Title

The Condensation of Electrons, Holes and Data: Imaging Non-Equilibrium Dynamics at the Microscopic, Mesoscopic, and Statistical Scales

Author

Arp, Trevor Benjamin

Publication Date

2020

UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Condensation of Electrons, Holes and Data: Imaging Non-Equilibrium Dynamics at the Microscopic, Mesoscopic, and Statistical Scales

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Trevor B. Arp

June 2020

Dissertation Committee:
    Dr. Nathaniel Gabor, Chairperson
    Dr. Vivek Aji
    Dr. Yongtao Cui

The Dissertation of Trevor B. Arp is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

Behind every PhD is a body of work that took years and the combined efforts of many people, both those contributing directly to the research and those supporting the graduate student. Getting a PhD is not an easy process, it requires countless hours of hard work, enormous mental discipline, lots of frustration, disappointment and self-doubt. Becoming a scientist is a hard road but I have been fortunate to have been helped along it by many wonderful people. They have helped me, motivated me and supported me when I needed it, and it is because of them that the trails of graduate school have not sapped my love of science. Even though getting through graduate school has been the hardest thing I have ever done, I firmly believe it has been a worthwhile use of several years of my life.

First, I must thank my advisor, Nathan Gabor, for all his effort teaching me, mentoring me, motivating me and being patient with me when I needed it. It is easy for a professor to be a mediocre advisor; all they have to do is not make any special effort to help their students. But Nathan was never this way. I felt that he invested in me, advocated for me, and wanted to help me succeed. Scientifically, he has shown me how to perform research in a creative and versatile way, he has listened to my ideas and helped guide me in expanding them rather than forcing me to go a particular way, and we have both benefited from it. Ultimately, he modeled the scientist that I want to be; flexible, collaborative, inventive and enthusiastic.

I thank Bahram Mobasher and the Fellowships and Internships in Extremely Large Datasets (FIELDS) program, for giving me a fellowship that broadened my

horizons greatly. I must also thank Mario De Leo Winkler for supporting me in that program and teaching me about outreach.

In the lab I was joined by an excellent team of talented people. Thank you to my lab mates Dennis Pleskot, Fatemeh Barati, Max Grossnickle, Jacky Wan, Jed Kistner-Morris, Dave Mayes, Farima Farahmand, and Esat Kilinc for the many ways, large and small, that you helped me. Ultimately, I would have gotten nowhere without samples fabricated by some very talented hands. I must thank Dennis Pleskot, who mastered the notoriously difficult material $MoTe_2$ and made the devices in which we discovered the electron-hole liquid. I must thank Jacky Wan for his artistry in making extremely high-quality graphene-hBN heterostructures and for working with me on understanding that subtle and difficult system. I must thank Fatemeh Barati for letting me play with her $MoSe_2$-WSe_2 heterostructures, working with me on the electron-phonon coupling in those heterostructures. Beyond samples, Max Grossnickle was my companion in MSE 179, helping me with many things and using an artistic talent for scientific visualization to make beautiful renders of some of my work and teach me about 3D visualization.

I could not have survived graduate school without family and friends to support me. Thank you to my mother, Gwen Fraley for always supporting me and for flying me up to Seattle to visit family and friends, to stay connected and to go on relaxing trips, which did much to prevent burnout. Thanks to my sister Julia for her constant support. Thanks to my friends from my cohort Nick Pollard, Jeremiah van Baren, and Christina Manzano-King, who went through this with me. Thanks to the many friends who helped keep my spirits up and occasionally had to drag me away from my work to socialize;

James, Stephen, Jed, David, Peter, John, Ella, Craig, Sheri, Brian, and Bobby. I would also like to thank Kurt, Eryn and Ayoka for helping me when I really needed it.

Finally, this work would simply not be possible without support from the United States taxpayer. Public investments in science are important, not just for their direct value producing knowledge and new technology, but also for providing education for graduate students like me. That education expands horizons, teaches skills, widens perspectives, gives experience, and creates opportunities that multiply an individual's ability to contribute to society. I am cognizant that hundreds of thousands of dollars have been put into my education and strive to make it a good investment.

Several chapters in this dissertation are based on my previously published work, and material from publications may be divided between chapters, modified and mixed with unpublished work, as appropriate to fit the format and narrative of this dissertation. Chapters 1, 2 and 3 are in part reproduced from T. B. Arp and N. M. Gabor "Multiple parameter dynamic photoresponse microscopy for data-intensive optoelectronic measurements of van der Waals heterostructures" *Rev. Sci. Instrum*. **90**, 023702 (2019). Chapters 4 and 5 are in part reproduced from T. B. Arp, D. Pleskot, V. Aji, N. M. Gabor "Electron–hole liquid in a van der Waals heterostructure photocell at room temperature" *Nature Photonics* **13**, 245–250 (2019). Chapter 6 is in part reproduced from T. B. Arp, J. Kistner-Morris, V. Aji, R. Cogdell, R. van Grondelle, N. M. Gabor "Quieting a noisy antenna reproduces photosynthetic light harvesting spectra" in press at *Science* as of the writing of this dissertation (arxiv:1912.12281). Furthermore, I must acknowledge Shanshan Su and Roger Lake for performing the DFT calculations discussed in chapter 4.

To all my teachers, who inspired my love of learning. This thesis is the culmination of their efforts and patience.

ABSTRACT OF THE DISSERTATION

The Condensation of Electrons, Holes and Data: Imaging Non-Equilibrium Dynamics at the Microscopic, Mesoscopic, and Statistical Scales

by

Trevor B. Arp

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, June 2020
Dr. Nathaniel Gabor, Chairperson

Physical systems driven out of thermal equilibrium may exhibit complex behavior before eventually relaxing to thermal equilibrium. The prototypical non-equilibrium state in quantum materials is the photoexcited electron, where an electron is imparted excess energy and may exhibit a variety of interesting phenomena before eventually recombining with a hole. In nanoscale systems, quantum confinement makes the behavior of non-equilibrium electrons more complex and experimentally accessible. To measure non-equilibrium dynamics, we develop a technique for data-intensively imaging photoresponse that efficiently samples phenomenological parameter space. The result is a large set of images which we condense down to dynamical parameters that can be visualized and physically interpreted. Using this data intensive methodology, we explore the non-equilibrium physics of photoexcited electrons on multiple scales. Firstly, on the microscopic scale we study the interactions of excitons and electron-phonon coupling in TMD heterostructures. Then, zooming out to the mesoscopic scale, we observe an electron-hole liquid phase in $MoTe_2$ and hot carrier regime in graphene heterostructures.

Finally, at the statistical scale we explore how a simple model for quieting a noisy antenna can decrease noise in the non-equilibrium states that power photosynthesis. In sum, we demonstrate that data intensive imaging is a powerful and versatile tool for exploring non-equilibrium dynamics of quantum materials and biophysical systems.

TABLE OF CONTENTS

Acknowledgments iv
Dedication vii
Abstract viii

Chapter 1: Motivation and Background

1.1 Non-Equilibrium Physics and the Meaning of Life 1
1.2 Introduction to 2D Quantum Systems 4
1.3 Excitation and Metastability in Quantum Systems 6
1.4 Electrons, Holes and Excitons in 2D Systems 7
1.5 Out of Equilibrium Mesoscale Physics in Quantum Confined Systems 10
1.6 Energy Transfer in Biological Systems 13

Chapter 2: Methodology and Metrology

2.1 The Metrology of Non-Equilibrium Complexity 15
2.2 The Metrology of Optoelectronic Measurements on 2D Materials 20
2.3 The Multi-Parameter Dynamic Photoresponse Microscopy Technique 22
2.4 MPDPM Optics 23
2.5 Apparatus and Cryostat 27
2.6 Integrated Data Acquisition System 29
2.7 hyperDAQ Software Design 32
2.8 Long Measurements and System Stability 38
2.9 Data Taking Process 40

Chapter 3: Data Handling, Analysis, and Visualization

3.1 The Data Analysis Process 43
3.2 Data Handling Concerns 43
3.3 Data Handling Policy 45
3.4 Experimental Metadata and Traceability 47
3.5 Software Toolbox 48
3.6 Hierarchical Analysis of MPDPM Data Sets 50
3.7 Dynamical Fitting 52
3.8 Image Analysis 54
3.9 Visualization Scripting 57
3.10 Visualization of Multi-Dimensional Data 59
3.11 Fitting of Data and the Visualization of Fit Parameters 66

Chapter 4: The Microscopic Scale

4.1 Excitons in 2D systems at the Microscopic Scale 72
4.2 Auger Recombination and Biexcitons in 2D Systems 72
4.3 Photocurrent from Graphene-$MoTe_2$ Heterostructures 74

4.4 Exciton Interactions and the Rate Equation Model 78
4.5 Two-Body Auger Recombination in $MoTe_2$ 82
4.6 Phonons and Interlayer Excitons in 2D Heterojunctions 84
4.7 The Photocurrent Parameter Space of $MoSe_2$-WSe_2 Heterostructures 88
4.8 The Phonon Modes of $MoSe_2$-WSe_2 Heterostructures 98
4.9 The Electron-Phonon Interaction in $MoSe_2$-WSe_2 Heterostructures 101
4.10 Conclusions: Controlling Excitons Unveils Interactions 106

Chapter 5: The Mesoscopic Scale

5.1 Correlated Phenomena at the Mesoscale 107
5.2 The High-Power Behavior of the Graphene-$MoTe_2$ System 108
5.3 Modeling the Photoresponse of an Electron Hole Liquid 115
5.4 Signatures of the Electron Hole Liquid 118
5.5 Implications of Room Temperature Electron-Hole Liquid 120
5.6 The Graphene-hBN-Graphene Heterostructure System 121
5.7 Interlayer-Tunneling Photocurrent Parameter Space 124
5.8 The Hot Carrier Model of Tunneling 129
5.9 Conclusions: Correlated Mesoscale Non-Equilibrium States 136

Chapter 6: The Statistical Scale

6.1 Biosystems as Model Systems in the Statistical Scale 137
6.2 Photosynthesis in a Noisy Environment 138
6.3 The One-Channel Noisy Antenna 139
6.4 The Two-Channel Noisy Antenna 140
6.5 The Discrete Noisy Antenna Model 147
6.6 Calculation of an Optimal Noisy Antenna in Diverse Light Environments 152
6.7 Natural and Artificial Experiments Based on the Noisy Antenna Model 155
6.8 Conclusions: Biophysics of Non-Equilibrium States 158

Chapter 7: Conclusions

7.1 Exploring the Physics of Non-Equilibrium Systems 160
7.2 Lessons Learned 161
7.3 The Future 164

References 166

Appendix A1: Sample Information

A1.1 Sample Information and Metadata 176
A1.2 Graphene-hBN-Graphite Heterostructure Sample 177
A1.3 Graphene-$MoTe_2$ Heterostructure Sample 178
A1.4 $MoSe_2$-WSe_2 Heterostructure Samples 179
A1.5 Graphene-hBN-Graphene Samples 181

Appendix A2:Calculations

A2.1 Estimate of Laser Absorption in $MoTe_2$ 183
A2.2 Steady State Solution to the Exciton Rate Equation Model 183
A2.3 Mean Distance Between Electron-Hole Pairs at P_C 184
A2.4 Modeling the spatially resolved photoresponse of an electron hole liquid 186
A2.5 Discrete Toy Model of Two-Channel Noisy Antenna 187
A2.6 Optimization Parameter for the Two-Channel Noisy Antenna 189
A2.7 Optimization of the Noisy Antenna in Diverse Light Spectra 197

LIST OF FIGURES

Chapter 1: Motivation and Background

1.1 Lattice Structure of Quantum Materials 6
1.2 Exciton Schematics 9

Chapter 2: Methodology and Metrology

2.1 Measurement and the Geometry of Phase Space 17
2.2 Schematic Diagram of the Principle Optics 25
2.3 Characterization of the Ultrafast Optics 26
2.4 Cross Section of the Scanning Optics and Cryostat 27
2.5 Photographs of the Optical Cryostat 28
2.6 MPDPM Data Flowchart 30
2.7 hyperDAQ Object Flowchart 34
2.8 hyperDAQ User Interface 36
2.9 hyperDAQ Working Directory Structure 37
2.10 Spatial Drift Correction 39

Chapter 3: Data Handling, Analysis, and Visualization

3.1 Schematic of the Hierarchical Data Analysis 51
3.2 Hierarchical Analysis of the *GBNGr* Imaging Dataset 53
3.3 Schematic of Image Analysis for the *GBNGr* Dataset 55
3.4 Diagram of the Composition of a Visualization Script 57
3.5 Comparison Different Types of Colorscale 61
3.6 Comparison of a Colormap and Line Plot 62
3.7 Tile Plot Representation of a 3D Photocurrent Dataset 63
3.8 Tile Plot Representation of a 4D Photocurrent Dataset 65
3.9 Fitting of a Line Scan 67
3.10 Representing a Scan with a Fit 69
3.11 Example of Filtering Fitting Parameters 71

Chapter 4: The Microscopic Scale

4.1 Schematic of Measurements on the *GMoTe2* Sample. 74
4.2 Photocurrent characterization of the *GMoTe2* heterostructure 76
4.3 Photocurrent Dynamical Fits in the Low Power Regime 77
4.4 Exciton Rate Equation Model Solutions 81
4.5 Fits to the Exciton Rate Equation Model 83
4.6 Schematic of the $MoSe_2$ -WSe_2 Heterostructure 85
4.7 Interlayer Excitonic Transitions of the $MoSe_2$ -WSe_2 Heterostructure 86
4.8 *MoSe2WSe2-2* Sample Schematic 87
4.9 Photocurrent spectroscopy on the *MoSe2WSe2-1* Sample 90
4.10 Data Intensive of the *MoSe2WSe2-2* Sample Versus Voltage 92

4.11 Data Intensive of the *MoSe2WSe2-2* Sample Versus Wavelength 93
4.12 *MoSe2WSe2-2* Photocurrent Derivatives 94
4.13 Determining Electronic Characteristics of the *MoSe2WSe2-2* Sample 96
4.14 Oscillations the *MoSe2WSe2-2* Sample at Room Temperature 97
4.15 DFT Calculation of $MoSe_2$-WSe_2 Phonon Modes 99
4.16 Visualization of $MoSe_2$-WSe_2 Optical Phonons 100
4.17 Determining Candidate Phonons 101
4.18 Summary of the 30 meV Oscillations from Multiple Experiments 102

Chapter 5: The Mesoscopic Scale

5.1 *GMoTe2* Spatial Photocurrent as a Function of Power 109
5.2 Spatial γ Maps Versus Two-Pulse Delay 110
5.3 Spatial Photocurrent and Gradient Maps Versus Power 111
5.4 Ring Feature Size Versus Power 112
5.5 Spatially Integrated Photocurrent Versus Laser Power 113
5.6 Electron-Hole Dynamics as a function of Density 115
5.7 Spatial Photocurrent Model Versus Experiment 116
5.8 Ring Feature Voltage Dependence 118
5.9 Electron-Hole Liquid Temporal Dynamics 119
5.10 Schematic of the Graphene-hBN-Graphene Heterostructure 121
5.11 Resistance Data for *GBNG* Devices 122
5.12 Spatial Photocurrent Maps of the *GBNG-3* Sample 125
5.13 *GBNG* Photocurrent Dynamic Behavior 127
5.14 *GBNG* Photocurrent and Photoconductance Versus V_g 128
5.15 Formation of the Hot Carrier Distribution 130
5.16 Schematic of Hot Carrier Tunneling 131
5.17 Simulated Tunneling Photocurrent for Uniform Temperature Profile 133
5.18 Simulated Tunneling Photocurrent for Gaussian Temperature Profile 135

Chapter 6: The Statistical Scale

6.1 Two-Channel Model Schematic 141
6.2 One Versus Two Channel Model Calculation 146
6.3 One Versus Two Channel Model Full Parameter Space 148
6.4 Discrete Two Channel Model 149
6.5 Discrete Two Channel Model in Three Cases 151
6.6 Optimal Noisy Antenna in Three Photosynthetic Niches 153
6.7 Underwater Noisy Antenna 157

Appendix A1: Sample Information

A1.2.1 Optical Image of the *GBNGr* sample 177
A1.3.1 Optical Image of the *GMoTe2* sample 178
A1.4.1 Optical image of the *MoSe2WSe2-2* sample 179

A1.4.2 Optical characterization of the *MoSe2WSe2-2* sample 181
A1.5.1 Optical images of the *GBNG* samples 182

Appendix A2: Calculations

A2.6.1 Schematic of Optimization Parameter Calculation 196
A2.7.1 Full Model Optimization Calculations for Prototypical Phototrophs 198
A2.7.2 Model Optimization for Variable *w* 199
A2.7.3 Model Optimization Degeneracy 202

CHAPTER 1:

MOTIVATION AND BACKGROUND

1.1 Non-Equilibrium Physics and the Meaning of Life

If pressed to give a one sentence definition for all of physics, I would say that physics is the study of the motion of matter and energy through spacetime. While this definition is vague, and unhelpful in determining what a physicist actually does, it has the advantage of being correct for almost any physics problem ones applies it to. Fundamentally, a physicist working in any given context will consider some situation where there is matter, with some energy, and determine how it evolves through some space over some time, or more generally, in some spacetime. This approach solves problems in wildly different contexts, from the first energetic moments of the universe, to avalanches in the mountains, to the molecular mechanisms that allow life to exist and human beings to contemplate the world around them.

Condensed matter physics is not different from any other sub-discipline of physics, it considers how matter in many body systems evolves through a system of molecules arranged in some form. But more than other sub-disciplines of physics condensed matter is limited by the pesky concept of thermal equilibrium. While thermal equilibrium is a fundamental phenomenon with deep and interesting physical implications, it is in another sense destructive. A system that is thermally equilibrated has no history; it does not matter how the system reached thermal equilibrium, it has the same properties any path it takes. Much work in solid state physics goes into determining the equilibrium properties of a material, and to the layperson this might seem incongruous with the definition of physics given above. After all, if thermal equilibrium depends on the state, not the history, of a system why is it worth studying the motion of matter and energy through a crystal to determine that history? Why bother determining the microscopic dynamics when all that matters to thermal equilibrium is the ensemble average? If every system inevitably ends up at thermal equilibrium, isn’t the determination of equilibrium properties all that matters in the end?

But this is a pessimistic illusion, there are many things in this universe driven by non-equilibrium physics. Consider that in order to read this dissertation neurons in the reader’s brain need to fire many billions of times, and each neuronal action potential exists in defiance of thermal equilibrium. Thermalization would mean death to any living thing. Life is, in some sense, a constant struggle against returning to thermal equilibrium, and while no individual organism ever wins their battle against thermal equilibrium our continued existence means that the war goes on[1]. Humans take this thermal rebellion even more seriously than other life, not only do we exist in spite of thermal equilibrium we make technology that either carefully manipulates or outright defies thermal equilibrium and use that technology to make our non-equilibrium lives longer and better. The field of Condensed Matter physics is perhaps the front line in this struggle, as we seek to not just leverage thermal equilibrium but also generate and explore novel non-equilibrium states.

The study of non-equilibrium states, especially experimentally, is difficult. Thermal equilibrium always wins in the end, and in doing so it erases any information about non-equilibrium states. But recent discoveries may help unveil these states, the isolation of nanoscale materials has created systems that, unprotected by the usual three-dimensional symmetries, can be strongly driven far out of equilibrium, such that rich dynamics can occur and be observed before they inevitably relax. Furthermore, our understanding of biological systems has advanced to the point that we can begin to understand how energy flows through these systems and is used to power the non-equilibrium state that is life. These systems are related at a deep level, they both are driven by the question of what physics occurs in an excited system prior to relaxation. That is, if energy is injected into a system what does it do, or what can it be made to do, before it ends up dissipating into thermal equilibrium?

This dissertation is about answering this question and understanding what happens in non-equilibrium states. While it is not possible to understand what happens in every non-equilibrium state, it is possible to answer that question in a number of model systems in which we can explore more general concepts. Furthermore, phenomena occur in these model systems on several scales, the microscopic, the mesoscopic and the statistical scales, each of which will be explored. Equally important to the question of what happens in non-equilibrium systems is the question of how we know what happens. Physics is an empirical science, it does not matter what we think happens, it matters what we can empirically show happens with controlled experiments. Thus, this dissertation will also focus on the metrology of these systems and how to perform experiments in these model systems. This is especially important since the field of nanoscience in general, and non-equilibrium nanoscience in particular, is in its early stages and therefore it is important to get the metrology right to enable future discovery. With this goal we embark on an attempt to image non-equilibrium dynamics in quantum materials at the microscopic, mesoscopic, and statistical scales.

1.2 Introduction to 2D Quantum Systems

In 2004 the first monolayer system was isolated[2]. Graphene, a single sheet of carbon atoms in a hexagonal lattice, as shown in Fig. 1.1, was fabricated by the simple application of scotch-tape to graphite and remains one of the most important and influential two-dimensional (2D) materials in the world. The key aspect that makes graphene so interesting is its atomic thinness, the electrons in a single sheet of graphene are confined to a single layer, well below their normal de Broglie wavelength. This quantum confinement makes them effectively two-dimensional, fundamentally changing the physics that can occur in them. In the intervening years, a vast number of new discoveries have been made and physics that was previously only accessible in carefully engineered quantum wells became widely accessible[3-5].

Since the discovery of graphene, scientists across physics, chemistry and engineering have developed rapidly evolving techniques to engineer novel quantum devices from atomically thin materials such as insulating hexagonal boron nitride (hBN, which has similar atomic structure to graphene) and semiconducting transition metal dichalcogenides (TMDs, show in Fig. 1.1b)[6,7]. These materials can be stacked vertically into van der Waals heterostructures that combine the electronic properties of the constituent materials in unusual ways[8,9]. Much recent research has focused on combining and engineering 2D materials to create designer properties that result from length scale engineering i.e. tuning the electronic properties by structuring the critical device length scales at or below the electron wavelength[10-14]. In loose analogy to optical metamaterials, engineering sub-wavelength structure in these quantum metamaterials may give unprecedented access to quantum material properties, allowing us to engineer custom unit cells, topological bands, and altered excited states. Intriguingly, length scale engineering of these materials may also allow us to tune interactions between charge carriers in the materials, creating novel correlated electronic phases[15-20]. Notably, many of these novel 2D properties allow for increased meta-stability and accessibility of excited states which gives rise to interesting non-equilibrium physics.

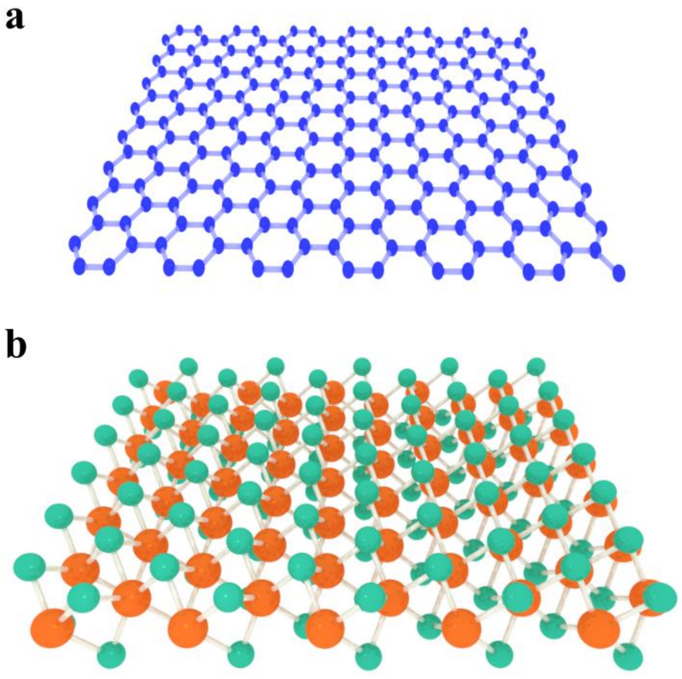


Figure 1.1: Monolayer lattices of graphene, **a**, and transition metal dichalcogenides (TMDs), **b**.

1.3 Excitation and Metastability in Quantum Systems

When energy is injected into a material system it spreads out into various microscopic modes until it thermalizes. For example, if one bangs a piece of metal with a hammer, the kinetic energy imparted initially travels through the metal as a shock wave but the strong bonds between the atoms disperse this pulse of energy until, after a few reflections, all that energy is spread out completely incoherently among the vibrational modes of the crystal lattice, that is until the energy becomes heat. Thus, an organized shock of kinetic energy becomes disorganized thermal energy, generating entropy.

In classical systems like this, it is difficult to define what a non-equilibrium state is since the transition from initial organized state to the final thermalized state is continuous. Furthermore, at any given time the energy is usually a perturbation from the equilibrium properties of the material, i.e. the only difference between the deformation of the lattice due to the shock wave and the deformation due to random thermal motion is the evanescent coherence. Quantum mechanics breaks this mold. Instead of a continuous transition to thermal equilibrium, quantum systems have discrete states and sudden transitions. Perturbations to equilibrium quantum states may form quasiparticles that can move, interact and have physics of their own. This adds a key concept that enables non-equilibrium physics: metastability. An excited quantum state, once created, may exist for a significant amount of time before relaxing to the ground state.

The process we will use to create non-equilibrium states in this thesis is photoexcitation, specifically because it generates a metastable state with a long lifetime. When a photon is absorbed in a material it excites an electron into a higher energy quantum state, from a valance band into a conduction band. This discrete transition creates two particles, the negatively charged electron and a positively charged hole, which a quasiparticle composed of the excess charge left over after an electron is promoted out of an atomic orbital. Eventually the electron must return to its ground state, that is it must recombine with a hole, with the excess energy becoming either heat or a photon (which will, inevitably, become heat later). But quite a bit can happen before recombination. If the band structure of a system has a gap or a constriction then the lifetime of the electron and hole may be quite long, and the electron and hole can move and interact in interesting ways.

Thus, the basic process that we will be exploring is as follows: when a quantum material is photoexcited, creating a population of electrons and holes, what do those electrons and holes do before they relax. Alternatively, it is usually convenient to think about this problem in terms of energy; when an electron is photoexcited it gains excess energy, which must be shed to reach its ground state. What happens to that excess energy and, importantly for applications, what can be done with it?

1.4 Electrons, Holes and Excitons in 2D Systems

When a quantum material is photoexcited, a negatively charged electron and a positively charged hole are created. Given the charges, and the laws of electromagnetism, the electron and holes experience an attractive Coulomb force. But the quantum nature of the system provides metastability, if there is a gap between the electron in the conduction band and the hole in the valence band then they cannot immediately annihilate, rather they exist, strongly interacting for some lifetime before a discrete transition brings relaxation. Similar to how an electron and a proton interact attractively, but cannot annihilate due to the symmetries of quantum mechanics, and instead form a stable hydrogen atom, an electron and hole, if prohibited from recombining by a gap (provided by some symmetry of the crystal lattice) may form a bound state called an exciton, shown schematically in Fig. 1.2a.

Excitons have long been studied in condensed matter systems[21] but until the advent of 2D materials they have largely been studied in careful engineered materials at very low temperatures, due to the fact that the binding energies of the exciton in a conventional material is very small compared to the 26 meV thermal energy scale room temperature. However, in 2D material systems the binding energies are significantly higher due to the reduced dimensionality[22]. Higher binding energy means it is easier to generate excitons and a wider variety of excitonic behavior can occur. Accordingly, excitons of various types have been a major topic of research in recent years[23-25].

There are many questions that can be asked about excitons, but the focus of this thesis is on their out-of-equilibrium dynamics, that is once an exciton is generated what happens to it? There are three general options: an exciton can spontaneously recombine releasing a photon, and exciton can be dissociated by an electric field resulting in a free electron and a free hole, or an exciton can interact with another particle or quasi-particle. The first option, spontaneous recombination, is relatively straightforward and we will not focus on it aside from noting that it is not always possible; indirect excitons have momentum constraints and require the exciton to interact with something in order to recombine. The second option, exciton dissociation, is the process which we use to generate a photocurrent signal in experiments. At finite electric field, excitons are polarized and have some probability of dissociating. The resulting electrons and holes are accelerated by the electric field, as shown in Fig 1.2b, they transit the device and are collected at the electrical contacts[26-28]. This is a linear process that is also straightforward. Therefore, the bulk of the work on excitons presented in this dissertation will focus on the third option, the interaction of excitons (Fig 1.2c).

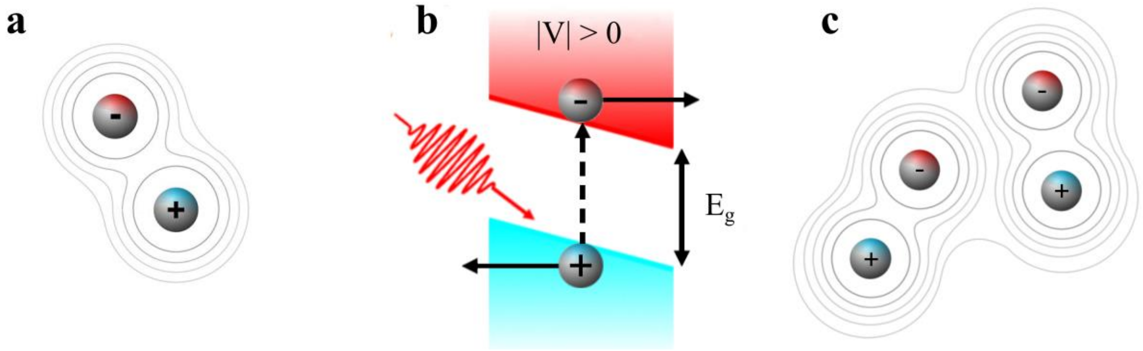


Figure 1.2: **a**, an exciton composed on an electron and a hole due to Coulomb attraction. Electric field represented with equipotential lines. **b**, the dissociation of an exciton in finite electric field, the resulting free carriers are accelerated by the field. **c**, interacting excitons.

Excitons may interact in a variety of ways, but we will focus on two prototypical interactions. Firstly, we will focus on interactions that result in recombination, that is when an exciton scatters off another body causing it to annihilate and, in the process, imparting energy or momentum to the other body. This broad class of processes is known as Auger recombination[29-31]. The second prototypical interaction is the exciton-phonon interaction, where an exciton undergoes a three-body process with a phonon, usually either in absorption or recombination. Ultrathin TMD materials are 2D systems that are frequently favorable to the formation of excitons in significant quantities, while also being relatively easy to fabricate and highly controllable. Therefore, we will use TMD heterostructures as model systems to study the interactions of excitons. The experimental results and findings on exciton interactions will be presented in Chapter 4.

1.5 Out of Equilibrium Mesoscale Physics in Quantum Confined Systems

It is a common adage in physics that a physicist counts as 1, 2, 3, N, with the implication that N is astronomically large. This is because while the one body and two body problems can usually be solved exactly, and the three-body problem can often be solved with approximations, above that the math begins to get prohibitively complicated for an exact solution. But at large N it is possible to forget about the individual bodies and focus on average behavior, simplifying the math once again. Solving the many body problem for intermediate values is mathematically difficult, usually requiring numerical methods if it's possible at all. The discussion up to this point has focused on individual charge carriers and their interactions as two or three-body interactions. Furthermore, we generally assume that the carriers act as a non-ideal gas, that is that they interact with each other only rarely, making higher-order interactions much less likely compared to the rate of two or three body interactions. This assumption is convenient, given the difficulties presented by higher order interactions, but like all simplifying assumptions we must examine and eventually relax it.

There exist regimes in which we may no longer assume that charge carriers rarely interact. In such strongly interacting regimes not only are many body interactions much more common, but the behavior of a charge carrier may be driven by a series of interactions instead of a single discrete interaction. Keeping track of what happens to individual charge carriers becomes impossible, if but charge carriers interact many times, then what happens at the microscopic scale is not very relevant. Rather the dynamics are driven by energy distributing itself among the many charge carriers. The result is a phase that is no longer a non-ideal gas but is correlated. A simple analogy being a liquid, where molecules interact constantly, versus a gas, where molecules interact rarely. Since energy can be efficiently transmitted between charge carriers in correlated phases, and individual microscopic interactions are meaningless, our focus then turns to a larger scale; the intermediate scale between the microscopic and macroscopic, known as the mesoscopic scale. This is distinct from the statistical scale (which we will discuss in the next section and chapter 6) as we are not averaging over time or assuming an infinite ensemble, rather for correlated phases we simply zoom out from the microscopic scale.

At the mesoscale quantum confinement makes the non-equilibrium dynamics more accessible than they would be otherwise. Firstly, the confinement can function to squeeze a larger number of carriers into a smaller volume. Consider a sheet of graphene absorbing light, which creates a number of electrons and holes which then move within the system. In a 3D material the carriers could expand in three dimensions, decreasing the probability that any individual carrier would interact with another carrier on any given timescale. In 2D the volume that the carriers can expand into is significantly decreased, increasing the probability of interaction. Thus, at a constant number of charge carriers the probability of interactions goes up in 2D systems. Quantum confinement in 2D systems can also lead to more long lived and more stable non-equilibrium states, allowing excited charge carriers to last longer and therefore interact more.

In this dissertation we will consider two types of mesoscale correlated states, the electron hole liquid in $MoTe_2$ and hot carrier regime in graphene. The electron-hole liquid is a semiconductor phase where electrons, holes and excitons may condense into a metallic Fermi liquid, usually through the formation of distinct droplets in a gas of excitons[32,33]. This process is in direct analogy to the condensation of a gas; as the density of a gas increases, a phase transition occurs, and droplets begin to condense and act as a liquid. The electron-hole liquid was observed in many conventional semiconductors, such as Si and Ge, but is found at low temperature because the binding energy of the excitons and the liquid phase are small compared to the thermal energy scale. As discussed above the dimensional confinement of 2D systems means the binding energies are larger, opening the intriguing possibility of room temperature correlated phases.

In graphene, though electrons and holes are not bound into metastable excitons they are constrained by the 2D band structure. The density of states is constricted at the charge neutrality point of graphene, creating a bottleneck for recombination. An electron that is excited by a photon from the valance band to the conduction band will interact with other carriers and thus begin to “fall” through the band to recombine with holes. But since there is a constriction, carriers may exist at higher energies for a longer timescale. In addition, excited carriers in graphene interact with each other on a faster timescale than they interact with the lattice, that is they can exchange energy and thermalize with other carriers much faster than with the lattice[34-38]. The result of this is a "hot carriers, cold lattice" regime where charge carriers in graphene form a hot quasi-thermal equilibrium state on timescales of hundreds of femtoseconds, which then cools through interactions with the lattice on timescales of picoseconds. This hot carrier state is strongly interacting and would be understood in the mesoscale regime.

1.6 Energy Transfer in Biological Systems

The microscopic states and macroscopic phases of non-equilibrium states exist for a short time before relaxing back to equilibrium. Macroscopic properties are usually determined by averaging over the short timescales on which these dynamics occur. This would represent a return to thermal equilibrium, so at first glance considering physics on this time averaged scale would seem to be antithetical to the study of non-equilibrium physics. However, what occurs on short non-equilibrium timescales affects what the time-averaged result will be, otherwise we would not be able to observe them. Therefore, we must consider non-equilibrium physics on the statistical scale, that is how non-equilibrium physics affects the eventual statistical time-averaged properties of the system. In the 2D material systems we discuss, this is fairly straightforward; the interactions of excited states either increases or decreases the number of charge carriers measured at a particular location on the mesoscopic sample. However, other systems can have more subtle statistical effects, particularly biological systems.

Nanoscale systems are not entirely artificial, biology has been generating them since life began on Earth. As discussed above, living organisms are non-equilibrium systems, and the basis of their non-equilibrium existence is large variety of nanoscale biomolecules, such as proteins and enzymes, that perform chemistry and manipulate energy. Though most of the mechanisms have complicated electro-chemical processes that would require detailed molecular models, we can gain insight from looking at the quantum processes involved and examining how the non-equilibrium physics of these systems effects the ability of these organisms to live.

We will focus on the prototypical optoelectronic biological process, photosynthesis. Photosynthesis occurs in molecular light harvesting complexes, and similar to 2D materials, when a photon is absorbed it creates an electron and hole pair with excess energy. The process of photosynthesis attempts to convert the energy of the non-equilibrium electron into useable chemical energy. To do this, photoexcitation energy is rapidly transferred through a molecular network before reaching the reaction center, where charge transfer converts excitation energy into an electrochemical potential gradient across the photosynthetic membrane[39]. This occurs within intricate networks despite both large variation in the external light source and internal fluctuations. A photosynthetic organism must regulate this noisy process, spending the least amount of time over or under the desired input power. In chapter 6 we will explore how the structure of the absorbers allows organisms to quiet a noisy photosynthetic antenna, demonstrating that the particular quantum structure of a non-equilibrium mechanism can influence the behavior of physical systems in profound ways.

CHAPTER 2:

METHODOLOGY AND METROLOGY

2.1 The Metrology of Non-Equilibrium Complexity

The non-equilibrium states that we wish to explore exhibit complex behavior that poses significant measurement challenges. We access non-equilibrium physics through the response of a heterostructure to optical excitation, or photoresponse, which is not straightforward to measure or interpret. Fundamentally, thermal equilibrium reverts quantities to an accessible average value, therefore non-equilibrium states may involve more fluctuations and noise than equilibrium states. Furthermore, the two-dimensional nature of 2D systems removes symmetries that may protect, but also limit, quantum states, and the atomically thin spatial profile reduces screening and makes 2D materials more sensitive to external fields. In sum, this means that 2D systems tend to depend on more experimental variables and have more relationships between those variables. The result is complexity in the measurement of photoresponse.

As nanotechnologists and materials scientists, how do we systematically assess complex electronic behavior that may arise in new material systems, particularly those with unusual synthetic properties? In solid-state physics, the answer has traditionally been to set up a single-parameter experiment that aims to cut through the complexity and capture quantum phenomena in as concise a measurement as possible. Typical experiments consist of well-established transport or spectroscopic measurements sampling over a *single* independent variable. Often, these measurements use commercially available instruments. Implicit in this approach is the assumption that all other experimental parameters have negligible effect on the variable of interest. In 2D materials, many properties are the result of atomic thinness, which also makes them sensitive to external conditions, defying the assumption that other independent variables do not contribute to the electronic behavior. Truly comprehensive characterization using standard measurement approaches would require prohibitively long times, due in part to the measurement rate and the numerous trials required to address variations across many material parameters. As the complexity of 2D systems increases, new data intensive approaches - taking inspiration from astrophysics, high-energy physics, and biomedical imaging - must be developed.

In this chapter, we lay out an elementary assessment of the most restrictive experimental parameter - experimental time - and discuss how multi-variable searches can be optimized to improve the search for correlations across experimental variables. Fundamentally, experimental time T is the dominant limiting factor in measuring complex device behavior. Simply stated, the total time of a measurement combines the hardware-limited time per point t_h with the sample response time t_s, multiplied by the total number of data points to be measured.

To illustrate how the total time can be evaluated for a simple experimental system, Fig. 2.1a shows a generic phenomenological response that depends on two experimental parameter dimensions, measured with single variable measurements. The experiment sweeps the X variable at constant Y, taking a series of line cuts through the experimental phase space. The time of such an experiment is $T = (t_h + t_s)\frac{\Delta X}{r_x}\frac{\Delta \mathrm{Y}}{r_y}$, where ΔX and ΔY are the ranges of X and Y defining the parameter space, and r_x and r_y are the resolutions of the X and Y variables. Generalizing to an N-dimensional parameter space spanned by N independent variables, $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_N)$:

$$T = (t_h + t_s) \prod_{i=1}^{N} \frac{\Delta \boldsymbol{e}_i}{r_i} \tag{1.1}$$

Here, equation 1.1 can be understood intuitively as the time spent per voxel multiplied by the volume of the parameter space, $\prod_i \Delta \boldsymbol{e}_i$, divided by the voxel volume,$\prod_i r_i$. For a fixed parameter space volume, as the voxel volume decreases (i.e. the resolution increases), the total experimental time will increase.

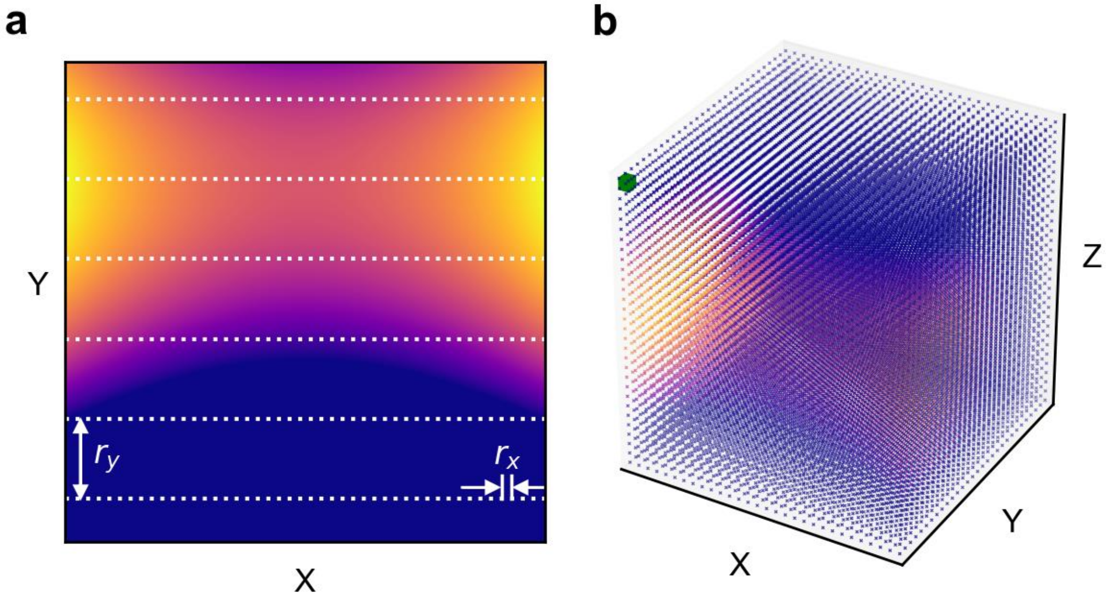


Figure 2.1: Measurement phase space. **a**, represents the phase space of a hypothetical phenomenon that depends on two independent variables, with the observable value represented by a color scale. Single variable measurements are represented as dashed white lines with Y held constant. **b**, represents the phase space of a hypothetical phenomenon that depends on three independent variables, each point in three-dimensional space has an observable value represented by color. The green cube in the upper left represents a single voxel.

Increasing the dimension, N, of a parameter space enforces greater limitations on total experimental time. To see this, Fig. 2.1b illustrates the same phenomena as Fig. 2.1a but in a 3D parameter space (N = 3), representing observables as colored points, and showing a voxel as a small green cube. Measurements of complex systems - those where non-trivial correlations exist between N > 1 independent variables - require significant values of $\Delta\boldsymbol{e}$ and r to obtain sufficient data for meaningful statistical analysis. In Fig. 2.1b, we see that due to the dimensionality of the phase space, the number of voxels is exponentially larger than for a two-dimensional experiment. Comprehensive measurements in a N-dimensional parameter space thus require exponentially more time.

High dimensional experimental phase spaces require making careful choices to minimize T while acquiring sufficient data for robust statistical analysis. Assume that, in general, T is large and constant, limited by experimenter (i.e., graduate student) time, sample lifetime or other resources. Optimizing high dimensional measurements involves optimizing the hardware, which decreases t_h, or optimizing the search of parameter space by making tradeoffs in $\Delta\boldsymbol{e}_i$ and r_i. However, the intrinsic sample response time t_s limits how fast a measurement can proceed, and if $t_s \gg t_h$, hardware optimization does little to increase measurement efficiency. Hardware optimization is application specific; it is discussed for the hardware used in this thesis in Sections 2.4 and 2.5.

The greatest gains in efficiency come from tradeoffs in resolution. Ideally, the experimenter can reduce excessive resolution in one parameter to gain resolution in another parameter. Less ideally, the experimenter can choose to restrict the range of one or more parameter(s) $\Delta\boldsymbol{e}_i$, or neglect certain parameters, resulting in a narrower but better resolved measurement. The latter is a common strategy, but has greater likelihood of missing or misrepresenting phenomena occurring within a complex parameter space.

In the large T limit, conventional single variable measurements are fundamentally inefficient. By their nature, single variable measurements explore one parameter, for example the X variable in Fig. 2.1a, with high resolution, and all other variables held constant, meaning $r_x \ll r_y, r_z, \ \ldots, \ r_N$. With hardware heavily optimized for only one variable it is difficult to effectively trade resolution in X for resolution in another variable and experimenters often deal with finite time by restricting $\Delta \boldsymbol{e}$ or omitting parameters. Single variable measurements become increasingly ineffective in identifying cross correlations between multiple parameters as the complexity of a measurement increases (i.e., as N increases), as higher resolution is needed, or as the relevant ranges increase.

Does a better understanding of multi-parameter measurement science translate into accelerated discovery? While this is impossible to precisely answer, we posit that experimentalists using only standard techniques risk falling prey to a version of the availability heuristic. By focusing on measurements that are easy to perform with off-the-shelf or commercial equipment, complex phenomena that correlate across multiple parameters are missed or misinterpreted. Expectation bias is a danger when choosing parameters for new materials: an experimenter may unconsciously select the parameters that are most likely to conform to expectations or established models[40-43]. Choosing which variables to hold constant can easily introduce selection bias that leads to compelling, yet incomplete, phenomenological knowledge complicating realistic interpretation. Comprehensive methods are therefore significantly advantageous in the search for new phenomena, particularly when a unique target system is probed using multiple non-standard experimental techniques.

2.2 The Metrology of Optoelectronic Measurements on 2D Materials

In optoelectronic materials, photo-excited electrons are promoted to high energies, leaving behind short-lived charge vacancies, or holes. In this way, electrons promoted across a semiconductor band gap result in long-lived electron-hole pairs, while those excited in a semimetal may result in short-lived excitations. The timescale over which the electron-hole pairs recover to equilibrium is determined by energy and momentum relaxation processes in the material, which in turn depend on electronic band structure, electronic interaction strength, and electron-phonon coupling.

In 2D semiconductors and semimetals, photoexcited electron-hole pairs may interact in unusual ways, giving rise to many body correlations that persist even at room temperature. In TMDs, charge carriers form hydrogen-like bound states with well-defined orbital and spin angular momentum[23,24]. Depending on the structure of the material, these strongly bound excitons may be influenced by non-trivial bands, such as topological or moirè bands, or have additional quantum numbers such as valley index or pseudospin[10]. In graphene, the electron-hole pairs form a rapidly evolving hot carrier distribution exhibiting unusual cooling pathways, with electron-electron and electron-phonon scattering processes competing to relax excess energy. Combining 2D semiconductors, 2D insulators, or semimetals into van der Waals heterostructures introduces additional degrees of freedom, for instance allowing excitons to form with the electron and hole in different materials[11]. All of these unique properties contribute to energy and momentum relaxation, giving rise to highly complex behavior over a large range of time scales, from femtosecond electron-electron scattering to nanosecond exciton recombination.

These unusual electron-hole interactions in van der Waals metamaterials result in part from reduced dimensionality, which increases the energy scales of electronic states and interactions (e.g., increasing the binding energy of excitons)[22]. Due to electron confinement, 2D materials allow correlated or interacting phases to exist at higher temperatures than in conventional materials. Such effects are less accessible in 3D materials, which exhibit high symmetry due to translation invariance of the unit cell in all three spatial dimensions. Not only does high symmetry constrain the possible phenomena in many ways, it also allows the experimenter to make several assumptions about the behavior based on the unit cell. 2D materials inherently break several exploitable symmetries, expanding the space of possible phenomena and increasing the phase space for electronic states and interactions.

In multiple respects the properties that make van der Waals heterostructure metamaterials interesting also make them difficult to measure and understand. Understanding electron-hole pair dynamics in 2D systems presents numerous experimental challenges since observable quantities - such as current, voltage, reflectivity or photoluminescence - are averaged in space and time. Purely electronic measurements only access low energy dynamics near the Fermi surface and average the electron dynamics over the spatial extent of the device. In the time domain, dynamics occur on timescales of femtoseconds to hundreds of picoseconds, and if an excitation persists significantly longer than those timescales it will give only steady state equilibrium values. Gaining experimental information about the dynamics and testing theoretical models requires optical techniques with high spatial and/or temporal resolution[25]. Moreover, in van der Waals heterostructures, multiple unusual electronic effects may overlap. Though individual effects could be exploited for manipulating electronic behavior, experiments must consider and carefully control for all overlapping effects. Separating out individual properties requires multiple experimental variables, so that the property of interest can be uniquely accessed.

2.3 The Multi-Parameter Dynamic Photoresponse Microscopy Technique

We describe a technique, called Multi-Parameter Dynamic Photoresponse Microscopy (MPDPM), that efficiently measures the optoelectronic response of van der Waals heterostructures. Utilizing diffraction limited optics, ultrafast lasers and scanning mirror optics, MPDPM excites the sample with a high intensity optical probe that drives the sample away from equilibrium, thus accessing correlated states, resolving short timescales, and producing high signal-to-noise photoresponse. The optical components are automated and controlled by an integrated, fully automated Data Acquisition (DAQ) program that simultaneously controls all other experimental parameters (such as applied voltage, magnetic field, temperature etc.). Such centralized control allows for efficient tradeoffs between parameters when exploring a large sample phase space. This technique acquires data rapidly, densely and systematically with respect to many experimental variables, resulting in high dimensional data arrays. The end result of MPDPM is a large set of photoresponse images spanning all relevant experimental variables that will ideally capture all the complexity of device's phase space. Although the data sets are more complex than in conventional measurements, these large and complex data sets can be efficiently handled through careful data analysis, described in Chapter 3.

The sections below discuss MPDPM as it applies to photocurrent measurements of van der Waals heterostructures, the experiments discussed in this dissertation. However, the technique is fairly general and could apply to a large variety of applications, thus it is useful to discuss what defines an MPDPM measurement in general, separate from a different data intensive technique. The core aspects of the technique are reflected in the acronym, Multi-Parameter Dynamic Photoresponse Microscopy, is designed to work with microscopy (of some form) that measures the photoresponse dynamics (that is the response of a sample to some local photo-excitation) in a multi-parameter view, in other words spanning several possible experimental variables. The multi-parameter nature of the measurements will require a sophisticated data handling and analysis procedures and will usually involve processing the data to reduce the dimensionality of the data set, as discussed in Chapter 3.

2.4 MPDPM Optics

MPDPM uses a local ultrafast optical probe to perform space-time resolved photocurrent and reflectance measurements. Incident light focused to the diffraction limit can resolve micron sized in-plane features, and the high incident intensity under a diffraction limited beamspot increases the signal and can drive the system well out of equilibrium. Using a scanning diffraction limited beamspot also allows light reflected back through the optics to be focused onto a single pixel detector, with much higher signal to noise than a CCD. The dynamics of charge carries often occur on timescales of order femtoseconds to picoseconds, so excitation by a continuous wave laser gives only equilibrium, steady state values, washing out the dynamics. Therefore, the optical probe must be localized in time as well as space. Ultrafast pulsed lasers can generate pulses on the order of the dynamics, giving access to phenomena that occur on those relevant timescales. In addition, the high peak pulse intensity increases the fluence of incident light, driving the system harder and increasing the signal.

To generate an optical probe that is local in space and time, we combine the techniques of scanning beam photocurrent and reflectance microscopy with ultrafast optical two-pulse measurements[44,45]. A schematic of the optical system is shown in Fig. 2.2. We use a Coherent MIRA 900 OPO ultrafast laser which generates 150 fs pulses with controllable wavelength from 1150 nm to 1550 nm at a 76 MHz repetition rate. The output of the laser is split into two paths by a 50/50 beamsplitter and a translation stage is used to controllably introduce a path length difference. The two beams are then recombined, and the path length difference splits a single laser pulse into two sub-pulses separated by a time delay, Δt.

The recombined beam is fed into scanning beam optics which consist of rotating mirrors and a system of two lenses that focus the beam onto the back of an objective lens. The objective lens is set at the focal length of the second lens such that, as the scanning mirror rotates the beam remains focused onto the same position on the back of the objective but arriving at different angles. The objective lens focuses the light down into a diffraction limited beamspot where the position of the beamspot depends on the incident angle. As the scanning mirror rotates, the beamspot moves over a large area of the sample without aberration, allowing for quick, high-resolution scanning. Many conventional optoelectronic measurements keep the optics fixed and translate the sample. While simple, this technique is too slow to sample phase space in a time efficient way. When focused, the laser beamspot spatial profile is an Airy disk, which can be approximated using a Gaussian point spread function. Fig. 2.3a shows the measured photoresponse of an absorber smaller than 1 μm using a wavelength of 1200 nm. The data is fit well by a Gaussian function (black line) with full width at half maximum of 1.67 μm, indicating that our system is at the diffraction limit.

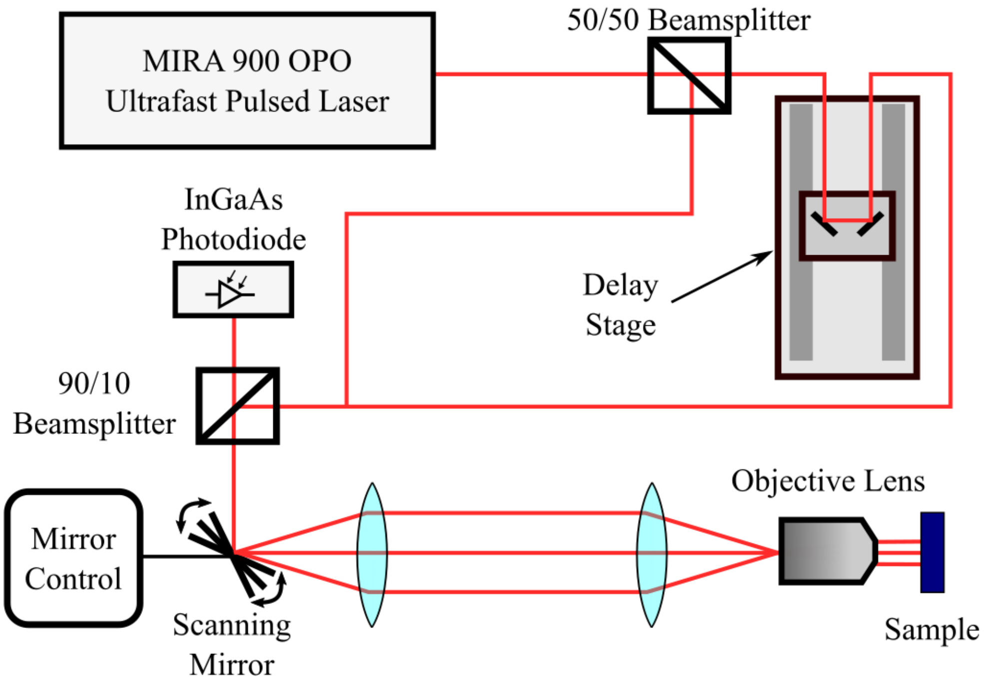


Figure 2.2: Schematic diagram of the principle components of the optics.

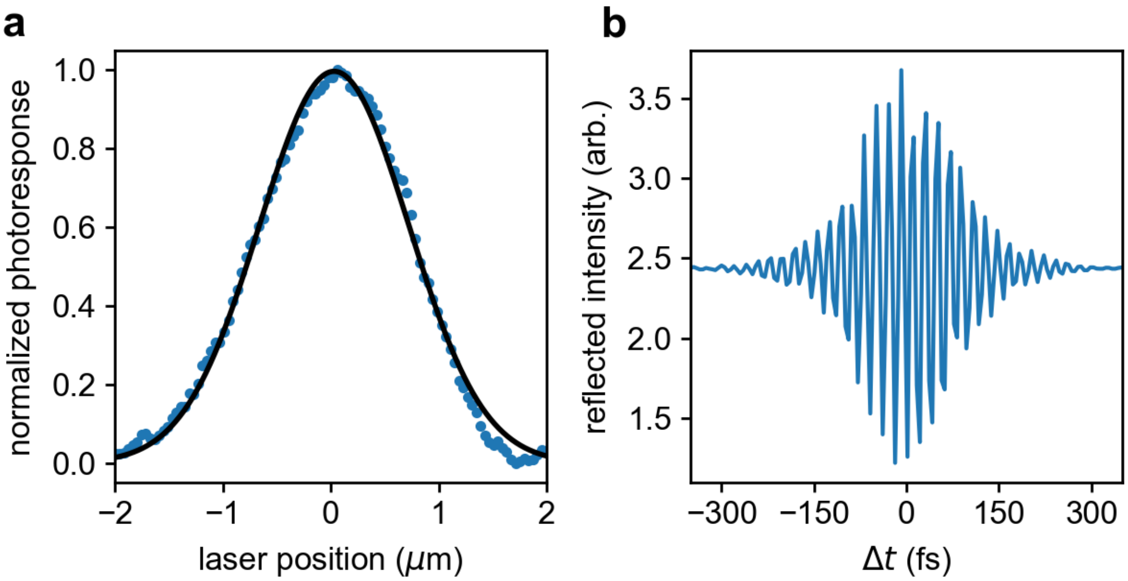


Figure 2.3: Characterization of the ultrafast optics. **a**, photoresponse to an absorber smaller than the diffraction limit. Black line is a fit to a Gaussian function with a full-width-half-max of 1.67μm taken at λ = 1200 nm. **b**, two pulse autocorrelation which is proportional to the pulse width at the sample.

To fully enclose our focusing optics inside the vacuum chamber, we use a Gradient Index of Refraction (GRIN) lens as an objective. A GRIN lens is a single small cylinder of glass with the index of refraction varied radially. Lacking the many interfaces of a conventional objective, a GRIN lens does not disperse laser pulses as dramatically as a traditional objective. Fig. 2.3b shows the autocorrelation of the reflected intensity due to two overlapping laser pulses, near $\Delta t = 0$. The autocorrelation width is approximately three times the pulse width. Our autocorrelation pattern is 570 fs wide, indicating that our pulses are 190 fs long at the sample, only 27% off the 150 fs laser specification. Low dispersion allows us to measure short timescales and gives high peak pulse intensity. However, a GRIN lens also has downsides compared to a traditional objective lens. When well aligned, the power throughput of the GRIN lens is very high, however the process of aligning the optic over the sample under vacuum introduces systematic uncertainty into the laser power. Also, the field of view for a GRIN lens is typically smaller than a traditional objective lens, which is no problem for micron sized samples, but can limit applications in some large area samples. MPDPM can be performed using a traditional objective lens at the cost of increased pulse dispersion and therefore decreased time resolution and peak pulse intensity.

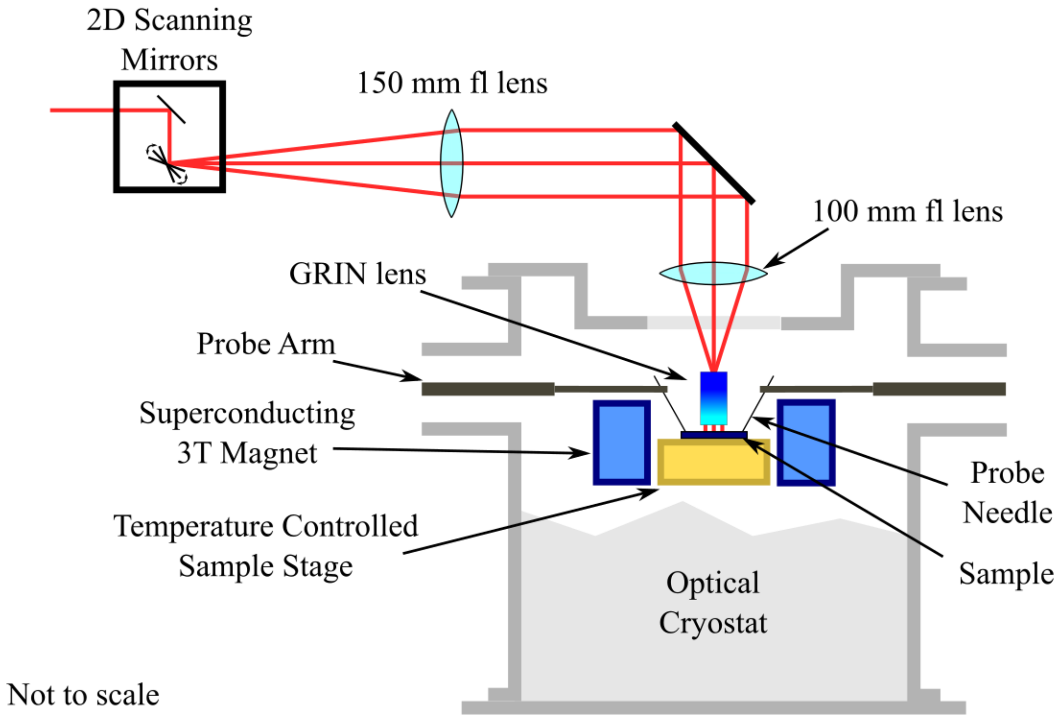


Figure 2.4: Schematic cross section of the scanning optics and the optical cryostat.

2.5 Apparatus and Cryostat

To perform experiments, we used a customized optical cryostat to control the sample environment to the maximum extent practical. Fig. 2.4 schematically details our specific scanning optics and the customized Janis Research ST-3T-2 optical cryostat (pictured in Fig. 2.4a) that we use in our experiments. The sample sits in vacuum on a sample stage, which can controllably vary the temperature from 4 K to 420 K. The sample stage is in the center of a 3 Tesla superconducting magnet. The sample is electronically probed using four probe needles which contact conductive pads on quartz chip carriers that are wire-bonded to fabricated titanium-gold contacts on the sample. There is also a specialized probe-arm that holds the GRIN lens (bonded to a ferrule) on a cantilever near the center of the sample stage, so that it can be focused and re-positioned by moving the probe arm. Fig. 2.4b shows the center of the cryostat, with an example sample on the sample stage, the GRIN lens is illuminated with green light.

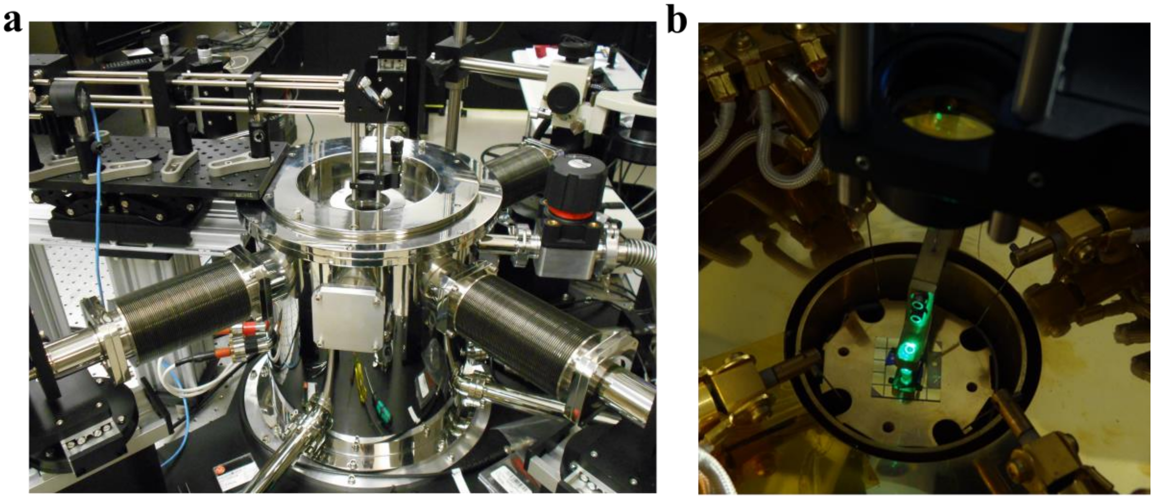


Figure 2.5: Photographs of the optical cryostat. **a** external view and scanning beam optics circa 2015. **b** detail of the sample stage, probe arms and GRIN lens (illuminated).

The sample is electronically contacted through the probe needles that connect to gold pads on the sample. The electrical connections are routed through a switchbox, which grounds all connections to the sample when not in use. We then amplify the electrical signal using a current pre-amplifier, and usually a lock-in amplifier, then measure the current resulting from the incident laser light, or photocurrent. We also measure the reflectance of the sample by measuring the intensity of the light that is reflected from the sample with a near-infrared photodiode. Signal is acquired by a National Instruments PCIe-6323 Data Acquisition card, run with the integrated data acquisition system discussed below.

2.6 Integrated Data Acquisition System

The goal of MPDPM is to time-efficiently sample as large of a parameter space as possible, using as many experimental parameters as are relevant and practical. To do this efficiently requires the ability to optimize the measurement time, as described in equation 1.1. The optics described in section 2.4 are designed to allow fast scanning and other hardware components to be optimized to work as rapidly as possible, decreasing t_h to a lower bound given by maximum hardware speed and amplifier time constants. Furthermore, the high signal-to-noise ratios can minimize t_s to its intrinsic limit. Well-designed optics improve the time efficiency of the experiment "for free." However, the largest increases in efficiency come from the ability to made tradeoffs in resolution. Optimal utilization of the optics and effective tradeoffs requires an integrated Data Acquisition (DAQ) system that automates all hardware components through one program. Such an integrated DAQ can control all hardware components at their optimum, in parallel, with minimal human input. The software allows the experimenter to choose the ranges and resolutions of various parameters in a scan in an intelligent manner, making appropriate tradeoffs. Finally, such a DAQ system allows data to be gathered densely, systematically, and repeatably, in a format that allows advanced data analysis.

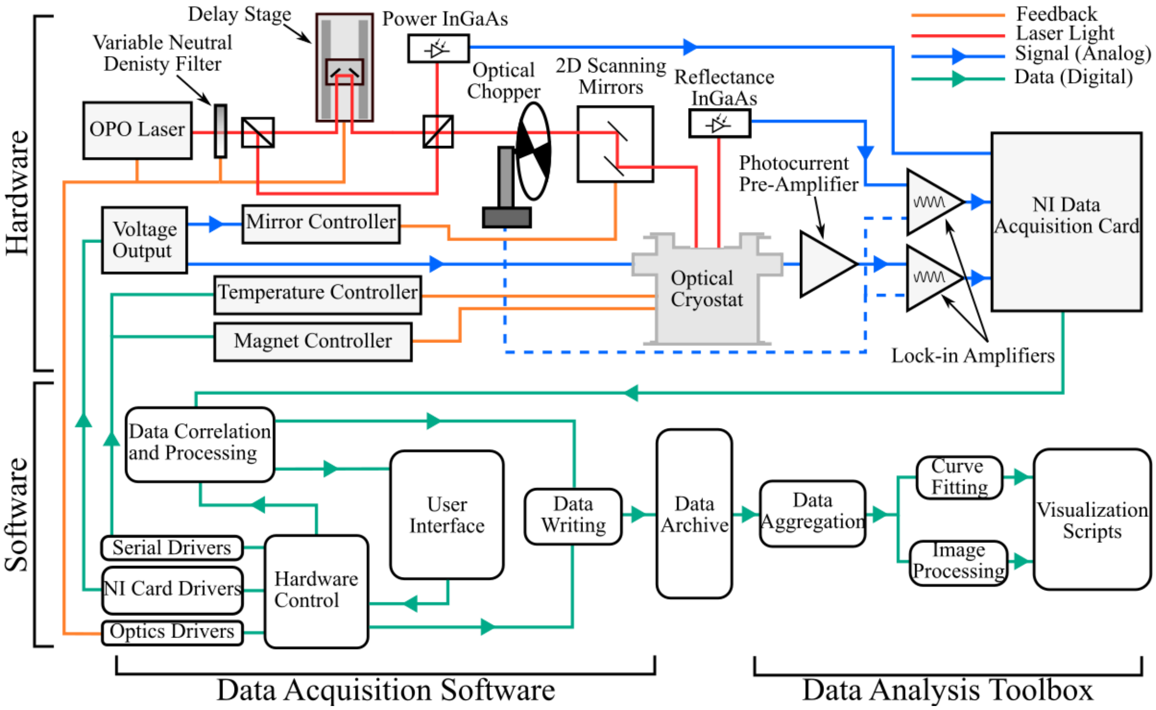


Figure 2.6: Data Flow between the various experimental hardware and software components showing the main functions of the hardware controllers, DAQ software, and Data Analysis Toolbox.

We developed an integrated DAQ program using a set of python modules that interface with equipment drivers and control all hardware components simultaneously with the maximum amount of automation possible. Our experimental setup can scan a beam in two dimensions, while applying voltages to the sample under various optical conditions. In addition, the optical cryostat that contains our samples can control the temperature of the sample and apply a magnetic field. Each of these components requires specialized hardware, which were designed and selected to allow for full automation. The flow of data is shown schematically in Fig. 2.6. The main hardware components of the optics and controllers, shown in the upper left, are controlled with feedback to the DAQ software, which is represented in the lower left. From the user interface, any of the hardware components can be changed or scanned, varying some output over a given range. If one or two of the components is set to scan, the rest will be held constant.

From the user interface the experimenter can define which parameters form the axes of a two-dimensional scan and define the scan's resolution in those parameters. The result is an array of data, or “data plane.” The experimenter can select a third parameter to scan over and the software will take successive 2D scans as a function of that parameter, constructing a 3D “data cube” out of many data planes stacked along the third axis. These data planes or data cubes form a “run”, the discrete unit of MPDPM image data. In addition to the data, each run saves all possible control parameters, on the order of a hundred individual parameters, of the hardware and software to ensure consistency and repeatability. Each run is assigned a unique run number and the files for that run are saved to disk in a data archive. To efficiently take many runs, the software allows the user to repeat a run varying another parameter, taking data cubes as a function of this fourth parameter. Put together, the runs form a four dimensional “data hypercube,” sampling a large volume of parameter space. This allows the experimenter to, with full control over the ranges and resolutions of all parameters, efficiently and fully automatically sample a four-dimensional parameter space. These fully automatic measurements can run for hours or days, collecting data with no human input needed.

2.7 hyperDAQ Software Design

The integrated DAQ system described in Section 2.6 is the key piece of software that enables MPDPM measurements to be performed. While analysis software (discussed in Chapter 3) is also important, it is more flexible and forgiving; analysis can be performed different ways and poorly designed data analysis will hamper the project but can be overcome. This is not true of the DAQ software, if it is not carefully designed MPDPM will not work or will take a prohibitively long time. In MPDPM, the limits of the DAQ software are fundamentally the limits of the experiment, therefore it is worth getting right. This section discusses the design of the DAQ software in detail. It is not presented as it was developed; the DAQ software was developed and improved in many iterations, and two major re-designs, over the course of five years, beginning in 2015 and continuing through 2020. Each time a significant improvement was made to the DAQ software, it opened up new experimental avenues and uses of MPDPM. The latest version of the DAQ software, hyperDAQ version 2.0, is discussed here it it's complete form following an upgrade in January 2020 which took special care to incorporate the lessons learned from previous experience, to make the design robust and extensible and to make it accessible for use in a wide range of experiments.

When designing integrated DAQ software there are several key considerations to ensure the software can meet the demands of MPDPM. Firstly, it must be able to accommodate a large volume of data. What qualifies as a high volume of data depends heavily on a given experiment, but in terms of MPDPM theory this means that the acquisition and processing time of the DAQ software should never significantly increase the total time T, described in equation 2.1. Secondly, the software should be versatile in the manner in which it takes data; in that it can make the necessary tradeoffs to optimize the total experimental time T. Third, the DAQ system needs to be as automated as possible, which will require controlling a variety of disparate equipment. Fourth, the data generated needs to be fully traceable; given the variety of the equipment that may be automatically controlled in MPDPM it is important that all metadata is saved with the data. Finally, the design needs to be extensible, because MPDPM can be used in so many ways, and the software requires careful design, it is necessary to have the DAQ be adaptable to many different types of experiment and when it cannot be adapted, be extended to include new functionality.

At its core, hyperDAQ is a python based object-oriented platform that allows for equipment control and data gathering from a Graphical User Interface (GUI). The hyperDAQ program should be thought of as a library of data acquisition code designed to be expanded and adapted to new experiments rather than a discrete program. The key to understanding how hyperDAQ works is to understand that the program is object oriented, components are objects that pass data and commands between each other. These objects may occur within several threads, each of which performs a key function of the program.

Fig. 2.7 illustrates the key data relationships between the objects. In the figure, the left column illustrates the various hardware control code, which is third party software that controls the hardware at a low level. Most hardware comes with drivers that can be used, but hyperDAQ can also communicate with hardware directly through serial ports. The objects that interface with these drivers (or serial ports) are in the control thread (serial ports may have their own threads however commands related to scanning are still sent from the control thread.) The most important object is the Card Controller which interfaces with the data acquisition card. It executes a scan of the system outputs and queues data from the analog inputs for processing. During a scan, the card control or control thread may use the device controller objects to execute parts of the scan as well.

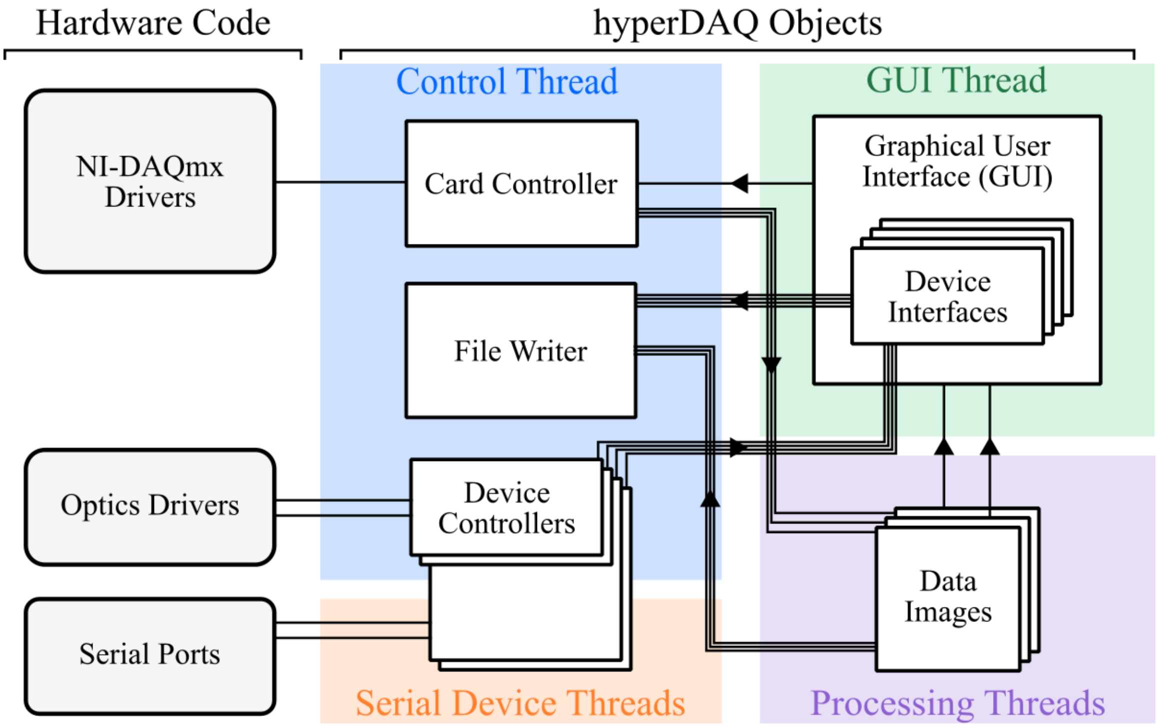
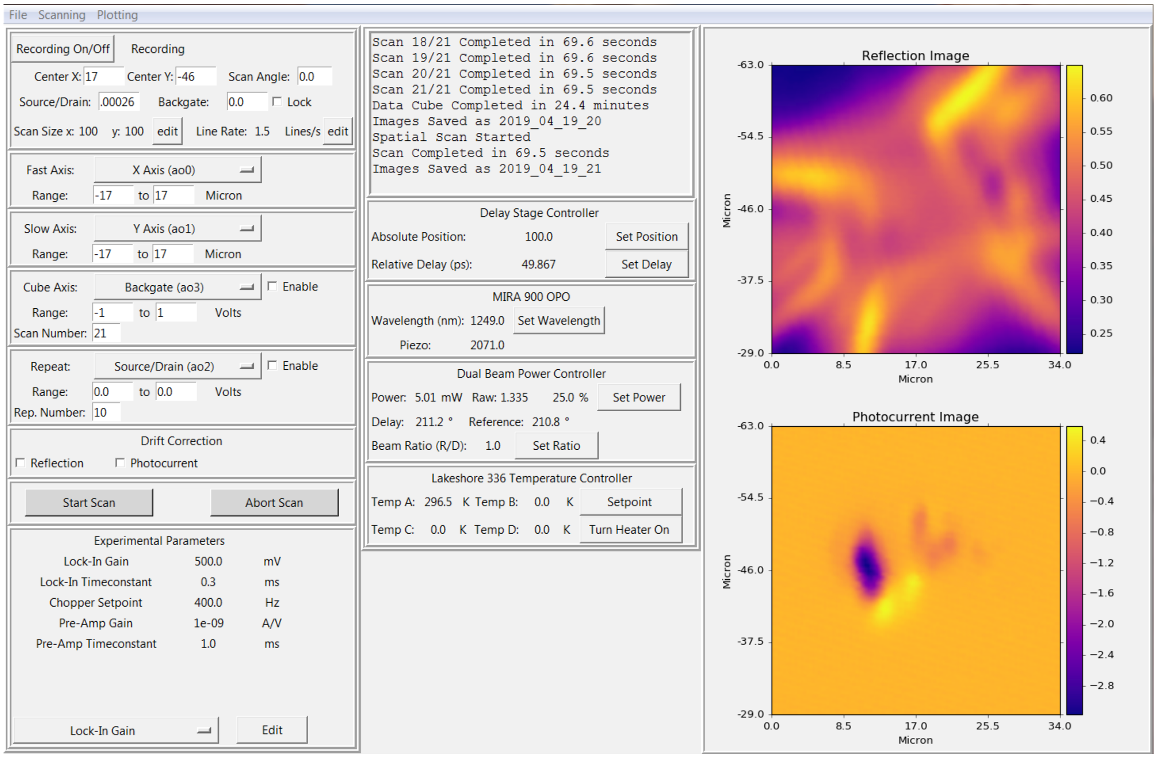


Figure 2.7: Flowchart showing the key data-centric relationships between objects within the hyperDAQ software. Does not fully illustrate all potential connections, only important connections.

Data acquired by the card is passed into the processing thread, where it is processed by the Data Images, which contain the primary imaging data. This data may be displayed on the GUI. When the scan is saved, the data is accessed by the File Writer object. The File Writer not only saves the data from the Data Image objects but also the status of all the hardware within hyperDAQ, the scan specification and all changeable software and equipment parameters to ensure reproducibility.

Finally, there is the GUI Thread which contains a tkinter main loop that displays the data and controls using the python tkinter software. Its main function is to setup scans (in 2, 3 or more dimensions) and display the primary data as it comes in. In addition, for each hardware controller beyond the card there is also an interface, which displays current status and may be used to send commands to the hardware when not scanning. These hardware interfaces are displayed as subframes within the GUI. From the user interface, shown in Fig. 2.8, the user can setup various MPDPM runs, view the status of the equipment and see the imaging data in real time.

Beyond this general data flow, hyperDAQ objects need to be setup in response to the particular needs of a given experiment. To do this hyperDAQ uses a working directory structure as illustrated in Fig. 2.9. The primary hyperDAQ code is contained within the “hyperdaq” subfolder, which is setup as a python package and tracked with the version management software git. The core of a given hyperDAQ program is the Main File, this is the file where all of the various hyperDAQ objects are instantiated and the program is started. The objects are drawn from two sources, either the main hyperDAQ code located within the hyperdaq package or any custom code that the user might like to use. In addition, there is that parameters.py file, this contains all the parameters and other information that hyperDAQ needs to be run a given experimental system and provides important metadata. The main file defines objects for all the hardware that an instance of hyperDAQ controls, and can be modified to add in new equipment and change the configuration and functionality of hyperDAQ. A properly setup hyperDAQ instance then performs scans and when appropriate will save them as files. To ensure that data is easily accessible each scan (or 3D data cube) is assigned a unique run number with a specific form, and all metadata is saved, see section 3.4 for details.



Recording

Center X: 17
Center Y: -46
Scan Angle: 0.0
Source/Drain: .00026
Backgate: 0.0
☐ Lock
Scan Size x: 100
y: 100
Line Rate: 1.5 Lines/s

Fast Axis: X Axis (ao0)
Range: -17 to 17 Micron

Slow Axis: Y Axis (ao1)
Range: -17 to 17 Micron

Cube Axis: Backgate (ao3)
☐ Enable
Range: -1 to 1 Volts
Scan Number: 21

Repeat: Source/Drain (ao2)
☐ Enable
Range: 0.0 to 0.0 Volts
Rep. Number: 10

Drift Correction
☐ Reflection
☐ Photocurrent

Experimental Parameters

Lock-In Gain	500.0	mV
Lock-In Timeconstant	0.3	ms
Chopper Setpoint	400.0	Hz
Pre-Amp Gain	1e-09	A/V
Pre-Amp Timeconstant	1.0	ms

```
Scan 18/21 Completed in 69.6 seconds
Scan 19/21 Completed in 69.6 seconds
Scan 20/21 Completed in 69.5 seconds
Scan 21/21 Completed in 69.5 seconds
Data Cube Completed in 24.4 minutes
Images Saved as 2019_04_19_20
Spatial Scan Started
Scan Completed in 69.5 seconds
Images Saved as 2019_04_19_21
```

Delay Stage Controller
Absolute Position: 100.0
Relative Delay (ps): 49.867

MIRA 900 OPO
Wavelength (nm): 1249.0
Piezo: 2071.0

Dual Beam Power Controller
Power: 5.01 mW Raw: 1.335 25.0 %
Delay: 211.2 ° Reference: 210.8 °
Beam Ratio (R/D): 1.0

Lakeshore 336 Temperature Controller
Temp A: 296.5 K Temp B: 0.0 K
Temp C: 0.0 K Temp D: 0.0 K

Figure 2.8: The user interface of the hyperDAQ software. The left column contains the controls for setting up a customized scan and an area to record information that cannot be automatically recorded. The middle column displays the status of the system, information about the runs it is executing and small interfaces for several external pieces of hardware. The right column displays data from two inputs, in this case the reflection image (top) and the photocurrent.

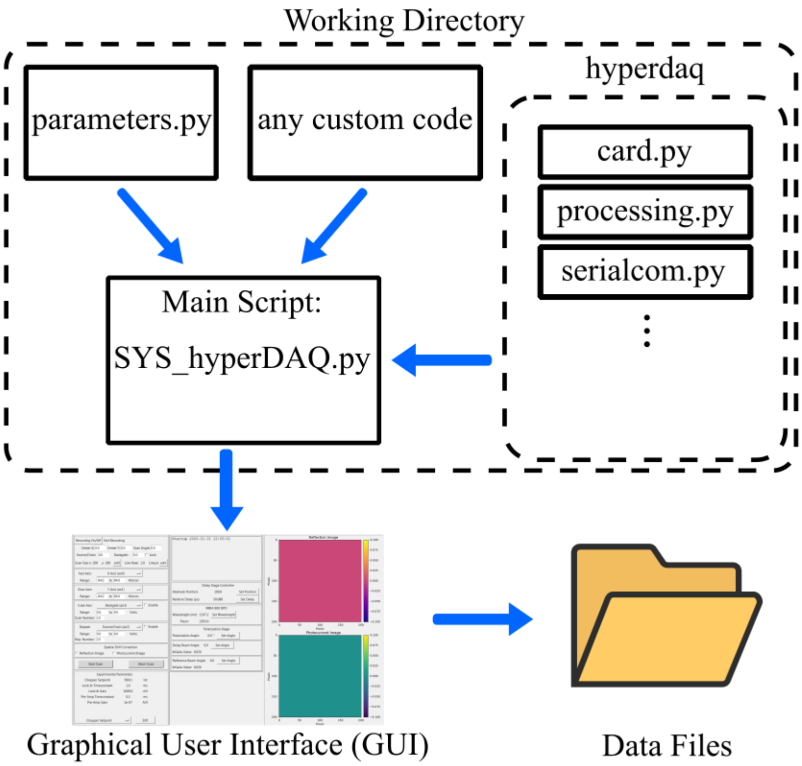


Figure 2.9: The working directory structure of the hyperDAQ software, showing the relationship between the various modules. Dashed boxes indicate directories, and the solid boxes indicate python files.

The hyperDAQ software satisfies all of the designed constraints mentioned above. First, it can scan hardware and acquire and process data at a rate faster that is hardware limited. Second, from the user interface the user can customize a scan to sweep (almost) any of the experimental variables in manner desired at any possible resolution. Furthermore, while hyperDAQ typically performs measurements in a standard recti-linear fashion (e.g. raster scanning with linear sampling) it is designed such that it can be extended to perform non-traditional types of scanning as appropriate. Third, it is designed to automate hardware to the maximum extend feasible so long as the equipment has, at a minimum, a serial port interface. Fourth, it saves all metadata, including the values in the parameter file with the data. Finally, hyperDAQ is specifically designed to work with a wide variety of experiments, and to be simple to add in customized code.

2.8 Long Measurements and System Stability

Given the maximally automated nature of the hyperDAQ and the multi-parameter nature of MPDPM some measurements may span days or weeks running fully automatically and be composed of hundreds of runs. For these longer measurements the main concern is laser stability, as any variation in the optics will cause the image location to drift. To counter this, careful attention must be paid to the alignment of the optics to avoid any geometrical aberration and the laser should be run continuously for several days prior to any long imaging measurement to stabilize its performance. When the system is stable, the main source of drift is thermal expansion of the various optical components; therefore, precautions should be taken to make the temperature and airflow of the lab as consistent as practical during the measurements.

Even with the extra considerations for stability there will inevitably be some drift in the images, which must be corrected for in the data. There are multiple options for drift correction dependent on the particular measurement, for example when the laser power is approximately constant throughout the measurement and the drift is small, the best method is usually to re-orient the data based on the reflection image. To find the correction for a given reflection image we use an algorithm that shifts the image by an integer number of pixels and calculates the difference between that image and the first image in the data set. By brute force, the algorithm determines the shift that produces the minimum difference and uses that as the correction. This allows for correction that is accurate to the pixel resolution.

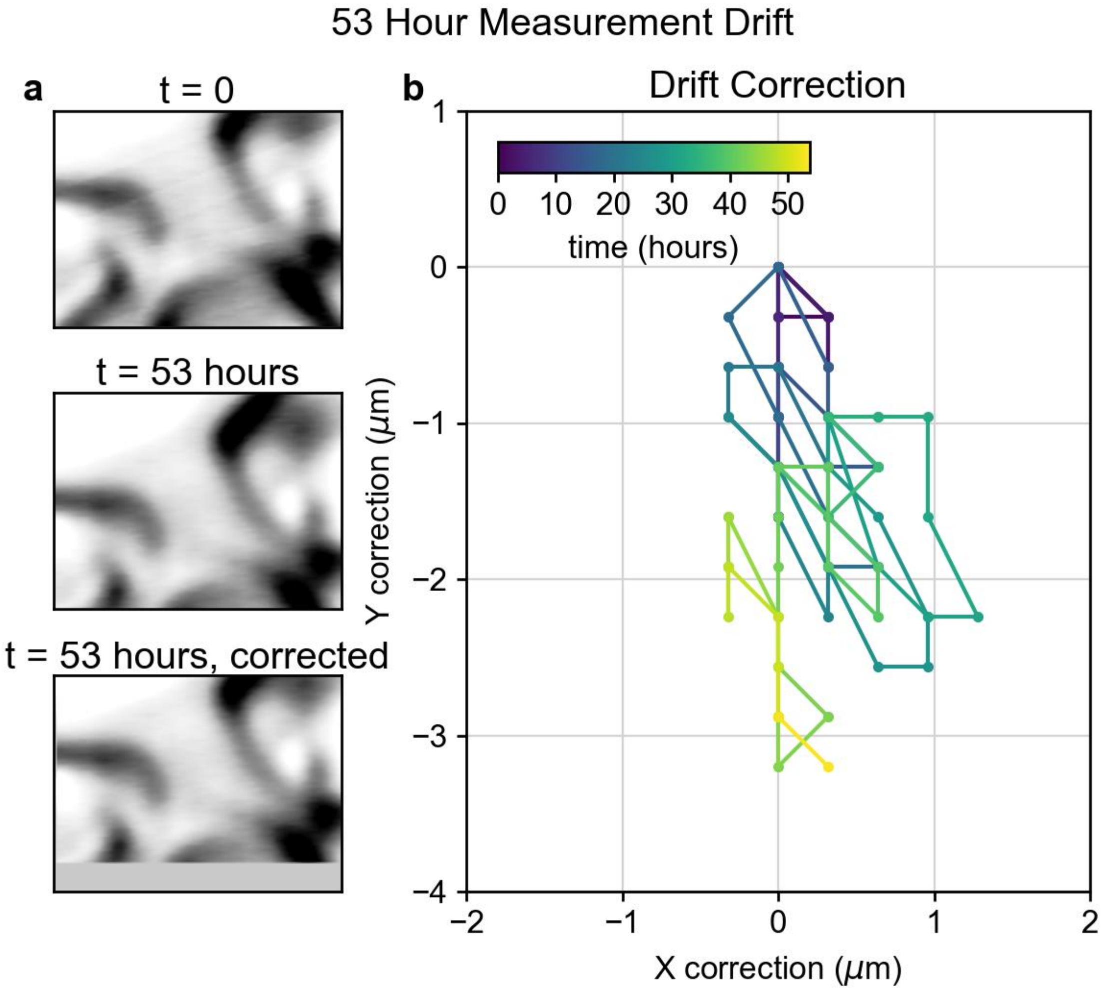


Figure 2.10: Drift correction in a 53-hour imaging measurement. **a** top and middle are reflection images taken first and last; bottom is the last reflection image corrected. **b** shows the applied drift correction (in microns) as a function of time.

As an example, Fig. 2.10 shows the drift correction over time for a 53-hour long imaging measurement taken on the *WSe2MoSe-2* sample. In Fig. 2.10a we see the first and last images in the data set, and then the last image with the correction applied. The correction removes some information near the edges, however the images were taken with a significant margin around the active area of the sample to prevent the drift correction from removing any important information. Fig. 2.10b shows the drift correction as a function of time over the 53 hour span of the measurement, critically we see that the drift random walks, as we would expect from thermally dominated drift, and that over the course of 53 hours it only shifts by about 3 microns (approximately 10 pixels), an acceptable level of drift that can be corrected for without harm to the data.

2.9 Data Taking Process

Each run that is acquired in an MPDPM measurement samples some amount of parameter space. Typically, a single sample will require hundreds of runs to fully examine its parameter space. The most common scan is a rectangular scan of the 2D scanning mirrors, which moves the laser beamspot spatially over the surface of the sample, observing the photoresponse. These spatial scans are designed to be high resolution with variable speed, so that resolution in space can be traded-off for resolution in other variables when needed while still spatially imaging. Depending on the measurement, the laser can be scanned in two dimensions, scanned along a line in a single spatial dimension or held spatially fixed. The automated multi-parameter nature of the process allows for long sets of runs to be measured together, for example a large imaging data set may run for several hours with no human input. Such long runs are usually run overnight, or over a weekend in order to optimize experimenter time.

The process of acquiring data from any given sample goes through several phases. Normally an experiment begins with an exploratory phase, which measures a new sample a set of relatively low-resolution runs over all possible parameters in order to determine the relevant parameters for a sample and the ranges that they vary over. During the exploratory phase it is important to be agnostic with respect to any expectations or hypothesis, so as to not bias the measurement. To this end it is usually helpful to define what range of parameters to explore over prior to the measurement and to stick to that plan during the exploratory phase, however this is not always practical as unexpected results frequently arise. After the exploratory phase has identified the significant behavior, a high-resolution set of data cubes is taken to densely sample the full parameter space, commonly spatial scans as a function of two parameters, generating a large set of images that is usually the main result. Finally, if any unusual or interesting features are seen in that data set, some high-resolution scans are taken to finely characterize those features, often continuing until the sample dies or degrades beyond usefulness. This process does not always proceed linearly in those phases, for example low temperature measurements are often done over a small window of time and after room temperature measurements due to practical concerns over the cost and difficulty of using cryogen.

During the data taking process, the workflow is driven by on the fly data analysis. When a run is taken it is assigned a run number and saved with meta-data which is all linked to the run number. Data is taken and recorded in the lab notebook and visualized using a set of visualization scripts that correspond to each type of measurement. These visualizations are also included in the digital lab notebook and are used to direct the course of data taking as new features are noticed. As with the overall analysis process, it is important to not rely on any one type of visualization as they all have advantages, disadvantages and potential biases. The process of handling, analyzing and visualizing this data is described in Chapter 3.

CHAPTER 3:

DATA HANDLING, ANALYSIS, AND VISUALIZATION

3.1 The Data Analysis Process

The data is only as good as what you can do with it. This simple concept takes on a special meaning when it comes to data intensive measurements, not only does the complexity of the data analysis increase there are a larger variety of pitfalls that can distract, obscure or confuse the interpretation of the data. This chapter describes all that is involved in turning raw data from a data intensive MPDPM measurement into a useable result. There are three primary steps to this process: handling, analysis and visualization. Data handling is the process of making the data available in useable form, analysis is the process of extracting information from the data, and visualization is the process of putting that information is a visual medium suited for human understanding.

In this chapter we discuss each step, starting with considerations that must be satisfied, then turn to a practical discussion of how to perform each step, along with discussion of problems that may arise at the scale of an MPDPM measurement. The concepts and process described in this chapter will be used implicitly and explicitly throughout all the discussion of the data and physics in later chapters. Example data will be taken from several different measurements that will be discussed later on.

3.2 Data Handling Concerns

The MPDPM technique is a powerful tool for taking data in versatile ways, at high volume, using a variety of disparate equipment. However, the experimental flexibility, the scale of data, and variety of the equipment involved demands careful handling of the data. Data Handling is a catch-all term for all the steps that are needed to ensure that data is useable. In experimental physics this can mean a variety of things, but in terms of “tabletop” experiments, i.e. those performed on the level of a single research group, the main concerns are as follows. *Retrieval*; data must be stored in a way where it can easily and efficiently be accessed. *Replicability*; steps must be taken to ensure that any experiment could, theoretically at least, be replicated using the same or similar equipment. *Calibration*; data must be turned from hardware values to real units. This often involves generating calibration data which must also be well documented and easily accessible. *Processing*; data must often be manipulated before it is useful, which can take many forms, for example, filtering out noise or correlating image sets. *Packaging;* if a measurement that is composed of many runs, the runs must be grouped together into a single unit, or package, for analysis and visualization. Care must be taken that the different runs can be directly compared. *Documentation*; all experiments need to be well documented, with metadata that allows any data referenced in the notebook to be easily accessed. *Traceability*; All the above steps must themselves be well documented such that not only can the data be replicated, but the entire process used to develop a result or visualization can be replicated if needed.

MPDPM works by transforming a problem of experimental phase space into a problem of data analysis. Therefore, there are some significant data handling considerations specific to an MPDPM measurement, as follows. Experiments usually involve multiple hardware sources, or even multiple experimental setups and a project will span many different measurements. The geometric nature of the phase space sampling is important in MPDPM data analysis and must be preserved. Measurements are often handled “on-the-fly” during the limited sample lifetime. Furthermore, inconsistent data handling or ill-considered visualizations can lead to confusion and wasted time. Data analysis often involves multiple passes through the data with different types of visualization scripts which may occur months or years after the measurement.

Overall, data handling practices break down into two categories: implicit and explicit data handling. Implicit data handling is all the steps of the data handling process that occur “behind the scenes” and work best when they are automated with well documented and traceable code so that the user does not have to modify them very often, thus reducing the chance for errors. Explicit data handling is all the data handling that is specific to a given experiment and needs to be modified by the user as they go through the data analysis and visualization process. This will usually be handled through scripting and will overlap with the visualization procedures discussed below.

3.3 Data Handling Policy

To satisfy all the considerations in the previous section, it is important that good data handling practices are used when working with MPDPM data. It is essential that experiments are well documented and have a carefully considered data handling policy that is applied consistently. The section details the data handling policy developed for the Gabor Lab. While the general practices have been followed for essentially all of the research discussed in this dissertation, the policy did not exist as currently constituted during much of the research. The policy is presented in its finished form and any significant deviations from it will be noted when data is presented.

The data handling policy is as follows, with the major principles in italics. *Traceable, well tested code, such as the Gabor Lab Toolbox, should be used to perform all implicit data handling.* Any deviations from the established data handling code must be well documented. Code that is developed to process data in a new way must be documented as well. The data handling components of the Gabor Lab Toolbox are specifically designed for this and should be used whenever practical (see section 3.5 for information on the Gabor Lab toolbox). *Calibration data needs to be traceable.* Any calibration that is applied to the data needs to be kept, organized so that it can be applied consistently, and the code that applies it should be traceable. *Data sets must retain metadata throughout the data analysis process.* For hyperDAQ measurements this means that the run number is used to refer to the data and all the information in the log file is retained. When data from multiple runs is packaged together to generate a data set, none of the metadata should be lost (through redundant metadata may be consolidated). *A digital lab notebook needs to document all significant data or analysis.* Any significant analysis that manipulates the data in a novel way should be similarly documented. *Scripts where data is processed and visualized in any nonstandard way need to be retained*. Code that significantly manipulates the data or performs any step in the data handling process should never disappear.

The data handling policy results in the following workflow. Each project has a project folder that contains all the data analysis and visualization scripts for a given project, this folder should have a descriptive name and the code in it should be retained. When data is acquired it is documented with its metadata in a digital lab notebook which records both the narrative of how the data was taken and includes basic visualizations. These visualizations often change frequently during data acquisition, as the data is looked at in new ways. Therefore, the run number and all relevant variables are directly included in the visualization, in addition data is normally accompanied by written descriptions of how the data was taken and why, what the visualization shows, and what the visualization tells you. The principle being that just from the notebook someone else should be able to fully understand and recreate a given experiment.

3.4 Experimental Metadata and Traceability

Metadata is information about a given piece of data, in MPDPM measurements taken with the hyperDAQ system there are two main pieces of metadata, a run number and a log file. To ensure it is easily accessible each scan (or data cube) is assigned a unique run number with a specific form, containing several strings of information separated by underscores, for example, SYS_2019_10_02_15. The first string is the "System Prefix" which identifies the experimental system where the scan was performed. The next three strings are the date on which the scan was taken as YEAR_MONTH_DAY and finally the scan number is indicated. For example, a run number of CBA_2020_01_23_32 was the 32nd scan taken by the "CBA" system on January 23rd, 2020.

For a given run number several files will be saved. A data file will be generated for each Data Image object defined within hyperDAQ and a log file will be generated containing all the scan information, including the status of all hardware and the contents of parameters.py, to ensure replicability. Each file is saved under a file name that is the run number and one additional string of information; their "extension" indicating what they are. For example, a 2D scan with two data images designated 'pci' and 'rfi' (standard hyperDAQ conventions for photocurrent and reflection data respectively) would generate three files: CBA_2020_01_23_32_pci.dat, CBA_2020_01_23_32_rfi.dat, CBA_2020_01_23_32_log.log, where the ".dat" files are numerical arrays of data and the ".log" file is the log file containing the scan information. In addition, to keep it organized and easy to find from the run number, the hyperDAQ software will save these files in subfolders organized by date, for example the above example file would be located in the folder DATADIR/2020/2020_01/2020_01_23.

3.5 Software Toolbox

The nice thing about implicit data handling is that it is largely automatic, the hyperDAQ software generates and data and packages it with the metadata, and code can be written to load the data consistent with the data handling policy. So long as an experimentalist uses the code consistently and correctly most of the implicit data handling will be followed. In addition to hyperDAQ, the main tools for implicit data handling are contained in the Gabor Lab Toolbox.

The Gabor Lab Toolbox (GLT) is a package of code that contains useful functions and classes for working with data. This includes data handling code, tools for making visualizations and a variety of other useful functions. Furthermore, the GLT was written to give a consistent codebase for the Gabor Lab and act as a jumping off point for lab members to develop their own custom toolboxes, containing code that they use frequently that adds onto the functionality in the GLT. The custom toolbox has much the same structure discussed below and mainly adds functionality on top of the GLT code. As with the data acquisition code, the toolbox was developed in stages over several years, and is discussed here in its completed form.

The GLT is divided up into modules. The GLT modules for data handling are: gaborlab.mpdpm, gaborlab.calibration, and gaborlab.utilities. Most of the time, these modules will automatically perform the implicit data handling and all that will be required in a visualization script is to load data using a Run or DataSet object from gaborlab.mpdpm. The GLT modules for visualization of data are: gaborlab.display, which contains functions that assist in generating visualizations and gaborlab.processing, which contains functions that filter or transform data in various useful ways. Finally, the GLT modules for the analysis and modeling of the data are gaborlab.fitting, gaborlab.math, and gaborlab.physics. Where gaborlab.fitting contains code for fitting the data, gaborlab.math contains a variety of useful mathematical code and gaborlab.physics contains various physical constants and functions useful for model building.

Data runs are the base unit of MPDPM data, and the data is packaged into runs by hyperDAQ, and a run number and log file provide the metadata for a given run. To keep all of the data and metadata together, and be consonant with pythons object-oriented design, the GLT treats a single run as an object. Each Run object loads the data on instantiation, and it can be accessed through several key attributes of a Run object: Run.log: a dictionary of all the values in the log file. Run.axes: a list of numpy arrays giving all the axes values. Run.units: a list of the units (as strings) of all the axes. Run.data: a dictionary of the data images. Finally, Run.shape: the shape of the data images. Runs also contain several functions to make the calibration, processing, re-scaling and stabilization of the data simple and efficient. The Run objects handle the implicit data handling and provide the basis for the complex data analysis and visualization discussed below.

3.6 Hierarchical Analysis of MPDPM Data Sets

MPDPM generates large sets of images varying across several experimental variables, requiring sophisticated analysis to extract and visualize results. While the analysis of these sets will vary based on the specific sample, in this section we will give a general procedure to hierarchically exploit data geometry in order to condense a multivariate data set down to a manageable amount of processed data. Fig. 3.1 illustrates this process for a hypothetical four-dimensional (hypercubic) data set. The raw data is a set of datacubes spanning three dimensions $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, incremented along a fourth dimension $\boldsymbol{e}_4$. Each datacube is processed to map out a dynamical parameter that represents the behavior of the datacube along one axis (in this case $\boldsymbol{e}_4$). There are multiple possible projections and representations, and although not all are useful the possibility space should be explored. Image analysis is used on the dynamical maps to identify key features that are then collected into a single visualization. In this hypothetical case ellipses enclosing the “bright” photoresponse are visualized as contours. Ideally, this visualization will represent the evolution of some physically interesting quantity within the four-dimensional parameter space.

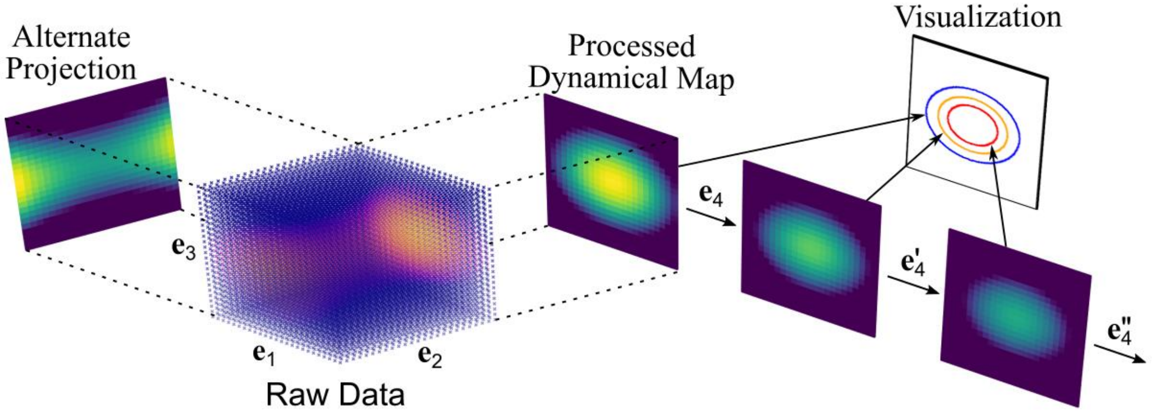


Figure 3.1: Schematic of the hierarchical data analysis of a hypothetical four-dimensional data set.

To illustrate this process in a van der Waals heterostructure device, Fig. 3.2 presents data and analysis from a graphene on boron nitride on graphite (*GBNGr*) stacked heterostructure photocell, detailed in Appendix A1.2. When photoexcited, a Fermi-Dirac distribution of hot carriers rapidly forms in the graphene layer and the exponential tail of this hot distribution may extend into the valence band of the boron nitride, resulting in interlayer photocurrent between the graphene and graphite (this physics is discussed in detail in Chapter 5). To measure this interlayer photocurrent, the graphene contacts were set at a fixed voltage, and current was collected and measured from the graphite

Using MPDPM on the *GBNGr* sample, we obtain an imaging data set that is a good example of the hierarchical analysis: photocurrent data cubes composed of 25 spatial scans at varying laser power, repeated as a function of voltage (applied to the top graphene) in 2 mV increments from -20 mV to 30 mV, for a total of 625 spatial photocurrent images, sampling a four dimensional parameter space (two spatial dimensions, laser power and voltage). Following the general procedure, we will condense the data by fitting it to a phenomenological power law that describes the sample's behavior, identify a physically interesting nodal feature in the resulting non-linearity maps, and visualize the sample's behavior by tracking that node as a function of space. The *GBNGr* data is instructive because it has very distinct features in the non-linearity dynamical parameter, making analysis straightforward, but this approach can easily be adapted to other experiments and MPDPM data sets without such clear spatial features.

3.7 Dynamical Fitting

After the data set is acquired and processed, the next step is to extract fitting parameters that can represent the dynamics occurring in the system. The data points are fit to a phenomenological law using a non-linear least squares fitting algorithm. The phenomenological law can be any function that parameterizes the data well. For photocurrent systems, we most commonly we use equations $I \propto P^{\gamma}$ and $I \propto e^{-\Delta t/\tau}$, for the photocurrent (I), versus laser power (P) or versus two pulse delay Δt, respectively. Phenomenological parameters, such as γ and τ, are extracted from these curve fits. These parameters should be dynamical quantities, so that they can represent changes in the underlying physics. For example, γ is related to the non-linearity of the photoresponse, similarly τ is the characteristic timescale of a process. Changes to γ or τ indicate a change to the character of the photoresponse, not simply a re-scaling of the data, making these parameters very useful proxies for the underlying physical phenomena. Important considerations when selecting dynamical a function and parameter are discussed in detail in section 3.11 along with the process for fitting.

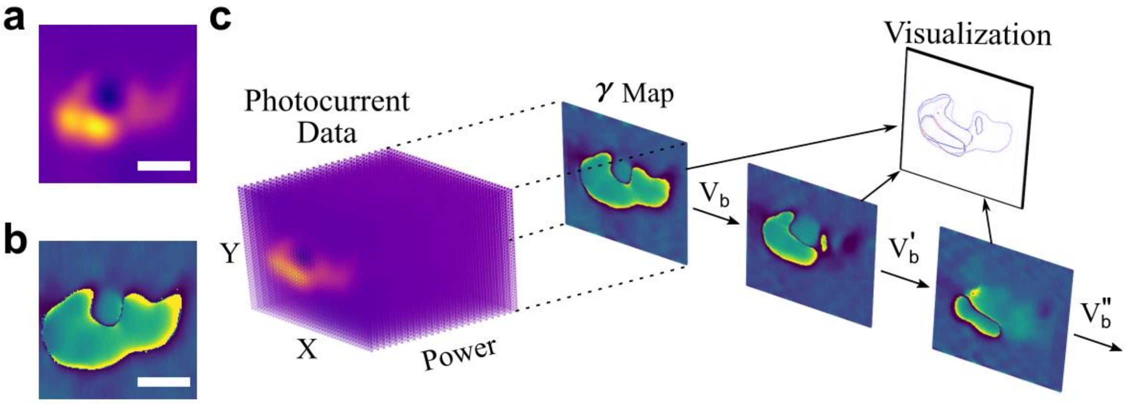


Figure 3.2: Hierarchical analysis of a set of datacubes from the *GBNGr* dataset. **a**, shows an example photocurrent image and **b** shows the γ map from a single datacube. **c**, is a schematic showing how the data is fit and processed into a workable visualization.

The dynamical fitting parameters are used to condense the data. For the *GBNGr* sample the raw data consists of a set of photocurrent images as a function of power, one such image is shown as a colormap in Fig. 3.2a. These images are correlated together, then the data at each point in space is curve fit along the power axis to the power law $I \propto P^{\gamma}$. The power law describes the data well in this case, and the parameter γ, acts as an index of the non-linearity, a useful dynamical quantity. The fitting gives a map of the fit parameter γ as a function of space, such as that shown in Fig. 3.2b. The processed γ image condenses the dynamics of the whole three-dimensional set of photocurrent images into a two-dimensional map. The entire *GBNGr* data set is four dimensional, with data cubes taken as a function of laser power at various values of applied voltage. Fig. 3.2c shows how the data set is processed, all of the data cubes in the set are processed into γ images, giving a three-dimensional set of γ images representing the sample non-linearity as a function of voltage. The resulting set of γ maps can then be analyzed using image analysis to condense them into a single visualization.

3.8 Image Analysis

Once processed, image analysis is used to identify, and algorithmically extract, physically interesting features from the processed images. Identified features can be projected onto the spatial axes (or taken as a function of some other variable). This further reduces the dimensionality, usually giving a result that is visualizable as data mapped in space, or even as a function of a single variable, which human intuition is more suited to handle. The algorithm used to perform image analysis is the most application specific component of the process, as the ability to quantitatively pick a feature out of an image depends highly on what features are present. However, there are many well established image processing algorithms, and a researcher with a solid foundation in programming and signal processing should be able to find a solution without much trouble.

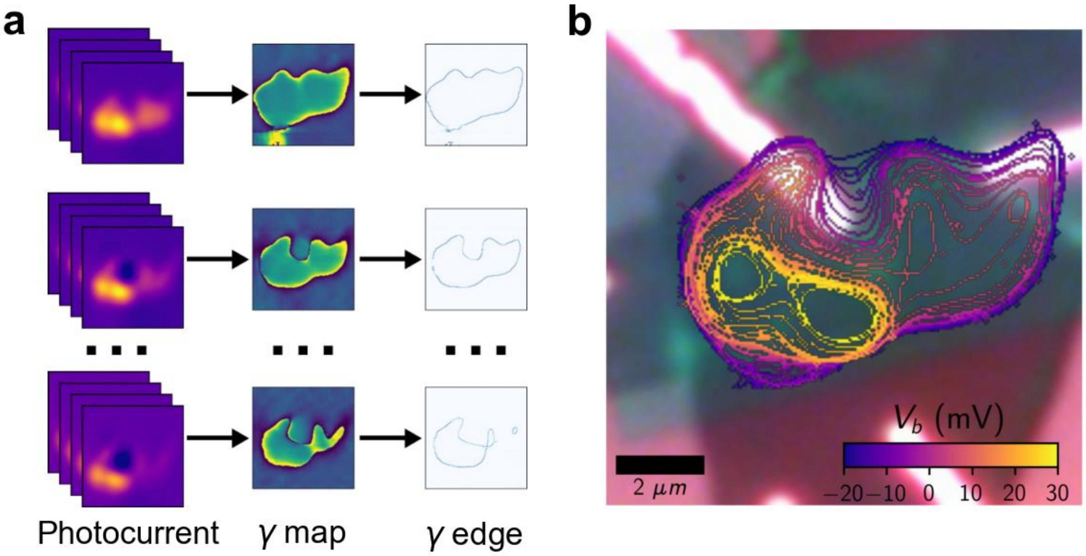


Figure 3.3: Schematic of image analysis for the *GBNGr* dataset. **a**, shows the processing from photocurrent maps to γ map representing a single datacube, then the edge features is isolated from each γ map. **b**, The result of the image analysis, the edge maps overlaid on an image of the sample, physics can be determined from this visualization by interpreting the edge features as a node in the sample electrochemical potential.

The *GBNGr* data provides a clear example of how to use image analysis to identify interesting features from image data. In the processed γ maps there are distinct regions with different γ values. The higher values of γ, (the green and yellow areas on Fig. 3.2b) are separated from the lower values of γ and the background (blue and dark purple areas) by a sharp boundary. The boundary is a physically interesting feature, because it indicates a node in the photocurrent versus power, which evolves as a function of applied voltage due to the internal electronic properties of the sample. We use a Laplace filter, a common image processing filter used for edge detection, to identify this feature. This is performed on each γ map at different values of applied voltage. The image analysis process is shown schematically in Fig. 3.3a, the raw datacubes yield maps of γ condensing the four-dimensional data set down into a three-dimensional data set.

Then the edge feature is extracted from each γ map, forming a highly condensed set of images showing only the feature of interest.

From the condensed data, which can be correlated to physical features of the sample, we can now develop an interpretation of the MPDPM data set. Fig. 3.3b shows the node versus applied voltage overlaid on an optical image of the *GBNGr* heterostructure sample. The edge is a node in the photocurrent, implying that charges excited at that location do not experience any force that would drive a current. This means that, on the node, the internal electrochemical potential of the sample is zero. Fig. 3.3b shows how the internal electrochemical potential of the sample is modified by an externally applied electric field. Of particular note, is the dipole-like feature in the top center of the nodal pattern which lies on top of an electrically floating metal contact. It has been predicted that a floating contact would modify the internal potential of graphene in a dipole pattern[46]. Similar MPDPM measurements can be used to further image the internal potential of conductors[47].

It would have been difficult to observe this data without using MPDPM. No single image, or dependence of a single parameter, contains a clear experimental signature of changing electrochemical potential. Only by sampling several experimental variables, observing the changing dynamics, and picking the right feature out of the complex photoresponse, could we identify this. In addition, there is no reason that the γ node was the only interesting feature in the data set.

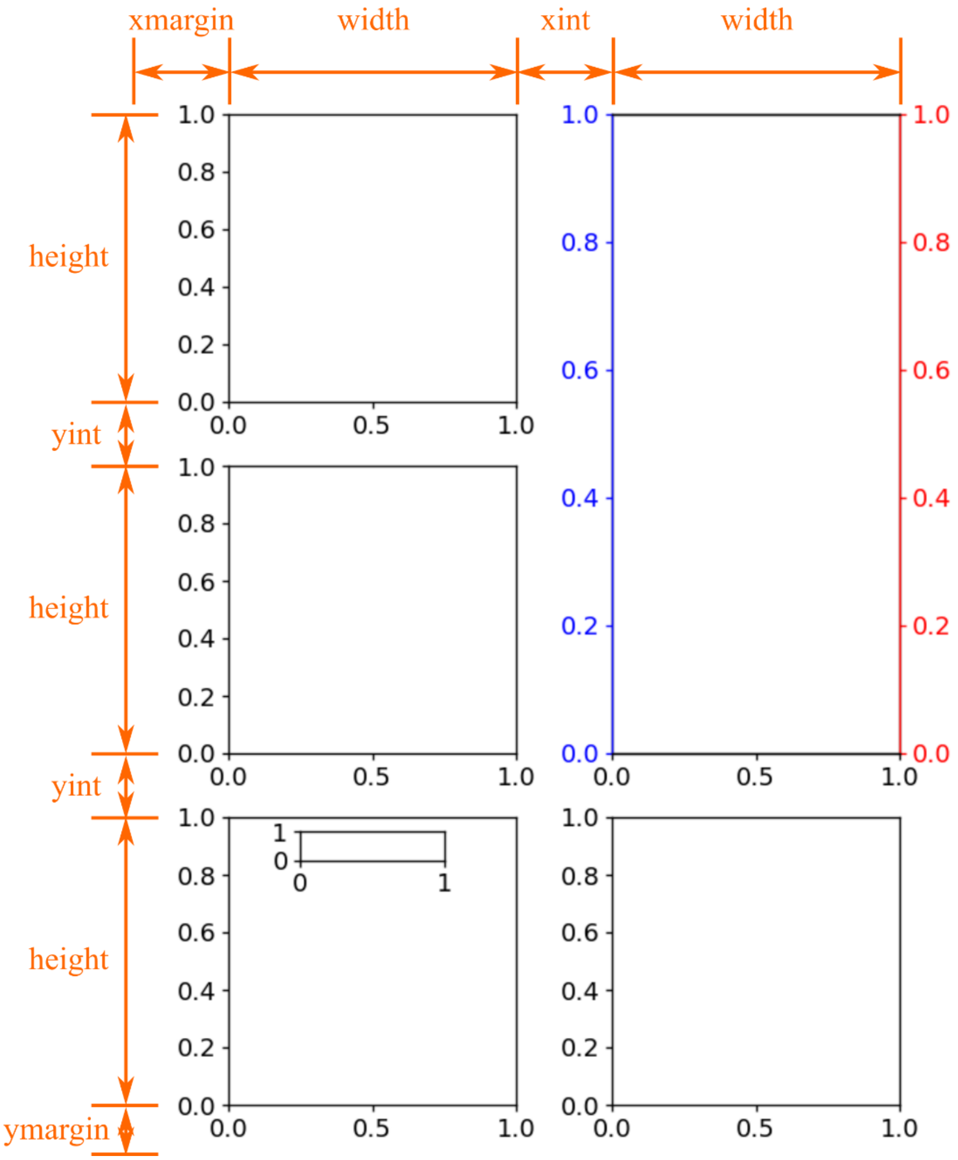


Figure 3.4: Diagram of the composition of a visualization script using the display.figure_inches tool, where the orange text are quantities passed to the tool to layout the figure in a cartesian grid, simple geometry can then be used for the layout.

3.9 Visualization Scripting

Visualization scripts are used to load, process and visualize the data, usually performing the explicit data handling. The basic structure of a visualization script is as follows. First relevant code is imported, then the data is loaded using a Run object from gaborlab.mpdpm, and the data attributes are given readable names. Then, the figure is defined and composed using code from gaborlab.display. Lastly the data is displayed, using some code from gaborlab.display but mostly standard matplotlib commands. Most of the display and axes manipulations are simply done with matplotlib, however certain matplotlib functions are used frequently enough a certain specific way that they can are encapsulated into functions in gaborlab.display for simplicity.

The most useful feature of gaborlab.display is the class display.figure_inches which assists in the composition of figures (that is, laying out the figures is a useful manner). Composition is a major area where someone can waste a lot of time and still end up with a bad visualization. The display.figure_inches object was designed to make composition easier by taking a geometric approach to the layout of figures. It generates axes (of various types) based on standard cartesian coordinates, allowing the user to use geometry to layout their figure. From the figure_inches object each call to figure_inches.make_axes creates a new panel with a geometric argument [left, bottom, width, height]. In the script various quantities of the grid, such as the width and height of the panels, the margins, and the intervals between panels are defined. Then, by adding together the quantities in the right way they are laid out in a grid. Fig. 3.4 shows an example of what can be generated with this tool, diagramming in orange the various geometric quantities that are added together to make coordinates. On top of a basic grid, various more complicated layouts can be made, such as creating an axis that spans two grid spaces, or make that axis have two y-axes (colored blue and red). A large variety of figure layouts can be accomplished with this simple grid style, and other geometries can be used if they can be defined using coordinates.

3.10 Visualization of Multi-Dimensional Data

One of the major challenges of MPDPM is how to visualize multidimensional data. Human vision and intuition are geared towards two dimensional depictions of data, which is best with one dependent and one independent variable[48-50]. MPDPM data is often multi-dimensional meaning it has two or more independent variables, complicating visualization. Developing a consistent way to visualize the data can prevent a researcher from becoming overwhelmed by the volume of data and provide a platform for deeper forms of data analysis, but care must be taken as some visualizations can inhibit understanding. Many possibilities should be explored and carefully considering how to present data in a presentation or paper, however that occurs after much consideration and with some understanding of what the data means. Such polished visualizations are usually a waste of time during the data acquisition and analysis process where scripting must be flexible. This discussion will focus on how to display multi-dimensional data in an exploratory context, both as the data is coming in and during initial analysis.

The use of matplotlib's colorscales in MPDPM is a good example of how the limitations of human perception influences visualization. Colorscales are important when working with images but, when looking at small differences, color is one of the worst ways to represent data because the complexities of how human perceive color can lead to perceived differences where there are none[51,52]. Human perception of color is neither uniform nor consistent, meaning that the perceived difference between two points depend heavily on what colorscale is used and perception of color is influenced by the colors of areas around them. If carelessly visualized, a researcher combing through a large set of images may waste time pursuing differences in contrast that appear to be significant but aren't. Thus, the selection of a colorscale is never an insignificant choice.

There are multiple options for colorscales that will display data while minimizing the adverse effects of human color perception, Fig. 3.5 shows several colorscales. Fig. 3.5a is the classic “rainbow” colorscale, simply to illustrate the problems inherent in such a colorscale. Firstly, compared to the others significantly less detail can be seen. Secondly, the rainbow colorscale seems to form rings around the prominent photocurrent areas, this banding is not due to any feature of the data but rather the non-linearity of human color perception. It is clear that such a colorscale should not be used.

A marginally better colorscale is a simple sequential colorscale, like that shown in Fig. 3.5b, which is a single color with varied lightness. The monochromatic nature reduces issues with color perception but limits the amount of available contrast. An alternative that uses more colors is a perceptually uniform colorscale, such as that shown in the Fig. 3.5c, much more detail can be seen because this scale is engineered to smooth out the non-linearities in human color perception allowing it to safely span many colors, thus increasing its contrast. Another option is to notice that the photocurrent shown is largely organized into lobes of positive and negative photocurrent and select a diverging colorscale that goes between two colors and is white in the middle, such as Fig. 3.5d. This is probably the best colormap for this type of photocurrent but may not be applicable to all situations. It is also important to note that to make a diverging colorscale work it is usually necessary to guarantee that zero is in the center of the colorscale. No matter which type of colorscale is used they all have difficulty discerning small differences in the data and transitions between types of behavior. Therefore, there is a rule of thumb when using colormaps to view data: if an interesting effect in the data is not obvious from a colormap, a colormap should not be the only visualization used.

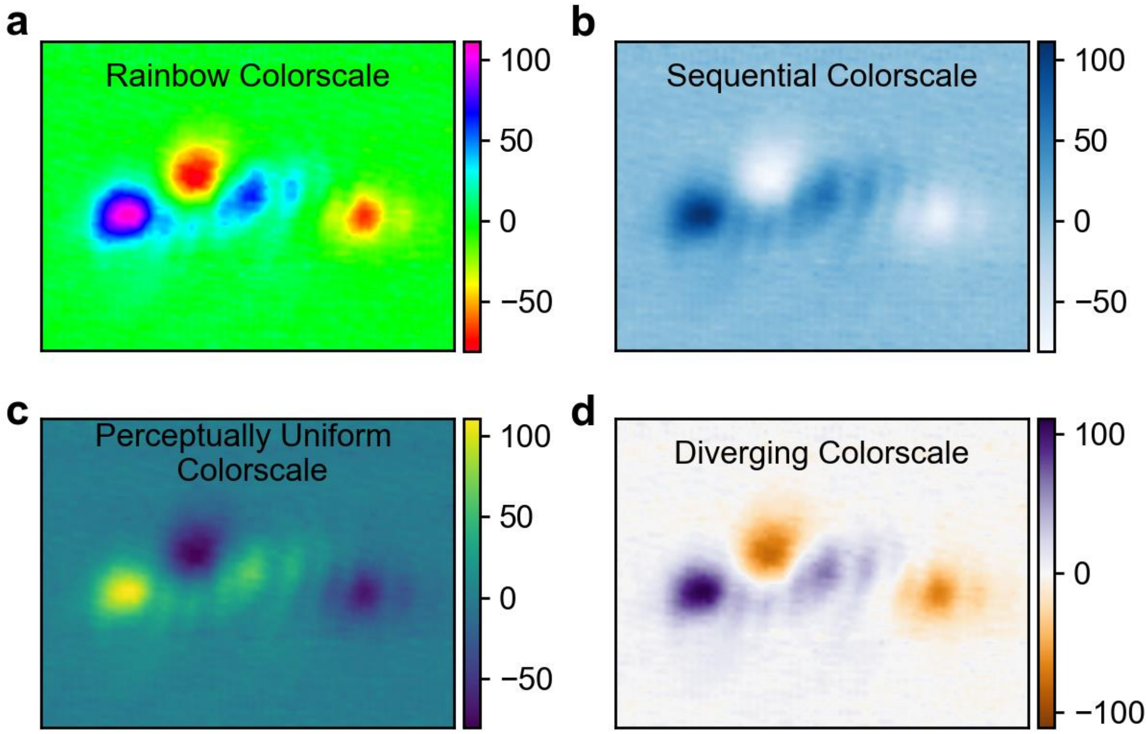


Figure 3.5: Comparison several different types of colorscale, all showing the same data. **a** is the rainbow colorscale, to demonstrate the problems associated with colorscales. **b** a single color sequential colorscale. **c** a perceptually uniform colorscale. **d** a diverging sequential colorscale.

There are numerous ways to represent two-dimensional data that aren't colormaps. A good example is a waterfall line plot, where data is represented as a series of dense colored line cuts, and leverage the human perception of distance, which is much more consistent than the human perception of color. Fig. 3.6 compares a colorscale and a waterfall line plot, the colorscale only really shows the behavior at large voltage and powers, features near zero voltage are washed out. Therefore, in the right panel we plot horizontal line cuts with power represented as color, from this we can see the nuances of the low voltage behavior. Another option for representing the data is to use a fit, provided you know what the dependence of the data is, fitting will be discussed in section 3.11.

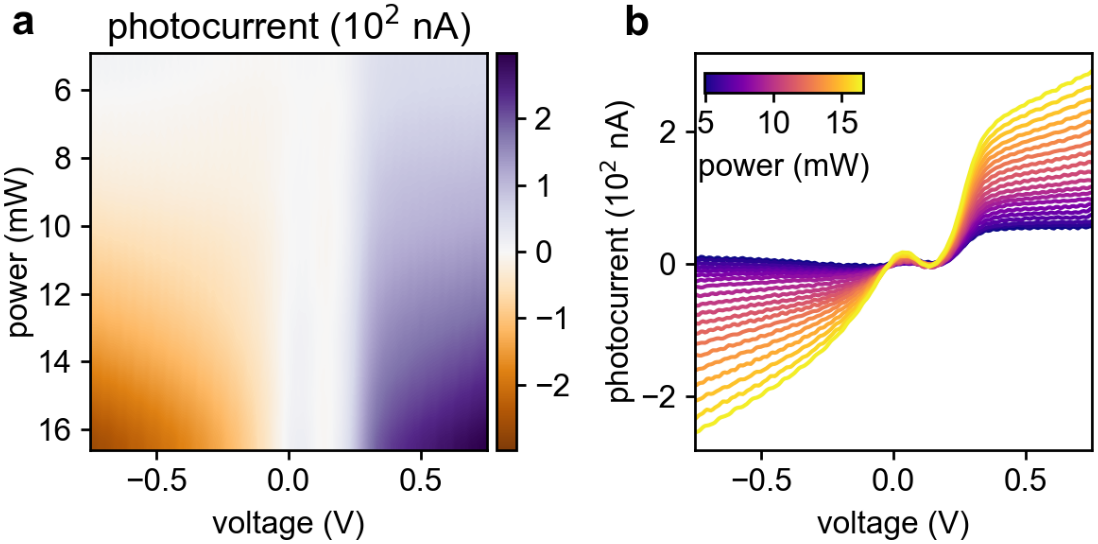


Figure 3.6: Comparison of a colormap, **a**, to a waterfall line plot, **b** for the same data.

Representing three-dimensional data sets (e.g. data cubes or series of scans) adds an additional level of complexity, given that it becomes difficult to view all the data at once (at least in high resolution). There are several options for representing this data, all of which have advantages and disadvantages. The most simple and intuitive way to represent a data cube is a tile plot, which simply displays all of the component images in a series, an example of this representation is Fig. 3.7 which shows all the component images of a datacube. This sort of visualization displays all the data and is very useful when looking for spatial changes or when looking at a large number of images. However, there are situations in which this type of visualization would not work, particularly when working with non-spatial maps or when the spatial features do not change much.

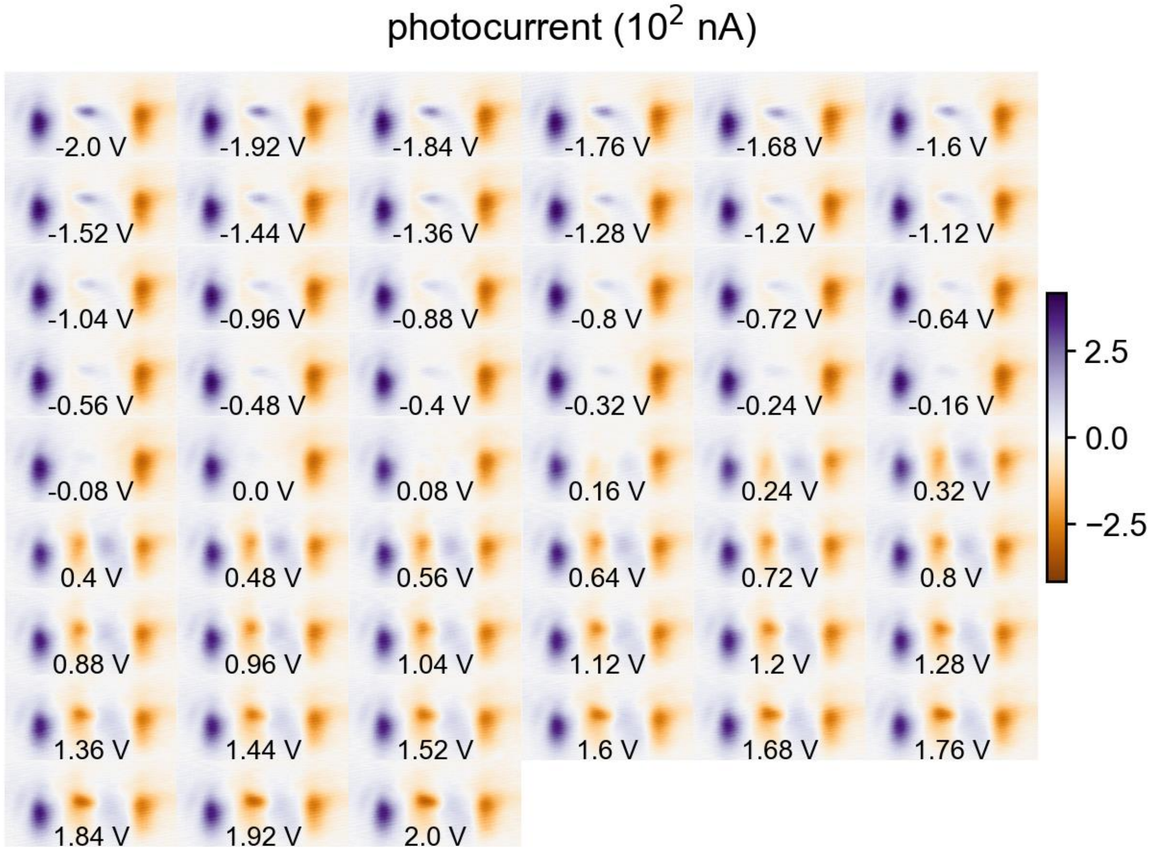


Figure 3.7: A tile plot representation of spatial photocurrent data, showing all images in a datacube (taken on sample *GBNG-2*) in a sequence of images.

Alternatives include movies and interactives. A movie would display the images in a timeseries, but while this may be useful when formally presenting the data it is usually only useful when the feature that is changing is obvious, in which case it isn't really any more useful than a series of images as shows above and has the downside that it is difficult to include in a notebook. Another alternative is an interactive, where, for example, a single image is displayed and there is an interactive control that lets the user select that image from the cube. While they can be used to pick out high resolution images and add them to the lab notebook, it is also limited. The main alternative to follow the general MPDPM procedure and lower the dimension of the data using a fit, as discussed in section 3.6.

Moving beyond three-dimensional data sets adds the additional complication that they are necessarily spread across many different runs. To satisfy the data handling policy and provide for easy access to a set of images gaborlab.mpdpm contains the DataSet object which takes a list of run numbers (from a local datasets.py file, see the documentation) and processes all of them into a single unified object, which most of the same attributes as an individual Run object. The series of runs in a data set can be a series of 2D scans, resulting in a 3D set, or a series of 3D data cubes resulting in a 4D data set. The most basic option is to display all the images in much the same way, but with the axes of the plot representing two of the 4D axes.

Fig. 3.8 shows a tile plot representation of a four-dimensional imaging data set, it is different from the datacube tile plot in that the images are arranged along the non-spatial axes. This representation shows the data in its full glory, from all the images the user can visually pick out large scale trends in the data. This is an excellent starting point; however, finer analysis will require the use of the MPDPM data analysis process, reducing data down through fitting. At or above four dimensions it becomes increasingly difficult to visualize raw data in a human compatible fashion.

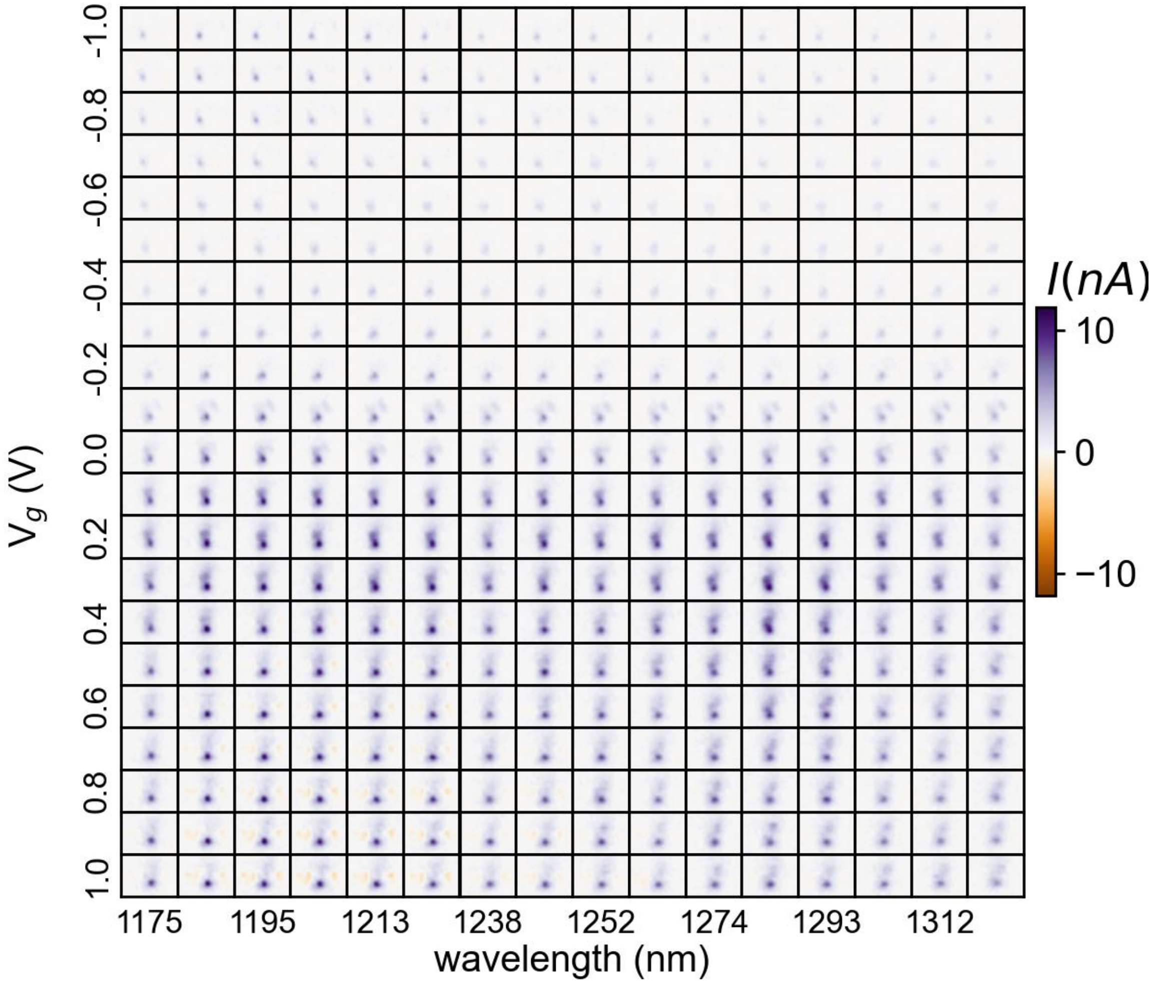


Figure 3.8: A tile plot representation of a four-dimensional set spatial photocurrent images, showing all images in a datacube (taken from sample *MoSe2WSe2-2*) laid out over the two non-spatial axes.

3.11 Fitting of Data and the Visualization of Fit Parameters

The final tool that is essential to working with multi-dimensional data sets (or any data set) is the ability to apply a fit to the data. Fitting is important in most applications of science, but essential in MPDPM data analysis where fitting is used to reduce an un-representable amount of data into a digestible plot in a publication or presentation.

The first, and most crucial, decision about a fit is what fitting function should be used. The user should look at the data and get a sense of the functional form that the data follows. Fig. 3.9a shows a line scan over the active area of a sample as a function of laser power. To examine this data, we simply take the average along the x-axis over the active area. Looking at the data we notice the key character of the data, firstly it is monotonically increasing and non-linear. That informs several potential fitting functions, the most obvious of which would be a polynomial such as a parabola. However, as we will discuss below, a polynomial is not a good choice to fit this data. There are two key concepts to keep in mind when selecting a fitting function.

The first concern when choosing a fitting function is that a fitting function should depend on at least one dynamical parameter. A dynamical parameter is a parameter which reflects the underlying dynamics of the system and is not simply a rescaling of the data. We want to use a parameter to represent the behavior of the fitted data and it is hard to see that any of the parameters in a polynomial do that, for example in a parabola, $I(P) = AP^2 + BP + C$, the parameters would be A, B and C but let's suppose that we represent the data using the parameter A. What does A mean? An increase in A rescales the data but nothing fundamental about the system changes. Furthermore, A or B could be negative, and even if they are not negative for a given fit the possibility doesn't reflect the monotonic characteristic of the data. Therefore, it's hard to see what A would actually represents because A is not dynamical.

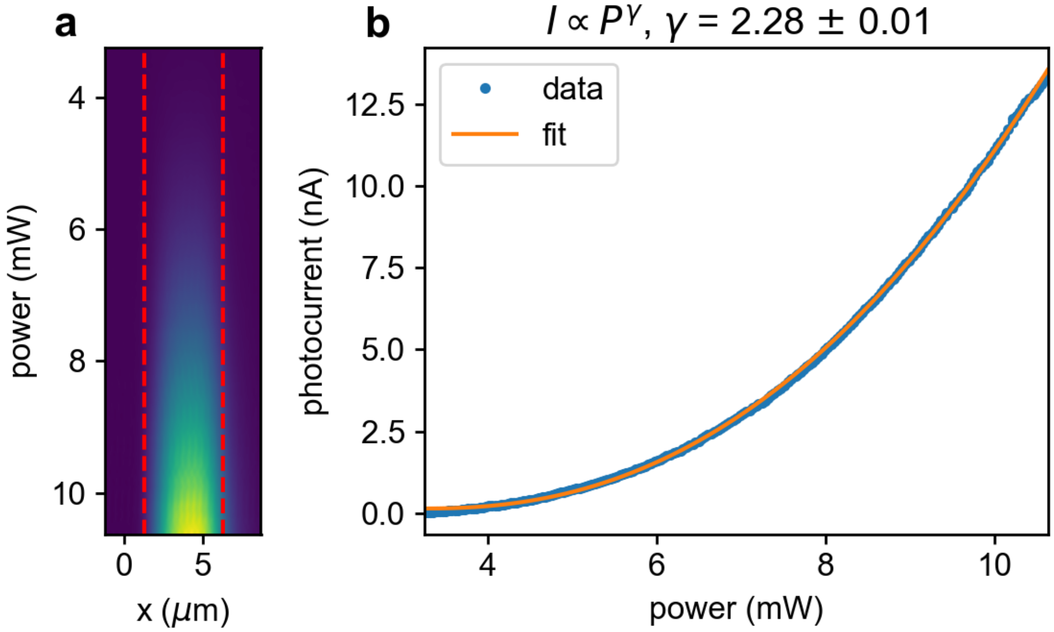


Figure 3.9: A simple line scan (taken on sample *GBNG-2*) fit to a power law.

The second concern when choosing a fitting function, is that a fitting function should have the minimum number of free parameters, each of which should be distinct. With a function like a polynomial there is a temptation to add additional free parameters, i.e. if a parabola doesn't fit well, one might add a cubic or quartic term to the equation. This would make the polynomial fit better, but it doesn't actually add anything. The way the fitting algorithms work adding more free parameters can, and often does, improve the fit superficially, meaning that the fit will more closely resemble the data even if the improvement isn't meaningful, a phenomenon known as overfitting. Consider adding a cubic term to the parabola above, $I(P) = DP^3 + AP^2 + BP + C$. Fitting to this will improve the fit artificially, but this improvement isn't meaningful. To see why, consider that if we fit the same function to similar data that is slightly more super-linear, the fitting could reflect the increased current scale by increasing A or by increasing D, with no way to determine which is better. In practice, what would happen is that the values of A and D would become inconsistent, fitting them to slightly different data (even if only different due to noise) would give significantly different values, because from the perspective of the fitting algorithm they both rescale the current, thus the algorithm can trade value in one for value in the other to further improve the resemblance to small features of the data. Thus, whatever meaning that the parameter A had is diluted when D is introduced.

Given these concerns, what fitting function should we use for the above data? Well we want a function that monotonically increases, is non-linear, and is simple, requiring only a few distinct parameters to capture a range of behavior. This points to a power law function, $I(P) = AP^{\gamma} + I_0$. There are only three parameters, each of which have different units and do different things, and there are no obvious opportunities to add more parameters and overfit. Furthermore, γ is a dynamical parameter, its value is an index of the non-linearity the data exhibits. If $\gamma < 1$ the system is sublinear, if $\gamma > 1$ the system is super-linear, and any notable change in value, for example going from 2 to 3 indicates a significant change in the underlaying behavior, not just a re-scaling of the data. If we apply a power law fit to the data, we get a single number, $\gamma = 2.28$, that represents the behavior of the sample with respect to power. The power law is one of the more common fitting parameters we use in MPDPM. The other option for selecting a fitting function would be to get it from a model of the data, but this comes along later in the process, in the exploratory phase use a simple and common fitting function with dynamical parameters, such as a power law.

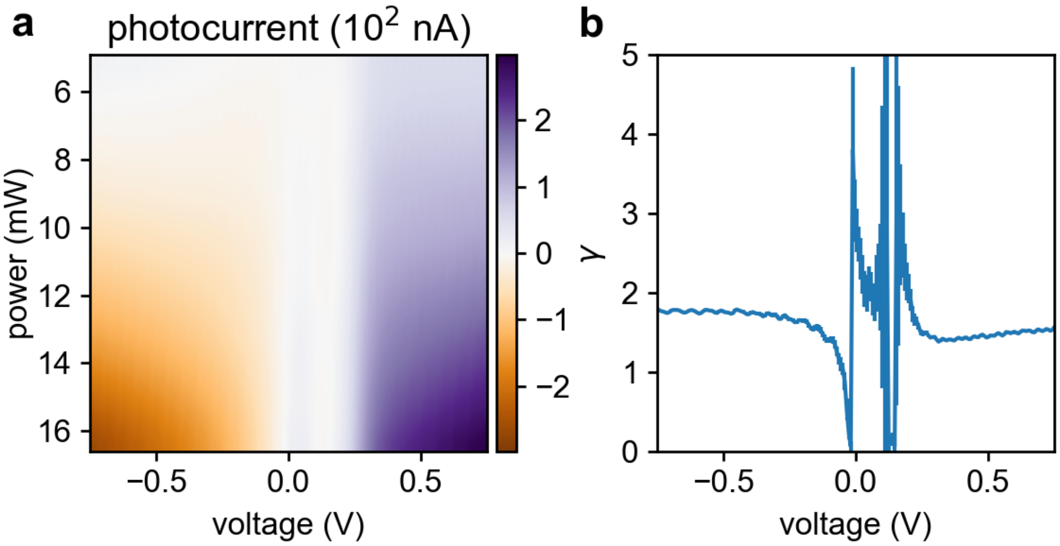


Figure 3.10: Representing a scan with a fit. **a**, photocurrent data, **b** the γ parameter of the power law function, acting as an index of the non-linearity of the sample.

When fitting MPDPM data sets, a single fit like that shown above is rarely sufficient, rather the whole Run or Set needs to be fit consistently. Performing such fits over multi-dimensional data quickly runs into several problems that the GLT has tools designed to solve. The main principle is to fit Runs (or Sets) as object, preserving the object orientation and allowing all the details of the fitting to take place behind the scenes such that the user can perform large fits quickly without getting lost in the details. The class that fits a cube is fitting.FittedData which takes a run and fits it along an arbitrary axis. Let's return to the data shown in Fig. 3.6, while the waterfall plot gives an idea of the trends of the data, the dynamics of that data can be represented by fitting to a power law and looking at the dynamical parameter γ. Thus, that data can be represented as shown in Fig. 3.10. This fitted representation gives a lot of insight that the waterfall plot does not, in it we observe two regimes and the points where the power law behavior breaks down (at the nodal points) can be identified.

A more advanced application of fitting involves fitting a data cube (or set) to represent, in two dimensions, what is happening in three dimensions. In this case it is even more important that the fitting function have a good dynamical variable, because the visualization will not include all of the data, and the user is relying on the fit to accurately represent it. Fig. 3.11a shows the γ parameter for a data cube, taken on a sample with a square geometry. However, it is apparent that there are features in the gamma map that do not correlate with anything. We see this in a vertical line near the left edge, the bright yellow edge around the green square and the several noisy green bands in the background. From knowledge of the system we know that this device is approximately square, and the image is larger than the sample size. These spurious features are due to noise in the system, the FittedData object fits at every point in the cube regardless of how well the fit works or how much signal there is at a given point. To deal with this we need to filter the γ result to include only significant points. Fig. 3.11b shows the γ data where points with fractional uncertainty greater than 10%, i.e. $\sigma_{\gamma}/\gamma > 0.1$, are filted out. Thus, only the green square is present, which is the response of the sample.

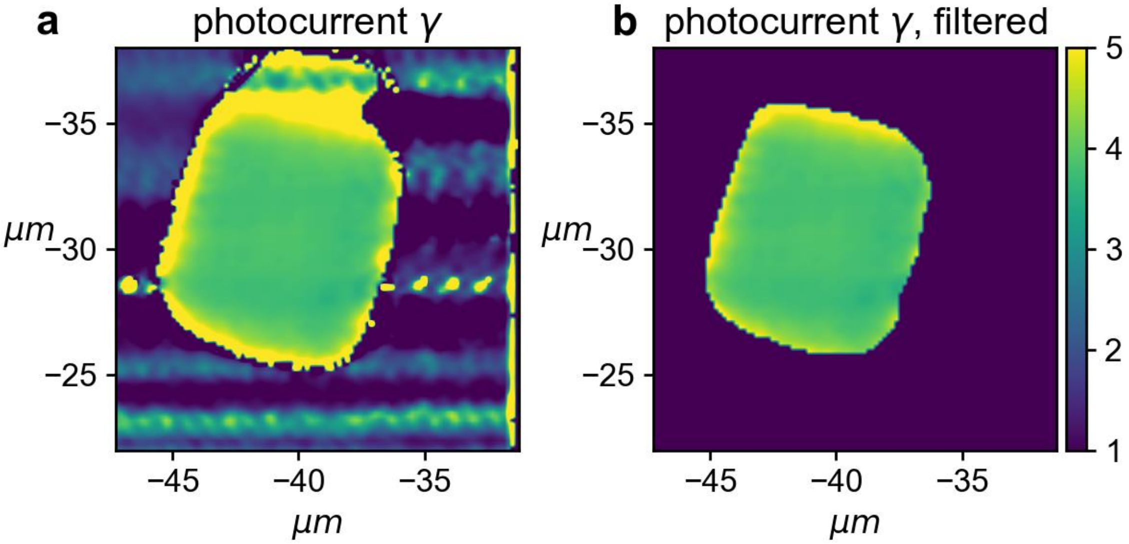


Figure 3.11: Non-linearity of a square sample (*GBNG-3*) representing a photocurrent datacube with a power-law fit. **a**, the γ parameter of the power law function, raw output of the fitting algorithm. **b**, the γ parameter of the power law filtered to include only point with low uncertainty.

CHAPTER 4:

THE MICROSCOPIC SCALE

4.1 Excitons in 2D systems at the Microscopic Scale

One of the fundamental and widely studied non-equilibrium excitations is the exciton. A bound state of an electron and a hole, an exciton is a quasi-particle comparable to a metastable hydrogen atom that can be used to study a variety of physics. As discussed in Chapter 1, excitons have higher binding energies in 2D materials, making them far easier to generate. TMDs in particular are conducive to excitons and we will use TMD heterostructures as a platform to study the interactions of excitons. These interactions are fundamentally microscopic, involving two, three or four body scatting. In the high exciton density limit, a strongly interacting phase can emerge where the individual microscopic interactions cease to be meaningful. This limit is discussed in Chapter 5 on the Mesoscale, in this chapter we will focus exclusively on the low-density limit where the exciton physics is dominated by few-body interactions.

There are two broad interactions we want to study, Auger recombination and the exciton-phonon interaction. In this chapter, we will use the graphene-$MoTe_2$ heterostructure as a model system to study Auger recombination and the $MoSe_2$-WSe_2 heterostructure as a model system to study the exciton-phonon interaction.

4.2 Auger Recombination and Biexcitons in 2D Systems

Photo-excited electrons and holes in semiconductors lose energy through rapid collisions, the types of collisions can vary but the common process is known as Auger recombination. Auger recombination is a relaxation process in which an electron and hole recombine by transferring energy to a third particle[53,54]. The usual assumption is that Auger recombination is a three-particle process between an exciton and a free carrier. This is the case in conventional electron transport: a free charge carrier collides with an electron-hole (e-h) pair, and only one charge carrier escapes. In low dimensional materials, this picture is expected to change slightly, allowing Auger processes to occur as two-particle processes, that are strongly enhanced compared to bulk materials due to relaxation in momentum conservation and strong spatial confinement[55].

At sufficiently high-density excitons can collide to form short-lived biexcitons - molecular-like complexes of two e-h pairs bound by the Coulomb force[56-58]. This is in strong contrast to conventional semiconductor materials (for example Si or Ge), in which biexcitons are highly unstable and exhibit very low binding energy[59]. In the limit that colliding exciton pairs are strongly bound, Auger recombination is a two-particle process: two excitons collide, and only one exciton escapes.

The biexciton process would normally compete with other Auger processes, particularly the exciton-free carrier three body process. However, the Graphene-$MoTe_2$ model system we use covers the $MoTe_2$ with graphene which has been shown to neutralize charge inhomogeneity in TMDs and reduce the influence of trionic states[60]. This suggests that three-body free carrier Auger recombination may be reduced in this system and two-body biexciton Auger recombination may be dominant. Thus, the goal of understanding the dynamics is to determine whether two or three body Auger processes dominate. In the following sections on the Graphene-$MoTe_2$ heterostructure we develop a simple yet comprehensive model that accounts for the complex photocurrent behavior resulting from ultrashort pulses complementing strong evidence that 2-particle Auger processes (exciton-exciton annihilation) dominate the infrared photoresponse.

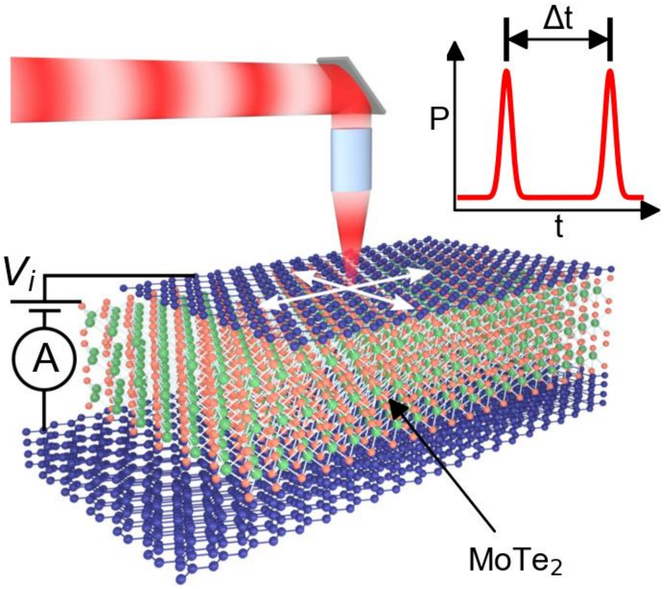


Figure 4.1: Schematic of measurements on the *GMoTe2* sample. Inset shows the power of the pulsed laser as a function of time, not to scale.

4.3 Photocurrent from Graphene-MoTe$_2$ Heterostructures

The first model system we will consider is the Graphene-MoTe$_2$ Heterostructure, specifically the *GMoTe2* sample (detailed in appendix A1.3), which has a “graphene sandwich” structure with graphene on the top and bottom and in the middle a 9 nm thick flake of MoTe$_2$. In this system the graphene layers are mostly inactive and are used as transparent contacts which collect local current. The broad area of overlap, combined with the relative thickness, and thus high absorption, of the MoTe$_2$ means that this system generates a large amount of signal that is proportional to the number of excitons, making it an excellent platform for understanding exciton dynamics in a TMD. Fig. 4.1 shows the basic experiment, an ultrafast pulsed laser locally injects light into the system, exciting electrons and holes which may form excitons. The graphene on the top and bottom is electronically contacted, and we measure the photocurrent as a function of the interlayer voltage V_i, the location of the beamspot and the properties of the laser including power and two-pulse delay, Δt. Here we will consider the $MoTe_2$'s excitonic behavior, though under strong excitation it exhibits a collective e-h liquid phase on the mesoscale, which will be discussed in chapter 5.

We begin by examining the photocurrent dependence of several relevant variables, with the laser at $\lambda = 1200$ nm. Fig. 4.2 shows the photocurrent behavior in the low excitation regime. Figs. 4.2a and b show the spatial photocurrent and an image of the device; we observe that the photocurrent is large and fairly uniform inside the heterostructure area (outlined in white), falling off outside the heterostructure area within a margin on the order of the beamspot width. From Fig. 4.2c we see that the device has a linear photoresponse as a function of interlayer voltage, slowly increasing in scale with laser power. The observed photoresponse, along with an examination of the laser wavelength and temperature dependence in other samples, is fully consistent with the picture absorbing at the A exciton of the $MoTe_2$ and getting photocurrent from dissociating excitons (for a discussion of laser absorption see appendix A2.1)[61]. The inset to Fig. 4.2c shows the mechanism of photocurrent generation, excitons are polar and when an electric field is applied the exciton has some probability of dissociating into an electron and hole which are collected at the top and bottom graphene.

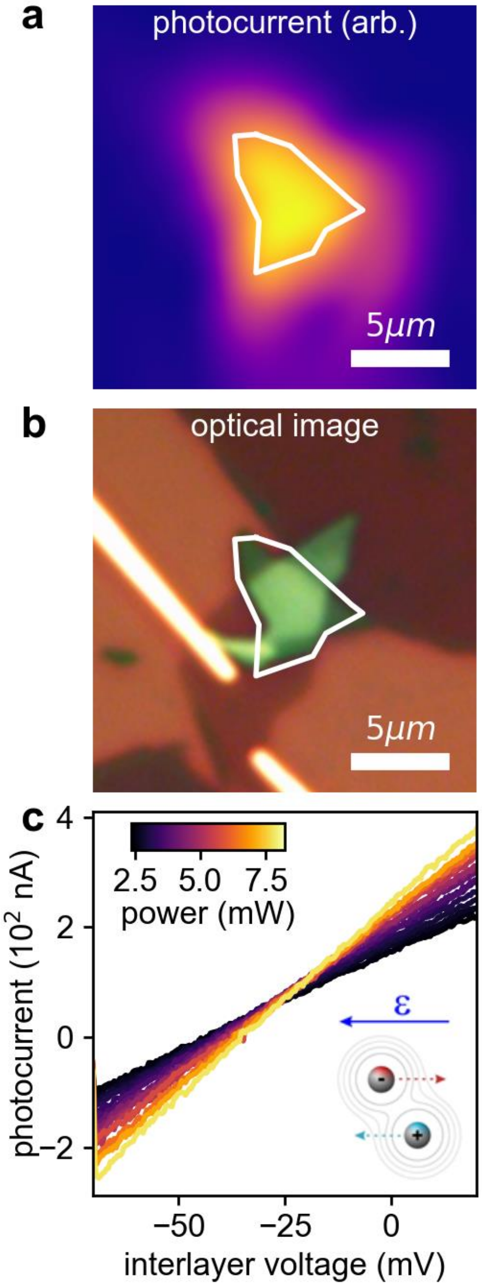


Figure 4.2: Photocurrent characterization of the *GMoTe2* heterostructure. **a**, the photocurrent as a function of space and **b** an optical image of the same space. The heterostructure overlap is outlined in white. **c**, the photocurrent versus interlayer voltage as a function of laser power.

To measure the dynamics of the excitons we will examine the dependence of the photocurrent upon two parameters that affect the density of excitons. First, we look at the power dependence of excitons, as the laser power affects the number of excitons created under constant absorbance. Fig. 4.3a shows the one-pulse infrared photocurrent with the laser fixed in the center of the sample and the laser power varied. The power dependence is accurately described using a power law relationship, $I \propto P^{\gamma}$ (red dashed line) where $\gamma = 0.44$ quantifies the sub-linearity of the photocurrent. Fig. 4.3b shows the photocurrent versus the two-pulse delay, Δt, for several low values of power. Importantly, when consider the two pulse at relatively low power. At short pulse delays and high powers, i.e. $\Delta t < 30$ ps and $P > 5$ mW, the photoresponse approaches the highly interactive regime at high powers washing out the dynamics we are interested in this chapter.

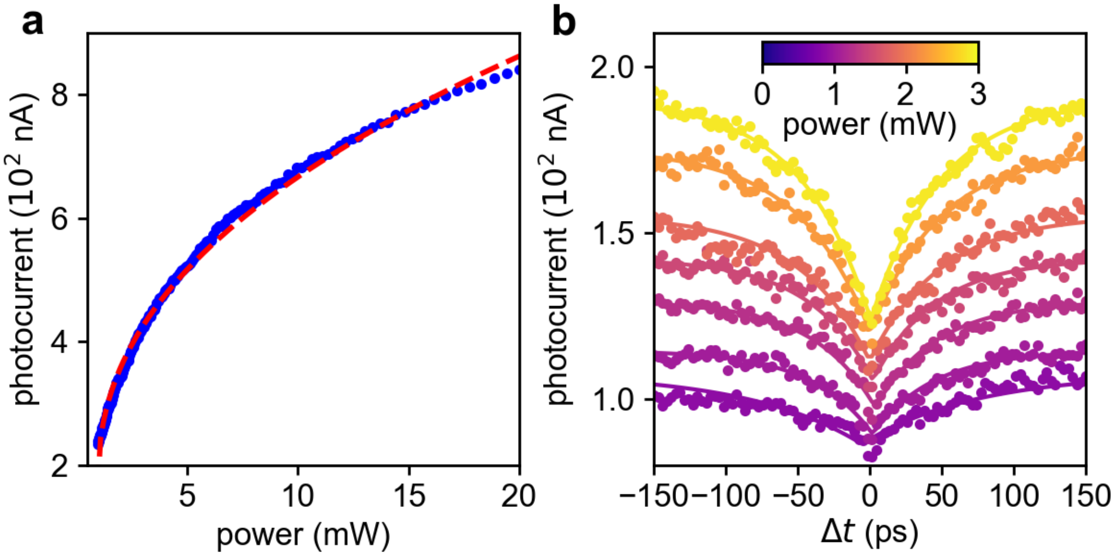


Figure 4.3: Photocurrent dynamical fits in the low power regime. **a**, the photocurrent as a function of laser power, the red dashed line is a fit to a power law $I \propto P^{\gamma}, \gamma = 0.44$. **b** photocurrent as a function of two pulse time delay, Δt, and power. Points are data and solid lines are fits to a symmetric exponential function $I \propto e^{-|t|/\tau}$.

In the two-pulse dependence we see that the photocurrent is strongly suppressed near $\Delta t = 0$ ps where the two pulses overlap but increases rapidly as the pulses become well separated ($\Delta t > 50$ ps). The photocurrent is fit by a symmetric exponential function (solid lines), $I \propto e^{-|t|/\tau}$, where the parameter τ represents the timescale on which photoexcited charge carriers linger in the sample. Again, we see a sub-linear dependence on density; two well separated pulses should double the current of a single pulse, but if those carriers are in the system when the second pulse arrives the result is less than double the current of a single pulse. The sub-linear behavior is comprehensive, dominating the photoresponse below the strongly interacting regime. To understand the charge carriers in this regime we develop a model for the sub-linear power dependence.

4.4 Exciton Interactions and the Rate Equation Model

We attribute the power law behavior $I \sim P^{1/2}$ in the *GMoTe2* sample to 2-particle Auger (exciton-exciton) annihilation: nonradiative decay of one e-h pair provides excess kinetic energy to the second e-h pair. After optical excitation by an ultrashort pulse, electron-hole pairs in $MoTe_2$ form a high-density gas of interacting charge carriers. In the absence of an electric field, photo-excited electrons and holes dwell within the junction sufficiently long to allow for two-body electron-hole annihilation. At finite electric field, dissociated electrons and holes transit the device and are collected at the graphene contacts[26-28]. Due to Coulomb and phonon-mediated scattering, free electrons and holes escape with velocity $v = \mu E = \mu \frac{2(V_i - \phi_0)}{L}$, where μ is the mobility, L is the $MoTe_2$ thickness, and ϕ_0 is the built-in potential (ϕ_0 = - 41 mV for device shown in this work, based on the voltage dependence in Fig. 4.2c). The time required for liberated charge carriers to escape the photocell junction τ_{es} can be tuned by the interlayer voltage $\tau_{es} = L^2 \,/\, 4\mu(V_i - \phi_0)$ [27].

In order to understand the dynamics, we model the influence of excitons, and exciton-exciton annihilation, using a basic rate equation to model the electron hole pair density in the sample as a function of time, $N(t)$:

$$\frac{dN}{dt} = G(t) - \frac{N}{\tau_{es}} - \alpha N^2 \tag{4.1}$$

At instantaneous time, the electron hole pairs generated by the exciting laser form into excitons with a generation rate G(t). Some of the excitons dissolve creating free carriers that are pulled out of the system and are observed as photocurrent. The linear term, $-N/\tau_{es}$, accounts for this, where τ_{es} is the timescale of carriers escaping the system. Some excitons interact with each other, forming higher order states (such as biexcitons) or annihilating. The quadratic term, $-\alpha N^2$, accounts for this where α is the overall rate of two-body interactions. Solving this equation accounts for the observed behavior. This model assumes a 2D sample, and N is the two-dimensional electron hole density, but can be used to model non-monolayer quasi-2D systems such as the sample used in this work, with a thickness of 9 nm.

Equation 4.1 describes the dynamics for recombination dominated by two-body Auger recombination, indicative of biexcitons. To proceed we solve this equation for pulsed excitation. It is common in the literature to work with similar rate equations in the steady state limit, i.e. $dN/dt = 0$, however we use a pulsed laser where the steady state limit is not valid. For a discussion of the steady state solution and a comparison to the pulsed solution derived below see appendix A2.2.

In the pulsed case, the laser deposits energy into the sample on the timescale of the pulse width, in our case ~190 fs, which is significantly faster than the observed dynamics. We assume that excitons initially form on a similar timescale. After the laser pulse, $G(t) = 0$ as the system returns to equilibrium. Thus the dynamics are determined by equation. 4.1 with $G(t) = 0$ and the initial condition $N(t = 0) = N_1$, where N_1 is the initial number of carriers from a single pulse, proportional to laser power. After that the density rapidly decays (Fig. 4.4a insets) according to the integrated eqn. 4.1:

$$N(N_1, t) = \frac{N_1 e^{-t/\tau_{es}}}{1 + N_1 \alpha \tau_{es} (1 - e^{-t/\tau_{es}})} \tag{4.2}$$

This gives the instantaneous density, but the observed quantity is the photocurrent arising from exciton dissociation, thus, to find the photocurrent from a single pulse, I_I, we integrate the linear term of equation 4.1:

$$I_I(P) \propto \int_0^{\infty} \frac{N(N_1, t)}{\tau_{es}} dt = \frac{1}{\alpha \tau_{es}} \ln(1 + N_1 \alpha \tau_{es}) \tag{4.3}$$

Fig. 4.4a shows this solution, plotting equation 4.3 as a function of laser power (proportional to N_1). The insets to Fig. 4.4a plot equation 4.2 for low and high power, the integrated area under the curves generates the values for the main plot.

The dynamics gets more complicated in the two-pulse picture but can still be treated with this model. When a second pulse arrives - delayed by Δt - newly excited excitons combine with those excited by the first pulse (Fig. 4.4b). The photocurrent counts the number of carriers that escape the photocell and is determined by the time-integrated dynamics of both pulses (Fig. 4.4b insets). The two-pulse photocurrent $I_{II}(\Delta t)$ is then written as:

$$I_{II}(\Delta t) \propto \frac{1}{\alpha\tau_{es}}\ln[1 + N_1\alpha\tau_{es}(1 - e^{-\frac{\Delta t}{\tau_{es}}})] + \frac{1}{\alpha\tau_{es}}\ln(1 + N_2\alpha\tau_{es}) \tag{4.4}$$

where the exciton population immediately following the second pulse is $N_2 = N_1 + N(N_1, \Delta t)$ (shown in Fig. 4.4b). When the delay time is much longer than the escape time ($\Delta t >> \tau_{es}$), Equation 4.4 reduces to the one-pulse photocurrent I_I:

$$I_{II}(\Delta t \to \infty) \approx 2I_I \propto \frac{1}{\alpha\tau_{es}}\ln(1 + N_1\alpha\tau_{es}) \tag{4.5}$$

as shown in Fig. 4.4a.

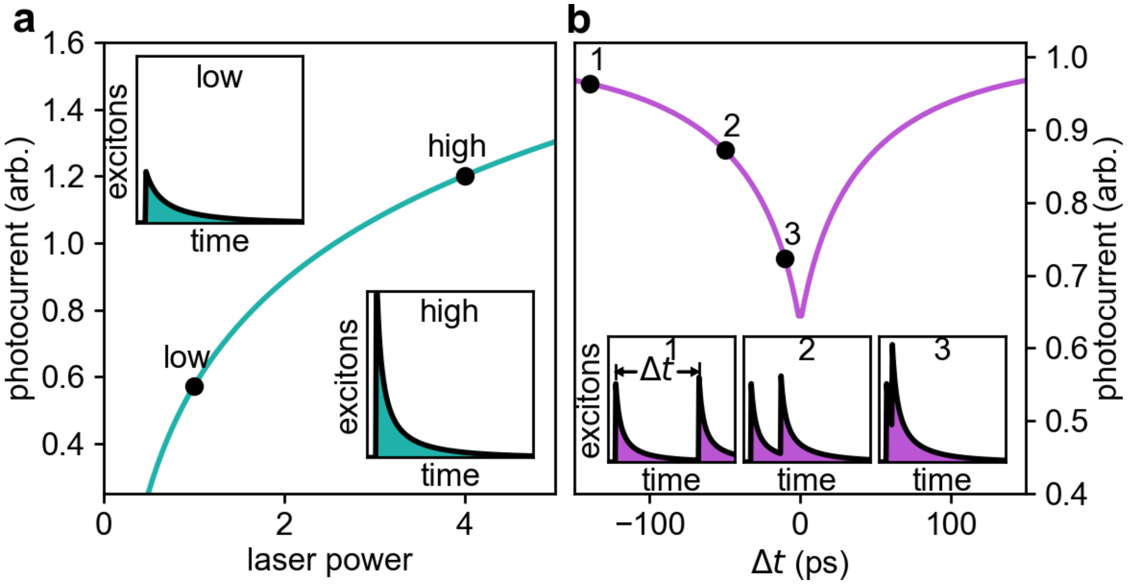


Figure 4.4: **a,** the solution of Eqn. 4.1 for a single laser pulse as a function of laser power, (Eqn. 4.3) and, **b**, for two pulses as a function of two pulse time delay, Δt (Eqn. 4.4). The photocurrent is proportional to the integral of the number of excitons versus time, shown in the insets, which is initially driven to a maximum value proportional to laser power, then exponentially decays.

4.5 Two-Body Auger Recombination in $MoTe_2$

The exciton-exciton annihilation model exhibits excellent qualitative and quantitative agreement with the time-resolved photoresponse of the *GMoTe2* sample. At long Δt, equation 4.5 indicates that the two-pulse photocurrent is twice that of a single ultrafast pulse $I_{II} \approx 2I_I$. By normalizing $I_{II}(\Delta t)/2I_I$, the two-pulse photocurrent should collapse at long time delays and be nearly independent of laser power. Fig. 4.5a shows the photocurrent as a function of Δt normalized by the photocurrent at $\Delta t \to \infty$ (obtained through fitting to a symmetric exponential). The data is shown as colored points and fits to equation 4.4 are shown as solid lines. We observe that for all subcritical powers the fits are in good agreement with the data.

In addition, we see that when plotted as $I_{II}/2I_I$ all of the data collapses onto roughly the same line, indicating that our model fully captures the dynamics of the photocurrent as a function of Δt and laser power. Similar models can be formulated that include three body terms, or treat excitons and free carriers as different species, however all of these models are both more complicated and demonstrate less qualitative and quantitative agreement with the data. Therefore, we conclude that the dynamics of Auger recombination in the *GMoTe2* sample is dominated by two-body collisions, indicative of biexcitons. Furthermore, the remarkable agreement with an analytic model means that we can use the fits to the data to extract numerical parameters characterizing this interaction.

Parameters α and τ_{es} can be obtained by fitting equation 4.4 to a data set that varies in both power and Δt. To help the fitting algorithm optimize without extraneous fitting parameters we divide the data by its long Δt value (obtained from a simple exponential fit) and fit that normalized data to $I_{II}/2I_I$. In addition, this function is difficult to fit due to the near degeneracy of α and τ_{es}, which are always found as the product $\alpha\tau_{es}$, except in the exponential factor $e^{-\Delta t/\tau_{es}}$. Thus, only measurements that have high data density in Δt can fit distinct values of α and τ_{es}. Given the above, we fit to the normalized data shown in Fig. 4.5a.

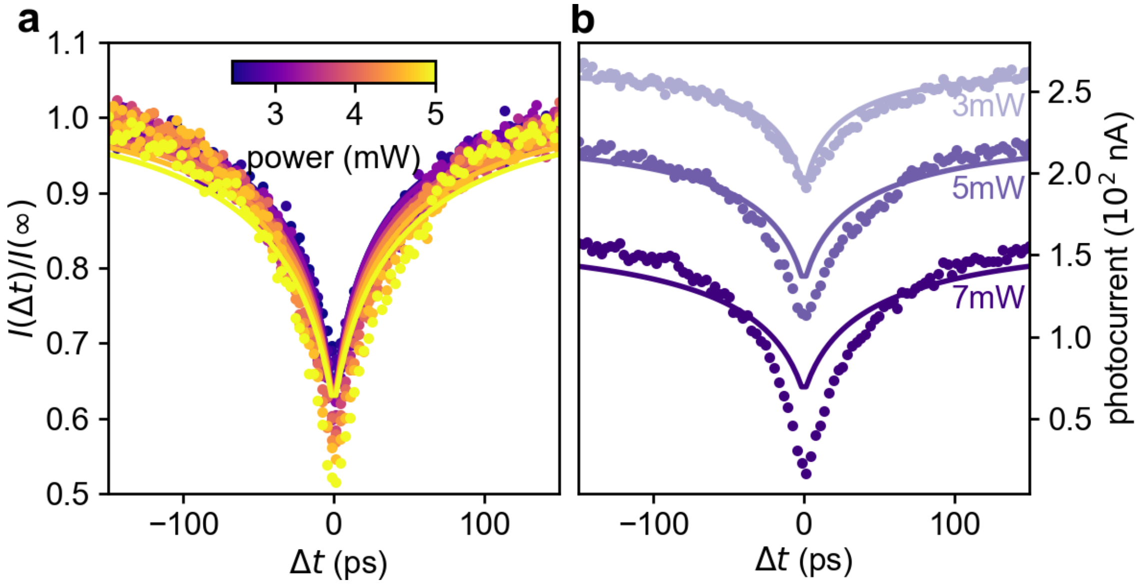


Figure 4.5: **a** photocurrent vs. Δt, normalized by $I(t \rightarrow \infty)$, as a function of power (points) fitted to the two-body model (solid lines). **b**, photocurrent versus delay, with a biexciton model fit, showing the model breaking down at high powers (>5 mW) and short delay, $|\Delta t| < 50$ ps, e.g. at high exciton density.

Since the fit values at each power were mostly consistent between all the different values of power and we average them to obtain the final values of $\alpha = (5.62 \pm 0.52)\ x\ 10^{-6}$ cm^2/s and the escape time at zero applied voltage of $\tau_{es} = 86 \pm 2$ ps. In addition, for simplicity this data was fit with the assumption of 100% absorption, i.e. A=1. Realistically, A~10% (see the discussion of absorption in appendix A2.1) but imperfect absorption does not essentially change the fitting, except that the value of α must be divided by A, i.e. $\alpha(A) = \alpha_{100\%}/A$, for A < 1.

All models have limits to their validity and the two-body Auger recombination model begins to break down as the density of excitons rises above a certain threshold. Fig. 4.5b shows three traces of photocurrent versus Δt (points) fit to the two-body model (solid lines) at three different values of power. We see that the 3 mW and the 5 mW photocurrent fit well to the two-body model, but that the model breaks down on the 7 mW photocurrent at $|\Delta t| < 50$ ps. This is because at high power and short Δt the density of excitons becomes so large that many-body effects are not negligible and begin to dominate the charge carrier dynamics. Not coincidentally, 5 mW is the critical threshold (at $\Delta t = 0$) in laser power for a phase transition to a strongly interacting state, which will be discussed in Chapter 5.

4.6 Phonons and Interlayer Excitons in 2D Heterojunctions

Having explored the interactions involved in Auger recombination let us turn to the other prototypical interaction of excitons, the exciton-phonon interaction. As discussed in Chapter 1 when electrons and holes form an exciton, they can be prohibited from recombining by momentum constraints. Such excitons are called indirect excitons because they require an indirect transition, that is a transition between two points in k-space that have different momenta. Due to the conservation of momentum indirect excitons fundamentally require a three-body interaction in order to form or recombine. The third body, besides the electron and the hole, is often a quantized vibrational mode of the lattice, i.e. a phonon. That makes indirect excitons an ideal system to study the interactions between electrons and phonons.

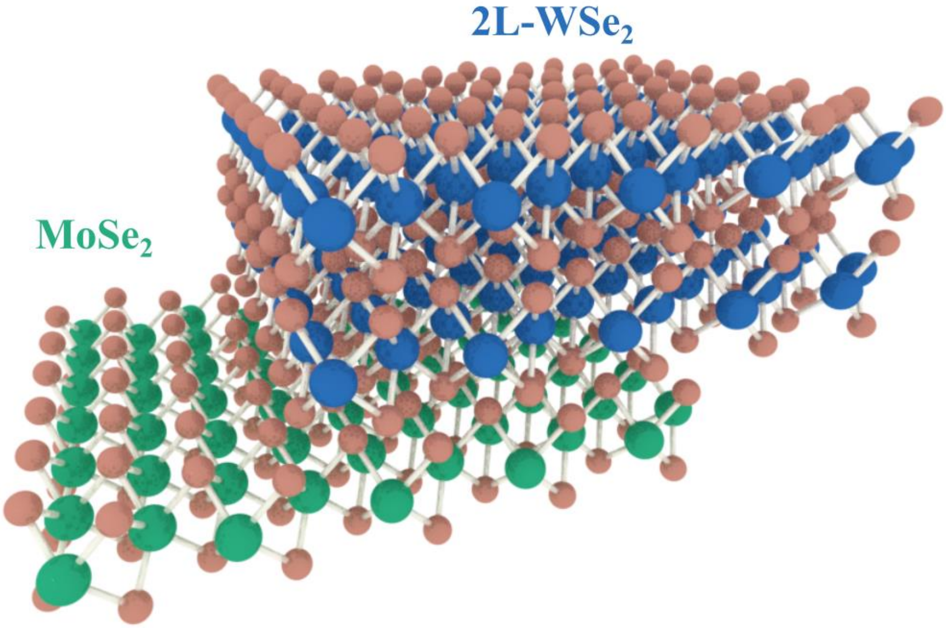


Figure 4.6: Schematic of the MoSe$_2$ -WSe$_2$ heterostructure that is the key component of the *MoSe2WSe2-1* and *MoSe2WSe2-2* samples.

There are multiple types of exciton that can occur in 2D systems, here we will focus on the interlayer exciton. An interlayer exciton is an exciton that occurs in a layered heterostructure where the electron and the holes are localized in different layers, a configuration which gives them a great variety of interesting physics[10,11]. Specifically, in our desire to study the electron phonon interaction interlayer excitons have two key properties, firstly a strong out of plane electric dipole moment that makes them highly controllable by external voltages. Secondly, the interlayer exciton is highly localized; by definition the interlayer exciton occurs at the interface between two materials, and if the materials are atomically thin then there are only a few layers for the electron and hole to exist in. Such a boundary condition means that the electron and hole are localized to only a few unit cells and are thus more sensitive to any changes on that same scale, such as the deformation due to a vibrating lattice.

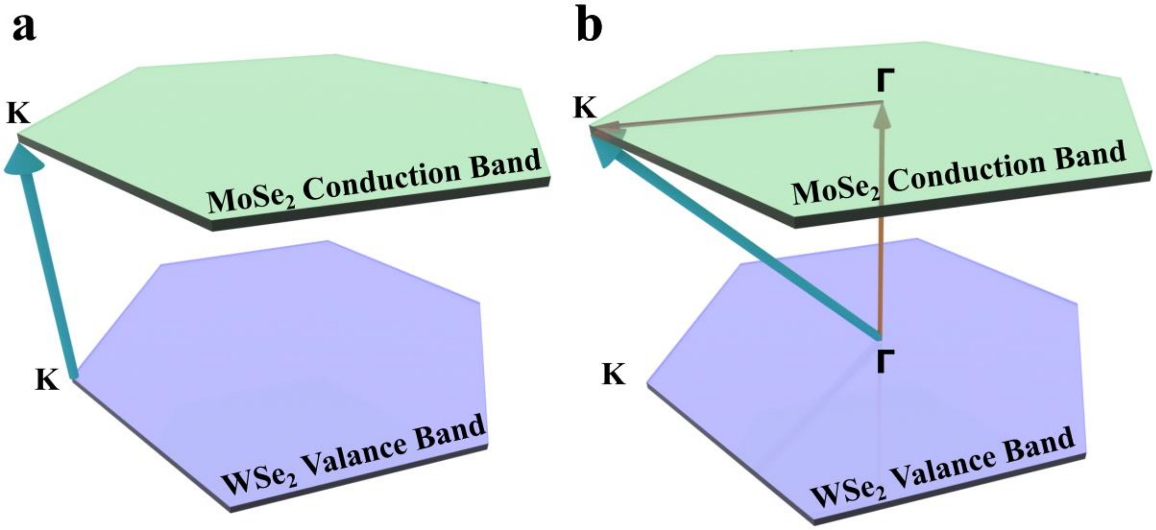


Figure 4.7: Two interlayer excitonic transitions, **a** direct and **b** indirect, of interest. Planes represent Brillion zones of TMD materials, with the high-symmetry points K and Γ marked.

The material system we will be using to study this is a stack of monolayer $MoSe_2$ and bilayer WSe_2, shown schematically in Fig. 4.6. This system is known to have a type II band alignment which enables electron-pair multiplication across an atomically sharp interface[62]. In this system we expect that there are two interlayer excitons relevant to our study the first is a higher energy (~1.3 eV) direct exciton corresponding to a transition between the K point of the WSe_2 valance band and the K point of the $MoSe_2$ conduction band, as shown in Fig. 4.7a [63,64]. Secondly, we expect a lower energy (~1.0 eV) indirect transition between the Γ point of the WSe_2 valance band and the K point of the $MoSe_2$ conduction band, as shown in Fig. 4.7a[62]. Being indirect the $\Gamma \rightarrow K$ phonon involves a third body, usually expected to be a phonon. Thus, we proceed with investigating the photoresponse of this exciton, with the $K \rightarrow K$ exciton useful for comparison.

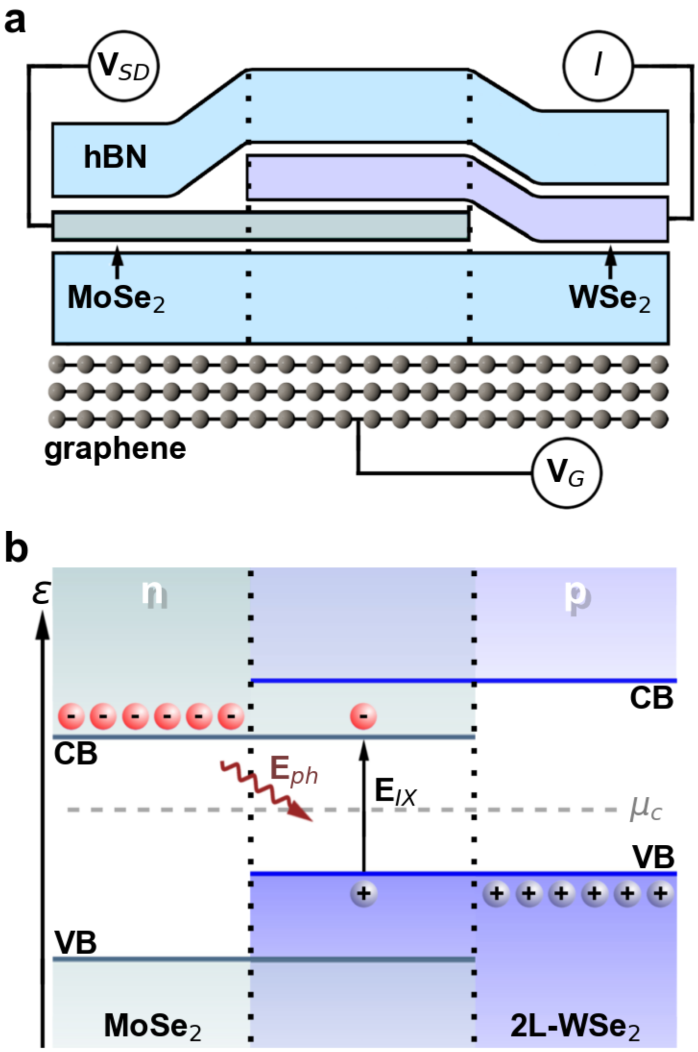


Figure 4.8: *MoSe2WSe2-2* sample schematic, **a**, and band diagram, **b**. The *MoSe2WSe2-1* sample used at low temperature is consistent except that it uses the Si substrate as the gate connected to V_G.

4.7 The Photocurrent Parameter Space of $MoSe_2$-WSe_2 Heterostructures

In this dissertation we will use two $MoSe_2$-WSe_2 samples to explore the photocurrent parameter space in search of the interlayer exciton. The fabrication and basic characterization of these samples, referred to as *MoSe2WSe2-1* and *MoSe2WSe2-2* is discussed in detail in appendix A1.4. Fig. 4.8a shows the *MoSe2WSe2-2* heterostructure schematically, indicating contacts used to apply an inter-layer voltage V_{SD}, and apply a gate voltage V_G to the graphene backgate (sample *MoSe2WSe2-1* is the same, except with a Si backgate.) Current I is measured on the WSe2 contact and includes photocurrent and dark current, the photocurrent, I_{PC}, is extracted wither with a lock-in amplifier or by subtracting the known dark current. As shown schematically in Fig. 4.8b, at equilibrium a WSe_2-$MoSe_2$ heterostructure forms a type II semiconductor interface: i.e., the energy difference between the highest WSe_2 valence band and lowest $MoSe_2$ conduction band is larger than the conduction (valence) band offsets between WSe_2 and $MoSe_2$. By applying a gate voltage V_G to the graphene electrode, we precisely control the charge carrier density in the heterostructure, carefully tuning the chemical potential μ_c. It is across this gap that we excite the interlayer exciton with energy E_{IX}.

To characterize the indirect $(\Gamma \rightarrow K)$ exciton we measure the photocurrent spectrum at low temperature. Fig. 4.9 shows the interlayer photocurrent vs. E_{PH} and V_G for the *MoSe2WSe2-1* device, measured at low temperature (T = 20 K). Strikingly, equally spaced photocurrent peaks are strongly pronounced below the interlayer *e-h* pair excitation energy E_{IX}. This highly unusual behavior is highlighted in Fig. 4.9a, which shows I_{PC} vs. E_{PH} and V_G at lower photon energies. In the range E_{PH} = 0.85 - 1.05 eV, we observed a set of evenly spaced photocurrent maxima, which increase in amplitude as E_{PH} increases. These peaks occur only in a narrow range of gate voltage near V_{G} = -5.0 V. The highest energy peak occurs at E_{PH} = 1.02 eV, and line traces of I_{PC} vs. E_{PH} (Fig. 4.9b) show additional regularly spaced peaks at energies below the dominant resonance. To characterize these oscillations in the photocurrent response we take the second derivative of photocurrent with respect to E_{PH} and calculate the Fourier transform, shown in Fig. 4.9c. The Fourier transform reveals that the oscillations consists of a primary component at $\Delta\varepsilon$ = 30 meV and a weaker secondary component at 22 meV.

Unexpectedly, the same oscillations occur when the direct $(K \rightarrow K)$ exciton is characterized. When E_{PH} was tuned between 1.24 eV and 1.40 eV (Fig. 4.9d), we observed a sequence of photocurrent peaks that occur only near V_{G} = -5.0 V. While the strongest peak occurs at E_{PH} = 1.31 eV, it is only slightly stronger than several equally spaced maxima at higher and lower photon energy. By calculating the Fourier transform of the second derivative (Fig. 4.9e), we find a clear periodic component at $\Delta\varepsilon$ = 30 meV.

The low temperature measurements on the *MoSe2WSe2-1* sample were taken using a standard spectroscopy approach, that is sampling a two-dimensional phase space, due to the difficulty of working at low temperature. However, there are a variety of other variables at play; the geometry of the sample and interlayer voltage V_{SD} also directly influence the photoresponse, and optical parameters such as laser power must be explored to understand if they influence the system. Given the number of parameters and the potential complexity, data intensive imaging is required to fully characterize this parameter space thus we proceed with the MPDPM methodology discussed in chapters 2.

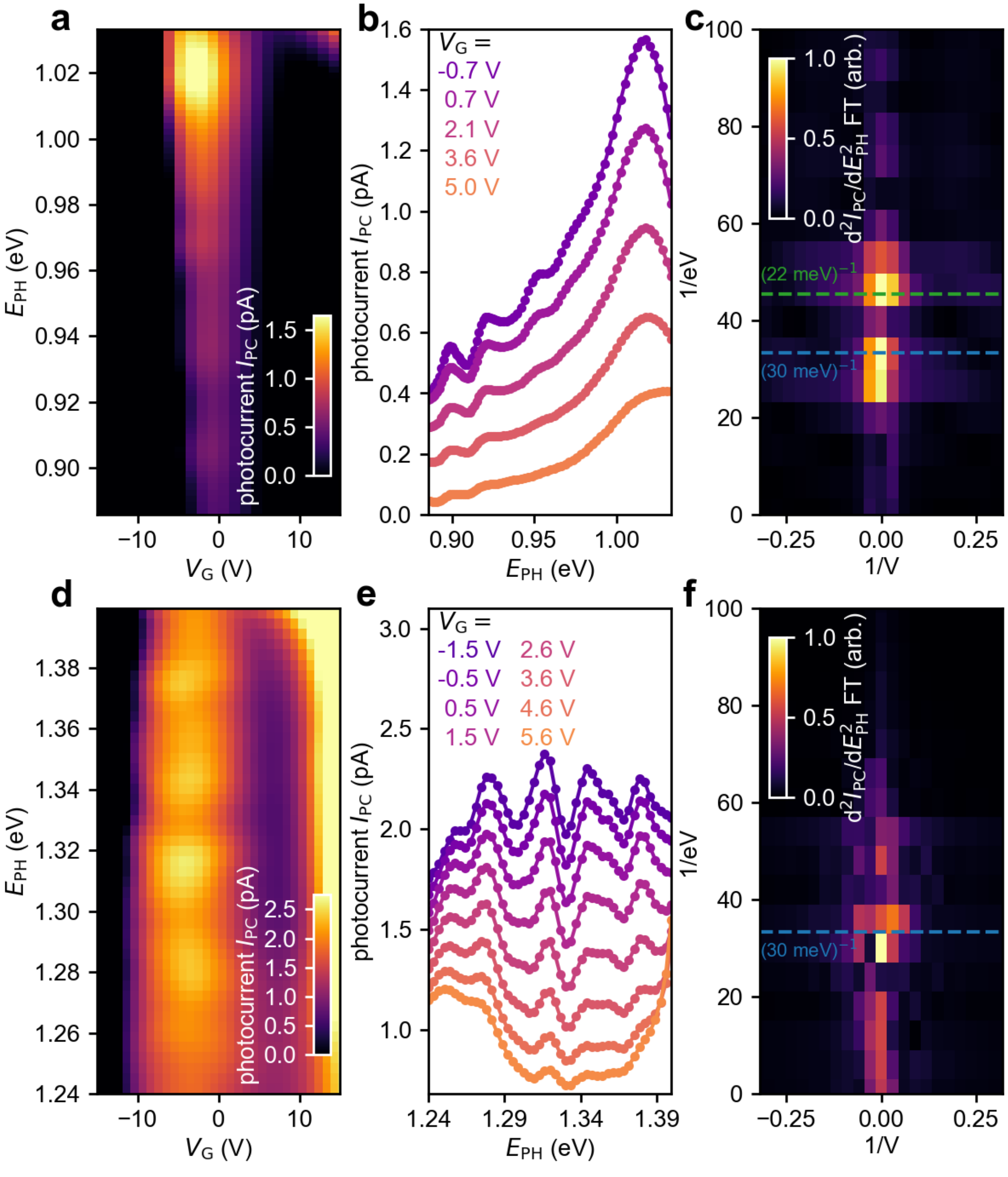


Figure 4.9 Photocurrent spectroscopy on the *MoSe2WSe2-1* sample at low temperature (T≈20 K). **a**, the photocurrent spectrum (E_{PH} vs V_G) near the indirect ($\Gamma \rightarrow K$) exciton. **b** line cuts of the photocurrent spectrum and **c** the Fourier transform of the second derivative of the photocurrent, showing two distinct modes at 22 meV and 30 meV. **d**, the photocurrent spectrum (E_{PH} vs V_G) near the direct ($K \rightarrow K$) exciton. **e** line cuts of the photocurrent spectrum and **f** the Fourier transform of the second derivative of the photocurrent, showing a distinct mode at 30 meV. Photocurrent shown is a modulation on top of a significantly larger dark current contribution which is subtracted off.

To begin, we data intensively sample the voltage parameter space. Fig. 4.10a shows a set of 2000 photocurrent images versus V_G and V_{SD} with the laser wavelength at 1250 nm (1.0 eV). This is dataset is spatially consistent, i.e. corrected for spatial drift, and in the tile plot format we see the large-scale behavior of the device. Zooming in on a subset of the images in Fig. 4.10b and superimposing the heterostructure area of the sample we see that the majority of the photocurrent originates outside the heterostructure, mainly from the semiconductor-metal interface at the device contacts. However, near V_G = 0 V we observe distinct and spatially uniform photocurrent originating from the heterostructure area, consistent with our expectation of photocurrent from an interlayer exciton near charge neutrality.

To verify that the photocurrent from the heterostructure is due to the interlayer exciton we take a similar set of spatial photocurrent maps at a fixed value of V_{SD} = -0.35 V and varying the laser wavelength, shown in Fig. 4.11a. Again, we identify the heterostructure region and consider the average photocurrent originating in that spatial area. The heterostructure responsivity, e.g. the photocurrent divided by the power, is shown in Fig. 4.11b left, and we see a bright stripe near V_G = 0 V. We average between V_G = 0.0 V and V_G = 0.4 V (right side of Fig. 4.11b) and observe a resonance in photon energy peaked at 0.96 eV, very close to our expectation of a 1.0 eV interlayer exciton though slightly shifted from the exciton seen in Fig. 4.9 due to a different electrostatic environment. Being able to confidently detect the excitonic photocurrent from the heterostructure we are able to explore the voltage dependence in more depth.

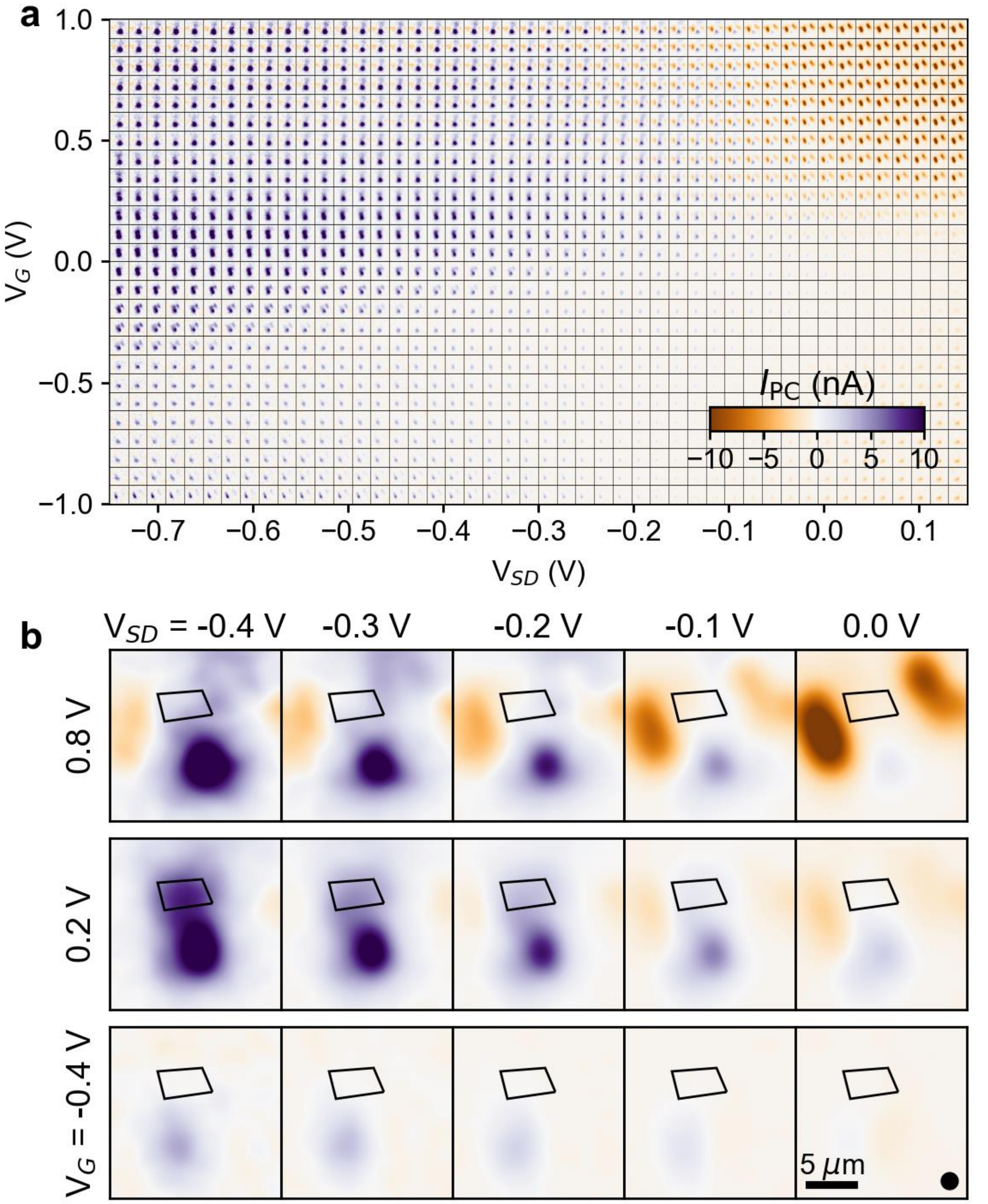


Figure 4.10: **a**, data intensive imaging of the *MoSe2WSe2-2* sample in the V_{SD} versus V_G parameter space. **b**, a series of photocurrent images showing the heterostructure area.

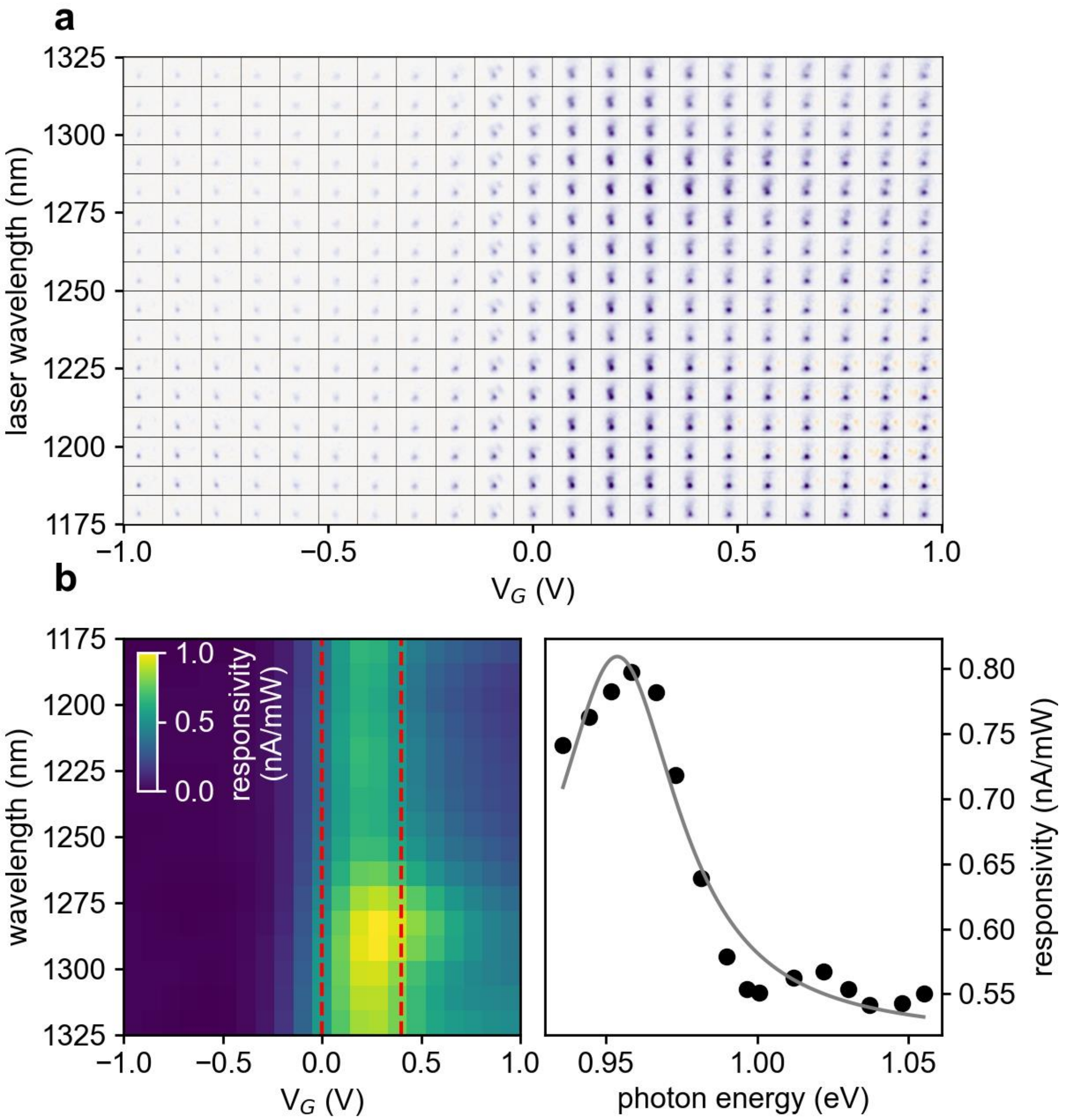


Figure 4.11: **a**, data intensive imaging of the *MoSe2WSe2-2* photocurrent versus wavelength and gate voltage. **b**, processed responsivity (photocurrent divided by laser power) from the heterostructure regions and **c**, the averaged responsivity (points) fit to a Lorentzian function (grey line).

In Fig. 4.12a we plot the average heterostructure photocurrent versus V_{SD} and V_G observing a bright feature of positive photocurrent near $V_G = 0$ as a function of increasing V_{SD}. To proceed, we will look at the photoconductance, that is dI_{PC}/dV_{SD}, which we plot in Fig. 4.12b. We observe a bright feature corresponding to the main feature, but we also see a number of small oscillations in the conductance. To further visualize these oscillations, we take the second derivative d^2I_{PC}/dV_{SD}^2, shown in Fig. 4.12c, where we see strong oscillations in the photocurrent. Observing these oscillations with respect to V_{SD} we want to compare them to the oscillations observed at low temperature with respect to E_{PH}. To proceed we must calculate the shift in energy shift across the heterostructure from a change in V_{SD}.

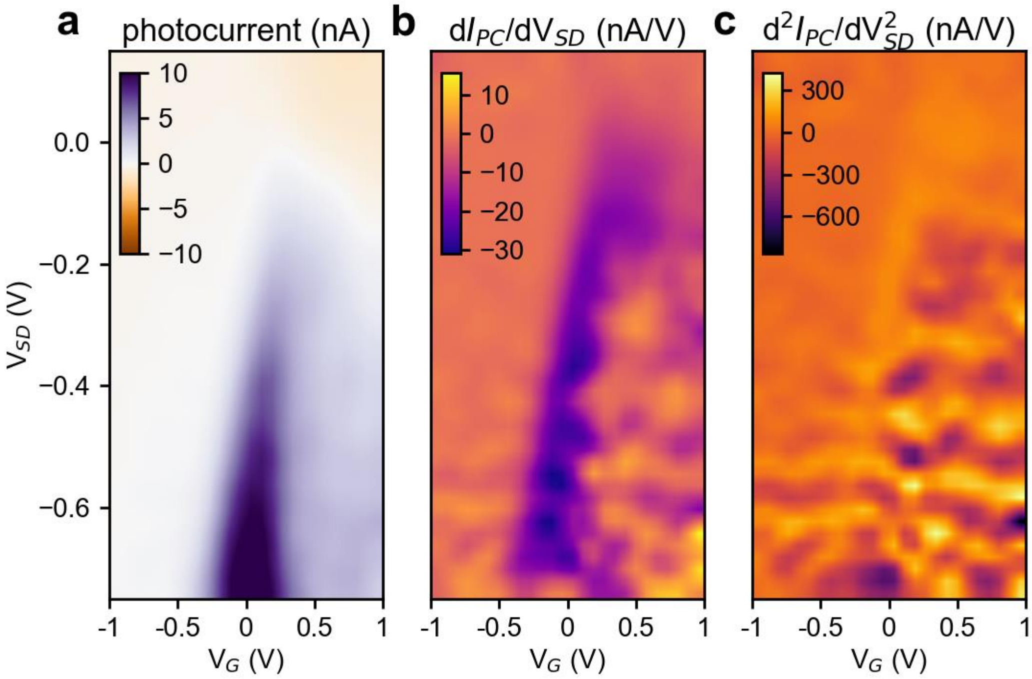


Figure 4.12: **a** heterostructure photocurrent, **b** the derivative of photocurrent with respect to V_{SD}, and **c** the second derivative of photocurrent with respect to V_{SD}.

From the photocurrent data (Fig. 4.13a) we plot the V_G-dependent characteristics taken at several V_{SD} values (Fig. 4.13b). We extract the gate voltage values of the photocurrent peaks V_G^*, and plot V_G^* as a function of V_{SD} (Fig. 4.13c). In addition to increasing the electric field across the $MoSe_2$-WSe_2 interface, the effective interlayer voltage $\Delta V_I \propto \Delta V_{SD}$ contributes to charging the individual layers, giving rise to a depletion capacitance. At charge neutrality, the charge density $n = 0$ and the change in gate voltage ΔV_G is related to a change in interlayer voltage ΔV_I by simple capacitive coupling $\Delta V_I = -(C_G/C_I)\Delta V_G$, where C_I and C_G are the capacitances per unit area for the interlayer depletion region and due to the bottom gate, respectively. From a linear fit to the data (red line in Fig. 4.13 c) we find that changes in V_{SD} are indeed proportional to changes in V_G: $\Delta V_{SD} = -\beta\Delta V_G \approx -2.1\Delta V_G$. Thus, along the diagonal line in Fig. 4.13a, the interlayer electric field increases while maintaining precise charge neutrality.

Remarkably, when the device is tuned to charge neutrality, I-V_{SD} characteristics are highly asymmetric, similar to a diode. In forward bias ($V_{SD} > 0$ V), the dark current increases exponentially (Fig. 4.13d), while in reverse bias the dark current is highly suppressed, consistent with van der Waals heterostructure p-n junction behavior[65]. We fit the dark I-V_{SD} characteristic to the usual diode equation $I = I_0(\exp(eV_{SD}/\alpha k_B T) - 1))$, where I_0 is the reverse bias saturation current, $k_B T$ = 26 meV is the thermal energy at room temperature, and $\alpha \geq 1$ is a phenomenological constant that relates changes in the interlayer electric field energy to a reduced voltage eV_{SD}/α [66]. The diode equation (red line in Fig. 4.13d) shows excellent agreement with the data, giving α = 1.8.

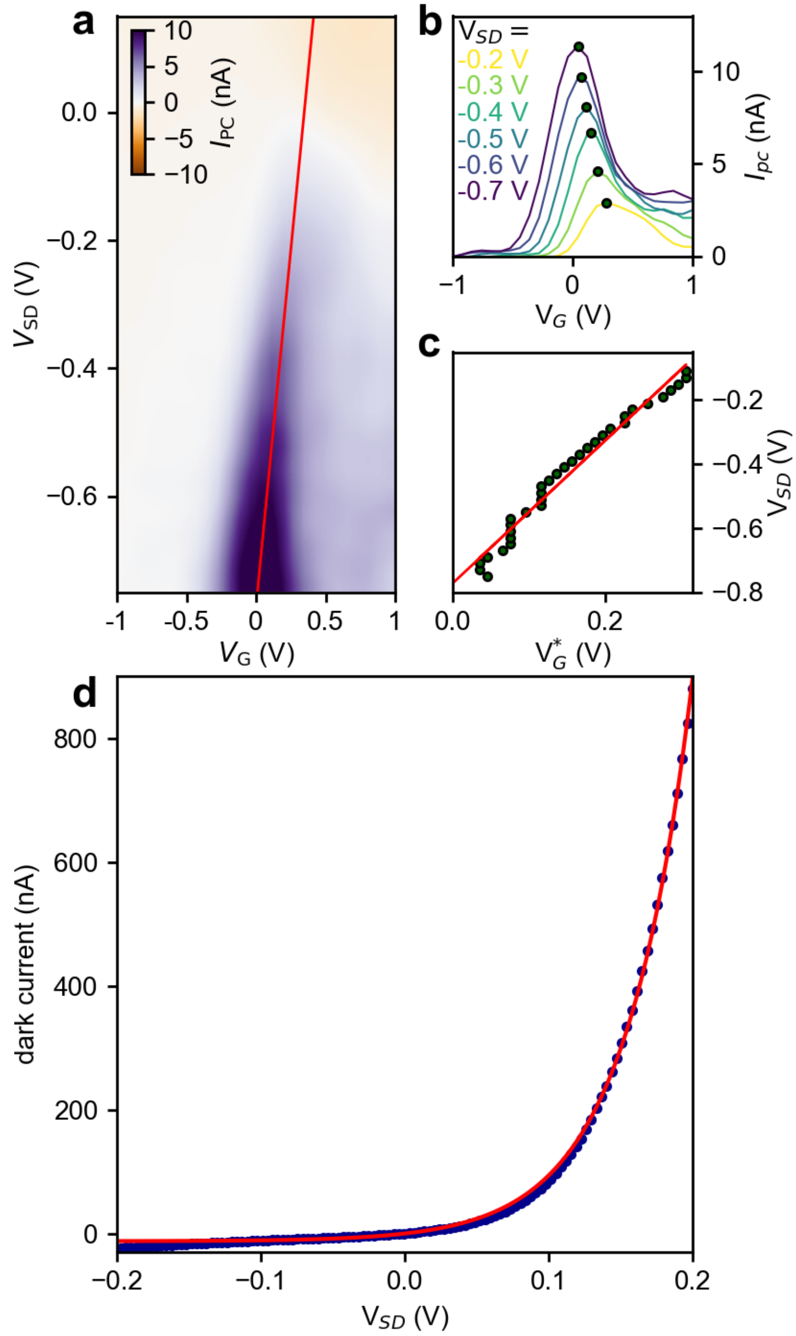


Figure 4.13: Determining electronic characteristics of the *MoSe2WSe2-2* sample. **a**, the heterostructure photocurrent. **b**, line cuts of the heterostructure photocurrent versus V_{SD} with maxima shown as points. **c**, locations of the maxima versus V_{SD}. Red line in **a** and **c** is the linear relationship between V_{SD} and V_G. **d**, dark current as a function of V_{SD} (blue points) fit to the Shockley diode equation (red curve).

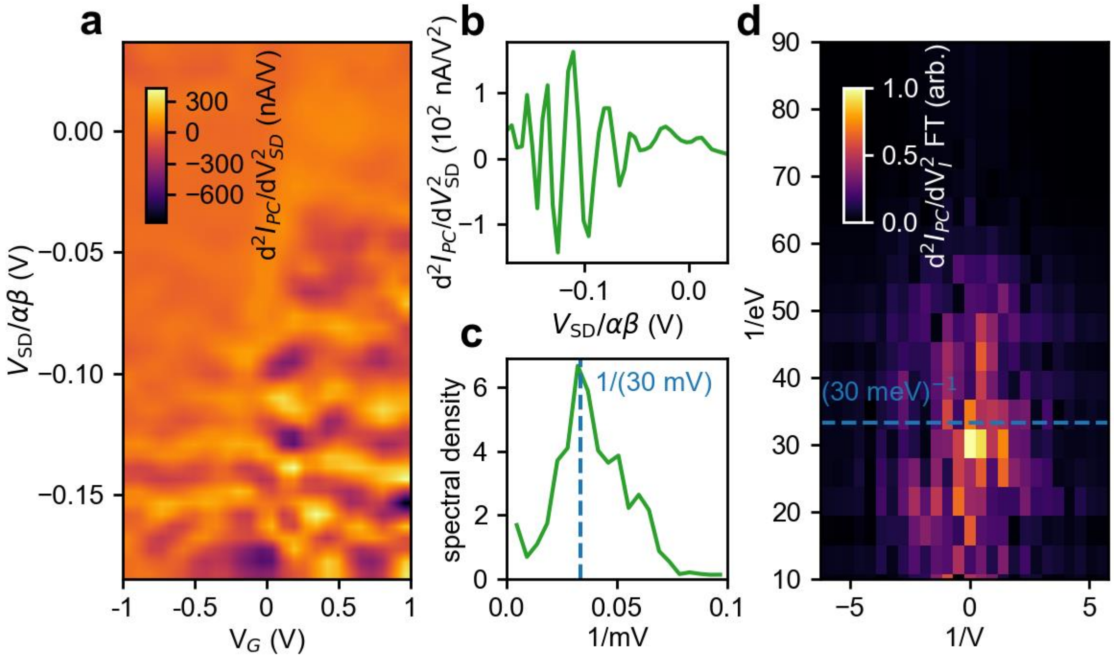


Figure 4.14: Oscillations in the room temperature data. **a**, the second derivative of the heterostructure photocurrent with respect to V_{SD} scaled by $1/\alpha\beta$. **b**, the photocurrent second derivative averaged between V_G = -0.3 and V_G = 0.4. **c**, the Fourier transform of the data in **b. d**, the two-dimensional Fourier transform of the data in a. Blue dashed lines mark the peak frequency at 1/(30 meV).

The above fitting characterizes the electrostatic properties of the data, next we calculate the energy change due to an effective interlayer voltage V_I. Writing the energy change as $\Delta E = \vec{p} \cdot \vec{E} = ed\, V_I/\beta t_{TMD}$ where $|\overrightarrow{p}| = ed$ is the dipole moment of the interlayer exciton with length d which we approximate as the heterostructure thickness. Then we relate changes in the interlayer electric field energy to the reduced voltage eV_{SD}/α, giving $\Delta E \approx e\, V_I/\beta = e\, V_{SD}/\alpha\beta$. Using the relationship $\Delta E = e\, V_{SD}/\alpha\beta$ to rescale V_{SD}, in Figure 4.14a we plot the derivative of the photoconductance d^2I_{PC}/dV_{SD}^2 as a function of effective energy. Fig. 4.14b shows d^2I_{PC}/dV_{SD}^2 averaged between V_G = -0.3 and 0.4 V to get a clear signal. We observe significant oscillations as a function of

ΔE. To obtain the period of oscillations we take the Fourier transform of $d^2 I_{PC}/dV_{SD}^2$ (Fig. 4.14c). Like the low temperature photocurrent spectra, voltage-dependent oscillations correspond closely to 30 meV. We can also take the two-dimensional Fourier transform of all the $d^2 I_{PC}/dV_{SD}^2$ data, shown in Fig. 4.14d in which we observe a broad peak centered at 30 meV. The broadness of this peak relative to the peaks in Fig. 4.9 c and f is due thermal broadening at room temperature, despite this the room temperature peak is consistent with the low temperature oscillations.

4.8 The Phonon Modes of $MoSe_2$-WSe_2 Heterostructures

To understand the oscillations in the photocurrent data we examine the phonon modes of this system, which Raman measurements tell us strongly active near 30 meV (see appendix A1.4). We use DFT calculations performed by our collaborators in Dr. Roger Lake's laboratory which were calculated using the Vienna ab initio simulation package (VASP) in the projected-augmented-wave method[66-70]. The dispersion of the phonon modes, and their density of states, is shown in Fig. 4.15 where the faded blue bar marks 30 meV, but is broadened by $k_b T$ to represent the range of energies that would be consistent with the 30 meV mode observed in the data at T = 20 K. From this we see that there are multiple strongly active modes with bands near 30 meV.

Each of the 26 bands in the dispersion corresponds to a particular mode of atomic vibration; next we explore what each of these modes corresponds to and which are most likely to affect the data. Fig. 4.16 visualizes the results of the DFT calculation for specific bands, showing a representative sampling of the phonon modes from the dispersion near 30 meV. The phonons mainly consist of in-lane vibrations, E_{2g} modes, with some out-of-plane vibrations, A_{1g} modes. This tells us that there are many optical phonons, especially in-plane modes, near the energy of the oscillations we observe. However, not all these phonons could result in the photocurrent oscillations.

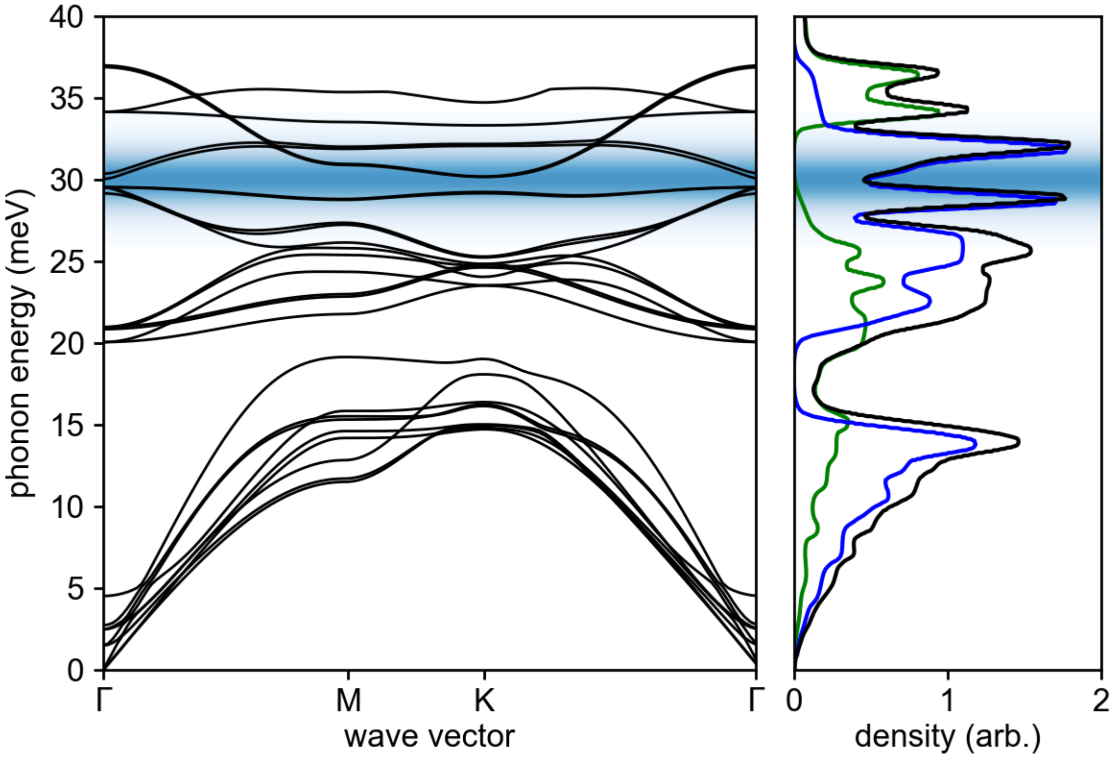


Figure 4.15: **left**, dispersion of DFT calculated phonon modes, and, **right**, phonon density of states, for the $MoSe_2$ (green line), WSe_2 (blue line) and heterostructure (black line). Blue bar shows 30 meV with thermal broadening of k_bT for T = 20 K.

By making some basic observations about our data we identify candidate phonon modes which may be related to the photocurrent oscillations. In the low temperature data, Fig. 4.9, we see sharp peaks at 30 meV across two different exciton transitions, that is involving transitions at different locations in the Brillion zone ($K \rightarrow K$ and $\Gamma \rightarrow K$). Therefore, if a phonon is responsible for these oscillations, we expect that it has an energy of approximately 30 meV at both the K and Γ points. In Fig. 4.17 we plot the seven phonon modes that are near 30 meV at the Γ point. We observe that most have a significantly different energy at the K point, except for two which are marked as blue lines. These bands corresponds to two nearly degenerate phonon modes, corresponding to the vibrations highlighted in Fig. 4.16 with the blue bars. We see that in these modes the tungsten and selenium atoms move oppositely in plane along either in-plane axis.

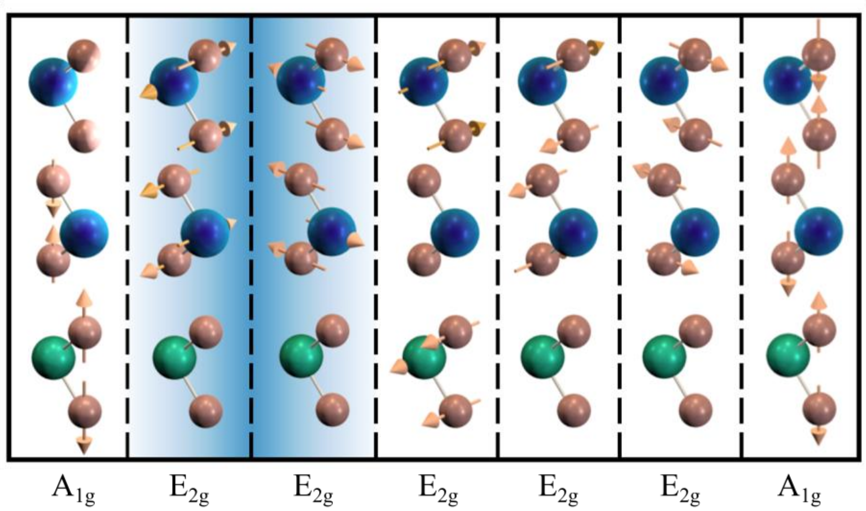


Figure 4.16: Visualization of optical phonons in the heterostructure along with their spectroscopy classification. Modes marked with the blue band correspond to nearly flat bands that are consistent with the oscillations observed in the photocurrent data.

Having identified candidate phonon modes that may correspond to the 30 meV oscillation, we observe two critical properties that will guide the formulation of a model. Firstly, we note that these bands are almost flat, meaning they have nearly the same value at every wavevector. This would imply that they are not constrained by momentum; that a 30 meV phonon in these modes can have almost any quasi-momentum. Secondly, we note that these phonon modes are localized on the same scale as the interlayer exciton. The interlayer exciton fundamentally has the electron localized in the WSe_2 layer, which vibrates in these modes. This implies that this phonon mode is vibrating the lattice over a significant part of the exciton wavefunction.

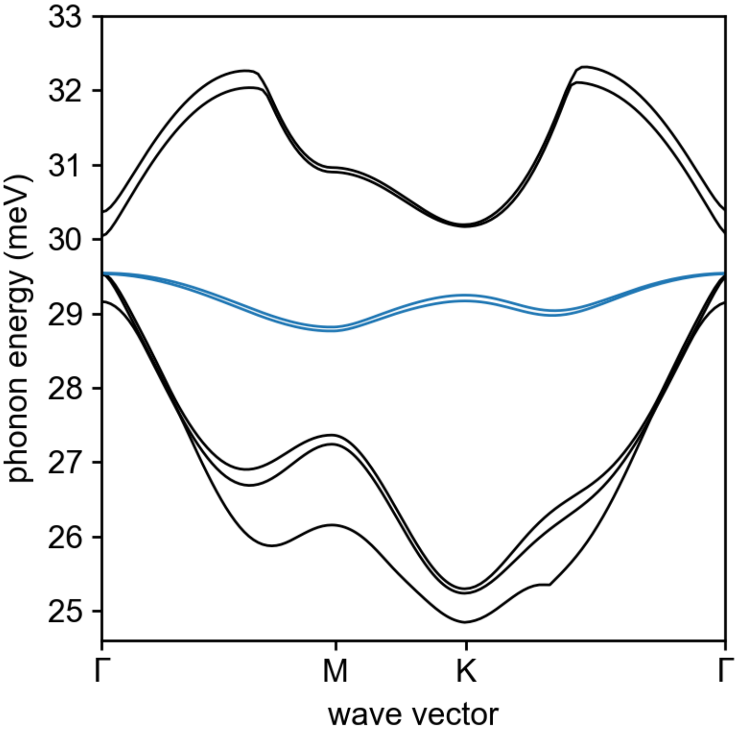


Figure 4.17: DFT calculated phonon modes zoomed in on modes that are near 30 meV at the Γ point. Blue lines are the candidate phonon modes.

4.9 The Electron-Phonon Interaction in $MoSe_2$-WSe_2 Heterostructures

Having clearly identified the 30 meV oscillation in the data and the properties of candidate phonon modes we develop a physical picture for how electron-phonon coupling can give rise to robust oscillations in the photocurrent. Fig 4.18 summarized the main experimental results, showing the 30 meV oscillation that is robust across several different regimes. To be consistent with the data for all three measurements a model must have an effect that occurs with respect to changes in energy from either E_{PH} or changes in the interlayer voltage $\Delta E = e\,V_{SD}/\alpha\beta$ and that it must be independent of which excitonic transition occurs, acting on either the lowest energy indirect $\Gamma \rightarrow K$ transition or the direct $K \rightarrow K$ transition. To develop and model that satisfies these constraints we use the two key features of the 30 meV phonon modes that we identified in the previous section.

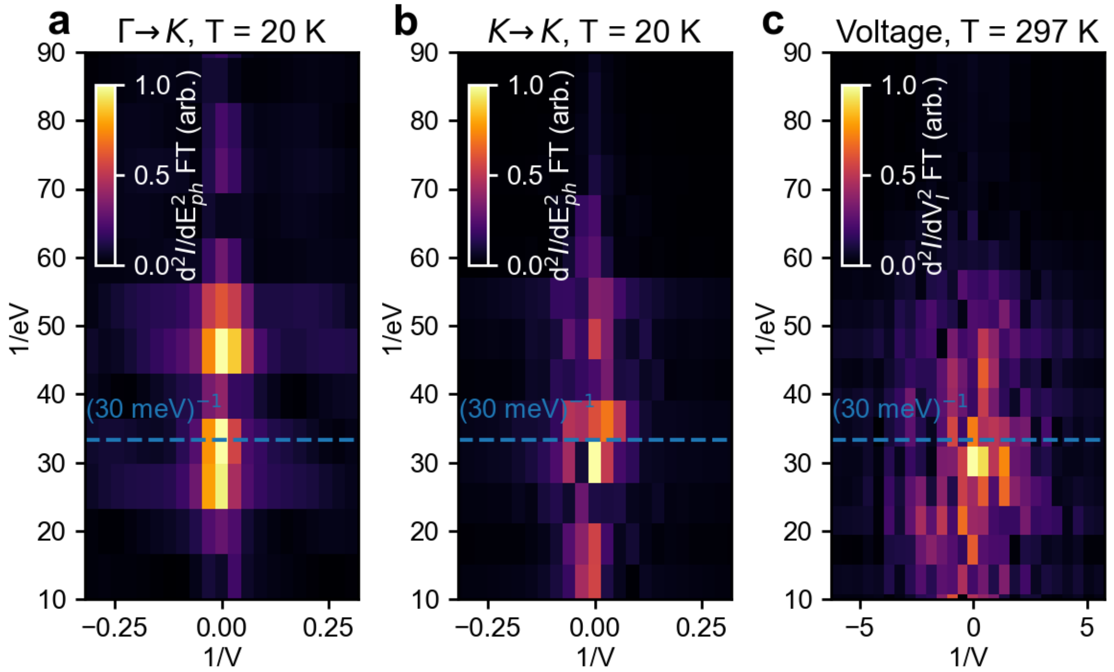


Figure 4.18: Summary of the 30 meV oscillations from multiple experiments. **a** and **b** taken from figure 4.9 c,f showing the exciton transition they measure. **c**, shows the voltage mode from figure 4.14d. In all cases there is a bright mode at (30 meV)$^{-1}$ marked with a blue dashed line.

To begin, we consider what happens when a phonon is localized on the same scale as the electron wavefunction. An exciton is confined to exist within the material that hosts it, in the case an interlayer exciton it is even more confined than normal as the electron and hole must exist in different materials. In a stacked van der Waals heterostructure that means that the exciton is highly localized, having a vertical extent approximately equal to the thickness of the heterostructure. Thus, an optical phonon, like the candidate phonon, is vibrating the lattice over a large amount of the exciton wavefunction. Therefore, the exciton experiences a potential that is oscillating in time at the frequency of the phonon.

To model the effect of a time varying potential on the exciton we follow the basic formalism that is used when considering the effect of a time varying electric potential in electronic transport (the Tien-Gordon effect) or an optically applied potential, (Floquet states)[71,72]. In either case, the answer comes from simple and general statements about the quantum mechanics of periodic time varying potential, regardless of the source of that potential. In the absence of time-variation, the time-dependent wavefunction of an electron with energy E is given by:

$$\Psi(\boldsymbol{r}, t) = \exp\left[-\frac{i}{\hbar}Et\right]\psi(\boldsymbol{r}) \tag{4.6}$$

Where $\psi(\boldsymbol{r})$ is the time-independent wavefunction, and the rest is a time varying phase related to the energy. This time varying phase is not observable, any observable quantity will integrate it out, and is usually not considered. The presence of a time-varying potential, $V(t)$, modifies this phase as such:

$$\Psi(\boldsymbol{r}, t) = \exp\left[-\frac{i}{\hbar}Et - \frac{ie}{\hbar}\int^{t} V(t')dt'\right]\psi(\boldsymbol{r}) \tag{4.7}$$

However, despite the shift in phase, if the potential is spatially symmetric the phase remains unobservable. Intuitively, this means that no matter what is happening to an electron in instantaneous time, if the spatial potential is uniform, it's motion will not be affected, thus no observable quantity will be affected. However, the un-observability of this phase depends on spatially symmetry, if there is a spatially variable potential this phase becomes relevant. Intuitively, if there is a potential difference across the wavefunction, then the electron can move at any instantaneous time, thus it matters what is happening to the electron wavefunction as a function of time.

In our system the electron and hole that make up an exciton experience both a time varying potential and a symmetry breaking spatial potential (in addition to the normal spatial potential provided by the crystal). The longitudinal optical phonon vibrates the lattice underneath the electron providing a potential, $V(t) = V_0 \sin(\Omega t)$ where Ω is the frequency of the phonon. Furthermore, as we calculated in section 4.7 there is a potential difference $\Delta E = e\, V_{SD}/\alpha\beta$ across the heterostructure, i.e. across the exciton wavefunction. This spatially varying potential is what allows the charge neutral exciton to dissociate and produce observable photocurrent, thus if it is related to the time varying phase, that could result in more or less photocurrent.

In summary, our system has all the necessary pieces for this phase to have an observable effect. If we put the potential on the exciton, $V(\boldsymbol{r}, t) = \Delta E(\boldsymbol{r}) + V_0 \sin(\Omega t)$, into equation 4.7 and series expand the phase factor, we expect that the interlayer excitons have a time-dependent wavefunction given by:

$$\Psi(\boldsymbol{r}, t) = \sum_{n=-\infty}^{\infty} a_n \exp\left[-\frac{i}{\hbar}(E_{IX} + eV_{SD}/\alpha\beta - n\hbar\Omega\,)\right]\psi_{IX}(\boldsymbol{r}) \qquad (4.8)$$

Where E_{IX} is the energy of the interlayer exciton given by the bands and ψ_{IX} is the time independent wavefunction of the interlayer exciton. From equation 4.8 we see a wavefunction made up of a mixture of components with energy $E_{IX} + eV_{SD}/\alpha\beta - n\hbar\Omega$, a difference in energy which could have observable consequences.

Even given this shift in the wavefunction, any effect would normally be difficult to observe since the energy of a phonon mode, $\hbar\Omega$, would normally depend on wavevector of the phonon, creating a severe momentum constraint. With that constraint any shift in energy would be disorganized, with many phonon energies contributing and resulting a wide dispersion likely unobservable from the broad spectrum normally observed for excitons. Fortunately, we identified that the candidate phonon modes have flat bands. In other words, those phonons will have approximately 30 meV of energy regardless of their wavevector. Therefore, the momentum constraint is relaxed, and the absorption of a photon depends only on the conservation of energy. Thus, absorption will occur, and an exciton will be produced, when the photon energy equals the energy of a component of the exciton wavefunction given by equation 4.8, thus the condition for absorption is:

$$E_{PH} = E_{IX} + eV_{SD}/\alpha\beta - n\hbar\Omega \tag{4.9}$$

Therefore, as either E_{PH} or V_{SD} is varied the photon energy will come into or go out of resonance with some component of the wavefunction. Since the components are evenly spaced in energy and there is no phonon momentum constraint, this resonance should result in peaks periodically spaced by $\hbar\Omega = 30$ meV and this should happen everywhere in the Brillion zone no matter what E_{IX} is. In other words, this phonon-assisted absorption process satisfies all the properties of the photocurrent oscillations that we observed from Fig. 4.18. Therefore, we conclude that the photocurrent oscillation are the result of strong exciton-phonon coupling in this system.

4.10 Conclusions: Controlling Excitons Unveils Interactions

In this chapter we explored the behavior of excitons in two model heterostructures containing TMDs and observed rich interactions between both the exciton and the lattice, and between excitons and other excitons. The diverse phenomena occurring in these systems at room temperature, and our ability to observe them through optical excitation, demonstrates that non-equilibrium excitons are more complex and more experimentally accessible in 2D materials. Furthermore, this chapter shows the usefulness of the data-intensive approach and how it can unveil subtle interactions from a parameter space filled with complex signal. The ability to observe and image these non-equilibrium excitons with a high degree of data density also gives a much greater degree of control for non-equilibrium states in these systems. This control is important, not simply because it allows us to explore these states, but also because we may be able to use it for scientific and technological applications in the future.

CHAPTER 5:

THE MESOSCOPIC SCALE

5.1 Correlated Phenomena at the Mesoscale

In the previous chapter we studied the non-equilibrium physics of excitons and their interactions with other excitons, leading to Auger recombination, and with vibrations of the lattice, i.e. phonons. This explored physics on the Microscopic Scale, that is we focused on the two or three body interactions of the excitons, or the effect of local lattice vibrations on a single exciton. Implicit in this was that interactions were confined to that scale, and that a few single discrete interactions determined their properties. But this assumption begins to break down at high density, if we increase the number of charge carriers in a constant volume they need to interact more. Eventually they start interacting so much that it's useless to consider what happens to a single electron or hole, rather we think about the properties of a phase consisting of many correlated electrons and holes. To do this we need to zoom out from the microscopic scale and study what happens at the mesoscopic scale.

There are two strongly interacting non-equilibrium phases that we will examine here, the electron-hole liquid in $MoTe_2$, and the hot carrier regime of graphene. In $MoTe_2$ the density of excitons can be increased to the point that the dynamics undergo a phase transition to a strongly interacting phase. In graphene excited charge carriers couple to each on a faster timescale than they couple to the lattice, resulting in a "hot carrier, cool lattice" phase where a population of high energy carriers can exist for several picoseconds before in cools.

5.2 The High-Power Behavior of the Graphene-MoTe$_2$ System

In the first half of chapter 4 we detailed the interactions of excitons in a Graphene-MoTe$_2$ heterostructure, concluding that the dynamics are dominated by two-body Auger recombination of biexcitons. In this picture, the excitons are essentially in a non-ideal gas phase, that is they interact with the container they are in and occasionally interact with each other. In this chapter we will discuss the breakdown of this limit, where excitons are densely packed and their dynamics become driven by numerous interactions, that is when it no longer makes sense to think about two-body or three-body interactions, but rather many-body physics takes over. Experimentally, this will correspond to the high-density limit which can be achieved by increasing the power of laser pulses. Thus, we will return to the *GMoTe2* sample (details in appendix 1.3) and explore the high-power parameter space.

We begin by examining the spatial power dependence of the *GMoTe2* sample. At long Δt the sample exhibits a photocurrent response that is spatially simple; it is maximized in the center of the sample and near the edges falls on, as shown as both color and height in Fig. 5.1a. This spatial profile expected due to the shape of the gaussian laser beamspot, and as we vary the power we see that the scale of the photocurrent increases but the profile does not change. This is further confirmed when we examine the non-linearity of the spatial photocurrent. From a large set of photocurrents images, the interlayer photocurrent vs. optical power is fit to $I \propto P^{\gamma}$ at each point in space, giving the spatial non-linearity $\gamma(x, y)$, shown in Fig. 5.1b, which we see if mostly uniform over the area of the sample. In this regime, the spatial non-linearity γ does a good job of capturing the dynamics of the system, i.e. those discussed in Chapter 4, thus following the MPDPM process we use the 2D maps of γ to represent the behavior of the sample from a 3D set of images as a function of power.

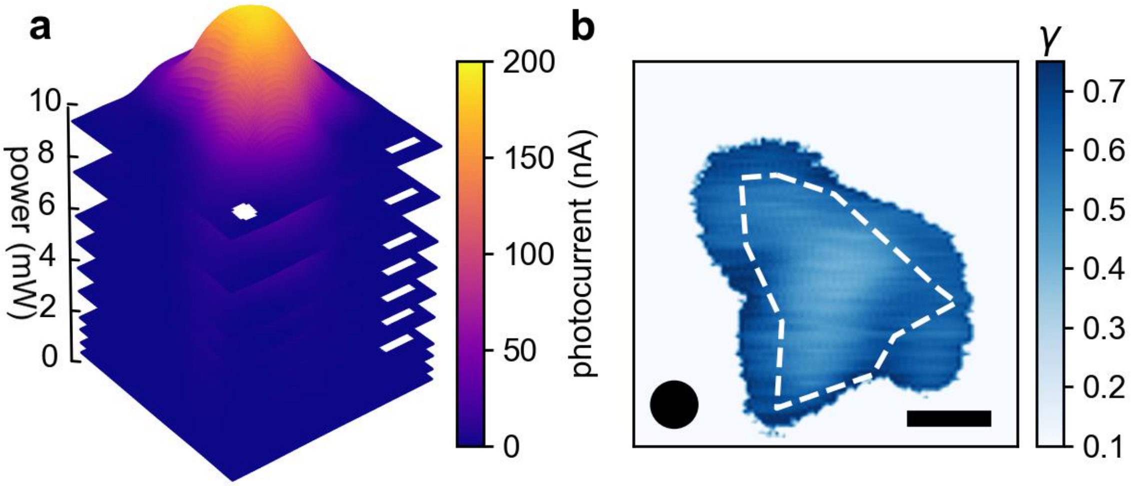


Figure 5.1: **a** Photocurrent images for increasing laser power; at room temperature, wavelength λ = 1200 nm, and time delay Δt = 120 ps. With photocurrent shown as both height above a plane and color. **b**, the power dependence versus space represented by the power law exponent, $\gamma(x, y)$ where the dashed line outlines the graphene-$MoTe_2$ heterostructure region. Scale bars are 3 μm and circles indicate the full width at half maximum (FWHM) of the diffraction-limited beamspot.

We want to explore the high exciton density regime, and there are two parameters we could tune to achieve high density. We could increase the density by increasing the power of the laser, and thus the carriers generated per pulse. Alternately, we could increase the density by decreasing $|\Delta t|$ thus causing carriers from the two pulses to overlap, increasing the instantaneous density of carriers (as shown schematically in the Fig. 4.4b insets). Given that γ extracted from the power dependence is a good dynamical parameter, we want to watch γ as a function of Δt to see if the dynamics change as the pulses overlap. Fig. 5.2 shows the spatial γ map as a function of Δt, reduced from a four-dimensional data set of spatial images versus power and two pulse delay.

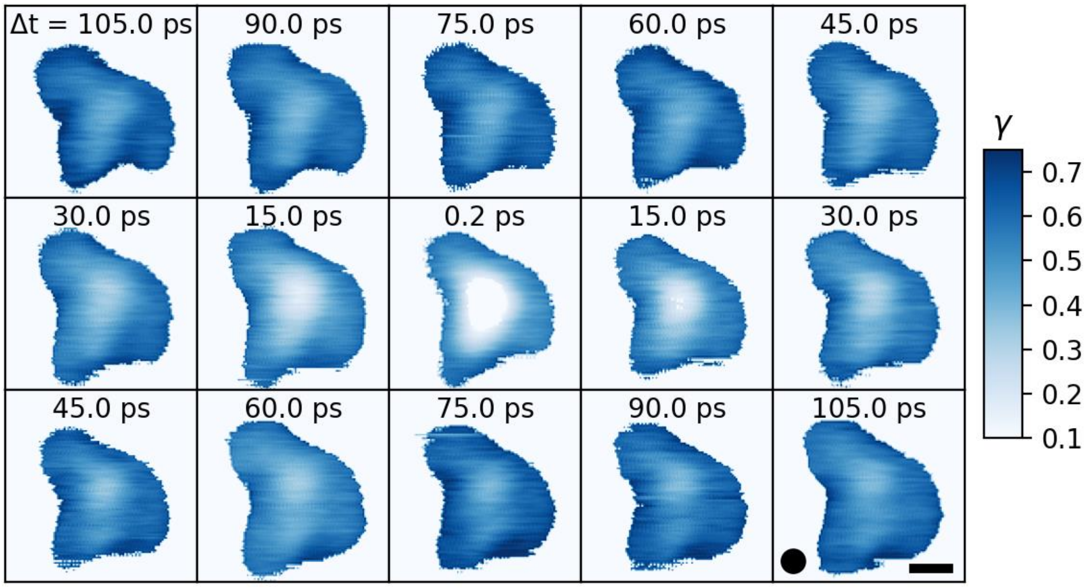


Figure 5.2: Spatial γ maps versus two-pulse delay Δt. Represents the laser power versus Δt parameter space. In the bottom right panel, the scale bars is 3 μm and the circle indicates the FWHM of the diffraction-limited beamspot.

Strikingly, when the time delay between laser pulses is very short, MPDPM reveals highly anomalous photoresponse. At long time delay (Δt > 30 ps) the spatial non-linearity is highly uniform over the sample area. But in the center row of Fig. 5.2, $|\Delta t| \leq$ 30 ps, we see a bright feature emerge in the center of the sample the non-linearity increases. At Δt = 0.2 ps, the power law behavior collapses near the center, resulting in a pronounced ring of sublinear photoresponse (γ ~ 0.5). The area of power law suppression significantly exceeds the beam spot size, indicating a global change in photoresponse.

This behavior is not a permanent change of device response. Instead, we observe γ suppression only at short time delay, while the ordinary photoresponse is recovered as soon as the laser intensity is reduced. In the following, we examine the space-time evolution of the MoTe$_2$ photoresponse, and extract detailed dependence of the spatial photocurrent features on optical power, interlayer voltage, and time delay.

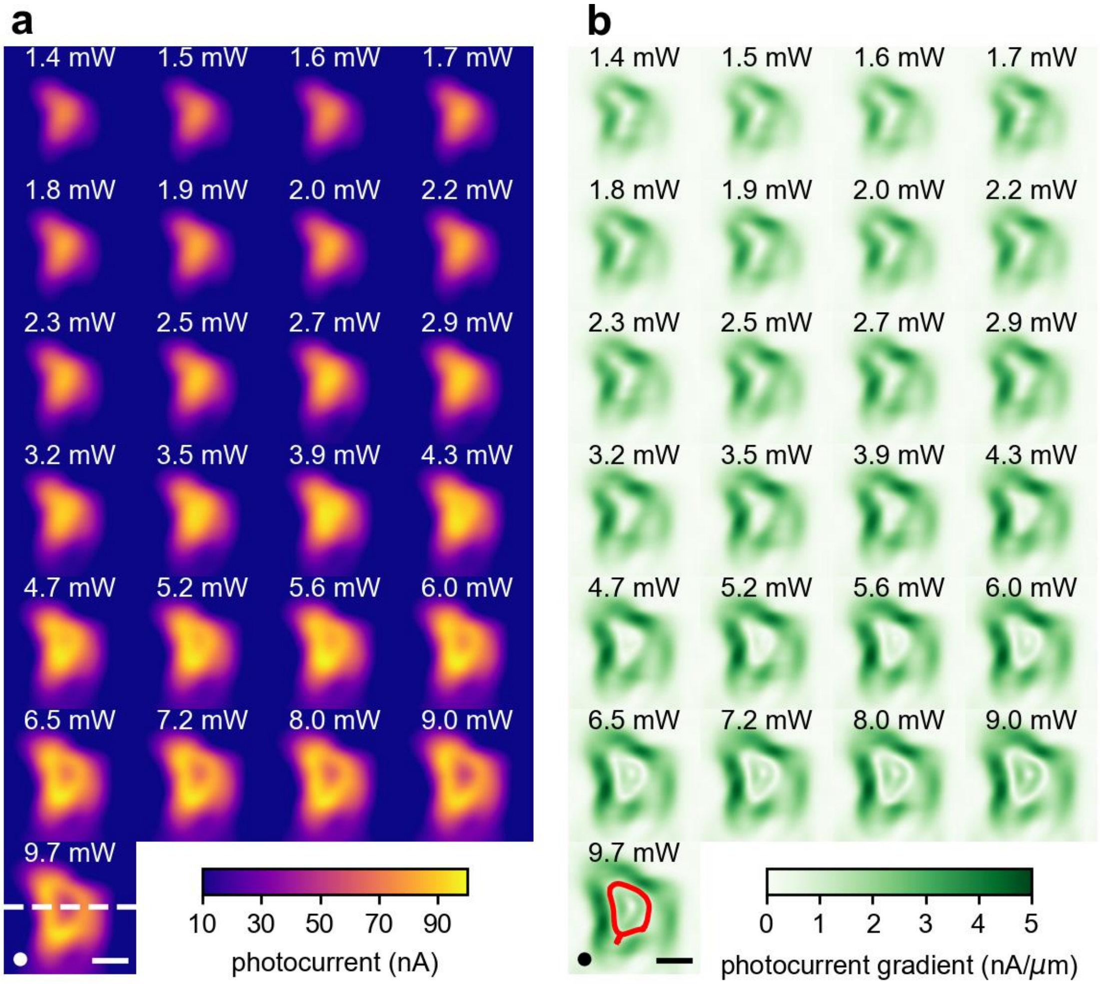


Figure 5.3: **a**, spatially resolved photocurrent maps. Dashed line indicates location of photocurrent line profiles. **b**, photocurrent gradient magnitude $|\nabla I|$. Photocurrent was measured at various powers (labeled) and Δt = 0.2 ps. The contour $|\nabla I| \approx 0$ encloses the photocurrent ring (red contour in the image at P = 12.2 mW). Scale bars 5 μm, circles indicate the beamspot FWHM.

In Fig. 5.3a we decompose the photocurrent image set measured at $\Delta t = 0.2$ ps (center image in Fig. 5.2) and examine the constituent photocurrent maps. At low optical powers, the photocurrent magnitude increases rapidly and monotonically as seen before in Fig. 5.1a. For $P > 5$ mW, however, the photocurrent at the center of the device suddenly decreases, forming a photocurrent ring of bright photoresponse. The photocurrent ring grows rapidly with increasing optical power. To see the ring expansion more clearly, Fig. 5.3b shows the magnitude of the spatial photocurrent gradient $|\nabla I|$, which visualize the local slope of the spatially resolved photocurrent landscape. At $P = 5$ mW, a clear edge begins to emerge and grows into a well-formed ring.

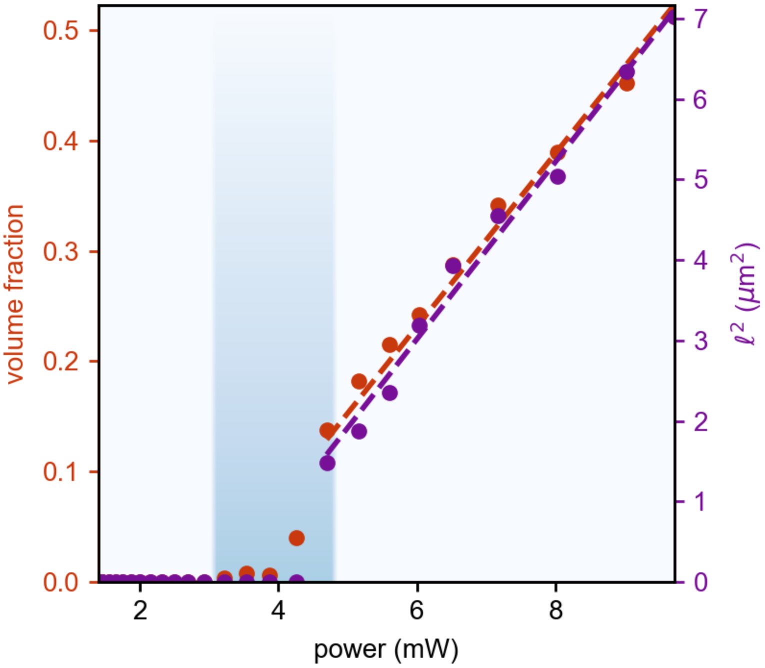


Figure 5.4: Photocurrent ring volume fraction (red data) vs. laser power, from the maps in Fig. 5.3. Volume fraction is the ratio of the volume enclosed by the $|\nabla I| \approx 0$ contours to the active photocell volume (area). Peak-to-valley distance ℓ (purple data) vs. optical power extracted from photocurrent line cuts in a. Red (purple) dashed lines are linear fits to volume fraction (ℓ^2) vs. optical power above $P = 5$ mW.

Remarkably, the anomalous photocurrent ring appears abruptly with increasing optical power. To quantify the ring feature, we note that in the gradient maps in Fig. 5.3b the ring feature corresponds to a contour where the gradient is zero, i.e. where the photocurrent stops increasing and starts to decrease. To pick out this contour we use an iterative algorithm, which starts at the edges of the photocurrent and at each iteration uses the gradient vector as a "force" to accelerate the points of the contour. This will cause the contour to "climb" the photocurrent, shrinking inwards as it moves up along the gradient. If there is no dip feature the contour will shrink until it has zero size at the maximum. If there is a dip feature, then the contour will get stuck along the edge of the dip, because the gradient is zero at that point. Eventually the contour converges either to zero or to a stable finite size. One such contour is shown in red in the bottom left of Fig. 5.3b.

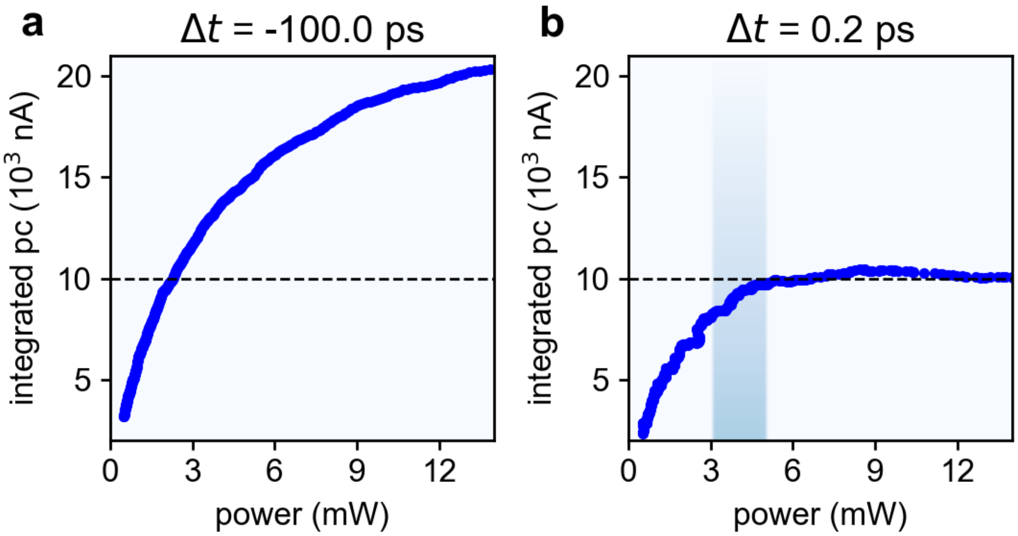


Figure 5.5: Spatially integrated photocurrent versus laser power at **a** Δt = -100 ps and **b** Δt = 0.2 ps. Dashed line indicates the level that the Δt = 0.2 ps trends to. Blue shaded area is the same as in Fig. 5.4.

Fig. 5.4 shows the power dependence of the ring volume (red points) as a fraction of the total heterostructure volume obtained by dividing the area in the contour by the sample area. At a critical power P_C = 5 mW, we observed a nearly discontinuous growth rate of the volume fraction. Above the transition $P > P_C$, the photocurrent ring, and thus volume fraction, expands linearly with optical power. To confirm this result we can also look at line cuts (for example the white dashed line in Fig. 5.3a lower left) and find the distance between the dip and the maximum, which we call ℓ. Taking ℓ^2, purple points in Fig. 5.4, we see that it follows the same linear trend as the volume fraction.

The sharp transition at P_C also manifests as a sudden deviation from power law behavior. Fig. 5.5 shows the spatially integrated photocurrent vs. power measured along the dashed line in Fig. 5.3a lower left. Fig. 5.5a shows the integrated photocurrent at Δt = -100 ps, and we observe sublinear but monotonically increasing power dependence, consistent with the behavior discussed in chapter 4. Fig. 5.5b shows the integrated photocurrent at Δt = 0.2 ps. The photocurrent increases rapidly at low power and exhibits ordinary sublinear growth, but above $P = P_C$, the spatially integrated photocurrent remains nearly constant as power increases. Thus, the abrupt formation and expansion of the photocurrent ring corresponds directly to the collapse of power law behavior observed in Fig. 5.2. For $|\Delta t|$ > 0.2 ps the same behavior occurs at higher P_C since the pulses become separated in time and the effect of each individual pulse is weaker than when they are combined.

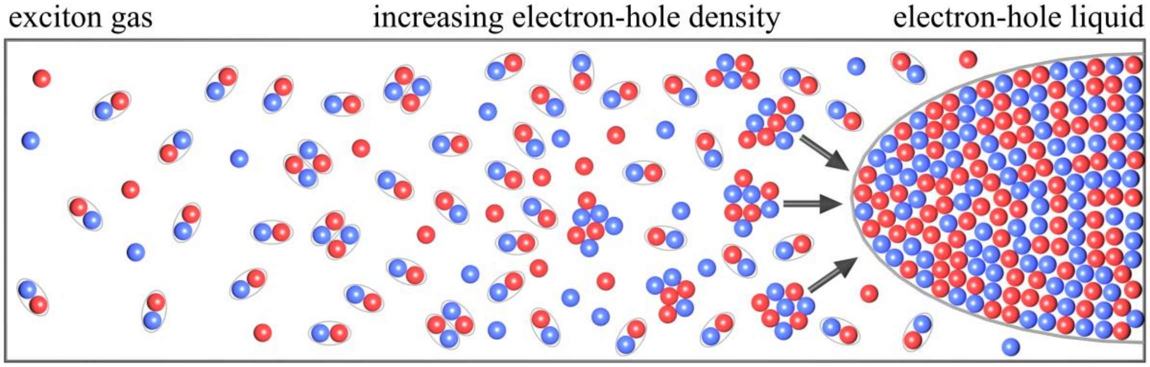


Figure 5.6: The dynamics electrons and holes in a TMD as a function of increasing charge carrier density

5.3 Modeling the Photoresponse of an Electron Hole Liquid

The anomalous photoresponse at room temperature is reminiscent of the gas-to-liquid phase transition of electron-hole (*e*-*h*) pairs in conventional semiconductors - such as Si, Ge, GaAs, and CdS - at low temperatures[32,33,73,74]. At low laser power, photoexcitation generates a gas of electrons and holes, many of which will form excitons. Enhanced Coulomb interactions bind electrons and holes into excitons with nanometer-scale Bohr radius a_B. The charge carrier density increases with laser power until exciton-exciton interactions become comparable to interactions within an individual exciton pair. Below the power threshold for the phase transition P_C, ordinary two-body exciton-exciton annihilation processes dominate the optoelectronic properties, as discussed in chapter 4.

At the critical laser power P_C, the electron-hole population merges into a non-equilibrium many-body phase (Fig. 5.6). The density of electrons, holes and excitons N becomes so large that the average spacing between pairs is nearly equal to the exciton radius. At $P_C = 5$ mW, the mean exciton-exciton separation, which we estimate to be $a_{xx} = 1\text{-}3$ nm in $MoTe_2$ (see appendix A2.3), is very close to the Bohr radius $a_B = 2.3$ nm extracted from magneto-optical measurements[75]. Once $a_{\mathrm{xx}} \sim a_{\mathrm{B}}$, the electron-hole (*e*-*h*) population reaches the critical density $N_C \sim 0.5/\mathrm{nm}^3$ *e*-*h* pairs. This density N_{C}, which is determined by the renormalized minimum energy per electron-hole pair, remains constant in the liquid phase. Above the phase transition, the renormalization of the energy per *e*-*h* pair results in a suppression of photon absorption within the *e*-*h* liquid[76,77]. The resulting *e*-*h* liquid exhibits a fixed *e*-*h* pair density N_{C}, is highly polarizable in an applied electric field, and forms a sharp, stable boundary that separates it from the gas phase[32,33,73].

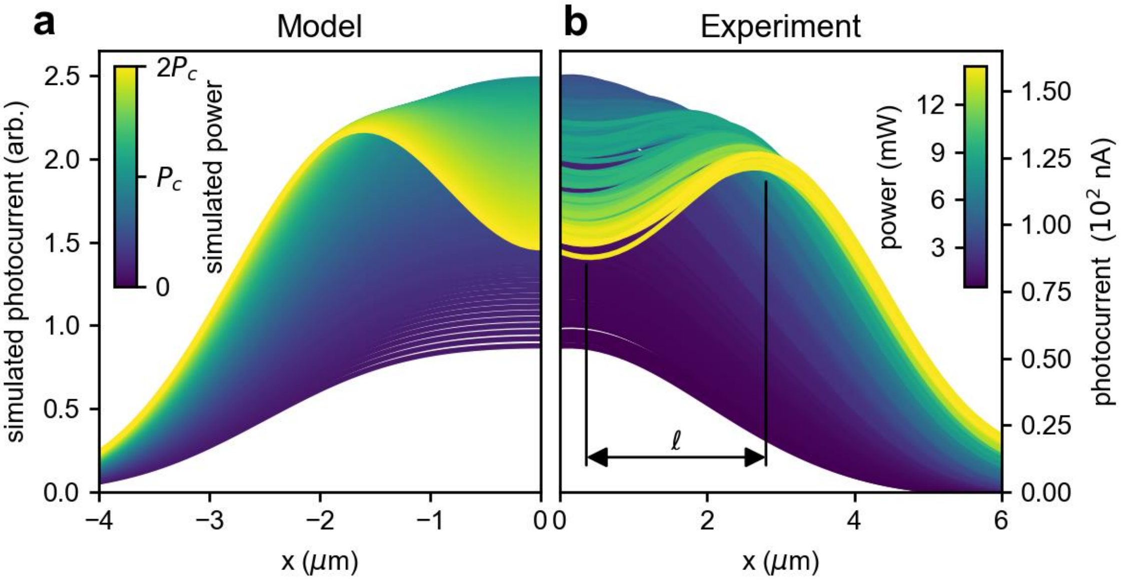


Figure 5.7: **a,** calculated spatially resolved photoresponse of the electron-hole condensate showing the suppression of photocurrent in the center of the sample above the critical power P_c. **b,** photocurrent line profiles measured across the center of the sample for increasing power; $T = 297$ K, $\Delta t = 0.2$ ps.

We attribute the anomalous photoresponse in the *GMoTe2* sample to an electron-hole liquid phase, the properties of which are readily revealed through MPDPM. First, the liquid phase is characterized by highly unusual ring-like interlayer photoresponse, which results from fast interlayer transit of liberated electrons and holes near the edge of the droplet. At the surface of the device, the excitation laser forms a diffraction-limited Gaussian beam spot with a full width at half maximum of 1.67 microns at $\lambda = 1200$ nm. When the local maximum power of the Gaussian beam exceeds the critical power threshold P_C, an electron-hole liquid droplet forms near the center of the beam spot. The observed photocurrent results from electrons and holes outside of the electron-hole liquid droplet as the recombination rate goes up inside the droplet, meaning that carriers inside the droplet do not contribute to the photocurrent. The characteristic ring shape then arises from the convolution of the beam spot with a sharply bound region of suppressed absorption (see appendix A2.4 for detailed calculation). Fig. 5.7a and 5.7b compare the *e-h* liquid model to interlayer photocurrent line traces as a function of increasing power (measured along the dashed line in Fig. 5.3a). The *e-h* liquid model in Fig. 5.7a shows excellent agreement with the photocurrent line profiles, indicating that the interlayer photocurrent is suppressed in the region of the anomalous phase.

Once the critical density N_C is reached, energy added to the *e-h* liquid contributes exclusively to expansion. We see this in both ℓ^2 which is the effective radius of the photocurrent ring, and in the volume fraction which both exhibit a sudden onset then linear expansion. From the data, we conclude that the state leading to suppressed photoresponse exhibits a well-defined volume that is highly localized, which increases as the number of electron-hole pairs (proportional to the power) increases. We thus deduce that, similar to a conventional incompressible liquid, the condensate density N_C remains fixed, and a linear increase of *e-h* pair density yields a linear increase in liquid volume.

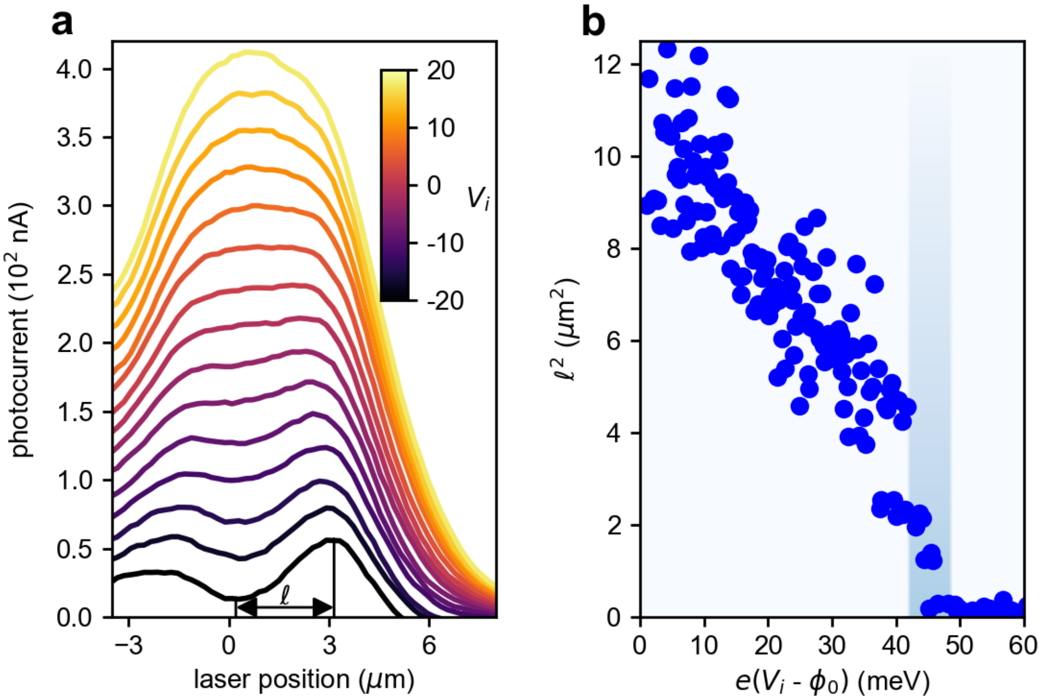


Figure 5.8: **a,** spatial line cuts of the photocurrent as a function of interlayer voltage V_i; $\Delta t = 0.2$ ps. **b,** photocurrent ring peak-to-valley distance ℓ^2 vs. interlayer voltage, where ϕ_0 is the built-in potential of the *GMoTe2* heterostructure.

5.4 Signatures of the Electron Hole Liquid

In addition to the spatial profile, there are other signatures of the electron hole liquid that we can look for in the MPDPM data. The *e-h* liquid phase is polarizable and can be dissociated with interlayer voltage. Fig. 5.8a shows spatial photocurrent line cuts as a function of voltage, where the size of the *e-h* liquid droplet can be tracked via ℓ. For interlayer voltages V_i above the built-in potential $\phi_0 = -41$ mV (extracted from the data in Fig. 4.2c), ℓ^2 decreases approximately linearly as voltage increases (Fig. 5.8b). When the total interlayer voltage exceeds the critical voltage $eV_\mathrm{C} = e(V_\mathrm{i} - \phi_0) > 45$ meV, electrons and holes become ordinary *e-h* pairs. Above the critical interlayer voltage V_C, spatially uniform photocurrent re-emerges as the 2D liquid dissociates in the electric field.

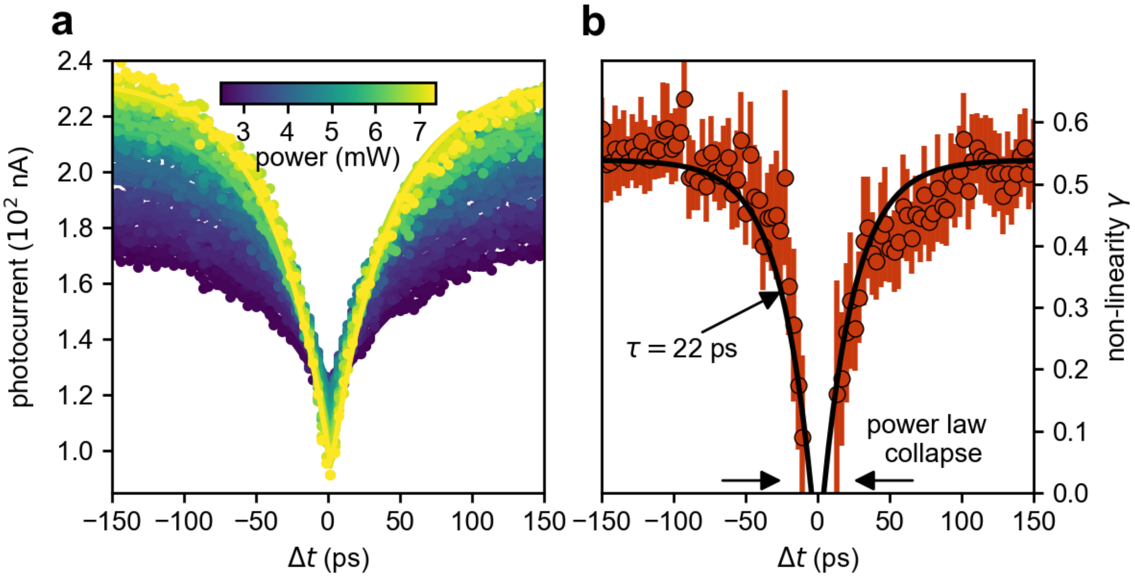


Figure 5.9: **a,** photocurrent vs. Δt for increasing optical power, with laser is fixed at the center of the device. Solid lines are exponential fits to the data at each power. **b**, power law exponent γ as a function of Δt, extracted by fitting the data in a to $I \propto P^{\gamma}$. Solid black line is an exponential fit with a characteristic timescale $\tau = 22$ ps.

Though a single pulse at high enough power could form the *e-h* liquid, two-pulse measurements can also reveal rich dynamic transition between the electron-hole liquid and gas phase. Fig. 5.9a and 5.9b show the time-resolved photocurrent and photocurrent nonlinearity γ. When the laser is fixed at the center of the device, the photocurrent vs. Δt exhibits remarkably different power dependence between short and long Δt. At short Δt, the photocurrent at the center of the device decreases with increasing power above P_C. This extreme sublinear photoresponse is fully consistent with power law collapse associated with the liquid phase onset. At long time delay, the photocurrent exhibits ordinary sublinear photoresponse. We fit γ vs. Δt to an exponential decay (black line Fig. 5.9b) to extract the charge density persistence time $\tau = 22$ ps. The timescale in the presence of the liquid is significantly smaller than the 86 ps recombination timescale determined in Section 4.5, which is consistent with our expectation that the recombination rate goes up inside the liquid. At and above the 5 mW threshold, for $|\Delta t| > \tau$, the pulses are independent and each is insufficient to drive the phase transition producing only a gas of ordinary e-h pairs. For $|\Delta t| < \tau$ the combined charge density produced by the two pulses is sufficient to cause the gas-to-liquid transition causing a suppression of the photocurrent.

5.5 Implications of Room Temperature Electron-Hole Liquid

In quasi-2D TMDs, the large binding energy and strong Coulomb interactions combine to allow for an exotic e-h liquid phase diagram at room temperature, suggesting new device applications that harness electronic fluids at room temperature[74]. The gas-to-liquid phase transition is set by the energy difference *ΔE* between the average energy per *e*-*h* pair in the gas phase and the reduced energy per *e*-*h* pair in the liquid phase. When *ΔE* is large compared to thermal energy at room temperature ($k_B T$ = 26 meV), the liquid is stable against thermal fluctuations. From the interlayer voltage dependence (Fig. 5.8), we estimate that $\Delta \mathrm{E} \sim eV_C \sim 45$ meV, approximately twice the thermal energy at room temperature. While this renormalization is comparable to conventional 2D electron systems, the *e*-*h* pair binding energy (~10^2 meV) in TMDs is several orders of magnitude larger[75-79]. Electrons and holes in the liquid phase move independently from another within the confined liquid volume, yet they are not able to escape without sufficient excess energy. Such stable collective excitations - potentially exhibiting very high mobility - could have applications in very high-power, high frequency Terahertz sources and detectors that can be manipulated with both electronic and optical control.

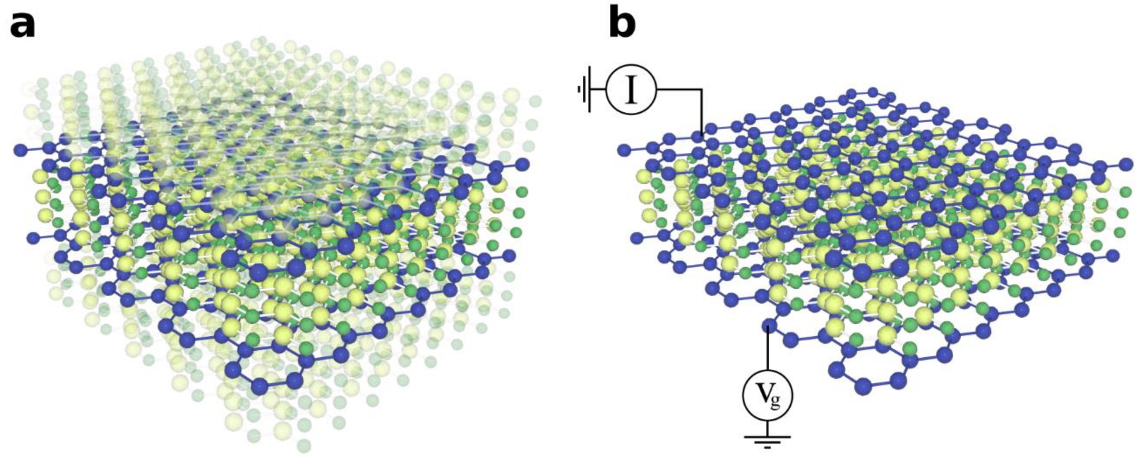


Figure 5.10: **a,** Schematic of the Graphene-hBN-Graphene heterostructure (samples *GBNG-1,2,3*), encapsulating layers semi-transparent. **b**, a *GBNG* sample contacted in the tunneling configuration, encapsulating layers omitted.

5.6 The Graphene-hBN-Graphene Heterostructure System

Despite being the first 2D material isolated and having the simplest composition, graphene continues to be an interesting material system to study in a variety of 2D materials research contexts. In optoelectronics, an interesting and important property of the graphene is the hot carrier regime, where carriers are optically excited and strongly couple to other carriers and couple more weakly to the lattice via electron-phonon interactions. The result is that a short time after light is absorbed, on the order of hundreds of femtoseconds, there is a population of electrons that has thermalized among itself, but not with the lattice, and is in a quasi-thermal equilibrium with an electronic temperature in the thousands of Kelvin, orders of magnitude larger than the lattice temperature[34-38]. This state then cools relatively slowly for several picoseconds before coming into equilibrium with the lattice. Much work has gone into understanding the cooling pathways that allow the graphene to relax and the photo-multiplicative interactions that occur during thermalization and cooling[80-83].

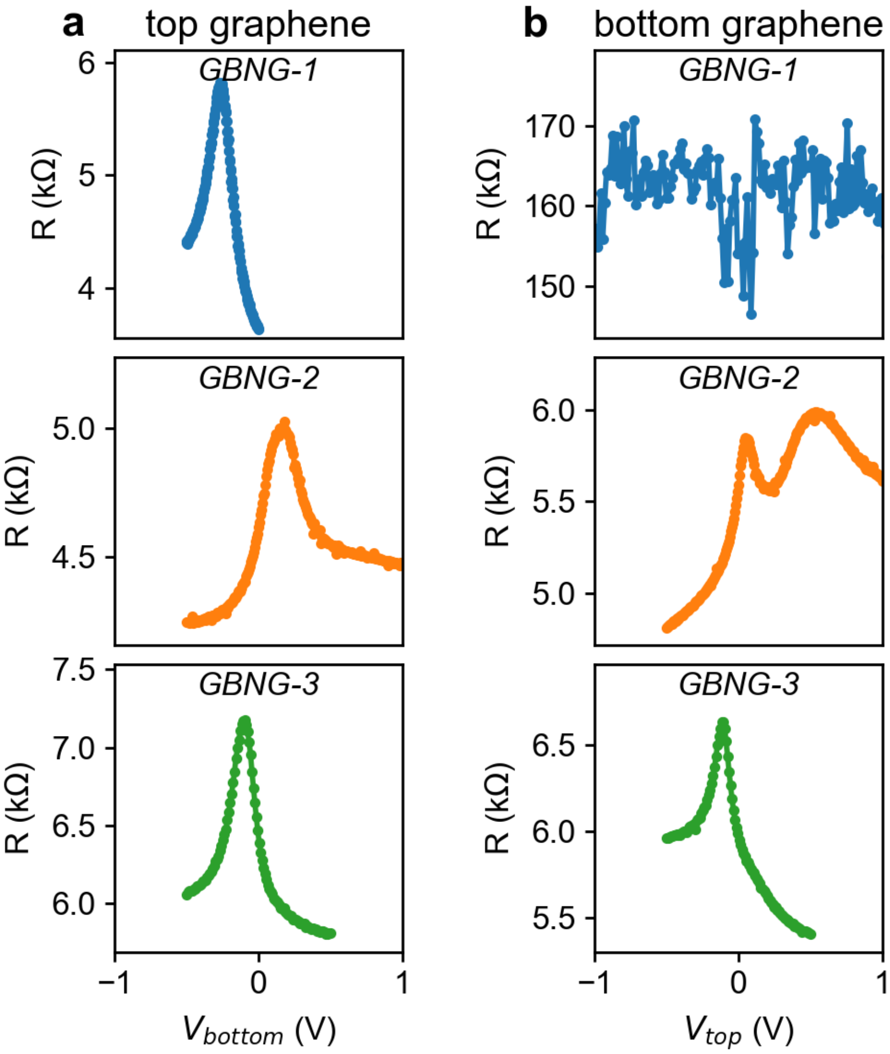


Figure 5.11: Resistance data for all three devices in the region of the Dirac point. **a,** the resistance of the top graphene, gated by the bottom. **b**, the resistance of the bottom graphene gated by the top.

This hot-carrier state exists on the mesoscopic scale, in that it is characterized by strong interactions of many charge carriers rather than the dynamics of individual carriers. Therefore, what we are interested in studying on this scale is not the collisions of the carriers but rather what the collective phase of hot carriers can do. To study the hot carrier phase, we explore a heterostructure composed of two sheets of graphene separated by a thin layer of hBN, as shown in Fig. 5.10a. This system is known to produce photocurrent between the graphene layers when photoexcited due to Fowler-Nordheim tunneling, direct tunneling, and tunneling due to thermal effects believed to be the result of the hot carrier phase[84]. We aim to explore this system in the low voltage regime while optically exciting well below the gap of the hBN, thus preventing Fowler-Nordheim and direct tunneling and leaving only hot-carrier tunneling.

Three graphene-hBN-graphene (*GBNG*) heterostructures will be explored in this chapter. The details of these devices can be found in appendix A1.6, but they all have the same basic structure; two sheets of monolayer graphene separated by hBN in the range of 7 to 10 nm thick. The ideal system will have identical graphene layers, with the only differences due to applied voltages, however due to the realities of fabrication this is difficult to achieve. For symmetry, the key is to encapsulate both layers in hBN, as shown with semi-transparent layers in Fig. 5.10a and have the fabrication of the graphene layers be highly clean and free of wet chemistry to keep the layers as close to their intrinsic properties as possible. Fig**.** 5.11 shows the resistance of the devices near their Dirac points. We see in the top row that *GBNG-1* is highly asymmetric, with the top layer exhibiting a Dirac peak, but no observable peak in the same range for the bottom layer.

This is due to the fact that the bottom layer rests directly on the SiO_2 substrate, which significantly degrades its quality compared to the top layer on hBN. The second device, *GBNG-2*, has both layers encapsulated with hBN but is geometrically imperfect, with part of the bottom graphene electrostatically screened by the Si substrate, likely resulting in the split Dirac peak in the bottom graphene. *GBNG-3* is the most symmetric, with the Dirac peak in both layers being near zero and near to the other.

5.7 Interlayer-Tunneling Photocurrent Parameter Space

We measure the *GBNG* samples contacted in the tunnel configuration, shown in Fig. 5.10b, which measures the current, I, that passes through the hBN as a function of the interlayer voltage V_g applied to the bottom layer. In this section, we discuss the photocurrent parameter space in the tunneling configuration which consists of spatial dependence, laser power, pulse delay and V_g with the laser wavelength fixed at 1200 nm. First, we examine the spatial dependence, which is fairly straightforward; in the tunneling configuration the heterostructure region exhibits spatially uniform photocurrent, the magnitude and sign of which depends on V_g, as shown in Fig. 5.12 for *GBNG-3*. Given the uniformity all subsequent data shown in this section was taken using spatial line scans, i.e. scanning the laser beam in one spatial dimension over the heterostructure area and averaging the photocurrent from the heterostructure for clarity. Spatial scans were taken between line scans to ensure that the laser remained centered on the heterostructure.

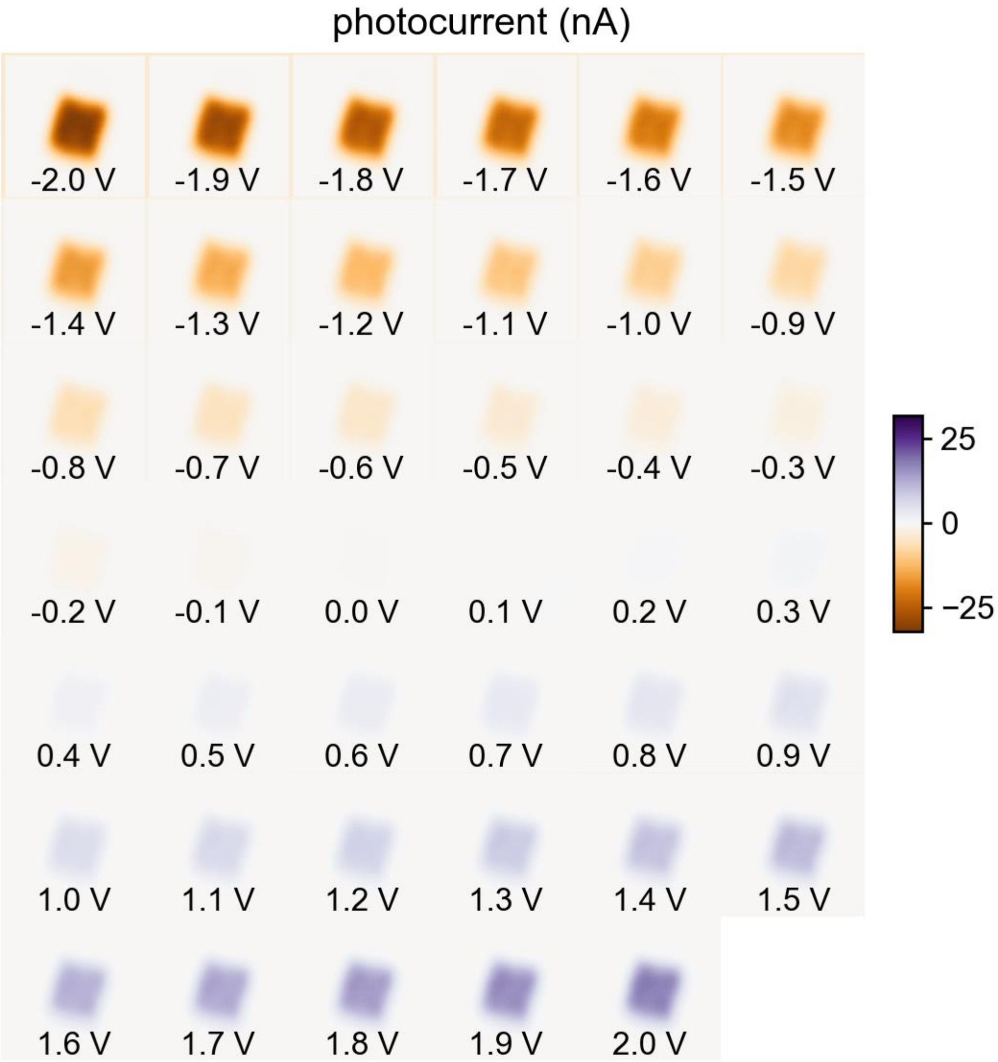


Figure 5.12: Spatial photocurrent maps of the heterostructure area of *GBNG-3* in the tunneling configuration as a function of V_g (labeled). Spatial photocurrent is similarly uniform in the tunneling configuration for *GBNG-1* and *GBNG-2*. Color near the edges in the upper left images are due to spatial drift correction.

Next we examine the photocurrent dependence as a function of the incident laser beam. Fig. 5.13a shows the power dependence of each sample and we observe that it is well described by a power law, $I \propto P^{\gamma}$, (red line) where γ is the key dynamical parameter. We note that γ is significantly higher in the high-symmetry device *GBNG-3*.

Fig. 5.13b shows the photocurrent as a function of two-pulse delay, Δt. We see that it is well described using two symmetric exponential timescales, i.e. $I \propto e^{-|t|/\tau_f} + e^{-|t|/\tau_s}$ where τ_s and τ_f are fast and slow timescales. The fast timescale τ_f is approximately the width of the autocorrelation pattern (see section 2.4) indicating it comes from the pulses overlapping and is not significant. The slow timescale, τ_s, is approximately 2 ps in all the devices indicating a population of excited carriers exists within the device for several picoseconds after initial photoexcitation. Data shown in Fig. 5.13 was taken at arbitrary values of V_g; though the dynamical parameters γ and τ_s vary a small amount as function of V_g the overall behavior is consistent across the V_g ranges discussed below.

The primary data-intensive exploration in the interlayer photocurrent versus interlayer voltage Vg, which is shown in Fig. 5.14. From Fig. 5.14a we see that the photocurrent monotonically increases as a function of V_g and crosses through zero near where the top gate Dirac point is in Fig. 5.11a. The differences in behavior are more clear when we examine the differential photoconductance, shown in Fig. 5.14b, where in all three cases we observe that at large positive and negative V_g the photoconductance increases. In the symmetric *GBNG-3* sample we see that photoconductance is minimized when the current goes to zero, however in the asymmetric sample *GBNG-1* we see a large peak in the photoconductance near zero, with the intermediately symmetric sample, *GBNG-2*, exhibiting a smaller, but clear, peak in the photoconductance.

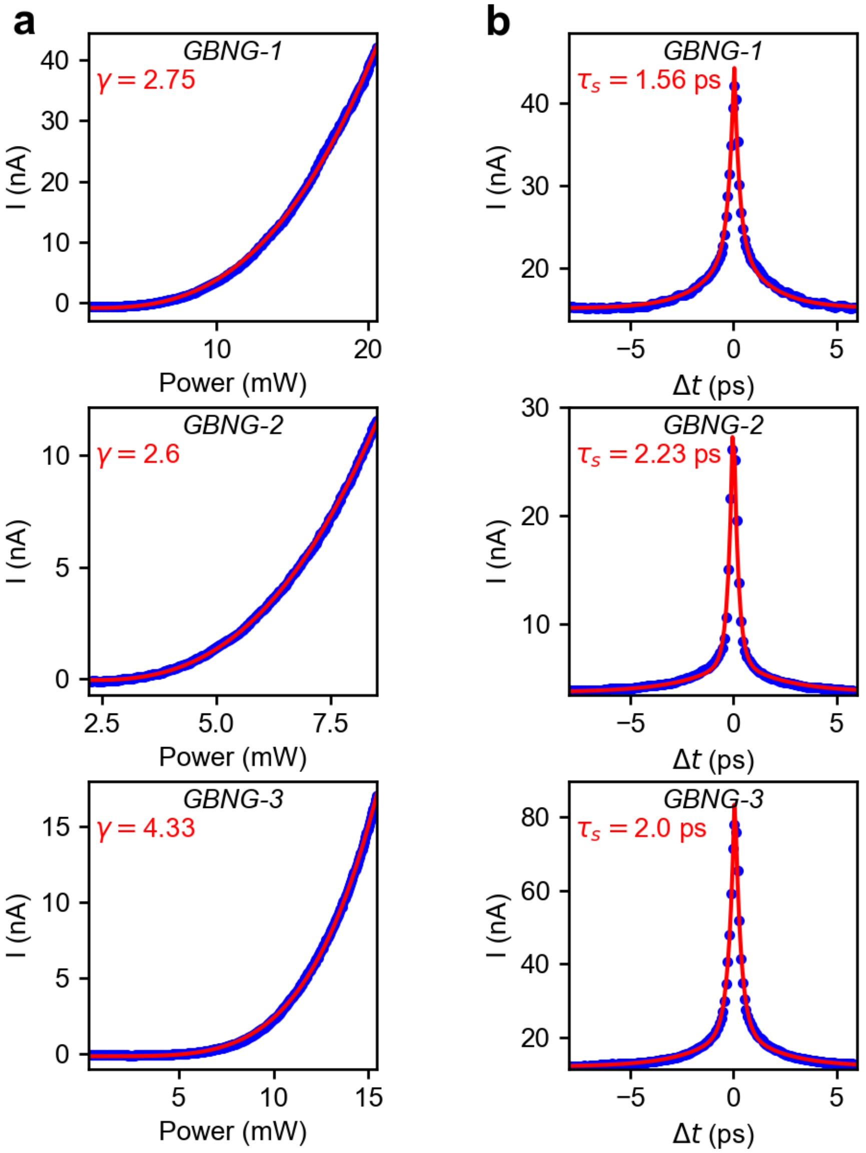


Figure 5.13: Photocurrent dynamic behavior. **a**, power dependence and **b**, Δt dependence in tunnel configuration. Blue points are data, red is a fit to a dynamical function. Scans were taken at values of V_g and laser power that give good signal to noise, variations in V_g may cause minor changes in the values of γ and τ, but overall behavior is consistent across the power-V_g or delay-V_g parameter space.

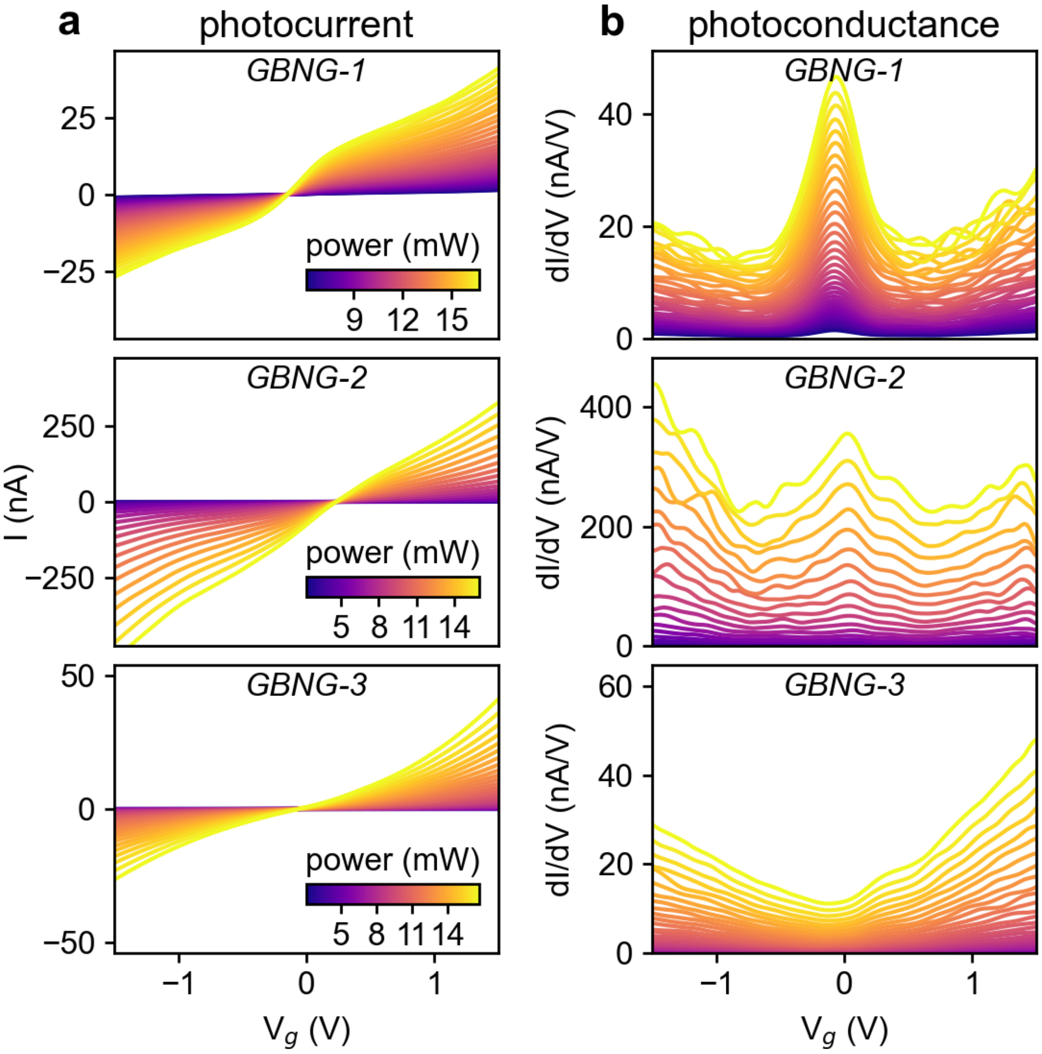


Figure 5.14: Photocurrent, **a**, and photoconductance, **b**, versus V_g for all three samples in the tunneling configuration. Colorscale indicates the incident laser power and power scales are consistent across rows.

5.8 The Hot Carrier Model of Tunneling

The photoexcitation of graphene forms a hot carrier non-equilibrium phase with a wide thermal distribution. The formation of the hot carrier state is described in the literature and Fig. 5.15 identifies how the distribution of photoexcited carriers evolves as a function of time[34,37]. When photons are absorbed at t = 0 the resulting electrons and holes are in a specific narrow set of states located at half the photon energy above and below the charge neutrality point. These carriers undergo two-body collisions and intra-band scattering that thermalize the population of carriers. A few hundred femtoseconds after the light is absorbed a thermal Fermi-Dirac distribution of holes is formed, which persists for up to a few picoseconds before the distribution cools and the system relaxes back to equilibrium.

The hot carrier picture is consistent with the dynamics we see in Fig. 5.13. From the power dependence we observe an increase in the number of absorbed carries due to an increase in power leads to a superlinear response. Then in the time dynamics a superlinear response is observed while the laser pulses are separated by ~2 ps. If the second pulse adds carriers on top of those dwelling in the system, we would expect to see a superlinear response, i.e. more photocurrent than that generated by two separated pulses. Therefore, the presence of a ~2 ps superlinear timescale indicates that hot carriers dwell in the system for approximately that length of time.

In graphene double layer structures interlayer charge transfer through the hexagonal Boron Nitride (hBN) can occur due to direct tunneling, Fowler-Nordheim tunneling, or thermal effects[84]. Fundamentally, all these processes can occur due to the specific band alignment between the graphene and hBN layers. Fig. 5.16 shows the basic band alignment between graphene and hBN, where the edge of the hBN valance band is $U_0 \approx 1.3$ eV away from the charge neutrality point (CNP) of the graphene, with the gap between the CNP and the hBN valance band being 3.5 eV[85]. In the direct tunneling case, holes excited below -U_0 can pass directly into the hBN valence band and conduct through the hBN. In the Fowler-Nordheim case the band-gap edge acts as a barrier modulated by a large interlayer voltage, allowing for tunneling of sub-barrier carriers at high voltage. In the thermal case sub-barrier holes can be excited by thermal processes to reach energies lower than the barrier and conduct through the hBN valance band.

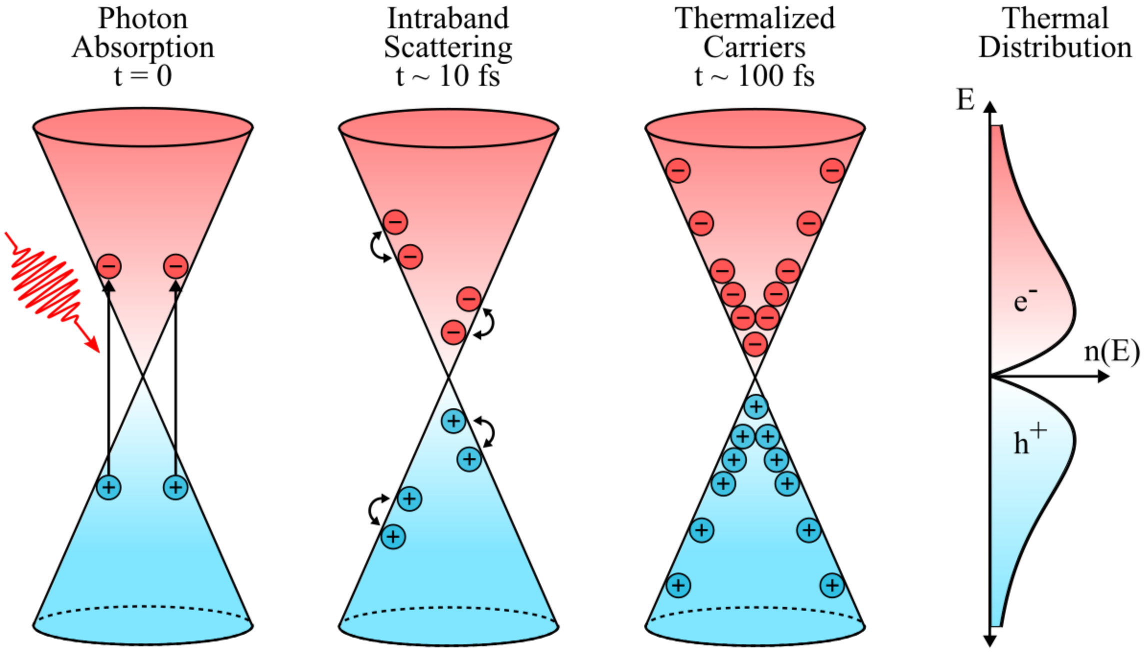


Figure 5.15: Schematic of the formation of a hot carrier state in graphene within a few hundred femtoseconds, resulting in a thermal distribution of charge carriers (right).

In this chapter, we wish to isolate the interlayer current due to thermal effects in order to study the mesoscale hot carrier state. Our measurements excite the system with photon energy of 1.03 eV (1200 nm), absorbed holes initially have an energy equal to half the photon energy, making direct tunneling from carriers initially above the barrier impossible and preventing Fowler-Nordheim tunneling except at extremely high voltages which are not used in this work. Therefore, in our interlayer photocurrent data is the result of thermal effects. In the time before the hot carrier state cools there is an exponential tail of the distribution that, if the distribution is sufficiently hot, may extend below -U_0, as shown in Fig. 5.16. Holes with energy below where some carriers may conduct into the hBN valance band and be observed as interlayer photocurrent.

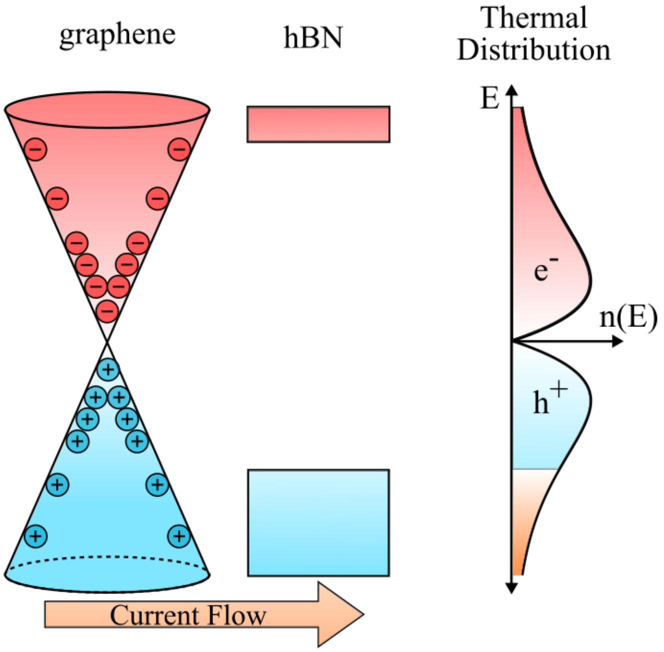


Figure 5.16: Schematic of tunneling of hot carriers due to the tail of the hot carrier distribution extending into the valance band of the hBN.

The carriers below the barrier -U_0 contribute to the interlayer current and can be calculated by integrating the carrier distribution below the barrier. Defining the distribution of holes as $1 - f_{FD}(E, \mu, T_e)$ where μ is the chemical potential and T_e is the effective electronic temperature of the carrier distribution, then the current from the i^{th} layer of graphene is proportional to:

$$I_i \propto \int_{-\infty}^{-U_0} \rho(E)[1 - f_{FD}(E, \mu_i, T_e)]dE \tag{5.1}$$

Then taking the approximation that near the barrier $1 - f_{FD}(E, \mu, T_e) \approx e^{-(\mu-E)/k_B T_e}$ and using the graphene density of states $\rho(E) = 2|E|/\pi v_F^2$ we can integrate by parts and get

$$I_i(\mu_i) \propto \int_{-\infty}^{-U_0} E e^{-\frac{\mu_i - E}{k_B T_e}} dE = k_B T_e(U_0 + k_B T_e) e^{-\frac{\mu + U_0}{k_B T_e}} \tag{5.2}$$

Both layers can contribute charge carriers and the interlayer current *I* is the difference between the two layers, $I = I_{top} - I_{bottom}$.

Once the hot carrier state forms, it cools within a few picoseconds, thus the tunneling photocurrent for a given laser power will not equal the current calculated from a single temperature. Despite this, we expect that the largest contribution will come from the highest temperature the carrier distribution reaches, which is related to the incident laser power, and should not complicate the voltage dependence. Thus, we expect qualitative agreement between the carrier tunneling model and power dependent photoconductance, with a full model of the cooling required for quantitative agreement.

Using equation 5.2 we can calculate what we expect from the tunneling photocurrent. We start with the simplest possible situation, two identical sheets of graphene that absorb the same amount and are thus heated to the same constant electronic temperature. Fig. 5.17a and 5.17b shows the simulated photocurrent and photoconductance as a function of electronic temperature and applied voltage V_g, which in the simple tunneling configuration should shift μ in each layer symmetrically. Comparing this to Fig. 5.14 we see that it qualitatively resembles the *GBNG-3* data from the highly symmetric device, and the increase in conductance at large positive and negative values of V_g for all devices.

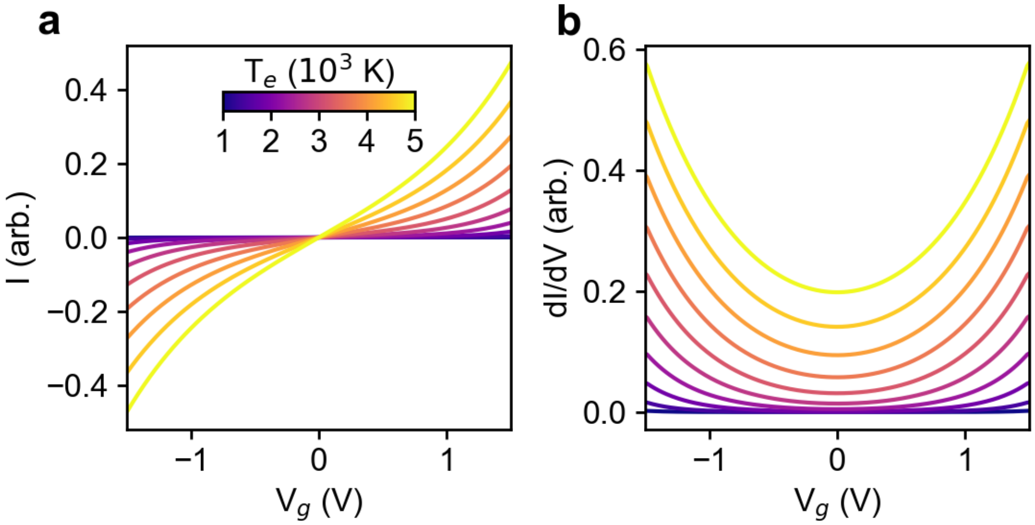


Figure 5.17: Simulated tunneling photocurrent, **a**, and photoconductance, **b**, from two symmetric graphene layers with constant electronic temperature, shows as color.

Non-uniform temperature profiles may explain the photoconductance features at low voltages. From Fig. 5.14b we see that for devices *GBNG-1* and *GBNG-2* there is a peak in the photoconductance near $V_g = 0$ that is largest for the asymmetric device *GBNG-1*. The calculation in Fig. 5.17 assumed that electronic temperature was independent of V_g, but this assumption is likely unrealistic. The hot carrier distribution is limited by the density of states of the graphene near the Fermi level. When the graphene is at or near the charge neutrality point the density of states at the Fermi level is very small and a constant number of carriers would be pushed further away from the Fermi level due to limited states. Essentially the density of states creates a bottleneck for cooling that results in a higher effective temperature. Both theoretical calculations and measurements of the temperature of hot carriers in graphene suggest that effective electronic temperature is highest when the graphene is charge neutral, though predicted temperature profiles may differ based on models of the cooling mechanisms[37,86].

We modify the tunneling model shown in Fig. 5.17 to feature a non-uniform temperature profile. Fig. 5.18 shows the same simulation with a temperature profile, shown in the Fig 5.18a inset, which increases from the base temperature T_e as a gaussian function centered at $V_g = 0$. Fig 5.18b shows the resulting photoconductance and we see a peak in the photoconductance near $V_g = 0$ and the overall profile agrees qualitatively with the conductance profiles of *GBNG-1* and *GBNG-2* as shown in Fig. 5.14b. Thus, a simple model in accordance with the hot carrier tunneling literature can qualitatively reproduce the observed photocurrent behavior in the tunneling configuration. Quantitative agreement would require a detailed dynamical model of the hot electron cooling process in each, likely different, graphene layer. Such a model is beyond the scope of this chapter and would require a different type of measurement.

The hot carrier phase is an interesting optoelectronic effect for both scientific interest and potential technological applications. The tunneling measurements discussed in the preceding sections established the existence and influence of the hot carrier phase, but are the simplest measurement possible with this system. Another measurement would be to apply an in-plane voltage to the graphene to examine the intralayer thermal effects and to measure at variable lattice temperature to explore questions about the various cooling mechanisms proposed in the literature, many of which are strongly temperature dependent. Technologically, one could imagine complex optoelectronic devices where hot carriers can be steered around the device for technological purposes or thermometry.

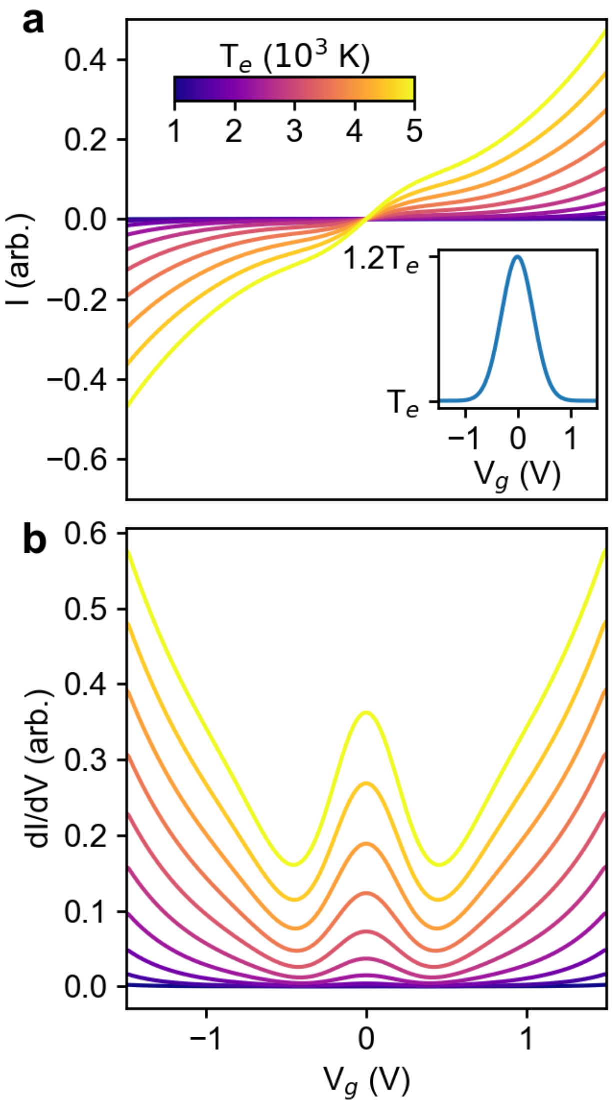


Figure 5.18: Simulated tunneling photocurrent, **a**, and photoconductance, **b**, from two symmetric graphene layers with temperature profiles as shown in the a inset.

5.9 Conclusions: Correlated Mesoscale Non-Equilibrium States

In this chapter we observed two mesoscale non-equilibrium phases, the electron-hole liquid and the hot carrier state in graphene. Notably, neither of these phenomena are exotic low-temperature effects that occur in extreme conditions; both occur at room temperature with optical excitations on the order of milliwatts. This underscores a critical point, in 2D materials these mesoscale interacting phases occur under relatively normal conditions because reduced dimensionality increases the accessibility of non-equilibrium mesoscale phases. This enables us to scientifically study these phases more easily and puts them within reach of technological applications. Furthermore, we see that data intensive imaging played an important role in studying these phases, allowing us to separate the mesoscale phases out of the normal optoelectronic behavior of these systems.

CHAPTER 6:

THE STATISTICAL SCALE

6.1 Biosystems as Model Systems in the Statistical Scale

Statistical mechanics provides the connection between the microscopic and mesoscopic dynamics we have explored in previous chapters and the macroscopic world of human existence. Therefore, to connect our understanding of physics to human existence it is important to consider how non-equilibrium states can affect time-averaged quantities, that is, when phenomena on the microscopic and macroscopic scales are viewed on the statistical scale. With electrons and holes in 2D heterostructures this is simple, the dynamics will increase or decrease observable quantities such as photocurrent. But there is more to the statistical behavior of non-equilibrium states than this, and an excellent place to understand the subtleties of this is biosystems.

Biological life is an intricate non-equilibrium state, it occurs because many tiny molecular systems conspire to manipulate energy, expel entropy, process information, assemble structures and perpetuate themselves. Given that the interruption of these mechanisms, that is a return to thermal equilibrium, means death to an organism, evolution has resulted in sophisticated and robust molecular machinery, much of which exists on the nanoscale and involves moving around excitation energy. Therefore, we look to biosystems as models and attempt to understand, not necessarily the details of these mechanisms, but rather how the overall structure of the systems allows the non-equilibrium processes to influence what happens to the organism on the statistical scale.

6.2 Photosynthesis in a Noisy Environment

In photosynthesis, light energy harvesting begins with the absorption of sunlight. Photoexcitation energy is rapidly transferred through an antenna network before reaching the reaction center, where charge transfer converts excitation energy into an electrochemical potential gradient across the photosynthetic membrane[39]. Even in the presence of dynamic light conditions, rapidly fluctuating molecular structure, and highly intricate energy transfer pathways[87-89], the light-to-electron conversion process exhibits near unity quantum efficiency. The delicate interplay of quantum effects with molecular mechanisms of energy management, e.g., non-photochemical quenching[90-92], has been explored across diverse phototrophs[93-96]. Yet, the elementary connection between highly robust light energy harvesting and energetic fluctuations is not well established.

Transforming noisy inputs into quiet outputs is a general design challenge in network architectures including multi-national energy grids, auditory and visual neural networks, and nanoscale photocells for next generation optoelectronics[97]. While network inputs exhibit statistical fluctuations (e.g., rapid changes in sunlight absorbed by a leaf or solar panel), network outputs may demand a steady rate of energy for optimal performance (e.g., constant power from the grid). Statistical fluctuations—arising from environmental variations and internal processes—fundamentally limit the throughput efficiency of any network. If the flow of energy (power) into a network is significantly larger or smaller than the flow out of the network required to optimally match the output demand, the network must adapt or be structured in such a way as to reduce the sudden over- or under-flow of energy. When the network fails to manage fluctuations, the results may be destructive, such as photo-oxidative stress in photosynthetic light harvesting or explosive damage to transformers due to fluctuations in the grid.

In this chapter, we examine the relationship between light harvesting organisms and their light environments, seeking to understand this using a simplified model. Such models aim to reduce a complex problem into a form for which calculations become more feasible. A famous example of such a model was devised by Watson and Lovelock to explain global temperature stability in the presence of biofeedback, known as the parable of Daisyworld[98]. Here, we ask whether there exists a simple network topology with *passive* noise reduction, and which requires *no external* assistance to function optimally. Remarkably, we find that such characteristics may emerge in exceedingly simple networks, and we show that carefully routed energy flow within a simple network architecture results in a robust system that inherently quiets internal noise.

6.3 The One-Channel Noisy Antenna

The premise of the model we construct is to achieve an optimal tradeoff between minimal noise in energy throughput versus robustness in a noisy environment. To understand why the two are antagonistic, consider the case of a single input node A that absorbs at wavelength λ_A with power $\mathcal{P}_A$. To minimize throughput noise the absorption rate has to match the output rate Ω, i.e. $\mathcal{P}_A = \Omega$. Such an architecture has no ability to regulate against external fluctuations. Any change in the ambient conditions alters $\mathcal{P}_A$ away from the optimal design. Thus, to gain any ability to adapt the absorber should be at a different power, i.e. $\mathcal{P}_A \neq \Omega$, which in turn introduces noise.

To construct a minimal model, we first consider the case of a single input node. For $\mathcal{P}_A > \Omega$ the absorbing channel switches on and off with probability p_A such that on average input matches output:

$$p_A \mathcal{P}_A = \Omega \tag{6.1}$$

This randomness gives rise to fluctuations. The input node is *on* some of the time injecting more power than is needed and *off* at other times leaving the network idle. The level of fluctuation is quantified by the standard deviation, which is the square root of the variance. As is conventional, throughout the following, we use the terms variance, fluctuations, and noise interchangeably. The variance, σ^2 is

$$\sigma^2 = p_A(\mathcal{P}_A - \Omega)^2 + (1 - p_A)\Omega^2 \tag{6.2}$$

Using Eq. 6.1 in we simplify the variance to

$$\frac{\sigma^2}{\Omega^2} = \frac{\mathcal{P}_A}{\Omega} - 1 \tag{6.3}$$

As anticipated, any mismatch between input and output power results in noisy throughput. Any external changes that lower $\mathcal{P}_A$ to approach the value of Ω in turn becomes beneficial (σ^2 approaches zero), while an increase cannot be regulated against.

6.4 The Two-Channel Noisy Antenna

Given the inability of the One-Channel model to regulate against any external fluctuation while operating in its lowest noise configuration we seek a model that will balance the antagonism between regulating external fluctuations and reducing internal noise while resembling the observed structure of real phototrophs. To do this we construct a model that employs generalizations of networks to extract essential aspects of photosynthetic light harvesting (Fig. 6.1a). A simple network of nodes (points at which lines intercept) connected by links (connecting lines) represent physical objects: excitation energy levels and intermolecular transfer events within the antenna system, respectively. In photosynthesis, light enters the antenna through a large number of pigment molecules, each of which is a member of a small set of distinct molecular species (e.g., chlorophyll a and b). Our model considers the advantage in having light entering the network through two classes of absorbing excitation energy levels, nodes *A* and *B*, which can absorb powers $\mathcal{P}_A$ and $\mathcal{P}_B$.

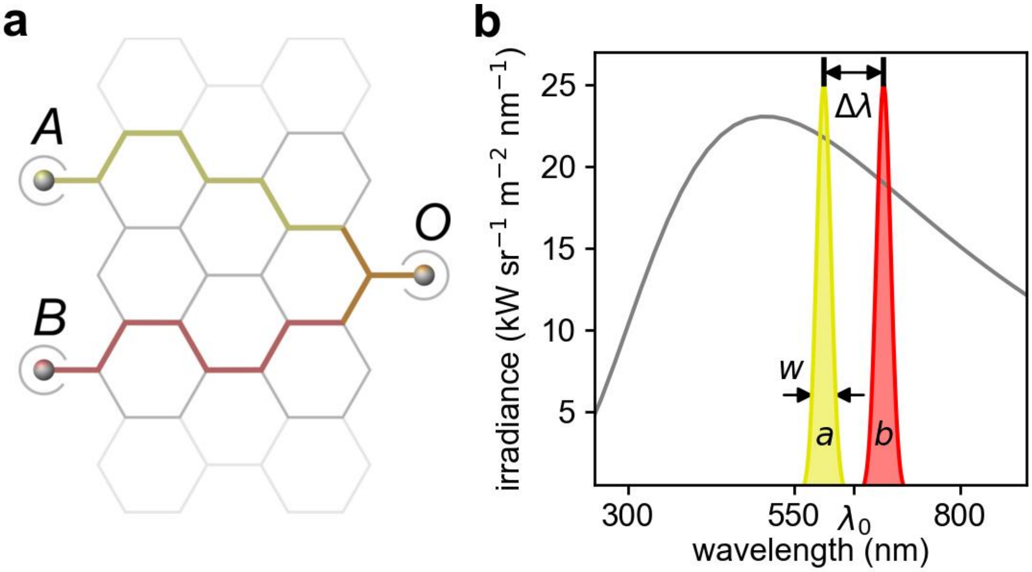


Figure 6.1: Two-channel model schematic. **a**, schematic of a photosynthetic antenna reduced into a network with two input nodes *A* and *B* with input rates $\mathcal{P}_A$ and $\mathcal{P}_B$, and output *O* with rate Ω. Energy is absorbed by molecules *a* and *b* (at rates $\mathcal{P}_A$ and $\mathcal{P}_B$) and is transferred to the output as usable energy. **b,** schematic two-channel antenna absorption spectra (yellow and red) and incident blackbody light source (grey). The quantities λ_0, $\Delta\lambda$, and w are, the center wavelength, distance between peaks, and width of the absorption peaks respectively.

Thus, a strategy based on a minimum of two input nodes straddling Ω is a natural next step. An implicit requirement is that the variations at the two input nodes be correlated. How to satisfy this condition depends sensitively on the nature of the external power spectrum from which the energy is being drawn. Below we derive the basic mathematical properties of such a model and demonstrate that a two-channel model can reduce the variance relative to a one-channel model. Then in subsequent sections we will discuss its statistical behavior in a variety of cases (section 6.5) and determine how to calculate the optimal two-channel noisy antenna for a given light spectrum (section 6.6) and then we will compare that to real biological systems.

Consider two input nodes at wavelengths λ_A and λ_B (Fig. 6.1b) with power $\mathcal{P}_A > \mathcal{P}_B$. We further set $\mathcal{P}_A > \Omega > \mathcal{P}_B$ as suggested by the argument above. At any given time only one of three possibilities occur: 1) power input from node A with probability p_A, 2) power input from node B with probability p_B, 3) no power absorbed with probability $1 - p_A - p_B$. We explicitly exclude the fourth possibility of power input from both nodes simultaneously, as such a process has input power that is much larger than the output, i.e. $\mathcal{P}_A + \mathcal{P}_B \gg \Omega$. This would add large fluctuations without any benefit in robustness due to the fact that the ambient power spectrum has a maximum, thus resulting in an upper bound on large fluctuations above Ω. The matching of input and output power gives the following relationship:

$$p_A\mathcal{P}_A + p_B\mathcal{P}_B = \Omega \,. \tag{6.4}$$

The variance is

$$\sigma^2 = p_A(\mathcal{P}_A - \Omega)^2 + p_B(\mathcal{P}_B - \Omega)^2 + (1 - p_A - p_B)\Omega^2 \,. \tag{6.5}$$

Using Eq. 6.4 and Eq. 6.5, the variance simplifies to

$$\frac{\sigma^2}{\Omega^2} = \left(\frac{\mathcal{P}_A}{\Omega} - 1\right) - p_B \frac{\mathcal{P}_B}{\Omega}\left(\frac{\mathcal{P}_A - \mathcal{P}_B}{\Omega}\right) . \tag{6.6}$$

When the second input node is absent, $p_B = 0$ and we recover the single node result. Adding a second node reduces noise for a given $\mathcal{P}_A$.

To find an optimal solution, one has to ensure that Eq. 6.4 is satisfied with the generic constraint that all probabilities must lie between 0 and 1, i.e. $0 \leq p_A \leq 1$, $0 \leq p_B \leq 1$ and $0 \leq p_A + p_B \leq 1$. A consequence of these restrictions is that the optimization process is rather nontrivial, as it is not possible to vary the probabilities and input powers independently. To make further progress we use the inequality $p_A + p_B \leq 1$. Thus

$$p_B \leq 1 - p_A \tag{6.7}$$

$$\leq 1 - \left(\frac{\Omega - p_B \mathcal{P}_B}{\mathcal{P}_A}\right) . \tag{6.8}$$

Multiplying both sides by $\mathcal{P}_A$ and solving for p_B gives

$$p_B \leq \frac{\mathcal{P}_A - \Omega}{\mathcal{P}_A - \mathcal{P}_B} . \tag{6.9}$$

Substituting the inequality in Eq. 6, we note that

$$\frac{\sigma^2}{\Omega^2} \geq \left(\frac{\mathcal{P}_A}{\Omega} - 1\right)\left(1 - \frac{\mathcal{P}_B}{\Omega}\right) . \tag{6.10}$$

Since $0 < \mathcal{P}_B < \Omega$, the *least* noise occurs when the input node exactly matches the output and generically goes up as it gets smaller. For a fixed $\mathcal{P}_A$ this is equivalent to the statement that the noise increases as the bandwidth, $\Delta = \mathcal{P}_A - \mathcal{P}_B$, increases. Eq. 10 gives a key insight: *An optimized network in an environment with correlated external* *fluctuations has Δ just large enough to quiet the noisy inputs. Increasing Δ further adds to the internal noise.*

Before further progress is made, we remark on several important conclusions. From Eq. 6.10 we can deduce that there is always a residual, which varies continuously with Δ. If the wavelengths λ_A and λ_B are far apart, the power spectrum in between need not be smooth and the fluctuations at each are in general uncorrelated. On the other hand, if λ_A and λ_B are close to each other, it is possible to create a robust network that adapts to smooth variation and correlated noise. Thus, one needs to find the region of the power spectrum that provides the largest bandwidth for adaptation (i.e. large Δ) with small $\Delta\lambda = \lambda_A - \lambda_B$. The larger the value of $\Delta = \mathcal{P}_A - \mathcal{P}_B$, the larger the available window for Ω to be within the bounds $\mathcal{P}_A > \Omega > \mathcal{P}_B$ in order to lower noise. It follows therefore that the upper limit of external fluctuations that the network can regulate against is Δ.

To understand the design parameters of the model, we must first ask: under what circumstances does the two-channel model give lower noise than the one channel model? At first look, equation 6.10 seems to imply that the second channel always suppresses the variance when compared to equation 6.3. But this comparison only holds if you compare a one channel model with input power $\mathcal{P}_A = \mathcal{P}_A^I$ to a two-channel model where the higher input power, $\mathcal{P}_A = \mathcal{P}_A^{II}$, is the same as the one-channel model, i.e. $\mathcal{P}_A^{II} = \mathcal{P}_A^I$. But this is only one possible comparison. How do we meaningfully compare the one-channel and two-channel models? It's always possible to pick a value of $\mathcal{P}_A^I$ that gives a lower variance than any given two-channel model (ignoring whether that value is stable against external fluctuation), so asking if there is a one channel model that is better than a given two-channel model is not useful.

The two-channel model defines an operable range of power $\mathcal{P}_A > \Omega > \mathcal{P}_B$ and attempts to regulate fluctuations for values of Ω within that range. Thus, to compare the one-channel and two-channel models we ask if - for a given one channel model - there is a two-channel operable range that improves the variance. More precisely: for a given one channel model defined by $\mathcal{P}_A^I$ and Ω is there a two-channel model with range $\mathcal{P}_B < \mathcal{P}_A^I \leq \mathcal{P}_A^{II}$ that has a lower variance for the same output Ω? To compare the one and two channel models we examine their respective variances in the parameter space with ordering $\mathcal{P}_B < \Omega < \mathcal{P}_A^I \leq \mathcal{P}_A^{II}$. For a quantitative measure to compare the internal noise of the models we subtract the one channel variance (equation 6.3) from the two-channel variance (equation 6.6, re-written in terms of the width of the operable range Δ), giving:

$$\Sigma = \frac{\sigma_{II}^2 - \sigma_I^2}{\Omega^2} = \left(\frac{\mathcal{P}_A^{II}}{\Omega} - 1\right) - p_B \frac{\Delta}{\Omega}\left(\frac{\mathcal{P}_A^{II} - \Delta}{\Omega}\right) - \left(\frac{\mathcal{P}_A^I}{\Omega} - 1\right) \tag{6.11}$$

When $\Sigma < 0$ the two-channel model has a lower variance. For a given one channel model, i.e. fixed $\mathcal{P}_A^I$ and Ω, Σ varies within the parameter space $(p_B, \mathcal{P}_A^{II}, \Delta)$.

Fig. 6.2a shows Σ for $(\mathcal{P}_A^{II}, \Delta)$ and arbitrarily chosen $p_B = 0.5$, $\mathcal{P}_A^I/\Omega = 1.5$. The yellow region is where $\Sigma \geq 0$ and darker colors indicate $\Sigma < 0$ where the two-channel model is advantageous. However, not all of this parameter space is allowed in the two-channel model or satisfies our ordering $\mathcal{P}_B < \Omega < \mathcal{P}_A^I < \mathcal{P}_A^{II}$. Therefore, we introduce constraints to the parameter space of Σ. Fig. 6.2b shows the meaningful parameter space, which is bounded by the constraints (shown as colored lines). The first constraints come from ordering, $\mathcal{P}_A^{II}/\Omega > \mathcal{P}_A^{I}/\Omega$ (blue line) and $\mathcal{P}_B < \Omega < \mathcal{P}_A^I$ or equivalently $\Delta/\Omega > \mathcal{P}_A^I/\Omega - 1$ (orange line). Then we have constraints that come from the two channel definition $\mathcal{P}_A^{II} > \Omega > \mathcal{P}_B$, which requires that $\mathcal{P}_A^{II}/\Omega > \Delta/\Omega$, (red line) and $\Delta/\Omega > \mathcal{P}_A^{II}/\Omega - 1$ (green line). Furthermore, equation 6.9 limits the probability p_B which translates to the constraint that $\Delta/\Omega \geq (\mathcal{P}_A^{II}/\Omega - 1)/p_B$ (magenta line).

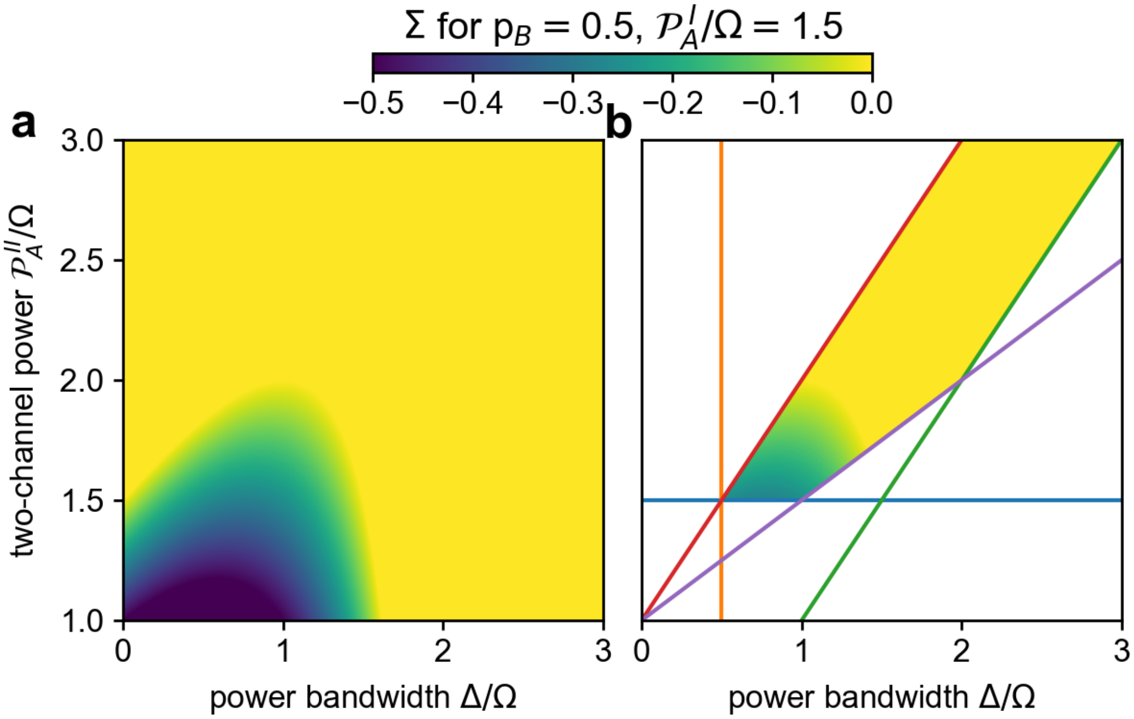


Figure 6.2: **a**, calculation of Σ in the unconstrainted parameter space $(\mathcal{P}_A^{II}, \Delta)$ for $p_B = 0.5$ and $\mathcal{P}_A^I/\Omega = 1.5$. Darker colors indicate the two-channel model has a noise advantage over the one channel model. **b**, the meaningful parameter space of Σ that satisfies all the constraints of the model and the ordering, constraints shown as colored lines.

Examining the parameter space of Σ reveals that there is a window in which the two-channel model has lower internal noise than the one channel model. Fig. 6.3 explores the full parameter space of Σ by showing $(\mathcal{P}_A^{II}, \Delta)$ for several values of p_B (constant along columns) as a function of various one channel models $\mathcal{P}_A^I/\Omega$ (constant along rows). Examining the limits of p_B we observe that for small $p_B \leq 0.1$ there is little advantage, and for $p_B \geq 0.9$ the parameter space that satisfies equation 9 is extremely narrow. This is expected given that in either limit the model barely uses one of the channels and is not very different from the one channel model. Looking at moderate range of probability, $0.3 \leq p_B \leq 0.7$, we see that for any given set of parameters there is a window of advantage in which $\Sigma < 0$ and that it generally grows larger as $\mathcal{P}_A^I/\Omega$ increases, which is to say, *as the one channel variance increases there is more room to improve.* Therefore, we conclude that if you have a noisy one channel antenna there is a two-channel antenna that will have lower internal noise for the same output. What we do not see in this parameter space is an absolute minimum in which the two-channel model is always best. Rather, for any given one channel model there is a range of possible two channel models which improve upon it.

6.5 The Discrete Noisy Antenna Model

By analyzing the stochastic flow of excitation energy, we can characterize the antenna network by statistical averages (power throughput) and fluctuations in the rate of energy flow. We take the two-channel model described in section 6.4 and generate a discrete simulation in order to understand how the model operates under different arrangement of absorbers (see appendix 2.5 for details).

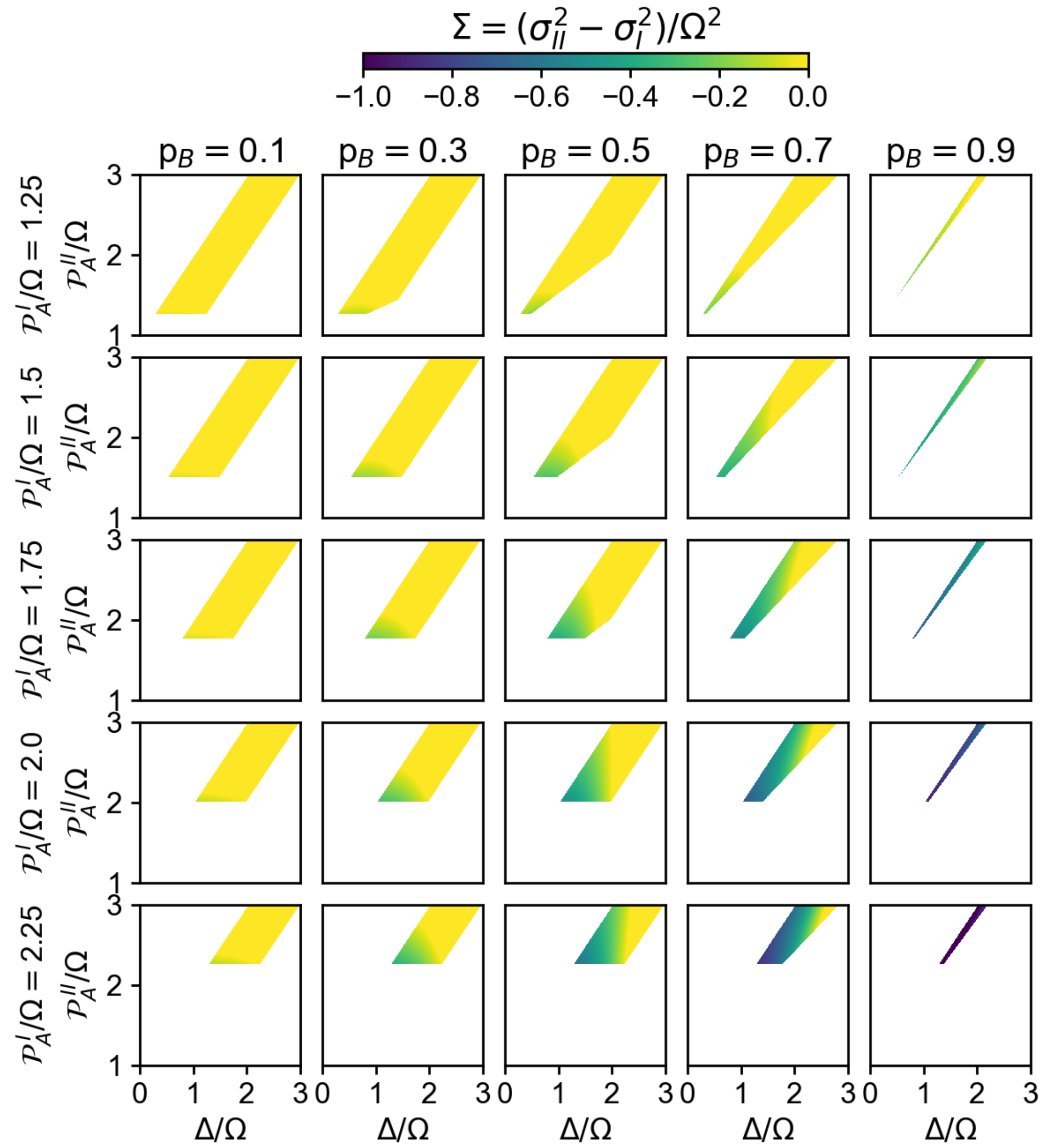


Figure 6.3: Calculation of Σ in the constrained parameter space $(\mathcal{P}_A^{II}, \Delta)$ for arbitrary values of p_B (constant across columns) and $\mathcal{P}_A^I/\Omega$ (constant across rows).

Since the absorbed solar power rarely matches exactly the rate of optimal output, a finely tuned network is that which most effectively balances minimizing the internal noise with robustness against external noise. Noise in the antenna arises from two main sources: inherent mismatch between inputs and output, which may arise due to fast dynamics in the protein structure and the corresponding electronic properties, and dynamic external light conditions. In photosynthesis, an over-powered antenna will produce excess energy that can drive deleterious back-reactions[99,100]. Conversely, a light harvesting network in an under-powered state produces non-optimal output, since the rate of energy transfer out of the network is fixed by electrochemical processes[101]. Over long periods of time, the degree to which the light harvesting network is over- or under-powered is measured by the mean-squared deviation (i.e., noise) of the total input power (through $\mathcal{P}_A$ and $\mathcal{P}_B$) from the optimal output power at Ω as shown in Fig. 6.4.

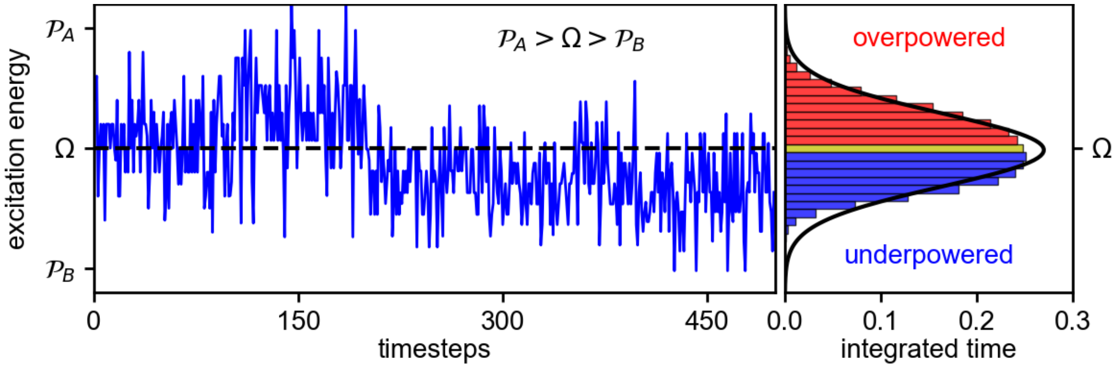


Figure 6.4: Simulated average excitation energy (**left**) as a function of time within a noisy antenna composed of 10 sets of *a* and *b* molecules. Time averaged histogram of the internal energy (**right**). The antenna is subject to internal (fast) and external (slow) fluctuations. Over long timescales the time averaged histogram resembles a normal distribution (black line).

Fig. 6.5 shows the behavior of three noise regimes within the antenna network: *over-tuned*, *fine-tuned*, and *poorly tuned*. For simplicity, Fig. 6.5 shows examples where $\Omega = (\mathcal{P}_A + \mathcal{P}_B)/2$, but the model is valid for the whole parameter space discussed in section 6.4. While the light conditions are identical for all three cases (grey lines Figs. 6.5a), we can examine how the noise changes with different absorption characteristics. When the absorbing peaks are spaced too closely (Fig. 6.5a top), the inherent antenna noise can be strongly reduced, and in the limit that $\mathcal{P}_A = \Omega = \mathcal{P}_B$ there are negligible fluctuations in the rate of energy flow (Fig. 6.5b top left). This lower bound to the internal noise cannot be reached in natural photosynthetic antennae, where protein dynamics will always drive fluctuations of intermediate excitation energy transfer events. Rather, the over-tuned antenna noise is directly proportional to, and thus dominated by, changes in the varying light spectrum (Fig. 6.5b top right). As shown in Fig. 6.5c top, in the presence of random external fluctuations, the distribution of time spent in an over- or under-powered state is flat. In the over-tuned antenna, the average input rarely matches the optimal output.

A poorly tuned antenna (Fig. 6.5a bottom) is similarly deficient. If the absorbing peaks are well separated, the antenna spends most of the time over- or under- powered. When the power sources $\mathcal{P}_A$ or $\mathcal{P}_B$ are significantly greater or less than the power sink ($\mathcal{P}_A >> \Omega >> \mathcal{P}_B$), the noise (as evidenced by a histogram of the excitation energy) in the poorly-tuned antenna becomes broader as the absorbing peaks become more separated (Fig. 6.5c bottom). When viewed over long times, the poorly tuned antenna spends too little time outputting the optimal power Ω.

The finely tuned antenna absorbs at specific positions on the spectrum that give rise to robust light harvesting even in the presence of both varying light conditions and substantial internal noise. When compared to the over- and under-tuned cases, the finely tuned antenna allows for intermediate internal noise levels (Fig. 6.5b middle) yet delivers a narrow distribution of power centered at the optimal output Ω (Fig. 6.5c middle). Robustness in light harvesting is thus the ability to output - on average - the optimal rate Ω, yet simultaneously allow for internal noise.

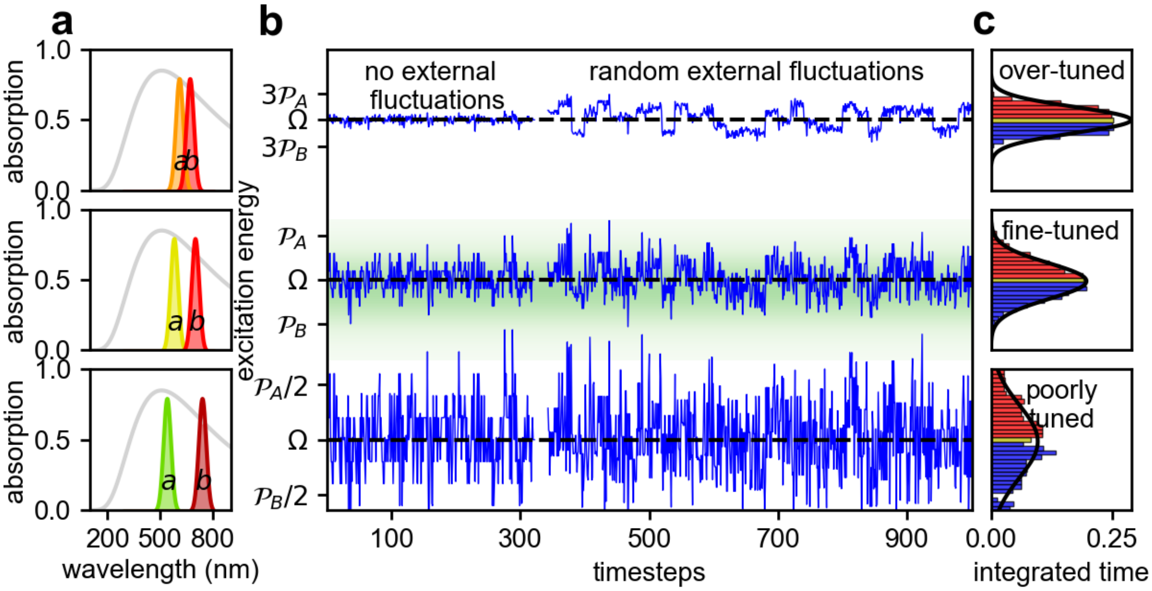


Figure 6.5: **a**, Absorption peaks for two absorbers *a* and *b* overlaid on an ideal blackbody solar spectrum (T = 5500 K, grey line) for three cases: **top**, two closely spaced absorbers; **middle**, two absorbers separated to optimize the noisy antenna; **bottom**, two widely separated absorbers. **b,** simulated excitation energy vs. time for a two-channel antenna with three different values of Δ, comparable to the cases shown in a. Left side shows the excitation energy time traces without external fluctuations. Right side includes random external fluctuations. **c,** histograms of time spent in over- (red) and under-powered (blue) states for the three series in b. **top**, the distribution is flat and favors no value. **middle**, the distribution is a sharply peaked normal distribution that favors Ω. **bottom**, the distribution is normal, but wider than in the middle panel.

6.6 Calculation of an Optimal Noisy Antenna in Diverse Light Environments

Tuning only the absorption characteristics, our goal is to find a network that spends the least amount of time in a state for which the input power is too large or too small compared to the output of the network, thus maximizing power conversion efficiency (Fig. 6.4). Within our mathematical model, probabilities p_A and p_B couple the inputs of the network $\mathcal{P}_A$ and $\mathcal{P}_B$ to the output Ω: $p_A\mathcal{P}_A + p_B\mathcal{P}_B = \Omega$. From this expression, we first evaluate the variance of the average distribution $p_A\mathcal{P}_A + p_B\mathcal{P}_B$. Minimizing this variance yields the optimal values of $\mathcal{P}_A$ and $\mathcal{P}_B$ to quiet the antenna. We then input the local spectral irradiance to a model optimization function, $\Delta^{op}(\lambda_0, \Delta\lambda, w)$, the maxima of which determine the optimal absorption characteristics for noise-cancellation (for a detailed derivation of Δ^{op} see appendix A2.6). The optimization function computes the spectral positions for which the peaks are as close as possible on the light spectrum (favoring reduced internal noise), yet the difference in the absorbed power $\Delta = \mathcal{P}_A - \mathcal{P}_B$ is maximized (supporting robustness against external variations). This condition is equivalent to maximizing the derivative of the light spectrum with respect to wavelength, thus resulting in absorption peaks in regions of steepest slope (see appendix A2.6). The absorption spectra, and thus the excitation transitions, are tuned so that the time averaged sum of input excitation energy is sharply peaked at the output rate (Fig. 6.5c middle).

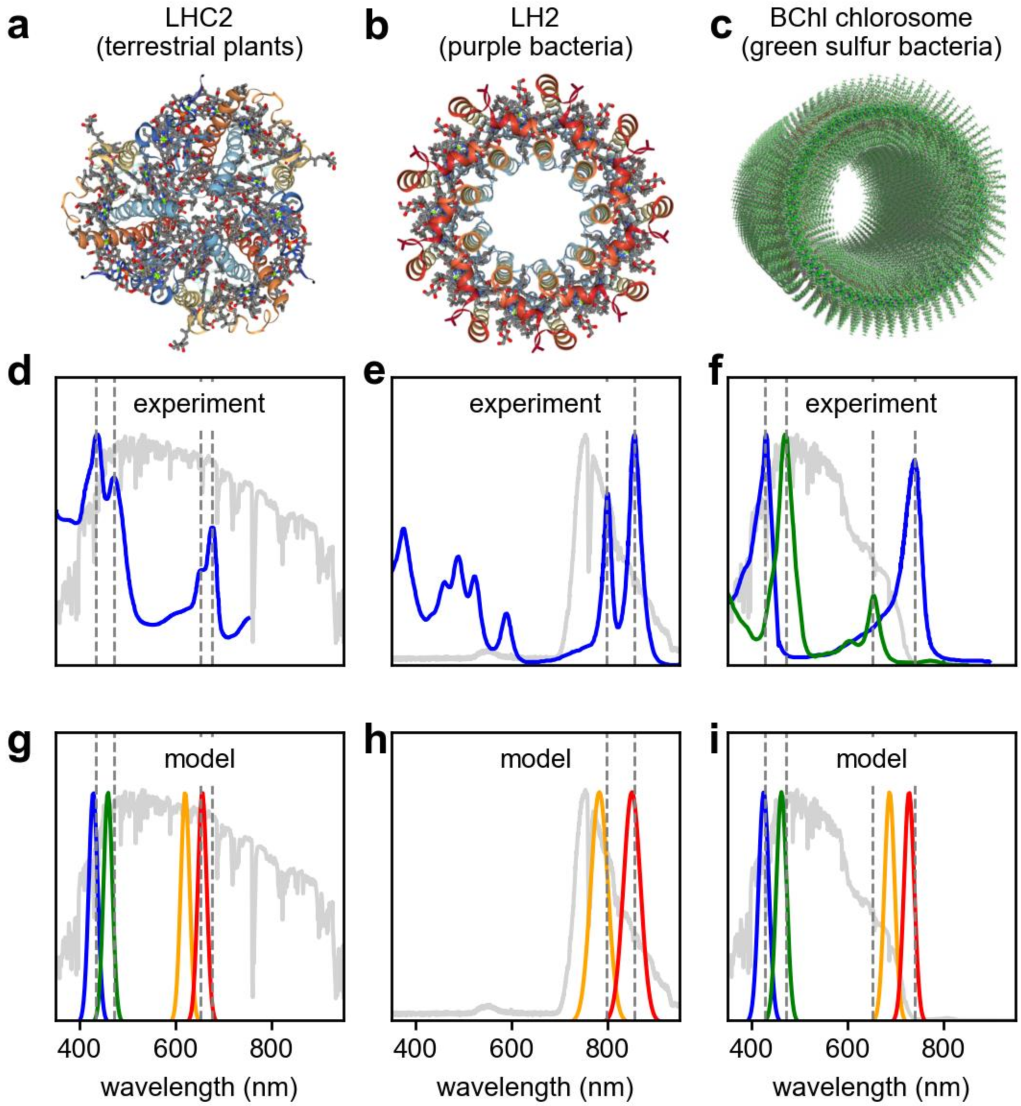


Figure 6.6: **a-c**, Structure of LHC2, LH2 and the green sulfur bacteria chlorosome **d**, absorption spectrum of LHC2 (blue) overlaid on the terrestrial solar spectrum (light grey). **e,** absorption spectrum of the LH2 complex overlaid on the solar spectrum measured below a canopy of leaves (light grey). **f,** absorption spectra of bacteriochlorophyll c (blue) and e (green) compared to the calculated solar spectrum at 2 m depth of water (light grey). **g-i,** predicted ideal absorption peaks from optimizing $\Delta = \mathcal{P}_A - \mathcal{P}_B$ for the full solar spectrum, solar spectrum attenuated through canopy, and solar spectrum attenuated through seawater. Photosynthetic absorption peaks are identified with dashed lines.

For three prototypical photosynthetic antennae, the light harvesting complex (LHC2) of green plants, the light harvesting complex (LH2) of purple bacteria, and the Bacteriochlorophyll (BChl) c and e pigments of green sulfur bacteria, the natural absorption spectrum (Figs. 6.6d,e,f)[102-105] can be compared to that predicted by our model (Figs. 6.6g,h,i) (see appendix A2.7 for full optimization details), which takes as input the local irradiance spectrum, shown as solid grey lines in Figs. 6.6d,e,f. The absorption peak positions and spectral separation predicted under light conditions in air[106], under canopy, or under seawater[107,108] (colored lines Fig. 6.6g,h,i) show striking agreement with the absorption spectra of these three phototrophs. Using only the external light spectrum and the linewidth w, the predicted peak center position λ_0 and separation $\Delta\lambda$ reproduce the measured absorption peaks with an average error of 2.1% (Table 6.1).

The noisy antenna model also reproduces a remarkable general feature of photosynthetic light harvesting observed across all three prototypical phototrophs. Photosynthetic pigments do not absorb at the maximum solar power. Instead, all three considered phototrophs exhibit pairs of closely spaced peaks in regions where the spectrum shows a steep rate of change with respect to wavelength. Photosynthetic plants look green because their antenna complexes absorb light across the visible spectrum including the blue and red portions yet reflect green wavelengths (Fig. 6.6d). Purple bacteria are aquatic phototrophs[109]. They have adapted to sunlight that is filtered through the canopy of trees and floating aerobic phototrophs (grey line Fig. 6.6e, sunlight measured through three leaves of various ash trees from on the UCR campus) and use a light harvesting complex in which bacteriochlorophyll dominates light absorption away from the visible, including green (Fig. 6.6e). Green sulfur bacteria are a geographically diverse group of bacteria that are adapted to solar light shining through seawater to depths where it is anaerobic[110]. They do not absorb the peak intensity of this attenuated light spectrum, instead absorbing in the region of steepest spectral rate of change.

Peak Name	Actual Value in nm [eV]	Calculated Value in nm [eV]	Relative % Error	Reference
Chl a 1	428 [2.90]	429 [2.89]	0.23 [0.34]	102
Chl b 1	440 [2.82]	459 [2.70]	4.32 [4.26]	102
Chl b 2	652 [1.90]	620 [2.00]	4.91 [5.26]	102
Chl a 2	660 [1.88]	656 [1.89]	0.61 [0.53]	102
LH2 band 1	801 [1.55]	783 [1.58]	2.25 [1.94]	103
LH2 band 2	857 [1.45]	851 [1.46]	0.70 [0.69]	103
BChl c 1	431 [2.88]	426 [2.91]	1.16 [1.04]	104
BChl e 1	461 [2.69]	462 [2.68]	0.22 [0.37]	105
BChl e 2	655 [1.89]	688 [1.80]	5.04 [4.76]	105
BChl c 2	740 [1.68]	728 [1.70]	1.62 [1.19]	104

Table 6.1. Absorption peak data versus model calculation for Chlorophyll (in an LHC2 complex), the LH2 complex and Bacteriochlorophyll.

6.7 Natural and Artificial Experiments Based on the Noisy Antenna Model

The qualitative and quantitative agreement that we see between the model and the spectra of real phototrophs, obtained using only the light environment, underscores the potential of this model to generate testable predictions. Underwater phototrophs provide an excellent natural experiment to test the predictive strength of our model since the solar spectrum is highly variable as a function of depth[111]. Fig. 6.7a shows the light spectrum at various depths below the seawater surface. The light intensity is attenuated as depth increases, particularly in the red and infrared, due to absorption and scattering in seawater. By comparing the absorption spectra of sub-surface marine phototrophs, such as green sulfur bacteria, to those predicted by quieting a noisy antenna, we can explore whether the natural photosynthetic absorption spectrum matches our model predictions for the relevant phototroph's preferred depth.

From the solar light spectra shown in Fig. 6.7a we calculate an optimization parameter Δ^{op} as a function of $\Delta\lambda$ and λ_0. Δ^{op} is a function modified from the calculation of $\Delta = \mathcal{P}_A - \mathcal{P}_B$ such that its maxima quiet a noisy antenna (see appendix A2.6). Fig. 6.7b shows an example color map of the magnitude of Δ^{op} at a depth of 1 m and w = 15 nm. Two maxima clearly emerge in the color plot near λ_0 = 400 and 750 nm. These maxima identify the wavelength characteristics of a finely tuned antenna under seawater. By extracting the value of $\Delta\lambda$ and λ_0 at the maximum in Δ, we obtain the characteristic absorption spectra of the quiet antenna as a function of seawater depth (Fig. 6.7c-f). We find that quieting a noisy antenna under 2 m of seawater accurately reproduces the absorption spectrum of green sulfur bacteria. Although highly adaptable, green sulfur bacteria are known to thrive at 1-2 m below the surface[112], coinciding with the conditions for which their light harvesting antenna is finely tuned for solar power conversion.

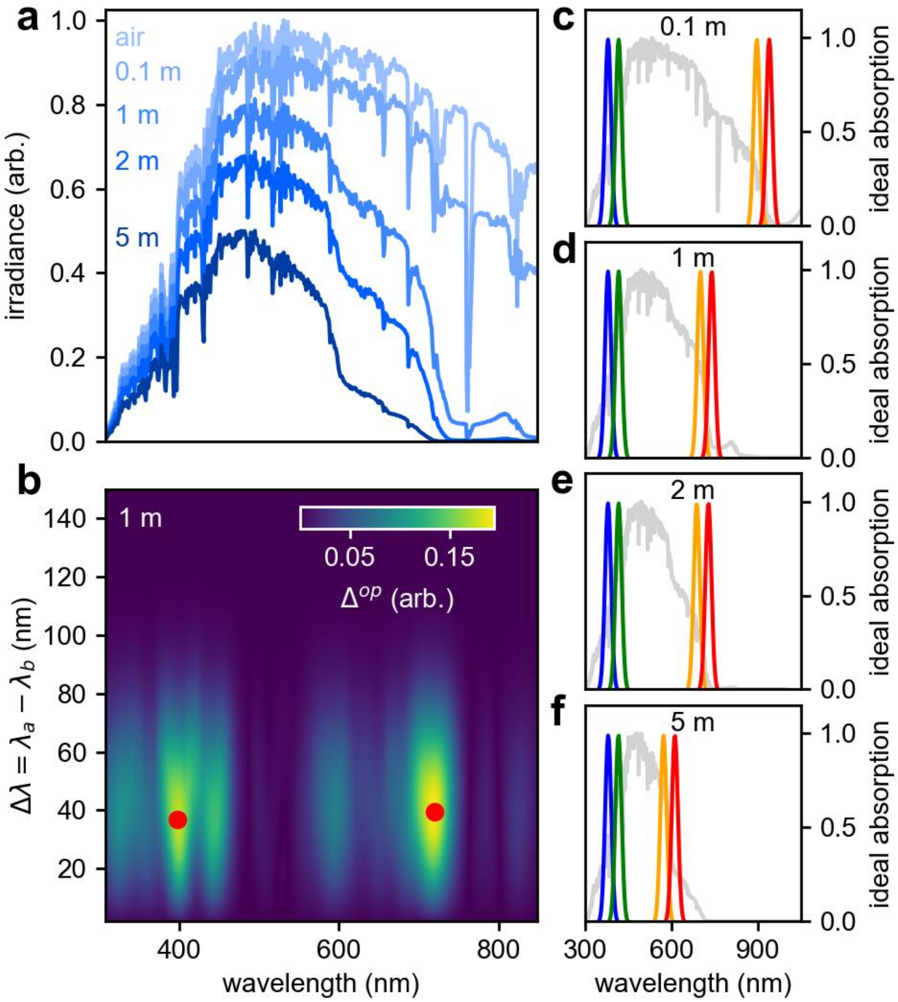


Figure 6.7: **a**, solar spectrum in air and attenuated by various depths of water (labelled). **b**, optimization landscape calculation of Δ^{op} versus center wavelength λ_0 and the peak separation $\Delta\lambda$ for solar spectrum under 1 meter of seawater (w = 15 nm). Red points identify two equally favorable maxima, corresponding to a set of peaks on either side of the spectral maximum. **c**-**f**, ideal absorption peaks predicted from the solar spectrum at each depth. Panel d shows the peaks extracted from the calculation in b, color coded blue, green, orange, red in order to track peak locations with depth.

Moreover, our findings could lead to comprehensive laboratory experiments, in which the light environment is carefully controlled while the absorption spectrum of adaptable model organisms is monitored. The simplest version of such a measurement would be to control the light environment by, for example, placing a controlled containers of Green Sulphur bacteria at variable depths and examining their growth rates, directly testing the result of Fig. 6.7. But moving beyond naturally occurring solar spectra, one could create controlled solar spectra and expose photosynthetic organisms to it, watching how they adapt. On the biological control side, mutant photosynthetic bacteria with different light harvesting spectra and exposed to a controlled spectrum they are not adapted for to observe the effect on the organisms. In any case, this model provides a jumping off point for potentially fruitful experiments.

6.8 Conclusions: Biophysics of Non-Equilibrium States

The remarkable degree to which we are able to reproduce photosynthetic absorption spectra is a surprising result, indicating an underlying organizing principle for light energy harvesting systems: Fluctuations fundamentally limit the efficiency of networks and must be avoided. Phototrophs must balance environmental inputs to sustain steady production and storage of fuel under remarkably different environmental conditions. Phototrophs across many photosynthetic niches may have adapted to build fluctuation-cancelling light harvesting antennae onto which other active mechanisms for reducing fluctuations can be added (e.g, non-photochemical quenching). Although the connection of our model to natural antenna systems requires detailed quantum models, our framework gives new insight into how extinction coefficients, delocalization lengths, and radiative rates conspire to reduce noise in natural antennae. By developing noise-cancelling antennae as a technological foundation, natural and artificial energy harvesting networks - from bacteria thriving underwater to extended power grids - could be adapted to efficiently convert noisy inputs into robust outputs.

Furthermore, this design principle in biological systems tells us the fluctuations and noise in non-equilibrium systems can be determined from their structure without difficult to acquire molecular details. This structural approach to noise suppression could apply to a wide variety of biosystems which must operate under significant noise. The fact that we are able to determine the optimal structure of a photosynthetic organism from only information about its light environment and absorber width suggests that properties of systems with highly complicated energy flow may be determined by simple dynamics.

CHAPTER 7:

CONCLUSION

7.1 Exploring the Physics of Non-Equilibrium Systems

The prototypical quantum transition, the excitation of an electron from a bound orbital to a free carrier, is a non-equilibrium system with a wide variety of phenomena on many scales. In the first chapter, we defined non-equilibrium states and identified the key physical properties that allow them to exist and to be more accessible in nanoscience. In chapter 2, we identified the challenges involved in measuring non-equilibrium states and how to overcome them with carefully considered high volume data acquisition. Then in chapter 3, we identified how to take a large set of photoresponse images and process it to extract key dynamical features which give insight into non-equilibrium phenomena. This defined an approach to measuring non-equilibrium systems that transforms one of the challenges of measuring these systems, the multi-dimensional complexity, into an asset, allowing for a comprehensive view of the parameter space requiring few assumptions.

Having established this process for handling the complexity of non-equilibrium systems we set out to explore non-equilibrium physics at multiple scales in multiple model systems. Chapter 4 begins this exploration at the microscopic level, considering interactions between individual charge carriers. This is explored in two types of TMD heterostructure. The first model system made of $MoTe_2$ with current gathered by transparent graphene gates which matched two-body Auger recombination in a non-ideal gas of excitons. Then we explored the interaction between non-equilibrium interlayer excitons and the lattice, observing extremely strong electron phonon coupling in a $MoSe_2$-WSe_2 heterostructure.

Having explored the individual interactions of photoexcited charge carriers we then explored their collective interactions. Zooming out spatially to the mesoscopic scale, in chapter 5 we explored phases where charge carriers strongly interact, both in the electron-hole liquid phase in $MoTe_2$ and the hot carrier phase in graphene-hBN heterostructures. From these we observe how the two-dimensional nature of these layered systems makes these interacting electron phases more accessible and able to occur under common laboratory conditions, such as at room temperature. Then we averaged over time to the statistical scale, and in chapter 6 observed the subtle effects that the quantum structure of an absorber can have on quieting the noise introduced by non-equilibrium states in varying environmental conditions. From this we established that the physics of non-equilibrium states can have profound effects on the biology of living organisms and that simple physics models can generate testable predictions even in very complex biological networks by considering the average statistical behavior.

7.2 Lessons Learned

Adam Savage, on the television show Mythbusters, famously quipped “the only difference between screwing around and science is writing it down.” While this quote was a joke there is some truth to this, and much of what a scientist does could reasonably be described as “screwing around” with a sample, experiment or theory in order to make it work. The difference between an amateur fiddling around, and science, is that at the end of the day a scientist attempts to learn something and improve their work in the future. More fundamentally, the entire scientific method revolves around testing ideas and then attempting to learn from those tests. Thus, it is important that at the end of a seven year long undertaking, we determine what we learned from performing the research presented in this dissertation.

The most important thing to understand is that the research presented in this dissertation bears only a passing similarity to the research as it was performed day by day. Scientific data or models are rarely presented as they were discovered; actual research is a messy process with lots of false starts, and much of the data gathered isn't explicitly useful in the end. Rather, the physics discussed here was pieced together from a variety of clues in the data, often long after the data was acquired. But one of the advantages of the data intensive approach is that it increases the amount of data that we have to draw from, even if that data does not seem very useful the moment it was taken, it may be useful down the road. Therefore, we learned that it is a good idea to invest in systematically gathering lots of data, so that you have more to work with later. A great deal of time was spent developing the data acquisition system discussed in chapter 2 and the analysis process discussed in chapter 3, but it was very much worth the time investment as it allowed us to get so much more out of the samples that we had and to consistently go back and compare our models to data, rather than wishing we had more or different data.

Furthermore, an important lesson in science is intellectual humility, particularly when planning experiments in new material systems. Sweeping exploratory searches are often derided as “fishing expeditions,” the implication being that they are higher risk and less useful than more focused measurements based on careful knowledge of the system, and more that any “accidental” discovery is not as worthwhile as a result as an experiment that confirms an existing hypothesis. Setting aside the objection that we are essentially fishing in a pond stocked with fish, i.e. that we knew there was interesting physics in these materials due to the literature and had good motivation to explore each material system, there is an unhelpful arrogance in assuming that one knows what happens in a material before we have tested it. Science is empirical, it does not matter how convincing your hypothesis seems, if it disagrees with experiment it is wrong. The human mind has a consistently terrible track record of hypothesizing about the natural world. It can take a physicist years of work to develop intuition for a physical phenomenon, and that intuition can easily be wrong. Furthermore, there is a psychological component, if one approaches an experiment believing they know the result, they can bias themselves and their conclusions about the outcome or miss findings.

Many important scientific discoveries were unexpected when they were first observed, and the research presented here is no exception. In particular, the electron-hole liquid described in chapter 5 was unexpected as such phases were thought to exist only in low temperature 2DEGs. This should not be dismissed as “accidental” with the implication that it was dumb luck. This discovery was the culmination of a lot of work to perform these types of experiments and required a great amount of effort to construct a definitive model from the data. The fact that we did not know it could happen beforehand did not reduce the amount of work required to obtain this result or its importance.

When approaching a new sample, we generally measure the system agnostic to any particular hypothesis, exploring parameter space as if we didn’t know what we know in order to avoid biasing ourselves. Afterwards we may go back and examine any interesting dependence we find in more detail or attempt to test specific hypotheses. That is to say, exploring unknown parameter space and testing a specific hypothesis are not antagonistic goals, one can look for new result as well as test expectations. The advantage of doing your measurements in a data intensive manner is that you can both look for what you expect and have the maximum chance of finding something unexpected if it exists. Science is not a straightforward path, and one should not assume that they know anything in advance even if they suspect it, rather it is advantageous to find a way to discover whatever is out there and remain flexible.

7.3 The Future

The future of research into nanoscale systems, both engineered and biophysical, is full of possibility. The subject is young and has rich systems to explore with many open questions. In this work, we have established the power of data-intensive measurements and MPDPM in exploring these systems, now it is time to implement these lessons. Being open about data intensive measurements and publishing information and code enabling data intensive experiments could shift the experimental paradigm in many related areas of research, not just nanoscale physics, but cellular and molecular biology and the characterization of newly engineered devices. The techniques discussed in chapters 2 and

3 are not tied to any particular type of 2D system and could be highly fruitful in other research contexts.

The various non-equilibrium states discussed in this dissertation also merit further investigation. The extremely strong electron-phonon coupling discussed in chapter 4 would provide a strong platform for investigations into this fundamental interaction in solid state physics. The electron-hole liquid discussed in chapter 5 raises many interesting questions: what materials can it occur in, how does the sample thickness affect its properties, what is its effective viscosity or surface tension? Lastly the results of chapter 6 on the statistical scale in biophysics were largely theoretical, with the empirical component coming from comparisons to naturally occurring phototrophs. But the structural principle described there is deep enough that it deserves to be thoroughly investigated, and controlled laboratory experiments could reveal it in far greater detail. Moving beyond photosynthesis, there are a great many biological processes that must function well in a noisy environment, the design principle of quieting a noisy antenna could be generalized to quieting noisy systems of many kinds. If the design principle is half as successful in explaining other processes as it was in predicting photosynthetic structure, then it could give enormous and fundamental insights into biological life, the ultimate non-equilibrium system.

REFERENCES

1. E. Schrödinger. *What Is Life? The Physical Aspect of the Living Cell*. (Cambridge University Press, Cambridge, 1944).

2. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, A. A. Firsov. Electric Field Effect in Atomically Thin Carbon Films. *Science* **306**, 666 (2004).

3. F. Schwierz. Graphene transistors. *Nat. Nanotechnol*. **5**, 487–496 (2010).

4. D. R. Cooper, B. D'Anjou, N. Ghattamaneni, B. Harack, M. Hilke, A. Horth, N. Majlis, M. Massicotte, L. Vandsburger, E. Whiteway, Victor Yu. Experimental Review of Graphene. *ISRN Condensed Matter Physics* **2012** 501686 (2012).

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

6. D. Pacilé, J. C. Meyer, Ç. Ö. Girit, and A. Zettl. The two-dimensional phase of boron nitride: Few-atomic-layer sheets and suspended membranes. *Appl. Phys. Lett.* **92**, 133107 (2008).

7. A. Ayari, E. Cobas, O. Ogundadegbe, and M. S. Fuhrer. Realization and electrical characterization of ultrathin crystals of layered transition-metal dichalcogenides. *J. Appl. Phys*. **101**, 014507 (2007).

8. A. K. Geim and I. V. Grigorieva. Van der Waals heterostructures. *Nature* **499**, 419 (2013).

9. Y. Liu, N. O. Weiss, X. Duan, H.-C. Cheng, Y. Huang, and X. Duan. Van der Waals heterostructures and devices. *Nat. Rev. Mater*. **1**, 16042 (2016).

10. N. M. Gabor and J. C. W. Song. Electron quantum metamaterials in van der Waals heterostructures. *Nat. Nanotechnol*. **13**, 986-993 (2018).

11. P. Rivera, H. Yu, K. L. Seyler, N. P. Wilson, W. Yao, X. Xu. Interlayer valley excitons in heterobilayers of transition metal dichalcogenides *Nat. Nanotechnol*. **13**, 1004 (2018).

12. H. Fang, et al. Strong interlayer coupling in van der Waals heterostructures built from single-layer chalcogenides *Proc. Natl. Acad. Sci. U. S. A*. **111**, 6198 (2014).

13. C. Zhang, C.-P. Chuu, X. Ren, M.-Y. Li, L.-J. Li, C. Jin, M.-Y. Chou C.-K. Shih. Interlayer couplings, Moiré patterns, and 2D electronic superlattices in MoS_2/WSe_2 hetero-bilayers. *Sci. Adv.* **3** 1 (2017).

14. J. C. W. Song, P. Samutpraphoot, L. S. Levitov. Topological Bloch bands in graphene superlattices. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 10879 (2015).

15. J. T. Ye, Y. J. Zhang, R. Akashi, M. S. Bahramy, R. Arita, Y. Iwasa. Superconducting Dome in a Gate-Tuned Band Insulator. *Science* **338**, 1193 (2012).

16. Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43 (2018).

17. B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden, W. Yao, D. Xiao, P. Jarillo-Herrero, X. Xu. Layer-dependent ferromagnetism in a van der Waals crystal down to the monolayer limit. *Nature* **546**, 270 (2017).

18. S. Jiang, J. Shan, and K. F. Mak. Electric-field switching of two-dimensional van der Waals magnets. *Nat. Mater.* **17**, 406 (2018).

19. X. Liu, K. Watanabe, T. Taniguchi, B. I. Halperin, and P. Kim. Quantum Hall drag of exciton condensate in graphene. *Nat. Phys.* **13**, 746 (2017).

20. J. I. A. Li, T. Taniguchi, K. Watanabe, J. Hone, and C. R. Dean. Excitonic superfluid phase in double bilayer graphene. *Nat. Phys.* **13**, 751 (2017).

21. J. Frenkel. On the Transformation of light into Heat in Solids. *Phys. Rev.* **37**, 17 (1931).

22. A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, T. F. Heinz. *Phys. Rev. Lett.* **113**, 076802 (2014).

23. T. Mueller, E. Malic. Exciton physics and device application of two-dimensional transition metal dichalcogenide semiconductors. *npj 2D Mater. Appl.* **2**, 29 (2018).

24. [Wang2018] G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, B. Urbaszek. Colloquium: Excitons in atomically thin transition metal dichalcogenides. *Rev. Mod. Phys.* **90**, 021001 (2018).

25. C. Jin, E. Y. Ma, O. Karni, E. C. Regan, F. Wang, T. F. Heinz. Ultrafast dynamics in van der Waals heterostructures. *Nat. Nanotechnol.* **13**, 994 (2018).

26. M. Kuiri, B. Chakraborty, A. Paul, S. Das, A. K. Sood, A. Das. Enhancing photoresponsivity using $MoTe_2$-graphene vertical heterostructures. *Applied Physics Letters* **108**, 063506 (2016).

27. M. Massicotte, P. Schmidt, F. Vialla, K. G. Schdler, A. Reserbat-Plantey, K. Watanabe, T. Taniguchi, K. J. Tielrooij, F. H. L. Koppens. Picosecond photoresponse in Van der Waals heterostructures. *Nature Nanotechnology* **11**, 42-46 (2016).

28. W. J. Yu, Q. A. Yu, H. Oh, H. G. Nam, H. Zhou, S. Cha, J. Y. Kim, A. Carvalho, M. Jeong, H. Choi, A. H. C. Neto, Y. H. Lee, X. Duan. Unusually efficient photocurrent extraction in monolayer van der Waals heterostructure by tunneling through discretized barriers. *Nature Communications* **7**, 13278 (2016).

29. M. Sze, K. K. Ng. *Physics of Semiconductor Devices*. (Wiley-Interscience publications, United States, 1963).

30. P. Landsberg. *Recombination in Semiconductors.* (Cambridge University Press, Cambridge, 1991).

31. D. Chepic, A. L. Efros, A. I. Ekimov, M. G. Ivanov, V. A. Kharchenko, I. A. Kudriavtsev, T. V. Yazeva. Auger ionization of semiconductor quantum drops in a glass matrix. *J. Lumin.*, **47**, 113 (1990).

32. C. D. Jeffries. Electron-hole condensation in semiconductors. *Science* **189**, 955-964 (1975).

33. L. V. Keldysh. The electron-hole liquid in semiconductors. *Contemp. Phys.* **27**, 395-428 (1986).

34. A. Tomadin, D. Brida, G. Cerullo, A. C. Ferrari, M. Polini. Nonequilibrium dynamics of photoexcited electrons in graphene: Collinear scattering, Auger processes, and the impact of screening. *Phys. Rev. B* **88**, 035430 (2013).

35. J. C. Johannsen et al. Direct View of Hot Carrier Dynamics in Graphene. *Phys. Rev. Lett.* **111**, 027403 (2013).

36. Gierz, I., Petersen, J., Mitrano, M. et al. Snapshots of non-equilibrium Dirac carrier distributions in graphene. *Nature Mater* **12,** 1119–1124 (2013).

37. A. Tomadin et al. The ultrafast dynamics and conductivity of photoexcited graphene at different Fermi energies. *Science Advances* **4** 5 (2018).

38. C. Ferrante, A. Virga, L. Benfatto, M. Martinati, D. De Fazio, U. Sassi, C. Fasolato, A. K. Ott, P. Postorino, D. Yoon, G. Cerullo, F. Mauri, A. C. Ferrari, T. Scopigno. Raman spectroscopy of graphene under ultrafast laser excitation. *Nat. Commun.* **9**, 308 (2018).

39. R. Blankenship. Molecular Mechanisms of Photosynthesis (John Wiley & Sons, 2002).

40. M. Jeng. A selected history of expectation bias in physics. *American Journal of Physics* **75**, 578 (2006).

41. A. Franklin. Forging, cooking, trimming, and riding on the bandwagon. *American Journal of Physics* **52**, 786 (1984).

42. A. Franklin. Selectivity and discord: two problems of experiment (University of Pittsburgh Press, Pittsburgh, PA, 2002).

43. F. Mulargia. Retrospective selection bias (or the benefit of hindsight). *Geophys. J. Int.* **146**, 489–496, (2001).

44. G. Steinmeyer. A review of ultrafast optics and optoelectronics. *J. Opt. A: Pure Appl. Opt.* **5**, R1 (2003).

45. M. Fushitani. Applications of pump-probe spectroscopy. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **104**, 272 (2008).

46. J. C. W. Song, L. S. Levitov. Shockley-Ramo theorem and long-range photocurrent response in gapless materials. *Phys. Rev. B* **90**, 075415 (2014).

47. D. Mayes, M. Grossnickle, M. Lohmann, M. Aldosary, J. Li, V. Aji, J. Shi, J. C.W. Song, N. M. Gabor. Current flux imaging of a micromagnetic electrofoil. arXiv:2002.07902 (2020).

48. E. R. Tufte. The Visual Display of Quantitative Information. (Graphics Press, Cheshire, Connecticut, 1983)

49. J. D. Hunter. Matplotlib: A 2D Graphics Environment. *IEEE Computing in Science & Engineering* **9**, 3 (2007).

50. P. K. Robertson. A methodology for scientific data visualization: choosing representations based on a natural scene paradigm. *Proceedings of the First IEEE Conference on Visualization: Visualization `90*, San Francisco, CA, USA, 114-123 (1990).

51. C. Ware. Color Sequences for Univariate Maps: Theory, Experiments, and Principles. *IEEE Comput. Graphics Appl.* **8**, 41 (1988).

52. K. Moreland. Diverging Color Maps for Scientific Visualization. *Proceedings of the 5th International Symposium on Advances in Visual Computing: Part II, ISVC'09* (Springer-Verlag, Berlin, Heidelberg, 2009), pp. 92–103.

53. M. Sze, K. K. Ng. *Physics of Semiconductor Devices*. (Wiley-Interscience publications, United States, 1963).

54. P. Landsberg. *Recombination in Semiconductors.* (Cambridge University Press, Cambridge, 1991).

55. D. Chepic, A. L. Efros, A. I. Ekimov, M. G. Ivanov, V. A. Kharchenko, I. A. Kudriavtsev, T. V. Yazeva. Auger ionization of semiconductor quantum drops in a glass matrix. *J. Lumin.*, **47**, 113 (1990).

56. F. Thouin, S. Neutzner, D. Cortecchia, V. A. Dragomir, C. Soci, T. Salim, Y. M. Lam, R. Leonelli, A. Petrozza, A. R. S. Kandada, C. Silva. Stable biexcitons in two-dimensional metal-halide perovskites with strong dynamic lattice disorder. *Phys. Rev. Mat.* **2**, 034001 (2018).

57. Y. You, X. X. Zhang, T. C. Berkelbach, M. S. Hybertsen, D. R. Reichman, T. F. Heinz. Observation of biexcitons in monolayer WSe_2. *Nature Physics* **11**, 477-482 (2015).

58. K. Hao, J. F. Specht, P. Nagler, L. Xu, K. Tran, A. Singh, C. K. Dass, C. Schuller, T. Korn, M. Richter, A. Knorr, X. Li, G. Moody. Neutral and charged inter-valley biexcitons in monolayer $MoSe_2$. *Nature Communications* **8**, 15552 (2017).

59. M. L. W. Thewalt, J. A. Rostworowski, Biexcitons in Si. *Solid State Communications* **25**, 991-993 (1978).

60. E. Lorchat, L. E. Parra López, C. Robert, D. Lagarde, G. Froehlicher, T. Taniguchi, K. Watanabe, X. Marie, S. Berciaud. Filtering the photoluminescence spectra of atomically thin semiconductors with graphene. *Nature Nanotechnology* **15**, 283–288 (2020).

61. C. Ruppert, O. B. Aslan, T. F. Heinz. Optical properties and band gap of single and few-layer MoTe2 crystals. *Nano Letters* **14**, 6231-6236 (2014).

62. F. Barati, M. Grossnickle, S. Su, R. K. Lake, V. Aji, N. M. Gabor. Hot carrier-enhanced interlayer electron–hole pair multiplication in 2D semiconductor heterostructure photocells. *Nature Nanotechnology* **12**, 1134–1139 (2017).

63. P. Rivera, J. Schaibley, A. Jones, et al. Observation of long-lived interlayer excitons in monolayer $MoSe_2$–WSe_2 heterostructures. *Nat Commun* **6**, 6242 (2015).

64. A. F. Rigosi, H. M. Hill, Y. Li, A. Chernikov, T. F. Heinz. Probing Interlayer Interactions in Transition Metal Dichalcogenide Heterostructures by Optical Spectroscopy: MoS_2/WS_2 and $MoSe_2/WSe_2$. *Nano Lett.* **15**, 5033-5038 (2015).

65. C. Lee, G. Lee, A. van der Zande, et al. Atomically thin p–n junctions with van der Waals heterointerfaces. *Nature Nanotech* **9,** 676–681 (2014).

66. W. Shockley. The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors. *The Bell System Technical Journal*. **28**, 435-489 (1949).

67. G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B* **54,** 11169 (1996).

68. G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals. *Physical Review B* **47**, 558 (1993).

69. Blöchl, P. E. Projector augmented-wave method. *Physical Review B* **50**, 17953 (1994).

70. A. Togo, I. Tanaka, First principles phonon calculations in materials science. *Scripta Mater*. **108**, 1 (2015).

71. Y. V. Nazarov, Y. M. Blanter. *Quantum Transport: Introduction to Nanoscience.* (Cambridge University Press, Cambridge, 2009).

72. F. H. M. Faisal, J. Z. Kamiński. Floquet-Bloch theory of high-harmonic generation in periodic structures. *Phys. Rev. A* **56**, 748 (1997).

73. L. V. Keldysh. *Proceedings of the 9th International Conference on Physics of Semiconductors*, (Nauka Leningrad, 1968) 1303.

74. A. Rustagi, A. Kemper. Theoretical Phase Diagram for the room-temperature electron-hole liquid in photoexcited quasi-two-dimensional monolayer MoS_2. *Nano Letters* **18**, 455-459 (2018).

75. Y. Sun, J. Zhang, Z. Ma, C. Chen, J. Han, F. Chen, X. Luo, Y. Sun, Z. Sheng. The Zeeman splitting of bulk 2H-$MoTe_2$ single crystal in high magnetic field. *Appl. Phys. Lett.* **110**, 102102 (2017).

76. M. M. Ugeda, A. J. Bradley, S. Shi, F. H. da Jornada, Y. Zhang, D. Y. Qiu, W. Ruan, S. Mo, Z. Hussain, Z. Shen, F. Wang, S. G. Louie, M. F. Crommie, Giant

bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor. *Nature Materials* **13**, 1091-1095 (2014).

77. A. Chernikov, C. Ruppert, H. M. Hill, A. F. Rigosi, T. F. Heinz, Population inversion and giant bandgap renormalization in atomically thin WS_2 layers. *Nature Photonics* **9**, 466-470 (2015).

78. A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, T. F. Heinz. Exciton binding energy and nonhydrogenic rydberg series in monolayer WS_2. *Phys. Rev. Lett.* **113**, 076802 (2014).

79. A. Arora, R. Schmidt, R. Schneider, M. R. Molas, I. Breslavetz, M. Potemski, R. Bratschitsch. Valley Zeeman Splitting and Valley Polarization of Neutral and Charged Excitons in Monolayer $MoTe_2$ at High Magnetic Fields. *Nano Lett.* **16** 3624–3629 (2016).

80. M. W. Graham, S. Shi, Z. Wang, D. C. Ralph, J. Park, P. L. McEuen. Transient Absorption and Photocurrent Microscopy Show That Hot Electron Supercollisions Describe the Rate-Limiting Relaxation Step in Graphene. *Nano Lett.* **13**, 11, 5497–5502 (2013).

81. M. Graham, S. Shi, D. Ralph, et al. Photocurrent measurements of supercollision cooling in graphene. *Nature Phys.* **9**, 103–108 (2013).

82. Q. Ma, N. M. Gabor, T. I. Andersen, N. L. Nair, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero. Competing Channels for Hot-Electron Cooling in Graphene. *Phys. Rev. Lett.* **112**, 247401 (2014).

83. A. Betz, S. Jhang, E. Pallecchi, et al. Supercollision cooling in undoped graphene. *Nature Phys* **9**, 109–112 (2013).

84. Q. Ma, T. I. Andersen, N. L. Nair, N. M. Gabor, M. Massicotte, C. H. Lui, A. F. Young, W. Fang, K. Watanabe, T. Taniguchi, J. Kong, N. Gedik, F. H. L. Koppens and P. Jarillo-Herrero. Tuning ultrafast electron thermalization pathways in a van der Waals heterostructure. *Nature Physics* **12**, 455–459 (2016).

85. N. Kharche, S. K. Nayak. Quasiparticle Band Gap Engineering of Graphene and Graphone on Hexagonal Boron Nitride Substrate. *Nano Lett.* **11**, 12, 5274–5278 (2011).

86. J. F. Rodriguez-Nieva, Mildred S. Dresselhaus, Leonid S. Levitov. Thermionic Emission and Negative dI/dV in Photoactive Graphene Heterostructures. *Nano Lett.* **15**, 3, 1451–1456 (2015).

87. R. Van Grondelle, J. P. Dekker, T. Gillbro, V. Sundstrom. Energy transfer and trapping in photosynthesis. *Biochimica et Biophysica Acta (BBA) – Bioenergetics* **1187**, 1-65 (1994).

88. G. D. Scholes, G. R. Fleming, A. Olaya-Castro, R. Van Grondelle. Lessons from nature about solar light harvesting. *Nature Chemistry* **3**, 763-774 (2011).

89. R. Croce, H. Van Amerongen, Natural strategies for photosynthetic light harvesting. *Nature Chemical Biology* **10**, 492-501 (2014).

90. C. Kulheim, J. Agren, S. Jansson, Rapid Regulation of Light Harvesting and Plant Fitness in the Field. *Science* **297**, 91-93 (2002).

91. A. A. Pascal, Z. Liu, K. Broess, B. van Oort, H. van Amerongen, C. Wang, P. Horton, B. Robert, W. Chang, A. Ruban, Molecular basis of photoprotection and control of photosynthetic light-harvesting. *Nature* **436**, 134–137 (2005).

92. A. V. Ruban, R. Berera, C. Ilioaia, I. H. M. van Stokkum, J. T. M. Kennis, A. A. Pascal, H. van Amerongen, B. Robert, P. Horton, Rienk van Grondelle, Identification of a mechanism of photoprotective energy dissipation in higher plants. *Nature* **450**, 575–578 (2007).

93. R. G. Fleming, R. Van Grondelle, The Primary Steps of Photosynthesis. *Physics Today* **47**, 48-55 (1994).

94. E. Romero, V. I. Novoderezhkin, R. Van Grondelle, Quantum design of photosynthesis for bio-inspired solar-energy conversion. *Nature* **543**, 355-365 (2017).

95. T. Mirkovic, et al., Light Absorbtion and Energy Transfer in the Antenna Complexes of Photosynthetic Organisms. *Chemical Reviews* **117**, 249-293 (2017).

96. G. Scholes, et al., Using coherence to enhance function in chemical and biophysical systems. *Nature* **543**, 647-656 (2017).

97. T. B. Arp, Y. Barlas, V. Aji N. M. Gabor. Natural Regulation of Energy Flow in a Green Quantum Photocell. *Nano Lett.* **16**, 7461-7466 (2016).

98. A. J. Watson, J. E. Lovelock, Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus B: Chemical and Physical Meteorology* **35**, 284-289 (1983).

99. P. Horton, A. V. Ruban, R. G. Walters, Regulation of light harvesting in green plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 655-684 (1996).

100. A. V. Ruban, M. P. Johnson, C. D. P. Duffy, The photoprotective molecular switch in the photosystem II antenna. *Biochemica et Biophysica Acta (BBA) - Bioenergetics* **1817**, 167-181 (2012).

101. G. D. Farquhar, S. Von Caemmerer, J. A. Berry, A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Plants* **149**, 78-90 (1980).

102. T.P.J. Kruger, R. van Grondelle, Design principles of natural light-harvesting as revealed by single molecule spectroscopy. *Physica B: Condensed Matter* **480**, 7-13 (2016).

103. R. Cogdell, N. Isaacs, et. al., The structural basis of light-harvesting in purple bacteria. *FEBS Letters* **555**, 1, 35-39 (2003).

104. D. C. Brune, T. Nozawa, R. E. Blankenship, Antenna Organization in Green Photosynthetic Bacteria. 1. Oligomeric Bacteriochlorophyll c as a Model for the 740nm Absorbing Bacteriochlorophyll c in Chloroflexus aurantiacus Chlorosomes. *Biochemistry* **1987**, 26, 8644-8652 (1987).

105. C. M. Borrego, J. B. Arellano, C. A. Abella, T. Gillbro, J. Garcia-Gil, The molar extinction coefficient of bacteriochlorophyll e and the pigment stoichiometry in Chlorobium phaeobacteroides. *Photosynthesis Research* **60**, 257-264 (1999).

106. Reference Air Mass 1.5 Spectra | Grid Modernization | NREL [Online]. 2003. https://www.nrel.gov/grid/solar-resource/spectra-am1.5.html. Accessed June 2018.

107. H. Buiteveld, J. H. M. Hakvoort, M. Donze, Optical properties of pure water. *Proceedings of SPIE* 2258 Ocean Optics XII **1994**, 174 – 183 (1994).

108. L. Kou, D. Labrie, P. Chylek, Refractive indices of water and ice in the 0.65 to 2.5-µm spectral range. *Appl. Opt.* **32**, 3531-3540 (1993).

109. C. N. Hunter, F. Daldal, M. C. Thurnauer, J. T. Beatty, *The Purple Phototropic Bacteria.* (Springer, 2009).

110. J. M. Olsen, Ed., *In Green Photosynthetic Bacteria* (Springer, 1988), pp. 315 – 319.

111. P. Jenkins, et al., High-Bandgap Solar Cells for Underwater Photovoltaic Applications. *IEEE Journal of Photovoltaics* **4**, 202-207 (2014).

112. A. Kharcheva, A. Zhiltsova, Studies of Green Sulphur Bacteria Using Spectral Methods: Comparison of Flourescence and Absorption Characteristics in Winter and Summer. *WDS'16 Proceedings of Contributed Papers – Physics* **2016**, 214-218 (2016).

113. F. Barati. Optoelectronics Investigations of Electron Dynamics in 2D-TMD Semiconductor Heterostructure Photocells: From Electron-Hole Pair Multiplication to Phonon Assisted Anti-Stokes Absorption. Thesis, University of California, Riverside, CA (2018).

114. D. L. Pleskot. Strongly Interacting Electrons and Holes in Ultrathin MoTe2 Heterostructures. Thesis, University of California, Riverside, CA (2019).

115. A. Castellanos-Gomez, M. Buscema, R. Molenaar, V. Singh, L. Janssen, H. S J van der Zant, G. A. Steele. Deterministic transfer of two-dimensional materials by all-dry viscoelastic stamping. *2D Mater. Lett*. **1**, 1-8 (2014).

APPENDIX 1:

SAMPLE INFORMATION

A1.1 Sample Information and Metadata

This appendix contains information and metadata on the various samples used throughout this dissertation. Each sample is given an identifier that abbreviates the names of its active materials that is italicized, for example a sample that is composed of graphene and WSe_2 would be called *GWSe2* and is referred to as such throughout the dissertation. This appendix will contain basic information about the composition of the samples, their characterization, and any special information about them. The samples will be listed in the order they are introduced in the text.

All samples were prepared using an exfoliation and a dry transfer stacking process by various members of the QMO Lab. The fabrication process has been developed over time allowing increasingly clean and sophisticated samples to be fabricated, a detailed description of the process can be found in the dissertations of Dennis Pleskot and Fatemeh Barati[113,114]. The component materials were exfoliated from high quality bulk crystals onto Si/SiO_2 substrates. The heterostructure was assembled in a custom-built transfer microscope using a well know dry transfer technique developed by Gomez et al.[115] Polydimethylsiloxane (PDMS)/polypropylene carbonate (PPC) stamps were used to pick up and controllably deposits the exfoliated flakes on top of each other. Titanium-gold (Ti/Au) electrical contacts were fabricated onto the devices using electron beam lithography. The person who fabricated a given sample and any variations from the standard process will be noted below.

A1.2 Graphene-hBN-Graphite Heterostructure Sample

The *GBNGr* sample discussed in Chapter 3 was fabricated by Dennis Pleskot and is composed of graphene on stacked on hexagonal boron nitride stacked on graphite, as shown in Figure A1.2.1. Electrical contacts were fabricated onto the graphene (on the top) and graphite (on the bottom), notably this sample contained electronical contacts on top of the heterostructure region, resulting in a more complex spatial photocurrent response. While this complex contact response provided a good example of the MPDPM data analysis discussed in chapter 3, in most samples we avoid contacts on the heterostructure to not contaminate signals with contact photocurrent.

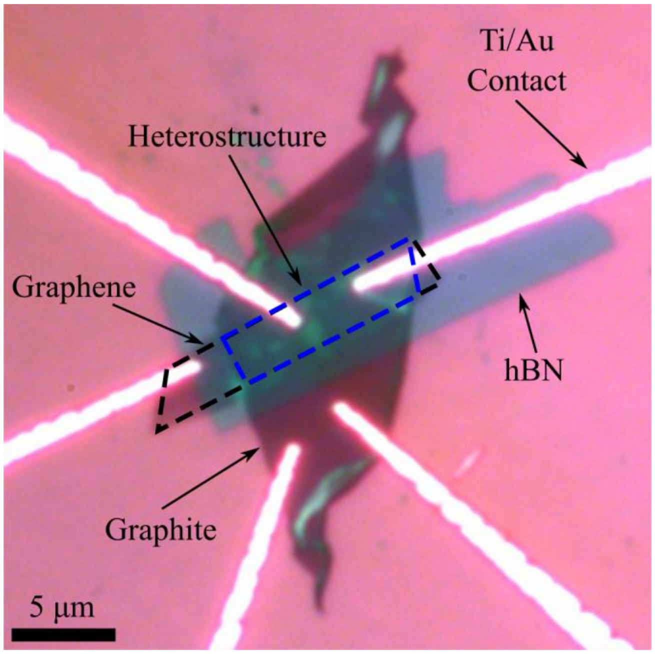


Figure A1.2.1: Optical Image of the *GBNGr* sample with component materials shown. Dashed black line outlines the graphene and dashed blue line outlines the heterostructure.

A1.3 Graphene-$MoTe_2$ Heterostructure Sample

The *GMoTe2* sample discussed in chapters 4 and 5 was fabricated by Dennis Pleskot out of graphene and exfoliated $MoTe_2$ flakes using the standard dry transfer method. The resulting heterostructure is shown in Figure A1.3.1. AFM measurements were performed using a Digital Instruments Nanoscope IV with a silicon cantilever in tapping mode to confirm the layer thicknesses and observed that the top graphene, $MoTe_2$, and bottom graphene layers display thicknesses of 5.88 nm, 8.96 nm, and 13.8 nm, respectively. Raman spectroscopy was used to confirm the identity and approximate thickness of the materials.

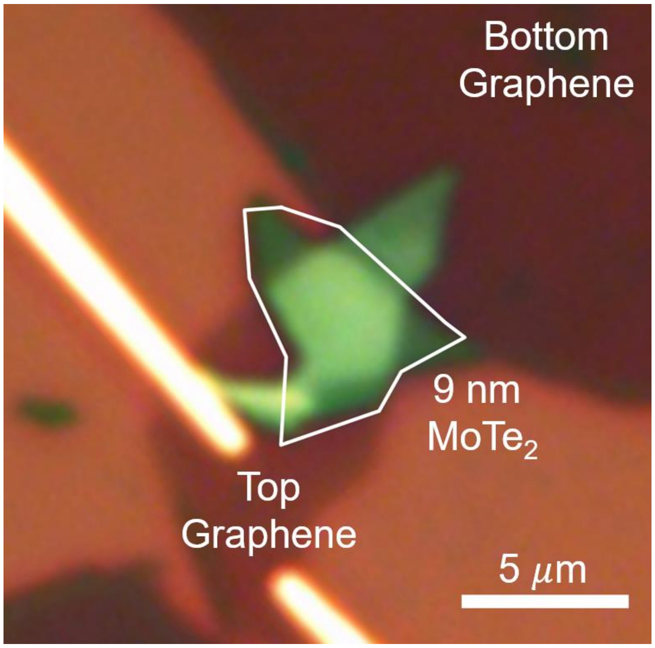


Figure A1.3.1: Optical Image of the *GMoTe2* sample. Outlined area is the approximate extent of the heterostructure.

A1.4 $MoSe_2$ - WSe_2 Heterostructure Samples

In chapter 4 we discuss two samples composed of stacked monolayer $MoSe_2$ and bilayer WSe_2, which were fabricated by Fatemeh Barati using mechanical exfoliation of WSe_2 and $MoSe_2$ flakes from bulk crystal onto prefabricated contacts. These contacts were fabricated on a Si wafer coated with 290nm-thick SiO_2. The heterostructure devices are assembled using a highly customized, temperature-controlled transfer microscope following the dry transfer process with special consideration to minimize the contamination of the interface. The dry pick-up transfer process was performed such that the two layers has no intentional contact to polymer films to preserve interface quality and was followed by two annealing processes to remove any contaminates.

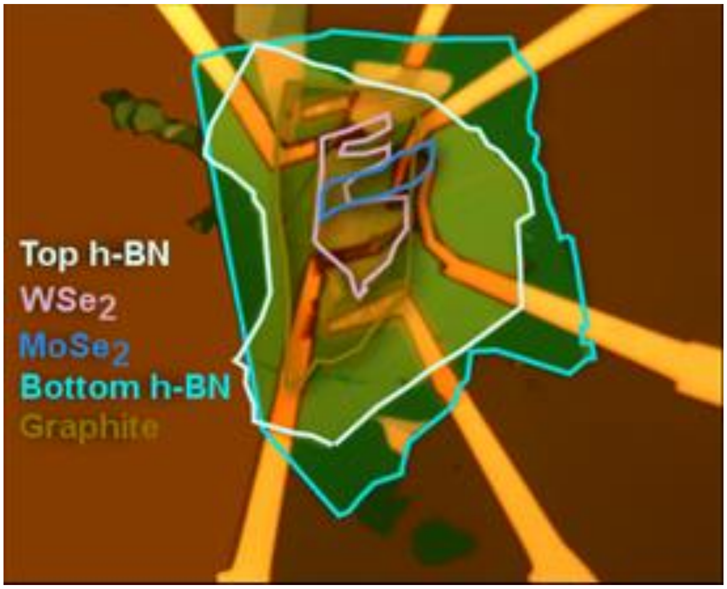


Figure A1.4.1: Optical image of the *MoSe2WSe2-2* sample with material outlines.

Two devices were fabricated and used, *MoSe2WSe2-1* was the first iteration was used for low temperature measurements, it uses an Si backgate instead of a graphene backgate but is otherwise consistent. The second sample *MoSe2WSe2-2*, show in in Figure A1.4.1, was the second iteration and featured a graphene backgate. It was used for all the room temperature measurements.

Devices were characterized using Raman, photoluminescence (PL), and photocurrent (PC) spectroscopy. As shown in Figure A1.4.2a, the Raman spectrum of the heterostructure (black line) exhibits peaks that are also evident in the $MoSe_2$ (green line) and WSe_2 (blue line). When comparing Raman peaks in the heterostructure region to those in the constituent layers, we observed negligible shift of the peak positions as a function of energy, consistent with previous work[62]. Importantly, among the various vibrational resonances, the most prevalent features in the Raman spectrum of the heterostructure occur at energies of 29, 31, and 32 meV.

Photoluminescence measurements reveal signatures of bound interlayer electron-hole (*e*-*h*) pairs, or excitons, at the WSe_2-$MoSe_2$ interface. Figure A1.4.2b compares the PL as a function of photon energy E_{PH} from $MoSe_2$ (green line), WSe_2 (blue line), and the heterostructure region (black line). The heterostructure exhibits similar spectral features to the individual layers, yet with a strong suppression of the lowest energy peak of WSe_2, consistent with ordinary behavior due to charge redistribution at the interface.

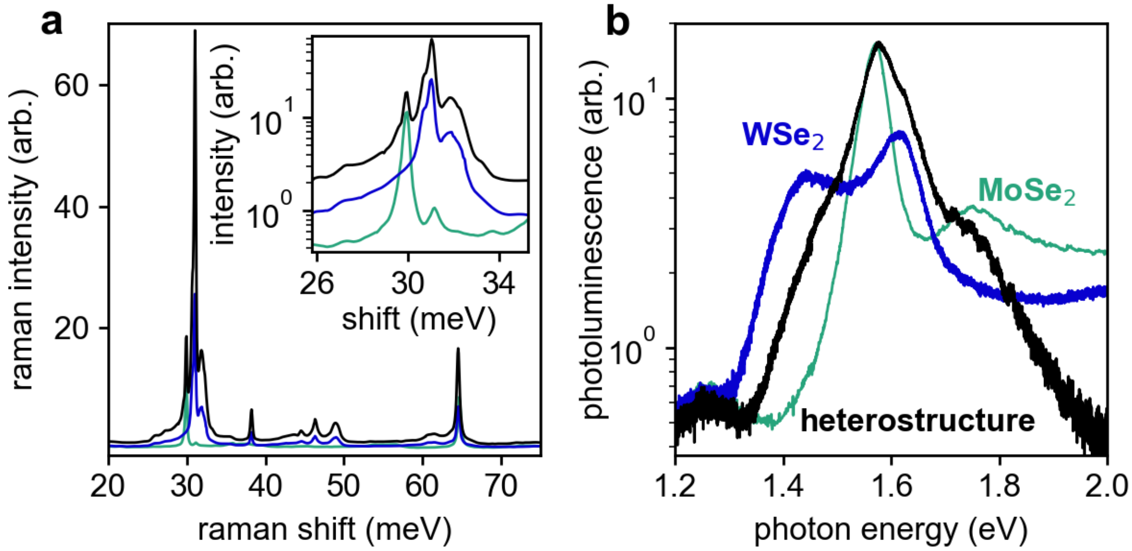


Figure A1.4.2: Optical characterization of the *MoSe2WSe2-2* sample. **a**, Raman and, **b**, photoluminescence spectra are consistent across both *MoSe2WSe2-1*.

A1.5 Graphene-hBN-Graphene Samples

In chapter 5 we discuss the behavior of three graphene-hBN-graphene (*GBNG*) heterostructures which were fabricated by Jacky Wan. The heterostructures are designated *GBNG-1*, *GBNG-2*, and *GBNG-3* and images are shown in Figure A1.5.1. The heterostructures were assembled through the normal dry transfer method with the variation that *GBNG-2* and *GBNG-3* were transferred onto prefabricated gold contacts allowing them to be stacked in an inverted order with the contacts underneath the materials and better coverage of the hBN. The prefabricated contacts allowed the graphene layers to be fully encapsulated inside hBN in the heterostructure region.

The order of the samples shows the consistent improvement in the fabrication and quality of the devices, which can be observed in the electronic characteristics shown in Figure 5.11. The first sample *GBNG-1* had the bottom graphene resting on the SiO_2

substrate which significantly degrades its quality compared to hBN encapsulated graphene. The second device, *GBNG-2*, has both graphene layers encapsulated with hBN but is geometrically imperfect, with part of the bottom graphene electrostatically screened by the Si substrate. The final device, *GBNG-3* has fully encapsulated graphene in an electrostatically perfect geometry, resulting in a GBNG pixel with a clean spatial photoresponse. The hBN layers range from 7 to 10 nm, except for the bottom layers in *GBNG-2* and *GBNG-3* which were much thicker, ~50 nm. In all cases steps were taken to minimize the wet chemistry the graphene was exposed to and keep the layers as close to intrinsic as possible.

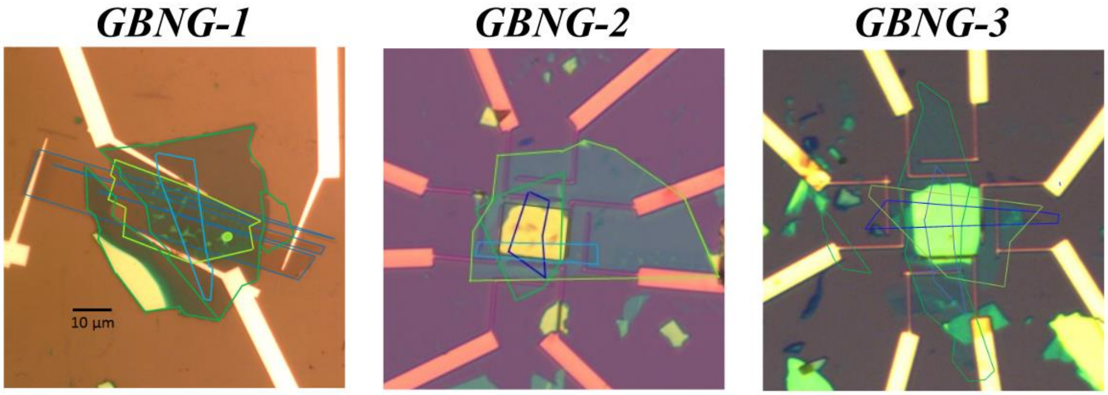


Figure A1.5.1: Optical images of the *GBNG* samples. Outlined areas identify materials, blue outlines are for graphene, green outlines are for hBN, and dark to light outlines are top to bottom of the heterostructure stack.

APPENDIX 2:

CALCULATIONS

A2.1 Estimate of Laser Absorption in $MoTe_2$

To consider the density of charge carriers generated by laser excitation in the *GMoTe2* sample we must first estimate what percentage of incident light is absorbed by the sample and converted into electron hole pairs. For a thick sample, on-resonance absorption is extremely high, approaching unity, however there are two complications. Firstly, ultrafast pulsed lasers are spectrally broad, meaning not all of the laser is on the resonance peak. Secondly, our laser is peaked at 1.03 eV (1200 nm) and the A exciton is peaked at 1.08 eV. The A exciton peak itself has a spectral width of approximately 50 meV[61]. Therefore, to estimate our laser's absorption we consider a gaussian absorption peak with 100% absorption at 1.08 eV and a 50 meV full width at half maximum (FWHM). Comparing the FWHM of our laser (~20 meV) to this distribution we see that our laser's spectrum has between 3% absorption and 24% absorption but is peaked at 9% absorption. From this we conclude that on the order of 10% of the laser is absorbed.

A2.2 Steady State Solution to the Exciton Rate Equation Model

For a continuous wave (CW) laser the excitation is constant and the current settles into a steady state. Though we do not use a CW laser in this work, this limit is still useful to consider. As discussed in Chapter 4, excitons generated by a laser undergo Auger recombination, which we model with Equation 4.1. In the steady state case, $dN/dt \rightarrow 0$, and $G(t)$ is a constant, G_0, that is proportional to the laser power. Thus Eqn. 4.1 is:

$$G_0 = \frac{N}{\tau_{es}} + \alpha N^2 \qquad (A2.2.1)$$

Which can be solved as $N = \left[\sqrt{1/\tau_{es}^2 + 4\alpha G_0} - 1/\tau_{es}\right]/2\alpha$. In general, the population of charge carriers that becomes photocurrent is small compared to the population of excitons, meaning that the linear term in Eqn. 4.1 can be neglected, i.e. $1/\tau_{es} \ll \alpha$. Assuming that, $N \approx \sqrt{G_0/\alpha}$. If we equate N with the steady state photocurrent and G_0 with laser power, then we see that the laser power should follow $I \sim P^{1/2}$.

Our data tends towards the steady state form $I \propto P^{1/2}$ at long Δt and high interlayer voltage. We can see that our model satisfies this limiting case by expanding I_I from Eqn. 4.3 to second order, in which case $I_I(N_1) \approx N_1 - \frac{1}{2}\alpha\tau_{es}N_1^2$, assuming $N_1\alpha\tau_{es} < 1$. We can also expand the steady state current around a non-zero power P_0, $I_{steady} \propto \sqrt{P} = \sqrt{(P - P_0) + P_0} \approx \sqrt{P_0}\left[\frac{3}{8} + \frac{3}{4}\left(\frac{P}{P_0}\right) - \frac{1}{8}\left(\frac{P}{P_0}\right)^2\right]$. If we equate P/P_0 with N_1 we see that to second order both the model and the steady state limit have the same basic structure of $I \sim P - P^2$, meaning that $\gamma \sim \frac{1}{2}$ is a generally good approximation in the single pulse limit.

A2.3 Mean Distance Between Electron-Hole Pairs at P_C

To calculate the average distance between the charge carriers at the critical point of the phase transition, we must find the density at the critical power P_C. The number of photons incident on the sample by a single pulse of the laser is given by $N = E_{pulse}(P)/\,E_{photon}(\lambda) = \lambda P/hcR$. Where R = 75 MHz is the repetition rate of the laser,

λ = 1200 nm is the laser wavelength, and P is the power of the laser beam. The density of electrons and holes generated by a single pulse is $n = 2AN/V$. Where A is the percentage of light absorbed by the sample (see section A2.1), and V is the total volume containing the electron-hole population. The average distance between charge carriers, d_{avg}, is then given by $\frac{4}{3}\pi d_{avg}^3 = \frac{1}{n}$ which gives the average spacing between electron-hole pairs

$$d_{avg} = \sqrt[3]{\frac{3}{8\pi}\frac{V}{AN}} \approx 1.4 \times 10^{-4}\sqrt[3]{\frac{V}{AP}} \qquad (A2.3.1)$$

To estimate d_{avg}, we must estimate the volume V in which the electron-hole population undergoes condensation. This volume depends on how fast the electron hole liquid forms, thus we can only place bounds on it. For the lower bound, the phase transition occurs quickly relative to the timescale of diffusion and the volume is a cylinder with diameter equal to the FWHM of the beamspot and height equal to the sample thickness. Therefore $V_{min} = \pi(1.67\ \mu m)^2(9\text{nm}) = 0.08\ \mu\text{m}^3$ which at the critical power of $P_c = 5$ mW gives an average distance of $d_{avg} = 0.35\ \text{nm}/\sqrt[3]{A}$. For the upper bound, carriers fully diffuse into the total volume of the sample prior to condensation. The area of the sample can be estimated from Figure 4.2 which gives an area of 66 μm^2, therefore $V_{max} = 5.94\ \mu\text{m}^3$ which gives an average distance of $d_{avg} = 1.48\ \text{nm}\ /\sqrt[3]{A}$. Assuming 10% absorption in $MoTe_2$ the lower bound is 0.8 nm and the upper bound is 3.2 nm. Thus, we conclude that at the critical point, charge carriers are separated by an average of 1-3 nm. This value is similar to values of the exciton Bohr radius determined through magneto-optical measurements of $MoTe_2$[75].

A2.4 Modeling the spatially resolved photoresponse of an electron hole liquid

To calculate the spatial distribution of the photocurrent we need to model what happens when a laser pulse illuminates the sample. At the surface of the sample the laser is a diffraction limited beam spot with a spatial profile that can be approximated as $P(x) = P_0 \exp(-x^2/2\sigma^2)$. Below the critical threshold, P_C, the photoresponse obeys a power law $I \propto P^\gamma$. Thus for $P_0 \leq P_{th}$ the observed photocurrent is given by taking the convolution of the Gaussian beam and the photocell profile (working in one dimension for simplicity but this can easily be generalized to two dimensions):

$$I(x) = \int_{-\infty}^{\infty} [P(x - x')]^\gamma f(x')dx' \ = \int_{-\infty}^{\infty} P_0^\gamma e^{-\frac{\gamma(x-x')^2}{2\sigma^2}} f(x')dx' \qquad (A2.4.1)$$

where $f(x)$ is a function describing the profile of the sample, $f(x) = 1$ on the sample and zero otherwise.

To calculate the spatial photoresponse in the electron hole liquid regime we consider the case where, for $P_0 > P_{th}$ an electron hole liquid droplet will form near the center of the beamspot. Once the droplet forms it can absorb nearby charge carriers and the *e*-*h* droplet size should increase linearly as the laser power is increased, thus the size of the droplet l is given by $l = L_0(P_0 - P_{th})$. Where L_0 is a free parameter that tunes the rate of expansion of the droplet. Since charge carriers inside the droplet recombine reapidly, the part of the beam spot that the droplet is under does not contribute to the observed photocurrent, thus the current is modeled by:

$$I(x) = \int_{-\infty}^{-l} P_0^{\gamma} e^{-\frac{\gamma(x-x')^2}{2\sigma^2}} f(x')dx' + \int_{l}^{\infty} P_0^{\gamma} e^{-\frac{\gamma(x-x')^2}{2\sigma^2}} f(x')dx' \qquad \text{(A2.4.2)}$$

The observed photocurrent is the convolution of the photocurrent due to gas-phase electron-hole pairs and the photocell active area profile; the volume of the liquid droplet does not contribute to the observed photocurrent, resulting in photocurrent suppression.

A2.5 Discrete Toy Model of Two-Channel Noisy Antenna

In chapter 6, we discussed tuning of the noisy antenna in terms of time spent over- and under-powered. To visualize the noise behavior under different choices of input parameters, we calculate the power throughput within a finite system of absorbers as the sum of the absorption events within discrete timesteps. By implementing our model using random trials in the discrete limit, detailed below, we can illustrate how the correct choice of Δ reduces the sensitivity to external noise while incurring the minimum increase in internal noise. The results of this calculation, shown in Fig. 6.4 and 6.5, provide intuitive visualization for a key statement of our analytical model: An optimized network in an environment with correlated external fluctuations has Δ just large enough to quiet noisy inputs. Increasing Δ further adds to the internal noise.

To explore the analytical model of Section 6.4 within a computationally discrete case, we consider a light harvesting network consisting of a small number of absorbing molecules falling into two classes - a and b - that undergo discrete absorption events. This illustrative computation considers a group of 10 absorbing pairs, which was chosen as an order of magnitude estimate within any physically relevant light harvesting network. Our model corresponds to a simplistic network of absorbers with a direct coupling to the output set by a single rate for each of the a or b type absorbers, similar to that shown schematically in Figure 6.1. Since the absorbers are simply connected within the network, the number of absorbing pairs sets the noise level.

At each timestep, the molecules will absorb either $\mathcal{P}_A$, with probability p_a, absorb $\mathcal{P}_B$ with probability p_b, or absorb nothing with probability $1 - p_a - p_b$. The probabilities are set by the equilibrium condition, $p_A \mathcal{P}_A + p_B \mathcal{P}_B = \Omega$ under the symmetric condition $\Omega = (\mathcal{P}_A + \mathcal{P}_B)/2$ for simplicity. There is a range of possible values of p_a and p_b that obey the equilibrium condition. To set the values we define a free parameter $0 \leq \phi \leq 1$ such that $p_B = p_B^{min} \phi + p_B^{max}(1 - \phi)$ where p_B^{min} and p_B^{max} are the minimum and maximum values of p_B; from there p_A is set by the equilibrium condition. For the calculations shown in Fig. 6.4 and 6.5, $\phi = 0.05$ was used, but all values of ϕ give fundamentally similar results.

The inputs to this calculation are the values of $\mathcal{P}_A$, $\mathcal{P}_B$, and Ω. To simulate external fluctuations in the light environment we add a slowly varying random fluctuation on top of $\mathcal{P}_A$ and $\mathcal{P}_B$, i.e. $\mathcal{P}_A \rightarrow \mathcal{P}_A + \mathcal{P}_A \delta P$ and $\mathcal{P}_B \rightarrow \mathcal{P}_B + \mathcal{P}_B \delta P$ for some random fluctuation δP that changes every 20 timesteps. The output of the calculation is the sum of the energy absorbed from all the absorbing pairs, which gives the excitation energy of the system for that timestep. Due to the equilibrium condition, we expect that the average output should be 10Ω, which we see when $\delta P = 0$ (in the figures, 10Ω is re-nomalized to Ω for simplicity). This calculation simulates excitation energy as a function of time, exhibiting fast stochastic noise and noise due the random external fluctuations. The histogram of this timeseries is what we show in Fig. 6.5c to illustrate the tuning of the model as a function of Δ.

A2.6 Optimization Parameter for the Two-Channel Noisy Antenna

In chapter 6, we discussed tuning of the noisy antenna and the key conclusion is that a two-channel model with a finite window in Δ is better at suppressing internal noise as compared to a one channel model. On the other hand, to protect against external variability, i.e. the fluctuation in the power spectrum incident on the antenna, Δ needs to be as large as possible. Our next step is to determine how these two contradictory properties are best satisfied and develop a quantitative way to find the optimal parameters for a Noisy Antenna in a given light environment. We start with a careful examination of the properties of the power bandwidth Δ.

To determine Δ, and the wavelengths of the input nodes for optimization, a more realistic model for absorption is needed. Unlike the ideal model described in section 6.4, where all the absorption happens at two fixed wavelengths, absorbers generically operate in a narrow window centered at the λ_A and λ_B as shown in Fig. 6.1b. The absorption of the two channels as a function of wavelength, $a(\lambda)$ and $b(\lambda)$, are parameterized by gaussian functions:

$$a(\lambda, \lambda_0, \Delta\lambda, w) = \frac{1}{w\sqrt{2\pi}} \exp\left\{-\frac{[\lambda - (\lambda_0 + \Delta\lambda/2)]^2}{2w^2}\right\} \tag{A2.8.1}$$

$$b(\lambda, \lambda_0, \Delta\lambda, w) = \frac{1}{w\sqrt{2\pi}} \exp\left\{-\frac{[\lambda - (\lambda_0 - \Delta\lambda/2)]^2}{2w^2}\right\} \tag{A2.8.2}$$

where λ_0 is the center wavelength between the two absorbing peaks, $\Delta\lambda$ is the separation between the peaks and w is the width of the peak functions (i.e. the standard deviation). Note that $\lambda_A = \lambda_0 + \Delta\lambda/2$ and $\lambda_B = \lambda_0 - \Delta\lambda/2$. For narrow absorbers, i.e. $w \ll \Delta\lambda$, the precise parametrization - such as Gaussian, Lorentzian, or equivalent - does not change the principal conclusions. The input powers, $\mathcal{P}_A$ and $\mathcal{P}_B$, are obtained in the usual way by integrating the product of the absorption with the irradiance of the solar spectrum, $I(\lambda)$. In other words, $\mathcal{P}_A = \int a(\lambda)I(\lambda)d\lambda$ and $\mathcal{P}_B = \int b(\lambda)I(\lambda)d\lambda$. From these simple definitions, the difference in the absorbed power is given by

$$\Delta(\lambda_0, \Delta\lambda, w) = \int [a(\lambda, \lambda_0, \Delta\lambda, w) - b(\lambda, \lambda_0, \Delta\lambda, w)]I(\lambda)d\lambda\,. \tag{A2.8.3}$$

Eq. A2.8.3 gives a key parameter as Δ sets the scale of correlated external fluctuations that the noisy antenna is robust against. This suggests a strategy for how to calculate the ideal absorbers for a given solar spectrum.

To insulate from external fluctuations, the first order optimization is a search for the largest value of Δ in the parameter space $(\lambda_0, \Delta\lambda, w)$. However, this will not quiet a Noisy Antenna as there is a spurious maximum that must be considered, yet discounted. In the case of large $\Delta\lambda$, peak *a* could sit on the maximum of $I(\lambda)$ and peak *b* could be on the far edge of $I(\lambda)$. In this scenario $\mathcal{P}_A$ is maximized, and $\mathcal{P}_B$ is effectively zero, thus Δ is the maximum possible. But this case is clearly outside the window in Δ where the two-channel model is able to reduce internal noise, in fact, this is the most extreme version of the *poorly tuned* case illustrated in the bottom panels of Fig. 6.5.

Fortunately, the apparent, but spurious maximum is automatically accounted for when realistic architecture of the network is implemented. The integrated *power* that is absorbed at a given wavelength is transferred to the output node with a finite transition probability, which in turn is proportional to the energy at which the absorption occurs. The energy corresponding to each absorber is $E_{A,B} = hc/\lambda_{A,B}$, where h is Planck's constant and c is the speed of light. If one channel has a significantly larger energy it will be preferred over the other, resulting in the *poorly tuned* case similar to the single input node scenario.

Put together our design considerations for the optimization of the Noisy Antenna are as follows: (1) The two-channel model is advantageous only for a finite range of $\Delta = \mathcal{P}_A^{II} - \mathcal{P}_B^{II}$ within a parameter space defined by $\mathcal{P}_A^{II}$ and p_B. (2) The larger the range Δ, the better the system can protect against external fluctuations. (3)The typical wavelength of absorption of the two channels should not be too different. The strategy we adopt to implement these considerations is to determine the operable range for which the absorbers are close in energy ($\Delta\lambda/\lambda_0 \ll 1$) and then determine an optimization function that gives the maximum possible Δ within this constrained subspace of parameters.

We can estimate the operable range through careful analysis of Eq. A2.8.3 with the line shapes specified in Eq. A2.8.1 and A2.8.2:

$$\Delta(\lambda_0, \Delta\lambda, w) = \int [a(\lambda, \lambda_0, \Delta\lambda, w) - b(\lambda, \lambda_0, \Delta\lambda, w)] I(\lambda) d\lambda \qquad \text{(A2.8.4)}$$

$$= \int \frac{1}{w\sqrt{2\pi}}\left[\exp\left\{-\frac{[\lambda-(\lambda_0+\Delta\lambda/2)]^2}{2w^2}\right\}-\exp\left\{-\frac{[\lambda-(\lambda_0-\Delta\lambda/2)]^2}{2w^2}\right\}\right]I(\lambda)d\lambda \quad \text{(A2.8.5)}$$

$$= \frac{2}{w\sqrt{2\pi}}\exp\left\{-\frac{\Delta\lambda^2}{8w^2}\right\}\int \exp\{-\frac{(\lambda-\lambda_0)^2}{2w^2}\}\sinh\left\{\frac{\Delta\lambda(\lambda_0-\lambda)}{2w^2}\right\}I(\lambda)d\lambda. \quad \text{(A2.8.6)}$$

To make further progress, we invoke empirical, but generic, facts of the spectral irradiance function to evaluate Eq. A2.8.6. Specifically, we first recognize that the spectrum $I(\lambda)$ is bounded both in magnitude, with a single maximum, and is limited to a finite window in wavelength. Combining this with the $\exp\{-(\lambda-\lambda_0)^2/2w^2\}$ factor in the integrand of Eq. A2.8.6, we conclude that $\Delta(\lambda_0, \Delta\lambda, w)$ is determined by the behavior of the integral near λ_0. Importantly, expanding Eq. A2.8.6 in the vicinity of λ_0

$$\Delta = \frac{2}{w\sqrt{2\pi}}\exp\left\{-\frac{\Delta\lambda^2}{8w^2}\right\}\int \mathrm{d}\lambda \exp\left\{-\frac{(\lambda-\lambda_0)^2}{2w^2}\right\}\sinh\left\{\frac{\Delta\lambda(\lambda_0-\lambda)}{2w^2}\right\}$$
$$\times\left[I(\lambda_0)+(\lambda-\lambda_0)\left.\frac{dI}{d\lambda}\right|_{\lambda_0}+\frac{1}{2}(\lambda-\lambda_0)^2\left.\frac{d^2I}{d\lambda^2}\right|_{\lambda_0}+\cdots\right] \quad \text{(A2.8.7)}$$

we see that all even derivatives in the expansion vanish since sinh is an odd function in $\lambda-\lambda_0$. The leading contribution to the integral comes from the first derivative of $I(\lambda)$. Therefore λ_0 is in the vicinity of the inflection points of $I(\lambda)$. This yields the intuitive result that maximizing the term puts λ_0 in the vicinity of the inflection points of $I(\lambda)$. The natural scale for $\Delta\lambda$ is $2\sqrt{2}w$ appearing in the leading exponential multiplying the integral. While an integration over all λ is not very meaningful, given that the expansion is only valid near λ_0, doing so yields

$$\Delta = \Delta\lambda\sum_{n=0}^{\infty}(2w^2)^n L_n^{\frac{1}{2}}\left(-\frac{\Delta\lambda^2}{8w^2}\right)\left.\frac{d^{2n+1}I}{d\lambda^{2n+1}}\right|_{\lambda_0} \quad \text{(A2.8.8)}$$

where $L_n^k(z)$ is the Laguerre Function which are polynomials in z with the highest power going as *n*. For negative z they are also positive definite. Thus, as *n* increases each successive term adds higher powers of $\frac{\Delta\lambda^2}{8w^2}$ with the sign determined by the *(2n+1)*th derivative evaluated at λ_0. As anticipated, for $\Delta\lambda < 2\sqrt{2}w$ successive terms become smaller and smaller allowing for a finite value of Δ consistent with the two channel antenna model. Therefore, we take $\Delta\lambda \leq 2\sqrt{2}w$ as the operable bandwidth that satisfies our design considerations discussed above.

To perform a parameter search for the values of λ_0 and $\Delta\lambda$ that quiet a Noisy Antenna we want a quantity that satisfies our design considerations, i.e. one maximized in the operable bandwidth while also maximizing power bandwidth within that range. As before, we start by considering Δ as our optimization parameter. Fig. A2.8.1a shows two sets of peaks, the blue peaks are when $\Delta\lambda \sim 2\sqrt{2}w$ and the green peaks when $\Delta\lambda$ is the maximum possible i.e. with one peak on the spectral maximum and one on the edge of the spectrum. Evaluating Eq. A2.8.3 involves integrating over the peaks, as indicated by the shading of the area under the curves, and the result is show in Fig. S2.8.1b where Δ is maximized when $\Delta\lambda$ is the maximum possible. As expected, simply maximizing Eq. A2.8.3 will *not* quiet a noisy antenna, but it provides a framework to do so.

To develop a better optimization parameter, we integrate Eq. A2.8.3 with modified bounds of integration:

$$\Delta^{\mathrm{op}}(\lambda_0, \Delta\lambda, w) = \int_{\lambda_0 - m}^{\lambda_0 + m} [a(\lambda, \lambda_0, \Delta\lambda, w) - b(\lambda, \lambda_0, \Delta\lambda, w)] I(\lambda) d\lambda \qquad \text{(A2.8.9)}$$

where *m* is an open parameter that sets a bound to the vicinity of λ_0. In other words, we are only considering the local contribution to the integral within an interval 2*m* wide, around a point λ_0. As $\Delta\lambda$ increases, the peaks will fall outside the operable bandwidth and not contribute to the integral. This is shown schematically in Fig. A2.8.1c, where only the 2*m* interval, indicated by the shading, is integrated and the wide peaks at $\Delta\lambda^{max}$ are excluded. The result is shown in Fig. A2.8.1d where we see that the optimization parameter Δ^{op} is maximized on the ideal bandwidth, and as $\Delta\lambda \rightarrow \Delta\lambda^{max}$ the optimization parameter $\Delta^{op} \rightarrow 0$ because green peaks are outside the bounds of integration. The choice of *m* is somewhat arbitrary, so long as the interval excludes the poorly tuned case and contains the maxima of the peaks when $\Delta\lambda = 2\sqrt{2}w$, then a change in *m* will not significantly change the location of the maxima. For computational convenience we choose *m* = 2*w* without loss of generality, and write

$$\Delta^{\mathrm{op}}(\lambda_0, \Delta\lambda, w) = \int_{\lambda_0-2w}^{\lambda_0+2w} [a(\lambda, \lambda_0, \Delta\lambda, w) - b(\lambda, \lambda_0, \Delta\lambda, w)] I(\lambda) d\lambda \ . \qquad \text{(A2.8.10)}$$

Equation A2.8.10 is the integral used to calculate all results within this work. Section A2.7 discusses this optimization for all of the spectra shown in Fig. 2.

From this analysis we can make two basic predictions about the model that can then be verified against real photosynthetic spectra. First, since the leading contribution comes from the first derivative of $I(\lambda)$ we expect Δ^{op} from Eq. A2.8.10 is maximized in the vicinity of the inflection point of the spectrum as a function of λ_0. We see that all of our optimizations, shown below in section A2.7, have maxima on the spectral inflection points. In addition, in all prototypical phototrophs shown in Fig 6.6, we find absorption peaks near inflection points in their solar spectra. Second, since we have shown that $\Delta\lambda \sim 2\sqrt{2}w$ is the operable bandwidth, then if the difference between the spectral minimum and maximum is of order $2\sqrt{2}w$ or less there are no optimal peaks because that section of the spectrum is not wide enough. In the case of the purple bacteria, shown in Fig. 6.6e,h, the left side of the spectrum rises from near zero at 700 nm to the spectral maximum at 750 nm, but this rise takes place over a range less than $2\sqrt{2}w \approx 70$ nm. We subsequently do not observe any peaks in the purple bacteria spectrum on the left side of the spectral maximum, consistent with our model.

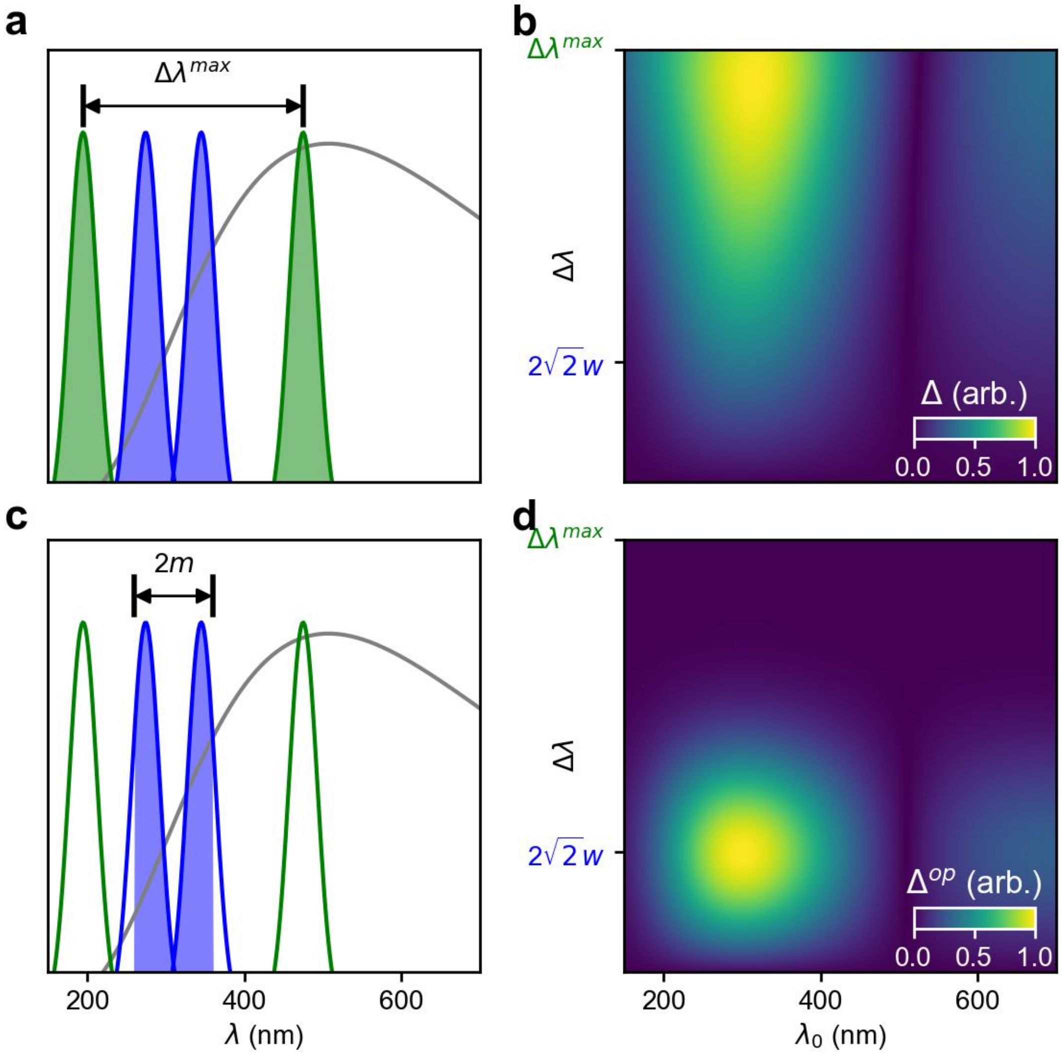


Fig. A2.6.1: **a** shows two pairs of absorption peaks over an ideal blackbody solar spectrum (grey line). The green peaks have the maximum peak separation, $\Delta\lambda^{max}$ and the blue peaks have the ideal bandwidth $\Delta\lambda = 2\sqrt{2}w$. **b**, the calculation of Eq. A2.8.3 for the ideal blackbody in the parameter space $(\lambda_0, \Delta\lambda)$, Δ is maximized at $\Delta\lambda^{max}$. **c**, shows the absorption peaks again but limits the integrated area to an interval $2m$ wide around λ_0, indicated by the shading, as in Eq. A2.8.9. **d**, the calculation of Eq. A2.8.10 for the ideal blackbody in the parameter space $(\lambda_0, \Delta\lambda)$ with $m = 2w$, maximized at $\Delta\lambda = 2\sqrt{2}w$.

A2.7 Optimization of the Noisy Antenna in Diverse Light Spectra

Taking the model discussed in chapter 6, and the optimization parameter derived in appendix A2.6, we perform the optimization for all the spectra described in chapter 6. In order to find the optimum peaks shown in Fig. 6.6g,h,i we use Eq. A2.8.10 to calculate Δ^{op} for the parameter space defined by $(\lambda_0, \Delta\lambda, w)$ using the solar spectrum in three distinct niches as an input. The solar spectrum is input as $I(\lambda)$, and we find the absorption peaks that correspond to the maxima of Δ^{op}. We first start with the solar spectrum at the surface of the Earth, shown as the grey line in Fig. 6.6d. Figure A2.9.1a shows the calculation of $\Delta^{op}(\lambda_0, \Delta\lambda, w = 13\text{ nm})$ for the solar spectrum (low pass filtered to eliminate high frequency spectral noise), and we examine this parameter space to determine the model prediction shown in Fig. 6.6g.

As discussed in appendix A2.6, for a smoothly varying, singly peaked spectrum, like that of a blackbody, there would be two clear maxima at the inflection points on either side of the spectral maximum. However, the complexities of real spectra make this optimization non-trivial. Examining Fig. A2.9.1a we see several local maxima on each side of the spectral maximum. This abundance of maxima is due to fine features of the solar spectrum, which absorbers with w = 13 nm are too narrow to average over. Figure A2.9.2 performs this calculation for arbitrary w. We see that at values of $w > 20$ nm, the calculation is not as sensitive to fine features and shows only two maxima of Δ^{op}, which occur when λ_0 is near the inflection points on either side of the spectral maximum and the peak separation is $\Delta\lambda{\sim}2\sqrt{2}w$.

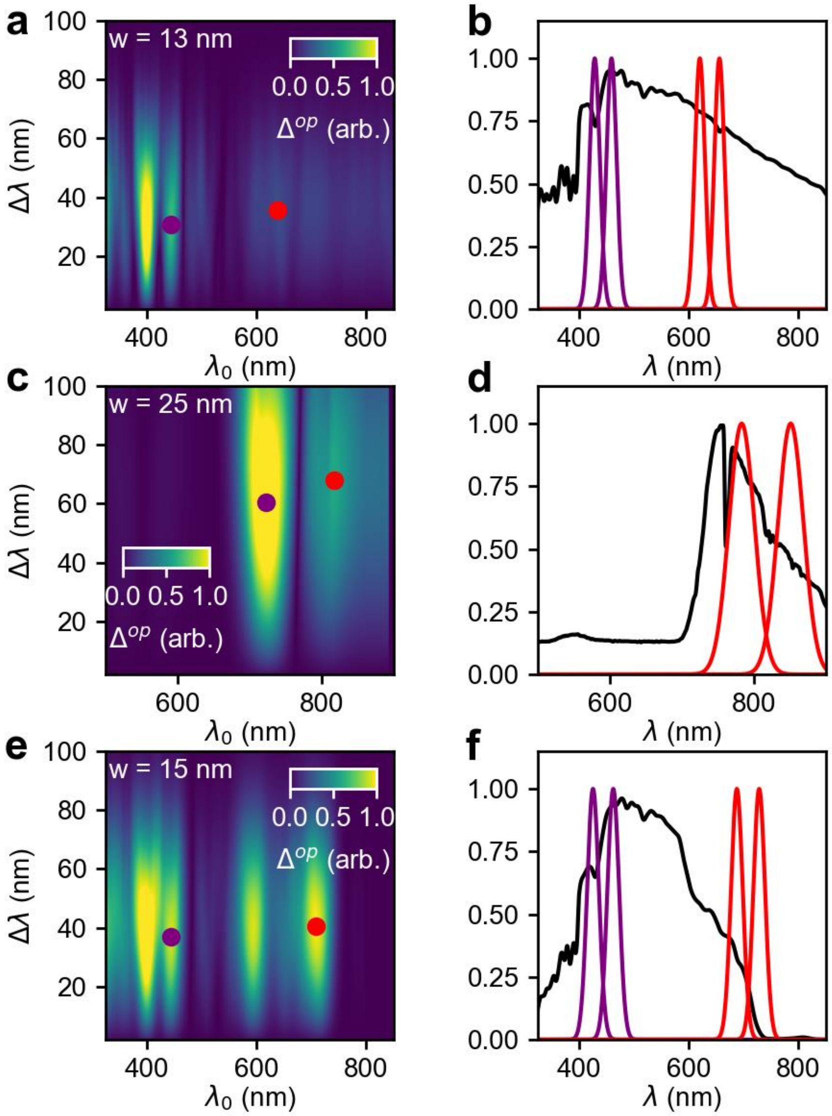


Fig. A2.7.1: The full model optimization of $\Delta^{op}(\lambda_0, \Delta\lambda)$, at a given width *w*, for a given solar spectrum (**left**) and the resulting peaks compared with the solar spectrum (**right**) for **a,b** the terrestrial solar spectrum, **c,d** the canopy-attenuated solar spectrum and **e,f** the solar spectrum under 2 meters of seawater.

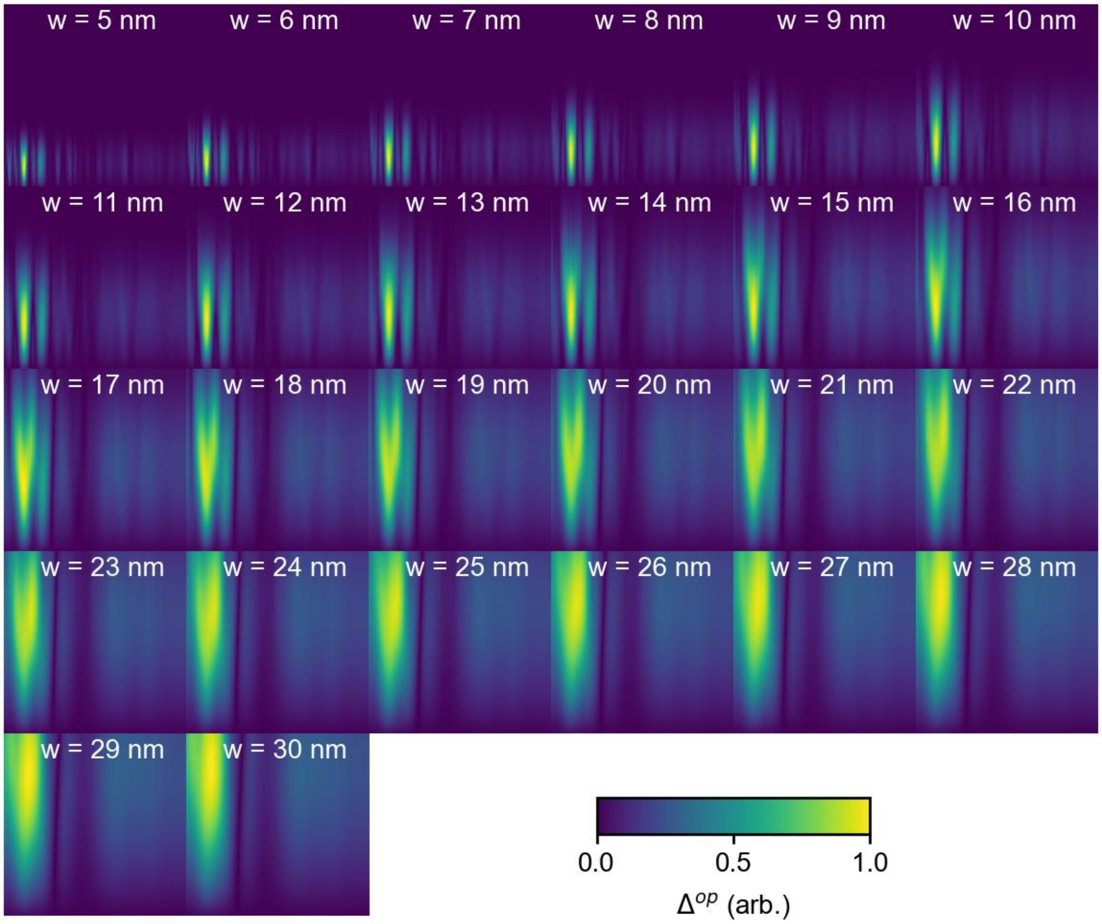


Fig. A2.7.2: Model optimization of the solar spectrum at the surface of the Earth for variable *w*. The axes of each individual panel are the same as the axes in Figure A.2.9.1a.

In photosynthesis, *w* is fixed by the intrinsic absorption of the absorbing pigment molecule, which is 13 nm for the LHC2. Lacking any way to calculate *a priori* the ideal w, we use $w = 13$ nm as an input to our model for chlorophyll and pick out optimal parameters by following the basic model prediction. Therefore, we search for maxima of Δ^{op} on either side of the spectral maximum, shown in Fig. A2.9.1a by the purple and red points. Notably in Fig. A2.9.1a we see two bright peaks in Δ in the 400-450 nm range, these peaks are degenerate for the purpose of our model and we pick the one closer to the spectral maximum, see the discussion below. The purple and red points correspond to two pairs of peaks, which are shown on top of the solar spectrum in Fig. A2.9.1b. As expected, the center wavelength of these peaks lies in the neighborhood of the maximum slope of the spectrum, and near the ideal bandwidth $\Delta\lambda \sim 2\sqrt{2}w \sim 37$ nm.

All of the results presented in Fig. 6.6 were calculated by choosing the spectral width of the relevant photosynthetic pigment, and then finding the maximum in $\Delta^{op}(\lambda_0, \Delta\lambda)$. The optimizations for the terrestrial solar spectrum, the solar spectrum under leaves and the solar spectrum under two meters of water are shown in Figs. A2.9.1a,c,e respectively. These optimizations generate the results shown in Fig. 6.6g,h,i.

The characteristic spectrum under canopy (Fig. 2.9.1c,d) confirms an interesting feature of the noisy antenna model: for optimal noise cancellation, the peak separation should be $\Delta\lambda{\sim}2\sqrt{2}w$. For the optimization shown in Fig. 2.9.1d only one pair of peaks is shown, corresponding to the red maximum in Fig. 2.9.1c on the right-hand side of the spectral maximum. The maximum on the left-hand side of the spectral maximum is shown as a purple point, however this maximum does not correspond to a *fine-tuned* case, due to the fact that the left-hand side of the spectrum rises too sharply. The left side of the spectrum rises from near zero at 700 nm to the spectral maximum at 750 nm, but this rise takes place over a range less than $2\sqrt{2}w$, the operable bandwidth. Thus, for the optimum peaks on the left side, one peak will be near the maximum, and the other will be near the edge, which is the *poorly tuned* case. Therefore, we do not expect to see a pair of left-side peaks. Indeed, if we look at the absorption spectrum of BChl a (Fig. 6.6e), we see only a pair of right-side peaks, which correspond well with the peaks found in Fig. 2.9.1d. In other words, on each side of the spectrum the range between the edge of the spectrum at the maximum is the available bandwidth in which an absorbing pair could exist. In order to have a fine-tuned noisy antenna, the operable bandwidth must be less than the available bandwidth. In this case, the available bandwidth is approximately 50 nm and the ideal bandwidth is $2\sqrt{2}w = 71$ nm, therefore fine-tuned peaks can't exist.

In the optimizations for the various solar spectra, shown in Figs. A2.9.1, we see that there are multiple peaks of Δ^{op}, meaning that there are multiple solutions for a given solar spectrum. Furthermore, in the optimization for the solar spectrum (Fig. A2.9.1a) and underwater (Fig. A2.9.1e) we do not pick out the largest peaks in Δ^{op} for the peaks on the left side of the spectral maximum. To explain this, we must examine the peaks on the left side of the solar spectrum in more detail.

Figure A2.9.3 shows the model optimization for two slightly different solar spectra. Fig. A2.9.3a shows the Δ calculation for the NREL data of the extraterrestrial solar spectrum[106]. This is the data we use for our terrestrial results to avoid any atmospheric features of the solar spectrum, which are usually variable and would themselves be a source of external fluctuation. Fig. A2.9.3b shows the Δ^{op} calculation for a similar NREL spectrum taken at the surface of the Earth (Direct Circumsolar spectrum from the Air Mass 1.5 measurement). Comparing the spectra (black lines) we see that there are some small differences between the spectra, but they have the same large-scale features. From the colored line traces of Δ^{op} we see, as in Fig. A2.9.2, that at small w there are always two prominent peaks, one near 400 nm and the other near 440 nm, and that as w increases, they merge together into a single peak near 425 nm. In Fig. A2.9.3a the 400 nm peak is clearly larger, but in Fig. A2.9.3b they are nearly the same for some values of w. Thus, there are two clear peaks in this wavelength range, and their relative amplitudes depend on fine features of the solar spectrum that can vary with atmospheric conditions, i.e. the raw amplitudes are not particularly meaningful. In other words, for the purposes of the model these two peaks are degenerate.

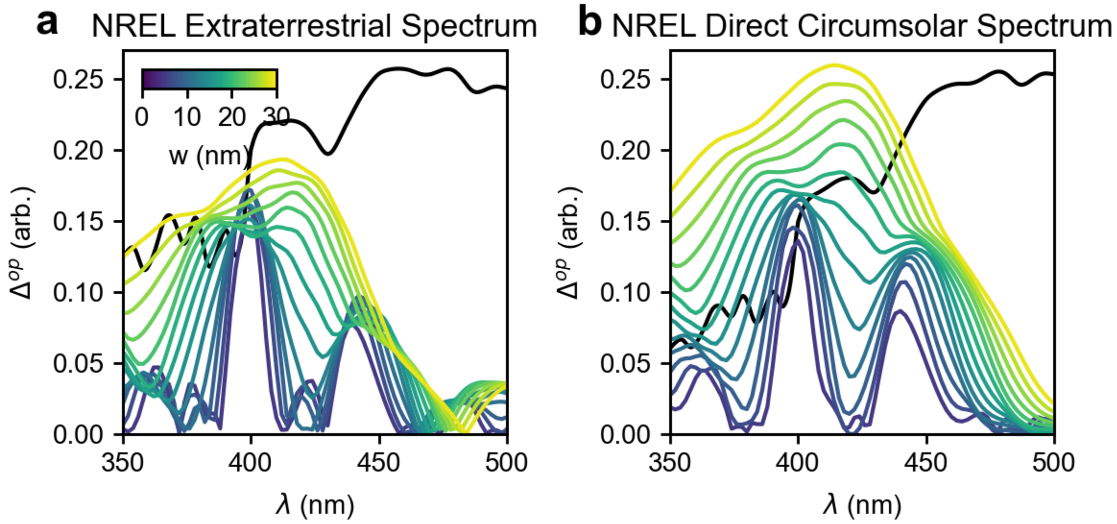


Fig. A2.7.3: **a**, the Δ^{op} optimization for the NREL Extraterrestrial solar spectrum. Colored lines show Δ^{op} for variable w at $\Delta\lambda = 2\sqrt{2}w$, compared with the solar spectrum (black line, arb units). **b**, the Δ^{op} optimization for NREL Direct Circumsolar solar spectrum, which is attenuated by Earth's atmosphere. Colored lines show Δ^{op} for variable w at $\Delta\lambda = 2\sqrt{2}w$, compared with the solar spectrum (black line, arb. units).

Given two degenerate peaks in the optimization, how do we choose which solution to include in the model prediction? Given that the two solutions are equal from the perspective of the model, we look at how the peaks line up with the data. When compared to multiple measurements in the literature the solution corresponding to the λ_0 ~ 400 nm peak exhibits a large and systematic error when compared with all of the measurements. In contrast, the λ_0 ~ 440 nm peak lines up well with the measured data across multiple values reported in the literature. Therefore, we conclude that nature uses the λ_0 ~ 440 nm peak. This analysis works equally well when applied to the left side peaks of the underwater spectrum for Green Sulphur Bacteria, which is expected given that the structure of the left side of the solar spectrum is largely unaffected by water (see Fig.6.7). There are several potential hypotheses for why the λ_0 ~ 440 nm peak is selected, the simplest being that if the two peaks are approximately equally advantageous for quieting a noisy antenna, nature might select the one with higher power throughput. But testing these hypotheses is beyond the scope of this work. From the perspective of our model there are two degenerate solutions for the left side of the solar spectrum and nature seems to use the one closer to the spectral maximum.