an argon-filled glovebox (see figure 3.26b). The argon atmosphere is circulated through a purification system which maintains the ambient oxygen and water levels below 0.1ppm. Stored within the glovebox are supplies for mechanical exfoliation of 2D materials onto SiO\textsubscript{2} including bulk crystals of 2D materials, annealed Si/SiO\textsubscript{2} substrates and tape. The glovebox is equipped with a spin-coater and a hot plate so that polymer films can be spun in inert atmosphere. For vdW heterostructure fabrication, a double-stage, hands-free optical microscope allows the user to align and transfer 2D materials with high accuracy and precision. The sample stage of the microscope sits on translational and rotational motors that allow precise positioning and orientation of the sample. A second stage has translational as well as z-axis positioning to align and place 2D materials on polymer films and PDMS handles (figure 3.26a).
Chapter 4
Seeing into van der Waals Heterostructures

This chapter follows closely with text and contains data reproduced from [196], written and collected by the thesis author.

Using heterostructure fabrication techniques, one can preserve or enhance a material’s intrinsic electronic properties by encapsulating it in an inert, insulating material such as hBN. Encapsulating 2D materials is a simple and elegant method to protect devices from interactions with the environment which negatively impacts their performance. This method was first applied to graphene encapsulated between flakes of hBN.

A common challenge in the fabrication of encapsulated graphene structures, or any 2D heterostructure, is maintaining and evaluating the cleanliness of the planar interfaces between materials. The quality of such devices depends, crucially, on the location and number of trapped contaminants within the structure and on the quality and reliability of the electrical contacts to the device. However, contamination due to tape and polymer residues and ambient particulates are ever-present, even when devices are fabricated inside a glovebox.

In this chapter, we demonstrate the use of non-contact AFM-based methods to image graphene encapsulated between stacks of hBN. These methods allow us to image and probe the electronic properties of the embedded graphene layer and the junctions between materials during and after sample fabrication. Details regarding trapped contamination, cracks in the buried graphene, and local doping modulations can be obtained on the sub-micron scale. Such a feature is very useful when creating
encapsulated transport devices, where identification of contamination-free regions and positioning of electrodes play a crucial role in device performance.
4.1 Seeing Into Heterostructures

Figure 4.1: **EFM Phase Shift Map**  
(a) an optical microscope image of an encapsulated graphene device showing the exposed SiO$_2$ substrate in the bottom left corner, the bottom hBN flake beneath the graphene layer (blue), and the top hBN flake (dark yellow). Additionally, bubbles within the structure (blue and brown spots) and a thin gold backgate (15 nm thick) which runs beneath the sample (orange stripe).  
(b) A schematic diagram of the measurement of an encapsulated bilayer graphene sample during the 2nd pass of the measurement. The tip is maintained at a distance above the sample surface, following the surface morphology, while a bias is used to probe the electrostatic properties of the structure.  
(c) AFM topography of the device imaged in (a) and (d) a linecut through the topography image along the dashed line in (c) gives the thicknesses of the encapsulating hBN layers  
(e) an EFM phase shift map in the same region as (c) displays a clear phase shift in the vicinity of the encapsulated graphene region  
(f) the dependence of the measured phase shift on applied tip bias in different regions of the device shown in (e).  
(g) a zoomed in EFM phase shift map at the edge of the encapsulated graphene region  
(h) a linecut through the EFM phase shift map along the dashed line in (g) [196]

Encapsulated graphene samples are fabricated using the vdW Pick-Up method described in section 3.6.3 and are placed on a 300nm SiO$_2$-capped Si substrate, some
with a pre-patterned gold electrode on the surface (which acts as a local backgate during transport measurements). An example of an encapsulated graphene sample is shown in the optical microscope image given in figure 4.1a. The final structure includes two stacked layers of graphene (with no twist between them) between two flakes of hBN on top of a gold electrode patterned on SiO$_2$. The SiO$_2$, bottom flake of hBN, and top flake of hBN can be seen in fig. 4.1a. The local gold backgate can be seen through the structure. The region where the graphene layers have been encapsulated exhibits bubbles of water and hydrocarbons that are trapped between the layers. The precise location of the embedded graphene layers are very difficult to see optically without filtering and/or heavy image processing due to the low contrast provided by the graphene layer.

When making heterostructure devices, it is common procedure to image the device surface and measure the component thicknesses using non-contact AFM. The AFM topography of the sample shown in figure 4.1a is given in 4.1c and a linecut at the top edge of the sample (red, dashed line) is given in 4.1c. The topography data shows that the top and bottom encapsulating flakes of hBN are 50nm and 71nm thick, respectively. The bubbles which are visible optically can also be seen to protrude from the sample surface but the location of the graphene layers are not resolved.

Next we measure the electrostatic force gradient above the sample surface during a second pass of the AFM tip (as described in section 3.2.3). Unlike typical EFM measurements, the reference ground is attached to the doped Si substrate rather than the conducting graphene layer (see figure 4.1b). The resulting plot of the cantilever phase shift in space, taken with a DC bias $V_b = -6V$, is given in figure 4.1e. Even without electrical contact to the embedded graphene, we can see a clear jump in the phase shift within the middle of the top hBN region, highlighting the location of the buried graphene layers.

We measure the phase shift versus bias voltage above several regions of the device surface: bare hBN (no graphene), encapsulated monolayer graphene, and where the
two graphene layers overlap (BLG) (figure 4.1f). From equations 3.10 and 3.12, the minima of the phase shift vs. bias curve is related to the surface potential while the curvature relates to the second derivative of the local capacitance in each region. In fig. 4.1f, we see that the phase shift minima nearly coincide in all regions of the device; however, there is a large enhancement of $\frac{\partial^2 C}{\partial z^2}$ above the conducting graphene layers both due to the screening ability of the graphene layers and due to the close proximity to the AFM tip compared to the Si backgate. The difference in the second derivative of the capacitance in the different regions of the sample is responsible for the difference in curvature of the plots in figure 4.1f and allows one to control the contrast between the regions of encapsulated graphene and the surrounding hBN stack with the application of a bias.

A zoomed in phase shift map, taken at the corner where the two graphene layers overlap within the device is given in figure 4.1g and a line cut across the edge of the buried graphene is plotted in figure 4.1h. The EFM phase shift decays quickly away from the graphene edge (within 500nm), defining the graphene edge location with high accuracy.
Figure 4.2: EFM Phase Shift as a Function of hBN Thickness and Tip-Sample Separation

a) an optical image of an encapsulated graphene sample with a folded hBN flake on top. b) An AFM topography scan taken at the same location shown in (a) showing the structure of the folded hBN. c) The EFM phase shift map taken at Vb = 6 V in the region shown in (a) and (b) shows an encapsulated graphene layer beneath the folded hBN flake. Regions of the graphene covered with different thicknesses of hBN are marked with stars. d) A linecut through the AFM topography scan along the teal line in (b). The graphene sits on top of a hBN flake 37 nm thick with three different thicknesses of hBN on top: 11.6 nm, 21.3 nm, and 30.7 nm. e) The cantilever phase shift recorded at Vb = 6 V for three different thicknesses of encapsulating hBN marked in (c) shows a monotonic decrease of the force gradient as thickness increases. f) The cantilever phase shift versus lift height taken at Vb = 3 V above another encapsulated graphene sample (not shown) displaying the sharp decrease in contrast between the encapsulated graphene and surrounding hBN as the tip-sample separation grows large. [196]

As the EFM phase shift is proportional to the second derivative of the local capacitance, it is highly dependent on the dielectric environment and the distance between the tip and graphene surface. We expect the thickness of the encapsulating hBN layer to have an effect on the contrast between the encapsulated graphene and surrounding regions. We make use of an encapsulated graphene sample which has a
folded sheet of hBN on top in order to demonstrate this effect. The optical image and AFM topography is shown in Figure 4.2a and b, respectively. The top hBN flake has folded on top of itself twice providing three separate regions with different encapsulating thicknesses determined from a linecut through the AFM topography (teal line in Fig. 4.2b) shown in Figure 4.2d. Three such regions are marked in the EFM phase shift map (teal stars in Fig. 4.2c) and their measured phase shift is plotted vs encapsulant thickness in Figure 2e and shows a monotonic decrease of the phase signal with hBN thickness. Similar to increasing the hBN thickness, we change the tip-sample separation (lift height) and monitor the phase shift above the encapsulated graphene and above bare hBN with a tip bias of $V_b = -3V$. We find that the contrast between encapsulated graphene and bare hBN decreases sharply for tip-sample distances between 10nm and few hundred nanometers, becoming immeasurably small for tip-sample distances larger than the thickness of the silicon oxide substrate.
Figure 4.3: **Cracks and Goo by EFM**

a) AFM topography of an encapsulated bilayer graphene sample fabricated in a glovebox 
b) EFM phase shift map in the same region shown in (a) showing a micron-sized crack in the graphene layer formed during device fabrication 
c) AFM topography of another encapsulated bilayer graphene fabricated in a glovebox 
d) EFM phase shift image corresponding to the region shown in (c) with non-uniform regions presumably due to trapped contaminants. These samples were measured without controlling the AFM chamber humidity. [196]

Using the method, we can identify contaminated and defective regions of the device early in the fabrication process, along with the sample topography prior to plasma etching through the hBN and placing electrical contacts. As a demonstration of this capability, we plot the AFM topography and EFM phase shift maps at the surface of two defective encapsulated bilayer graphene samples in figure 4.3. The sample shown in fig. 4.3a,b exhibits a tear or crack in the buried graphene layer which occurred during the fabrication procedure. The sample in figure 4.3c,d looks flat and clean with AFM topography, however, the EFM phase shift map shows splotches of residue which has been trapped within the structure. Only with EFM imaging are we able to identify these defective regions before performing time consuming device fabrication procedures and transport measurements.
4.2 Charging and Charge Fluctuations

Figure 4.4: Surface Potential Map of Encapsulated Graphene a) The charge of the buried graphene layer locally modifies the measured surface potential. b) A zoom in to a small region outlined in (a) where two encapsulated graphene layers overlap. The band structure differences between monolayer and bilayer graphene leads to an observable difference in the measured charge density when the overall graphene layer is charged. c) Raman spectrum taken at the location of the stars in (b) confirms that the contrast that we observe is due to the boundary between monolayer and bilayer graphene. Note that the graphene D-peak is not shown due to convolution with the hBN G-peak. d) A histogram of surface potential values measured in the monolayer and bilayer regions gives a measure of surface potential fluctuations.

In contrast to previous measurements of graphene by KPFM, we measure the surface potential variation without electrical contact to the graphene layer. With this method, we probe the location and properties of the graphene layer before performing any lithography or electrode deposition. Instead, electrical contact and the applied bias is made between the coated tip and the highly doped silicon backgate beneath the SiO$_2$ substrate. Without electrical contact to the graphene layer, the signal depends on a combination of the work function difference between the gold coated tip and the doped silicon backgate modified by the charges and conductors between them.
\[ V_s = \frac{-\Delta \Phi}{e} + \frac{q}{C_b} \quad (4.1) \]

where \( q \) is the total charge trapped within the structure beneath the tip, and \( C_b \) is the effective capacitance between the localized charges and the grounded backgate \([181, 180]\).

Next we employ AM-KPFM to map the surface potential of the sample shown in Figure 4.1. The results are displayed in Figure 4.4. The region where graphene is encapsulated shows a lower surface potential with respect to the tip than the surrounding boron nitride indicating a difference in local charge density given by equation 4.1. By determining the thickness of the bottom hBN flake (from Fig. 4.1d) and evaluating using equation 4.1, we can determine the 2D charge density within the graphene layer to be \( n_g = 0.92 \times 10^{10} / \text{cm}^2 \). This level of doping is typical for encapsulated graphene devices fabricated in a dry atmosphere. Despite fabrication in a controlled environment using fresh, clean materials, we always observe some small amount of charge bound to the graphene layer. This charge density is large enough to provide the contrast observed in the surface potential map.

Due to the charge trapped in the embedded graphene, we observe a surface potential contrast between the encapsulated monolayer and bilayer graphene. The difference in doping dependence of monolayer and bilayer graphene stems from the difference in low energy band structures. To maintain the same Fermi level, the bilayer graphene requires a larger charge density, as seen in Figure 4.4b. With the surrounding boron nitride stack as background, we find a charge carrier concentration in the bilayer region of \( n_{BLG} = 1.16 \times 10^{10} / \text{cm}^2 \). Note that from AFM topography (Figure 4.1c), neither the graphene edge nor the monolayer/bilayer boundary are visible, however, using KPFM we are able to select device regions to produce monolayer, bilayer, or a device across the boundary.

Upon fabricating and attaching electrodes to the graphene layer, the trapped
charge will change due to the work function differences between the graphene layer and the contacting metal. Thus, the absolute charge density is not a good metric for device quality. Instead, we may look at the surface potential fluctuations within the encapsulated graphene region. The potential within the graphene layer is expected to be uniform and thus, any surface potential fluctuations must originate from unscreened charges within the structure. Surprisingly, despite the presence of bubbles observed in sample topography (Fig. 4.4a), the measured surface potential seems locally unaffected. However, we do observe fluctuations in the surface potential across the device surface reflected in the RMS spread of the surface potential distribution (Fig. 4.4d). Additionally, in some regions of the device, we observe local potential fluctuations (up to 70mV) which are not observed to be associated with any topographical features (Fig. 4.4e,f). These features provide strong electronic scattering centers which might accidentally be included in a processed device if using AFM topography alone. It is well-known that charge traps, which are typically found within the insulating SiO2 substrate, the hBN encapsulant and at the various interfaces, provide a random potential and scattering centers which reduce the devices transport properties. We measure the RMS variation in surface potential in several regions of the device. For the monolayer (g) and bilayer (BLG) regions, we find RMS variations of $\delta V_g = 26.09mV$ and $\delta V_{BLG} = 18.73mV$ corresponding to charge variations of $\delta n_g = 1.51 \times 10^9/cm^2$ and $\delta n_{BLG} = 1.08 \times 10^9/cm^2$, respectively. The values obtained are comparable to those measured by transport in similar devices [192]. Thus, we can use this metric to identify pristine regions of the sample that are suitable for further processing so as to maximize device quality.
To remove the charge from the sample shown in figures 4.1 and 4.4, we anneal it in forming gas (10% hydrogen in argon) at 220°C for 3 hours. After annealing, we again probe the sample surface potential in the same region with KPFM (see Figure 4.5a). We find that the contrast has inverted, indicating a change in sign of the charges in graphene (and also the gold backgate). Again, comparing to the potential measured at the surface of the bottom layer of hBN, we determine a charge
density \( n_g = -2.08 \times 10^{101}/cm^2 \) within the encapsulated graphene monolayer and \( n_B LG = -2.11 \times 10^{101}/cm^2 \) within the encapsulated bilayer region (negative sign indicates electron doping) with charge fluctuations of \( \delta n_g = 8.04 \times 10^{81}/cm^2 \) and \( \delta n_B LG = 7.55 \times 10^{81}/cm^2 \) measured within the most pristine regions of the device. This finding is contrary to the expectation that thermal annealing would reduce the amount of contamination in our encapsulated samples. Instead, we find that the sign of dominant dopants reverses, the carrier concentration increases, and the charge fluctuations within the device reduce some. After annealing, the surface potential map (Fig. 4.5c) reveals many strong, localized fluctuations. Comparing with Figure 4.5b, we see that these fluctuations remain predominantly localized to the edges of the bubbles within the heterostructure with some fainter features connecting between them. Further zoomed-in measurements of one defect near a small bubble (Fig. 4.5d,e) reveals that the charge measured in KPFM can be associated with a charged defect which has attached itself to a pit formed in the hBN surface at the high-curvature region near the bubble. The degradation of the hBN surface and formation of large, charged defects during annealing is unexpected and further work is required to verify the origin of these defects and their formation.

Thus, the key to high quality encapsulated device fabrication may not be so simple as avoiding the bubbles. Even typical annealing is shown to have a negative impact on sample uniformity. Similar effects on local potential fluctuations due to further fabrication techniques such as plasma etching and electrode deposition are yet to be investigated, however, we have shown that our measurement techniques are perfectly suitable for probing these effects and will be the topics of future work.
Chapter 5

Measurements of Graphene Covered TaS$_2$

5.1 Avoiding Oxidation of TaS$_2$ Surface

This chapter contains theoretical calculations and modeling which were created by co-authors of, as yet, unpublished works [197] in addition to experimental data and text written by the author of this thesis.

A common challenge in studying novel materials is protecting the sample surface from chemical interactions with the environment. Many 2D materials, especially transition metal chalcogenides, oxidize readily in air [198]. In bulk (3D) materials, the oxidation process is self-limiting. That is, the oxidation only begins at the surface and as it builds in thickness, prevents more oxygen from interacting with the material. This is detrimental to characterization of such materials, especially by scanning probe methods.
5.1.1 Identification of Oxidation

Figure 5.1: **Oxidation Rate of 1T–TaS$_2$ in Ambient Conditions** a) AFM topography of the bulk TaS$_2$ surface. The region outlined by a dashed box has been previously scanned multiple times, removing the loose, oxidized surface there which piles up at the edges of the scan. b) KPFM surface potential map taken in the same region shown in (a). The freshly exposed TaS$_2$ surface outlined by the dashed box has a lower surface potential (blue) than the oxidized surface (yellow). c) Evolution of surface potential with time (circles) in the freshly exposed region outlined in (a, b). (error bars reflect the standard deviation). From the exponential fit (solid red line) we obtain a characteristic surface lifetime of 46.8 minutes. d) AFM topography (top) and KPFM surface potential map (bottom) taken on a small section of a G/TaS$_2$ flake shortly after fabrication. Three regions of interest are marked in the topography image: graphene on SiO$_2$ (outlined in dashed gray), G/TaS$_2$, and air exposed TaS$_2$. (outlined in dashed red). G/TaS$_2$ shows a lower surface potential than graphene on SiO$_2$ and exposed TaS$_2$. e) AFM topography (top) and KPFM surface potential map (bottom) taken in the same area as in (d) after a long exposure to air (greater than one month) and after STM measurement. The surface potential increases near the edges of the graphene cover layer where the seal is not perfect. To compare the KPFM maps in (d) and (e), the surface potential is calibrated with respect to a flake of graphite near the sample (not shown in this image).

To probe the oxidation effect at the surface of TaS$_2$, we repeatedly scanned the surface of an oxidized bulk TaS$_2$ crystal using AFM. After many scans, the loosely bound oxide gets swept away by the AFM tip, revealing a fresh TaS$_2$ surface. Figure 5.1a
shows the AFM topography map of the corner of such a region that has been freshly exposed. Piles of swept away surface and debris can be seen outlining the square region that was previously scanned multiple times. The removal of the surface is not surprising as even the surface of pristine TaS₂ has been reported to be easily swept away by scanning with probes [199]. The controlled removal of the oxidized surface is, perhaps, a useful method to clean the surface of 2D devices prior to measurement (similar to [200, 201]) or even to thin down or remove regions of a device. Furthermore, the ability to locally induce oxidation at the surface of metallic TMDs using the AFM tip [202] potentially allows one to pattern devices on the scale of tens of nanometers as well as control their thickness; which has yet to be demonstrated.

Next, we perform Kelvin probe force microscopy (KPFM) to measure the work function difference between the sample surface and the gold-coated AFM tip, fig. 5.1b. Within the newly exposed square, the contact potential difference between the tip and surface is lower than the surrounding, oxidized region. Assuming the work function of the gold tip remains constant, we measure the work function change of the TaS₂ surface as a function of time while exposed to ambient conditions. As time elapses, the measured surface potential begins to increase toward (and beyond) the value measured in the surrounding (oxidized) region, see figure 5.1c. By fitting the results with an exponential decay function, we can determine a characteristic lifetime of the material surface, \( \tau = 46.8\text{min} \).

5.1.2 Early Transport Measurement of G/TaS₂

The time dependence of the oxidation process that we measure has allowed previous STM groups to cleave bulk samples in air before quickly loading into vacuum. Similarly, transport measurements which have placed electrical contacts while minimizing the exposure to air. To combat the negative effects of oxidation on our early transport measurements, we make sure that the TaS₂ surface is not exposed to air for more than a few minutes (\( << \tau \)) before spin-coating with PMMA and e-beam
lithography and electrode deposition are performed as quickly as possible. An example of an early transport measurement of a bulk TaS$_2$ is shown in figure 5.2. By performing quick sample preparation, we are able to make electrical contact with the TaS$_2$ flake. Additionally, we have quickly placed a small piece of monolayer graphene onto the TaS$_2$ surface and find good contact between the two flakes, again suggesting that we have avoided oxidation of the surface prior to device fabrication. This results in two regions of the device which we probe with the gold electrodes shown in figure 5.2a: a region of bare TaS$_2$ and another containing the graphene flake (G/TaS$_2$). The temperature dependent resistance of these two regions is given in 5.2b during one thermal cycle (Room temperature→∼77K→room temperature).
Figure 5.2: **Transport Measurement of Graphene on 1T–TaS$_2$**

a) Optical image and schematic diagram of a transport measurement of 1T–TaS$_2$ with a small piece of graphene on top (outlined with a green dashed line). The sample sits on top of a SiO$_2$ substrate and is contacted with gold electrodes. Current is passed between the top and bottom electrodes while the voltage across a region of bare TaS$_2$ ($V_T$) and a region containing graphene ($V_G$).

b) Resistance vs. temperature for both the bare TaS$_2$ (black) and graphene on TaS$_2$ (red) regions shown in (a) during cooling and heating. c) A zoom in to the transport data shown in (b) in the temperature range of the NCCDW to CCDW phase transition. d) A zoom in to the transport data shown in (b) in the temperature range of the CCDW to T-phase transition. e) A zoom in to the transport data shown in (b) in the temperature range of the T-phase to NCCDW phase transition.

The resistance vs. temperature curve for bare TaS$_2$ (black points) shows the expected first order transition, and corresponding hysteresis window, between the nearly commensurate (at room temperature) and commensurate (at 77K) CDW phases. As temperature is reduced, the TaS$_2$ undergoes the NCCDW to CCDW phase transition at about 192K, where the resistance suddenly jumps up by more than a factor of two. The resistance continues to rise more slowly as temperature reduces further. This regime where the system seems to be gradually entering the commensurate phase is
likely due to the reordering of domains which have been quenched into metastable configurations during the initial, sudden commensuration of the majority of the sample at 192K. Whether the domains are ordering within the atomic lattice planes or between them and the relative role they play in determining the total sample resistance has yet to be seen (refer to section 2.5). After the TaS$_2$ CDW is fully commensurate (below $\sim 150$K), we can see that the resistance continues to increase, consistent with the expected semiconducting properties of the CCDW phase.

Now, as temperature increases from 77K, the resistance of the bare TaS$_2$ region initially decreases due to thermal activation of carriers across the Mott gap before leveling out and then reducing faster with increasing temperature. This downturn at $\sim 173$K coincides with the expected Debye temperature of TaS$_2$ and is thus attributed to a decrease in electron-phonon scattering; impurity and Umklapp scattering become the dominant contributors [203]. Then at 240K, the resistance drops suddenly as the CDW system enters the T-phase. This sudden drop in resistance has been shown to coincide with a collapse (or, at least, a partial collapse) of the Mott gap [131] as well as a increase in interlayer spacing [204]. Even though the resistance is nearly the same as in the nearly commensurate phase ($R_T(243K) - R_{NC}(243K) = 16\Omega$ for this sample), recall that the T-phase has a markedly different ordering of topological CDW defects than the NCCDW phase (refer back to section 2.3.5). Finally, below the CCDW to T-Phase transition, the resistance decreases linearly with increasing temperature by $2\Omega/K$ until around 280K where there is a small shoulder which marks the weak, first-order transition from the T-phase to the NCCDW phase (see figure 5.2c).

In the region with a flake of graphene on top, the current flows through both the graphene flake and the TaS$_2$ between the leads (figure 5.2a) making deconvolution of the effect of graphene and the bare TaS$_2$ difficult. Instead, we simply compare the resistance vs temperature curves and speculate on possible effects of the interaction between the graphene layer and TaS$_2$. Starting, again, from the cooling curve of
graphene on TaS$_2$ starting at room temperature (red data points in 5.2b,c) the resistance increases linearly with a slope that follows that of bare TaS$_2$. This behavior contrasts that of doped graphene (without TaS$_2$) which shows typical metallic behavior of decreasing resistance with decreasing temperature indicating that the resistance contribution from the TaS$_2$ dominates in this temperature region. As the sample nears the commensuration transition temperature from above, the G/TaS$_2$ resistance begins turn downward, precluding the CCDW transition. The crossover from TaS$_2$-like to graphene-like $dR/dT$ behavior, without any corresponding obvious change in charge ordering of the TaS$_2$, is surprising. It seems that graphene "knows" about the impending phase transition, possibly related to the increase in electron scattering within the TaS$_2$ near commensuration. At the temperature which bare TaS$_2$ suddenly transitions (190K), the G/TaS$_2$ system transitions gradually with temperature. One would expect that the sudden enhancement of electron-electron interactions and loss of carriers at the Fermi level within the TaS$_2$ would cause a corresponding abrupt change in scattering and density of carriers in graphene, reflected in a sharp jump in resistance. However, it seems that screening and/or charge transfer between graphene and TaS$_2$ provides feedback which hinders commensuration near the graphene layer. Of course, in the CCDW regime, we observe metallic behavior of the G/TaS$_2$ system where bare TaS$_2$ is semiconducting due to graphene shorting the Mott gap of TaS$_2$.

As G/TaS$_2$ warms up from 77K, the resistance increases due to increased phonon scattering within the graphene layer before turning over, following the behavior of TaS$_2$. At the CCDW to T-phase transition (figure 5.2d), we observe a much sharper transition than in the bare TaS$_2$. Either the graphene layer somehow reduces the amount of metastable configurations which are trapped within the TaS$_2$ and released during commensuration, or the G/TaS$_2$ resistance is highly sensitive to the surface configuration of the TaS$_2$ CDW and the sudden drop in resistance corresponds to the transition of the topmost CDW layer. Within the T-phase, the G/TaS$_2$ resistance decreases linearly with temperature, following that of bare TaS$_2$, however, we notice
that the shoulder in R vs. T which marks the T-phase to NCCDW transition occurs at a lower temperature in G/TaS$_2$ than TaS$_2$ which indicates that, perhaps, the graphene layer reduces the energy barrier between the two CDW phases. Finally, we note that at the temperature where a mysterious, second shoulder is seen in TaS$_2$, a peak occurs in the resistance of the G/TaS$_2$ which supports the possibility that a buried, metastable T-phase configuration transitions into the NCCDW phase at this point.

In summary, our measurements reinforced the notion that rapid fabrication of device leads and heterostructure construction can beat the negative effects of ambient oxidation, allowing metallic contact between them. Additionally, they have opened novel questions about the interaction of graphene carriers and the equilibrium phases of the CDW at the surface of TaS$_2$ as well as how graphene mediates the transition between these phases. Clearly, temperature dependent, microscopic measurements and analysis of charge transfer at the heterojunction between graphene and the correlated CDW states of TaS$_2$ are needed to elucidate these mysteries.

5.1.3 Constant Work Function vs. Thickness

To further combat the effects of interactions with the ambient environment, we have developed a heterostructure fabrication system inside of a glovebox (see section 3.6.4). The atmosphere of the glovebox maintains < 0.1 ppm of molecular oxygen and water using argon as an inert balance. The glovebox is equipped so that spin-coating and curing of PMMA is performed inside the glovebox as PMMA can slowly (>> $\tau$) absorb and diffuse oxygen. Additionally, we modified our commercial AFM system by adding a slow, controllable flow of dry nitrogen gas into the measurement chamber. We increase the flow of nitrogen until a humidity gauge within the chamber reads less than 5%.

Within the glovebox, we exfoliate 1 T – TaS$_2$ thin flakes on SiO$_2$ and spin-coat with PMMA followed by electron beam lithography and evaporation of Ti/Au (5nm/45nm)
to define electrical contacts. After lift-off, samples were submerged in isopropanol and blown dry with ultra-pure nitrogen gas. Immediately afterwards, the sample is placed onto a sample stage using double-sided tape and electrical contact is made between the sample electrode and the AFM stage using a drop of silver paint before the sample stage is inserted into the climate-controlled AFM system. Minimization of the sample surface’s exposure to air and substantial reduction of oxygen and water inside the measurement chamber ensures that we observe no change in work function at the $1\text{T} - \text{TaS}_2$ surface within the measurement time, indicating the prevention of oxidation of mechanically exfoliated samples.

As $1\text{T} - \text{TaS}_2$ has been shown to exhibit different electronic ground states when thinned to few layers, it would be informative to know if there is a corresponding signature to this transition in the work function of the $1\text{T} - \text{TaS}_2$ surface. A benefit of using mechanically exfoliated thin films is that a single sample can present a wide range of thicknesses of the 2D material. With protection from oxidation during measurement, this experiment is now feasible.

Within the glovebox, we exfoliate $1\text{T} - \text{TaS}_2$ thin flakes on SiO$_2$ and spin-coat with PMMA followed by electron beam lithography and evaporation of Ti/Au (5nm/45nm) to define electrical contacts. After lift-off, samples were submerged in isopropanol and blown dry with ultra-pure nitrogen gas. Immediately afterwards, the sample is placed onto a sample stage using double-sided tape and electrical contact is made between the sample electrode and the AFM stage using a drop of silver paint before the sample stage is inserted into the climate-controlled AFM system. Minimization of the sample surface’s exposure to air and substantial reduction of oxygen and water inside the measurement chamber ensures that we observe no change in work function at the $1\text{T} - \text{TaS}_2$ surface within the measurement time, indicating the prevention of oxidation.
Figure 5.3: Work Function vs Thickness of Exfoliated 1T–TaS$_2$ a) AFM topography of an exfoliated 1T–TaS$_2$ flake on the surface of SiO$_2$. The 1T–TaS$_2$ flake is broken such that several thicknesses are exposed. The bottom right of the image is a gold electrons (approx. 50nm thick). b) Surface potential map of the region shown in (a). The SiO$_2$ surface is charged with many random variations due to trapped charges both within the SiO$_2$ and at the interfaces. The gold electrode (bottom right) show uniform surface potentials as they are both metallic. Despite various thicknesses presented, the surface potential of the 1T–TaS$_2$ flake is uniform across its entire surface. c) Work function of 1T–TaS$_2$ (calibrated against HOPG) as a function of thickness shows a constant value across the thicknesses measured in this work. d,e) A zoomed in view of the 1T–TaS$_2$ flake shown in (a) and (b) better demonstrates the constant work function at all thicknesses.

AFM topography images of one such sample which displays various thicknesses is shown in figure 5.3a,d. The sample sits on an insulating SiO$_2$ substrate (Si backgate is grounded) and has a gold electrode attached to it (bottom right of fig. 5.3a). The KPFM surface potential maps corresponding to the topography images in figure 5.3a,d are given in 5.3b,e. In 5.3b, we can see that the gold electrode appears dark as there is a small work function difference between the gold electrode and the gold-coated AFM tip. On the silicon oxide surface, the potential fluctuates in space due to local charges trapped within, on top, and below the SiO$_2$ layer. The surface potential at the 1T–TaS$_2$ surface remains very flat, despite the large changes in thickness. Probing several similar samples using two different gold-coated tips (calibrated using freshly cleaved HOPG), we obtain a plot of the 1T–TaS$_2$ work function as a function
of thickness displayed in figure 5.3c. From image 5.3b and the plot in 5.3c, we determine a mean value of the 1 T–TaS$_2$ work function to be $5.261 \pm 0.066$ eV and does not deviate from this value down to the lowest thickness measured here (10 nm). The obtained workfunction value is consistent with previous reports from photoemission measurements, supporting our previous claim that quick measurement after electrode fabrication in controlled atmosphere can prevent the oxidation of thin 1 T–TaS$_2$ films. There have been contradictory reports of the critical thickness at which 1 T–TaS$_2$ exhibits electronic properties that deviate from bulk behavior. Our work presented here provides local evidence that supports claims of the critical thickness being below 10 nm.
5.1.4 Protection from Oxidation by Graphene Encapsulation

As was demonstrated in section 4.2, KPFM is an ideal tool for measuring nano-to mesoscopic charge distributions in van der Waals heterostructures. We use this probe to locally investigate the contact between graphene and TaS$_2$ as well as the ability of graphene to prevent the oxidation of the TaS$_2$ surface. Unlike the G/TaS$_2$ transport sample, we fabricate these graphene/1T–TaS$_2$ heterostructures inside a glovebox using the PMMA transfer film method (described in section 3.6.3) to place
a graphene layer on top of 1T-TaS$_2$ exfoliated on SiO$_2$.

Next, we probe the oxidation in a graphene covered TaS$_2$ (G/TaS$_2$) sample. Figure 5.4a shows the AFM topography image (top) and a surface potential map (bottom) of the freshly fabricated heterostructure. There are three important regions of the device: graphene on SiO$_2$ (outlined with a dashed, gray line), G/TaS$_2$, and an exposed TaS$_2$ region (outlined with a dashed, red line). In regions of G/TaS$_2$, we observe wrinkles and bubbles in the graphene layer. These are commonly found on graphene heterostructures where the graphene adheres strongly to the material surface beneath it, pushing debris and interfacial contamination into bubbles and localizing strain into wrinkles [205, 206]. The presence of these bubbles and wrinkles is a good indicator of close contact between the graphene and its substrate. Comparing the surface potential of G/SiO$_2$ and G/TaS$_2$, we find that the G/TaS$_2$ is electron doped by 102±2mV due to the work function difference between the two materials (see figure 5.4c). When spatially separated, the Fermi levels of graphene and 1T-TaS$_2$ are not aligned in energy due to the work function difference between the two materials. After making contact, electrons flow from the graphene layer into the TaS$_2$, moving the Fermi level of both materials. The amount of energy that the Fermi energy moves during this charging process depends on the energy dispersion as well as the carrier density of each material. Because the carrier concentration of graphene is at least three orders of magnitude smaller than that of TaS$_2$ (per unit area), we expect the Fermi level within graphene to move much more dramatically than that of TaS$_2$. The literature reported values of the work functions of graphene and 1T-TaS$_2$ are 4.6eV and 5.2eV, respectively. This would indicate a measured surface potential difference of ~0.5V which is larger than the value obtained here, presumably due to the presence of trapped charges within the SiO$_2$ and water and other contaminants which adhere to the material interfaces and contribute to an initial doping and work function shift of the graphene layer. In the region where TaS$_2$ is exposed, we find that the surface has already partially oxidized during the electrode fabrication, as indicated by the larger
surface potential value compared to fresh TaS$_2$, leaving a sharp boundary between oxidized and graphene covered TaS$_2$.

Repeating the measurement after more than 1 month (1000 surface lifetimes, defined in figure 5.1) of exposure to air, in the same region of the device (figure 5.4b), we find that, barring a small surface potential increase near the graphene edge (suggesting oxidation has occurred there), no measurable oxidation has occurred over large areas in the center of the device (toward the left side of fig. 5.4a,b). This shows that graphene acts as a highly effective gas barrier, protecting the TaS2 from exposure to oxygen. This is consistent with earlier reports showing that graphene is impermeable to gasses and liquids [207, 208]. Our work provides a demonstration of this effect using local probes and reveals that, over time, oxidation still occurs near the graphene edge, where oxygen and water may intercalate into the van der Waals gap between the two materials. Using this knowledge, we can imagine improvements to the device by placing the structure on top of hBN and completely covering the TaS$_2$ surface with graphene. In this configuration, there is no TaS$_2$ exposed to air, preventing the initial oxidation of the surface and the particularly strong vdW bond between graphene and hBN will be able to better prevent intercalation of oxygen and water between the two materials, further improving on the device lifetime.
5.2 Room Temperature STM Measurement of Graphene on TaS$_2$

Figure 5.5: **Room Temperature STM of G/TaS$_2$**

a) Constant-height atomic resolution topography of G/TaS at room temperature show the periodic structures of the graphene lattice and the charge density wave of TaS2 (Vb=100mV, I). b) Schematic diagram of the ambient STM measurement scheme: a metallic STM tip is brought close to the G/TaS2 heterostructure which is supported on an SiO$_2$ substrate. The bias voltage (Vb) is applied to a gold contact (Au) connected to the sample (described in Methods). The presence of atmospheric oxygen does not affect the sample surface during measurement. c) Fourier transform of the topography image displays two sets of coherent peaks corresponding to the graphene lattice ($a_g = 0.245(3)$nm) and the CDW in TaS$_2$ ($\lambda_{CDW} = 1.20(1)$nm).

With protection from oxidation by a graphene layer, we are able to image the CDW state(s) of TaS$_2$ by STM at room temperature both in vacuum and in air. Figure 5.5a gives an atomically resolved, constant-height STM image of the heterostructure surface. As illustrated in the schematics of the experimental setup (figure 5.5b), the STM measurement is performed at the surface of the G/TaS$_2$ flake supported on an insulating SiO$_2$ substrate. Figure 5.5c displays the fast Fourier transform (FFT) of the topography image showing two sets of 6 peaks which correspond the periodic structures seen in figure 5.5a. The outer set of peaks in fig. 5.5c (circled) correspond to the honeycomb lattice expected of graphene ($a_g = 0.245(3)$nm) while the inner set of peaks (marked with arrows) correspond to the CDW modulation expected in TaS$_2$.
The observation of the CDW modulation through the graphene layer is striking, but the mechanism by which we observe it is not immediately apparent. Imaging conducting materials through atomically thin insulating barriers can be explained, trivially, by typical quantum tunneling through a dielectric barrier. When the cover layer is conducting, observation of charge modulations of the encapsulated material is not guaranteed [209]. The observed charge modulation in this work indicates a surprising doping modulation of the graphene layer consistent with a proximity induced CDW. Analogous to the superconducting proximity effect, the charge ordering of the CDW in TaS$_2$ cannot abruptly change to the uniform charge density of the metallic graphene layer at their interface. Instead, the ordering persists into the graphene layer and a non-zero CDW amplitude is observed. Conversely, we expect electrons from the graphene layer to enter the TaS$_2$, screening the electron-phonon interaction and reducing the CDW gap. The question whether the induced CDW exhibits features of typical CDWs such as phonon softening, phase coherent transport, or single particle gap formation remains to be explored.
5.2.1 CDW Domains of TaS$_2$

Figure 5.6: **CDW Domains in the NCCDW Phase** a) Large scale STM topography image of G/TaS$_2$ showing the CDW lattice through the graphene layer. Inset: a zoomed in image of the CDW lattice (scale bar is 2 nm) b) Fourier transform of (a) shows six peaks corresponding to the NCCDW ($Q_{NC}$) with satellite peaks nearby ($Q_{sat}$), associated with the long range domain structure with period $2\pi/|k_d|$ c) Slope map of the topography image (a) highlights domain walls where the CDW amplitude is reduced (some are highlighted in white). Intersection of domain walls form topological vortices and anti-vortices (highlighted with red dots) d) Orientational (black) and translational (red) correlation functions are measured (squares) and fit to a power law (dashed lines) indicating quasi-long range ordering.

Figure 5.6a displays a large area STM topography map of G/TaS$_2$ taken at room temperature exhibiting a periodic modulation consistent with that of the CDW of bare/bulk TaS$_2$. The expected modulation from the NCCDW phase is subtle and can be difficult to see in large area scans with the naked eye. The fast Fourier transform of the drift corrected STM topography map is shown in figure 5.6b. There is a main set of peaks associated with the NCCDW wavevectors as well as nearby satellite peaks which are indicative of the presence of CDW domains [210]. The wavevector which links adjacent domains ($\vec{k}_d$) is given by the difference between the NCCDW wavevector ($\vec{Q}_{NC}$) and the satellite wavevector ($\vec{Q}_{sat}$) so that the domain period is given by $\lambda_d = 2\pi/k_d = 6.305(8)$ nm. In figure 5.6c, we plot a slope map (the gradient) of the topography image (figure 5.6a). In the slope map, individual
CDW peaks appear as small white rings while the periodic reduction of the CDW amplitude at domain walls can be seen as dark lines with spacing approximately equal to $\lambda_d$. Adjacent domains (highlighted with white hexagons in figure 5.6c) interlock with each other and are rotated with respect to the CDW wavevector. The corners of the domain walls form topologically protected vortices and antivortices which, like their analogs in type II superconductors, can be removed from the system by annihilating one another. Note that these domains are fundamentally different from the extended domains and domain walls which have been induced within the low temperature, Mott/CCDW phase using pulses from an STM tip [146, 120]. The domains in the NCCDW phase are more ordered and arise from commensuration strain-induced CDW rearrangements rather than quenched disorder within a Mott insulting background. The CDW domains observed here are consistent with previous X-ray scattering [109] and STM [105, 210] data as well as the theoretical model proposed by Nakanishi and Shiba describing the NCCDW phase of TaS$_2$ (see section 2.3.4). It has been previously shown that the equilibrium CDW phases of TaS$_2$ are distorted or lost when the film is thinned below a critical thickness, however, the flake presented here (approximately 14nm thick) is above this thickness and expected to display intrinsic, bulk qualities [211, 212, 144, 138]. The preservation of the intrinsic CDW structure in G/TaS$_2$ suggests that it is possible to study the collective excitations and complex CDW structure of this thin, air-sensitive material through the graphene layer.

Taking advantage of the imaging technique, we perform a correlation analysis of the vortex lattice produced by the network of domain walls to investigate its proposed two-dimensional character. The nature of ordering in two dimensions is an old problem beginning with Berenzinski-Kosterlitz-Thouless (BKT) theory which shows the existence of oppositely charged topological defects in any 2D lattice and their unbinding at the melting transition. KTHNY theory extends BKT theory by considering the interactions between topological defects, identifying a theoretically predicted but
rarely observed intermediate hexatic phase between the 2D solid and liquid, which is characterized by quasi-long range orientational order and short-range translational order [213]. The topic continues to draw interest as the interaction between topological defects and pinning impurities leads to novel, glassy phases like those observed in the vortex lattices of type II superconductors [214, 215]. Following methods used for analyzing flux line lattices of type II superconductors, the vortex and antivortex positions are determined from the Fourier filtered slope map following standard Delauney triangulation and Voronoi decomposition of the domain lattice [216]. Using these positions, we calculate the translational and six-fold bond-orientational correlation functions, $G_T$ and $G_6$ respectively, as

$$G_T(\vec{r}) = \langle \rho_T(\vec{r})\rho_T^*(0) \rangle$$

and

$$G_6(\vec{r}) = \langle \psi_6(\vec{r})\psi_6^*(0) \rangle$$

where $\rho_T(\vec{r}) = \sum \exp(-i\vec{k}_d^{(i)} \cdot \vec{r})$ and $\psi_6(\vec{r}) = \sum_{nn} \exp(i6\theta(\vec{r}))$ are the appropriate order parameters and $\theta(\vec{r})$ is the bond angle between neighboring lattice sites [217, 218]. The decay of the correlation functions versus inter-particle spacing gives the range of ordering: $G_{T,6}(\vec{r}) \sim \text{const.}$ for long range order, $G_{T,6}(\vec{r}) \sim r^{-\eta}$ for quasi-long range order, and $G_{T,6}(\vec{r}) \sim \exp(-r/\xi)$ for short-range order.

The resulting correlation functions for the CDW vortex lattice are plotted in figure 5.6d on a log-log scale. The bond-orientational correlation function, $G_6(\vec{r})$, decays as a power law ($\sim r^{-\eta}$) with $\eta_6 = 0.03(1)$ indicating quasi-long range orientational order. This is consistent with a 2D hexatic phase, however, in such a case, the translational order correlations are expected to decay exponentially. The translational correlation function, $G_T(\vec{r})$, does decay much faster than $G_6(\vec{r})$, however the decay shows a power-law dependence with $\eta_T = 0.77(1)$. The decay of the correlation functions observed here are consistent with those of a hexatic glass, a 2D crystal or hexatic
liquid which is pinned to random impurities (and/or interacting with the atomic lattice) [215, 219, 220, 221]. Indeed, CDW lattice-sized defects can be seen in the real space topography and slope images (fig. 5.6a,c and Supporting Information) which cause local distortions of the vortex lattice. Such weak pinning defects have been shown to have a small effect on $G_6(\vec{r})$ but a large effect on $G_T(\vec{r})$ in 2D, consistent with the observations made here [222, 223, 224].
5.2.2 Room Temperature Spectroscopy of the G-TaS$_2$ Surface

Figure 5.7: **Room Temperature Spectroscopy of G/TaS$_2$**

a) (Left image) Fourier filtered STM topography image of graphene on SiO$_2$ shows periodicity of the graphene lattice with disorder induced by the SiO$_2$ substrate. (Right) Tunneling conductance spectra (d$I$/d$V$) shows the V-shaped spectrum consistent with the density of states of clean, monolayer graphene. Spectrum is taken with a bias of 500mV and a current set point of 80pA approximately at the location marked with a square.

b) (Left image) Fourier filtered STM topography image of G/TaS$_2$ shows both the periodicity of the graphene lattice and the CDW. The graphene lattice remains honeycomb in appearance due to the atomically flat interface between the two materials. (Right) d$I$/d$V$ measured at the surface of G/TaS$_2$ shows a dip near zero bias and a shoulder near 550mV, consistent with previous reports of bare TaS$_2$ (without graphene). Spectrum is taken with a bias of -200mV and a current set point of -120pA approximately at the location marked with a square.

To investigate the electronic properties of the heterostructure, we perform scanning tunneling spectroscopy (STS). A typical spectrum for graphene on SiO$_2$ taken near the G/TaS$_2$ device, shown in figure 5.7a for reference, reveals the well-known V-shaped density of states of graphene. The fact that the minimum of the spectrum,
which marks charge neutrality, is located near zero bias (Fermi level), indicates only a small amount of extrinsic doping due to contaminants. A Fourier filtered atomic resolution topography image taken near the location of the tunneling spectrum shows the periodicity of the graphene lattice, however, the typical honeycomb registration is distorted by charge inhomogeneities and strain induced by the SiO₂ substrate. At the surface of G/TaS₂ (figure 5.7b), a typical spectrum shows two low energy features: a dip in density of states near zero bias and a shoulder (or a shallow dip) around 550mV. The dip near zero bias is a stable feature across the entire surface, however, the dip at positive bias varies slightly in depth between different spectra. Importantly, the spectra observed here are consistent with previous STM experiments taken on bare TaS₂ at this temperature [225, 224]. The fact that the spectrum of G/TaS₂ reflects that of bare TaS₂, and not the expected shifted Dirac spectrum of doped graphene, suggests that the graphene layer does not significantly interfere with the intrinsic electronic properties of TaS₂. The small role of the graphene carriers in the tunneling measurement is perhaps due to the large difference in density of states near the Fermi level (approximately 1000 times larger for TaS₂). The commonly adopted picture of conduction in this complex CDW phase asserts that conduction occurs within the domain walls surrounding insulating domains. However, it should be noted that in contrast to this notion, we do not observe gap-like features anywhere on the sample surface. The lack of a fully gapped spectrum is consistent with previous STM works [225, 224] and early ARPES measurements50 which show a strong reduction in Mott gap size within the NCCDW phase and a non-zero density of states at the Fermi level. This suggests that the Ta 5d orbitals are not strongly localized in this phase as is the case in the low temperature, commensurate CDW phase [146, 226, 120]. Further, despite the observed proximity-induced CDW in the graphene overlayer, the electronic spectrum is not fully gapped as expected of a Peierls-type CDW.
5.3 Comparison of Bare and Graphene Covered TaS$_2$ at 77K

5.3.1 Topography Comparison

Figure 5.8: **Comparing Bare and Graphene Covered TaS$_2$ at 77K** a) Ball and stick model of the $\sqrt{13} \times \sqrt{13}$ CDW reconstructed 1T-TaS$_2$ lattice (left panel) LDA computed low energy charge density shows how charge distributes in the commensurate CDW phase (right panel) b) STM topography image of the bare surface of 1T-TaS$_2$ shows the CDW charge modulation as well as the atomic lattice (marked with blue dots) (V$_b$=1.2V, ISP=40pA) inset: FFT of topography image c) Ball and stick model of graphene placed on 1T-TaS$_2$ in the commensurate phase (left panel) and LDA computed low energy charge density within the graphene layer shows a local modulation of doping with the periodicity of the CDW of TaS$_2$ (right panel) d) STM topography image of graphene placed on 1T-TaS2 shows both the graphene lattice as well as a larger modulation of charge density associated with the CDW of TaS$_2$ inset: FFT of topography image

Graphene on 1T-TaS$_2$ samples are cooled to approximately 77K inside a homebuilt STM under high vacuum (below $10^{-5}$ Torr). At this temperature we partially remove the encapsulating graphene layer using the STM tip, resulting in a region that allows us to differentiate the properties of bare and graphene covered 1T-TaS$_2$. The comparison of DFT calculated, low energy charge densities and measured STM topographies are given in figure 5.8. The 1T-TaS$_2$ surface exhibits the well-known $\sqrt{13} \times \sqrt{13}$ CDW with 13 Ta atoms per unit cell arranged in a Star of David (SD)
pattern shown in figure 5.8a (left). The lattice reconstruction leads to a modulation of the low energy charge density within the Ta atom plane shown in figure 5.8a (right). The STM topography measurement of the bare 1 T–TaS$_2$ surface (figure 5.8b) measures a mixture of the charge densities at the top and bottom sulfur layers as well as the Ta lattice evidenced by the observed honeycomb sulfur lattice with darker spots in between. The amplitude of the darker spots, associated with the Ta lattice, modulates in the $\sqrt{13} \times \sqrt{13}$ CDW pattern with the darkest spot at the center of each star. This observed pattern seems to be the inverse of the calculated low energy charge density, however, the higher energy charge density (near the bias voltage used in the measurement) is expected to have a larger electron density near the periphery of the SDs and is consistent with previous STM measurements of bulk 1 T–TaS$_2$ taken at large, positive biases [133].

With graphene on top, STM topography measures the honeycomb graphene lattice as well as a modulation with the period of the CDW in 1 T–TaS$_2$ (see figure 5.8d). The two periodic modulations are evident in the fast Fourier transform (FFT) of the topography image (figure 5.8d inset). It is known that in the commensurate CDW phase of 1 T–TaS$_2$, the sulfur lattice is slightly buckled out of plane on the order of 20 pm, but this does not account for the large corrugation observed experimentally. As STM topography measures a convolution of both the sample topography and charge density, the modulation observed with the period of the 1 T–TaS$_2$ CDW is due, in part, to local modulations of the charge density at the graphene layer surface.

To elucidate the observed behavior, we construct the $\sqrt{13} \times \sqrt{13}$ supercell and carry out geometry optimization of bare 1 T–TaS$_2$ and the G/TaS$_2$ heterostructure lattice model based on DFT. Fig. 5.8(a) left panel shows the geometrically optimized $\sqrt{13} \times \sqrt{13}$ CDW reconstructed 1 T–TaS$_2$ supercell. The David star CDW pattern given from DFT calculations agrees well with experimental observations. As shown in section 2.3.1, the origin of the CDW formation in 1 T–TaS$_2$ is the Kohn anomaly
around the middle of GM in the phonon spectrum. This Kohn anomaly in the reciprocal space corresponds to $\sqrt{13} \times \sqrt{13}$ supercell in the real space. The formation of the star-shape atomic arrangement involves the local lattice contraction around the center of the David star, in which the bond lengths between Ta ions are shorter in comparison with those between Ta ions outside the star. To compare with STM image, we calculated the partial charge density of the energy bands within [-0.1, 0.1] eV as shown in Fig. 5.8(a) right panel. The electron cloud mainly distributes over the stars with the highest charge density locating at the center Ta ion of the star. The atomic charge analysis shows much higher number of electrons accumulated at the center Ta ion.

Fig. 5.8(c) left panel shows the side view of graphene $5 \times 5$/TaS$_2$ $\sqrt{13} \times \sqrt{13}$ heterostructure. By subtracting the charge density of freestanding graphene monolayer from the charge density of graphene/TaS$_2$, Fig. 5.8(c) right panel illustrates the 1 T–TaS$_2$ induced hole-doping in graphene cover layer. The dark (negative) iso-charge contour indicates the missing electron density of graphene due to 1 T–TaS$_2$. The contour on the graphene layer shows the same periodicity as the CDW in 1 T–TaS$_2$ layer with the strongest charge transfer located around the center of the 1 T–TaS$_2$ CDW stars, demonstrating the correlation between graphene and 1 T–TaS$_2$. This star shaped charge transfer in graphene layer, as will be further explored later, provides clear evidence of the proximity effect.
5.3.2 Moiré Pattern between Graphene and TaS₂

Figure 5.9: Moiré Pattern between Graphene and 1 T–TaS₂ a) STM topography displaying a Moiré interference pattern as well as charge density wave modulation \((V_b = 0.550 \text{ V}, I_{SP} = 150 \text{ pA})\) b) Plot of the expected Moiré pattern wavelength between the graphene and 1 T–TaS₂ lattices as a function of angle. The experimentally determined value is marked with dashed lines c) Fourier transform of the topography image in (a) shows peaks associated with the graphene lattice \((a_g = 0.246 \text{ nm})\), TaS₂ lattice (measured \(a_{TaS_2} = 0.336 \text{ nm}\)), the induced Moir wavelength (measured \(\lambda_M = 0.906 \text{ nm}\)), and the commensurate CDW (measured \(\lambda_{CDW} = 1.22 \text{ nm}\))

In a different region of the structure and at a different bias voltage, we more clearly resolve the graphene and TaS₂ lattices (figure 5.9a). The structure observed is different than that of the expected commensurate CDW. The differences can be identified more clearly by viewing the FFT of the real space image (figure 5.9c). In the Fourier transform we can find 4 sets of Bragg peaks with different wavevectors. For accuracy, we compare the measured wavevectors to that of the graphene lattice (the outermost set of peaks in figure 5.9c). We find that the other 3 sets of peaks correspond to the periodic modulations of the TaS₂ lattice \((a_{TaS_2} = 0.336 nm)\), the commensurate charge...
density wave ($\lambda_{CDW} = 1.22\text{nm}$), as well as a modulation with period $\lambda = 0.906\text{nm}$. It is known that interference between two misaligned periodic patterns leads to a Moiré pattern, observable by STM [79]. The well-known relationship between the Moiré period and the crystal lattice mismatch is given by:

$$
\lambda_M = \frac{(1 + \delta)a_g}{\sqrt{2(1 + \delta)(1 - \cos \theta) + \delta^2}}
$$

(5.3)

For the lattice mismatch between the graphene and TaS$_2$ lattice, $\delta = (a_{TaS2}/a_g) - 1 \approx 0.366$, the Moiré wavelength as a function of angle is plotted at low angles in figure 5.9b. We can use the positions of the Bragg peaks associated with crystalline lattices to determine the angle between graphene and TaS$_2$ in this region to be approximately $\theta = 3\text{ deg}$. Thus, we identify the unknown periodic modulation as the Moiré pattern between the graphene and TaS$_2$ lattices.
5.3.3 CDW Proximity Effect in Graphene

Figure 5.10: Band Structures of $1\text{T}–\text{TaS}_2$ and Graphene on $1\text{T}–\text{TaS}_2$ a) LDSA+U band structure of $\sqrt{13} \times \sqrt{13}$ CDW reconstructed TaS$_2$ with $U=2.27\text{eV}$. Hubbard bands, associated with the localized electron state at the center of the David stars, are highlighted. b) LDSA+U band structure of G/TaS$_2$ using a phenomenological value $U=1.70\text{eV}$. Owing to the charge transfer from graphene to $1\text{T}–\text{TaS}_2$, the Fermi level (zero energy) moves from the lower Hubbard band to the upper Hubbard band with a graphene-associated Dirac cone intersects around $0.3\text{eV}$ above the Fermi level ($E_F$) at the K-point of the superstructure Brillouin zone. Hubbard bands and Dirac cone are highlighted for clarity. c) Zoomed in view of the crossing points between the Dirac cone and Upper (top) and Lower (bottom) Hubbard bands. d) Top view of G/TaS$_2$ heterostructure. Blue, brown, and yellow spheres indicate C, Ta, and S atoms, respectively. Black, gray, and red rhombuses show the $5\times5G/1313$ $1\text{T}–\text{TaS}_2$ supercell, graphene 1x1 unit cell, and $1\text{T}–\text{TaS}_2$ 1x1 unit cell, respectively. The graphene and $1\text{T}–\text{TaS}_2$ layer are twisted by $\sim13.9^\circ$ in this CDW phase. e) Side view of G/TaS$_2$ heterostructure overlaid with the charge density map corresponding to the states at the two crossing points of the Dirac cone and Lower Hubbard band.

The side and top views of the geometrically optimized lattice are depicted in Figs. 5.10c and d, respectively. To take the strong correlation of Ta-d electrons into consideration, we perform LSDA+U calculations with $U=2.27\text{eV}$ for bare $1\text{T}–\text{TaS}_2$ in
accordance with previous DFT calculations [133, 117] as shown in Fig. 5.10a. The CDW-induced isolated half-filled spin-degenerate flat band at $E_F$ splits into occupied spin up lower Hubbard band (LHB) and empty spin down upper Hubbard band (UHB) with a Mott gap of $\sim 0.25\text{eV}$ in between. The on-site Hubbard $U$ significantly enhances the Mott gap from $\sim 0.12\text{eV}$ to $\sim 0.25\text{eV}$ (Fig. 5.10a) so that the UHB touches the lowest conduction bands. While experiments observe a localized Hubbard band more separated from the conduction band. This discrepancy implies the on-site Hubbard $U$ of 2.27eV might be somewhat overestimated.

Graphene is known as a wide bandwidth metal with high-mobility itinerant carriers. With graphene on top of 1 T–TaS$_2$, there exist nontrivial charge transfer from graphene to 1 T–TaS$_2$ as shown in the LDSA+U band structure in Fig. 5.10b. These highly mobile electrons from graphene somewhat suppress the localized picture of the narrow Hubbard d-bands in 1 T–TaS$_2$ and thus reduce the on-site $U$ value of Ta. We have compared our band structures from different $U$ values with our STS results, suggesting that the graphene layer screens the Coulomb interaction in 1 T–TaS$_2$ and the Hubbard $U$ of Ta is lowered by $\sim 0.5\text{eV}$ due to the itinerant electrons from graphene. Therefore, we show in Fig. 5.10c the G/TaS$_2$ band structure with a phenomenological value $U=1.70\text{eV}$, which reasonably reproduces the trend observed in our STS measurements including the more separated UHB from the conduction band minimum.

As shown in Fig. 5.10b, both the 1 T–TaS$_2$ Hubbard bands and the graphene Dirac cone preserve in the heterostructure. While the Fermi level moves from the lower Hubbard band up to the upper Hubbard band due to the charge transfer from graphene to 1 T–TaS$_2$. This is also accompanied with an energy raise of the graphene Dirac point from $E_F$ up to $\sim 0.3\text{eV}$ above $E_F$ at the K-point of the superstructure Brillouin zone. Careful comparison of the Dirac cone of G/TaS$_2$ with that of bare graphene conclude that the Fermi velocity is not changed by 1 T–TaS$_2$. The highly dispersive $p_z$-orbital derived Dirac cone occupies the CDW and Mott gaps and crosses
both the d$_z^2$-orbital derived narrow LHB and UHB. They indeed interact with each other, though in a gentle manner. Four small gaps in meV order of magnitude can be found at the four crossing points by zoom-in the band structure shown in Fig. 5.10d. These small gaps originated from the weak interlayer couplings indicate finite hybridization between the Dirac and Hubbard states, giving rise to the proximity effect between graphene and 1T-TaS$_2$.

As mentioned previously, the contour in Fig. 5.8c right panel shows the hole density in graphene induced by 1T-TaS$_2$, which corresponds to the charge transfer from graphene to 1T-TaS$_2$. The missing electron density in graphene shows clear correlations with the CDW pattern in 1T-TaS$_2$, indicating a proximity-induced charge density wave in graphene. Figure 5.10e demonstrates the real-space interlayer hybridization given from the state around the Dirac-Hubbard bands crossing points in the reciprocal space. The interlayer charge density overlap shows the electron tunnel connecting graphene and 1T-TaS$_2$ layers, paving the way for proximity effect as discussed below.

The local charge transfer observed in the DFT calculation can be understood in the context of proximity-induced charge density waves. We can qualitatively describe such CDW proximity effect via a simple mean-field Hamiltonian.

\begin{equation}
H = H_d + H_c + H_t,
\end{equation}

\begin{equation}
H_d = \sum_{<i,j>,\sigma} -t_{ij}^d c_{i,\sigma}^\dagger c_{j,\sigma} + h.c.
- \sum_{<i',j'>,\sigma} \Delta_{i',j'}^{CDW} d_{i',\sigma}^\dagger d_{j',\sigma} + h.c. + \sum_{<i',j'>} |\Delta_{i',j'}^{CDW}|^2,
\end{equation}

\begin{equation}
H_c = \sum_{<i,j>,\sigma} -t_{ij}^c c_{i,\sigma} c_{j,\sigma} + h.c. = \sum_{k,\sigma} (\epsilon_k - \mu) c_{k,\sigma}^\dagger c_{k,\sigma},
\end{equation}

\begin{equation}
H_t = -t \sum_{i,\sigma} c_{i,\sigma}^\dagger d_{i,\sigma} + h.c.,
\end{equation}

(5.4)
where $H_d$ ($H_c$) stands for the simplified Hamiltonian of the $TaS_2$ (graphene) layer, respectively, and $H_t$ describes a weak charge transfer (hoping) term between these two layers. The insulating $TaS_2$ layer shows CDW order with the order parameter $\Delta^{CDW}_{d}(i', j') \equiv \sum_{\sigma} \langle d^\dagger_{i',\sigma} d_{j',\sigma} \rangle$ defined in the $\sqrt{13} \times \sqrt{13}$ unit cell with the "star" shape of CDW pattern and $i', j'$ are sites in this unit cell with CDW order. The graphene layer shows the electronic energy with linear Dirac spectrum: $(\epsilon_k - \mu) \sim v_F |k - k_F|$. Here, $i, j$ refer to the nearest-neighbour sites of the corresponding lattices, and $t^{(d)}_{i,j}$ refers to the nearest-neighbour tight-binding hoping terms on the graphene ($TaS_2$) layer, respectively.

Via the 2nd order perturbation in $H_t$ term over the CDW unit cell, the following exchange term $H^{(2)}_t$ is generated:

$$H^{(2)}_t = t^2 \sum_{\langle i', j' \rangle, \sigma} c^\dagger_{i',\sigma} d_{i',\sigma} d_{j',\sigma} + h.c. \tag{5.5}$$

A simple mean-field decoupling of $H^{(2)}_t$ in terms of $\Delta^{CDW}_d(i', j')$ (considering only $\sigma = \sigma'$ and assuming spin-isotropic CDW order $\langle d^\dagger_{i',\uparrow} d_{j',\uparrow} \rangle = \langle d^\dagger_{i',\downarrow} d_{j',\downarrow} \rangle$) gives $H^{(2)}_t \rightarrow H^{MF}_{t^2}$ with:

$$H^{MF}_{t^2} \approx -t^2/2 \sum_{\langle i', j' \rangle, \sigma} (\Delta^{CDW}_d(i', j'))^* c^\dagger_{i',\sigma} c_{j',\sigma}$$

$$- t^2/2 \sum_{\langle i', j' \rangle, \sigma} (\Delta^{CDW}_d(i', j'))^* \langle c^\dagger_{i',\sigma} c_{j',\sigma} \rangle + h.c. \tag{5.6}$$

where the mean-field decoupling $-t^2 \sum_{\langle i', j' \rangle, \sigma} \langle c^\dagger_{i',\sigma} c_{j',\sigma} \rangle d^\dagger_{j',\sigma} d_{i',\sigma} + h.c.$ term in $H^{(2)}_t$ is neglected since we expect $|\langle c^\dagger_{i',\sigma} c_{j',\sigma} \rangle| \ll |\langle d^\dagger_{i',\sigma} d_{j',\sigma} \rangle|$. The CDW proximity effect is manifested in $H^{MF}_{t^2}$ as a weak CDW order $\sum_{\sigma} \langle c^\dagger_{i',\sigma} c_{j',\sigma} \rangle$ is induced on graphene layer by the 2nd-order charge transfer between the two layers with the following
identification:

\[
\Delta_{c}^{CDW}(i',j') \equiv -\frac{1}{2} \sum_{\sigma} \langle c_{i',\sigma}^{\dagger} c_{j',\sigma} \rangle^* = -\frac{t^2}{2} (\Delta_{d}^{CDW}(i',j'))^*, \tag{5.7}
\]

or equivalently, \( \sum_{\sigma} \langle c_{i',\sigma}^{\dagger} c_{j',\sigma} \rangle = t^2 \Delta_{c}^{CDW}(i',j') \). Via the above identification, the Hamiltonian \( H_{t_2}^{MF} \) can be expressed as:

\[
H_{t_2}^{MF} = \sum_{\langle i',j' \rangle,\sigma} \Delta_{c}^{CDW}(i',j') c_{i',\sigma}^{\dagger} c_{j',\sigma} + h.c. + 2 |\Delta_{c}^{CDW}(i',j')|^2, \tag{5.8}
\]

which leads to

\[
\Delta_{c}^{CDW}(i',j') = -\frac{1}{2} \sum_{\sigma} \langle c_{i',\sigma}^{\dagger} c_{j',\sigma} \rangle^* \text{ identified in Eq. (5.7) via minimizing the free energy associated with } H_{t_2}^{MF} \text{ in Eq. (5.8) with respect to } \Delta_{c}^{CDW}(i',j').
\]

Note that from above derivations, we indeed find that \(|\langle c_{i',\sigma}^{\dagger} c_{j',\sigma} \rangle| \sim t^2 |\langle d_{i',\sigma}^{\dagger} d_{j',\sigma} \rangle| \ll |\langle d_{i',\sigma}^{\dagger} d_{j',\sigma} \rangle|\), as expected. Note also that the CDW order parameters induced on graphene layer shows the opposite sign with respect to that on TaS2 layer, consistent with the hole-like (particle-like) CDW intensity on graphene (TaS2) layer obtained from DFT calculations, respectively.
5.3.4 Spectroscopy at the Surface of G/TaS$_2$

Figure 5.11: **Density of States of Graphene on Commensurate 1T–TaS$_2$**

a) LDSA+U calculated density of states of bare 1T–TaS$_2$ (U=2.27eV) b) Measured tunneling spectroscopy of the bare 1T–TaS$_2$ surface shows a gap at zero bias with peaks associated with upper and lower Hubbard bands (highlighted with violet and blue, respectively) in agreement with the calculated result (a). c) LDSA+U calculated density of states of the G/TaS$_2$ heterostructure showing features resembling the Hubbard bands, conduction band, and valence band of 1T–TaS$_2$, but with a reduced Mott-Hubbard interaction strength (U=1.70eV). States which lie between the Hubbard-like bands project to the Dirac cone states within the graphene layer. d) Measured tunneling spectroscopy of G/TaS$_2$ qualitatively agrees with the calculated result in (c), displaying Hubbard-like peaks, with a gap size that is reduced with respect to the Mott gap observed in bare TaS$_2$, and mid-gap states which we associate with the graphene layer. Spectra in (b),(d) are taken at a set point of [$V_b = 1.2$V, $I_{SP} = 80$pA] with AC modulation $V_{AC} = 8$mV. Lower and upper Hubbard peaks are highlighted in blue and violet, respectively.

Next, we compare tunneling spectra (STS) in both the bare and encapsulated 1T–TaS$_2$ regions. Tunneling spectroscopy measured at the bare 1T–TaS$_2$ surface
is given in figure 5.11b. In accordance with previous STS measurements, we find a Mott gap at the Fermi level flanked by lower and upper Hubbard peaks at -256mV and 144mV, respectively. Careful inspection shows mid-gap states near (but not directly at) the Fermi level which are typically associated with out-of-plane transport. Additionally, two large dips in density of states can be seen at -378mV and 218 mV are associated with the known charge density wave gap. We note that our thin, exfoliated 1T–TaS$_2$ flake (approx. 40 layers) displays a similar spectrum to that reported on bulk, vacuum-cleaved samples indicating that the exfoliation and heterostructure fabrication process is not highly damaging to the 1T–TaS$_2$ surface.

With graphene on top, the tunneling spectrum (figure 5.11d) shows features reminiscient of bare 1T–TaS$_2$ as well as some notable differences. Namely, we observe two large peaks near the Fermi level at -280mV and 88mV, flanked by dips in the density of states. Due to their similarity to the Hubbard peaks observed in bare 1T–TaS$_2$ we label those seen at the graphene surface as lower and upper Hubbard peaks. Unlike the bare 1T–TaS$_2$ surface, we observe a finite density of states at the Fermi level, presumably due to states confined within the graphene layer. Unlike pristine graphene on an insulating substrate, the density of states does not linearly go to zero at the Fermi level. This is consistent with room temperature measurements which have shown substantial charge transfer between the two materials when placed in contact. The change in carrier concentration shifts the Dirac point of pristine graphene to higher energies. Comparing to the bare 1T–TaS$_2$ case, we see that the Hubbard peaks are shifted in energy and their separation is reduced. We expect that the graphene layer contributes to screening the Coulomb interaction near the surface of 1T–TaS$_2$, reducing U and the separation between Hubbard peaks.

LSDA+U density of states (DOS) of bare 1T–TaS$_2$ shown in Fig. 5.11(a) reflect the key features of 1T–TaS$_2$ band structures in Fig. 5.10(a). The lower and upper Hubbard bands locate respectively below the Fermi level and around 0.35 eV above the Fermi level with a $\sim$0.3 eV Mott gap in between. The dip below LHB and above
UHB indicate the scale of the CDW gap (Supplementary Figure S7). These main features demonstrate good consistency with those observed in our STS measurement shown in Fig. 5.10(b). The dip above UHB given by on-site \( U = 2.27 \) eV is relatively narrower in comparison with the dip below LHB and with that observed in STS (Fig. 5.11b). The overestimated Mott gap implies that the correct \( U \) value might be smaller than the previously used value of 2.27 eV.

With graphene on top of 1 T–TaS\(_2\), the LSDA+U DOS plotted in Fig. 5.10(c) also show UHB, LHB, Mott gap, and CDW gap similar to the DOS of bare 1 T–TaS\(_2\). The main differences from the bare 1 T–TaS\(_2\) case include 1. The relative energies among the Fermi level, UHB, LHB, and Dirac point. 2. The linear-like dispersion states within the Mott gap. Owing to the relatively stronger electron negativity of 1 T–TaS\(_2\), there exist notable charge transfer from graphene to 1 T–TaS\(_2\). Consequently, the Fermi level moves from the LHB top to the UHB bottom, meanwhile the graphene Dirac point shifts to a higher energy about 0.3eV above \( E_F \). On the other hand, the mid-gap linear band comes from the Dirac cone states of the graphene cover layer. Because of the energy shift of the Dirac point to \( \sim 0.3 \) eV above \( E_F \), the V-shape DOS thus moves to \( \sim 0.3 \) eV, leaving a linear-like band within the Mott gap below \( E_F \). As mentioned previously, we adopt a smaller on-site \( U = 1.70 \) eV in G/TaS\(_2\) case to mimic the screening effect observed in our STS caused by the itinerant electrons from graphene, the UHB is more separated from the conduction band minimum with a wider dip above UHB. The overall trend given from LSDA+U are in good agreement with our STS results shown in Fig. 5.11(d).
Chapter 6
Conclusions

In summary, this thesis work provides novel instrumentation, techniques, and measurements which locally probe charge transfer and distributions in van der Waals heterostructures built from 2D materials.

Toward this goal, we construct an environmentally controlled electrostatic AFM measurement system which can measure electrostatic properties of van der Waals heterostructures. A versatile scanning tunneling microscope system is developed and noise optimized to measure local electronic properties of conducting and semiconducting 2D materials supported on insulating substrates. We apply heterostructure fabrication techniques, borrowed from the transport community, in air and in an inert atmosphere glovebox. With these advancements, we are poised to study local charge distributions in heterostructures on insulating substrates and relate these to global properties observed in transport measurements.

We image atomically thin conductors through an insulating encapsulant using electrostatic AFM. Applied to graphene encapsulated with hexagonal boron nitride, the conducting graphene is observed through tens to hundreds of layers of hBN. The electrostatic force gradient at the heterostructure surface varies strongly above the edge of the encapsulated graphene, allowing the identification of the atomically thin conducting layer on sub-micron length scales. Thus, the precise location of the graphene edge, including the presence of cracks and contamination, can be obtained at higher resolution and contrast than possible using light-based microscopy. The surface potential variations, measured by Kelvin probe force microscopy, are directly
related to the charge distribution within the heterostructure. Doping variations due to charge impurities, traps at the interfaces, and material differences (i.e. monolayer vs. bilayer) are all readily and quickly attainable with KPFM. These methods are performed without electrical contact to the sample and can thus monitor the location and orientation of the embedded graphene sheets as well as the presence of electronic disorder at various stages of device fabrication. These details are directly beneficial in the construction of high-quality encapsulated graphene transport devices which are currently receiving great interest from the scientific research community.

These methods and instrumentation are applied to the study the electronic properties at the surface of the correlated 2D material, 1T-TaS$_2$. Measuring the KPFM-derived surface potential as a function of time, we measure a characteristic timescale on which the surface of 1T-TaS$_2$ oxidizes in ambient conditions, an age-old problem which plagues the 2D material and scanning probe research communities. Using simple heterostructure fabrication techniques, placing a layer of graphene on top of 1T-TaS$_2$ extends the surface lifetime by several orders of magnitude by acting as a gas barrier. Similar to hBN-encapsulated graphene samples, it is demonstrated that electrostatic AFM modes probe charge transfer and charge disorder within the heterostructure on sub-micron length scales. Local electrostatic AFM measurements reveal that the slow oxidation of the surface of 1T-TaS$_2$ persists beneath the graphene layer due to intercalation of air into the van der Waals gap between the two materials. Not only does this work demonstrate the versatility and value of scanning probe measurements toward novel discoveries and device optimization but also provides a demonstration of the ability of graphene to be used as an inert encapsulating material. These capabilities are particularly useful for fabrication of devices made from thin, exfoliated films of air-sensitive materials which require exposure to the environment during fabrication.

On atomic length scales, STM measurements reveal that the charge density modulations of the correlated phases of 1T-TaS$_2$ can be imaged through the covering
graphene layer. Detailed variations of the charge density wave order parameter can be mapped in space. At room temperature, we identify the TaS$_2$ in the nearly commensurate charge density wave state which exhibits a nearly-periodic domain structure. As the domain walls are marked by phase shifts, the 1D topological defects of the 2D CDW state, they are related to vortices in type II superconductors. We extend the analogy by demonstrating that the interacting domain wall lattice orders similar to the observed Abrikosov vortex lattices in type II superconductors forming a two dimensional hexatic glass state, a rarely observed state of matter. As 1T-TaS$_2$ hosts a variety of exotic, correlated physics, accessible at room temperature, our method of protecting the surface without requiring vacuum makes the study of these effects much more readily available to researchers and educators.

Observing a charge density wave state through a thin conductor is unprecedented and opens new questions about the electronic states induced at the interface of the two materials. The charge density wave is again seen within the graphene layer at 77K, where 1T-TaS$_2$ is in a commensurate charge density wave state which is made semiconducting by strong Coulomb interactions. With the commensurate structure, we connect experiment with calculations based on density functional theory. In accordace with STM measurements, DFT calculations show a non-trivial, local charge transfer between graphene and 1T-TaS$_2$ which modulates with the periodicity of the host CDW. This effect is similar to the superconducting proximity effect where correlated electronic states are induced in a material placed in contact with a superconductor. A toy model describing the charge density wave proximity effect is introduced which can be easily extended to consider other systems in contact with the surface of charge density wave materials.

Tunneling spectroscopy at the heterostructure surface further explores novel effects due to the interaction of the two systems at their interface. The tunneling spectrum reveals neither of the spectrum expected of graphene or that of bare 1T-TaS$_2$, but a convolution of the two. As contact proximity effects work both ways, mid-gap
states are induced in TaS$_2$. Together with itinerant carriers within the graphene layer, these screen Coulomb interactions at the surface of TaS$_2$, reducing the size of the intrinsic Mott gap. Thus, we not only identify the effect of contact between graphene and the correlated charge density wave state, but also the effect of contact with a Mott insulating state.

The ideas presented above open new research avenues and pave the way for novel experiments which further probe our observations. For the G/TaS$_2$ system in particular, low temperature transport and Raman spectroscopy measurements are needed to determine the existence of a periodic lattice distortion, single particle gap, or sliding of the proximity induced CDW. Further, scanning probe measurements at the edge of the structure might reveal the coherence length of the induced CDW as it extends into the graphene layer beyond the edge of the 1T–TaS$_2$ flake. Similar to superconducting Josephson junctions, phase coherence in a proximity CDW would lead to new logic devices based on correlated physics. Additionally, the effect of various external knobs such as electrostatic gating (doping), electronic and optical pulses, in-plane electric fields, and magnetic fields has yet to be explored in such a system. Finally, the methods presented are kept general and can be adapted and applied to new materials for future discoveries at the interfaces of various electronic systems.
Bibliography


[56] A. Jain and M.A. Alam. Prospects of hysteresis-free abrupt switching (0

[57] Asif Islam Khan, Debanjan Bhownik, Pu Yu, Sung Joo Kim, Xiaoqing Pan,
Ramamoorthy Ramesh, and Sayeef Salahuddin. Experimental evidence of fer-
roelectric negative capacitance in nanoscale heterostructures. *Applied Physics
Letters*, 99(11), 2011.

[58] Victor V. Zhurinov and Ralph K. Cavin. Negative capacitance to the rescue?

[59] S. Salahuddin and S. Datta. Use of negative capacitance to provide voltage

[60] Adrian M. Ionescu and Heike Riel. Tunnel field-effect transistors as energy-

[61] Isao H. Inoue and Marcelo J. Rozenberg. Taming the mott transition for a novel

[62] K. Gopalakrishnan, P.B. Griffin, and J.D. Plummer. I-mos: a novel semicon-
ductor device with a subthreshold slope lower than kt/q. In *Electron Devices


[64] Dean Moldovan. Tight-binding package for python.

[65] Sinisa Coh and David Vanderbilt. Python tight binding (pythtb).

[66] Lon Van Hove. The occurrence of singularities in the elastic frequency distri-


stability of fermi points in multilayered graphene. *Physical Review B*, 75(15),
2007.

[69] J. L. Maes. Symmetry-based approach to electron-phonon interactions in

[70] Vitor M. Pereira, A. H. Castro Neto, and N. M. R. Peres. Tight-binding ap-

P. Kim, J. Y. Choi, and B. H. Hong. Large-scale pattern growth of graphene


[182] Nikhil Tilak, 2019. Thanks to Nikhil Tilak and Xinyuan Lai for showing me how to etch tungsten tips.


[197] Sheng-Hsiung Hung, Horng-Tay Jeng, and Chung-Hou Chung, 2020. We thank Prof. Horng-Tay Jeng and Sheng-Hsiung for providing DFT calculations of bare and graphene covered 1T-TaS2 as well as Prof. Chung-Hou Chung for providing his model describing the CDW proximity effect.


