Exploring Valleytronics in 2D Transition Metal Dichalcogenides

by
Brian Joseph Modtland

B.S., Electrical Engineering, Iowa State University (2012)
M.S., Electrical Engineering, Iowa State University (2013)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2018

© Massachusetts Institute of Technology 2018. All rights reserved.
Exploring Valleytronics in 2D Transition Metal Dichalcogenides

by

Brian Joseph Modtland

Submitted to the Department of Electrical Engineering and Computer Science on January 31, 2018, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering

Abstract

Monolayer transition metal dichalcogenides (TMDs) exhibit distinct electrical and optical properties according to the relative occupation of each of two valleys in their dispersion relation. The resulting valley degree of freedom is robust, linked to a large spin-orbit splitting between valence bands, and shows promise in electro-optical devices or as an information token for logic applications. In order to explore applications of these properties, monolayer crystals are required that have reduced intervalley scattering. To date, the majority of valley-related studies have focused on exfoliated samples for their quality and ease of production. In this thesis, valley polarization is explored in monolayer tungsten disulfide (WS$_2$) synthesized by chemical vapor transport (CVT). This novel method of bottom-up growth relies on halide-driven vapor transport commonly utilized in bulk crystal growth. Using a small amount of sodium chloride salt as a source of chlorine, non-volatile WS$_2$ can react to form gaseous tungsten chloride and sulfur. With an open tube system, a controlled reaction generates mono- and few-layer WS$_2$ crystals. These crystals have excellent optical properties and exhibit a degree of valley polarization near 50% at 77 K and up to 30% at room temperature. This surpasses previous values reported in WS$_2$. By decoupling pump photon and thermal energy, valley depolarization shows the characteristics of an electron-hole exchange interaction rather than nonradiative scattering. These results offer the initial groundwork for future devices that use the coupled valley-spin degree of freedom as a robust token of information, promising reduced power consumption compared to conventional MOSFET-based electronics.

Thesis Supervisor: Marc A. Baldo
Title: Professor of Electrical Engineering and Computer Science
Acknowledgments

This thesis is dedicated to my wife, Samantha. She was the rock of support I always needed to make it through tough days, months, and years. She delayed some of her plans to let me chase my passion and I love her dearly for it. I hope I can repay her never-ending patience and kindness as we grow old together. Per Lee Brice, "I don’t know what I’d do if I’d never met a woman like you."

My Ph.D. was full of growth, professionally and personally. I worked on projects I could have never imagined when I started at MIT. I would like to thank my advisor Prof. Marc Baldo for letting me slug through the mud to find a project that I was passionate about. I carved my own path and experienced a lot of rough terrain, but Marc was always optimistic of what was around the next turn. My ability to conduct science and generate interesting questions worth solving developed by leaps and bounds under his guidance. I am hopeful another Cyclone finds their way to his group; the legacy should not stop with me.

I want to also thank Dr. Jing Kong and Dr. Will Tisdale for acting as members of my thesis committee. Jing was invaluable in the growth work contained in my thesis and her discussions were critical in changing the direction of the project when it was not proceeding as planned. Will also had a knack for asking questions during our discussions that led to critical findings and I thank him for his critiques.

Next, I thank Tony Wu and Farnaz Nirouei for acting as sounding boards for most of the ideas I had during my Ph.D. They both taught me a lot during our conversations, but they also acted as amazing friends during my 4.5 years at MIT. Additional students who deserve thanks for their assistance and collaboration includes Efren Navarro-Moratalla, Joe Finley, Phil Reusswig, Xiang Ji, Dan Congreve, and Jean Ann Incorvia. I also need to thank members of Dr. Bulovic’s optics subgroup for teaching me all about optics, especially Tony Zhu and Tom Mahony. Many staff members around MIT were vital to my ability to complete projects, including Mark Belanger (Edgerton Machine Shop), Mark Mondol (NSL), Jim Daley (NSL), Kurt Broderick (MTL), Wilfred Mbah (EHS), and Tim McClure (CMSE).
Last but not least, I want to thank my parents, Denise and David, who instilled in me the values of respect, hard work, and humility. My mother taught me to think outside the box and chase my dreams, while my father taught me how to problem solve and take pride in your work. Together they supported me, encouraging me to ask questions and be curious, which led me to science and engineering.
Contents

1 CMOS and Beyond .............................. 17
   1.1 CMOS Technology .......................... 18
      1.1.1 Conventional (Dennard) Scaling ............ 19
      1.1.2 Recent CMOS Updates ........................ 20
      1.1.3 Challenges to Moore’s Law .................. 21
   1.2 Extended CMOS ............................... 25
      1.2.1 Tunneling FETs ............................. 26
      1.2.2 Negative Capacitance with Ferroelectrics .... 27
   1.3 Beyond CMOS ................................. 29
      1.3.1 MEMS/NEMS ................................. 29
      1.3.2 Spin-based Logic ............................ 30
   1.4 Valleytronics ................................. 32
      1.4.1 Early Work on Valleytronics ................. 33
      1.4.2 The Valley Degree of Freedom in 2D Hexagonal Crystals ... 34
   1.5 Summary and Outlook .......................... 35

2 Valley Physics in 2D Transition Metal Dichalcogenides ...... 37
   2.1 Crystal Structure and Brillouin Zone ............. 38
   2.2 Electronic Band Structure - Energy Dispersion .... 40
      2.2.1 Other Interesting TMDs ....................... 42
   2.3 Excitons and Trions in 2D TMDs .................... 43
   2.4 Valley Physics ................................. 47
      2.4.1 Valley-Spin Coupling ........................ 47
2.4.2 Berry Curvature and the Valley Hall Effect .................................. 50
2.4.3 Valley Connection to Circularly Polarized Light .............................. 53
2.4.4 Valley Zeeman Effect ................................................................. 55

3 Growth of High-quality Monolayer Tungsten Disulfide (WS₂) via Chemical Vapor Transport 59

3.1 Synthesis of Monolayer TMDs ....................................................... 60
3.1.1 Exfoliation ................................................................................. 60
3.1.2 Chemical Vapor Deposition ..................................................... 61
3.1.3 Physical Vapor Deposition ....................................................... 62
3.2 Chemical Vapor Transport of Bulk Crystals ..................................... 62
3.3 Synthesis of Monolayer WS₂ via CVT ............................................ 64
3.3.1 Process Details for WS₂ Growth ............................................... 68
3.3.2 Thermodynamic Calculations .................................................. 69
3.3.3 Evidence of Growth by Chlorine Transport .................................. 72
3.3.4 Growth with Different Alkali Salts .......................................... 74
3.3.5 XPS of Impurity Circles ............................................................ 74
3.3.6 Attempts with Other TMD Crystals .......................................... 76

4 Characterization of WS₂ Crystal Quality 79

4.1 Optical Microscopy ........................................................................... 79
4.2 Atomic force microscopy (AFM) ...................................................... 80
4.3 Raman Spectroscopy ....................................................................... 81
4.4 Photoluminescence Spectroscopy ................................................... 83
4.5 Photoluminescence Mapping .......................................................... 85
4.6 Time-dependent Photoluminescence .............................................. 85
4.6.1 Details of TCSPC ...................................................................... 87
4.6.2 Modeling of Time-dependent Photoluminescence ....................... 88

5 Valley Polarization Physics in CVT-grown WS₂ 91

5.1 Brief Review of the Valley Pseudospin in Monolayer TMDs .............. 93
5.2 Measuring Valley Polarization via Polarization-dependent Photoluminescence Spectroscopy ........................................... 94
   5.2.1 Polarization Calibration ............................................. 95
5.3 Valley Polarization as a Function of Pump Photon Energy .......... 96
   5.3.1 Model for Valley Polarization Dynamics ....................... 98
   5.3.2 Valley Depolarization Mechanisms ............................... 99
5.4 Valley Polarization at Fixed $\Delta E$ versus Temperature .......... 102
   5.4.1 Note on Exciton-Exciton Exchange ............................... 103
5.5 Summary of Valley Polarization ....................................... 104

6 Valleytronics: Applications of the Valley Pseudospin .................. 105
   6.1 Valley Splitting on Magnetic Substrates .......................... 106
      6.1.1 Proposal to Use Rare Earth Iron Garnets .................... 108
      6.1.2 Measurement Techniques ....................................... 109
      6.1.3 Preliminary Results and Discussion .......................... 114
   6.2 Valley Optoelectronics ............................................. 116
   6.3 Valley-Spin FETs ................................................... 116
   6.4 Valley Optical Modulators ......................................... 118
   6.5 Prediction on the Future of Moore’s Law .......................... 121
   6.6 Conclusion .......................................................... 121

A Fabrication Techniques for 2D Materials ............................... 123
   A.1 Micro-transfer of 2D Samples ..................................... 123
      A.1.1 PMMA Transfer with Hydrofluoric Acid ....................... 124
      A.1.2 PDMS Liftoff with Water ...................................... 125
      A.1.3 Polycarbonate Pickup ......................................... 126
      A.1.4 Transfer Setup ................................................. 127
   A.2 Electron Beam Lithography and Metal Liftoff ..................... 127
List of Figures

1-1 Moore’s Law observed as the number of transistors on each integrated chip. .................................................. 18
1-2 Table of MOSFET parameters for ideal scaling. ............... 19
1-3 Concept of a FinFET transistor. .................................. 20
1-4 Concept of a Gate All Around FET (GAA-FET). ................ 21
1-5 CMOS power density and supply voltages over time. .......... 22
1-6 Subthreshold swing for MOSFETs. .............................. 24
1-7 Tunneling FETs to reduce the subthreshold swing. .......... 26
1-8 Negative Capacitance FETs using ferroelectrics. ............. 28
1-9 Taxonomy of Beyond CMOS technology as presented in the ITRS 2015 report .................................................. 28
1-10 Concept of a N/MEMS Switch. .................................. 30
1-11 Domain wall logic based on spin-transfer torque. ........... 31
1-12 Concept of logic based on the relative occupation of valleys in k-space. ................................. 33
1-13 Summary of CMOS and next-gen logic technologies, showing their current status. .................. 36

2-1 Periodic table of elements, highlighting those that form layered TMDs. .................. 37
2-2 Common crystal phases for monolayer TMDs .................. 39
2-3 Reduced Brillouin Zone (BZ) for monolayer TMDs. ........... 40
2-4 Calculated Electronic Band Structure for WS₂ with spin-orbit coupling included. ................. 41
2-5 Periodic table of layered TMDs from group IV to group VII of the transition metals. .......................... 42
2-6 Excitons in 3D semiconductors versus 2D. .......................... 44
2-7 Absorption spectra for WS$_2$ for different number of layers. .......................... 44
2-8 Photoluminescence vs. Gate Voltage for WS$_2$ .......................... 45
2-9 Exciton dynamics for monolayer TMDs at various pump fluences, both before and after chemical treatment with TFSI superacid. .......................... 46
2-10 Crystal lattice for 2H-phase TMDs. .......................... 48
2-11 Coupled spin and valley physics in monolayer TMDs. .......................... 49
2-12 Berry phase in a closed path on the surface of a sphere. .......................... 50
2-13 Berry curvature of bands in monolayer TMDs. .......................... 51
2-14 Valley Hall effect in MoS$_2$. .......................... 52
2-15 Optical selection rules for circularly polarized light. .......................... 54
2-16 Valley Zeeman Effect in Monolayer TMDs. .......................... 56

3-1 General concept of CVT growth in a closed ampule system. .......................... 63
3-2 Concept of CVT growth of WS$_2$ with an open-tube design. .......................... 65
3-3 Optical micrographs of WS$_2$ growth. .......................... 67
3-4 Growth using chlorine gas diluted in argon. .......................... 73
3-5 Micrographs of CVT growth using different alkali halide salts. .......................... 74
3-6 XPS data from defect circles. .......................... 75
3-7 Growth of WSe$_2$ using NaCl salt. .......................... 76
3-8 Growth of WTe$_2$ using AgCl. .......................... 77

4-1 Optical micrographs of monolayer WS$_2$ samples grown by CVT with NaCl. .......................... 80
4-2 Atomic Force Micrograph of a typical WS$_2$ flake. .......................... 81
4-3 Vibrational modes common in few layer TMDs. .......................... 81
4-4 Raman spectrum for exfoliated and CVT monolayers .......................... 82
4-5 Photoluminescence from CVT and exfoliated monolayers. .......................... 83
4-6 Normalized photoluminescence spectra before and after flake transfer. 84
6-11 Concept for Valley-Spin PFET. .......................... 117
6-12 Integrated silicon electro-optical modulator. ................. 119
6-13 Proposal for Valley-based Electro-optical Modulators. ....... 120

A-1 PMMA-mediated transfer process using HF acid. ............... 125
A-2 Direct pickup transfer process using polycarbonate (PC) films. 126
A-3 Photographs of the home-built transfer setup. ................. 127
A-4 Ebeam Lithography. ............................................. 128
List of Tables

3.1 Standard thermodynamic data used in reaction calculations. [1] 71
3.2 Thermodynamic calculations of various reaction pathways involving chlorine gas. 72
4.1 Table of previous exciton lifetimes reported for WS$_2$ 86
Chapter 1

CMOS and Beyond

Complementary Metal-Oxide-Semiconductor Field-effect Transistor (MOSFET) fabrication has been one of the defining technologies of the last half-century with dramatic implications for our computer-driven society. Since the 1960s, our society has advanced to the point where the computational power that took Americans to the moon can be had for pennies and exists in children’s toys. The ubiquity and proliferation of powerful general-purpose computer processors is now so common that applications such as Artificial Intelligence and self-driving cars are becoming mainstream. Smartphones went from expensive novelty to permanent placement in nearly every person’s pocket in just a few years. The Internet and the social constructs that were formed as a result of its formation were a direct result of this cheap access to computing power.

This trend of decreasing cost for computational power roughly every 18-24 months was noted by Gordon Moore, co-founder of Intel, in 1965. For the past 50 years, Moore’s Law has rung true. Semiconductor companies have been able to double the number of transistors on every chip, at roughly the same chip cost, every couple of years.

There is little doubt that the ability to continuously improve the cost of computational power has changed the world forever. Yet, for all its success, CMOS technology is reaching a physical limit that must be confronted if we hope to keep up with this rapid pace of technological advancement. While some minor-changes can be made to
Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Moore's Law describes the empirical observation that the number of transistors on integrated circuits doubles approximately every two years. This observation is important as other aspects of technological progress, such as processing speed or the price of electronic products, are strongly linked to Moore's Law.

Figure 1-1: Moore's Law observed as the number of transistors on each integrated chip. Since the area and cost of computer chips has remained somewhat constant, this can be thought of as the inverse of the cost of each transistor, which was the original observation by Gordon Moore. In 2017, chips contain around 10 billion transistors, an unbelievable number. This plot is available from OurWorldInData.org, created by Max Roser and licensed under CC-BY-SA using data available on wikipedia.org/wiki/Transistor_count.

keep CMOS marching on, eventually new tokens of information will be required to reduce the power dissipated on-chip. One such proposal is to use carrier valleys in a material's E-k dispersion relation as discrete bearers of information. By choosing materials with distinguishable valleys that can be accessed electronically or optically, this degree of freedom has the potential for low power computation.

1.1 CMOS Technology

Complementary Metal Oxide Semiconductor (CMOS) technology is the dominant technology for creating digital logic circuits. The complementary part of CMOS means that actual gates are built with electron-rich (n-type) and hole-rich (p-type) devices which act as complements to each other. The n-type devices act as the active-high, pull-down components while the p-type transistors are active-low, pull-up devices. Acting together, ideal CMOS gates have zero static power dissipation as
no path from $V_{DD}$ to ground exists once the load capacitance is fully (dis)charged. This is the real advantage of CMOS versus other MOS-based schemes. For the matter of this thesis, MOSFETs are the primary concern, not the gate architecture, but it is worth noting as the problems with modern MOSFETs relate back to the idea of power dissipation.

1.1.1 Conventional (Dennard) Scaling

Up until the mid-2000s MOSFETs were shrinking, allowing them to operate faster due to the decreasing channel width. Clock rates increased from a few MHz in the 1970s to a maximum of 3-4 GHz in the 2000s. During this time, scaling of the transistor's dimensions followed from advances in fabrication technology. As the gate oxide decreased in thickness by the scaling factor (roughly 1.4 per technology node to produce a 2x improvement in transistor density), the electric field was kept constant by decreasing the supply voltage and on-state current. Frequency would increase by the same scaling factor. These scaling relationships were known as Dennard scaling after Robert Dennard who noted the changes to MOSFET parameters under ideal scaling (see Figure 1-2). [2]

<table>
<thead>
<tr>
<th>Device or Circuit Parameter</th>
<th>Scaling Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device dimension tox, L, W</td>
<td>1/K</td>
</tr>
<tr>
<td>Doping concentration Na</td>
<td>K</td>
</tr>
<tr>
<td>Voltage V</td>
<td>1/K</td>
</tr>
<tr>
<td>Current I</td>
<td>1/K</td>
</tr>
<tr>
<td>Capacitance eA/t</td>
<td>1/K</td>
</tr>
<tr>
<td>Delay time per circuit VC/I</td>
<td>1/K</td>
</tr>
<tr>
<td>Power dissipation per circuit VI</td>
<td>1/K^2</td>
</tr>
<tr>
<td>Power density VI/A</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 1-2: Table of MOSFET parameters for ideal scaling where K>1 for shrinking transistor dimensions. Table is adapted from Dennard et al. [2].

As a side effect of Dennard scaling, power dissipation per device decreased by $K^2$ as the area decreased by the same factor. Thus, power density was constant. However, the ideal MOSFET scaling could not continue forever. In the mid-2000s, voltage
scaling was outpaced by frequency scaling due to the thirst for speed. Reliability became a problem and power densities climbed as a result. Eventually, the limit on power dissipation for a fan-cooled CPU was reached, around 100 mW cm$^{-2}$, and clock frequencies stopped scaling (see Figure 1-5). This was known as the power wall. Ever since, computational power has improved not by increasing speeds, but by changing architecture, namely with parallel processor cores. This allows multiple computations to be done simultaneously. The end of Dennard scaling was a warning that the next 10-15 years of semiconductor fabrication was going to require a lot of complicated tricks to keep the train on the tracks in regards to Moore’s Law.

1.1.2 Recent CMOS Updates

Starting around the 22 nm node, a dramatic change occurred in MOSFET fabrication. In prior process nodes, high-K dielectrics were introduced to allow the gate oxide to effectively shrink in thickness without resulting in large leakage current. To maintain the necessary electrostatic behavior, the gate oxide needed to thin even further. Rather than come up with another dielectric, fabs instead changed the channel geometry. By extending the channel upward and having the gate surround it on three sides, a pseudo-nanowire was created. This improved the gate control of the channel enough to reduce threshold voltages and keep power density within limits. These devices were called Tri-gates or FinFETs due to the three-sided gate structure that resembled a fin sticking up. [3]

![Figure 1-3: Concept of a FinFET transistor. Starting around the 22 nm node, FinFET (Tri-gate FETs) were introduced to improve electrostatic gate control once the gate oxide thickness could not be decreased any further. Image is courtesy of Intel Corp.](image)

As semiconductor companies reach the 10 nm node in 2018 (about 2 years behind
schedule) and sub-10 nm in the future, the problem of gate control is expected to again force the geometry to change. Rather than a single fin, channels are predicted to go to the ideal MOSFET limit of a nanowire surround on all sides by the gate. [4,5] This has been called a GAA-FET, standing for gate all-around FET. The challenge of growing Si (or possibly III-V) nanowires and surrounding it with oxide and a gate electrode seems daunting on the scale of billions of devices.

Figure 1-4: Concept of a Gate All Around FET (GAA-FET), where nanowire channels (either vertical or horizontal) are grown with the gate (and oxide) completely surrounding it. This structure approaches the theoretical limit of MOSFET technology based on a 1D channel completely surrounded by the gate. Image is courtesy of Leti Corp.

1.1.3 Challenges to Moore’s Law

The challenges to cramming more transistors into the same chip area come from a few different fronts that are somewhat related. Ignoring issues with interconnect density or problems with parasitic resistances and capacitances, MOSFET technology itself has some physical limits that make scaling difficult. One of the primary problem facing engineers today is power dissipation and operating chips under low load to avoid thermal runaway. [6] The ability to decrease the power density has become a problem of reducing the supply voltage, i.e. the threshold voltage of the FET. Lowering the threshold voltage for a fixed on/off current ratio is a matter of reducing the subthreshold swing ($S$). That problem requires improved electrostatic gate control, which has already neared its geometrical limit using FinFETs and GAA-FETs. The minimum subthreshold swing is thermally-limited to 60 mV/decade by Boltzmann
statistics and the independent nature of electrons in the channel. This limits the threshold voltage, and ultimately the power consumption of MOSFET technology. In addition, continued shrinking places extreme pressure on the ability to fabricate these devices, such that process yields decrease. This increase in the cost per chip will eventually cause the demise of CMOS, as the economics will no longer make sense.

Power Density

The real root of the current problem with MOSFET technology is that power density on-chip is not scaling with transistor size. The dynamic power consumed by a MOSFET during a complete on/off cycle is given by the following equation, easily derived by assuming the switch is driving a capacitive load from the next FET’s gate.

\[ P_{\text{dynamic}} = \alpha C V_{DD}^2 f_{CLK} \]  \hspace{1cm} (1.1)

The utilization factor (\(\alpha\)) is the amount of clock cycles in which the switch undergoes a switching operation, which is more or less constant for a given architecture.

Figure 1-5: On-chip power density over time for high-performance CPUs and mobile CPUs. Power increased quickly with clock frequency during the 1990s, eventually hitting a power wall. Ever since, power dissipation has been a large consideration, with the main goal being the reduction of the supply voltage. Figure is from [7].
Per Dennard scaling (Section 1.1.1), power should decrease by the square of the scaling, while power density remains constant. Instead, until the mid 2000s, power density steadily increased as frequency scaled with rapid pace (Figure 1-5). Eventually, fan cooling could not dissipate the heat and thermal runaway became a consideration. This limit is around 100 mW/cm², about the same as an incandescent light bulb. In addition, cellular phones and laptops were becoming more popular than desktop computing. Thus, new power considerations were required for these mobile applications. The new trend for improving computing was to add computing cores rather than faster operation. The ideal of parallel computing in some way stemmed from the limitations of fan cooling.

Subthreshold Current and Electrostatic Gating

With clock frequencies limited to 3-4 GHz, reducing power consumption requires decreasing the supply voltage. To decrease the supply voltage, the MOSFET threshold voltage needs to be decreased in order to operate with similar levels of performance, i.e. similar on/off current ratios. This requires improved electrostatic gate control, which for decades meant decreasing the gate oxide thickness. However, as mentioned in Section 1.1.2, recent CMOS process nodes have changed the profile of the channel, as simply decreasing the oxide thickness is no longer effective. The reason is mainly that the oxide would be so thin that large leakage current would dominate and static power dissipation would be prohibitively large. A large $I_{OFF}$ current also means reduced device performance in regards to distinguishing 0’s and 1’s. Thus, three-sided and nanowire channels are being used to alleviate this issue.

An underlying problem with MOSFET technology, regardless of channel shape or material, is subthreshold switching behavior. Rather than switching as an ideal step function from $I_{ON}$ to $I_{OFF}$ at the threshold voltage ($V_T$), the drain-source current ($I_{DS}$) instead climbs exponentially with gate-source voltage ($V_{gs}$) according to Boltzmann statistics. The inverse slope of this exponential characterizes the 'sharpness' of this transition. A sharper slope is better, as similar performance could occur at lower supply voltages. This quantity is defined as the subthreshold swing, or subthreshold
Figure 1-6: The subthreshold swing for MOSFET technology is thermally-limited by the independent nature of carriers in the channel to no less than 60 mV/decade. A sharper subthreshold swing is desired to reduce the threshold voltage, and power dissipation of devices.

slope, and is given by Equation 1.2. The value is usually given in millivolts per decade of current increase, as an on/off ratios of many orders of magnitude is desired.

\[
S = \frac{\partial V_{gs}}{\partial (\log I_{DS})} = \ln 10 \frac{kT}{q} \left[ 1 + \frac{C_d}{C_{ox}} \right] \geq 60 \text{ mV/dec}
\]  

From the equation, the subthreshold slope depends on the depletion capacitance and the oxide capacitance. In an ideal FET, the oxide capacitance dominates and the ratio of \(C_d/C_{ox}\) approaches zero. Even in the ideal case, the minimum subthreshold swing is still 60 mV/decade. This is a fundamental limit of MOSFET technology that cannot be altered. For a desired \(I_{ON}/I_{OFF}\) ratio of \(10^6\), the minimum threshold voltage is 0.36 V. Due to noise problems, the practical limit is closer to 0.5 V. [8]

The underlying reason for this limit is that electrons in a FET are independent of one another when looking at the probability of being injected into the channel. Each electron has to overcome an energy barrier from the source to the channel that is lowered by the gate voltage. If instead the behavior of electrons was correlated and grouped together, the subthreshold swing could be lowered. Examples would be that in ferromagnetic materials; once a threshold of electrons have a certain orbital spin, the remaining electrons will match that state, creating a strong magnetic moment.
Provided the need for lower power switching, the key is to find a way to improve the sharpness of the switching.

**Process Yields**

In addition to problems with device operation, MOSFET fabrication continues to get more difficult from a yield standpoint. Yield is defined as the ratio of good chips versus the total number of chips produced on a wafer. Considering the sub-nanometer dimensions required it is no surprise that hitting yield goals is a lofty challenge.

Lithography of the lateral features is one of the primary considerations. 193 nm light is currently used with water immersion to improve the optics. Future sub-10nm nodes plan to use some form of extreme ultraviolet lithography (EUV), proposed at 13.5 nm. This technology has many issues since almost all materials used in the optical path would absorb at this energy.

Another trend in CMOS is incorporating heterogeneous materials and complicated alloys. With each change, considerations such as lattice matching and edge passivation must be made. Extensive materials development is required with each process node at this point, adding to yield pressures.

### 1.2 Extended CMOS

An extension of traditional CMOS technology is likely the next step forward due to its relative ease in incorporating without major architectural changes. Two likely technologies exist. First, FETs with tunneling junctions from source to channel reduce the subthreshold slope below the 60 mV/decade limit, although the cost is vastly reduced operating current. Second, ferroelectric materials can be used in series with current gate oxides to produce FETs with enhanced gate capacitance. This also drives the subthreshold slope lower, but issues remain in meshing these materials with high yield.
1.2.1 Tunneling FETs

A normal n-type MOSFET would be made with n-doped source and drain in a lightly p-type background. Injection from the source into the channel is driven by gate-induced lowering of the n-p barrier. Thermionic emission of electrons increases as the barrier decreases. Instead, if the source is replaced with p-type doping and the region under the gate is intrinsic, the device can now operate as a gate-controlled PIN junction. If the p-doping is large, a tunneling barrier is created (see Figure 1-7a). The gate is now able to modulate the tunneling current by altering the junction width. In addition, the gate can change the number of overlapping states in the P and I regions. Since current modulation is not determined by thermal excitation above a barrier, but rather by band-to-band tunneling determined by the overlap of states, the subthreshold slope can be lower than 60 mV/dec (Figure 1-7b). These devices are known as tunneling FETs (TFETs).

One review of Lu and Seabaugh summarizes 14 reports of TFETs with subthreshold swings that are less than 60 mV/dec. [9] For the most part, silicon and other Group IV materials are used, but better performance exists with direct gap materials such as III-V alloys. While subthreshold swings are reported as low as 20 mV/dec, the cost is vastly reduced operating current. The largest current for a device operating...
less than 60 mV/dec is only around 1 nA/μm, about three orders less than low-power CMOS.

Improvements are needed to increase the performance of TFETs before they could replace MOSFETs, but some power-hungry paths that can use lower currents could make use of TFETs in the near future since most of the processing is similar for CMOS.

1.2.2 Negative Capacitance with Ferroelectrics

One way to improve the subthreshold swing is to design devices with altered electrostatics such that the applied gate voltage is ’multiplied’, acting on the channel with a larger effective capacitance. One proposal to implement this idea is to use ferroelectric materials in series with a normal dielectric gate oxide. These ferroelectric materials exhibit a negative differential capacitance that act similar to an inductor. [10] By replacing the gate oxide stack with a ferroelectric and dielectric, the effective result is a step-up transformer. The net capacitance is larger across the two materials than the oxide itself. Thus, a lower supply voltage would be required to generate a similar charge density, allowing the subthreshold swing to be less than 60 mV/dec, as sketched in Figure 1-8. [11]

One of the main advantages of negative capacitance FETs (NC-FETs) is that the technology is basically CMOS with the addition of a ferroelectric insulator. Recent reports have shown these materials can be grown via atomic vapor deposition (ALD) by doping HfO₂, a common high-K dielectric. [13] High $I_{ON}$ can be achieved relative to TFETs given the basic device is unchanged. Although structures appear similar to conventional CMOS, incorporating these materials into the CMOS process flow remains the biggest challenge.

1.3 Beyond CMOS

At some point, even the improvements from TFETs and NC-FETs will not be enough to reduce power consumption. Alternative technologies will be needed, and in many
Figure 1-8: Concept of Negative Capacitance FETs using ferroelectrics in the gate insulator stack, acting as a gate voltage amplifier. This allows the subthreshold swing to be less than 60 mV/dec. Image courtesy of [12].

In some cases this technology will not be based on charge as the carrier of information. Instead, other quantum states should be used. The idea is that the token of information will be utilized to create low threshold switching. Some technologies are presented in this section, and then the valley degree of freedom will be introduced as a novel token of information that is especially relevant when using monolayer transition metal dichalcogenides.

Figure 1-9: Taxonomy of Beyond CMOS technology as presented in the ITRS 2015 report, "Beyond CMOS". [14]
The International Technology Roadmap for Semiconductors (ITRS) has created a taxonomy for Beyond CMOS technology, starting with a choice of information token at the bottom, per Figure 1-9. This leads to a choice of materials with the chosen property, and then the design of a switching device. The two top levels will be mostly neglected in this thesis, but are equally important to consider. In this work, we assume binary digital information will continue to be used and the architecture will be similar to the CPUs designed today. With that said, new ways to design architecture is desperately needed to make efficient use of any new devices. This is no more clear than with quantum computing, where data is no longer strictly binary or determinate.

1.3.1 MEMS/NEMS

Micro or Nano Electro-Mechanical Switches (N/MEMS) offer an interesting alternative to transistors that still uses charge as the information token. The motivation is that a mechanical switch offers a near perfect switching behavior from high to low (i.e. $0 \text{ mV/dec}$ subthreshold swing). When the switch is disconnected no current flows, but once it is latched closed, current is determined by the load resistance (see Figure 1-10). This enables very low threshold voltages with zero static power dissipation. At the nanometer scale, the idea usually involves a cantilevered or suspended conductor that comes into physical contact with the drain contact when a gate voltage is applied. The gate voltage creates an ambipolar attractive force that latches the switch.

Although ideal in principal, a few lingering issues remain. First, at these scales, materials tend to have a hysteretic stiction force. This means the threshold voltage is not the same for low-to-high as high-to-low (see Figure 1-10b). Over time, wear can cause the device to become difficult to turn off and may eventually become stuck in the 'on' position. Second, switching times are slow compared to MOSFETs. Due to the mechanical action required, switching times in the nanosecond regime are observed, capping clock speeds to hundreds of megahertz. Combining the low power aspects of NEM switches for clock gating with the speed of CMOS for vital processing cores...
may be a way to improve overall power dissipation without major changes to CMOS fabrication.

Figure 1-10: Concept of a N/MEMS Switch. At the nanoscale, devices are normally created such that a conductor is suspended above another conductor. When gate bias exceeds a threshold, the conduction path is closed, with near-ideal switching. Image courtesy of [15]

1.3.2 Spin-based Logic

There are many proposals for logic based on carrier spin that is generically labeled spintronics, but the general idea is the same: electrons have an intrinsic spin degree of freedom that can be up/down lending to their usage as binary tokens of information. Due to the connection of spin with ferromagnets, all spintronic devices include magnetic material. Due to the collective threshold-type switching of ferromagnets, this technology offers hope of low threshold switching for low power dissipation.

Spin-based logic is roughly divided into two categories: 1) spin injection by ferromagnets into non-magnetic channels, and 2) spin injection into ferromagnetic channels. In both cases, the 'source' and 'drain' electrodes are ferromagnets, enabling spin injection. The output is a magnetoresistance. When the 'drain' or read structure is parallel with the spin, resistance is low; when antiparallel, resistance is high and current is lowered.

Initial proposals for spin-based switching included replacing the source and drain contacts in a MOSFET with ferromagnetic materials. Spin-polarized current is injected into the channel of a non-ferromagnetic channel. Depending on the particular design, the gate either modulates charge density (similar to a conventional FET) or can change spin via spin-orbit coupling. In both proposals, the drain’s magnetization
is fixed, acting as a spin analyzer. The result is either a large current (parallel) or small current (anti-parallel).

Using spin transfer torque (STT), magnetic domain walls (DWs) can be moved along a magnetic racetrack (nanowire). By injecting spin-polarized current more than a threshold current, determined by intrinsic and extrinsic properties of the device, the injected spin will provide a force that shifts the domain wall to the next meta-stable position. Using a magnetic tunnel junction to read out the underlying magnetization via magnetoresistance, nonvolatile binary states are created. Initial simulations and device measurements for this STT-DW logic have been reported by Currivan et al. in recent years, as displayed in Figure 1-11. [18,19]

While demonstrated to have low threshold switching with sharp transitions, the technology is currently limited by the poor on/off ratios due to low magnetoresistance materials. Whereas MOSFETs expect an on/off ratio of $10^6$, magnetoresistance read-out gives ratios on the order of less than 10. Also, drive currents are low, limiting fan-out and cascading. Improvements will require advancement in materials with larger magnetoresistance.

Figure 1-11: (a) Concept of STT-DW logic. A domain wall in a ferromagnetic nanowire can be moved by spin-transfer torque to either side of a magnetic tunnel junction (MTJ), acting as a spin analyzer. (b) Magnetoresistance versus applied voltage (current) for fabricated STT-DW devices, showing a sharp transition above threshold. Image courtesy of [18,19]
1.4 Valleytronics

In this thesis, I will discuss a new information token that has been proposed in the past, but has greater merit using the properties of 2D transition metal dichalcogenides (Discussed further in Chapter 2). The idea starts by looking at the E-k dispersion relation in crystalline materials. For electronics and optics, we often consider only the conduction band minimum and valence band maximum, but we know this is not the whole picture. These bands of allowed Bloch states are usually parabolic in shape, with maxima/minima at certain points in momentum-space (k-space). These 'valleys' in k-space act as low-energy sites for carriers to 'pool'. Due to crystal symmetry, there are often multiple, energetically-degenerate valleys in the dispersion relation that are related through some sort of symmetry. If we label two degenerate valleys as -K and +K, we expect them both to be populated equally when excited electronically or optically (Figure 1-12a).

However, if there is a property that distinguishes these valleys, we can treat them as binary holders of information, i.e. a new token of information. The key is to find a distinguishable property that can be modulated externally, such as with electromagnetic fields or electro-mechanically. Once a given valley can be selectively written to and read from, we have the basis for a new form of information based around this valley degree of freedom in the dispersion relation (Figure 1-12b).

The initial goal is then obvious: finding a material with distinguishable valleys in k-space that are energetically degenerate but linked to properties that can be manipulated via external stimuli. This property should be able to be read as well as written, with high selectivity. Finally, the polarization of this state should be maintained for a long enough time that operations can be performed. This thesis demonstrates initial work on the validity of the valley degree of freedom as an information carrier in monolayer transition metal dichalcogenides, specifically tungsten disulfide (WS₂). The scope of this thesis is mainly the first two steps in the Beyond CMOS taxonomy (refer to Figure 1-9), with discussion pertaining to future valleytronic applications (Chapter 6).
Figure 1-12: Concept of logic using the relative occupation of excited carriers in unique k-space valleys in a crystal’s dispersion relation (E-k diagram). (a) Generally, k-space valleys are indistinguishable and excitation equally populates valleys that are energetically degenerate. (b) In some materials, these valleys can be distinguished by an intrinsic property. If the valley index is externally accessible, we can read/write to select valleys, providing binary information storage.

1.4.1 Early Work on Valleytronics

Manipulation of valleys in solid-state crystals should not be foreign to those studying semiconductor technology. Strained silicon and germanium have been used for at least a decade in MOSFETs to increase carrier mobility by altering the valley curvature. In fact, early work pertaining to the valley degree of freedom dates back to the 1970s, where valley splitting and degeneracy were studied in silicon inversion layers (2D electron gas). [20–22].

After much work on the related field of spintronics in the 2000s, the field of valleytronics and the ability to control the valley index gained attention. Confined geometries have primarily been used to enhance the ability to observe these many body effects. For instance, valley control using strain and magnetic field in aluminum
arsenide (AlAs) quantum wells was demonstrated by work at Princeton. [23–25] Other materials with familiar crystal structure, such as silicon and diamond, have also been reported with small valley polarization using applied electrical and magnetic fields to break crystal symmetry. [26–29] Finally, valley degeneracy breaking was observed in crystalline bismuth at low temperatures in an applied magnetic field. [30]

However, all previous reports had a common problem that needs to be alleviated for efficient valleytronic devices. The ability to control valley populations is limited in these materials as the valley index is not linked to an intrinsic property that can be externally controlled. Valley polarization via valley splitting (i.e. changing the energy of different valleys) is the only control for accessing certain valleys and that lever is a weak one in these materials. By contrast, carrier spin can be manipulated easily via magnetic fields (magnetic moment), electric fields (spin-orbit coupling), and optical excitation (spin selection rules). A preferred material platform for valleytronics would have valleys linked to easily accessible properties (such as spin) and these valleys would be robust.

1.4.2 The Valley Degree of Freedom in 2D Hexagonal Crystals

Expanded research in valleytronics owes itself to the discovery of 2D materials, such as graphene and transition metal dichalcogenides. Having a hexagonal crystal lattice, these materials exhibit a hexagonal Brillouin zone with both conduction band and valence band valleys in their $E-k$ dispersion at the six corners, labeled K (see Chapter 2 for more information). Due to three-fold symmetry of the crystal, every other K valley is indistinguishable. But nearest neighbor K valleys are unique and time-reversed of each other, which can be labeled K$^+$ and K$^-$ to clarify their uniqueness. However, these valleys are all energetically degenerate.

In graphene, the bands meet at K, forming a massless linear dispersion with zero band gap. Due to the crystal's inversion symmetry, any property in K$^-$ can be expected to be the same in K$^-$, making these valleys degenerate. By breaking inversion symmetry, through external stimuli, this valley degeneracy can be broken as well. Inversion breaking is also shown to cause graphene to have an intrinsic
magnetic moment and opposite Berry curvature between valleys. [31] Details are explained further in Chapter 2, but this leads to a transverse valley current (i.e. Hall effect), whereby carriers in opposite valleys move in opposite direction. Based on this observation, active valley devices (valves/filters) have been theoretically proposed in graphene, the first step toward true valleytronic systems. [32,33]

While schemes can give graphene a small, infrared bandgap, optically probing the valleys in graphene is nearly impossible. Thankfully, a new group of semiconductor 2D materials exist with direct band gaps in the visible regime. These are the transition metal dichalgonides (TMDs). Similar to graphene, monolayers have a hexagonal lattice and their band gap exists at the K points in the Brillouin zone. Interestingly, their hexagonal lattice intrinsically lacks inversion symmetry, and thus has nondegenerate valleys. Combined with a large spin-orbit coupling that is also present, these unique valleys are linked to spin and circularly polarized light. [34] Obviously, TMDs provide an intriguing platform for valleytronics that needs to be explored more carefully, which will be the task of this dissertation.

1.5 Summary and Outlook

Computational power has roughly doubled every two years in no small part thanks to the constant improvements in CMOS technology and processing. However, after 50 years of Moore’s Law, the technology is starting to bump up against the fundamental thermal limits of MOSFET devices due to the independent nature of electrons in the channel. Improved charge-based technologies may provide lower subthreshold swings and power, but technologies that use different tokens of information must also be looked. As per Figure 1-13, the key is finding technologies that operate with minimal energy consumption and also maximum speed. At this point, no Beyond CMOS technology is ready for mass commercialization (see Figure 1-13). To the mix of N/MEMS and spin-based logic, I also propose the valley degree of freedom as a valid token of information. Using monolayer transition metal dichalcogenides (TMDs) with broken inversion symmetry, the goal of this thesis is to explore valley polarization and
depolarization mechanisms. In the end, the goal is to add to the current work related to valleytronics and provide ideas for future exploration based on this work.

Figure 1-13: Summary of CMOS and next-gen logic technologies, showing their current status. More research and work is required to get Beyond CMOS technologies to the point where they can replace CMOS. Figure from the 2015 ITRS Report on Beyond CMOS technologies. [14]
Chapter 2  

Valley Physics in 2D Transition Metal Dichalcogenides

Transition metal dichalcogenides (TMDs) are layered crystals with the chemical formula MX₂, where M is a transition metal and X is a chalcogen other than oxygen. The periodic table in Figure 2-1 highlights the three chalcogens, sulfur, selenium, and tellurium, as well the known transition metals that have a layered phase. Three common phases exist which lead to electronic properties from metallic to insulating with a band gap.

![Periodic Table of the Elements](image)

Figure 2-1: Periodic table of elements, highlighting the transition metals (magenta) and chalcogens (blue) that form the family of layered TMDs.

Of the many TMDs, the four containing the combination of molybdenum/tungsten
and sulfur/selenium will be the focus of this work: MoS$_2$, MoSe$_2$, WS$_2$, and WSe$_2$. Specifically, the hexagonal 2H phase of these four compounds is of interest to the field of optoelectronics for their semiconducting property with a bandgap in the visible range. At the monolayer thickness limit, these TMDs transition from an indirect band gap to a direct one, improving light-matter interactions as observed by enhanced absorption and luminescence.

This chapter will describe the main properties of 2D TMDs that will be of interest to this work, starting with the physical crystal structure and Brillouin representation in k-space. This leads directly to useful electronic and optical properties. Due to the confined nature and reduced dielectric environment for excited electron-hole pairs, well-bound excitons dominate the optoelectronic properties and have implications for device applications. Finally, these crystals lack inversion symmetry in the monolayer limit and have large spin-orbit interactions due to the heavy metal orbitals, leading to a unique valley degree of freedom in k-space that acts similar to carrier spin. Valleys become linked to the helicity of light that is absorbed or emitted, and can energetically shift in the presence of a magnetic field. The relative population of these valleys is the basis of this thesis, and this chapter provides the foundation for understanding the later discussion regarding valley polarization and depolarization mechanisms in Chapter 5.

2.1 Crystal Structure and Brillouin Zone

Transition metal dichalcogenide crystals have a few different phases, leading to differing properties. There are three main phases: 2H, 1T, and 1T'. A fourth phase sometimes occurs and is referred to as T$_d$ because it is a distorted 1T structure. Figure 2-2 shows the crystal structures for the 2H and 1T phases, from the top and side view. 1T' and T$_d$ are not displayed as they do not occur for the Group VI sulfides and selenides. From the figure, the 2H phase is hexagonal from above, similar to graphene. The 3D unit cell is a trigonal prism. The 1T phase has one layer of the chalcogen twisted to fill the gaps in the hexagonal lattice from above. The unit cell
becomes octahedral.

![Diagram of 2H and 1T crystal phases](image)

Figure 2-2: The two common crystal phases for monolayer TMDs are 2H and 1T, as labeled, courtesy of [35]. The 2H phase has a hexagonal structure from above and a unit cell that is trigonal prismatic, whereas the 1T unit cell is octahedral.

In general, 1T phase TMDs are metallic or semimetallic, whereas 2H phase TMDs are often semiconducting. Thus, our focus is on the 2H phase monolayers, and specifically the TMDs combining W/Mo with S/Se, which possess bandgaps in the visible and NIR. By default, the remainder of this dissertation will use TMDs to refer to these four semiconductors with similar properties, unless explicitly stated.

Looking more closely at the 2H crystal in Figure 2-2, we see that although hexagonal from above like graphene, the lattice's two-atom basis consists of the metal and chalcogen atoms, whereas both basis atoms are carbon in graphene. From the side, each layer consists of the transition metal atoms sandwiched between layers of the chalcogen. Every transition metal is covalently bonded to six chalcogen atoms, whereas each chalcogen has three nearest metal atoms. Layers are bonded through weak van der Waals forces, leading both to good lubrication properties (graphite and MoS$_2$ are both common lubricants) and the ability to isolate layers by various means of exfoliation. A good review of the structure of these monolayer TMDs and their properties can be found in Ref [36].

The reduced Brillouin zone (BZ) for the hexagonal 2H lattice is a hexagon prism (Figure 2-3a), with the $\Gamma$ point in the center. Rather than treating the whole BZ, a hexagonal projection is taken to capture the interesting physics (Figure 2-3b). At the
six corners are the K points, six energetically degenerate locations where the bandgap exists for monolayer TMDs (the same location as the Dirac cones in graphene). The reciprocal lattice vectors are shown from the center of the reduced BZ. By using linear combinations of the two reciprocal vectors, you can see the three-fold rotation symmetry of the BZ.

A consequence of the symmetry is that every other K point in the Brillouin zone is identical. Adjacent K points, on the other hand, are inequivalent and are often labeled K+ (K) and K- (K'). In particular, any K+ point is equivalent by a shift of the reciprocal vectors. Oppositely, going from K- to K- is not possible with the reciprocal lattice vectors, showing that these two points are inequivalent. These two valleys are energetically degenerate and exhibit time-reversal symmetry though. As will be shown in later sections, time reversal shows odd symmetry with spin and other spin-like properties.

Figure 2-3: The reduced Brillouin zone (BZ) is also hexagonal, with the monolayer band gap appearing at the six energetically-degenerate, yet inequivalent, K points due to the monolayer’s crystal symmetry. Figure (a) is adapted from [37].

2.2 Electronic Band Structure - Energy Dispersion

One of the most interesting properties of the Group VI TMDs is their electronic band structure. Similar to graphene, the conduction band minimum and valence band maximum occur at the K points of the reduced Brillouin Zone (Figure 2-4). However, due to interactions between the orbitals, a finite band gap exists. In bulk
MX₂ TMDs (M = Mo, W and X = S, Se) this gap is not the lowest band gap as the conduction band minimum is part-way between Γ and K points, denoted Δ (see Figure 2-4) and the valence band maximum occurs at Γ. Interestingly, when thinned to a single layer, the band gap transitions from indirect to direct, occurring at the energetically degenerate K points. Primarily, this means monolayer TMDs are excellent optical materials compared to their bilayer (and thicker) cousins. Per unit thickness, monolayers can absorb as much as 15% of incident light and emission can be bright enough to see by eye when material quality is improved. [38, 39]

Figure 2-4: Calculated electronic band structure for WS₂ with spin-orbit coupling included, showing a spin-splitting in the valence band of 0.43 eV. Figure is adapted from [40].

The direct band gap also suggests optoelectronic devices that emit light, such as LEDs and lasers. While this thesis does not focus on such electrically-pumped photon emission, many others have shown that monolayer TMDs can produce reasonably good light emitters. [41, 42] Similar to traditional light-emitting semiconductors, efficient devices require a junction that is either generated by electrostatic doping [41, 43–45], chemical doping [46, 47], heterostructures [48–50], or tunneling barriers [42, 51]. One limitation of light-emitting devices is the small photoluminescence quantum yields often reported. [52–54] Chemical treatments have shown promise in improving quantum yields [53, 55, 56], but these treatments come with their own negative consequences, including prohibitive device fabrication steps.

The other consequence of this direct band gap is that excited electron-hole pairs exist primarily in these valleys. As will be the topic of this dissertation, this valley degree of freedom is somewhat unique to monolayer TMDs and makes for novel
technology proposals and future device applications. Section 2.4 below includes a theoretical introduction to this valley pseudospin degree of freedom and the implications for electro-optics. Combining the optical and valley-spin properties of these materials is central to the work presented throughout and many of the measurements require this direct bandgap at the K points.

### 2.2.1 Other Interesting TMDs

While this thesis is concerned primarily with the aforementioned semiconducting TMDs, it was previously mentioned that other layered TMDs exist using the other transition metals in groups IV to VII. A brief summary of their electrical properties is cataloged in Figure 2-5.

![Periodic table of layered TMDs from group IV to group VII of the transition metals. Direct gap semiconductors from group VI will be the focus of this work, but it is interesting to note the properties of other TMDs, such as WTe$_2$, a Weyl semimetal, and TaS$_2$, a superconductor at low temperatures.](image)

Of the metals and semimetals, the reduced dimensionality again presents interesting properties of interest to physicists and engineers alike. For example, WTe$_2$ has been shown to have incredibly large magnetoresistance that increases with decreasing layer number.[2,21] At the monolayer limit, WTe$_2$ is believed to be a 2D topological insulator.[22] It has also been shown to support Weyl fermions, and is
considered a type-II Weyl semimetal, of interest to condensed matter physics with potential applications in low-energy electronics and quantum computing.[3]

Finally, many of the metals, including TaS₂ and NbSe₂ are known superconductors, with $T_C$ values that change as the number of layers changes.[4,23,24] Monolayer superconductors have not been synthesized to date, but could provide a platform for true 2D superconductivity without complicated molecular beam epitaxial (MBE) growth. Quantum computing and information via superconducting qubits is again one of the heralded applications of such material.

2.3 Excitons and Trions in 2D TMDs

Due to charge confinement in the 2D plane and reduced dielectric screening, Coulomb interactions in 2D TMDs are extremely strong. Excited electron-hole pairs are tightly bound with a binding energy greater than 100 meV, much larger than $k_B T$ at room temperature. [57–60] These hydrogen-like bound electron-hole pairs are referred to as excitons, a quasi-particle representing an excited state. Excitons in 2D TMD semiconductors are reported to have a Bohr radius of approximately 1 nm, larger than that of molecular excitons of Frenkel-Poole type, but smaller than delocalized Wannier-Mott excitons common in bulk semiconductors. [60–62] As shown in Figure 2-6, the excitonic gap is less than the extended band-to-band gap by the amount of the exciton binding energy.

Excitons dominate the optical properties of monolayer TMDs, providing a link between photons and free charge carriers. Figure 2-7 shows the absorption spectra for various layer thicknesses of WS₂. The dominant feature is a nearly Lorentzian peak that is dissimilar to the band edge absorption prevalent in bulk semiconductors. These peaks are exciton states. From the plot, the A and B exciton peaks are clearly visible (as labeled), which correspond to the two spin-orbit split valence bands. Given the hydrogenic model, excitons can have higher 'orbital' energies, often labeled $A'$ and $B'$ respectively (not shown in the figure). Beyond that, higher energy states blend into the continuum of the band-to-band states, with some prevalent band nesting leading
Figure 2-6: Cartoon comparing excitons in 3D semiconductors versus 2D (side view). Due to the reduced dimensions in 2D, carriers interact more strongly, creating a tightly-bound excited state, called an exciton. The excitonic gap is less than the band-to-band gap by the amount of the binding energy. Image courtesy of [60].

to the labeled C peak. At different temperatures and layer number, the energy of these peaks shift, as expected by solid-state band theory.

Figure 2-7: Typical absorption spectra for WS₂ for different number of layers. Also, absorption for 1L-WS₂ at 10 K and 300 K. The absorption shows the obvious signs of exciton states, labeled A and B. The C peak comes from band nesting. Figure courtesy of [57].

In addition to neutral exciton with one electron and one hole, charged excitons are prevalent in 2D TMDs due to the large Coulomb interactions. Upon doping with
either electrons or holes, a stable state can be created that involves an additional free charge bound to the neutral exciton. This three-body ionic state is referred to as a trion (triple particle ion). The physics of trions in 2D TMDs is still poorly understood. Besides the oft-reported redshift in photon emission with doping by various means (Figure 2-8), the mechanism for trion formation and recombination is unclear. [61,63,64] Although a trion is treated as a discrete particle with a reported binding energy on the order of 20 meV, no reports on trion transport in TMDs have been reported to substantiate this claim. [61,65,66]

![PL vs. Gate Voltage](image)

Figure 2-8: Photoluminescence versus gate voltage for monolayer WS₂ in a field-effect structure, providing electrostatic doping. At large gate voltages, negative doping is induced, showing a redshift in energy and decrease in intensity consistent with trion emission. Negative bias lowers the Fermi level, eventually showing neutral exciton emission. The zero gate bias emission indicates the monolayer is naturally n-type prior to doping.

Provided the optical properties of 2D TMDs are dominated by excitons, any optoelectronic or photonic device relying on TMDs will depend on exciton physics. Of the various properties, knowing the time-dependent and space-dependent dynamics of excitons in monolayer TMDs is crucial to the design of practical devices.

First, excitons are generated by energy, which excites an electron into the conduction band, leaving behind a hole in the valence band. These particles have a Coulomb interaction that is enhanced by the reduced geometry in 2D sheets, as well as the reduced dielectric environment since field lines can extend into space above or
Figure 2-9: Exciton dynamics for monolayer TMDs at various pump fluences, both before and after chemical treatment with TFSI superacid. The plots show superacid treatment removes much of the nonradiative recombination, but nonlinear dynamics still exist even at very low fluence levels. Figure courtesy of [54].

below the plane of the flake (Figure 2-6). When these excitons recombine, they release this energy as either radiatively as a photon or nonradiatively. For a given material, excitons have a characteristic radiative lifetime ($\tau_r = 1/k_r$) governed by the coupling of their excited and ground states (i.e. Fermi’s Golden rule). In addition, faster nonradiative paths ($k_{nr}$) rob excitons from recombining radiatively, usually by scattering with a phonon, defects, or an ion (e.g. dopants, free charge, etc.). The total exciton lifetime measured is the sum of the two rates ($\tau_X = 1/[k_r + k_{nr}]$). When nonlinear mechanisms that are three-body or higher are involved, the measured exciton dynamics are no longer mono-exponential. Common sources of nonlinear recombination include Auger (exciton-charge) and exciton-exciton interactions. [67–70] Biexciton recombination can be radiative, but is usually nonradiative. [71] Figure 2-9 shows how the pump fluence (i.e. exciton density) and chemical doping from superacid treatment change the exciton dynamics in monolayer WS$_2$ by reducing nonradiative
recombination. [54]

2.4 Valley Physics

Other interesting properties also arise at these reduced dimensions. For odd numbers of layers, specifically the monolayer, the crystal lattice is non-centrosymmetric and lacks inversion symmetry. Combined with the large spin-orbit coupling in group VI TMDs, thanks to the large transition metal elements, the K valleys in momentum space couple to carrier spin to create interesting valley physics. This pseudo-spin degree of freedom is linked to orbital angular momentum, and can be accessed via circularly polarized light following optical selection rules. [72–74] In this manner, pumping with circularly polarized light can selectively excite carriers in certain valleys in k-space. Likewise, emission from these valleys maintains the selection rules, and luminescence is circularly polarized, oppositely for differing valleys. What follows in this section is an introduction to the theory that is needed to establish this valley degree of freedom, and some implications, such as the valley Hall effect and valley Zeeman effect.

2.4.1 Valley-Spin Coupling

The missing property of previous valleytronic materials (Section 1.4.1) was that these valleys were not linked to an intrinsic property that could be controlled. With TMDs, that is no longer true and the key is the crystal’s non-centrosymmetry, i.e. it does not have spatial inversion symmetry. The theory for how this links the valley degree of freedom to carrier spin is summarized here.

First, let’s start by looking at the real space honeycomb lattice of the 2H-phase TMDs of interest in Figure 2-10. Unlike graphene, the A/B basis sites are occupied by different atoms, specifically a transition metal and a chalcogen (minus oxygen). The primitive cell (purple dashes) indicates this spatial symmetry breaking. In addition, the figure shows the two primitive lattice vectors that define the crystal. The simplified model ignores the three-dimensionality of TMDs compared to graphene,
but the experimental results can be explained conceptually using this model.

\[ a_1 = \frac{3a}{2} (\hat{x} + \frac{1}{\sqrt{3}} \hat{y}) \]
\[ a_2 = \frac{3a}{2} (\hat{x} - \frac{1}{\sqrt{3}} \hat{y}) \]

Figure 2-10: Hexagonal (honeycomb) lattice and primitive lattice vectors for 2H-phase TMDs, with different basis atoms at the A/B lattice sites. The primitive cell (purple dashes) shows the lack of inversion symmetry.

A consequence of the two atom basis and broken spatial symmetry is that a direct band gap opens at the K points, different than the Dirac point in graphene. Using appropriate basis wavefunctions for the conduction and valence bands based on the crystal symmetry, a two-band Hamiltonian can be written using the Pauli matrices \( \hat{\sigma} \). [34]

\[ \hat{H}_0 = at(\tau k_x \hat{\sigma}_x + k_y \hat{\sigma}_y) + \frac{\Delta}{2} \hat{\sigma}_z - \lambda \tau \frac{\hat{\sigma}_x - \hat{\sigma}_y}{2} - \delta_z \]

(2.1)

In the Hamiltonian, \( \tau = \pm 1 \) is the valley index, \( a \) is the lattice constant, \( t \) is the hopping integral, and \( \Delta \) is the energy gap. The basis wave functions chosen for the conduction and valence bands at K are related by time-reversal symmetry, consisting of hybridized d-orbitals from the transition metal with magnetic quantum numbers of \( m = 0 \) and \( m = 2\tau \), respectively.

\[ |\phi_c \rangle = |d_{z^2}\rangle, \quad |\phi_v^\tau \rangle = \frac{1}{\sqrt{2}}(|d_{x^2-y^2}\rangle + i\tau |d_{xy}\rangle) \]

(2.2)

The last term of the Hamiltonian is due to spin-orbit coupling (SOC), where \( 2\lambda \) is the spin splitting in the valence band. Unlike graphene with weak SOC, TMDs exhibit large SOC due to the heavy metal atoms. [75] While the chosen conduction
band basis is spin degenerate at the K points, the valence band-edge state will split. By time reversal, K+ and K- will have opposite spin splitting. To emphasize, this spin splitting asymmetry is a result of the broken crystal inversion symmetry, not the choice of model.

![Figure 2-11: Coupled spin and valley physics in monolayer TMDs. Large spin-orbit splitting and inversion symmetry breaking force the valley index to be linked to carrier spin for band-edge carriers.](image)

Due to the combined result of time reversal and the broken spatial symmetry, large spin orbit splitting, on the order of 100-500 meV in TMDs \cite{75}, is opposite in K+ and K- valleys. This links the valley degree of freedom to carrier spin for band-edge carriers, specifically holes (Figure 2-11). Using more extensive electronic modeling, the conduction band also exhibits a small spin splitting, but only on the order of $k_B T$ at room temperature. \cite{76} The valence band spin splitting, combined with the large momentum separation between valleys, should suppress scattering between valleys as spin flips would be required in addition to atomic-scale scatters. As demonstrated in Chapter 5, this spin-valley coupling allows large valley polarization values and is key to most valleytronic proposals (Chapter 6).
2.4.2 Berry Curvature and the Valley Hall Effect

The Berry phase is the geometric quantum phase from a round-trip in a quantum mechanical system. Conceptually, if a tangential vector takes one round trip on a closed surface of a sphere, the vector changes direction (phase) even though the starting and ending points are the same (Figure 2-12).

![Figure 2-12: Berry phase in a closed path on the surface of a sphere.](image)

This geometric phase is generally useful in many areas of physics and mathematics, but can be applied to Bloch electrons moving adiabatically in an energy band to explain anomalous Hall effects.

The associated Berry curvature of the Berry phase is defined by,

\[ Q_n(k) = \hat{z} \cdot \nabla_k \times \langle u_n(k) | r \nabla_k | u_n(k) \rangle \]  \hspace{1cm} (2.3)

where \( |u_n(k)\rangle \) is the periodic part of the Bloch wavefunction and \( n \) is the band index.

Berry curvature can also be written in terms of interband elements of the momentum operator \( \hat{\mathbf{p}} \) as,

\[ \Omega_n(k) = \frac{i \hbar}{m} \sum_{i \neq n} \frac{P_{n,i}(k) \times P_{i,n}(k)}{[E_n^0(k) - E_i^0(k)]^2} \]  \hspace{1cm} (2.4)

where \( P_{n,i}(k) \equiv \langle u_{n,k} | \hat{\mathbf{p}} | u_{i,k} \rangle \)
For the Hamiltonian in Equation 2.1, the Berry curvature in the conduction band is reduced to. [31, 34]

\[ \Omega_c(k) = -\frac{2a^2t^2\Delta'}{[\Delta^2 + 4a^2t^2k^2]^{3/2}} \] (2.5)

Within a single valley, the Berry curvature is dependent on the spin-orbit-dependent band gap, defined as \( \Delta \equiv \Delta - \tau s_z \lambda \). From the dependence on \( \tau \), Berry curvature is equal and opposite in opposite valleys. As well, Berry curvature has the opposite sign for the valence band compared to the conduction band in the same valley. These key symmetries, shown in Figure 2-13, eventually lead to the valley Hall effect.

Figure 2-13: Berry curvature calculated for monolayer MoS\(_2\) along the K–\(\Gamma\)-K path in the Brillouin Zone. The red curve is the bottom of the conduction band, with blue being the top of the valence band. Figure courtesy of [73].

Equations of motion can now be written for the Bloch electrons in each band, starting with the familiar Lorentz force. We know from classical electromagnetics, that the magnetic force is transverse to the direction of motion and also the applied magnetic field, leading to the ordinary Hall effect.

\[ \hbar \dot{\mathbf{k}} = -e\mathbf{E} - e\dot{\mathbf{r}} \times \mathbf{B} \] (2.6)

A symmetric and coupled equation for Bloch electrons in k-space can be written, as in Equation 2.7. The velocity of the wavefunction is the group velocity and a transverse term related to the Berry curvature, \( \Omega \).
\[ \mathbf{r} = \frac{1}{\hbar} \frac{\partial E_n(k)}{\partial k} - \mathbf{k} \times \Omega_n(k) \] 

(2.7)

The Berry curvature acts as an effective magnetic field in k-space and creates a transverse velocity component to the electron. Thus, instead of traveling in the same direction as the group velocity when an electric field is applied, the wavefunction now travels a path that curves. This leads to a valley Hall effect, whereby carriers in opposite valleys move to opposite edges in an in-plane electric field due to their opposite Berry curvature (Equation 2.5). Since the same carriers from each valley would travel in opposite directions equally, we would expect no charge current to flow. Due to the established connection to spin, we would expect there to be an ability to detect this valley hall conductivity using a magnetic signal.

Using the coupling of valleys to circularly polarized light (see the following section), the valley Hall effect has been detected in MoS$_2$ monolayers as a transverse voltage using optical excitation. [78] As will be shown, circularly polarized light excites electron-hole pairs in a single valley. Due to their opposite Berry curvature in the valence and conduction bands, holes and electrons move in opposite transverse directions when an electric field is applied. The result is a transverse Hall voltage that can be measured, with opposite slopes for the opposite helicity of excitation. (Figure 2-14)

![Figure 2-14: (a) Valley Hall transverse voltage detected for excitation at different polarizations, showing circularly polarized light shows a Hall voltage. (b) Concept of the valley Hall effect for optically-generated excitons in a single valley from circularly polarized light. Figures adapted from [78]](image-url)
2.4.3 Valley Connection to Circularly Polarized Light

In addition to Berry curvature, an orbital magnetic moment also exists due to the broken inversion symmetry in monolayer TMDs. Using the similar interband momentum matrix elements as Equation 2.4, the orbital magnetic moment is, \[ M(k) = \frac{e\hbar}{2m^2} \sum_{i \neq n} \frac{P_{n,i}(k) \times P_{i,n}(k)}{E_{n}(k) - E_{i}(k)} \] \[ (2.8) \]

The orbital magnetic moment \( m \) arises from the self-rotation of the electron wavepacket and is responsible for the \( g \) factor in semiconductor physics, i.e. the Kramer's pair. Again using the massive Dirac fermion model from Equation 2.1, the magnetic moment can be derived near the valleys based on material parameters. \[ m(k) = -2e2a^2t^2 \Delta \Delta e \] \[ (2.9) \]

Unlike Berry curvature, \( m \) has the same value in the conduction and valence band. However, the value is still opposite in opposing valleys, given the \( \tau \) dependence. Thus, the valley index acts as a sort of spin with intrinsic magnetic moment, and is often called a pseudospin with many of the same properties as spin. For instance, the \( m \) allows the energy of valleys to be changed in an out-of-plane magnetic field. To avoid confusion with the atomic orbital magnetic moment, I will refer to \( m \) as the valley orbital magnetic moment, or the valley moment.

\[ E_n(k) = E_{n}^0(k) - m_n(k) \cdot B \] \[ (2.10) \]

More details of this valley Zeeman effect will follow in the next section, with applications discussed in Chapter 6.

One of the more interesting features of the magnetic moment is the coupling to circularly polarized light. The coupling strength with optical fields with circular polarization are found using the transition matrix element of circular polarization \[ 34,73]: \]
\[ P_\pm^c(k) \equiv P_x^c(k) \pm iP_y^c(k) \]  
(2.11)

Where the conduction-valence interband matrix elements are given by \( P_\alpha(k) \equiv m_0(u_\alpha(k) | \hbar^{-1} \frac{\partial H}{\partial k_\alpha} | u_\alpha(k)) \)

For transitions near the K points, the massive Dirac Hamiltonian results in the following optical strength,

\[ |P_\pm(k)|^2 = \frac{m_0^2 a^2 l^2}{\hbar^2} \left( 1 \pm \frac{\Delta}{\sqrt{\Delta^2 + 4a^2 l^2 k^2}} \right)^2 \]  
(2.12)

Since the band gap is large (i.e. \( \Delta \gg atk \)), the transition strengths simplify and become:

\[ |P_\pm(k)|^2 \propto \left( 1 \pm \frac{\Delta}{\sqrt{\Delta^2 + 4a^2 l^2 k^2}} \right)^2 \]  
(2.13)

The conclusion is that light with positive helicity \((\sigma^+, \text{RHCP})\) only couples to \( \tau = +1 \), whereas light of negative helicity \((\sigma^-, \text{LHCP})\) only couples to \( \tau = -1 \). The light only couples to the orbital part of the wavefunction and does not perturb spin, causing the valley-dependence to also become spin-dependent.

Figure 2-15: Calculating the interband transition strengths for circularly polarized light shows robust valley-dependent selection rules, such that circularly polarized light can be used to selectively populate a chosen valley.
The inversion symmetry breaking in monolayer TMDs results in valley-dependent optical selection rules with circularly polarized light (Figure 2-15). The result is two-fold: 1) Circularly polarized light can be used to populate select valleys, and 2) When excitons recombine in a given valley, circularly polarized light of a given helicity is emitted. These two properties have tremendous application towards valleytronics, especially connected with carrier spin, as will be discussed in Chapters 5 and 6.

2.4.4 Valley Zeeman Effect

As was shown in the previous section, the valley index is linked to a magnetic moment \( \mathbf{m} \) that we can call the valley magnetic moment. This allows the energy of bands to be shifted in the presence of an out-of-plane (\( \hat{z} \)) magnetic field (parallel to the magnetic moment) per Equation 2.10. Energy shifts will be equal and opposite in the K+ and K- valleys, but due to the assumed symmetry between electrons and holes in the Hamiltonian, the net shift between the valence and conduction bands will be zero between valleys. If the actual asymmetry between carriers in the two bands was approximated as differences in the effective masses between electrons and holes, the net energy difference between valleys would be roughly,

\[
\Delta_v^{K^+} - \Delta_v^{K^-} \approx 2 \left[ \frac{m_0}{m_e^*} - \frac{m_0}{m_h^*} \right]
\]

(2.14)

While the valley magnetic moment is equal in magnitude to first approximation, the atomic orbital magnetic moment for the conduction band and valence bands are not equivalent. From the prior discussion (Section 2.4.1), the conduction band of monolayer TMDs consists mostly of \( d_{z^2} \) orbitals from the heavy metal atom, having a magnetic quantum number \( m = 0 \). The valence band consists of hybridized \( d_{x^2-y^2} \) and \( d_{xy} \) orbitals with magnetic number \( 2\tau \). Application of an out-of-plane magnetic field will have zero effect on the conduction band, while shifting the valence band edges of each valley by \( \Delta_a = 2\tau \mu_B B \), where \( \mu_B \equiv e\hbar/2m_0 \) is the Bohr magneton. The net shift in exciton energy between the two valleys would thus be \( \sim 4\mu_B B \), which is linear with \( B \).
To be complete, there is a third contribution from the intrinsic spin magnetic moment, but similar to the valley moment, the net change between valleys should be zero.

\[ \Delta_{\alpha}^{K^+} - \Delta_{\alpha}^{K^-} = 4\mu_B B \]  

(2.15)

Figure 2-16: (a) There are three main contributions to the valley Zeeman effect, but only the atomic orbital magnetic moment results in a energy splitting between valleys using our 2-band model. (b) Valley Zeeman splitting, showing a splitting of PL emission energy with output light of opposite helicity. [80, 81]

A summary of these three components that couple to magnetic fields is provided in Figure 2-16a from [80]. In back-to-back reports, two groups reported the predicted valley Zeeman effect. [80, 81] While their results vary somewhat, they confirm an energy splitting between the optical transitions of opposite valleys that is linear with the applied B-field. The results show this splitting is weak, only 1-2 meV Tesla\(^{-1}\), which would require large fields for normal applications. Last, both reports showed that the splitting is significantly different from \(4\mu_B B\), indicating a significant deviation from the simple two-band model with electron-hole symmetry that has been assumed...
in this section. The simple model has its limits, but the insight it provides is important for the majority of experiments and applications that will be discussed (Chapter 5 and 6).
Chapter 3

Growth of High-quality Monolayer Tungsten Disulfide (WS$_2$) via Chemical Vapor Transport

TMD monolayers are typically produced via top-down micromechanical exfoliation from a bulk crystal, as was first explored in graphene. While giving the best material quality, typical flakes are small and irregular, making device fabrication tedious. For large-scale device fabrication and improved research throughput, a bottom-up synthesis technique is desired. To date, the most reported methods are chemical vapor deposition (CVD) and physical vapor deposition (PVD).

A common CVD approach involves the chemical reaction of a metal oxide (e.g. WO$_3$) with a chalcogen vapor (e.g. sulfur) in a tube furnace. Although successful, CVD becomes challenging for TMDs that are less stable due to rapid oxidation (e.g. NbSe$_2$). For those TMDs, it would be preferable to employ a method that avoids oxygen-containing precursors while growing in an oxygen-depleted environment.

PVD, on the other hand, is simpler. A bulk form of the desired material (e.g. WS$_2$ powder) is heated, producing a vapor at a high enough temperature. PVD has been used to successfully grow some monolayer TMD crystals, but a variety of limitations exist. For example, it is challenging to grow TMDs that lack a sufficient
vapor pressure for transport at high growth temperatures under vacuum.

Therefore, I developed a method combining the benefits of CVD and PVD, namely using the bulk product as a single precursor, while including a chemical process to sidestep the lack of vapor pressure. In fact, for bulk crystal growth, chemical vapor transport (CVT) is widely used for just this purpose. By adding a reactive and volatile transport agent, often a halogen gas (Cl₂, Br₂, or I₂) or a halide compound, the non-volatile compound (WS₂) can chemically produce products in the gas phase that undergo vapor transport. A temperature gradient is typically established between the source zone and the growth zone in a sealed ampule using values from thermodynamic calculations. As a result, bulk crystals are slowly grown in the zone where formation is favored. [98–101]

3.1 Synthesis of Monolayer TMDs

A few techniques have been explored for growing the more stable disulfides, MoS₂ and WS₂. Chemical vapor deposition (CVD) is the most popular, specifically using elemental sulfur and metal oxides (i.e. MoOₓ or WOₓ) as precursors. [84–86] Other precursors have also been reported, including metal chlorides (e.g. MoCl₅ and WCl₆) or carbonyls (e.g. Mo(CO)₆ or W(CO)₆) with hydrogen sulfide gas (H₂S). [102–109] Although successful, these methods become challenging for TMDs that are less stable due to rapid oxidation (e.g. NbSe₂). [87–90] For those TMDs, it would be preferable to employ a method that avoids oxygen-containing precursors while growing in an oxygen-depleted environment.

3.1.1 Exfoliation

Micromechanical exfoliation is often referred to as the Scotch™ tape method in reference to the initial demonstration in isolating single layers of graphite, now known as graphene. [82] While other tapes, or polymers in general, are often chosen to reduce residue and improve flake production, the method is inherently the same. Bulk crystals, either natural or synthetic, are pressed between two sticky surfaces
and repeatedly peeled, splitting layers. Once reduced to a satisfactory thickness on the tape (based partly on experience and partly on speculation), the sticky surface containing multi-layer TMDs is pressed onto a receiving substrate. TMD layers are transferred to the substrate by slowly removing the sticky surface. With luck, you can transfer and identify a single layer (see Chapter 4).

The process of exfoliation is fraught with dogma and is poorly understood overall. Few studies have identified variables that improve the likelihood or size of deposited monolayers. [110] I have seen numerous students deserving ridicule when performing exfoliations, convinced their elaborate method helped but was, in fact, little more than blind faith. In the end, exfoliation is a game of chance and the odds are relatively poor no matter the steps taken. If not for the quality of bulk crystal growth relative to direct monolayer growth so far, researchers would disregard exfoliation. Flakes produced via exfoliation are random in size, shape, location, and occurrence, all of which makes fabrication difficult and serial in nature. It cannot scale for mass production, and it does not work for the majority of novel TMDs. Yet until a bottom-up growth technique can achieve exfoliation quality in a consistent manner, academic research will continue to struggle with exfoliation.

### 3.1.2 Chemical Vapor Deposition

The standard bearer for bottom-up growth of monolayer TMDs is chemical vapor deposition (CVD). The idea of CVD for TMD growth is to choose two or more precursors containing the constituent atoms desired in the final compound, and then chemically react them near a substrate where they can deposit.

CVD research is aplenty, but the method of choice for the majority of studies is to reduce metal oxides (e.g. MoO₅) with the help of the desired chalcogen vapor (S₂) at high temperature. The whole process takes place in a tube furnace, usually under inert gas flow at atmospheric pressure. [84–86] I use MoS₂ as an example here because it is the most mature CVD-grown material, but you can replace Mo with W for a similar process to grow WS₂. The only difference is that WO₅ requires higher temperatures and the growth is less stable. In general, most CVD methods grow
more difficult moving from S to Se to Te, and from Mo to W. Thus MoS$_2$ is relatively easy to grow while WTe$_2$ is extremely difficult due to its relative instability.

Other precursors have been used, including replacing the metal oxides with metal chlorides or carbonyls [102–109], and replacing pure chalogen with extremely dangerous, but reactive, dihydrogen chalcogenides (e.g. H$_2$S) in gas form. [103,106,109] Chemically, these reactions are usually more favorable, but the associated dangers tip the scales in favor of the oxide precursors for most academic labs.

3.1.3 Physical Vapor Deposition

One commonly reported alternative to CVD is physical vapor deposition (PVD). PVD is simpler, as it uses the bulk form of the desired material (e.g. WS$_2$ powder) as the only precursor. By heating the material under vacuum, a vapor is produced and the desired mono- to few- layers can be deposited on a cooler substrate. PVD has been used to successfully grow some monolayer TMD crystals [91–97], but a variety of limitations exist. For example, it is challenging to grow TMDs that lack a sufficient vapor pressure for transport at high growth temperatures under vacuum.

3.2 Chemical Vapor Transport of Bulk Crystals

A method combining the benefits of CVD and PVD is desired, namely using the bulk product as a single precursor, while including a chemical process to sidestep the lack of vapor pressure. In fact, for bulk crystal growth, chemical vapor transport (CVT) is widely used for just this purpose. By adding a reactive and volatile transport agent, often a halogen gas (Cl$_2$, Br$_2$, or I$_2$) or a halide compound, the non-volatile compound (WS$_2$) can chemically produce products in the gas phase that undergo vapor transport. A temperature gradient is typically established between the source zone and the growth zone in a sealed ampule using values from thermodynamic calculations. As a result, bulk crystals are slowly grown in the zone where formation is favored. [98–101] Figure 3-1 summarizes the reaction where pink crystals are grown at temperature T$_1$ with the help of a volatile transport agent (violet) that reacts with the original
precursor at temperature $T_2$.

![Diagram](image)

Figure 3-1: General concept of CVT growth in a closed ampule system. The temperature gradient from $T_2$ to $T_1$ is calculated from thermodynamics. Figure courtesy of [101].

The key to CVT is a forward and backward reaction that co-exist in the established temperature gradient. The forward reaction generates volatile species while the backward reaction slowly builds up the crystal while regenerating the reactive species needed for the forward reaction. This cycle continues until the original non-volatile powder is completely consumed on the reactive side of the ampule. Generally, a CVT reaction would have the following chemical reaction, where the right side could be a single product or more; here assumed two reactants as is common. It is also important to note that depending on if the reaction is exothermic or endothermic (displayed), the location of the resulting crystal will be in the hotter or colder zone, respectively.

$$A(s) + B(g) \rightleftharpoons C(g) + D(g) \quad (3.1)$$

A traditional example of CVT would be the process by which tungsten halogen light bulbs work. When the tungsten filament heats, it emits black-body radiation, some in the visible spectrum that we visibly see. But much of the energy is lost as heat and acts to slowly evaporate the tungsten, which reacts with oxygen to form tungsten oxide. If a halogen, commonly iodine ($I_2$) is also added to the bulb, the W instead can produce $WO_2I_2$, a gaseous species. An exothermic reaction near the filament can decompose the metal oxyhalide to redeposit tungsten back on the filament. In this way, the filament degrades much slower, improving the lifetime of the light bulb.

$$WO_2(s) + I_2(g) \rightleftharpoons WO_2I_2(g) \quad (3.2)$$
3.3 Synthesis of Monolayer WS$_2$ via CVT

Although CVT is ubiquitously used to synthesize bulk crystals, it was only recently reported that atomically-thin TMDs, including WS$_2$, can be grown using CVT in a sealed ampule by carefully controlling mass transport. [111, 112] While using I$_2$ and metal chlorides as transport agents, the presented closed-tube technique still requires metal oxides and sulfur as precursors, akin to CVD where stoichiometry is always a consideration. In this work, we explore the possibility of using CVT in an open chamber under vacuum to achieve mono- and few-layer TMDs. In addition, we improve on previous work by using bulk TMDs as their own oxygen-free precursors, simplifying the process by guaranteeing initial stoichiometry.

We selected WS$_2$ for our investigation; to start, WS$_2$ growth has been previously reported using both CVD and PVD methods, so those results can serve as a comparison. Second, PVD growth of WS$_2$ bears the challenges mentioned above: WS$_2$ lacks a sufficient vapor pressure even at 1000 °C under vacuum and actually dissociates at 1250 °C rather than evaporating. [113] Furthermore, monolayer WS$_2$ itself is a material of great interest as was discussed in the previous chapters.

Figure 3-2a illustrates our experimental setup. In order to limit growth to single and few layers, we use an open tube design, in contrast to the closed ampule employed for bulk crystal growth. For bulk crystals, a closed system allows recycling of the transport agent, producing large crystals. By using an open design with a limited dose of halide transport agent, we have found that the reaction can be slowed to produce crystals down to a single layer. We also perform growth at low process pressure (100 mTorr - 10 Torr) under regulated gas flow, which adds a level of control to crystal nucleation and growth.

Figure 3-2b shows the measured temperature profile and temperature gradient of our single zone tube furnace. During a typical growth run, a carrier gas of argon with 5% hydrogen flows at 150 sccm while the temperature of the furnace is increased to 950 °C at 30 °C min$^{-1}$. At the center of the furnace sits a quartz crucible containing WS$_2$ powder mixed with sodium chloride (NaCl). Downstream, a SiO$_2$/Si (300 nm
Figure 3-2: Concept of CVT growth of WS$_2$ with an open-tube design. (a) Schematic of our open-tube chemical vapor transport growth system. At the center of a single-zone tube furnace, a crucible contains WS$_2$ powder mixed with NaCl that is heated to 950 $^\circ$C, forming tungsten chloride(s) and sulfur gases. The reverse process occurs downstream, resulting in monolayer WS$_2$. (b) The measured temperature and temperature gradient as a function of distance from the furnace's edge, indicating where optimal growth occurs.

oxide) growth substrate is located in the optimum location for crystal deposition, where the temperature is approximately 700 $^\circ$C - 750 $^\circ$C when the set point is reached (see calculation in Section 3.3.2). At the same time, the temperature at the edge of the furnace (18 cm from the center) is roughly 300 $^\circ$C. The temperature gradient, especially at locations close to the edge of the furnace, is quite large with values $>$50 $^\circ$C cm$^{-1}$. It is worth noting that achieving large monolayer flakes by PVD is challenging in such large gradients. Thus, optimal PVD synthesis uses furnaces with three heating zones. [94,97]

In order to obtain monolayer deposition on the growth substrate, it is important to choose a suitable transporting agent. Based on thermodynamic calculations, chlorine is predicted to be a suitable transport agent for WS$_2$ (see calculations in Section 3.3.2)). The proposed reaction includes the formation of volatile tungsten chlorides (of various oxidation states) and diatomic sulfur gas as products at temperatures above 700 $^\circ$C, with slight variation depending on the assumed oxidation state of tungsten (x=1-6).
\[ WS_2(s) + \frac{x}{2} Cl_2(g) \rightleftharpoons WCl_x(g) + S_2(g) \] (3.3)

We found the use of sodium chloride (NaCl) as the source for chlorine to work extremely well in our system. Figure 3-3a and b compare the typical growth results with (Figure 3-3a) and without (Figure 3-3b) NaCl added to WS\textsubscript{2} powder. In the absence of NaCl, the growth process is the same as PVD, which we found to be unreliable for monolayer WS\textsubscript{2} growth. Presumably, the main challenges are WS\textsubscript{2}’s low vapor pressure, indicated by the need for a 1050 °C growth temperature, and the large temperature gradient in a single zone furnace, as mentioned previously. When NaCl is mixed with WS\textsubscript{2} (20 wt%), the production of monolayer WS\textsubscript{2} becomes significantly more reproducible. To verify NaCl as a transport agent, we grew bulk crystals via conventional CVT by placing 1 g of WS\textsubscript{2} powder and 200 mg of NaCl in an evacuated, sealed quartz ampule. The ampule was placed in a three-zone furnace with a temperature gradient from 800 °C to 730 °C. Total growth time was 15 days. Figure 3-3c shows the resulting crystals from the low temperature region with similar morphology to our atomically-thin samples, indicating the same mechanism. Additional experiments agree with this conclusion, as discussed in Section 3.3.3.

Of the various parameters, the amount of NaCl mixed with WS\textsubscript{2} is the most important for creating monolayers, acting as the limiting factor for crystal growth. Once consumed, growth halts. In addition, the relative ratio of WS\textsubscript{2} to NaCl greatly affects growth quality. For our setup, we found the best ratio of WS\textsubscript{2} to NaCl was 4:1 by weight (Figure 3-3e). Adding excess salt resulted in broken WS\textsubscript{2} crystals and undesired particles (Figure 3-3f, also see 3.3.5 for XPS analysis and discussion of these particles). Too little salt resulted in smaller or nonexistent crystals (Figure 3-3d).

Moreover, we discovered the process pressure range of 3-5 Torr to work well for producing large monolayers (Figure 3-3h). Higher pressures result in a larger occurrence of pyramidal growth with monolayer edges (Figure 3-3g). Lower pressures resulted in few monolayers; instead usually resulting in uniform bilayers and trilayers (Figure 3-3i). These structures have been difficult to achieve by CVD, thereby
Figure 3-3: Optical micrographs of WS$_2$ growth. (a) Growth using NaCl added to the WS$_2$ precursor crucible, as compared to (b) growth without, requiring a higher temperature and growth time to produce smaller crystals. (c) Bulk crystal growth via conventional CVT in a sealed quartz ampule using WS$_2$ and NaCl as the transport agent. (d-f) Comparison of CVT growth with different WS$_2$:NaCl ratios. (g-i) Comparison of CVT growth with different process pressure, noting the uniform bilayers in (g). Unless otherwise denoted, growth occurs with the furnace temperature set to 950 °C, with a process pressure of 3 Torr, and a WS$_2$:NaCl ratio of 4:1. All scale bars are 20 μm unless denoted.

offering an added benefit of this method. Using optimized parameters, we are able to achieve some triangular monolayer flakes with edges over 50 μm in length (Figure 3-3h), while typical flakes have edges that are 10-30 μm long.

Besides sodium chloride, other transporting agents were investigated as well. Chlorine gas (100 ppm in argon) was attempted directly, but only thick crystals could be obtained, seemingly because even that amount of Cl$_2$ is still too high compared to the residual amount generated from NaCl at high temperatures (Figure 3-5). Iodine (I$_2$), on the other hand, was too difficult to control due to its low sublimation temperature,
and no significant deposition occurred. In the end, halide salts were the easiest and most consistent transport agents (see Section 3.3.4 for results from various transport agents). Interestingly, NaCl and other alkali halide salts were previously shown to chemically assist with the growth of WS$_2$ and WSe$_2$ monolayers, but in that study tungsten oxide and sulfur were still used as precursors, with the limitations previously discussed. [114]

3.3.1 Process Details for WS$_2$ Growth

At the center of a 1" tube furnace (Thermo Scientific Lindberg/Blue M Mini-Mite), 99.99% pure WS$_2$ powder (Alfa-Aesar) is mixed with NaCl salt (Sigma-Aldrich) and placed in a fused-silica crucible (MTI Corp.) inside a quartz tube. Empirically, adding only 20 wt% NaCl to 10 mg of WS$_2$ resulted in the best growth. Downstream from the crucible, a 1" x $\frac{1}{2}$" SiO$_2$/Si substrate (300 nm oxide) was placed face up. Before growth, the entire tube is pumped to low vacuum ($1 \text{ mTorr}$), followed by an extensive process to remove oxygen and moisture from the tube. Since oxidation of WS$_2$ readily occurs at high-temperatures, high-quality growth requires the removal of oxygen and water from the system, both during and between runs.

Upon closing the compression fittings as tightly as possible, the isolation valve to a scroll pump (Oerlikon Leybold ScrollVac SC15D) was opened. To obtain base pressure quickly, ultra-high purity argon (Airgas Grade 6 Ar) was used as a purge gas. At least 5 pump/purge cycles were used before every run. Next, argon at 150 sccm flowed through the tube for at least 5 minutes; the longer the better. After, base pressures were $<5 \text{ mTorr}$ with rise rates $<5 \text{ mTorr min}^{-1}$. To further remove system impurities, pumping continued for a minimum of 60 minutes or more, until based pressure was approximately 1 mTorr and rise rates were $<1 \text{ mTorr min}^{-1}$, indicating minimal leakage and outgassing. During unloading and loading of material, the furnace temperature is never dropped below 100 °C to minimize water from condensing while the tube is open. When the system is in standby between runs, vacuum is maintained and temperature is set to 300 °C to prevent condensation of impurities. Finally, every experimental run begins with a dummy run to bake out
the tube. The dummy run consists of heating the empty furnace to 1050 °C for 1 hour with pure argon flowing at 150 sccm. Empirically, runs following this dummy run were improved compared to runs without, showing poorer optical quality.

As the entire process occurs under vacuum, equilibrium is reached between the injected and exhausted gases. Using 150 sccm Ar/H$_2$ flow, process pressure is approximately 1 Torr. To give another degree of freedom to the system, a butterfly valve can dynamically control pressure independent of gas flow. Process pressures during typical growth runs is manually adjusted between 3-5 Torr. Growth occurs at a set point temperature of 950 °C for 20 minutes, at which point the furnace cools naturally to 600 °C. Finally, the lid is opened and the furnace cools further until 200 °C when the substrate and crucible are cool enough to handle.

### 3.3.2 Thermodynamic Calculations

The use of NaCl as the transport agent proceeds from the thermodynamic calculations for the reaction of WS$_2$ with chlorine gas (Cl$_2$). The goal is to find whether or not chlorine acts as a valid transport agent for WS$_2$. If it is, the reaction’s free Gibbs energy, either product-favored or reactant-favored, will be close to zero and flip signs in a temperature gradient that is practical for our system design. This is equivalent to saying the reaction’s equilibrium constant $K_{eq}$, switches from being greater than 1 (product-favored) to less than 1 (reactant-favored) on either side of an optimal temperature $T_{opt}$.

\[
\Delta G_r^0 = \Delta H_r^0 - T \Delta S_r^0
\]  
\[
\Delta G_r^0 = 0
\]  
\[
T_{opt} = \frac{\Delta H_r^0}{\Delta S_r^0}
\]

The first step is to assume a few different possible reaction pathways and calculate
the standard enthalpies and entropies for each. This is done by applying Hess's law to the balanced chemical reaction formula using the standard enthalpies and entropies of formation for the constituent compounds, found in thermodynamic data tables. [1] Table 3.1 provides the values used. Note the standard enthalpy and entropy data is at standard conditions of 300 K and 1 atm. Our tube is at low pressure (<10 Torr), but we assume standard pressure is a good approximation.

\[
\Delta H_r^0 = \sum_{\text{products}} \Delta H_f^0 - \sum_{\text{reactants}} \Delta H_f^0
\]  

(3.7)

\[
\Delta S_r^0 = \sum_{\text{products}} S_f^0 - \sum_{\text{reactants}} S_f^0
\]  

(3.8)

We present one such calculation and discuss the results. The presented pathway assumes chlorine reacts with WS\(_2\) to produce WCl\(_4\) and S\(_2\). The balanced equations is:

\[
\text{WS}_2(s) + 2\text{Cl}_2(g) \rightleftharpoons \text{WCl}_4(g) + \text{S}_2(g)
\]  

(3.9)

Using the previous equations and data from Table 3.1, we can find the reaction enthalpy and entropy. Using those values, the optimum temperature where the reaction equilibrium flips is found.

\[
\Delta H_r^0 = \sum_{\text{products}} \Delta H_f^0 - \sum_{\text{reactants}} \Delta H_f^0 = (-335.98+128.6)-(-298.3+2*0) = 90.92 \text{ kJ/mol}
\]  

(3.10)

\[
\Delta S_r^0 = \sum_{\text{products}} S_f^0 - \sum_{\text{reactants}} S_f^0 = (379.27 + 228.2) - (67.78 + 2*223) = 93.69 \text{ J/K.mol}
\]  

(3.11)
\[ \Delta G_r^0 = 0, \quad T_{opt} = \frac{\Delta H_r^0}{\Delta S_r^0} = \frac{90.92 \text{ kJ/mol}}{93.69 \text{ J/K.mol}} = 970.4 \text{ K} = 697.28^\circ C \] (3.12)

Looking at the absolute Gibbs energy values at 950 °C (-23.66 kJ mol\(^{-1}\)) and 25°C (63.00 kJ mol\(^{-1}\)), we see that above 700 °C, the reaction produces WCl\(_4\) and S\(_2\). Below the optimum temperature, the reaction is reactant favored and WS\(_2\)–Cl\(_2\) can reform. Thus, chlorine looks to be a good transport agent if this reaction can occur. Looking at our growth results, our tube has a center temperature of 950 °C, but the substrates are located where the temperature has decreased to around 700 °C, in good agreement with the previous calculation.

Tungsten chloride has other stoichiometric mixes as the oxidation state of tungsten varies. In addition, alternative pathways include the formation of H\(_2\)S, but calculations indicate these are not good for vapor transport given our experimental conditions. Table 3.2 includes a few reactions and their respective calculated quantities.

**Table 3.1: Standard thermodynamic data used in reaction calculations. [1]**

<table>
<thead>
<tr>
<th>Component</th>
<th>( \Delta H_r^0 )</th>
<th>( S_r^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl(_2)(g)</td>
<td>0</td>
<td>223</td>
</tr>
</tbody>
</table>
Table 3.2: Thermodynamic calculations of various reaction pathways involving chlorine gas.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>$\Delta H^0$</th>
<th>$\Delta S^0$</th>
<th>$T_{opt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{WS}_2(s) + 2\text{Cl}_2(g) \rightleftharpoons \text{WCl}_4(g) + \text{S}_2(g)$</td>
<td>90.92</td>
<td>93.69</td>
<td>970.43</td>
</tr>
<tr>
<td>$\text{WS}_2(s) + \text{Cl}_2(g) \rightleftharpoons \text{WCl}_2(g) + \text{S}_2(g)$</td>
<td>414.35</td>
<td>246.83</td>
<td>1678.7</td>
</tr>
<tr>
<td>$\text{WS}_2(s) + 3\text{Cl}_2(g) \rightleftharpoons \text{WCl}_6(g) + \text{S}_2(g)$</td>
<td>-66.81</td>
<td>-89.44</td>
<td>746.98</td>
</tr>
<tr>
<td>$\text{WS}_2(s) + 2\text{Cl}_2(g) + 2\text{H}_2 \rightleftharpoons \text{WCl}_2(g) + \text{H}_2\text{S}(g)$</td>
<td>245.55</td>
<td>169.23</td>
<td>1450.98</td>
</tr>
</tbody>
</table>

3.3.3 Evidence of Growth by Chlorine Transport

The use of sodium chloride as a source of chlorine for vapor transport is not an obvious one, but our evidence points to that conclusion. First, attempting PVD growth in the same tube furnace, at the same temperatures and under the same conditions, results in vastly different results than when we use NaCl (Figure 3-3a-b). Second, growth runs were performed with the NaCl in a separate, upstream crucible from the WS$_2$ powder and we were still able to obtain WS$_2$ triangular crystals, pointing towards a vapor phase reaction. To ensure NaCl is not simply acting as nucleation sites for WS$_2$ vapor, experiments where salt from water solution was dropcast onto substrates were performed. Results showed no WS$_2$ growth. Fourth, XPS data (Figure 3-6a) was obtained for areas on the substrate deemed 'salt circles' because they only appeared with NaCl added as a precursor. XPS surprisingly did not show equal amounts of sodium and chlorine in these areas, interestingly showing no chlorine, but only sodium and tungsten as the primary components (aside from silicon and oxygen) (see Figure 3-6). This indicates sodium and chlorine are separate on the growth substrate, agreeing with the chlorine-driven mechanism.

In attempting to use pure chlorine gas, diluted in argon to roughly 100 ppm, we saw thick crystals forming under the same growth conditions, but we were unable to grow few layer crystals in this manner (Figure 3-4). Visually, it was obvious the chlorine was reacting with the WS$_2$ to form a chloride or sulfur chloride (e.g. WS$_2$Cl$_2$),
Figure 3-4: (a-b) Micrographs of growth using chlorine gas diluted in argon (100 ppm), resulting in thick hexagonal crystals. Scale bars are 20 μm. (c) Photograph of growth tube after a run using chlorine, showing brown-orange band suspected to be tungsten chloride. This band turns blue-black upon exposure to moisture, consistent with tungsten chloride (WCl₆).

which was never clear with NaCl (Figure 3-4c). The conclusion is that chlorine gas changes the rate of reaction to a point that controlled, slow growth required for monolayers is not possible. Similar growth occurred using silver chloride (AgCl), which is known to easily decompose to release chlorine at elevated temperatures.

Two other experiments add to our hypothesis of the mechanism using NaCl as a source for chlorine. First, we were able to confirm chlorine gas was produced by heating NaCl to 950 °C under the sample conditions. A spot test for Cl₂ gas was performed in-situ using filter paper soaked in a 10 wt% solution of 4,4'-bis(dimethylamino)thiobenzophenone (DMATBP, Sigma) in benzene (Sigma). DMATBP acts a colorimetric reagent for halogens, as detected by a change from pale yellow to blue. [115] Second, WS₂ powder was mixed with NaCl per the usual procedure, but instead of putting upstream in a crucible, it was deposited directly on the substrate and placed in the tube’s center. Heating to 800 °C for 1 hour, similar, but low-quality growth occurred, with some monolayers present.

The final note is that chemical vapor transport, with NaCl as a transport agent, has been used to purify beryllium (Be), vanadium (V), uranium (U), and titanium (Ti) metals, per previous patents. [100,116]
3.3.4 Growth with Different Alkali Salts

In addition to NaCl, other alkali halide salts were tested. Potassium chloride (KCl), potassium iodide (KI), and sodium iodide (NaI) all show successful CVT growth of lesser quality, but conditions were not optimized. Knowing each has a different melting point, temperature would be the first parameter to adjust for improved growth. Due to similar initial run quality, we decided NaCl was preferred, as it is less hazardous and better controlled. Interestingly, sodium bromide (NaBr) did not work under the same conditions. Optical micrographs of the iodide and chloride salts is given below (Figure 3-5). The conditions used for all salts were the same as the best NaCl runs, namely 950 °C center temperature and a 4:1 WS₂ to salt ratio, with a process pressure 3 Torr.

![Micrographs of CVT growth using different alkali halide salts. All scale bars are 20 μm.](image)

3.3.5 XPS of Impurity Circles

Often after growth we noticed nearly perfect circles that appeared only after NaCl (or other salt) was added to the system. We suspected these circles were salt that condensed as a liquid and cooled to form a circular coffee stain pattern. Interestingly, these circles could not be removed when dipped in water, nor acetone, chloroform, or isopropyl alcohol. These 'circles' also appeared in the bulk CVT growth ampule near the source zone, indicating they are not related to the use of hydrogen or other impurities in the system, but rather a product of the known compounds. This includes
the quartz furnace (SiO$_2$), or substrate. Tape used to remove the bulk crystals for transfer and exfoliation was unable to adhere to these features. Furthermore, when performing flake transfers with hydrofluoric acid as a SiO$_2$ etchant, we observe these circles do not lift off with the flakes, but have become part of the substrate, etched alongside the SiO$_2$ all the way down until the silicon is revealed. XPS was employed to further study, indicating these circles are a compound that contains Na and W, with possible additions from Si and O in the substrate (Figure 3-6a). Our best guess is sodium tungstate (Na$_2$WO$_4$) or another compound that also incorporates silicon. In any case, the WS$_2$ crystals do not have these XPS peaks and show a tungsten-to-sulfur ratio of 66.11 : 33.89 (~2:1), indicating high quality (Figure 3-6b).

![Figure 3-6](image)

**Figure 3-6**: XPS data from (a) Circles that appear with NaCl usage (optical micrograph of example circles, 20 $\mu$m scale), and (b) WS$_2$ triangles, including corresponding atomic concentration

<table>
<thead>
<tr>
<th>Atomic Concentration Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2s</td>
</tr>
<tr>
<td>0.426</td>
</tr>
<tr>
<td>9.306</td>
</tr>
<tr>
<td>66.11</td>
</tr>
</tbody>
</table>
3.3.6 Attempts with Other TMD Crystals

In addition to WS$_2$, we also attempted to test this CVT technique with other TMDs, namely the tungsten-based TMDs WSe$_2$ and WTe$_2$. We also tried to grow TaS$_2$, a superconductor with increasing Curie temperature when thinned to fewer layers. Unfortunately, TaS$_2$ showed no signs of success with NaCl under the conditions we attempted. On the other hand, we saw some success with the W-based TMDs, indicating that the tungsten chemistry was likely a key to our choice of NaCl for growing high-quality WS$_2$.

**Tungsten Diselenide (WSe$_2$)**

Moving from WS$_2$ to WSe$_2$, growth is more difficult as WSe$_2$ is more unstable and easier to oxidize. However, simply replacing WS$_2$ powder with WSe$_2$ powder resulted in thick flake growth on the same SiO$_2$/Si substrates. Figure 3-7 shows an optical micrograph from an unoptimized run under the same conditions as WS$_2$. Aside from the triangular geometry, the Raman signal from this flake clearly indicates it is WSe$_2$ as desired. No other tests or optimizations were conducted aside from this proof of concept showing similar chemistry could be used.

![Optical micrograph of WSe$_2$ grown via CVT technique with NaCl under similar conditions as WS$_2$.](image)

**Tungsten Ditelluride (WTe$_2$)**

While WS$_2$ and WSe$_2$ are close cousins, WTe$_2$ is vastly different in many ways. First, WTe$_2$ is a Type-II Weyl semimetal. This property has implications in ultra-fast...
electronics, due to the formation of massless, topologically-protected Fermions in the bulk material. Second, WTe$_2$ has been demonstrated to have a gigantic, non-saturating magnetoresistance (450,000%) up to 14.7 Tesla, useful in spintronics applications. [117] Finally, at the monolayer limit, WTe$_2$ has recently been shown to be a true 2D topological insulator, of keen interest to condensed matter physicists pursuing the quantum spin hall effect (QSHE). [118]

Trying WTe$_2$ growth with NaCl did not result in WTe$_2$ crystals, at least under the conditions that were attempted. Instead, the result was only tellurium nanowires and tellurium platelets, presumably due to WTe$_2$ dissociation. By changing the salt to silver chloride (AgCl) with a lower melting point, I could lower the growth temperature to 500 °C. Under those conditions, WTe$_2$ crystallites are present. In the Raman spectrum, the peaks near 160 cm$^{-1}$ and 209 cm$^{-1}$ are indicative of bulk WTe$_2$. Beyond this initial demonstration, further optimization was not pursued, but it is clear that the proper choice of salt can permit the growth of WTe$_2$, which is normally very difficult.

![Raman SI](image)

Figure 3-8: Optical micrograph of WTe$_2$ grown via CVT technique with NaCl under similar conditions as WS$_2$. 

77
Chapter 4

Characterization of WS$_2$ Crystal Quality

In the previous chapter, a method for growing WS$_2$ monolayers was demonstrated. Motivation for developing an improved growth technique included the need for exfoliation quality crystals with improved throughput. To examine how these CVT-grown monolayers compare to exfoliated ones, common materials characterization is used. Optical microscopy can identify monolayer thick samples, confirmed with atomic force microscopy. Raman and photoluminescence (PL) spectroscopy show CVT flakes are bright and of similar quality compared to exfoliated samples. Finally, time-resolved PL (TRPL) is used to examine exciton dynamics, a primary concern for optical and electronic applications.

4.1 Optical Microscopy

The most basic characterization involves an optical microscope with white light illumination. The most important parameters established from the optical microscope are flake size, flake shape, and flake density. In addition, growth occurs on substrates consisting of 300 nm silicon oxide on silicon, creating a thin-film stack for enhanced contrast of monolayer flakes. A trained eye can discern single layers with nothing more than a microscope. Figure 4-1 gives some example microscope images of mono-
layers from the CVT growth presented in Chapter 3. On the images, you can also find thicker regions, which show up brighter and of different colors.

![Image of optical micrographs of monolayer WS2 samples grown by CVT with NaCl. All scale bars are 20 µm.]

**Figure 4-1:** Optical micrographs of monolayer WS2 samples grown by CVT with NaCl. All scale bars are 20 µm.

4.2 Atomic force microscopy (AFM)

Single layers of WS2 have a thickness of around 0.7-0.8 nm. Aside from optical contrast in an optical microscope, atomic force microscopy is the best way to verify the number of layers that a flake contains. Atomic force microscopy in tapping mode is most common. Due to interactions with the substrate, step heights are often measured around 1 nm for exfoliated monolayers, but growth samples are better bonded to the substrate so heights are usually less than 0.8 nm. An AFM image from a CVT sample is shown in Figure 4-2. The image displays good height uniformity matching the underlying substrate (RMS roughness = 0.31 nm) with no visible nanoscale defects.
4.3 Raman Spectroscopy

Raman spectroscopy is a common tool for crystal analysis as it probes lattice vibrations, or phonons, characteristic of a given structure. In TMDs, a common feature of the Raman spectra is two peaks, both due to optical phonons. The first mode is the out-of-plane $A_{1g}$ mode where the two chalcogen atoms in each layer oscillate opposite of each other perpendicular to the extended plane. The second is again an optical phonon, labeled $E_{2g}^1$, which is the in-plane mode, where opposite basis atoms (W and S) vibrate opposite each other. Figure 4-3 illustrates these two modes, and the other modes common in 2D TMDs.

Figure 4-3: Vibrational modes common in few layer TMDs, including modes than appear in Raman spectroscopy. The $E_{2g}^1$ and $A_{1g}$ modes are easily seen in the Raman spectrum of WS$_2$. Image is courtesy of [119].
Interestingly, Raman spectroscopy can also be used to identify flake thickness since the energy difference between $E_{2g}^{1}$ and $A_{1g}$ peaks decreases monotonically as the layer number decreases. For WS$_2$ monolayers, these peaks are separated by about 61 cm$^{-1}$. Finally, Raman spectroscopy is a sensitive probe of crystal quality, indicating inhomogeneities in the lattice, often due to defects, strain, or doping. To characterize the CVT grown flakes, Raman spectroscopy with a 532 nm laser is used and compared to samples produced by exfoliation. The results, provided in Figure 4-4, show a peak separation of 61.0 cm$^{-1}$ between the in-plane ($E_{2g}^{1}$) and the out-of-plane modes ($A_{1g}$). Comparing to the exfoliated sample, the peak wavenumbers are the same and peak ratios are similar, validating our monolayer thickness from AFM and indicating comparable crystal quality. [120]

Raman data is acquired using a Horiba LabRAM HR800 system with a 532 nm CW laser. An 1800 line mm$^{-1}$ grating is used with a 100x objective (NA=0.9). PL emission to $A_{1g}$ Raman mode ratios (see Section 4.4) were calculated using data from the same setup by scanning from 200 cm$^{-1}$ to 5000 cm$^{-1}$.

![Raman spectrum](image)

Figure 4-4: Raman spectrum for exfoliated (green) and CVT (blue) monolayers, showing separation between the $E_{2g}$ and $A_{1g}$ peaks of 61.0 cm$^{-1}$. 
4.4 Photoluminescence Spectroscopy

Photoluminescence (PL) is the radiative recombination of electron-hole pairs (EHPs) initially generated by absorption of photons larger than the band gap. The relative intensity and line shape of PL indicates crystal quality, while the emitted photon energy (wavelength) indicates the exciton energy.

Photoluminescence from the CVT-grown WS$_2$ flakes is bright relative to exfoliated samples (Figure 4-5a). Compared to the other common direct gap group VI TMDs (MoS$_2$, MoSe$_2$, and WSe$_2$), WS$_2$ is the brightest with a PL quantum yield reported up to 20% [54]. WS$_2$ also has the bluest emission of the common TMDs, with the peak occurring between 620-640 nm at room temperature. With this improved light emission and absorption compared to other 2D semiconductors, monolayer WS$_2$ has found interest in atomically-thin optoelectronic devices, such as LEDs, lasers, and photodetectors [42,121–125].

![Figure 4-5: (a) Relative and (b) normalized photoluminescence from as-grown CVT monolayers versus exfoliated monolayers. CVT crystals appear brighter with sharper linewidth than exfoliated samples. The 10 nm shift in peak wavelength is attributed to strain generated during growth.](image)

While bright, the peak wavelength of 635 nm from our CVT grown sample is red-shifted by 10 nm compared to the exfoliated sample (Figure 4-5b), but still within the range observed by others for samples produced via CVD and PVD [96,126–130]. The hypothesis is that this is a result of crystal strain during growth due to the difference in thermal expansion between the oxide substrate and WS$_2$ material, mostly during
the cool-down phase. This proposal has been explored with CVD samples showing transfer relieves the strain and blue-shifts the PL peak. [39] Using a poly(methyl methacrylate) (PMMA)-mediated technique with hydrofluoric acid to etch the SiO₂, we transferred as-grown flakes to a similar donor substrate (see Appendix A). A blueshift in PL emission is observed that closely matches the exfoliated sample (Figure 4-6). In addition, the peak for monolayers exfoliated from the salt-assisted thick crystals (Figure 3-3c) was 625 nm, matching closely with reference flakes exfoliated from bulk crystals synthesized with iodine. Finally, the linewidth of PL emission is narrow, with a FWHM of 48 meV, which compares to a FWHM of 61 meV for the exfoliated flake. The difference in FWHM could be attributed to either inhomogeneous broadening and/or the presence of trions and other states that are related to crystal defects. [131]

![Normalized photoluminescence spectra before and after flake transfer, showing a blueshift in peak emission after transfer that matches closely with exfoliated samples.](image)

Following the work of Splendiani et al. on MoS₂, the ratio of PL emission to the $A_{1g}$ Raman peak is calculated as a proxy for PL quantum efficiency. [132] Previous reports give values from 20 to as high as 880 for WS₂ monolayers using 488 nm excitation. [128,133] We measure a peak ratio around 550 for the displayed monolayer in Figure 3 using 532 nm excitation at 1 μW power. Due to potential complications, such as intensity-dependent nonlinearities in PLQY and Raman modes that vary with excitation wavelength, my opinion is that this ratio should be interpreted cautiously.
as an estimate of sample quality, rather than for quantified comparison. [54]

4.5 Photoluminescence Mapping

To test uniformity of photoluminescence over the entire monolayer, a spatially-dependent PL measurement is obtained by scanning with a piezo stage. The uniformity and relative brightness of different areas allows defects and poor samples to be identified relative to good ones. Figure 4-7a presents a PL map from the same CVT-grown flake as Figure 4-2. The PL intensity slowly decreases from center to edge, with about a factor of three decline. Others have reported relatively bright edges compared to the center, apparently from localized edge defects and chemical heterogeneity. [127–129,134] The PL maps of a two other grown samples are displayed in Figure 4-7, showing defects that quench PL are easily visible and these samples should be avoided.

![Figure 4-7: Spatially-dependent photoluminescence of three CVT-grown flakes, showing good sample uniformity in (a). (b-c) Poorer quality flakes show areas where PL is quenched.](image)

4.6 Time-dependent Photoluminescence

Transient photoluminescence offers a direct probe of exciton dynamics in TMDs, unaffected by uncertainties concerning absorption and pump alignment in a confocal micro-photoluminescence experiment. Assuming the emitting state comes from radiative recombination of excitons with an intrinsic lifetime, any deviation indicates
undesired pathways that are stealing potential photons and reducing PLQY. For light-emitting applications, a high ratio of radiative events is preferred over nonradiative events, typically associated with trap states from defects and dopants. Previously reported lifetime values for monolayer WS₂ are provided in Table 4.1. WS₂ is less often studied than MoS₂, but the reported values are different by orders of magnitude, indicating a wide range of crystal quality.

<table>
<thead>
<tr>
<th>Study</th>
<th>Reported Lifetime</th>
<th>Material Source</th>
<th>Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>He et al. [135]</td>
<td>22 ± 1 ps</td>
<td>Exfoliated</td>
<td>Pump-Probe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Differential Reflection</td>
</tr>
<tr>
<td>Yuan et al. [69]</td>
<td>806 ± 37 ps</td>
<td>Exfoliated</td>
<td>TCSPC</td>
</tr>
<tr>
<td>Amani et al. [54]</td>
<td>3.4 ± 0.3 ns (treated)</td>
<td>Exfoliated</td>
<td>TCSPC</td>
</tr>
<tr>
<td></td>
<td>1.2 ns (untreated)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To characterize the exciton dynamics of flakes grown by CVT, we measure the time-dependent photoluminescence compared to the same exfoliated sample (Figure 4-8a). The third and final curve gives the instrument response function (IRF), indicating the delta response of the entire system that limits temporal resolution. All data is taken using a time-correlated single photon counting (TCSPC) setup with a pulsed 405 nm laser and a single-photon avalanche detector (see Section 4.6.1 below). The laser fluence of 80 nJ cm⁻² was the lowest level where the average photon count rate was still one order of magnitude above the dark count rate (~50 cps).

The PL transient of the CVT sample is observed to be faster than the corresponding response of the exfoliated flake, which suggests that the CVT-grown material may possess more non-radiative losses. Interestingly, the CVT sample is almost entirely dominated by a quadratic intensity dependence, consistent with recombination by multi-particle annihilation such as exciton-exciton or exciton-charge interactions. The distinction between these two annihilation mechanisms cannot be made with this measurement. The implication of such strong multi-particle interactions, even at the lowest excitation intensities available, depends on whichever process dominates and
is an active area of research. [68,69,136] The dynamics of both processes are not well understood in 2D TMDs, although the reduced charge screening and dimensionality that binds excitons seem to cause strong multi-particle interactions.

Figure 4-8: (a) Time-dependent photoluminescence comparison between CVT (blue) and exfoliated (green) monolayer samples at low intensity excitation. (b) Power-dependent PL transient for the CVT sample, indicating strong nonlinear exciton dynamics in WS$_2$ due to multi-particle annihilation, such as exciton-exciton or exciton-charge interactions.

4.6.1 Details of TCSPC

Time-dependent photoluminescence was acquired using the time-correlated single-photon counting (TCSPC) technique with a Picoharp 300 (Picoquant GmbH) in T3 time-tagged mode. An inverted microscope (Nikon Eclipse Ti) with a PLANFLURO 40x (NA = 0.95) objective is used as the platform for measurements, having a spot size of ~1 μm. Photon detection uses a single photon avalanche diode with a response time of 50 ps (SPAD, Micro Photon Devices PDM Series). The excitation source is a 405 nm pulsed laser (Picoquant LDH-P-C-405) with a 100 ps pulse width, operating at a repetition rate of 40 MHz. Average excitation power is initially set to 10 μW, low enough to avoid photo degradation and after-pulsing from the laser. Using a reflective neutral density filter wheel, intensity is further reduced until the detection rate is around 500 cps, sufficiently above the dark counts of ~50 cps.

The instrument response function (IRF) is the combined delta response of the optical and electronic components, coming mostly from the convolution of the laser
pulse and SPAD response. We measured the IRF of our system by using a fast (picosecond) emitter in solution, namely Allura Red (AR) in DI water (0.5 mM concentration). [137] On a coverslip, a droplet of solution acts as the sample. With the same 10 μW full power (to avoid power-dependent pulse shape changes), the differential detection rate is made to be almost the same as the WS\textsubscript{2} samples using neutral density (ND) filters. Allura Red is used rather than detection of the scattered laser due to the wavelength-dependent timing response of the SPAD, since WS\textsubscript{2} emits around 625 nm and the laser emits at 405 nm. As the PL lifetime is similar in magnitude to the IRF pulse width, any errors associated with the color shift are minimized. Data plotted in Figure 4-8 is normalized after background subtraction using data taken with no laser excitation for the same acquisition time.

### 4.6.2 Modeling of Time-dependent Photoluminescence

Figure 4-8b confirms the power-dependence of exciton dynamics in our CVT samples, with the lowest fluence still not showing monoexponential behavior. This is true of the exfoliated sample as well. By least-squares fitting with a simple dynamical model that includes two-body annihilation, we observe that our CVT samples remain in the exciton-exciton annihilation regime even at pump intensities of 80 nJ cm\textsuperscript{-2} (Figure 4-9b). Effectively, the dynamics are modeled well with a quadratic dependence on exciton density. The rate model and ODE solution is provided, where \( N_x(t) \) is the areal population of excitons, \( k = k_r + k_{nr} \), from the radiative and nonradiative rates, respectively, and \( k_{XXA} \) is the exciton-exciton annihilation rate.

\[
\frac{dN_x(t)}{dt} = -(k_r + k_{nr})N_x - k_{XXA}N_x^2 = -kN_x - k_{XXA}N_x^2
\]  

\[
N_x(t) = \frac{k/k_{XXA}}{\left(k_{XXA} + N_x(0)\right)e^{kt} - 1}
\]  

Fitting to the model is accomplished using iterative reconvolution with the measured instrument response (IRF). The assumed second-order model is convolved with the IRF and compared to the measured data from each sample. Using non-linear
optimization by least-squares, model parameters are adjusted iteratively until the resulting fit matches best with the measured data. The fit for the reported PL transient data in Figure 4-8a is given in Figure 4-9. Note, that the model used does not include a background offset. In practice, a background constant can be added to the fit model or the background offset can be subtracted prior to fitting. We subtract the background prior to fitting by taking a measurement without the laser for the same acquisition time as the PL signal.

![Figure 4-9: Transient PL emission from CVT sample, fit to a dynamical model that includes two-body processes by iterative reconvolution with the setup's IRF.](image)

The fit when \( k = 0 \) is indistinguishable from \( k \neq 0 \), pointing to dynamics that are not in the linear regime. The quality of fit (\( \chi^2 = 6.644 \)) indicates the exciton dynamics is dominated by a quadratic density dependence, assumed in our model to be exciton-exciton interactions. The extracted exciton-exciton annihilation rate \( (k_{XXA}) \) is \( 0.39 \pm 0.05 \text{ cm}^2 \text{s}^{-1} \), which agrees well with values previously reported for monolayer WS\(_2\), MoS\(_2\), WSe\(_2\), and MoSe\(_2\). \[67-70, 136\] At the lowest laser intensity used, the exciton areal density is approximately \( 10^{10} \text{ cm}^{-2} \), meaning two nearest excitons approximately separated by over 100 nm. Assuming exciton-exciton annihilation, this indicates strong exciton-exciton interaction over hundreds of nanometers, as previous work on monolayer WS\(_2\) has already discussed. \[69\] Further, the same model does not fit well with the exfoliated samples, as it is unable to reproduce the large curvature around 5 ns (Figure 4-8).
Chapter 5

Valley Polarization Physics in CVT-grown WS$_2$

Valley-based applications require a large degree of valley polarization to operate efficiently. While molybdenum-based TMDs have been the primary focus of valley polarization studies, tungsten-based TMDs have larger spin-orbit splitting between valence bands; on the order of 400 meV has been reported compared to 150 meV in MoS$_2$. [52,57,58] Provided the larger energy barrier needed to flip spins within a valley, photo-generated holes are expected to exhibit a long spin coherence time on the scale of nanoseconds. [138] Surprisingly, this has not been the case. Steady-state, polarization-dependent photoluminescence (PL) measurements have measured valley polarization in MoS$_2$ with values approaching 100%, and values above 40% have often been demonstrated at room temperature when pumped with photon energies closely matching the exciton emission. [72,73,91,139,140] By contrast, the largest value for WS$_2$ is roughly 40% under near-resonant pumping conditions at intermediate temperatures (200 - 250 K). [141] At room temperature, the degree of valley polarization is reported as large as 35% for CVD-grown samples when pumped at 633 nm, slightly below the exciton resonance. [142]

In this chapter, I show 48% valley polarization in monolayers of WS$_2$ directly grown by chemical vapor transport using pure WS$_2$ powder and sodium chloride as the only precursors.[28] An optical image of a typical flake, and the one that was
used for this study, is shown in Figure 5-1a. We study valley polarization on this sample as a function of sample temperature, pump energy, and pump fluence to better understand the underlying physics that depolarizes valley excitons.

Figure 5-1: (a) A typical WS₂ flake grown by CVT using WS₂ powder and NaCl as the only precursors. (b) Differential reflection for a monolayer of WS₂ grown in the same manner, and transferred to a transparent substrate. The differential reflection is proportional to absorbance, showing the A and B exciton transitions created by the large spin-orbit splitting of the valence bands. (c) The exciton resonance shifts as a function of temperature, which is often neglected when measuring valley polarization as a function of temperature with a fixed pump photon energy.

Increasing temperature and pump energy both act to reduce valley polarization in TMDs, per earlier studies. [72,74,140,143] However, as temperature increases, the direct band gap decreases in energy in a nonlinear fashion (see Figure 5-1c), an effect that is often neglected. We address this and decouple the effects of excess kinetic energy versus thermal energy on valley depolarization by fixing the pump energy to a value above the temperature-dependent exciton resonance. By comparing trends in circular polarization with relative PL quantum yields, we discuss two common valley depolarization mechanisms, specifically non-radiative exciton scattering and
5.1 Brief Review of the Valley Pseudospin in Monolayer TMDs

As discussed in Chapter 2, transition metal dichalcogenide (TMD) monolayers exhibit a direct band gap at the time-reversed corners of the hexagonal Brillouin zone, labeled K and K-. Each layer, with its trigonal prismatic structure, also lacks inversion symmetry. In combination with large spin-orbit coupling that splits the valence bands, carrier spin becomes locked with one of these energetically degenerate valleys in k-space. Consequentially, optical selection rules for the orbital moment allow optical probing of these valley populations by the helical polarization of light. Put simply, a given circular polarization of light will excite carriers in only one valley and radiative recombination should produce the same polarization (see Figure 5-2). These valley physics are summarized in Figure 5-2.

Figure 5-2: Review of coupled valley-spin physics with select valley excitation provided by circularly polarized light, per Chapter 2.
5.2 Measuring Valley Polarization via Polarization-dependent Photoluminescence Spectroscopy

To measure steady-state valley populations we use polarization-dependent PL. Exciting the monolayer semiconductor above the band gap energy (~1.98 eV at RT) with circularly polarized light should populate only one valley with excitons, per the optical selection rules. Upon radiative recombination, the relative intensity of right-versus left-handed circularly polarized photoluminescence measures the relative valley populations, per Equation 5.1. The figure of merit is the degree of valley polarization which is the difference in populations (i.e. intensity) normalized by the total population. To be complete, Equation 5.1 assumes all radiative recombination takes place only from a K valley. The nearly Lorentzian PL lineshape we measure justifies this assumption.

\[
P_{\text{circ}} = \frac{N_{K^+} - N_{K^-}}{N_{K^+} + N_{K^-}} = \frac{I(\sigma^+) - I(\sigma^-)}{I(\sigma^+) + I(\sigma^-)}
\] (5.1)

In our experiments, we use a tunable pulsed laser source (40 MHz repetition rate) to excite the CVT-grown sample at different pump energies. The tunable laser is a SuperK Extreme supercontinuum laser from NKT Photonics coupled to an acousto-optical tunable filter (AOTF) having an emission linewidth of ~4 nm. The laser is collimated and passed through a linear polarizer (LP) and Barek’s compensator (Newport). After a 50:50 beam cube, the incident laser passes through a superachromatic quarter-wave plate (QWP) to generate circularly polarized light on the sample. Due to birefringence and polarization-dependent components along the beam path, polarization at the sample is not very circular using just a LP and QWP. Calibration is required as detailed below.

Average pump power is varied from 1 μW to 20 μW for selected temperature/energy combinations, corresponding to fluence values of 200 nJ cm\(^{-2}\) to 4 μJ cm\(^{-2}\) (spot diameter = 4 μm). This order of magnitude is equivalent to \(10^{10}-10^{11}\) excitons cm\(^{-2}\) pulse\(^{-1}\), realizing both photon flux density and absorption are wavelength-dependent.
Figure 5-3 overviews the details of polarization-PL setup, using an inverted microscope (Nikon Eclipse Ti-U) as a platform. A high NA objective (50x, NA 0.7) is used to excite and collect PL from the sample in reflection geometry. The sample is situated on a custom cold finger inside of a Janis ST-500 continuous-flow cryostat to allow temperature-dependent measurements down to 78 K. The cryo is attached to a third-party XY stage (Ludl) with custom-built hardware to reduce vibrations. Finally, wavelength-dependent luminescence is collected and measured with a dispersion-type spectrograph using a Si-CCD camera (Princeton Instruments SP2300).

![Polarization-dependent PL setup diagram](image)

**Figure 5-3:** Polarization-dependent PL setup. The Barek’s compensator is used to calibrate the input light’s polarization to ensure circular polarization at the sample.

### 5.2.1 Polarization Calibration

The key to accurate valley polarization measurements is accounting for all of the different optical components that depolarize the input and output circularly polarized light. To ensure polarization is the same, each pump wavelength is separately calibrated with a Barek’s compensator (Newport) to ensure greater than 98% cir-
cular polarization to the sample. A superachromatic quarter-wave plate (Thorlabs) is placed just before the objective lens to convert linearly input light to circular, performing the reverse operation to the emitted PL. Output polarization is detected through a linear analyzer mounted on a rotating mount to distinguish between right- and left-handed components. The output path is also calibrated by tuning the laser to the temperature-dependent emission wavelength and performing a similar operation.

The Barek's compensator allows arbitrary wave retardation and axis rotation to pre-compensate the input beam to result in $>98\%$ circularly polarized light at the sample, as measured using a rotating LP and amplified photodiode. This calibration is done every time the pump wavelength is adjusted.

5.3 Valley Polarization as a Function of Pump Photon Energy

Figure 5-4 shows the polarization-dependent PL that is measured when excited with polarized light of positive helicity ($\sigma^+$) at different pump wavelengths (energies) from 500 nm (2.48 eV) to 610 nm (2.03 eV). The left pane is measured at 78 K and the right pane at 300 K, showing in general that higher temperatures result in lower polarization. From the plots, it is clear that there is a large dependence for valley polarization on the pump energy.

The primary interest is not excitation energy itself, but photon energy in excess of the temperature-dependent exciton resonance, defined as $\Delta E = E_{\text{pump}} - E_{\text{exciton}}$. Extracting the integrated PL intensities of Figure 5-4 and using Equation 5.1, we compare valley polarization as a function of $\Delta E$ at 78 K (green squares) and 300 K (blue circles) and at a few different values of average power (Figure 5-5). The key features are the dramatic decline in valley polarization above 0.1-0.15 eV for both temperatures, with a near zero polarization by $\Delta E=0.4$ eV. Excited carriers with excess kinetic energy, so-called hot carriers, are a noticeable cause for low polarization. This trend was reported with exfoliated WS$_2$ and various MoS$_2$ samples, but never
Figure 5-4: Valley Polarization as a function of excitation photon energy for the sample at 78 K and 300 K. Using input laser excitation with positive helicity (σ⁺), the emitted light is largely the same polarization for pump energies close to the exciton resonance at both temperatures, but falls as the pump is tuned farther from resonance.

in grown WS₂ to our knowledge. [139, 141, 143]

Figure 5-5: Valley polarization versus excess photon energy, defined as \( \Delta E = E_{\text{pump}} - E_{\text{exciton}} \), at 78 K and 300 K for various average power values.
5.3.1 Model for Valley Polarization Dynamics

Clearly, excitation near the exciton resonance can limit the sampling of k-space and produce higher polarization values. Yet, even under these conditions, we see a degree of valley polarization that is no larger than 30% at room temperature or 50% at 78 K. Evidently, valley carriers can depolarize efficiently in WS$_2$, on the same timescale as exciton lifetimes. Two major mechanisms have been proposed: scattering to optically-forbidden states by phonons or defects, and electron-hole (exciton) exchange coupling between valleys due to the large Coulombic interactions. A simple rate model, visualized in Figure 5-6, will aid our discussion in comparing these mechanisms.

![Diagram](image)

Figure 5-6: Model for valley polarization and depolarization dynamics. Valley excitons can recombine or scatter to the opposite valley. The mechanism of intervalley scattering ($\gamma$) is discussed throughout this chapter.

The dynamical model can be used to derive a quantitative equation for valley polarization. The time-dependent population of excitons in either valley, K+ or K−, includes the generation rate to that valley (determined by the optical selection rules) minus the recombination rate for excitons (either radiative or nonradiative) and the intervalley scattering. We assume that the exciton recombination rate is equal in both valleys.

\[
\frac{dN_{K+}(t)}{dt} = g_+ - k_{K+}N_{K+}(t) - \gamma(N_{K+}(t) - N_{K-}(t)) \quad (5.2)
\]

\[
\frac{dN_{K-}(t)}{dt} = g_- - k_{K-}N_{K-}(t) - \gamma(N_{K-}(t) - N_{K+}(t)) \quad (5.3)
\]
The polarization-PL experiment is pseudo-CW, indicating steady-state conditions. Thus, the valley populations are constant with respect to time.

\[
\frac{dN_{K+}(t)}{dt} = 0 \quad \frac{dN_{K-}(t)}{dt} = 0
\]

(5.4)

Using the definition for circular polarization based on the valley populations (Equation 5.1), we can algebraically find the expression as a function of the relevant rates. By inverting the exciton recombination and intervalley rates as the exciton lifetime and intervalley scattering time, the valley polarization can be written by a simple equation.

\[
P_{\text{circ}} = \frac{P_0}{1 + 2 \frac{\tau_{\text{exciton}}}{\tau_{\text{valley}}}}
\]

(5.5)

Here, \(P_0\) is the initially polarized populations, \(\tau_{\text{exciton}}\) is the exciton lifetime, and \(\tau_{\text{valley}}\) is the intervalley scattering time. \(P_0\) is defined as the normalized difference in the valley generation rates. For circularly polarized light with positive helicity we assume \(P_0\) to be 100% and the degree of circular polarization to depend on the relative ratio \(\tau_{\text{exciton}}/\tau_{\text{valley}}\).

\[
P_0 = \frac{g_+ - g_-}{g_+ + g_-}
\]

(5.6)

While high-energy photons may produce initially depolarized excitons \((P_0)\) in the presence of crystal disorder [145], excitation up to the energy of the B exciton creates selectively populated K valleys. [146] Thus, for our range of wavelengths, we can neglect this mechanism and assume \(P_0\) is nearly 100%.

### 5.3.2 Valley Depolarization Mechanisms

Direct intervalley scattering of excitons from large momentum quanta, such as low-wavelength acoustic phonons or atomic-scale defects is possible, but unlikely. Such a transition requires not only a large change in momentum the size of the Brillouin zone, but also a spin flip for both carriers. [34] This mechanism would directly couple
radiative K-valleys, reducing $\tau_{valley}$, but we can expect it to be negligible compared to other relaxation pathways.

Alternatively, phonons and defects may flip the spin of a single carrier, or scatter the exciton to a dark state in k-space, where optical emission is forbidden by selection rules for spin and/or momentum. [147] The result of these transitions would be a lowering of the exciton lifetime, $\tau_{exciton}$, by increasing the nonradiative rate (i.e. $\tau_{exciton}^{-1} = k_{exciton} = k_{rad} + k_{non-rad}$). Looking at Equation 5.5, $P_{circ}$ is expected to increase. This effect has also been proposed as the reason for the large sample-to-sample variation that exists, with an inverse relationship between relative PL intensity and circular polarization. [148] The decline in PL is associated with the same increase in nonradiative recombination. Equivalently, this corresponds to a decline in PL quantum yield, defined as the ratio of emitted photons to absorbed photons and related to the relative rates of exciton decay.

$$PLQY(\lambda_{pump}) = \frac{k_{rad}}{k_{rad} + k_{non-rad}} = \frac{PL}{\alpha(\lambda_{pump})n_{photon}(\lambda_{pump})} \quad (5.7)$$

To test the idea of non-radiative scattering being related to circular polarization, we calculate relative PLQY versus pump energy using relative absorption data ($\alpha(\lambda)$) using differential reflectance from a similarly-grown sample transferred to a transparent substrate (Figure 5-1b). The result is included in Figure 5-7 (red triangles). Even as circular polarization decreases, PLQY is nearly constant from $\Delta E=0.1-0.5$ eV. This indicates nonradiative events are not significantly changing as a function of pump energy in this energy range, as supported by work on WS$_2$. [146] Further, the result supports the proposal that initial depolarization can be neglected, as a similar drop in PLQY would be expected as photon energy increased.

**Exciton Intervalley Exchange Interaction**

Excitons are well-bound in 2D TMDs due to enhanced Coulomb interactions between the electron and hole, a result of confinement, a large carrier effective mass, and reduced dielectric screening. [60] Likewise, we should expect enhanced Coulomb in-
Figure 5-7: Valley polarization compared to PLQY as a function of excess energy ($\Delta E$) above the exciton resonance. The nearly constant PLQY as a function of photon energy, while $P_{circ}$ is decreasing, indicates exciton valley exchange is a dominant depolarization mechanism. The dotted lines act as guides to the eye.

Interactions to affect valley coupling. Based on theory developed by Maialle-Silva-Sham for spin depolarization in quantum wells [149], a similar exchange interaction has been proposed for valley pseudospin depolarization. [150–152] This exciton exchange can be visualized as a virtual photon process between valleys (see Figure 5-8). The result is that a K+ valley exciton can exchange couple to produce a K- exciton. Following the theoretical work of Wu and Yu, the long-range component of this exchange interaction acts as an efficient intervalley scattering process on the timescale of 100's of femtoseconds. [150]

Specifically, this mechanism predicts the intervalley coupling to be dependent on the center-of-mass momentum of the exciton. The larger the momentum, the faster the intervalley rate, i.e. the smaller the circular polarization per Equation 5.5. Pump energy relates directly to the exciton momentum, explaining the observed dependence in Figure 5-7. In addition, exchange coupling occurs efficiently between bright exciton states, thus PLQY would remain constant even as the depolarization rate increases, exactly as we demonstrate.
5.4 Valley Polarization at Fixed $\Delta E$ versus Temperature

It is apparent that hot excitons are detrimental to a large degree of valley polarization at constant temperature. However, temperature itself is often reported as a source of depolarization. One experimental challenge is that as temperature changes, so does the exciton resonance. Using our continuously tunable laser, we are able to decouple the two variables by fixing $\Delta E=0.1$ eV above the exciton resonance as a function of temperature. Said another way, at each temperature, the pump wavelength is chosen to keep $\Delta E$ constant. In doing so, we can show how temperature influences polarization by itself at a nearly constant value of center-of-mass kinetic energy. The results are presented in Figure 5-9. Near low temperatures, polarization is just under 50%, and decreases to around 30% at 275 K. At 300 K, the required pump wavelength starts to leak through our filters and bleeds onto the CCD detector, preventing clean data from being collected. The results clearly show polarization values near 30% can be maintained even at room temperature, which is shown for exfoliated WS$_2$ at similar pump wavelengths. [141, 148, 153] However, we do not observe the rise in polarization values at intermediate temperatures that was observed and attributed
to trion recombination from a three-body Auger process. [141] As we observe neutral exciton emission ($X_0$), we would not anticipate observing the same Auger process.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure5-9}
\caption{Valley polarization compared to relative PL intensity as a function of temperature at a fixed $\Delta E=0.1$ eV. Since PL trends with $P_{\text{circ}}$ instead of oppositely, depolarization is not primarily due to increasing $k_{\text{non-rad}}$, but rather alludes to the Coulombic exchange interaction between valley excitons.}
\end{figure}

Similar to Figure 5-5, the result can also be explained in two different ways. First, excitons that have relaxed from excitation with small center-of-mass momentum can be scattered by phonons to higher momentum states allowing long-range exchange. Second, temperature acts to increase scattering events to dark states, which can actually decrease the exciton lifetime, and therefore PLQY. Just as we did in Figure 5-5, PL intensity can be extracted as an analog to PLQY, as absorption the same $\Delta E$ from resonance should be roughly the same, within a few percent. The result is also displayed in Figure 5-9. The plot shows that circular polarization trends similar to PL intensity, not opposite. This rules out the decrease of exciton lifetime and alludes again to Coulomb exchange as the primary mechanism, although an explanation for any relationship between PL intensity and circular polarization is not understood.

### 5.4.1 Note on Exciton-Exciton Exchange

Finally, we should note that exciton-exciton exchange was recently observed in the femtosecond transient dynamics of CVD-MoS$_2$ and proposed as a mechanism of depo-
larization. [154] We do not, however, observe a power (exciton density) dependence to the valley polarization (Figure 5-5) perhaps because the femtosecond transient does not contribute appreciably to the overall photoluminescence. Alternatively, it is possible that exciton-exciton exchange is active at much higher exciton densities than studied here. We note that in previous time-resolved PL measurements, including ones we have done on the same CVT material, a strong nonlinearity is observed in the time-resolved PL, often attributed with exciton-exciton annihilation (see Section 4.6). We therefore know exciton-exciton effects are important at these fluence levels, but steady-state measurements cannot discern this effect on valley polarization.

5.5 Summary of Valley Polarization

In this chapter, a valley polarization is shown near 50% at 78 K in monolayer WS$_2$ grown by chemical vapor transport. By separately characterizing polarization as a function of pump energy and temperature, we provide evidence for exciton exchange interactions being the primary source of valley depolarization. Our results further show that temperature has an additional effect on intervalley coupling that cannot be explained by thermal energy alone, likely due to momentum provided by phonons. These findings provide insight into valley polarization in monolayer TMDs, which is critical to the design and production of low-power valleytronics devices.
Chapter 6

Valleytronics: Applications of the Valley Pseudospin

In Chapter 1, the valley degree of freedom was proposed as a 'Beyond CMOS' technology for lower power switching. In order to achieve this goal, a material with binary valleys with linked intrinsic properties was required to access these valleys externally. In Chapter 2, I showed that monolayer TMDs exhibit a valley index linked to carrier spin and the helicity of optical transitions. With motivation to explore the valley properties of this material, a monolayer growth technique was presented in Chapter 3, along with basic material characterization in Chapter 4, demonstrating the production of high-quality samples. The next step in developing any valley-based logic is to understand how well valleys remain polarized once selectively populated, as was done in Chapter 5.

With a better understanding of the depolarization mechanisms and limitations of the material, it is time to propose the next set of experiments to actively control this valley index. In this chapter, I propose a platform based on the valley Zeeman effect to actively change the lowest energy valley. Using the magnetic exchange bias from a ferromagnetic insulator, the first experiment is to test which materials are compatible. Assuming valley splitting can be observed, I propose two major device platforms that should be explored. One is an electro-optic modulator, based on the resulting difference in refractive index between left- and right-handed circularly
polarized light. The second is a valley-spin FET whereby a valley current is injected into a 2D channel.

![Optical selection rules for valley and spin in monolayer TMDs](image)

Figure 6-1: Optical selection rules for valley and spin in monolayer TMDs, presented as a summary of Chapter 2.

### 6.1 Valley Splitting on Magnetic Substrates

The valley Zeeman effect was discussed in Chapter 2 (Section 2.4). Due to the asymmetry in the orbital magnetic moment between the conduction and valence bands, and also the flipping of sign between different valleys, the energy degeneracy is broken between valleys. The implications are numerous, stemming from the fact that the lower energy valley should populate more readily than the one with higher energy. Both electronic and optical applications arise, and the need for a circularly polarized input is also loosened. There are two caveats though. First, the generation and usage of the required external magnetic fields is not easy in integrated devices, whereas electric fields are easily produced. Second, the Zeeman effect is weak (2 meV T⁻¹), requiring superconducting magnets with field strengths in the 10s of Tesla in order to produce the magnitude of splitting needed for most applications.

Together, the valley Zeeman effect seems doomed for applications on-chip. However, large effective magnetic fields can be created through proximity exchange bias.
with a ferromagnetic substrate. This exchange coupling is a short-range effective field, requiring atomic-scale proximity to the surface of a magnet. Thankfully, monolayer TMDs are inherently on this length-scale. Using ferromagnetic insulating substrates, on-chip devices may be possible that use the unique valley-spin physics of TMDs.

Prior work confirms that ferromagnetic insulators can exchange couple to 2D materials to produce physics that is equivalent to large external magnetic fields. Relevant to TMDs is recent work showing that WSe$_2$ on the ferro-insulator europium sulfide (EuS) shows an enhanced valley splitting. [155] Comparing to a sample on a non-magnetic substrate, the effective field is calculated to be approximately 10 Tesla (Figure 6-2). In addition, chromium iodide (CrI$_3$) was recently used as a ferromagnetic insulating substrate with WSe$_2$, and shown to cause valley splitting at low temperatures. [156] This prior work has demonstrated the necessary physics of magnetic exchange bias as an effective way to induce the valley Zeeman effect without large bias fields.

![Figure 6-2: Enhanced valley splitting in WSe$_2$ on an EuS substrate, showing an effective coupling of around 10 Tesla. Figures courtesy of [155].](image)

The primary goals going forward are to improve on the prior work by demonstrating valley energy splitting with magnetic insulators having a Curie temperatures larger than 300 K for room temperature switching. Moreover, using out-of-plane anisotropic magnetic substrates instead of in-plane, non-volatile switching behavior is desired to avoid external bias fields as will be needed for integrated devices. The non-volatility is also desired as it promises lower threshold switching. From the concept in Figure
6.3. the initial goal is to show valley splitting on a ferromagnetic insulator at room temperature.

\[ \Delta E \]

Figure 6-3: General concept for the initial measurement to show valley energy splitting using a ferromagnetic insulating substrate. The ideal result is a valley energy splitting that closely follows the magnetization of the underlying substrate.

### 6.1.1 Proposal to Use Rare Earth Iron Garnets

While ferromagnetic metals are common and plentiful, applications would require ferromagnetic insulators to avoid quenching of valley excitons. Of the known ferromagnetic insulators, most have a Curie temperature below 100 K and/or are in-plane magnetically anisotropic (IMA). In order to break the valley degeneracy, magnetic fields need to be applied perpendicular to the plane of the 2D sample. While external fields can rotate the magnetization of an IMA magnet, useful devices would operate without an external bias. Thus, the need is for an out-of-plane (perpendicular) anisotropic (PMA) ferromagnetic insulator. Finding an insulator with a Curie temperature above 300 K would permit room temperature operation. Finally, our wish list includes a soft ferromagnet with low switching fields (low coercivity). This allows minimal switching energy and makes for easier experimentation.

Recent work on rare earth garnets has provided a few possible materials that fit our requirements. Specifically, thulium iron garnet (TmIG, Tm\(_2\)Fe\(_5\)O\(_{12}\)) is an ferromagnetic insulator that exhibits a perpendicular anisotropy with low switching fields at room temperature. [157, 158] As such, my proposal is to use these rare earth iron garnets as the ferromagnetic substrate to exchange couple to TMD monolayers.
Figure 6-4: Proposal for initial experiment using thulium iron garnet (TmIG) as a ferromagnetic insulator with CVT-grown WS$_2$. Using external magnetic fields to switch the magnetization of the ferro-insulator, absorption measurements are taken to detect the change in valley energies. Future devices can switch the magnet electrically via spin-transfer torque or similar.

### 6.1.2 Measurement Techniques

To test the validity of the proposed ferro-insulator/TMD platform, the first experiment is to confirm exchange coupling by measuring valley splitting. While the valley Zeeman studies used polarization-dependent PL, similar to Chapter 5, we chose to use differential reflectivity to probe the absorption of the valley states. Due to valley depolarization, absorption is a better probe of the valley energetics.

In addition, initial measurements of PL on the iron garnet substrates showed weak signal (Figure 6-5). The suspected reasons are three-fold: 1) Optically, the substrates are transparent compared to SiO$_2$/Si, meaning no reflection occurs in the incident or emitted paths, 2) Damage occurs during flake transfer, and 3) the substrate contributes to quenching due to the presence of the heavy metal atoms. By taking PL before transfer, after transfer, and after heat annealing, the dim PL seems to mostly come from substrate quenching as the heat annealing step resulted in $\sim$100x decrease in intensity. The transfer step also leads to degradation of signal, but the effect is around 3x.
Building a Magneto-optical Cryostat Setup

Performing the desired experiment requires a setup that can apply a magnetic field perpendicular to the sample (Faraday geometry) while exciting the sample with a laser or incoherent source. Given the microscopic size of the 2D flakes, a microspectroscopy setup is needed, meaning the sample has to be mounted very close to an optical window to enable the use of a high numerical aperture objective lens. Finally, temperature control is ideal, given the strong dependence of the valley polarization and magnetization with temperature.

A diagram of the setup for differential reflectivity is shown below (Figure 6-6), but the same platform is flexible enough to perform polarization-dependent photoluminescence (see Chapter 5) and also magneto-optical Kerr effect measurements to confirm the performance of the magnetic substrate in-situ. Details of both measurements are found below.

The base platform for the entire setup was an inverted microscope (Nikon Eclipse Ti-U) with a third-party XYZ stage (Ludl). On top, a custom-fabricated mount was made for a Janis ST-500 continuous-flow, open-cycle cryostat. The cryostat can achieve temperatures down to 4 K with liquid Helium (LHe), but we only performed initial characterization to 77 K with liquid nitrogen (LN2) due to our goals of high-temperature switching.
Figure 6-6: Magneto-optical Cryostat Setup. The sample is mounted in a continuous-flow cryostat that is surrounded by an electromagnet. The optical path allows precise control of the polarization of light to ensure selective polarization and detection of valley energetics.

Figure 6-7: Photographs of completed magneto-optical cryostat setup. (a) The cryostat is mounted on an inverted Nikon microscope with a custom-built stage. (b) A hand-wound solenoid surrounds the objective lens and cryostat cold finger/window, with the ability to apply up to 0.1 Tesla of magnetic field under pulsed conditions.

**Differential Reflectivity - Absorption**

As previously discussed, absorption is preferred over PL for a few reasons. The simplest way to measure absorption on a transparent (non-absorbing) substrate is to measure the differential reflectivity. The basic premise is to measure reflection on the sample with the substrate underneath, and then compare that to the reflection of the
substrate by itself. The difference normalized to the substrate signal is proportional to the absorption when the absorbing sample is thin. [159, 160]

\[
\frac{\Delta R}{R} = \frac{R_{\text{flake}}(\lambda) - R_{\text{sub}}(\lambda)}{R_{\text{sub}}(\lambda)} = \frac{4n_{TM}(\lambda)}{n_{sub}(\lambda)^2 - 1} \alpha(\lambda)
\]

(6.1)

To measure differential reflectivity, the excitation source is normally a broadband light source such as a white LED or tungsten-hallide lamp. The key to low noise is choosing a stable light source such that the spectrum changes little during the span of multiple measurements. The light needs to be carefully calibrated to be circularly polarized at the sample after passing through many optical components. To do so, a Barek’s compensator is used as a variable wave retarder on a rotation mount. The polarization is measured before mounting the sample using a rotating linear polarizer in front of a photodetector. When circularly polarized, the difference between the maximum and minimum intensities is less than 1%, indicating a polarization \( \geq 99\% \). One complication exists due to broadband nature of the source. Since wave retardation is inherently wavelength-dependent, a bandpass filter centered on the wavelength of the exciton resonance is used during polarization calibration. This ensures circularly polarized light near the wavelengths of concern.

From the diagram in Figure 6-6, the calibrated light is focused by the objective lens onto the sample and reflected back. The reflected light is collimated by the objective and eventually dispersed by a monochromator and imaged with a Si:CCD (Princeton Instruments). The grating in the spectrometer is chosen to be 1200 lines/mm for the final measurements, equaling a pixel resolution of about 0.13 nm/pixel when using a slit width of 50 \( \mu m \). A 1 meV shift in energy at 620 nm corresponds to approximately 0.3 nm so there should be resolution in the hardware.

Typical raw reflection data is shown in Figure 6-8a, one line for the sample/substrate and one for the substrate alone. In Figure 6-8b, the differential reflection is taken (Equation 6.1), clearly showing the A and B exciton peaks. This data is taken with a 300 lines/mm grating to see both peaks, as a 1200 lines/mm grating only allows a window of about 50 nm.
Magneto-optical Kerr Effect (MOKE)

To compare the switching fields for the magnet versus the valley splitting, a way to measure the (relative) magnetization of the substrate is desired. One such way is using the magneto-optical Kerr effect (MOKE). A MOKE setup is similar to measuring Faraday rotation except in reflection mode. Based on the magnetization of the substrate, linearly polarized light will undergo a small rotation due to the chirality of the material. This rotation is opposite when the magnetization is flipped. Thus, you can use a linear analyzer to completely block the signal when the magnet is in one state. This would mean the polarizer is perpendicular to the reflected polarization (after rotation). When the magnetization is flipped by the external bias, the rotation is opposite such that the total change is $2\theta$. This small change is enough to be detected by an increase in the intensity (Malus' Law) through the analyzer.

While the concept is simple, signal-to-noise is an issue. The rotation angles are often small ($< 0.05$ degrees) meaning small changes in intensity ($\cos^2 \theta \approx \theta^2$). To improve signal-to-noise, the incident laser is chopped and the output is measured with a lock-in amplifier. Next, the size of the photodetector must be carefully chosen to reduce noise and parasitics. A large photodetector has more noise. In the end, I went with a Thorlabs PDA36 Si detector for simplicity and low noise equivalent power (NEP). Compared to the PDA100 model, the noise was reduced by about an order of magnitude. Finally, the laser drift becomes a consideration for weakly magnetic material. The solution is a double lock-in setup to track the drift separate from the
photocurrent. The ratio has much less noise, dropping the noise by another 5-10x.

A proof-of-concept measurement was taken on a cobalt terbium sample (CoTb) having perpendicular anisotropy (PMA) and a coercivity of around 400 Oersted (measured externally with a VSM). The MOKE signal in Figure 6-9 clearly shows a switching in intensity between 300 and 400 Oe. Unfortunately, similar measurements on TmIG failed due to lingering issues with noise.

![Figure 6-9: MOKE signal on CoTb, a PMA ferromagnetic material, clearly showing switching between 300 and 400 Oe.](image)

### 6.1.3 Preliminary Results and Discussion

Samples were fabricated by transferring WS$_2$ via HF/PMMA method (Appendix A) to the TmIG (10 nm) grown on gadolinium gallium garnet (GGG). After transfer, samples are heat annealed in a low vacuum chamber ($< 10^{-6}$ Torr) at 350 °C for 1 hour to try and improve substrate-flake bonding. Samples were then loaded into the crystat and pumped down to less than $10^{-4}$ Torr. Differential reflectivity was then acquired on multiple samples from 77 K to 300 K. At each temperature, magnetic fields were swept back and forth (double sweep) using a LabVIEW program created specifically for the task. The sweep parameters changed with temperature based on the magnetic properties of the substrate. For room temperature measurements, the coercivity is less than 100 Oe so measurements are acquired with bias on to avoid accidental switching near 0 Oe. At low temperatures, the fields were swept to the set point and back to zero before measuring, in order to avoid Faraday effects in the objective lens.
Data was analyzed using Python and MATLAB. The differential reflectivity curves at each magnetic field were fit to a Lorentzian function. The peak of the fit was extracted and plotted versus the applied field. Typical results for monolayer WS$_2$ on TmIG are shown in Figure 6-10, indicating no valley splitting was observed, or at least no splitting larger than the noise of the system, which is around 0.25 meV. In an attempt to see splitting, samples of WSe$_2$ and exfoliated WS$_2$ were used, as well as attempting various parameters for annealing.

![Figure 6-10: The extracted peak of differential reflection versus magnetic field for RHCP and LHCP incident light at 300 K. The sample is exfoliated WS$_2$ transferred in an argon glovebox to a TmIG (10 nm)/GGG substrate. The results indicate no valley splitting is detected as peak-to-peak changes are less than 0.6 meV.](image)

The reason for this result are not understood. Initially, the belief was that interactions with the substrate were poor due to lingering water molecules and other matter that is difficult to remove. Annealing at ultra-low vacuum at elevated temperatures did not improve results. In the end, a dry transfer was conducted with polycarbonate (Appendix A) inside an argon glovebox to produce the cleanest interface. The results were the same: no detectable splitting above the noise. The remaining explanations include the structure of the iron garnet having weak exchange coupling due to excess oxygen on the interface or physics which does not match the current understanding. To be complete, TmIG has a low magnetization and a complicated structure. More experiments are needed, but based on prior work, some combination of material and processing will get this to work. Recent work using MoS$_2$ on a related yttrium iron garnet (YIG) substrate indicates some exchange coupling, so hope is not lost. [161]
6.2 Valley Optoelectronics

The ability to excite select valleys via circularly polarized light is an exciting way to study valley-based physics. However, the opposite is just as intriguing for optoelectronics devices. If electrons and holes of opposite spin recombine in the same valley, the result is a circularly polarized photon matching the selection rules established in Chapter 2. A few light-emitting devices have been reported which demonstrate this fact. [121, 162–164] Two of the reports use spin-injecting contacts/substrates to control the spin of the hole [163, 164]. The other two claim in-plane electric fields can change the shape of the valley overlap in k-space between electrons and holes, thereby modulating the valleys where emission originates. [121, 162] Interestingly, the electroluminescence polarization reported for the spin-polarized devices is less than 20%, versus larger than 50% for the other two. Since the number of publications is low, it is hard to draw conclusions, but the trend would say that spin-injection is less effective than electrical tuning. The paper from Yang et al. is especially exciting as they report a polarization value of 80% at low temperatures using a PN heterostructure.

The use of these circularly polarized emitters is unclear, but applications must exist. For instance, rather than using horizontal and vertical polarizers for 3D TVs, a better option is LHCP and RHCP orthogonal polarizations as they are not as sensitive to the angle of viewing. This could also be a new form of channel multiplexing for communication based on the polarization of light. Assuming circular polarized lasers can be made, they will be useful in polarization spectroscopy. This of course assumes a high degree of circular polarization can be established.

6.3 Valley-Spin FETs

Based on the idea of spin FETs, I propose a similar device that adds the valley degree of freedom as a means of maintaining spin coherence. Given the large spin-orbit splitting and large momentum shift required to flip hole spin, the device should work
best as a p-channel device (Figure 6-11). Using spin-injecting contacts, one of the valleys can be selectively populated. [164] Assuming the polarization of the TmIG is the same in the channel, hole current should flow normally. However, when the TmIG insulator is flipped, the valence band shifts and holes are no longer able to flow, resulting in low current at the output. The gate is a heavy metal, able to impart spin-torque to switch the polarization of the iron garnet film.

One implication is that the drain contact may not need to be a ferromagnet, as spin filtering is done in the channel based on valley energetics, whereas it is necessary in a spin FET. This should allow improved on/off ratios as spin FETs act as magnetoresistive devices, whereas the valley-spin FET is a current modulating device. The key to good on/off ratios is good filtering which requires a large valley Zeeman splitting, at least greater than $k_B T$, which has currently not been observed.

Figure 6-11: Concept for a valley-spin PFET. Spin-injection source populates one valley with holes depending on the magnetization direction. This hole current can then be modulated by the exchange coupled ferromagnetic insulator, which is in turn controlled by spin-transfer torque from a heavy metal gate. When the magnetizations are all aligned and equal, hole current passes. Oppositely, holes are blocked when the magnetization of the TmIG is flipped.
A caveat about the proposal. While some of the ideas have been verified experimentally, such as spin-torque switching of a magnet [158], valley splitting due to exchange bias [155], and spin injection to a TMD [164], these components require integration, with a pathway that is unclear. Though challenges exist, the novelty and potential improvements over spin FETs requires the idea to have some consideration.

6.4 Valley Optical Modulators

Although integrated photonics is quickly improving, the basic switching component is still fraught with problems. Silicon electro-optical modulators can be made small and low-energy, but they must use sensitive resonant structures in order to do so. This limits their bandwidth and makes them very sensitive to small temperature fluctuations. The inherent problem is that silicon does not have the ability to modulate its refractive index with enough magnitude. The current state-of-the-art modulators make use of the fact that silicon’s refractive index changes when carrier density changes (Franz-Keldysh effect). [165] However, even using depletion-mode PN junctions the change in refractive index is only on the order of $\Delta n = 10^{-4}$. This puts a burden on fabrication as high-Q resonators are required.

Based on the Kramers-Kronig relation, a change in peak absorption results in a change in refractive index, with the largest slope in refractive index occurring where absorption peaks. Since TMDs are excitonic, they have sharp absorption peaks, rather than a gradual band edge. Using a similar scheme as Figure 6-11, the absorption peak can be swapped for RH vs. LH circularly polarized light, i.e. the refractive index is shifted. This difference in refractive indices between orthogonal circularly polarizations causes monolayer TMDs to act as a chiral material. If a resonant structure is built using monolayer TMDs as a waveguide cladding on at least one side, the evanescent wave can couple with the TMD and the refractive index differences would change the resonant frequency of the device. (Figure 6-13)

Any polarization of light can be deconstructed into a linear combination of RHCP and LHCP components. Thus, an electro-optical modulator can be built that is res-
Figure 6-12: State-of-the-art silicon electro-optical modulator (EOM) based on depletion mode physics at a PN junction. Figure courtesy of [165].

Onant to an input light frequency using one polarization’s refractive index, say $n_{HI}$, while being out-of-resonance for the opposite polarization’s refractive index $n_{LO}$. By modulating the refractive index between $n_{HI}$ and $n_{LO}$ via the underlying magnetization of the substrate, you can switch between letting a chosen circular polarization pass or be filtered. Thus, a single resonator acts as a circular polarizer. Putting two in series and operating them independently, or setting the first to a fixed value, allows the system to act as an intensity switch.

A quick approximation is made using the Kramers-Kronig relation for a shift in energy of only $\Delta E \approx 1 \text{ meV}$, reasonable based on prior results showing a 10 Tesla exchange field. Using 10% absorption with a linewidth of 50 meV, and a distance of 10 nm below resonance, the calculated $\Delta n$ is on the order of $10^{-4}$, similar to depletion-mode silicon. Thus, larger $\Delta E$ is required for larger $\Delta n$. To increase $\Delta n$ by an order of magnitude, $\Delta E$ needs to be around 5 meV. These calculations make quite a few assumptions, but orders of magnitude are useful for comparison.

Big open questions remain to be answered and explored:

1. Can the device be built to operate as desired? Some of the fabrication between
Figure 6-13: Concept of valley based electro-optical modulator. Right- and left-handed circularly polarized light has different resonances that can be swapped by changing the polarization of the exchange coupled TmIG layer. When cascaded with devices in the opposite configuration, the first device creates circularly polarized light from unpolarized light and the second acts as an intensity modulator.

layers is untested and unclear. Individual pieces are worked out, but putting it all together is rarely a trivial task.

2. What is the coupling coefficient for such a thin cladding on top of a waveguide? How large of a Q-factor is needed? This number is crucial to justify pursuing this technology.

3. What wavelength would this device operate at? The exciton resonance is in the orange-red range for TMDs, which is far from the 1550 nm or 1350 nm wavelengths used in communications.

4. You cannot operate directly on resonance or self-absorption will attenuate the signal too much. Can you operate off-resonance far enough to avoid attenuation, but still close enough for large Δn? Perhaps an EOM based on absorption would work just as well.
6.5 Prediction on the Future of Moore’s Law

As someone who has spent almost 10 years paying close attention to the semiconductor industry and related academic research, I feel that Moore’s Law is finally within its final decade. My belief is that Intel and others will figure out how to make CMOS work down to the 5 nm node, but that the true benefits of this effective scaling will not be reached due to limitations with reliability. In addition, each node will take longer to get up to a point where top end CPUs are able to take full advantage of each improvement. The entire industry looks to be reworking the map itself and moving the goalposts as the game is being played. While I predict a version of Moore’s Law will continue past 2030, the typical improvement in MOSFET fabrication every 2 years is basically dead. Photonics continues to improve, but the connections between photons and electrons is still tricky without wasting too much time and energy. This is where spin and pseudo-spin systems may find their niche, by bridging photonics and electronics without the energy cost. In addition, quantum computing is quickly improving and may start replacing traditional CPUs in roughly 10 years for applications that best utilize their probabilistic nature.

6.6 Conclusion

Monolayer TMDs have a set of unique properties which makes them intriguing for a variety of applications, from next-generation logic, to integrated photonics, and even quantum computing. Yet, many of these applications are full of competition and TMDs have their issues in regards to material quality and novelty. In order to be commercialized as a product, there must be a niche application that cannot be solved with current technology, or is too expensive to do so. As far as charge transport and basic optical (emission and absorption) are concerned, there is not enough to likely spark commercialization versus other nanomaterials. I believe the valley degree of freedom is the unique property that will lead to the commercialization of TMD-based technology. In what capacity is yet to be determined, as research is just getting to
the point of creating novel devices, but in this chapter I’ve proposed a few options that take advantage of this unique valley physics.
Appendix A

Fabrication Techniques for 2D Materials

A.1 Micro-transfer of 2D Samples

The transfer of 2D materials from their growth substrate to the final device is a critical step in the fabrication process. As such, there has been quite a bit of work on developing methods which allow easy transfer with no wrinkles, cracks, or other flake damage. In addition, some processes are not compatible and different transfer methods must be used depending on the needed application. For instance, for devices sensitive to solvents, dry transfer techniques are used.

In general, the transfer techniques can be broken into three categories: (1) Liftoff by etching the underlying substrate, namely, silicon dioxide, (2) Liftoff using water that separates hydrophobic 2D flakes from hydrophilic silicon oxide, and (3) Liftoff with Polymers using thermoelastic properties. Depending on how the flake is removed from the growth substrate, determines which method is used to finally place it on the device. I have explored a plethora of techniques to establish those that work best for different needs required by the different devices and projects. A few are compared here that will be useful for the final devices that we intend to fabricate.
A.1.1 PMMA Transfer with Hydrofluoric Acid

The poly(methyl methacrylate) (PMMA) method has become the standard amongst the 2D community. There are variations, but the main idea is submerging a substrate containing 2D flakes in a liquid etchant that removes the substrate material, leaving behind a PMMA layer with 2D crystals attached. For my WS2 samples grown on SiO2/Si, the process is outlined in Figure 4.11. First, PMMA is spun onto the oxide surface and baked at 180 °C for 2 minutes. The sample is dipped in diluted hydrofluoric acid (20% in DI water) to etch the oxide layer. Once this occurs, the PMMA layer releases, producing a clean transfer onto the PMMA that can be transferred to another substrate.

HF is chosen instead of the popular hot potassium hydroxide (KOH) bubbling method for three main reasons. First, the KOH method takes many hours to days, whereas HF etching occurs within a few minutes. Second, hot KOH produces bubbles that damage and wrinkle the 2D flakes. Third, KOH leaves a residue on the PMMA and final substrate that is difficult to remove. HF leaves little to no residue and evaporates completely without damage to the flakes.

An improvement to the typical PMMA transfer technique is used so that we can later transfer the flakes under a microscope to device structure with micrometer precision. By putting a PDMS stamp in contact with the PMMA film, we end up with a TMD/PMMA/PDMS stack that can be put on a glass slide and used with our transfer setup (Figure 4.13). Since the PDMS has a tendency to stick to the substrate and prevent further etching, a two-step HF dip was invented, whereby the first dip does most of the etching under the PMMA, and the second dip finishes the etch after PDMS is added on top of PMMA. In the end, the 2D flake is transferred to the final substrate by bringing into contact and melting the PMMA using a heated stage (120 °C). The PDMS stamp is peeled away, and the PMMA can be dissolved in acetone to complete the transfer process.
A second process makes use of the stickiness of PDMS, and the hydrophobicity of TMDs to pick up grown flakes. Given the substrate oxide is hydrophilic, adding a drop of water can induce a capillary force that separates the 2D samples from the growth substrate. While other groups have reported similar methods using polystyrene (PS) or poly l-lactic acid (PLLA), I developed a method using commercially available PDMS from Gel-pak that has a proprietary sticky coating. The process uses a small piece of Gel-pak X4 on a microscope slide with our transfer setup. A drop of deionized water is added to the PDMS stack, which is upside down under the microscope objective. The PDMS with water is slowly engaged to the surface to ensure no wrinkles. The sample stage is then heated to around 90 Â°C to promote the delamination of the 2D layers from the surface without generating a vicious boil. Last, the PDMS stack is slowly raised. If done carefully, the water creeps under the 2D layers, separating them from the growth substrate and leaving them on a semi-rigid PDMS stamp for further processing. The advantage of this method is the lack of solvent exposure compared to the other techniques and the ability to pick up only select flakes from the growth surface, in comparison to the whole substrate with the PMMA etching method. The downside is the yield. Very often this method fails, and is difficult to reproduce. Even successful liftoff often results in torn or holey flakes. Finally, PDMS transfer to the final substrate is done by taking advantage of
its viscoelastic properties. When pulled quickly, it tends to stick, but when pulled slowly, it releases. This method is my preference for clean transfer of exfoliate flakes when I have the option to exfoliate on PDMS directly, but the frequent damage to my CVT-grown flakes is a disqualifier for device work.

### A.1.3 Polycarbonate Pickup

A method commonly used for exfoliated flakes is known as polycarbonate (PC) pickup. Using the thermoelastic properties of PC, one can grab 2D flakes that are not very well held to the substrate and be physically lifted. The method begins with a stamp created on a microscope slide consisting of a thin layer of PC stretched over a piece of PDMS. The PC/PDMS stack is brought nearly in contact with a desired flake at room temperature and heated to 110 °C. The PC expands when heated, engaging the desired flake. As it is heated further, it surrounds the flake. After 5 minutes at 110 °C, the heater is turned off and the stage cools. The PC film shrinks and creates a hold on the flake. As the stage cools further, the film slowly disengages the surface, lifting the 2D flake. The PC/PDMS stack can now act as a stamp to transfer to a donor substrate, where the stage is heated to 150 °C to fully melt the PC. Chloroform is used to remove the PC, completing the transfer.

While impressive and useful, it still requires a final solvent bath to remove the PC, similar to the PMMA method. It also seems limited to exfoliated flakes. Numerous attempts to try this method with grown flakes has failed, presumably because the flakes are more strongly bound to the surface.

Figure A-2: Direct pickup transfer process using polycarbonate (PC) films.
A.1.4 Transfer Setup

To accomplish selective transfer of 2D flakes, I fabricated a platform based on a probe station microscope with long working distance objectives. A picture of the completed equipment is given in Figure 4. Beneath the probe station, a heated stage with vacuum chuck allows substrates to be held in place. This stage can be lifted in the z-direction, moved in x- and y-directions, and rotated in the xy plane. Above the stage is a cantilevered microscope glass holder that is attached to a 3-axis manipulator. The z-axis is connected to a motorized actuator for fine control when engaging and disengaging, also preventing vibrations due to manipulation of manual actuators. The stage can be heated by either a PID loop controller or, more carefully, using a high-voltage DC power supply. The various degrees of freedom and flexibility of the platform allow different transfer techniques to be used, all with micrometer accuracy. Custom software was written in Python to control most of the setup’s features.

![Figure A-3: Photographs of the home-built transfer setup used to move 2D materials between various substrates. The close up image on the right shows a heated XYZ stage, with vacuum chuck, and a cantilevered glass slide on a XYZ manipulator.](image-url)

A.2 Electron Beam Lithography and Metal Liftoff

One of the most repeated steps in device fabrication is the creation of metal layers for gates, drain/source contacts, and grid markers. Photolithography would normally be used, for instance, in a CMOS process, but in our devices, flake dimensions and
locations are not determinate. Thus, using a fixed mask as with a standard photolithography process is not possible. Electron beam lithography machines, such as the Elionix ELS-F125, allow CAD files to be used for completely custom designs with each sample or device. In addition, PMMA is used as the ebeam resist, which is already part of our normal transfer process, so we can avoid any further contamination removing PMMA residue.

A typical e-beam lithography process with metal liftoff is outlined in Figure A-4. Based on a pattern designed in CAD software, the spun on PMMA is exposed to a dose of electrons which break the polymer, making it susceptible to development with 'gentle' solvents, such as a 3:1 ratio of isopropyl alcohol (IPA) to methyl iso-butyl ketone (MIBK). After development, the substrate is exposed in areas we wish to fill with metal. This is accomplished by evaporation of the chosen metal to the desired thickness. After evaporation, the entire surface is covered in metal, but the majority has PMMA resist underneath. By dipping in acetone for a few hours, usually at slightly elevated temperatures (e.g. 40-50 °C), the solvent undercuts the pattern in the openings/edges and removes the PMMA, lifting the metal away from the surface. The remaining metal is left only where we originally patterned with electrons.

![Figure A-4: Scheme of electron beam lithography followed by metal evaporation and liftoff. This process is used repeatedly in device fabrication.](image)
Bibliography


134


