Quasiparticle Energy and Excitons in Two-Dimensional Structures

Yufeng Liang
Washington University in St. Louis

Follow this and additional works at: http://openscholarship.wustl.edu/etd

Recommended Citation

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Physics

Dissertation Examination Committee:
  Li Yang, Chair
  Erik Henriksen
  Cynthia Lo
  Richard Loomis
  Alexander Seidel
  Stuart A. Solin

Quasiparticle Energy and Excitons in Two-Dimensional Structures

by

Yufeng Liang

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2014
St. Louis, Missouri
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 THEORETICAL FRAMEWORK</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Density Functional Theory (DFT)</td>
<td>12</td>
</tr>
<tr>
<td>2.1.1 Hohenberg-Kohn Theorems</td>
<td>12</td>
</tr>
<tr>
<td>2.1.2 Kohn-Sham Equation</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Modeling Dielectric Functions</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Polarizability</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2 Asymptotic Behaviors</td>
<td>16</td>
</tr>
<tr>
<td>2.3 Quasiparticle Properties</td>
<td>19</td>
</tr>
<tr>
<td>2.3.1 GW approximation</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2 Generalized Plasmon-Pole Model</td>
<td>23</td>
</tr>
<tr>
<td>2.3.3 Self-Consistency</td>
<td>24</td>
</tr>
<tr>
<td>2.4 Optical Properties and Excitons</td>
<td>25</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Bethe-Salpeter Equation</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Optical Spectra and Excitonic Wavefunctions</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Hydrogen Model of Excitons</td>
</tr>
<tr>
<td>3</td>
<td>QUASIPARTICLE AND EXCITONIC PROPERTIES OF GATED BILAYER GRAPHENE</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>Methods and Computational Settings</td>
</tr>
<tr>
<td>3.3</td>
<td>Quasiparticle Properties</td>
</tr>
<tr>
<td>3.4</td>
<td>Optical Properties and Absorbance</td>
</tr>
<tr>
<td>3.5</td>
<td>Exotic Excitonic States</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusions</td>
</tr>
<tr>
<td>4</td>
<td>STRONGLY BOUND EXCITONS IN TWISTED BILAYER GRAPHENE</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>4.2</td>
<td>Structural Properties</td>
</tr>
<tr>
<td>4.3</td>
<td>Computational Settings</td>
</tr>
<tr>
<td>4.4</td>
<td>Optical Spectra</td>
</tr>
<tr>
<td>4.5</td>
<td>New Mechanism for Strongly Bound Excitons in Gapless 2D Structures: Ghost Fano Resonance</td>
</tr>
<tr>
<td>4.6</td>
<td>Conclusions</td>
</tr>
</tbody>
</table>
5 QUASIPARTICLE ENERGIES IN TRANSITION METAL DICHALCOGENIDES ................................................. 65
  5.1 Introduction ................................................................. 65
  5.2 Computational Settings .................................................. 67
  5.3 Results and Discussions ................................................ 68
  5.4 Conclusions .................................................................. 78

6 TOWARDS MANY-ELECTRON EFFECTS IN ELECTROSTATICALLY DOPED REDUCED-DIMENSIONAL STRUCTURES .................................................. 79
  6.1 Introduction ................................................................... 79
  6.2 Doping-Induced Variation in Screening Effects ......................... 83
    6.2.1 The static dielectric function ....................................... 85
    6.2.2 The dynamical dielectric function .................................. 90
  6.3 Quasiparticle Properties: $G_0W_0$ Approximation ....................... 99
  6.4 Quasiparticle Properties: $GW_0$ Approximation ...................... 116
  6.5 Conclusions .................................................................. 119

7 Conclusions ............................................................................ 121

REFERENCES ............................................................................. 124

VITA .............................................................................................. 138
ACKNOWLEDGMENTS

First of all, I would like to express my deep gratitude to my thesis advisor, Prof. Li Yang, who provides me with enormous support and intelligent guidance throughout my journey towards a physicist researcher. Listening to his explanations on physics has been very enjoyable as he often pulls me through the bog of a large sum of equations and leads me to sound physical results. I particularly appreciate his advice for my work: keep calm and focusing on one thing at a time. This piece of advice has not only been very helpful in promoting my research progress but also quite influential to my way of dealing other issues in my life.

I would like to give my thanks to the members of committee, Prof. Stuart A. Solin, Alexander Seidel, Erik Henriksen, Richard Loomis, and Cynthia Lo, from whom I’ve learned how to view problems from different perspectives. Meanwhile, I am also thankful to a lot of professors in this department who excels in teaching and greatly broaden my horizon. They include but not limited to Prof. Mark G. Alford, Zohar Nussinov, Anders E. Carlsson, Carl Bender, and Willem H. Dickhoff.

I would also like express my thanks to my peer group members, Shouting Huang, Ryan Soklaski, Ruixiang Fei, Vy Tran, and Shiyuan Gao, for being true good friends and ready to help. Without them, my graduate study would have been much harder and more monotonous.
Also thank Runfei Luo for her long-distance accompany during my last year staying here in Washington University and her critical advise when I was hunting for a postdoctoral position.

Lastly, I own much to my parents and my grandmother who raised me up and made my dream being a scientist possible. My parents have been very supportive to what I am pursuing, giving me advises but not affecting my decisions. During the composition of this thesis, my grandmother is still in a ICU and hereby I also send my bless to her.
Dedicated to those who have a dream.
ABSTRACT

Quasiparticle Energy and Excitons in Two-Dimensional Structures

by

Yufeng Liang

Doctor of Philosophy in Physics,

Washington University in St. Louis, August, 2014.

Professor Li Yang, Chair

Two-dimensional materials, such as graphene-related structures, transition metal dichalcogenides, are attracting enormous interest in nowadays condensed matter physics. They not only serve as ideal testbeds for rich physics in reduced-dimensional electron systems but are also of particular importance in nanoelectronic technology. Their electronic, transport, and optical properties are largely determined by the nature of excited states, such as quasiparticles and excitons. Understanding how these excited states emerge from a many-electron system is an intriguing intellectual process, which gives insight into experimental observation and sheds light on manipulating the materials’ properties. From this aspect, it is highly desirable to introduce many-body perturbation theories, which do not rely on data from experiments, to study these excited-state properties and their relations to experimental measurements. In this
thesis, I will present a comprehensive study on a variety of two-dimensional materials using first-principles calculation with many-body effects taken into account. Particular attention is given to the impact of electrical gating, stacking order, and doping on the quasiparticle and excitonic properties.

In chapter 3, we obtain the accurate quasiparticle (QP) band gap and optical absorption spectra of gated bilayer graphene (GBLG), which is a prototypical 2D field effect transistor. Enhanced electron-electron interactions dramatically enlarge the QP band gap; infrared optical absorption spectra are dictated by bright bound excitons. In particular, the energies of these excited states can be tuned in a substantially wider range, by the gate field, than previous predictions. Our results clearly explain recent experiments and satisfactorily resolve the inconsistency between experimentally measured transport and optical band gaps. Moreover, we predict that the most deeply bound exciton is a dark exciton which is qualitatively different from the hydrogenic model, and its electron and hole are condensed onto opposite graphene layers, respectively. This unique dark exciton will not only impact the exciton dynamics but also provide an exciting opportunity to study entangling exchange effects of many-body physics.

In chapter 4, we study the excitonic effects in twisted bilayer graphene and reveal a novel mechanism for bound exciton formation in a metallic system. Common wisdom asserts that bound excitons cannot form in high-dimensional ($d > 1$) metallic structures because of their overwhelming screening and unavoidable resonance with nearby continuous bands. However, we will illustrate that this prevalent assumption
is not true. A key ingredient that has been overlooked is the destructive interference effect that thwarts the formation of resonances. We use twisted bilayer graphene as an example to investigate the possibility of this exciton formation mechanism. Excitonic effects in the system are examined by both first-principles methods and a low-energy effective model.

In the following chapters, we turn to study the excited properties in transition metal dichalcogenides. In chapter 5, we report on the quasiparticle band-edge energy of monolayer of molybdenum and tungsten dichalcogenides, MX$_2$ (M=Mo,W; X = S, Se, Te). Beyond calculating bandgaps, we have achieved converged absolute band-edge energies relative to the vacuum level. Compared with the results from other approaches, the GW calculation reveals substantially larger bandgaps and different absolute quasiparticle energies because of enhanced many-electron effects. Interestingly, our GW calculations ratify the band-gap-center approximation, making it a convenient way to estimate band-edge energy. The absolute band-edge energies and band offsets obtained in this work are important for designing heterojunction devices and chemical catalysts based on monolayer dichalcogenides.

In chapter 6, we report on a first-principles study of the band gap renormalization, using the doped MoS$_2$ monolayer as a test material. A parameter-free generalized plasmon pole model is proposed to capture the low-frequency dynamical screening that is featured by the carrier plasmon arising from intraband transitions. Our calculations show that the quasiparticle band gap of monolayer MoS2 exhibits a unique evolution according to the doping density; it drops quickly at the low doping den-
sity but saturated at 2.3 eV, which is around 300 to 400 meV smaller than that of undoped MoS2. This result is crucial for understanding and explaining current experimental measurements and our proposed approach can be employed to general doped 2D semiconductors. Meanwhile, this tunable band gap via doping may work as an efficient tool to tailor the electronic and optical properties of two-dimensional semiconductors for broad applications.
1. INTRODUCTION

Ever since the discovery of intriguing low-dimensional physics in monolayer graphene [1] in 2004, there has been a fast-growing interest in exploring the remarkably varied and exotic properties of two-dimensional (2D) materials among the physics and engineering society. Not only did the fabrication of monolayer graphene defy the famous Mermin-Wagner theorem [2], asserting that exact 2D materials are impossible to exist, but it also gave way to an amazing class of excitations in solids, Dirac fermions, that mimic massless relativistic particles by virtue of graphene’s honeycomb lattice structure [3–5]. The massless fermions possess a very high mobility (up to $\mu = 200,000 \text{cm}^2\text{V}^{-1}\text{s}^{-1}$ at room temperature), making graphene a particularly attractive candidate for realizing high-speed integrated circuit that hopefully outperforms the silicon-based circuit widely used nowadays [6]. This also ignites a journey off searching for materials that have comparable intriguing properties. Along the route are the discoveries of a series of graphene-like structures, such as silicene [7, 8] and phosphorene [9].

Despite its high carrier mobility, monolayer graphene does not have a sizeable electronic band gap, making the structure unsuitable for use in bipolar junction transistors or photovoltaic applications. Substantial efforts have been made towards gen-
erating a band gap in graphene-related structures by breaking the lattice symmetry. Such efforts include passivating graphene’s surface with alien atoms, leading to the formation of graphane [10] and fluorographene [11] atoms, and applying voltage biases a Bernal-stacking bilayer graphene sheet [12–14]. Unfortunately, the size of the band gap in these materials are either too large or too small, which is less desirable compared with the traditional bulk semiconductors.

The isolation of monolayer molybdenum disulfide (MoS$_2$) [15] in 2010 started a new paradigm using layered transition metal dichalcogenides for nonelectronic applications. Unlike graphene, monolayer MoS$_2$ is an intrinsic 2D semiconductor with a direct band gap, as verified by means of photoluminescence spectroscopy [15, 16]. By contrast, strong photoluminescence does not emerge in a bulk of MoS$_2$ that has an indirect band gap. Furthermore, it is mechanically stable. These properties make monolayer MoS$_2$ an excellent candidate for optoelectronic application in the visible-light region. More excitingly, this material has already been used to create field-effect transistors with high on/off ratio and appreciable mobility MoS$_2$ [17–20]. Similar electronic and optical properties have been experimentally demonstrated in other 2D dichalcogenides, such as monolayer WS$_2$, WSe$_2$, and MoSe$_2$ [21, 22].

The electronic, transport, and optical properties of the aforementioned materials are determined by a family of elementary excitations that arise in complex many-electron systems. For example, an electron travelling a solid can interact with other electrons, forming an electron quasiparticle excitation that behaves like noninteracting electrons but with modified properties, such as carrying a different effective mass and
gaining an additional amount of interaction energy. Likewise, a hole quasiparticle results from a vacant state in the valence band. The energies of quasiparticles in a semiconductor determine the size of its band gap, and control the band alignment at a bipolar junction structure. An exciton is produced when an electron is kicked out from the valence band by incident photons, and is attracted by the positively charged hole that was left behind. This electron-hole pairing is an important mediator between electricity and light. A plasmon is a many-body excitation in which the electrons in a crystalline structure oscillate collectively. Just like the exciton, a plasmon can also resonate with the electromagnetic field of light, which represents another underlying mechanism for optoelectronics via light-matter interactions [23–25].

The aforementioned excitations are a consequence of complicated many-electron interactions that are governed by the laws of quantum mechanics. Solving a many-body system is a persistent and difficult problem even on the level of classical mechanics, not to mention the complex quantum entanglement in the world of quantum mechanics. However, remarkable simplifications to many-body systems that precisely describes the nature of excited states are available and readily utilized. The landau-fermi liquid theory provides a convenient means for describing quasiparticles [26, 27]. It states that a weakly interacting electron system can be mapped onto a non-interacting system, which is conceptually simple. In an exciton, although electron and hole are correlated through delicate many-body effects, their motion and energy spectrum can often times be understood through a mapping to the hydrogen model [28, 29]. This is true in many semiconductors. The two kinds of excitations,
quasiparticles and excitons, play key roles in understanding most of the electronic and optical properties of solids and are the main themes of this thesis. It is of particular importance to these excited states to understand how screening effects, which arise from the migration of the charge inside the material, impact these excited states.

When electrons are confined in a low-dimensional material, many-body effects can be profoundly enhanced. For example, while the exciton binding energies are of a few tens of meV in bulk semiconductors, it can reach a few hundred to several thousand meV in reduced-dimensional structures. Such structures include a series of 1D materials, such as carbon nanotubes [30–32], boron nitride tubes [33,34], and graphene nanoribbons [35–37] as well as a number of 2D materials, such as layered boron nitride [38], graphane [39,40], fluorographene [41], graphyne [42], silicene [43], germanene [44], silicane, germanane (passivated silicene and germanene respectively) [44,45], and a variety of dichacogenides [46–50]. The enhancement of excitonic effects is partially due to a reduction in screening effects in low-dimensional structures because of their surrounding vacuum. Additionally, the density of states (DOS) at the band edges for 1D and 2D semiconductors are proportional to \( \frac{1}{\sqrt{E}} \) and roughly a constant respectively, which are significantly larger than the bulk case (\( \sqrt{E} \)). This also leads to enhanced many-body interactions because quantum states that are close in energy tend to correlate.

The divergence in the DOS of \( \frac{1}{\sqrt{E}} \) in a 1D system is a kind of van Hove singularity (vHs) and is an important indicator of possible enhanced excitonic effects. Interestingly, this divergence in DOS can also occur in a 2D system. For example,
the saddle-point in the band structure of graphene [5] produces a vHs, which manifests itself as a prominent absorption peak in the optical spectrum. In spite of the semimetallic nature of graphene, the redshift of the absorption peak due to excitonic effects can be as large as 600meV [51,52]; this is even comparable to excitonic redshifts in semiconducting nanotubes. Very intriguing examples of vHs in 2D systems are also discovered in multilayered graphene structures. Bilayer graphene systems produced via chemical vapor deposition (CVD) may have two graphene layers that are stacked in different orientations. This has given rise to a new class of materials called twisted bilayer graphene (tBLG) [53, 54]. The band structure of the respective layers do not lay on the top of each other; their energy band intersection yields to new vHs’. Depending on how far apart one layer is twisted away from the other, the energy level and the strength of these vHs’s can be tuned in reference to the Fermi level. This enables one to control the many-body effects in a 2D system via simply manipulating the stacking order, which is less likely to achieve in a bulk structure. Together with the original saddle-point, these band intersections form a complex system of vHs that can not be explained within the simple hydrogen model.

When a bulk material is exfoliated into a few-atom-thick layer, electrical gating provides an important means of controlling its electronic and optical properties because the gate field is not as effectively screened as in the bulk case. While the electric field can modify the band structure via symmetry breaking [12,55–58], they can also shift the Fermi level from the charge neutral point, doping the material with free charge carriers. Doping introduces new intraband transitions that can bolster the screen-
ing considerably. This offers a special opportunity for modifying the many-electron effects that are dictated by screened Coulomb interactions. More importantly, the electrons that occupy the doped band behave like a 2D electron gas that hosts a plasmon excitation. This excitation is markedly different from a 3D one: the plasmon oscillation is quenched at the long-wavelength limit in a 2D case whereas it persists in a 3D case. The 2D plasmon dispersion relation scales as $\omega \sim \sqrt{q}$ which has the same small-momentum limit with the quasiparticle dispersion relation $\omega \sim q^2$ [27, 59, 60]. The energy proximity of the dispersion relation implies that there could be a strong coupling between the quasiparticle and the carrier plasmon in a doped 2D material that can largely reshape the excited-state properties.

In this thesis, I will present a theoretical and computational study of the excited states in a variety of new 2D materials that are potentially viable for incorporation in optoelectronic or transistor devices. Our calculations are mainly based on the first-principles methodologies [61–65, 65], which do not rely on fitting parameters from experiments. The density functional theory (DFT) provides a very accurate description of the ground-state properties of materials, including realistic crystal structures, electronic charge density, and non-interacting electron wavefunctions. However, it is well known that the DFT systematically underestimates the band gap of a semiconductor [62, 64, 66], To obtain reliable band gaps, which is quasiparticle property, electron-electron (e-e) interactions need to be captured by many-body perturbation theory. The widely used GW approximation [62, 64, 67] is a successful treatment of the e-e interaction for many medium correlated materials. The approximation is an
aptly named one: the self-energy is a product of green function $G$ and the screened Coulomb interaction $W$, meaning that the electron propagating in a solid is dressed by the screening effects of the background as formed by the other electrons. The exciton is a two-body excitation that substantially impacts the optical responses of materials. This excitation is not described by the DFT or GW levels of approximation. The Bethe-Salpeter equation (BSE) [63], which was originally formulated for electron-positron bound pairs, is used to include $e$-$h$ pairing in the theoretical model.

On a personal note, what quantum many-body physics brings to my mind is the *Starry Night Over the Rhone* painted by Vincent van Gogh in 1888. I have long imagined that the richness of physics in an interacting particle system that resembles the beautiful scene that the artist painted on the canvas with his boundless imagination. This bolstered my motivation for investigating the "fine art" of many-body physics. That being said, the goal of this study is not only to provide benchmark parameters to further the intense study of reduced-dimensional materials but also explore new interesting excited-state properties in these materials within the framework of the $GW+BSE$ method. I will demonstrate two examples of excitonic states that can not be simply understood by the hydrogen model. The first one occurs in the gated bilayer graphene, which possesses a mexican-hat dispersion feature in its band structure, leading to the failure of the effective mass approximation that is essential in the hydrogen model. The unique band structure gives rise to an exotic excitonic state in which electrons and holes are separated onto opposite layers. The second example occurs in twisted bilayer graphene, which is semimetallic and has multiple
Figure 1.1. Starry Night over the Rhone, Vincent van Gogh, 1888
vHs’s. For the first time, we predict that a strongly bound exciton state can exist in this gapless 2D system in spite of its semimetallic screening. This is made possible by virtue of quantum coherence, an effect that is not captured by the ordinary hydrogen model. The other part of the thesis will focus on the excited-state properties of metal dichalcogenides. Monolayer MoS$_2$, which is naturally n-doped in experiments, will be used as a test material to tackle the challenging band gap renomalization problem in doped 2D systems whose many-body interactions are exhibited by quasiparticle-plasmon coupling.
2. THEORETICAL FRAMEWORK

All the physics in an interacting many-electron system can be described within the full Hamiltonian \[68, 69\]

\[
\hat{H}_{\text{el}} = -\frac{\hbar}{2m_e} \sum_i \nabla_i^2 + \sum_i V_{\text{ext}}(r_i) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|r_i - r_j|}
\]  

(2.1)

where \(i\) goes over all the electrons in the system. The first two terms are the kinetic energy and external potential of independent electrons respectively whereas the third term is the Coulomb interaction between all pairs of electrons. The external potential \(V_{\text{ext}}\) incorporates the electron-ion interaction in a crystal as well as the effect of an external electric field.

Solving the all-electron Hamiltonian by brute-force is a formidable computational task since for a system with \(N\) electrons the Hamiltonian gives \(3N\) partial differential equations. The density functional theory (DFT) which emerged in the 60s \[61, 70\] is a remarkable breakthrough for dramatically simplifying the many-electron problem and lays the foundation for a variety of modern excited-state calculations based on supercomputers. The success of the DFT started with the two Hohenberg-Kohn theorems \[61\] on the total electron density in a many-electron system.
2.1 Density Functional Theory (DFT)

2.1.1 Hohenberg-Kohn Theorems

The first Hohenberg-Kohn theorem states that the external potential \( V_{\text{ext}}(r) \) is uniquely determined by the corresponding ground-state charge density \( n(r) \), except for an additive constant. This means all the ground-state properties are merely a functional of \( n(r) \), which profoundly simplifies the \( N \)-electron problem with \( 3N \) degrees of freedom into a problem with only 3 spatial degrees of freedom. The second Hohenberg-Kohn theorem states that for any external potential \( V_{\text{ext}} \), there is a universal energy functional \( E[n] \), which is minimized at the correct ground-state charge density distribution \( n(r) \). Kohn and Sham proposed an ansatz to find the \( n(r) \) that assumes the many-electron system can be mapped onto an auxiliary non-interacting-electron system with orthogonal Kohn-Sham orbits \( \phi_{nk} \) that has the same set of good quantum numbers \( nk \). The auxiliary system has the same number of valence orbits as the system without \( e-e \) interactions. Thus, the total charge density can be viewed as the summation of the charge density contributed by each occupied Kohn-Sham orbit \( \phi_{nk} \): \( n(r) = \sum_{nk}^{\text{occ}} \phi_{nk}^*(r)\phi_{nk}(r) \). A general form of the energy functional is given by

\[
E[n] = T[n] + \int dr V_{\text{ext}}(r)n(r) + E_{\text{Hartree}} + E_{\text{II}} + E_{\text{xc}}[n] \tag{2.2}
\]

where \( e-e \) interaction functional breaks down into a Hartree term \( E_{\text{Hartree}} \) and a exchange-correlation term \( E_{\text{xc}} \) and \( E_{\text{II}} \) is the ion-ion interactions.
2.1.2 Kohn-Sham Equation

The procedure of solving the many-electron problem is substantially advanced by another ambitious approximation over the exchange-correlation functional $E_{xc}$, presuming that $E_{xc}$ is only dependent on the charge density at each point in space

$$E_{xc}[n] = \int n(r) \epsilon_{xc}(n(r)) dr,$$  \hspace{1cm} (2.3)

which is the local density approximation (LDA) [61]. Minimizing the energy functional 2.2 with LDA yields a Schrödinger-like equation for the auxiliary electron system

$$\{-\frac{1}{2} \nabla^2 + [V(r) + \mu_{xc}(n(r))]\} \phi_{nk}(r) = \epsilon_{nk} \phi_{nk}(r)$$  \hspace{1cm} (2.4)

where $V(r)$ is the summation of the external (ionic) potential and the Hartree energy.

$$V(r) = V_{ext}(r) + \int d\mathbf{r'} \frac{n(r')}{|\mathbf{r} - \mathbf{r}'|}$$  \hspace{1cm} (2.5)

and $\mu_{xc}$ the exchange-correlation potential as a functional of charge density

$$\mu_{xc}(n) = \frac{d(n \epsilon_{xc}(n))}{dn}$$  \hspace{1cm} (2.6)

Eq. 2.4 is the famous Kohn-Sham equation, which can be solved by a self-consistent computational scheme. For example, using the atomic orbital wave functions to construct the initial charge density and thus the potential functional. And then keep
solving the equation and updating the orbital energies and wavefunction until the
total energy converges. The final orbital energy $\epsilon_{nk}$ will be the DFT-LDA energies.

The Kohn-Sham equation can be used to accurately determine a number of ground-
state properties [66, 71–73], such as charge density distribution, total energies, atomic
forces, and equilibrium crystal structures that are in excellent agreement with exper-
iments. However, the Kohn-Sham equation is inadequate for predicting the band gap
of semiconductors which is a quasiparticle property [63, 71]. Neither can the Kohn-
Sham equation be used for studying the excitonic effects for optical properties, which
is a two-body excitation. In general, there is no such “Hohenberg-Kohn” theorems cor-
responding to these excited states so that density functional theory can be extended
straightforwardly to solving these problems. To precisely capture the nature of quasi-
particle and excitonic states, it is essential to incorporate many-body perturbation
theory [62, 63] as introduced in the following sections.

2.2 Modeling Dielectric Functions

The key to the understanding the excited states is the screened Coulomb inter-
action $W$ which reflects how strongly electrons are correlated. $W$ is the product of
the inverse dielectric function $\epsilon^{-1}$ and the bare Coulomb interaction $W = \epsilon^{-1}v$. It is
very important to accurately model the dielectric function before it is finally plugged
into the excited-state calculation. From the basis relations in electromagnetism, the
dielectric function can be expressed in terms of the polarizability $P$ [27]

$$\epsilon = 1 - vP$$

(2.7)

The polarizability $P$ is a measure of how strong an electron system is polarized in the
presence of an electric field. The full expression of the dielectric function for a crystal
structure should take lattice symmetry and local field effects [74,75] into account and
incorporate frequency-dependence

$$\epsilon_{G'G}(q, \omega) = \delta_{G'G'} - \nu(G + q)P_{G'G}(q, \omega)$$

(2.8)

Now $\epsilon$ and $P$ are two-point correlation functions relating the scattering process be-
tween the momentum $q + G$ and $q + G'$, where $G$ is a reciprocal-lattice vector. The
scattering process between $q + G$ and $q' + G'$ with $q \neq q'$ is forbidden by lattice sym-
metry and hence is not present in Eq. 2.8. The problem of modelling the dielectric
function is thus converted into formulating the polarizability $P_{G'G}$ for crystals.
2.2.1 Polarizability

The polarizability can be calculated within the random phase approximation (RPA) [27, 69, 74, 75], which is a successful theory for describing the crystal screening effect on a microscopic level. The polarizability given by RPA reads

\[ P_{G\bar{G}'}(q, \omega) = \frac{1}{V} \sum_{nn'} M_{nn'}^*(k, q, G) M_{nn'}(k, q, G') \frac{f_{nk} - f_{n'k+q}}{\omega + \epsilon_{nk} - \epsilon_{n'k+q}} \]  

(2.9)

where \( f_{nk} \) is the occupation number, and \( \epsilon_{nk} \) is the single-particle energy. \( M_{nn'}(k, q, G) = (nk + q|e^{iG \cdot r}|n'k) \) being the envelope function of a Bloch state \( nk \) (This is equivalent to the notation \( M_{nn'}(k, q, G) = \langle nk+q|e^{iG \cdot r}|n'k \rangle \), where \( |nk\rangle = e^{iq \cdot r} |nk\rangle \)). It is intuitive to understand the expression of the polarizability. It sums over all possible transitions from occupied states to unoccupied states that can contribute screening effects. The contribution of a particular transition to screening is determined by the transition amplitude between the initial state and the final state, namely the matrix element \( M_{nn'}(k, q, G) \), as well as the inverse transition energy \( 1/(\omega + \epsilon_{nk} - \epsilon_{n'k+q}) \), which is the price it takes for that transition to occur.

2.2.2 Asymptotic Behaviors

The asymptotic behavior of \( \epsilon^{-1}(q) \) in the long-wavelength limit \( q \to 0 \) is crucial for determining the excited-state property. The Coulomb potential \( v(q) \) diverges at \( q = 0 \), making the value of \( \epsilon^{-1}(q) \) near \( q = 0 \) very important for deciding the screened Coulomb interaction \( W = \epsilon^{-1}v \). Moreover, \( \epsilon^{-1}(q) \) may vary too rapidly
near \( q = 0 \) for a discrete sampling grid to capture. From this perspective, one can greatly improve the numerical accuracy by performing the so-called mini-BZ average [62, 65] near \( q = 0 \) based on the asymptotic behavior of \( \epsilon^{-1}(q) \). It can be ruled out by analyzing all the \( q \)-dependence ingredients involved in the dielectric functions, including the Coulomb interaction \( v(q) \), the matrix elements \( M_{nn'}(k, q, G) \), and the energy gap \( \epsilon_{nk} - \epsilon_{n'k+q} \). For the Coulomb interaction, while the bulk case is

\[
v(q) = \frac{4\pi}{q^2}
\]

the 2D one is truncated to avoid spurious interaction between unit cell images. In real space, it reads

\[
V_{2D}^t(r) = \frac{\theta (L_z/2 - z)}{r}
\]

This truncation results in a weaker divergence in the long-wavelength limit \( q \to 0 \).

When transformed into the momentum space, it takes the form

\[
v_{2D}^t(q) = \frac{4\pi}{q^2} \left[ 1 - e^{-q_{xy}L_z/2} \cos (q_x L_z/2) \right]
\]

where \( q_{xy} \) is the in-plane component of \( q \) and \( L_z \) is unit cell parameter perpendicular to the 2D material being investigated. It should be noted that the 2D-truncated
versus the crystal momentum $q$.

<table>
<thead>
<tr>
<th></th>
<th>semiconductor</th>
<th>metallic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk Const.</td>
<td>$\propto q^2$</td>
<td></td>
</tr>
<tr>
<td>Two – Dimensional</td>
<td>1</td>
<td>$\propto q$</td>
</tr>
</tbody>
</table>

Coulomb interaction scales as $\sim \frac{2\pi}{q}$ [76, 77], significantly slower than the 3D case.

The asymptotic behavior of the transition amplitude is given by

$$M_{nn'}(k, q, 0) = (nk + q|n'k) = \begin{cases} 1 & n = n' \\ \propto q & n \neq n' \end{cases}$$

Therefore, dependent on whether the system is semiconducting or metallic and whether it is a bulk material or a reduced-dimensional one, the asymptotic behavior of $\epsilon_{00}^{-1}$ varies. The limits can be evaluated by the formula

$$\epsilon_{00}^{-1}(q) \propto \frac{1}{1 + v_{\text{eff}}(q)M(q)^2}$$

where $v_{\text{eff}}(q)$ is effective Coulomb interaction dependent on dimensionality. $M(q) = 1$ for semiconductors whereas $M(q) \propto q$ for metals. Under different circumstances, the limits of $\epsilon_{00}^{-1}$ are compared in Table 2.1 [65].
2.3 Quasiparticle Properties

A quasiparticle is a one-particle excitation in which an electron is dressed by the screened Coulomb interaction. Its energy is solved from the Dyson equation [27,62,69], which is a one-body Green function method

$$\left[-\frac{1}{2}\nabla^2 + V(r)\right] \Phi_{nk}(r) + \int dr' \Sigma(r,r'; E_{nk}) \Phi_{nk}(r') = E_{nk} \Phi_{nk}(r) \quad (2.15)$$

The Dyson equation is highly analogous to the Kohn-Sham equation Eq. 2.4 except that the exchange-correlation functional is replaced by the non-local, energy-dependent self-energy operator $\Sigma$. The formulation of the self-energy plays an important role in determining quasiparticle energies.

2.3.1 GW approximation

According to Hedin [67], the full self-energy can be expressed in terms of a perturbation series of propagator $G$ and screened Coulomb interaction $W$. Taking the first-order in the series leading to the $GW$ approximation to the self-energy operator

$$\Sigma = iGW \quad (2.16)$$

For practical use in calculation, one can represent the self-energy operator in real space [62]

$$\Sigma(r, r'; E) = i \int \frac{dE'}{2\pi} e^{-iE'} G(r, r'; E - E') W(r, r'; E'), \quad (2.17)$$
with the propagator constructed from the single-particle energies $\epsilon_{nk}$ and wavefunctions $\phi_{nk}(r)$

$$G(r, r'; E) = \sum_{n,k} \frac{\phi_{nk}(r)\phi_{nk}^*(r')}{E - \epsilon_{nk} - i\delta_{nk}},$$

(2.18)

where $\delta_{nk}$ is positive (negative) infinitesimal for occupied (unoccupied) state $nk$. In real space, the screened Coulomb interaction $W$ can be expressed as

$$W(r, r'; E) = \sum_{q,G,G'} e^{i(q+G)\cdot r} W_{GG'}(q, E) e^{-i(q+G')\cdot r'}$$

(2.19)

The screened Coulomb interaction is related to the dielectric function previously obtained (Eq. 2.8) via

$$W_{GG'}(q, \omega) = \epsilon_{GG'}^{-1}(q, \omega) \upsilon(q + G'),$$

(2.20)

The real part of self-energy $\Sigma$ can be decomposed into two parts

$$Re\Sigma = \Sigma_{SX} + \Sigma_{CH}$$

(2.21)

where the screened-exchange (SX) term $\Sigma_{SX}$ stems from the poles in the Green’s function and the Coulomb-hole (CH) part $\Sigma_{CH}$ from the poles in the screened Coulomb interaction. Therefore, the SX part involves the summation over all occupied states

$$\Sigma_{SX}(r, r'; E) = -\sum_{n,k}^{\text{occ}} \phi_{nk}(r)\phi_{nk}^*(r') W(r, r'; E - \epsilon_{nk})$$

(2.22)
whereas the CH part involves the “hole” part of $W$

$$\Sigma_{\text{CH}}(r, r'; E) = \sum_{n,k} \phi_{nk}(r) \phi^*_{nk}(r') P \int_0^\infty dE' \frac{B(r, r'; E')}{E - \epsilon_{nk} - E'} \tag{2.23}$$

where $B(r, r'; E) = \frac{1}{\pi} \text{Im}[W(r, r'; E)]$ is the spectral function. Here, the SX term is analogous to the bare exchange energy

$$\Sigma_{\text{X}}(r, r') = - \sum_{n,k} \phi_{nk}(r) \phi^*_{nk}(r') v(r - r') \tag{2.24}$$

except the screened Coulomb interaction $W$ is replaced by the bare Coulomb interaction $v$. It is more convenient to express the self-energy in terms of the Bloch states that are compatible with the Dyson Equation as given by Eq. (2.15)

$$\langle n k | \Sigma_{\text{SX}}(E) | n' k \rangle = - \sum_{n_1} \sum_{q, G, G'} M^*_{n_1 n}(k, -q, -G) M_{n_1 n'}(k, -q, -G') \times \epsilon^{-1}_{G G'}(q, E - \epsilon_{n_1 k - q}) v(q + G') \tag{2.25}$$

$$\langle n k | \Sigma_{\text{CH}}(E) | n' k \rangle = \sum_{n_1} \sum_{q, G, G'} M^*_{n_1 n}(k, -q, -G) M_{n_1 n'}(k, -q, -G') \times \left[ \epsilon^{-1}_{G G'} \right]^h(q, E - \epsilon_{n_1 k - q}) v(q + G') \tag{2.26}$$

where $\left[ \epsilon^{-1}_{G G'} \right]^h$ only contains the positive poles in the inverse dielectric function

$$\left[ \epsilon^{-1}_{G G'}(q, E) \right]^h = \frac{1}{\pi} P \int_0^\infty dE' \frac{\text{Im} \epsilon^{-1}_{G G'}(E')}{E - E'} \tag{2.27}$$
In practice, it has been proved sufficient to just use the diagonal term to estimate the self-energy correction. Then the Dyson equation for a specific state \( n_{\mathbf{k}} \) boils down to a single-variable equation with respect to \( E_{n_{\mathbf{k}}} \)

\[
E_{n_{\mathbf{k}}} = \epsilon_{n_{\mathbf{k}}} - \langle n_{\mathbf{k}}|V_{\text{xc}}|n_{\mathbf{k}} \rangle + \langle n_{\mathbf{k}}|\Sigma(E_{n_{\mathbf{k}}})|n_{\mathbf{k}} \rangle
\]  

(2.28)

From this expression, it can be seen that the quasiparticle energy \( E_{n_{\mathbf{k}}} \) is determined based on the single-particle energy \( \epsilon_{n_{\mathbf{k}}} \). The exchange-correlation energy from the DFT is now replaced by the self-energy that incorporates many-body correlations.

In principle, Eq. (6.17) is nonlinear and needs to be solved iteratively. However, for a semiconductor, the self-energy is typically linear near the single-particle energy \( \epsilon_{n_{\mathbf{k}}} \) and the quasiparticle energy \( E_{n_{\mathbf{k}}} \) is not too far away from \( \epsilon_{n_{\mathbf{k}}} \). Therefore, it is legitimate to expand the self-energy to the first-order so that the quasiparticle energy can be solved from a linear equation.

\[
E_{n_{\mathbf{k}}} = E_{n_{\mathbf{k}}}^0 + \frac{d\Sigma_{n_{\mathbf{k}}}(\epsilon_{n_{\mathbf{k}}})/dE}{1 - d\Sigma_{n_{\mathbf{k}}}(\epsilon_{n_{\mathbf{k}}})/dE}(E_{n_{\mathbf{k}}}^0 - \epsilon_{n_{\mathbf{k}}})
\]

(2.29)

where

\[
E_{n_{\mathbf{k}}}^0 = \epsilon_{n_{\mathbf{k}}} - \langle n_{\mathbf{k}}|V_{\text{xc}}|n_{\mathbf{k}} \rangle + \langle n_{\mathbf{k}}|\Sigma(\epsilon_{n_{\mathbf{k}}})|n_{\mathbf{k}} \rangle
\]

(2.30)

It has to be pointed out that this first-order expansion does not apply to the situation in which the self-energy exhibits strong nonlinearity near the solutions. For example, as will be shown in Chapter 7, doping can cause a fluctuation near the
energy range of interest. In that circumstance, a numerical root-finding scheme has to be employed instead.

2.3.2 Generalized Plasmon-Pole Model

The inverse dielectric function $\varepsilon_{GG'}^{-1}(q, \omega)$ does not only depend on momentum but also depend on the frequency. Calculating and storing the full $\varepsilon_{GG'}^{-1}(q, \omega)$ requires a large amount of computational resources. It would be more convenient if $\varepsilon_{GG'}^{-1}(q, \omega)$ is first calculated in the static limit $\omega = 0$ and then extended to finite frequencies. To simplify the energy-dependence, Hybertsen and Louie [62] proposed a generalized plasmon-pole (GPP) model that yields very little error from the full-frequency calculation.

In the GPP model, the inverse dielectric function is assumed to be a single-simple-pole function, the real part of which can be represented by

$$\text{Re} \varepsilon_{GG'}^{-1}(q, \omega) = \delta_{GG'} + \frac{\Omega_{GG'}^2(q)}{\omega^2 - \tilde{\omega}_{GG'}^2(q)} \quad (2.31)$$

There are two parameters in the model, $\tilde{\omega}_{GG'}(q)$, the plasmon-pole frequency, and $\Omega_{GG'}(q)$, the plasmon-pole strength. The two parameters can be determined by two constraints. One is the inverse dielectric function in the static limit, which gives

$$\varepsilon_{GG'}^{-1}(q, \omega = 0) = \delta_{GG'} - \Omega_{GG'}^2(q)/\tilde{\omega}_{GG'}^2(q) \ .$$

The other constraint comes from the
\( f \)-sum rule which relates the integration properties of the dielectric function to the plasma frequency \( \omega_P = \sqrt{\frac{4\pi n_e e^2}{m_e}} \) and the crystalline charge density \( \rho \)

\[
\Omega_{GG'}^2(q) = \omega_P^2 \frac{(q + G) \cdot (q + G') \rho(G - G')}{|q + G|^2 \rho(0)}.
\] 

(2.32)

Together with the previous condition

\[
\tilde{\omega}^2_{GG'}(q) = \frac{\Omega_{GG'}^2(q)}{\delta_{GG'} - \epsilon_{GG'}^{-1}(q, \omega = 0)}.
\] 

(2.33)

these relations completely fix the two parameters in the GPP model so that it is ready to be used in the self-energy calculation in Eq. (2.25). Meanwhile, with the GPP model \([\epsilon_{GG'}^{-1}]h\) in Eq. (2.26) can be expressed as

\[
[\epsilon_{GG'}^{-1}]h(q, \omega) = \frac{1}{2} \frac{\Omega_{GG'}^2(q)}{\tilde{\omega}_{GG'}(q)[\omega - \tilde{\omega}_{GG'}(q)]}
\] 

(2.34)

2.3.3 Self-Consistency

In the most widely used GW calculation, the single-particle energies and wave-functions are directly taken from a mean-field calculation, such as DFT-LDA [62], to construct the Green function \( G \) (Eq. (2.18)) and the screened Coulomb interaction \( W \) (Eq. (2.8), (2.9), and (2.20)). But this is not the original GW approximation proposed by Hedin. The full GW approximation require that the Green function \( G \) should represent the motion of the dressed particle and the screening effects, which is embedded in \( W \), is also contributed by dressed particles rather than bare particles.
Typically, this requires further iterations in the GW calculation: keep substituting the band energies $\epsilon_{nk}$ in G and W with the solved ones until self-consistency is achieved. But this iteration procedure is not affordable for most of the materials. A calculation scheme with partial self-consistency could be used instead, in which only the energies in the Green function are updated. This is the so-called $GW_0$ approximation, which saves us from recalculating the time-consuming dielectric function. It is more reliable when there is low-energy plasmon excitations in which one need to align the plasmon resonance profile with the whole band structure. This self-consistency problem will be encountered in Chapter 7.

2.4 Optical Properties and Excitons

The optical properties are strongly reshaped by the excitonic effects. When electrons are excited onto conduction bands, they can bind with the holes left in the valence bands via the attractive Coulomb interaction between opposite charge, leading to the formation of excitons. In general, the resulting excitonic state is not a single $e-h$ pair but a coherent superposition of a number of such pairs. The excitonic state is a two-body excitation which can not be described by the previous theoretical formalism. The Bethe-Salpeter equation (BSE) based on the many-body perturbation theory [63,64] is very powerful tool for quantitatively studying these excitations.
2.4.1 Bethe-Salpeter Equation

Just like the Schrödinger equation, the BSE is also an eigenvalue problem [78] but the “single-particle” state is now a vertical transition from a valence state \( v\mathbf{k} \) to a conduction state \( c\mathbf{k} \).

\[
(E_{c\mathbf{k}} - E_{v\mathbf{k}})A_{v\mathbf{k}}^S + \sum_{v'c'\mathbf{k}} K_{v\mathbf{k},v'c'\mathbf{k}'}^{eh}(\Omega_S)A_{v'c'\mathbf{k}'}^S = \Omega_S A_{v\mathbf{k}}^S. \tag{2.35}
\]

where \( E_{c\mathbf{k}} \) and \( E_{v\mathbf{k}} \) is the quasi-electron and quasi-hole energy, \( K^{eh} \) the \( e\)-\( h \) interaction kernel. \( \Omega_S \) and \( A_{v\mathbf{k}}^S \) are the exciton energy and amplitude of a state \( S \) respectively.

According to the Tamm-Dancoff approximation [78], an exciton state \( S \) is expressed in terms of the linear superposition of \( e\)-\( h \) pair states

\[
|S\rangle = \sum_v \sum_{c} \sum_{\mathbf{k}} A_{v\mathbf{k}}^S |v\mathbf{k}\rangle \tag{2.36}
\]

The rich physics of \( e\)-\( h \) correlations is embedded in the interaction kernel \( K^{eh} \). It can be separated into a direct term \( K^d \) and an exchange term \( K^x \). \( K^d \) describes \( e\)-\( h \) attractive interactions, which is captured by the previously formulated screened Coulomb interactions \( W \). \( K^x \) describes the repulsive exchange energy arising from the cross transition process \( v \to c' \), \( v' \to c \) and is only related to bare Coulomb interactions. It is the origin of the exciton singlet-triplet splitting. For a semiconductor, dynamical effects are unimportant for determining the excitonic effects because the
plasmon frequency is generally much larger than the exciton binding energy. The
above two terms are then represented by

\[
\langle vc\mid K^d\mid v'c'\rangle = -\int dx dx' \psi^*_c(x)\psi_{v'}(x)\psi_v(x')\psi^*_{v'}(x')W(r, r', \omega = 0)
\]
\[
\langle vc\mid K^x\mid v'c'\rangle = \int dx dx' \psi^*_c(x)\psi_{v'}(x)v(r - r')\psi_v(x')\psi^*_{v'}(x')
\]

(2.37)

Including the exchange term \(K^x\) in the BSE or not decide whether the solutions are
singlet or triplet states. From the expression of \(K^x\), one can see the magnitude of
the singlet-triplet splitting of a certain exciton is closely tied to the wavefunction
overlap integral between the involved valence and conduction state, \(v\) and \(c\). If the
most important valence and conduction state have a very weak overlap, then the
singlet-triplet splitting will be very small and \(K^x\) can be omitted in the calculating
the exciton energy.

For a plane-wave basis calculation, the \(e-h\) interaction Kernel can be represented
with the reciprocal lattice vectors \(G\). According to different degree of divergence, the
direct term \(K^d\) is decomposed into three terms

\[
\langle vck\mid K^d_{eh}\mid v'c'k'\rangle = \frac{a_{vck,v'c'k'}}{q^2} + \frac{b_{vck,v'c'k'}}{q} + c_{vck,v'c'k'}
\]

(2.38)

where

\[
a_{vck,v'c'k'} = M_{cc'}(k, q, 0)\epsilon^{-1}_{00}(q)[M_{vv'}(k, q, 0)]^*
\]

(2.39)
\[
b_{vck, v'c'k'} = \sum_{G} \left\{ M_{cc'}(\mathbf{k}, \mathbf{q}, G) \frac{\epsilon_{G0}^{-1}(\mathbf{q})}{|\mathbf{q} + G|} [M_{vv'}(\mathbf{k}, \mathbf{q}, 0)]^{*} \right. \\
+ M_{cc'}(\mathbf{k}, \mathbf{q}, 0) \frac{\epsilon_{G0}^{-1}(\mathbf{q})}{|\mathbf{q} + G|} [M_{vv'}(\mathbf{k}, \mathbf{q}, G)]^{*} \right\} \\
\]

\[
c_{vck, v'c'k'} = \sum_{G \neq 0} \sum_{G' \neq 0} M_{cc'}(\mathbf{k}, \mathbf{q}, G) \frac{\epsilon_{G0}^{-1}(\mathbf{q})}{|\mathbf{q} + G||\mathbf{q} + G'|} [M_{vv'}(\mathbf{k}, \mathbf{q}, G')]^{*} \\
\]

They are called the “head”, “wing”, and “body” term respectively, which reflect the location of the involved dielectric matrix elements relative to the whole dielectric matrix. Likewise, the crystal representation of the exchange term \(K^x\) is

\[
\langle vck|K_{eh}^x|v'c'k'\rangle = \sum_{G \neq 0} \sum_{G' \neq 0} M_{ce}(\mathbf{k}, \mathbf{q}, G) \frac{1}{|\mathbf{q} + G||\mathbf{q} + G'|} [M_{cv'}(\mathbf{k}, \mathbf{q}, G')]^{*} \\
\]

In this case, there is no need to decompose the Kernel. It has no singular behavior near \(q \to 0\) because the wavefunction matrix element \(M_{ce}(\mathbf{k}, \mathbf{q}, 0)\) vanishes as \(\propto q\) and will eventually cancel out the divergence in the Coulomb interaction. The absence of divergence also indicates that \(K_{eh}^x\) is a less prominent contribution compared to \(K_{eh}^d\).

2.4.2 Optical Spectra and Excitonic Wavefunctions

Without excitonic effects, the absorption spectra are given by summing up the oscillator strength from independent vertical transitions

\[
\epsilon_2^0(\omega) = \frac{16\pi e^2}{\omega^2} \sum_{v,c} |\bar{\lambda} \cdot (\langle v|\bar{v}||c\rangle)|^2 \delta(\omega - (E_c - E_v)), \\
\]

28
where the oscillator strength is derived from the well-known Fermi Golden rule. \( \lambda \) points to the light polarization direction and \( \mathbf{v} \) is the single-particle velocity operator. When excitonic effects are taken into account, the coherent superposition of \( e-h \) pairs needs to be considered.

\[
e_2(\omega) = \frac{16\pi e^2}{\omega^2} \sum_S |\mathbf{\lambda} \cdot \langle 0|\mathbf{v}|S\rangle|^2 \delta(\omega - \Omega_S).
\] (2.44)

For 2D materials, in order to obtain an optical absorption spectrum that can be directly compared with experiments, one may convert \( e_2 \) into optical absorbance \( A(\omega) \) via the relation.

\[
A(\omega) = \omega e_2(\omega) d / c
\] (2.45)

where \( d \) is the vacuum distance in the 2D supercell configuration and \( c \) is speed of light. This quantity reflects the fraction of incident photon energy absorbed by a 2D material.

The BSE also yields detailed information about how the electron and hole are spatially correlated that is important for understanding the nature of the exciton. From Eq. (2.36), an exciton wavefunction is expressed in terms of single-particle wavefunctions

\[
\chi(x_e, x_h) = \sum_{\text{occ}} \sum_{\text{empty}} \sum_c A_{vc}^{S} \phi_c(x_e) \phi_v^*(x_h)
\] (2.46)

In general, this exciton wavefunction is a six-dimensional function which is not very intuitive to understand. In practice, it is more convenient to visualize the exci-
ton wavefunction by fixing the hole $x_h$ at its most probable position $x_h^{\text{max}}$ and plot
\[ \chi(x_e, x_h^{\text{max}}). \]

2.4.3 Hydrogen Model of Excitons

The electron orbiting around the hole is very similar to a situation in which a single-particle is travelling near a a scattering, attractive center. This can be reflected by the formalism of BSE. Although the equation is sophisticated at a first glance, for a direct-band-gap semiconductor it can be boiled down to the ordinary Schrödinger equation for an electron in a hydrogen atom, which is analytically solvable.

The simplification procedure of the BSE will be carried out in several steps. First, using the effective-mass approximation to the band edged where excitons are most likely to form, one can replace the kinetic term in the BSE with the quadratic terms

\[
E_{ck} - E_{vk} = \left( \frac{k^2}{2m_e} + E_g \right) - \left( -\frac{k^2}{2m_h} \right) = \frac{k^2}{2\mu} + E_g
\]

(2.47)

where $m_e$ and $m_h$ (positive) are the conduction and valence effective masses respectively, $E_g$ the direct band gap. The reduced effective mass of the $e$-$h$ pair is naturally introduced here: \[ \frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h}. \]

The second step is to approximate the $e$-$h$ interaction Kernel. The momentum-space representation from Eq. (2.39) to (2.42) has provided a convenient starting-point for the approximation. In many crystalline materials, the excitons typically
span over a decent number of primitive cells. For example, the radius of the wannier exciton in bulk semiconductors can reach up to several nanometers. This means the higher-order contributions from shorter wavelength (large $q$, finite $G$), $b_{vck,v'c'k'}$ and $c_{vck,v'c'k'}$, are averaged out on the exciton-radius scale. The most prominent contribution to the $e-h$ binding is from the “head” term (Eq. 2.39) at small $q$, which scales up most rapidly at longer wavelengths $q \to 0$. Neglecting the exchange term for simplicity, we can approximate the Kernel with only the “head” contribution

$$
\langle vck | K_{eh} | v'c'k' \rangle \approx M_{cc'}(k, q, 0)[M_{vv'}(k, q, 0)]^* \frac{\epsilon_{oo}^{-1}(q)}{q^2} \quad (2.48)
$$

Again, for small $q$, we may replace $\epsilon_{oo}^{-1}(q)$ with the inverse dielectric constant $1/\epsilon$. Meanwhile, the wavefunction overlap matrices have well known limits as given by Eq. (2.13). The kernel is further simplified into

$$
\langle vck | K_{eh} | v'c'k' \rangle \approx \delta_{vv'} \delta_{cc'} \frac{1}{\epsilon q^2} \quad (2.49)
$$

This approximation decouples the transitions that are originated from different valence band or conduction band. Thus the full BSE matrix is divide up into many non-interacting subsystems according to the valence and conduction band index. For a specific pair of $v$ and $c$, the approximated BSE now reads

$$
\left( \frac{k^2}{2\mu} + E_g \right) + \int d^3k' \frac{1}{\epsilon |k' - k|^2} A_k = \Omega A_k \quad (2.50)
$$
When Fourier-transformed into real space, it is quite obvious that this becomes the Schrödinger equation of a hydrogen atom with screened Coulomb interaction $1/\varepsilon r$.

The exciton binding energy is then easily given the Rydberg Series

$$E_b = \frac{\mu}{\varepsilon n^2} \quad (2.51)$$

It can be seen clearly how strong the excitonic effect depends on two factors: the reduced effective mass and the dielectric screening. Flatter bands will give large effective masses and strongly bound excitons. A set of parallel bands, such as in silicon and the twisted bilayer graphene that will be discussed in Chapter 4, can also result in a very large reduced effective mass $\mu$ and hence significant excitonic effects.

One the other hand, a stronger dielectric screening will reduce the effective Coulomb interaction and weaken the $e-h$ binding, leading to a depressed excitonic effect.
3. QUASIPARTICLE AND EXCITONIC PROPERTIES OF GATED BILAYER GRAPHENE

3.1 Introduction

Despite its intriguing electronic, thermal and optical properties [4, 5, 12], intrinsic graphene is a gapless semimetal, impeding its utility in bipolar devices, high-performance field-effect transistors and subsequent broad applications. Therefore, huge efforts have been made to overcome this barrier by generating a finite band gap in graphene or its derivatives [11, 55, 56, 79]. One promising approach is to apply the gate electric field perpendicular to the AB (Bernal) stacked bilayer graphene (BLG) to break the inversion symmetry of sublattices [12, 14, 55, 56, 80]. Such an induced band gap of GBLG can be tuned in a wide range by field strength [12–14, 81], offering an important degree of freedom to optimize performance of graphene devices.

A satisfactory understanding of fundamental properties of GBLG is still lacking, such as its QP band gap and optical excitations. For instance, electrical conductance experiments [13, 14, 81] have confirmed the existence of a finite QP band gap but their measured value is disturbed by many extrinsic factors, e.g., the inevitable
contact resistance between the electrodes and graphene sheet. While noncontacting optical measurements \([12, 80, 82]\) have revealed a tunable band gap in GBLG, these results are indirect because the optical absorption peak (edge) is not conceptually equivalent to the QP band gap \([63, 64]\). Particularly, enhanced excitonic effects often dramatically shift the optical absorption peak as we have seen in many other reduced dimensional semiconductors \([30–32]\), making this inconsistency even more serious. Therefore, an accurate calculation with many-electron effects is crucial for settling the above inconsistency.

Conventional density functional theory (DFT) simulations cannot answer the above questions because of their known deficiencies of handling excited-state properties \([62–64]\). Tight-binding models \([83]\) have revealed appealing properties of excitons in GBLG, but it must rely on parameters. In particular, recent \textit{ab initio} GW-BSE simulation has successfully predicted enhanced many-electron effects on intrinsic graphene \([52]\), which are confirmed by subsequent experiments \([51, 84, 85]\). Therefore, a reliable first-principles calculation with many-electron effects included is also promising.

More importantly, beyond providing reliable parameters for device design, learning about the excited states of GBLG and how they evolve with gate field will be of fundamental interest because it will fill our knowledge gap on many-electron interactions in two-dimension (2-D) narrow-gap semiconductors, a field that has not been well understood yet. In fact, it is challenging for first-principles simulations to accurately capture the nearly metallic electronic screening of narrow-gap semiconductors.
For this purpose, an improved algorithm has to be developed and shall be of broad interest for the electronic-structure community.

In this chapter, we employ the modified model to accurately describe the screening and conclude four important remarks about the excited states of GBLG: 1) The QP band gap and its dependence on the gate field are obtained. The self-energy correction is significant because of the enhanced electron-electron (e-e) interactions; the calculated QP band gaps and their tunable range are more than 150% of previous DFT predictions [86, 87], which are beneficial for device applications since a wider band gap implies higher working temperature. 2) Optical absorption spectra of GBLG are dominated by excitonic effects. With electron-hole (e-h) interactions included, our calculated absorption peaks are in excellent agreement with recent experiments [12], explaining the inconsistency between QP gap and optical gap. 3) e-h interactions are so sensitive to the gate field that we can efficiently tune the exciton binding energies and even the order of exciton levels by the gate field. 4) Excitons in GBLG exhibit a number of unusual features. For example, the electron and hole of the lowest-energy dark exciton are completely separated onto opposite layers of graphene, giving rise to an optical approach to polarize BLG. Moreover, this separation of electron and hole offers a neat opportunity to evaluate entangling effects, such as the exchange interaction, of many-electron systems.
3.2 Methods and Computational Settings

To reveal the significance of many-body correlations in GBLG, we calculate the excited states using the following procedure. First, ground-state energy and wavefunctions are obtained by DFT within the local density approximation (LDA). Secondly, the QP energy is calculated within the single-shot $G_0W_0$ approximation [62]. Finally we obtain the exciton energy, wavefunctions and optical absorption spectra by solving the following BSE [63]

$$
(E_{ck} - E_{ek})A_{vck}^S + \sum_{v'c'k'} \langle vck|K^{e\hbar}v'c'k'\rangle A_{v'c'k'}^S = \Omega^S A_{vck}^S
$$

(3.1)

where $A_{vck}^S$ is the amplitude of excitonic state $S$, consisting of single-particle hole state $|v\rangle$ and electron state $|c\rangle$. $K^{e\hbar}$ is the $e-h$ interaction kernel and $\Omega^S$ is the exciton excitation energy. $E_{ck}$ and $E_{ek}$ are QP energy of electrons and holes, respectively.

All calculations are based on a plane-wave basis and norm-conserving pseudopotential approximations with a 60-Ry energy cutoff. To eliminate the spurious interaction between neighboring BLG, the slab-truncation scheme is applied to mimic isolated GBLG [76, 77]. The electric field is applied, via periodic sawtooth potential, perpendicular to graphene layers.

The crucial part of describing many-electron interactions is to obtain the dielectric function. For GBLG with the truncated Coulomb interaction, the inverse static dielectric function $\epsilon^{-1}(q)$ rapidly changes within the long wave-length regime $q \rightarrow 0$, which is similar to what has been noticed in recent first-principles simulations of
carbon nanotubes (CNTs) [31, 32]. A brute-force way to capture this feature is to use a dense q-grid, which demands formidable computational resources. To solve this problem, we deliberately employ the mini Brillouin zone (BZ) sampling scheme to account for this sharply-varying character as motivated by Refs. [76, 77], and use it to both evaluate the QP energies and solve the BSE (see the supplementary material for details). As a result, a $72 \times 72 \times 1$ coarse k-grid sampling is adequate for the GW calculation. In addition, we employ a partial $1440 \times 1440 \times 1$ fine k-grid sampling around the Dirac cone for a dependable BSE calculation.

3.3 Quasiparticle Properties

The LDA and $GW$ band structures near the BZ corner (the K point) are plotted in Fig. 3.1 (a) for GBLG, respectively. The applied gate field induces a finite band gap and changes the band dispersion to the Mexican-hat feature. After including the self-energy correction via the GW calculation, the Mexican-hat-shaped feature remains intact; nevertheless, the fundamental band gap is significantly enlarged due to the depressed screening of isolated GBLG. Moreover, the slope of band dispersion is sharpened by the self-energy correction, implying a smaller effective mass of free carriers.

We also investigate the QP band gap dependence on the applied gate field as shown in Fig. 3.1 (b). The QP band gap can be varied from zero up to 300 meV under experimentally reachable gate field, which is also more than 150% of previous DFT predictions. These features are desired for device applications because a wider
Figure 3.1. (Color online) (a) DFT/LDA and GW calculated band structures around the Dirac point of BLG under a gate field of 2 V/nm. (b) Comparison of the "gap" values obtained from different approaches and their dependence on the gate electric field. The value of the optical gap is defined by the position of the first bright peak of the optical absorption spectrum. The experimental values are extracted from Ref. [12].

gap means a higher working temperature and lower noise. Moreover, when listing the ratio of the self-energy correction to their DFT/LDA value, we see the progression, 56%, 67%, 78%, and 81%, respectively, as the applied field is decreased. This growing trend of the correction ratio for the smaller gap is of particular interest because recent experiments [81] shows a possible small band gap (around a few meV) even for BLG in the absence of gate field. However, due to the limited accuracy of our numerical simulation, we cannot resolve those energy differences below 10 meV and hence a more advanced simulation technique needs to be developed.
3.4 Optical Properties and Absorbance

Additionally, the recent optical measurements of the optical gap are plotted in Fig. 3.1 (b) as well. The key feature is that the QP gap is substantially larger than both previous DFT predictions [86,87] and measurements from optical experiments, [12,88]. The inconsistency of the QP band gap with the optical measurements has also been observed in several other semiconducting nanostructures [30–32], which manifests enhanced excitonic effects and motivates the following calculation on optical spectra of GBLG.

Fig. 3.2 depicts the optical absorption spectrum and its evolution subject to the increasing field magnitude. We first focus on absorption spectra in the absence of the e-h interactions (blue lines in Figs. 3.2 (a)). In the low-energy regime, the absorption is mostly contributed to by the transition from the highest valence band (v1) and the lowest conduction band (c1). As expected, the absorption onset displays a blueshift as the electric field increases the band gap magnitude. Meanwhile, the prominent absorption feature is gradually broadened and split into a double-peak structure ($I_1$ and $I_2$) which stems from the two one-dimensional-like von-Hove singularities [83,89] at opposite “Mexican-hat brims” (Fig. 3.1 (a)), which is consistent with previous DFT results [86].

Surprisingly, the von-Hove singularity at the K point does not contribute greatly to the absorption and therefore is not resolved in the spectra. This is because the relevant valence state $|v_k\rangle$ and the conduction state $|c_k\rangle$ at the Dirac point K are strongly...
localized on different layers upon field-induced symmetry breaking. Therefore, the overlap of wavefunctions is very small and thus leads to a negligible oscillator strength. This can be seen in Fig. 3.3 (a), in which we present the contour plot of the oscillator strength around the corner of the first BZ. The strongest oscillator strength is actually from the “Mexican-hat brims” regime while it is almost zero at the K point.

With e-h interactions included, a different optical absorption picture emerges. As shown in Figs. 3.2 (a), the exciton effect dramatically reshapes the spectra; the broad asymmetric, absorption peak in the single-particle picture is replaced by a symmetric prominent absorption peak. This peak lies below the QP band gap, indicating the existence of bound e-h pairs. The binding energy vary significantly with the gate voltage. They are 35, 54, 76 and 80 meV under four sampling voltages, respectively, fairly close to previous tight-binding calculations [83]. Remarkably, these peak positions are in excellent agreement with the previous infrared microspectroscopy experiment [12] as shown in Fig. 3.1 (b). Under realistic experimental setups, both self-energy corrections and e-h interactions will be reduced by the screening effect of dielectric substrates. Additionall, these reductions may cancel each other more or less [90]. This yields an excellent agreement of our calculations with experimental data.

In the higher energy regime (around 0.4 eV) next to the first optical active peak, the absorbance maintains a constant on the whole (∼ 3%), which is significantly smaller than 4.6%, which is the ideal value of the optical absorbance of BLG [91,92]. This is due to the sum rule of oscillator strength [63] in that e-h interactions drain the absorbance from the high-energy regime to enhance the exciton peak.
Figure 3.2. (Color online) (a) Optical absorption spectra of GBLG. The vertical black dashed line indicates the GW fundamental gap. The incident light is polarized parallel to the graphene plane. A 10 meV Gaussian smearing is applied. (b) Optical activity and eigenenergy of excitons. Each bar represents one exciton state and its height (plotted in the logarithmic scale) indicates the corresponding optical activity. The lowest-energy dark exciton $D$ and the prominent exciton $A$ are particularly outlined by widened dark and red bars, respectively.
It has to be pointed out that electron-phonon coupling is another important factor in determining the infrared optical spectra of GBLG. For example, a G-mode phonon at 195meV has been found to be in Fano interference coupled with e-h excitations in GBLG [93]. Therefore, we may expect this dip feature from such a G-mode phonon may impact the lineshape of our studied exciton absorption peaks.

3.5 Exotic Excitonic States

A close inspection of solutions of the BSE reveals an intriguing exciton picture that has not been observed by experiments. We plot the oscillator strength of excitons in a logarithmic scale in Figs. 3.2 (b). The isolated exciton state with the largest oscillator strength, \( A \), is responsible for the symmetric, prominent absorption peaks in the spectra. Surprisingly, there is one lower-lying excion, \( D \), with a much weaker
oscillator strength for most gated fields (except 4V/nm). This is contrary to the usual
effective-mass model, in which the lowest singlet exciton shall be the brightest one
involved with two bands.

Furthermore, we observe that both the position and oscillator strength of this
dark exciton \(D\) are more sensitive to the gate field than those of bright exciton \(A\).
As plotted in Figs. 3.2 (b), the energy of \(D\) progressively approaches that of \(A\) with
an increasing gate field strength and its optical activity is strongly quenched simult-
aneously. In particular, when the gate field is more than 3V/nm, the order of the
bright and dark excitons is switched as shown in Fig. 3.1 (b). This tunable energy
difference can surely affect the thermal population of exciton states and their lumini-
scence performance. The tunability of the order of exciton energies is in qualitative
agreement with previous tight-binding studies [83].

To understand the brightness of these exciton states, we need to further investigate
the origin of their optical activity. For a typical field strength (2 V/nm), Figs. 3.3
(b) and (c) display the distribution of the square of exciton amplitude \(A^S_{\text{vec}}\) for the
excitons \(A\) and \(D\). Since the optical activity of an exciton \(i\) [63] is

\[
|\langle 0|\vec{v}|i\rangle|^2 = |\sum_{vck} A^i_{vck} \langle v k |\vec{v}| c k \rangle|^2,
\]

which is roughly the product of the single-particle oscillator strength shown in Fig.
3.3 (a) and exciton amplitude shown in Figs. 3.3 (b) or (c), we immediately see
Figure 3.4. (Color online) (a) and (b) Side views of the isosurface plot of the square of wavefunctions of the excitons $D$ and $A$. (c) to (f) Top view of these exciton wavefunctions for top and bottom layers, respectively. The hole position is marked by the open circle in (a) and (b) while it is located at the center of the bottom layer in (d) and (f).

the product of exciton $D$ is much bigger than exciton $A$, suggesting their markedly different brightness.

Fig. 3.4 visualizes the exciton wavefunctions in the real space. As is readily seen, both excitons $A$ and $D$ are strong charge transfer excitons but with distinct characters. In particular, the electron and hole of the dark exciton $D$ almost become disentangled. As shown in Figs. 3.4 (a), (c) and (d), the electron and hole
wavefunctions of exciton $D$ are nearly completely separated into two layers. This is very dissimilar to the $e$-$h$ correlation in other 2-D semiconductors [39, 52]. From the perspective of optoelectronic applications, exciton $D$ could yield the interesting possibility of efficient $e$-$h$ separation and polarize BLG by optical excitations. For the exciton $A$, the degree of $e$-$h$ separation is much lower. In Figs. 3.4 (b), (e) and (f), the electron distributes over a ring on the top layer while on the bottom layer the electron distributes on a disk centered at the hole.

Moreover, these excitonic wavefunctions will be crucial to understand why the dark exciton $D$ and the dark exciton $A$ respond very differently to the electric field. As concluded in Fig. 3.1 (b), the energy level of exciton $D$ exhibits an approximately linear relationship with the field strength, whereas that of exciton $A$ shows a nonlinear behavior. This can be rationalized by the fact that exciton $D$ can be viewed as a plane of dipoles composed of dissociated electron and hole, as revealed in Fig. 3.4 (a), whose energy levels of the positive and negative poles linearly depend on the applied gate field. In contrast, the electron and hole for the exciton $A$ are spatially entangled and therefore the energy level is less sensitive to the gate field and does not follow a simple linear trend. This explains the origin of the energy order switch when the applied gate field is more than 3 V/nm.

3.6 Conclusions

In conclusion, we have provided first-principles calculations for the QP energy and excitonic effects of GBLG. $e$-$e$ and $e$-$h$ interactions are significant and must be
considered to understand the electronic structure and optical excitations of GBLG. Moreover, our calculation clearly explains recent experiments and reveals more of the physics associated with many-electron effects. Finally, we have observed an exotic dark exciton structure that is not likely to present itself in conventional direct band gap semiconductors. The different degree of charge transfer for different exciton states may be useful in optoelectronic applications.
4. STRONGLY BOUND EXCITONS IN TWISTED
BILAYER GRAPHENE

4.1 Introduction

Bound excitons, electron-hole \((e-h)\) pairs, are of particular interest because of their neat physics picture and intrinsic long lifetime that makes broad applications, including photovoltaic and photocatalytic \([30,94-96]\). However, the formation of bound \(e-h\) pairs had been thought to be impossible in metallic (gapless) systems due to their overwhelming screening effects. Moreover, \(e-h\) pairs in gapless structures tend to hybridize with continuous transitions nearby, forming resonant states, whose intrinsic lifetime is substantially shorter. To date, the only exception was found in metallic carbon nanotubes (mCNTs), in which the depressed one-dimensional (1D) screening together with the unique optical symmetry gap lead to the formation of a bound \(e-h\) pair \([31,32,97,98]\). Meanwhile, these studies ignite many obvious but fundamental questions: besides 1D metals, can we observe bound excitons in structures with stronger dielectric screening, e.g., higher dimensional \((d > 2)\) gapless materials?
In addition to the symmetry-related reason revealed in mCNTs, are there any other general mechanisms responsible for bound exciton formation in gapless systems?

Graphene, as a 2D semimetal, may serve as an excellent testbed to answer these outstanding questions. Unfortunately, due to a broad Fano resonance [51, 85, 90, 90, 99], no evidence of bound excitons has been observed, despite the presence of significant e-h interactions in graphene. Recently, twisted bilayer graphene (tBLG) [53,54,100–108], a 2D semimetal, has ignited substantial interest since a twist between graphene sheets introduces new van Hove singularities (vHSs) [53,54,105,108,109] that emerge at the intersections of Dirac cones on opposite layers. From the perspective of excitons, this unique band structure with several vHSs (see Fig. 4.1) has a particular implication for unusual excitonic effects. As shown in Fig. 4.1 (b) and (c), the outlined bands in each schematic are parallel to each other, due to the proximate group velocities of electrons and holes, which lead to a large joint density of states (JDOS). This special band topology enhances e-h interactions and therefore sheds new light on the potential existence of bound e-h pairs in 2D metallic systems.

In this chapter, we predict the existence of strongly bound excitons in higher-dimensional gapless structures by a new decoherent effect, the Ghost Fano resonance. As an example of realistic material, we focus on excitonic effects of tBLG. Through first-principles GW-Bethe Salpeter Equation (BSE) simulations, we successfully observe a bound (though less bright) exciton with a significant binding energy of 0.5 eV in tBLG, which is an order of magnitude larger than that found in mCNTs [31, 32, 96–98] and is even comparable to those in semiconducting nanostructures.
Figure 4.1. (a) Low-energy band structure of tBLG. α (light red) is the plane passing both axes of the Dirac cones whereas β (light blue) is the bisector plane of the two cones. (b)(c) Schematic formation of exciton $X_13(b)$ and $X_24(c)$ illustrated on the α-plane. The energy bands are labeled with 1 to 4 in ascending energy order. The involved bands are outlined in black while the states that mainly compose the exciton are enclosed by the ellipses. (d) Bands plotted on the β-plane. $E_T^z = \hbar v_F |\Delta K|$ is the transition energy gap between 1st (2nd) and 3rd (4th) band.
With the help of the low-energy effective model, we found that the formation of this unusual bound exciton is explained by the ghost Fano resonance \cite{112,113}, a unique destructive coherence between two sets of resonant states with similar energies. This represents a new mechanism for forming bound excitons in gapless systems. In particular, because of its coherent origin, our proposed mechanism gives hope to creating bound excitons in many other metallic systems, despite their strong screening.

4.2 Structural Properties

For a AB-stacking bilayer graphene, if one layer is twisted away from the other around the c-axis of system, a sophisticated moire pattern will emerge, leading to a novel multilayered graphene structure called twisted bilayer graphene. Dependent on the resulting structure is periodic or not, it is called a commensurate structure or an incommensurate structure. The commensurate structures can be viewed as a honeycomb superlattice and it has all the point-group symmetries as graphene except for the inversion operation. Because of their periodicity, they are compatible with the first-principles calculations using plane-wave basis. Fig. 4.2 lists two examples of commensurate structures with the smallest number of atoms in their unit cells.
Figure 4.2. (Color online) Two example of commensurate structures with the rotation angle $\theta = 21.8^\circ$ and $32.2^\circ$. The red parallelograms mark the unit cells in the two commensurate structures.
4.3 Computational Settings

We perform first-principles calculations by employing the many-body Greens function theory for tBLG. We focus on two commensurate structures [15] with 21.8° and 32.2° rotated from the AB-stacking order. Our study begins with a density-functional-theory (DFT) calculation within the local density approximation (LDA) [61] Next, the dielectric function is calculated using the random-phase approximation with a \(30 \times 30 \times 1 \ (18 \times 18 \times 1)\) k-grid [108] over the first Brillouin zone. Meanwhile a slab-Coulomb-truncation scheme [76,77] is also employed. We then obtain the quasi-particle (QP) band energies within the \(G_0W_0\) approximation [62]. The vital step in describing the many-body excitonic effects is to solve the BSE [63]. To ensure a smooth and accurate optical spectrum, we incorporated a fine \(60 \times 60 \times 1 \ (36 \times 36 \times 1)\) k-grid in solving the BSE. Seven (Twelve) valence bands and seven (twelve) conduction bands are included to cover a broad range of the optical absorption spectrum up to 6.0eV.

4.4 Optical Spectra

Both optical spectra with and without e-h interactions are presented in Figs. 2 (a) and (b) with three distinct peaks (marked by \(E_1\), \(E_2\), and \(E_3\) in non-interacting spectra). Our calculation yields an excellent agreement with recent optical conductivity measurements [101,108]; the first two peaks, \(E_1\) and \(E_2\), stem from the two intersections between the Dirac cones from opposite layers, and the third one, \(E_3\), results
Figure 4.3. (Color online) (a)(b) Optical absorbance obtained by the GW+BSE method. The blue dash-dotted curves are the non-interacting spectra while the red solid ones are the spectra with the e-h interactions included. (c)(d) e-h attractive energy $E_a^S$ (blue bars, in arbitrary unit) plotted versus the exciton energy $\Omega^S$ for graphene (c) and tBLG (d) within an identical energy window from 2.2-5.0eV. For references, the absorbance spectra of both structures are also plotted (red dashed curves).

from the perturbed saddle-point vHSSs intrinsic to monolayer graphene [106,108]. We observe enhanced excitonic effects in the absorbance. e-h interactions cause peaks $E_1$ and E2 to redshift by $\sim 0.2$eV for both twist angles.
4.5 New Mechanism for Strongly Bound Excitons in Gapless 2D Structures: Ghost Fano Resonance

The fundamental mechanism forming the corresponding excitonic states in these new prominent peaks in tBLG ($E_1$ and E2), however, may be substantially different from our knowledge learned from usual BLG [52, 114]. At the band intersection between two Dirac cones, only two sets of optical transitions with similar energies are allowed due to the selection rule, as shown in Fig. 4.1(b) and (c), producing double resonance [104, 105]. From the point view of two-particle excitations, the parallel sets of bands give rise to significant JDOS and potentially unusual bound e-h pairs.

The most direct approach to examine whether an excitonic state is bound or resonant is to investigate its wave function in real space. We plot the wave functions of two typical bright excitons, R and S, located around peak $E_1$ (marked in Fig. 4.3(a)). Here, R is the brightest excitonic states around the absorption peak. However, as shown in Figs. 4.4 (a) and (b), the electron is distributed loosely around the hole and even extends beyond our simulation range. These wave functions manifest a signature of resonant states, as observed in graphene [52] and CNTs [31, 32, 97]; the binding feature of excitons is substantially weakened by hybridization with continuous Bloch states that are spatially periodic and extended. In conclusion, these prominent peaks in Fig. 4.3 are dominated by resonant excitons, instead of bound ones.

So far we have focused on the brightest exciton, which often corresponds to the most bound state. However, bound states are not necessarily bright [115]. In order
Figure 4.4. (Color online) $|\chi_S(x_e, x_h)|^2$ of excitons R, S, and A in 21.8° -tBLG plotted on the top layer (a1) (c1) and the bottom layer (a2) (c2) with the hole fixed at the most-probable position on the top sheet. The distribution of the electron is normalized to the maximal probability of the two layers so that it ranges from 0 (blue) to 1 (red). The details within the primitive cell are less important and thus have been smoothened out.
to find possible bound exciton states that are not optically active, we have to, in principle, scan all excitonic states solved by BSE and inspect their real-space wave functions, which is implausible because of a huge number of excitonic states (more than 170,000). Motivated by the fact that e-h interactions of bound excitons are typically more significant than those of resonant ones, we evaluate the e-h attractive energy for a given excitonic state S, by calculating $E^S_a$, expectation value of the e-h interaction kernel $K^\text{eh}$ sandwiched by that state S

$$E^S_a = \langle S|K^\text{eh}|S \rangle = \sum_{veck} (E_{ck} - E_{vk}) |A_{veck}^S|^2 - \Omega^S \quad (4.1)$$

Although $E^S_a$ is not the binding energy, it can be understood as the difference between the excitons “kinetic energy” and its eigenenergy, roughly reflecting the degree of e-h attractions.

Using this e-h attractive energy analysis, an intriguing comparison can be made between monolayer graphene and tBLG. For both cases, we plot the e-h attractive energy spectra ($E^S_a$ versus $\Omega^S$) for all exciton states in Fig. 4.3 (c) and (d). Surprisingly, the $E^S_a$ spectrum of graphene (see the blue-bar plot in Fig. 4.3 (c)) exhibits no distinct features up to 5.0eV, even for the prominent absorption peak at 4.6eV. This indicates that all its excitonic states are broadly resonant [52]; however, the $E^S_a$ spectrum of tBLG (Fig. 4.3 (d)) clearly shows several distinct spikes over a broad energy range, implying the existence of excitonic states with stronger e-h interactions. Following this idea, we select the most bound excitonic state, A (marked by an arrow
in Fig. 4.3(d)), and plot its real-space wave function in Fig. 4.4(c1) and (c2). For this case, we obtain an isotropic distribution with significant localization. For the first time, our calculation predicts the presence of a bound exciton state in tBLG, a 2D gapless material.

More questions are raised regarding exciton A. First, its energy is not at the prominent absorption peak \(E_1\) but approximately 0.38eV below it. Moreover, its optical oscillator strength is weak, roughly one fifth of that of brightest excitonic state R. These are in conflict with the conventional wisdom; the most bound state is usually the most optically active one according to the hydrogen model. Second, since the position of the peak \(E_1\) in the non-interacting spectrum indicates the transition energy \(E^T_G\) between the valence and conduction vHSs, the bound exciton A emerges 0.49eV below the energy \(E^T_G\) in Fig. 4.3(a). Such a surprisingly large binding energy is an order of magnitude larger than that found in mCNTs \([31, 32, 97]\) and it is even comparable to those exciton binding energies of semiconducting nanostructures \([31, 32, 34, 35, 37, 110, 111]\).

Unfortunately, it is challenging to directly analyze the results of our above first-principles simulation. Here, we use a low-energy effective model \([116]\) for simplifying the analysis:

\[
H(k) = \begin{pmatrix}
H_0(k, 0) & T^\dagger \\
T & H_0(k - \Delta K, \theta)
\end{pmatrix}
\] (4.2)
where the intralayer dispersion and the interlayer interaction are respectively

$$H_0(k, \theta) = \hbar v_F \begin{pmatrix} 0 & e^{-i\theta(k_x - ik_y)} \\ e^{i\theta(k_x + ik_y)} & 0 \end{pmatrix}, T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$  \hspace{1cm} (4.3)

The factor $e^{i\theta}$ creates a rotation of the crystal momentum $k$. $H_0$ has been established as a reasonably good approximation for each layer. The matrix $T$ describes the average interlayer interaction between AB and BA stacking order, where $\Delta$ is the interlayer coupling strength. Previous works [116] have shown that the matrix with dominant off-diagonal terms represents a more realistic interlayer interaction than does the all-ones matrix. With this low-energy effective model, we are able to reproduce the energy bands and the wave functions needed for solving the BSE. To proceed, we approximate the screened Coulomb interaction in the direct term $K_{x_{rh}}$ with the 2D Coulomb potential $v_c(q) = \frac{2\pi}{q}$ but drop the exchange term $K_{x_{eh}}$ because of its lessened importance in the graphene-related systems [83]. The solution to the BSE is obtained on a uniform k-grid with approximately 2000 k-points in proximity of the two Dirac cones, so that all the states with energy ranging from $-E_{T_G}$ to $E_{T_G}$ are included in the calculation. As an example, we choose a tBLG with 5°-rotation with interlayer coupling strength $\Delta$ of 130meV.

Following the analysis via Eq. 4.1, we scan the $e$-$h$ attractive energy spectrum obtained by our model BSE calculations. Now, the transition energy gap $E_{T_G}$ is 1.05eV and the energy regime below it will be of primary interest. Interestingly, as displayed in Fig. 4.5(a), a series of discrete excitonic states $X_n(n = 1, 2, \cdots)$ are
found with distinct $e\cdot h$ attractive energies $E_S^a$ alongside a background of resonant excitons (marked by grey bars). The energy spectrum of $X_n$ exhibits a few important characteristics. For instance, with ascending exciton energy, their population becomes denser towards $E_T^G$ whereas $E_S^a$ decreases monotonically. If $E_T^G$ and $E_S^a$ are regarded as a “band gap” and “binding energies”, respectively, they exhibit standard features of bound excitons in a semiconductor. The corresponding wave functions of the lowest few states are plotted in the reciprocal space, as shown in Fig. 4(b), which evidently exhibit features of bound excitonic states. For example, the distribution of $X_1$ is highly analogous to 1s state as solved by the hydrogen model. Also given the fact that $X_1$ possesses the largest $e\cdot h$ attractive energy, it can be inferred that $X_1$ corresponds to exciton A in our first-principles simulations.

To further explain the emergence of bound exciton series, $X_n$, in tBLG, one can analyze the BSE solution given by the low-energy effective model. Surprisingly, it can be shown that each state $X_n$ is only composed of two branches of double-resonant transitions ($1 \rightarrow 3$ and $2 \rightarrow 4$), indicating that the exciton state is free of resonance with the Dirac continuum transition ($2 \rightarrow 3$) occurring at low energies.

The above observation also inspires further investigation into the relation of $X_n$ to the excitonic states solved on the ($1 \rightarrow 3$)- and ($2 \rightarrow 4$)-transition subspaces. For each subspace, we respectively obtain a set of subband bound excitons $X_{13,n}$ (Fig. 4.1(b)) and $X_{24,n}$ (Fig. 4.1(c)), emerging at identical energies $\Omega_{X_{13,n}} = \Omega_{X_{24,n}}$. For each $n$,
Figure 4.5. (Color online) (a) $e$-$h$ attractive energy from the low-energy effective model. The blue bars mark the exciton states with prominent $e$-$h$ attractive energies whereas the gray bars represent the background states of less interest. (b1) (b3) Modulus squared wavefunction of exciton $X_1$, $X_2$, and $X_3$. 
our calculation shown the state $X_n$ are in fact purely antisymmetric superposition of two subspace excitons $X_{13,n}$ and $X_{24,n}$

$$|X_n\rangle = \frac{1}{\sqrt{2}}(|X_{13,n}\rangle - |X_{24,n}\rangle) \quad (4.4)$$

Because of the anti-phase coherence between $|X_{13,n}\rangle$ and $|X_{24,n}\rangle$, the optical oscillator strength of $X_n$ is largely diminished. Additionally, the symmetrically superposed states between $X_{13,n}$ and $X_{24,n}$ contribute to a set of higher-energy states, which are resonant and bright excitons. This insight can be evidenced by Fig. 4.6(a), in which we present the projected density of states (PDOS) of the subband exciton $X_{13,1}$ (or $X_{24,1}$) over the full space $X_f$. Both $X_{13,1}$ and $X_{24,1}$ found near 0.78eV have 50%-overlap with $X_1$ occurring around 0.74eV, which is seen as a single spike in the PDOS. Meanwhile, they overlaps with a number of excitonic states at higher energy (around $\Omega^1 = 0.82$eV), suggesting they have resonant components there. Moreover, although the oscillator strengths of $X_{13,n}$ and $X_{24,n}$ are individually bright, the destructive interference of the two components in exciton A renders its net oscillator strength relatively weak compared to the optically active higher-energy excitons, such as R and S.

The above model calculation provides a surprising picture of excitonic interference as displayed in Fig. 4.6(b), in analogy with the so-called Ghost Fano resonance discovered in the model of quantum dot molecules [112,113]. First, although subband excitons $X_{13,1}$ and $X_{24,1}$ might hybridize with the (1 → 4) and (2 → 3)-transition
Figure 4.6. (Color online) (a) PDOS $|\langle X_{13,1} | X_f \rangle|^2$ ($|\langle X_{24,1} | X_f \rangle|^2$) where $X_f$ spans over the full exciton space. (b) Exciton hybridization diagram in tBLG. The outlined circles represent the excitons formed on either the $(1 \rightarrow 3)$- or $(2 \rightarrow 4)$-transition subspace while the grey ellipses represent the continua. The plus (minus) sign indicates the symmetric (antisymmetric) superposition of exciton states. (c) Coupling scheme of a quantum dot molecule with two leads (left and right) as introduced by [112].
continua, they are also subject to mutual hybridization and are thrown into a symmetric state and an antisymmetric one. In the symmetric state, the coupling of $X_{13,n}$ and $X_{24,n}$ with the two transition continua interfere constructively, broadening into a group of bright excitonic states at higher energies via a conventional Fano resonance. Meanwhile, in the antisymmetric state, the couplings with the two continua cancel each other exactly, resulting in a dark and localized state $X_n$ at lower energy via the so-called Ghost Fano resonance [112,113]. It should be noted that the above model is appropriate for small twist angle, and it may not be fully compatible with the quantitative results of our first-principles simulation, in which the twist angles are large (21.8° and 32.2°). However, the essential physics, such as the double resonance of transitions and the related destructive interference should still play an important role in shaping the strongly bound exciton A, even though the imperfect symmetry of conduction and valence bands could weaken the deconstructive coherent effect, making exciton A not completely dark and not perfectly bound.

4.6 Conclusions

In conclusion, we have demonstrated a novel mechanism for the formation of strongly bound excitons in 2D (semi-) metallic nanostructures via the Ghost Fano effect. We introduce the analysis of the $e\cdot h$ attractive energy for locating these bound excitons in gapless systems and show that they can be efficiently identified in the continuous exciton spectrum, even through the scope of first-principles simulations. A strongly bound exciton with a 0.5eV binding energy is identified in tBLG with a
twisting angle of $21.8^\circ$. Our results also reveal a class of exotic excitations in the graphene related system, paving the path for a new degree of freedom for protecting the $e$-$h$ correlations in solids.

The experimental confirmation of the above bound excitons in tBLG is challenging because of their weak optical activity. Other factors might need to be included to break the selection rules for observing them. Two-photon techniques or applying magnetic field may provide a means to detect them, as what had been done to observe dark excitons in CNTs [117,118]. In particular, the double-resonant picture holds better for tBLG with small twist angles because of the better $e$-$h$ symmetry. Therefore, we expect the lifetime of these bound excitons will become longer as the twist angle is reduced. Despite these, the coherent formation mechanism of bound excitons in gapless systems is conceptually important and it should be possibly observed in many other band structures characterized with double resonance. More importantly, this formation mechanism is a coherent effect that is not strongly affected by the screening and $e$-$h$ interaction strength. Therefore, we expect this phenomenon to be robust in many other metallic systems.
5. QUASIPARTICLE ENERGIES IN TRANSITION
METAL DICHALCOGENIDES

5.1 Introduction

Recently, two-dimensional (2D) semiconducting monolayer and few-layer dichalcogenides have drawn significant interest from researchers in light of the materials’ exciting chemical, electrical, and optical properties [3, 15–17, 119–123]. For example, enhanced spin-orbital coupling and unique optical selection rules make these materials promising for spintronics applications [124–127]. Accordingly, the electronic structure and, in particular, the quasiparticle energy of 2D dichalcogenides have been intensively studied to date. It is of particular interest that first-principles GW calculations have shown that enhanced many-electron effects dictate bandgaps of these 2D semiconductors [47, 50, 128, 129].

Many important properties of semiconductors are not solely decided by the bandgap. For instance, the relative band-edge energies between different semiconductors and corresponding band offsets are of fundamental interest in solid state physics and are indispensable for the design of heterojunction devices [130–132]. Dichalcogenides have
been known catalysts for years [121–123]. Understanding the ways in which quantum
confinement modifies the absolute band-edge energy and associated charge-transfer
processes of chemical reactions in these monolayer or few-layer semiconductors is of
critical importance for their catalytical applications. Recently, substantial advances
have been achieved in obtaining qualitative band offsets [133,134], but there has been
extremely limited progress towards overcoming the bandgap problem and including
enhanced many-electron effects in order to achieve accurate quasiparticle energies in
monolayer dichalcogenides.

In this chapter, we employ the well-established first-principles GW approach to
solve the aforementioned problems. Usually obtaining the absolute band-edge energy
and band offsets of semiconductors requires, at least, two conditions: 1) a reference
energy level and 2) an accurate quasiparticle energy. Because we are considering
isolated samples of these 2D dichalcogenides, it is natural to choose the surrounding
vacuum as the reference energy. The more sophisticated challenge, however, is obtain-
ing the quasiparticle energy. In particular, we must ensure that the absolute energies
are well-converged; this process is significantly more costly than is the bandgap cal-
culation [135,136]. In this vein, approximations that estimate the absolute band-edge
energy but avoid a fully-converged GW calculation have been proposed [137] but their
validity has not yet been verified in dichalcogenides. Given this context, the simple
atomic structures and relatively inexpensive cost of fully-converged GW calculations
for monolayer dichalcogenides make these systems excellent vehicles for obtaining
reliable absolute band-edge energies and, moreover, ratifying the aforementioned approximations.

5.2 Computational Settings

In this study, our calculation provides the quasiparticle energy and corresponding band offsets of monolayer dichalcogenides via the single-shot $G_0W_0$ calculation. The enhanced many-electron interactions in such confined 2D semiconductors significantly enlarge the bandgap and change the absolute band-edge energy accordingly. The absolute band-edge energy and band offsets from the GW calculation are substantially different from those of density functional theory (DFT) and hybrid functional theory (HFT). On the other hand, the types of band alignments from DFT and HFT qualitatively agree with the GW results, meaning DFT and HFT are valuable for band-alignment estimations, especially given their inexpensive simulation costs. Interestingly, we find that the band-gap-center model gives a surprisingly accurate absolute band-edge energy without requiring a converged GW calculation. Ultimately, the absolute quasiparticle band-edge energies and band offsets obtained in this work will be helpful for designing heterojunctions and catalysts comprised of these materials.

We apply the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof (PBE) as the exchange-correlation functional in the DFT calculation [138]. The single-shot $G_0W_0$ calculation is employed to obtain the quasiparticle energy. The spin-orbital coupling is not considered in this study. The plane-wave cutoff for the DFT calculation is set to be 80 Rys. The norm-conserving pseudopotentials [139]
of molybdenum and tungsten include the 4s4p and 5s5p semi-core electrons, respectively. The k-point sampling is 12x12x1 for both DFT and GW calculations. The dielectric-function energy cutoff is set to be 10 Rys and the generalized plasmon-pole model (GPP) is applied to obtain the dynamical screening [62]. A slab Coulomb truncation is applied to mimic the isolated monolayer structure with a vacuum spacing of 23 Å between adjacent layers. All structures are fully relaxed according to the force and stress by the DFT/PBE calculation. Their relaxed lattice constants, listed in Table 5.1, are well consistent with previous results [128, 133].

5.3 Results and Discussions

The general features of the band structures of studied monolayer molybdenum and tungsten dichalcogenides are similar. As an example, we plot the bandstructure of MoS$_2$, which exhibits a direct bandgap, via DFT calculation in Fig. 5.1. As revealed by many other works, there is another local maximum of the valence band at the Γ point, whose energy is very close to the valence band maximum (VBM) at the K point. Interestingly, this local maximum at the Γ point will increase to become the VBM in few-layer MoS$_2$ and thus the overall band structure turns out to possess an indirect bandgap, which dramatically changes the photoluminescence. [16] In order to track the change of this subtle but important change in band structure, we denote the energy difference as $\Delta_v$, which is marked in Fig. 5.1, and list its values in Table 5.1. For most monolayer dichalcogenides, $\Delta_v^{DFT}$ is positive and larger than 20 meV. The only exception is WS$_2$, whose $\Delta_v^{DFT}$ is almost zero. However, the further inclusion of
Figure 5.1. (Color online) DFT-calculated band structure of monolayer MoS2. The top of valence band is set to be zero. The energy difference between the conduction band minimum at the K point and the local minimum at the Σ point is denoted by $\Delta_c$. The energy difference between the valence band maximum at the K point and the local maximum at the Γ point is denoted by $\Delta_v$. 
Table 5.1
Structure and electronic properties of calculated monolayer dichalcogenides: lattice constant $a$, band gap $E_g$, the energy difference $\Delta_v$ and $\Delta_c$ defined in Fig. 5.1. All the energies are in the unit of eV. (The HSE result is read from Reference [133].)

<table>
<thead>
<tr>
<th>Material</th>
<th>$a$ (Å)</th>
<th>$E_g^{DFT}$</th>
<th>$E_g^{HSE}$</th>
<th>$E_g^{GW}$</th>
<th>$\Delta_v^{DFT}$</th>
<th>$\Delta_v^{GW}$</th>
<th>$\Delta_c^{DFT}$</th>
<th>$\Delta_c^{GW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MoS$_2$</td>
<td>3.18</td>
<td>1.69</td>
<td>2.02</td>
<td>2.75</td>
<td>0.02</td>
<td>0.16</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>MoSe$_2$</td>
<td>3.31</td>
<td>1.43</td>
<td>1.72</td>
<td>2.33</td>
<td>0.23</td>
<td>0.34</td>
<td>0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>MoTe$_2$</td>
<td>3.51</td>
<td>1.10</td>
<td>1.28</td>
<td>1.82</td>
<td>0.59</td>
<td>0.83</td>
<td>0.15</td>
<td>0.34</td>
</tr>
<tr>
<td>WS$_2$</td>
<td>3.20</td>
<td>1.78</td>
<td>1.98</td>
<td>2.88</td>
<td>&lt;0.002</td>
<td>0.06</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>WSe$_2$</td>
<td>3.33</td>
<td>1.50</td>
<td>1.63</td>
<td>2.38</td>
<td>0.26</td>
<td>0.34</td>
<td>0.21</td>
<td>0.36</td>
</tr>
<tr>
<td>WTe$_2$</td>
<td>3.52</td>
<td>1.10</td>
<td>1.03</td>
<td>1.77</td>
<td>0.65</td>
<td>0.79</td>
<td>0.45</td>
<td>0.39</td>
</tr>
</tbody>
</table>

The spin-orbital coupling usually increases the VBM at the K point [133], preserving monolayer WS$_2$ as a direct bandgap semiconductor. Moreover, from Table 5.1, we can see that the value of $\Delta_v^{DFT}$ increases as the group-VI element of dichalcogenides changes from S, Se, to Te. Meanwhile, we have marked the energy difference, $\Delta_c$, between the lowest conduction band at the K point, which is the CBM, and that at the $\Sigma$ point. The corresponding data are also listed in Table 5.1. At the DFT level, $\Delta_c$ is around a few hundreds meVs for all studied monolayer dichalcogenides. Having applied the single-shot $G_0W_0$ approach to calculate the quasiparticle energy of those six monolayer dichalcogenides, the results are summarized in Table 5.1. First, the GW correction significantly enlarges the bandgap for all studied dichalcogenides. This enhanced many-electron effect has been widely observed in many reduced dimensional semiconductors as a result of depressed screening and stronger electron-electron (e-e) interactions [31, 32, 35]. Our GW calculated bandgaps are in good agreement with
previous results [47, 128]. Meanwhile, we have listed the bandgaps as calculated by HFT with the HSE functional [140] read from Ref. [133]. It can be seen that the GW bandgaps are significantly larger than those from HFT/HSE.

From Table 5.1, one can see that the direct bandgap feature is preserved for all of our calculated monolayer dichalcogenides: the signs of all energy differences, $\Delta_v$ and $\Delta_c$, remain unchanged after GW corrections. We find that the inclusion of the 4s and 5s semi-core electrons is crucial for preserving the direct bandgap feature in the GW calculation; otherwise, the local minimum of the lowest conduction band at the $\Sigma$ point would be the CBM, resulting in an indirect band gap. Our result is also slightly different from another previous work, in which an indirect bandgap of the WSe$_2$ structure is observed. [128] This difference could be from the spin-orbital coupling.

Beyond the quasiparticle bandgap, we have calculated the absolute band-edge energy relative to the vacuum level [134, 136]. The absolute band-edge energy at the DFT level is referred to the vacuum level that is defined by the potential energy in the vacuum between dichalcogenide monolayers in our supercell arrangement, as shown in Fig. 5.2. Then we preform the GW calculation and superpose the self-energy corrections to the DFT eigenvalues, obtaining the absolute quasiparticle energy relative to the vacuum level. In Fig. 5.2, the final absolute quasiparticle band-edge energies are $E_{cQ}^{QP}$ and $E_{vQ}^{QP}$ for the CBM and VBM, respectively. Unlike the bandgap calculation, the convergence of the absolute quasiparticle band-edge energy with respect to the number of unoccupied conduction states included in the self-energy calcula-
Figure 5.2. (Color online) Schematic illustration of the absolute band energy at the DFT and GW levels, respectively, relative to the vacuum level.
Figure 5.3. (Color online) The convergence of the quasiparticle energy of the CBM and the VBM, respectively, according to the number of conduction band included in the GW calculation.

tion is very slow. For example, we present the convergence of the CBM and VBM of monolayer MoS$_2$ in Fig. 5.3. Although the quasiparticle bandgap is reasonably converged at a value of 2.75 eV after including around 200 conduction bands, the absolute values of CBM and VBM do not reach their converged values until we include around 1500 conduction bands. In order to understand the slow convergence of
the absolute quasiparticle energy, we must examine the details of self-energy in the GW calculation. Usually, the self-energy correction is comprised of two contributions according to their physical origins, the Coulomb-hole (COH) term and the screened-exchange (SEX) term [62]. The aforementioned slow convergence is mainly due to the COH term that involves the summation of an infinite number of conduction bands, in principle [62]. We find that the SEX term also converges slowly, although it is faster than the COH term. Thus we use around 500 conduction bands for the calculation of the static screening and around 1500 conduction bands for the final self-energy calculation.

Finally, our calculated absolute quasiparticle band-edge energies are summarized in Fig. 5.4 (a), in which the DFT results are also listed for reference. The enhanced many-electron interactions in monolayer dichalcogenides substantially changes the absolute band-edge energy from the DFT results. However, the general trend of the evolution of the band-edge energies are similar for both DFT and GW results. For instance, the band-edge energy of MX\textsubscript{2} gradually increases as X varies from S to Te or M varies from Mo to W. A particularly interesting point is that the self-energy corrections modify both valence band and conduction band-edge energies similarly, as seen from Fig. 5.4 (a). This is substantially different from the corrections found by previous HFT studies, in which the corrections mainly affect the VBM [133]. The band alignments in Fig. 5.4 exhibit several unusual features. First, even after the costly GW calculation, except the WSe\textsubscript{2}/WTe\textsubscript{2} interface, the qualitative types of band alignments for these materials from DFT and HFT/HSE have not
Figure 5.4. (Color online) The absolute band-edge energy of calculated monolayer dichalcogenides relative to the vacuum level to the vacuum level. (a) The blue and red dashed lines stand for the DFT/PBE results while the grey-shadow regions stand for the GW results. The water reduction ($H^+/H_2$) and oxidation ($H_2O/O_2$) are marked by the cyan dashed lines, respectively. (b) The absolute band-edge energies by the fully converged GW simulation (grey-shadow regions) and the band-gap-center approximation (solid dark lines).
changed. For example, all of these calculations consistently predict that the interface of MoS\textsubscript{2} and MoSe\textsubscript{2} has a type II (staggered) band alignment. Secondly, the values of GW-calculated band offsets are larger than those from DFT or HFT, mainly due to larger bandgap corrections. Therefore, a sophisticated calculation, such as the GW method, may be necessary in order to obtain the quantitative band offset for heterojunctions of 2D chalcogenides, while DFT or HFT calculation can be convenient when assessing the type of the band alignment or other properties \[141\]. On the other hand, for heterojunctions of our studied 2D chalcogenides with other semiconductors, our calculated absolute band-edge energy shall be crucial to decide the band offsets and even alignments.

Earlier work predicts that monolayer MoS\textsubscript{2} and WS\textsubscript{2} may work for water splitting \[133\]. Hereby we have marked the energy levels for the oxidation and reduction processes of water splitting in Fig. 5.4 (a). The GW calculation yields a similar conclusion although the VBMs are usually lower and the CBMs are generally higher than those of DFT and HFT results.

Previously, in order to avoid the slowly converging absolute band energy, the band-gap-center approximation was proposed to estimate the absolute band-edge energy with the assumption that the self-energy correction shifts both CBM and VBM by similar amounts but in inverse directions \[134, 137\]. Interestingly, we find this model works very well for our studied monolayer dichalcogenides. As shown in Fig. 5.4 (b), the band-gap-center approximation gives nearly the same band-edge energy as the costly direct GW calculation. As discussed above, the reason for this agreement
is that our converged GW calculation yields similar corrections for both the DFT-calculated VBM and CBM. Therefore, this band-gap-center approximation may be particularly useful for estimating the band offset of 2D dichalcogenides because it only requires the quasiparticle bandgap.

One must to be cautious when applying our absolute band-edge energy towards realistic applications. Here we only consider the isolated monolayer structures surrounded by vacuum. However, for realistic conditions, the environmental effects will be extremely important for these ultra-thin layer of semiconductors. For example, the background dielectric response may substantially reduce the self-energy corrections, affecting the band gap and band offset dramatically [135, 142]. Moreover, for photocatalytic processes, such as water splitting, excitonic effects must be included since such processes are driven by optically-excited excitons. In particular, electron-hole \((e-h)\) interactions are known to be enhanced in monolayer chalcogenides and in many other reduced-dimensional semiconductors [114, 128]. Thus \(e-h\) interactions will substantially reduced the energy of the optical absorption edge, making it significantly different from the quasiparticle bandgap. In this sense, more sophisticated calculations including the environment effects and the impact of excitons are desirable, which is a major thrust in the field nanotechnologies as well. However, our calculation serves as a valuable foundation for such studies.
5.4 Conclusions

In conclusion, we have employed a first-principles GW calculation to obtained the quasiparticle band structure and absolute band-edge energy of monolayer dichalcogenides. Our converged GW simulation not only produces the bandgaps but also provides the band offsets of relevant heterojunctions. Both the bandgap and absolute band-edge energy are substantially different from previous DFT and HFT/HSE results. Surprisingly, the band-gap-center model works very well for obtaining the absolute band-edge energy without a fully-converged GW simulation, making it a convenient way to estimate the band offsets and chemical activity of monolayer dichalcogenides.
6. TOWARDS MANY-ELECTRON EFFECTS IN ELECTROSTATICALLY DOPED REDUCED-DIMENSIONAL STRUCTURES

6.1 Introduction

It is computationally challenging to quantitatively obtain the QP properties and excitonic effects for doped materials. For doping density typically involved in experiments, one might have to employ unrealistically dense k-grid to sample the tiny fraction of BZ that is occupied by extra carriers introduced from doping. We can use monolayer MoS$_2$, whose unit cell area is $\sim 8.8 \, \text{Å}^2$, to exemplify this smallness. Even a relatively high doping density of $n = 10^{11}\, \text{cm}^{-2}$ means there is only $\sim 10^{-4}$ extra charge carrier in a single unit cell, which at least requires a $100 \times 100 \times 1$ uniform k-grid to sample the BZ in order to capture the doping effects. Performing *ab initio* many-body calculations on such a k-grid has already challenged the limits of nowadays supercomputers.

When turn to reduced-dimensional structures, the dynamical screening features introduced by doping leads to additional heavy computational effort. First, Dynamical screening effects have to be taken into accounted for correctly understanding the impact of doping on the many-body effects of materials or it will lead a paradox-
ical conclusion as follow. In the static limit, we can infer the screening is roughly proportional to DOS at the fermi level because higher DOS provides more electronic transitions to screen electric field. Also from the expression of static polarizability,

$$P(q, \omega = 0) = \frac{1}{V} \sum_k \frac{n(\epsilon_k) - n(\epsilon_{k+q})}{\epsilon_k - \epsilon_{k+q}}$$  \hspace{1cm} (6.1)$$

we immediately see transitions with smaller energies have larger contribution to screening effects. As the system is doped away from the neutral point, the fermi level comes across the band edge and enters the doped band. This gives arise to new intraband transitions, which should dominate the static screening effects. Since both the DOS in a 1D ($\sim \frac{1}{\sqrt{E}}$) and a 2D ($\sim C$) semiconductor increase sharply at band edges, one will conclude that both screening effects and the properties of excitations are impacted abruptly once the system is doped, regardless of the doping density. The 1D-case is even more catastrophic due to the divergence of DOS at band edges: the smaller the doping density is, the more the system will deviate from the intrinsic one. Of course, these conclusions violate our common sense of perturbation theory. The key to solve the paradox is to consider the dynamical screening effects. Viewed from the full expression of the dynamical polarizability Eq. [6.1], a large frequency $\omega$ in the denominator $\omega + \epsilon_k - \epsilon_{k+q}$ can in fact slight the contribution of the intraband transitions from $\mathbf{k}$ to $\mathbf{k}+\mathbf{q}$ to the screening of the system.

Upon doping, the emergence of low-energy acoustic plasmons also complicates the treatment of dynamical effects of reduced-dimensional structures. In these structures,
the electrons experience more quantum confinement than in bulk systems and cannot move freely in every direction. The restoring Coulomb force for the collective oscillation of electron liquid is significantly weakened due to low dimensionality. As a consequence, a new branch of softened plasmon mode emerges in the low-energy region, whose frequency $\omega$ approaches zero in the zero-wavevector limit $q \to 0$. While this plasmon dispersion relation differs substantially from a 3D one, whose plasmon frequency varies slowly with different wavevectors, it is highly analogous to the acoustic phonon mode in solids and sometimes referred to as “acoustic plasmon” [143, 144].

From previous results on a low-dimensional gas, the acoustic plasmon mode for a 1DEG

$$\omega_p(q) \propto q \sqrt{\frac{N}{m^*}} |\log(q)|$$  \hspace{1cm} (6.2)

and for a 2DEG [59, 60]

$$\omega_p(q) = \sqrt{\frac{2\pi N e^2}{m^*}} q$$  \hspace{1cm} (6.3)

For most reduced-dimension structures, one may estimate that the plasmon-pole energy varies within 1 eV with various crystal momentums. This is far below the prominent spectral feature in dielectric response due to interband transitions, which is typically above 10 eV [62].

While a routine calculation with dense uniform grids on momentum and frequency space demands huge computational resources, effort can be made to specifically deal with the region that is mostly affected by doping. In fact, it can shown that the doping effects are highly concentrated in the small-momentum and low-frequency dielectric

81
response. One can combine the normal first-principles methods with ultra fine grids over the “doped” $q$-$\omega$ region to efficiently calculate the excited-state properties for reduced-dimensional materials. Using this philosophy, Catalin [143,144] have studied the QP doping-renormalized band gap in carbon nanotubes. However, their calculation scheme of the self-energy is still expensive because it relies on a dense k-grid and a huge number of conduction bands, which is even less feasible when extended to a 2D system.

In this chapter, we systematically study the impact of electrical doping on the dielectric response of reduced-dimensional systems and propose efficient calculation methods that combine first-principle simulations and analytical approaches. We consider the extra charge carriers from the view of perturbation theory. For instant, the self-energy in doped systems can be expressed in terms of quantities of intrinsic ones perturbed by doping:

\[
\Sigma = iGW = i(G_{\text{int}} + \delta G)(W_{\text{int}} + \delta W) = i(G_{\text{int}}W_{\text{int}} + \delta GW_{\text{int}} + G_{\text{int}}\delta W + \delta G\delta W)
\]  

where the subscript “int” and “$\delta$” represents the charge neutral point and the variation caused by doping respectively. In section 6.2, we analyze the doping effects on the screened Coulomb interaction $\delta W$ in $q$-space and frequency domain, which lays the foundation for further excited-state calculation. In section 6.2.2, we will em-
ploy the modelled screened Coulomb interaction to actually calculate the self-energy corrections for doped 2D systems.

6.2 Doping-Induced Variation in Screening Effects

Our discussion begins with the variation in screened Coulomb interaction \( \delta W = \delta \epsilon^{-1} v \) as introduced by doping. Since the Coulomb interaction \( v \) remained the same after doping, the goal is to find the variation in dielectric function \( \delta [\epsilon_{GG' q, \omega = 0}]^{-1} \) for crystal structures. In order to quantitatively describe the aforementioned acoustic plasmon, it is essential to perform first-principle simulations to study a realistic 2D systems. Monolayer MoS\(_2\) has been chosen because it has a direct band gap at the K point and excellent isotropy near band edges. The system is of particular interest for potential devices applications and would undoubtedly be an ideal test bed for further analytical treatment. At this stage, only the case of n-doping is considered and the issue of spin-orbital splitting in the valence band is set aside for simplicity. To understand the evolution of the screening properties related to doping level, five electron-doping densities \( N = 0.002, 0.005, 0.01, 0.02, \) and \( 0.03\) electrons/unit cell are examined. Several key parameters are listed in Table 6.1 to assist further analysis. Given the number of spins, \( N_s = 2 \), the number of valleys in monolayer MoS\(_2\), \( N_v = 2 \), and the BZ area \( A = \frac{\sqrt{3}}{2} b_1^2 \), the corresponding fermi wave-vectors are evaluated by \( k_F = \sqrt{\frac{NA}{\pi N_s N_v}} \). The dimensionless Wigner-Seitz radii \( r_s \) are obtained by \( r_s = \left( \frac{1}{\pi n} \right)^{\frac{1}{2}} \frac{m^*}{\kappa} \) [27] with the electron effective mass \( m^* = 0.54 \) and the dielectric constant \( \kappa = 4.2 \) [47].
Table 6.1
Fermi wave-vectors $k_F$ for the doping densities examined (in the unit of $b_1$).

<table>
<thead>
<tr>
<th>N (per unit cell)</th>
<th>0.002</th>
<th>0.005</th>
<th>0.010</th>
<th>0.020</th>
<th>0.030</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$ (per $10^{12}$cm$^2$)</td>
<td>2.28</td>
<td>5.71</td>
<td>11.4</td>
<td>22.8</td>
<td>34.3</td>
</tr>
<tr>
<td>$k_F$</td>
<td>0.012</td>
<td>0.019</td>
<td>0.026</td>
<td>0.037</td>
<td>0.045</td>
</tr>
<tr>
<td>$2k_F$</td>
<td>0.023</td>
<td>0.037</td>
<td>0.052</td>
<td>0.074</td>
<td>0.091</td>
</tr>
<tr>
<td>$r_s$</td>
<td>9.1</td>
<td>5.7</td>
<td>4.1</td>
<td>2.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>

The ground-state calculation is performed within the DFT using the GGA PBE exchange correlation functional [138]. The wavefunctions are expanded in a plane-wave basis with a 80 Ry energy cutoff. The intrinsic and doped atomic structures are relaxed until forces are less than 0.02eV/Å and the unit cell pressure is less than 1kbar. For the highest doping level $N = 0.03$/unit cell, the extra charge carriers cause the lattice constant of MoS$_2$ sheet to dilate by 0.5%, which in turn causes the whole band structure to shift upwards almost rigidly by 300meV relative to the vacuum level except that the band gap is lowered by 50meV. The effect of dilation is similar to tensile biaxial strain [46,145,146] and the band gap remains direct at this doping level.

As below, we demonstrate how fast the convergence of $\delta \epsilon^{-1}$ can be achieved by varying different parameters, followed by the simulation results that lead to a dramatic simplification of treatment on the dynamical screening effects.
6.2.1 The static dielectric function

To obtain accurate occupation number on the conduction band, a $120 \times 120 \times 1$ k-point sampling has been employed over the entire BZ, which gives $\frac{28}{120^2} \approx 0.001944e$/unit cell for the lowest doping level at 0.002e/unit cell. As is shown in Fig. 6.1, the convergence of the static polarizability for modestly small momenta ($q0 = q \leq 0.2b_1$ and $G = G' = 0$) can be achieved by including only a few tens ($\sim 20$) of conduction bands in the calculation, where $b_1$ is the length of the reciprocal lattice vector. Fig. 6.2 compares several leading-order matrix elements for the intrinsic system and the 0.03e-doped one. The inverse dielectric function $\epsilon^{-1}_{GG'}(q, \omega = 0)$ are evaluated within
a sufficiently high energy-cutoff of 15 Ry. The head $\epsilon_{00}^{-1}(q, \omega = 0)$ is largely depressed and exhibit a markedly different asymptotic behavior as $q \to 0$ upon doping. In that limit, the $\epsilon_{00}^{-1}(q, \omega = 0)$ approaches to 1 in a undoped 2D system whereas it vanishes in a doped one. The large discrepancy reflects the strengthened screening introduced by doping in the long-wavelength limit. On the other hand, all the other matrix elements, $\epsilon_{GG'}^{-1}(q, \omega)$ with $G \neq 0$ or $G' \neq 0$, changed by $\sim 0.01$ at most for a wide range of $q$ after doping, which means the change can be safely neglected and only $\delta \epsilon_{00}^{-1}(q, \omega)$ needs to be considered in further treatment of doping effects. This is because the extra charge only plays a dominant role on a long-wavelength scale that is much larger than the unit-cell size.

Also, it can be demonstrated that in MoS$_2$ the dielectric function is highly isotropic for small enough $q$. Fig. 6.3 compares the dielectric functions calculated along two perpendicular directions, which differ by $< 0.001$ for every $q$ examined. Provided this excellent isotropy, it should be sufficient to investigate the dielectric function along one-dimension instead of over a 2D-plane. This leads to a substantial reduction in computational cost and further facilitates the description of the screened Coulomb interaction in the system.

To understand how the static screening varies with increasing doping levels, $\epsilon_{00}^{-1}(q, \omega = 0)$ are calculated and plot in Fig. 6.4. From $q = 0$, $\epsilon_{00}^{-1}(q, \omega = 0)$ rises as $\alpha q$ with the same slope $\alpha$ for all doping densities since the 2DEG polarizability of remains a constant that is independent of $k_F$ for $q \leq 2k_F$ (see Table 6.1 for the actual values). As the momentum $q$ increases and comes across $2k_F$, $\epsilon_{00}^{-1}(q, \omega = 0)$ begins to rise faster.
Figure 6.2. (Color online) Comparison of the dielectric matrix elements for the intrinsic (solid line) and 0.03e-doped system (dashed-dot line). $\epsilon^{-1}_{GG'}(\mathbf{q}, \omega = 0)$ with $\mathbf{q}$ along the (0,1,0)-direction in crystal coordinates are computed and matrix elements with different combinations of $\mathbf{G}$-vectors are plot in different colors.
Figure 6.3. (Color online) Comparison of the dielectric matrix elements $\varepsilon_{00}^{-1}(q, \omega = 0)$ evaluated along $b_1$ (solid curves) and perpendicular to $b_1$ (solid squares).
and merge into the dielectric function for the undoped system, which results from the rapid decay of the 2DEG polarizability for $q > 2k_F$. These results are in consistence with the discontinuity in the Lindhard function of a 2D electron gas. The above discussions indicate that for the static screening the doping effects are dominantly concentrated on the head dielectric function $\epsilon_{00}^{-1}(q, \omega = 0)$ in the small-momentum region, which enables one to efficiently obtain the static dielectric function for doped 2D systems.
6.2.2 The dynamical dielectric function

Another primary concern regarding doping in low-dimensional structures is its impact on the dynamical screening effects. To fully study its effects, dielectric functions $\epsilon_{00}^{-1}(q, \omega)$ are directly calculated within the RPA approximation for a broad frequency range.

A coarse calculation is performed on a $12 \times 12 \times 1$ k-grid and a uniform frequency-grid from 0 to 40eV for yielding an overall spectrum feature. In light of a huge number of interband transitions within this energy regime, more than 100 conduction bands are included to ensure the convergence. Fig. 6.5 contrasts the calculated dielectric functions before and after 0.03 electron is added to the bottom conduction band. Even at this doping level, the high-energy feature ($> 2$eV) yet remains intact and thus one can extend the previous plasmon-pole model for the undoped systems to treat the dynamical effects in this region. Meanwhile, major changes to the screening effects are exclusively concentrated in the lower energy region, which is manifested by the emergence of distinctly new plasmon poles in $\Im[\epsilon_{00}^{-1}(q, \omega)]$, as displayed in the inset of Fig. 6.5(b).

Substantial computational effort can now be focused on the low-energy region where doping effects dominate. In order to obtain a reliable, highly resolved dielectric function, a $120 \times 120 \times 1$ k-grid and a 30meV-broadening is employed whereas the energy region between 0 to 0.6eV is sampled with a 0.01eV-spacing grid. The inclusion of large number of bands is, however, not necessary in this case because intraband
Figure 6.5. (Color online) Real (a) and imaginary (b) of $\varepsilon^{-1}_{00}(q, \omega)$ calculated for the intrinsic (black solid lines) and doped (blue squares) monolayer MoS$_2$ with five representative momenta $q$. The low-energy region of most interest is magnified and showed in each inset.
transitions play a prominent role in this energy region. Also given the fact that the doping effects are effectively damped at large momentum, neither a large number of \( G \) vectors are needed. To testify this fast convergence with respect to the band number and energy cutoff of \( G \) vectors, calculations are done with two schemes with different combinations of parameters. In scheme 1 (2), 26 (4) conduction bands and a energy cutoff of 8 Ry(2 Ry) are used. Fig. 6.6 demonstrates that although scheme 2 is much more computationally inexpensive, it still reproduces very close results, particularly the pole positions, as scheme 1. Therefore, one can employ scheme 2 to obtain the low-energy dielectric response for all the doping levels with low cost. For doping density \( n = 0.03e/\text{unit cell} \), the calculated dielectric functions with different \( q \) are showed in Fig. 6.7 and the imaginary part can be directly compared with the corresponding electron energy loss spectroscopy.

The plasmon dispersion relation at different doping levels are summarized in Fig. 6.8. With increasing doping density, the blue-shift of the plasmon frequency \( \tilde{\omega}_d(q) \) is clearly observed in the simulation. At a fixed doping density, the plasmon frequency increases monotonically with larger momentum \( q \) and converges to \( \sim 0.5\text{eV} \) at \( q = 0.2b_1 \). It has to be pointed out that this plasmon dispersion relation apparently deviates from the “\( \sqrt{q} \)-law” for the perfect 2DEG given by Eq. (6.3). Although at long wavelength the plasmon energies obtained from the \textit{ab initio} calculation are roughly proportional to \( \sqrt{q} \), they are still considerably lower than the prediction from Eq. (6.3), as is shown in Fig. 6.8(b). There are multiple reasons that can cause the deviation. First, Eq. (6.3) is derived from a classical treatment of electron gas and does
Figure 6.6. (Color online) Fast convergence of the dielectric function with respect to band number and energy cutoff. Scheme 2 yields highly similar \( \Im[\epsilon_{\omega}(q, \omega)] \) as the scheme 1, which is much more costly.
Figure 6.7. (Color online) Real and imaginary part of the RPA dielectric function calculated with different momentum $q$ at doping density $n = 0.03e$/unit cell.
not account for the fast damping of the acoustic plasmon mode into the e-h continuum. Instead, one has to use a fully quantum-mechanical formula to find the correct plasmon energy at shorter wavelength \([27, 60, 147, 148]\). The plasmon in monolayer MoS\(_2\) has been studied using a effective two-band model [149], which clearly evidences the modification of plasmon dispersion by the intraband e-h excitations. Moreover, the interband transitions arising from crystal structures, in particular those near the band edges, also strengthen the screening effects. This can in turn lower the energy of the acoustic plasmon mode, which has been revealed by previous theoretical [150–152] and experimental [153, 154] work on 2D electron systems. The above dynamical calculation has taken the crystal effects and all possible transitions into account and should reproduce an accurate description of the acoustic plasmon modes introduced by doping.

Provided the simple-pole nature of the acoustic plasmon, one can extend the previous plasmon-pole approach to model the low-energy dynamical effects. Hence the variation in the dielectric function due to doping can be described using a form of response function

\[
\delta \epsilon_{oo}^{-1}(q, \omega) = \frac{\Omega_{d}^2(q)}{\omega^2 - \tilde{\omega}_{d}^2(q)}
\]

(6.6)

where the parameter \(\Omega_{d}(q)\) and \(\tilde{\omega}_{d}(q)\) are the plasmon pole strength and position as functions of momentum \(q\) respectively. Here, only the variation in the head term needs to be modelled because all other \(\delta \epsilon_{GG}^{-1}(q, \omega)\) is negligible, as previously discussed. To determine the two parameters, one can first find the differ-
Figure 6.8. (Color online) (a) Acoustic plasmon energies determined from the *ab initio* RPA dielectric function. (b) Comparison of the low-energy region between the *ab initio* calculation and the classical electron gas model.
ence of the static dielectric functions of the undoped and doped system, which gives
\[ -\frac{\omega_d(q)}{\tilde{\omega}_d(q)} = \delta \epsilon_{00}^{-1}(q, 0) = \epsilon_{00}^{-1}(q, 0) - \epsilon_{\text{int,00}}^{-1}(q, 0). \]
Meanwhile, the pole position \( \tilde{\omega}_d(q) \) is determined from the imaginary part of \( \epsilon_{00}^{-1}(q, 0) \), which can be obtained by a fast full-frequency calculation with undemanding parameter settings. In Fig. 6.9, it is demonstrated that the proposed plasmon-pole model is able to reproduce the dielectric functions from the full-frequency calculation with high fidelity once the static limit \( \delta \epsilon_{GG'}^{-1}(q, \omega = 0) \) and the pole-position \( \tilde{\omega}_d(q) \) are known. One may expect this model greatly facilitates the numerical treatment of the dynamical effects by avoiding employing a very dense frequency grid to capture the rapidly varying dielectric function.

In summary, it has been shown in this section that the dynamical effects of electrostatically doped reduced-dimensional materials can be modelled accurately with modest computational effort. Upon doping, only the head part of the dielectric function, \( \epsilon_{00}^{-1}(q, \omega) \), undergoes significant variation in the small-momentum and low-frequency region whereas other components remain almost unaltered. The dynamical effects due to the acoustic plasmon can be described satisfactorily with a plasmon-pole model, which makes the numerical calculation convenient and efficient.
Figure 6.9. (Color online) Comparison of real(a) and imaginary(b) part of the dielectric functions obtained from a full-frequency calculation and the proposed plasmon-pole model (PPM) at $N = 0.03e/\text{unit cell}$. 
6.3 Quasiparticle Properties: $G_0W_0$ Approximation

For a doped structure, the self-energy operator can be decomposed into the intrinsic one and the perturbation introduced by the doping effects [143, 144, 155], and further grouped into three major contributions:

$$\Sigma = iGW = i(G_{\text{int}} + \delta G)(W_{\text{int}} + \delta W)$$

$$= i(G_{\text{int}}W_{\text{int}} + \delta GW_{\text{int}} + G\delta W)$$

$$= \Sigma_1 + \Sigma_2 + \Sigma_3$$

(6.7)

where $\Sigma_1 = iG_{\text{int}}W_{\text{int}}$ is the self-energy of the intrinsic system, $\Sigma_2 = i\delta GW_{\text{int}}$ arises from purely the change of occupation number on the doped band with intrinsic screening, and $\Sigma_3 = iG\delta W = i(G_{\text{int}} + \delta G)\delta W$ is contributed by the change in the screening effects induced by doping. While the first two terms only involve the intrinsic screening that is linearly varying in the energy range of interest, the third term exhibits strong nonlinear frequency-dependence due to the emerging acoustic plasmon. Thus these contributions of self-energy will be evaluated in distinctly different approaches.

The evaluation of $\Sigma_1 = iG_{\text{int}}W_{\text{int}}$ is straightforward. However, particular attention should be given to its slow convergence in a reduced-dimensional structure since the solutions to the Dyson equation depend sensitively on the absolute value of the total self energy $\Sigma$, which is nonlinear in the doped situation. $\Sigma_1$ is a primary background contribution to $\Sigma$ and is normally much large than the size of band gap. Therefore a large number of empty states have to be included. As for monolayer MoS$_2$, 500 and
1500 conduction bands have been used to achieve convergence in the static screening and self energy calculation. On the other hand, the first-principles calculations show that the variation in $\Sigma_1$ caused by the lattice dilation arising from doping is of a few tens of meV, which is tiny.

The second term $\Sigma_2 = i\delta GW_{\text{int}}$ can be directly obtained by subtracting the total self energy of a undoped system from that of a doped one, both of which are evaluated with the intrinsic dielectric function $\epsilon_{\text{int}}^{-1}$. Because in the self-energy decomposition only the screened-exchange term involves the summation over the occupied states, the contribution of $\Sigma_2$ to a specified state $|nk\rangle$ reads

$$
\langle nk | \Sigma_2(E) | nk\rangle = - \sum_{q,G,G'} f_{n_1k-q} M^*_{cn}(k, -q, -G) M_{cn}(k, -q, -G') \\
\times \epsilon^{-1}_{\text{int},GG'}(q, E - \epsilon_{n_1k-q}) v(q + G')
$$

(6.8)

where $c$ is the doped band index and $f_{nk}$ is the occupation number of the state $|nk\rangle$. Since the doped region is a small fraction of the BZ, the calculation has to be performed a sufficiently dense k-grid. For n-doped monolayer MoS$_2$, a $36 \times 36 \times 1$ k-grid has been employed. $\langle \text{VBM} | \Sigma_2 | \text{VBM}\rangle$ and $\langle \text{CBM} | \Sigma_2 | \text{CBM}\rangle$ are evaluated with on-shell energies at discrete doping densities and are displayed in Fig. 6.10. While $\langle \text{VBM} | \Sigma_2 | \text{VBM}\rangle$ almost remains a constant at various doping levels, $\langle \text{CBM} | \Sigma_2 | \text{CBM}\rangle$ drops considerably up to several hundred meV because adding electrons (holes) to the doped band significantly change the amount of exchange energy. This could lead to a major contribution to band gap narrowing. To find the $\Sigma_2$ at
other doping levels, one can use a simple power law function $\alpha N^\beta$, where $N$ is the doping density, to fit the first-principles results (the solid curve in Fig. 6.10). At light doping, one may even just keep the head dielectric function in the summation of Eq. (6.8) for a state $|n\mathbf{k}\rangle$ near the band edges

$$
\langle n\mathbf{k} | \Sigma_2(E) | n\mathbf{k}\rangle \simeq - \sum_{\mathbf{q}} f_{\mathbf{k}-\mathbf{q}} |M_{\alpha\alpha}(\mathbf{k}, -\mathbf{q}, 0)|^2 \epsilon_{\text{int},00}(\mathbf{q}, E - \epsilon_{\mathbf{k}-\mathbf{q}}) v(\mathbf{q}) \quad (6.9)
$$

because local field effects are negligible in the long-wavelength limit $q \to 0$. Together with other approximations used for mini-BZ averaging, Eq. (6.8) for CBM in a n-doped system eventually boils down to an expression with a clear physical meaning

$$
\langle CBM | \Sigma_2 | CBM\rangle \simeq - \int_{\text{doped}} \frac{d^2\mathbf{q}}{(2\pi)^2} \epsilon_{\text{int},00}^{-1}(\mathbf{q}, 0) v_{2D}(\mathbf{q}) \quad (6.10)
$$

which is simply the integration of the 2D screened Coulomb interaction over the doped region. For the tiny doped region that is difficult to sample, Eq. (6.10) is an efficacious estimation to $\Sigma_2$ and it yields excellent agreement with the full first-principles calculation for $N \to 0$ (the dashed curve in Fig. 6.10).
Figure 6.10. (Color online) Variation of $\langle \text{VBM} | \Sigma_2 | \text{VBM} \rangle$ and $\langle \text{CBM} | \Sigma_2 | \text{CBM} \rangle$ with doping density. The black squares and red circles denote the values from first-principles calculation for VBM and CBM respectively. The solid curve is a fit for the drop in CBM with the power law $\alpha n_d^\beta$ and the dashed curve is estimated using the head of the dielectric function only.
The third term $\Sigma_3 = iG\delta W$ comes from the contribution of both the screened-exchange (SX) and Coulomb-hole (CH) part. The full expression reads

\[
\langle n|\Sigma_3, SX (E) | n \rangle = - \sum_{n_1} \sum_{q, G, G'} f_{n_1 k - q} M^*_{n_1 n}(k, -q, G) M_{n_1 n}(k, -q, G')
\times \delta^{-1}_{GG'}(q, E - \epsilon_{n_1 k - q}) v(q + G')
\]

(6.11)

\[
\langle n|\Sigma_3, CH (E) | n \rangle = \sum_{n_1} \sum_{q, G, G'} M^*_{n_1 n}(k, -q, G) M_{n_1 n}(k, -q, G')
\times [\delta^{-1}_{GG'}]^{h}(q, E - \epsilon_{n_1 k - q}) v(q + G')
\]

Since doping effects are exclusively concentrated on the head component at small $q_c$, the above summation can be approximated into

\[
\langle n|\Sigma_3, SX (E) | n \rangle \simeq - \sum_{n_1} \int_{q < q_c} \frac{d^2 q}{(2\pi)^2} f_{n_1 k - q} |M_{n_1 n}(k, -q, 0)|^2 \delta^{-1}_{00}(q, E - \epsilon_{n_1 k - q}) v_{2D}(q)
\]

(6.12)

\[
\langle n|\Sigma_3, CH (E) | n \rangle \simeq \sum_{n_1} \int_{q < q_c} \frac{d^2 q}{(2\pi)^2} |M_{n_1 n}(k, -q, 0)|^2 [\delta^{-1}_{00}]^{h}(q, E - \epsilon_{n_1 k - q}) v_{2D}(q)
\]

where $q_c$ is the momentum cutoff for doing the BZ summation. In our case of doped monolayer MoS$_2$, excellent convergence can be achieved by choosing $q_c = 0.2b_1$ (see Fig. 6.4).

$\Sigma_3$ involves the rapidly varying dielectric function due to the low-energy acoustic plasmon and has to be evaluated separately from $\Sigma_1$ and $\Sigma_2$ on a ultra dense grid.
Thanks to the plasmon-pole model proposed by Eq. (6.6), the numerical treatment is made simple and efficient. Substituting the $\delta \epsilon^{-1}_{00}$ with the model in Eq. (6.12) gives

$$\delta \epsilon^{-1}_{00}(q,E - \epsilon_{n_1 k - q}) = \frac{\Omega^2_d(q)}{(E - \epsilon_{n_1 k - q})^2 - \tilde{\omega}^2_d(q)}$$

and

$$[\delta \epsilon^{-1}_{00}]^h(q,E - \epsilon_{n_1 k - q}) = \frac{\Omega^2_d(q)}{2\tilde{\omega}_d(q)[E - \epsilon_{n_1 k - q} - \tilde{\omega}_d(q)]}$$

where the pole energy $\tilde{\omega}^2_d(q)$ is now as small as a few hundred meV in the doped case. As $E$ approaches a certain energy band $n_1$, both $\delta \epsilon^{-1}_{00}$ and $[\delta \epsilon^{-1}_{00}]^h$ becomes significant and could possibly contribute a spectral feature to $\langle n k | \Sigma_3(E) | n k \rangle$ near the energy range of $n_1$. Given the summation over $n_1$, $\langle n k | \Sigma_3(E) | n k \rangle$ should in principle exhibit a rich spectral feature with satellites widely distributed through the entire range of the band structure. However, the contribution from each band $n_1$ is also weighted by the square of wavefunction overlap integral between that band and the band of interest, $|M_{n_1 n}(k,-q,0)|^2$, where $M_{n_1 n}(k,-q,0) = \delta_{n_1 n} + O(q)$ at small $q$. For light doping in which a small momentum $q_c$ suffices for convergence, the contribution from the diagonal term with $n_1 = n$ will dominate the band summation in Eq. (6.12). In contrast, the contributions from the off-diagonal terms with $n_1 \neq n$ are of high order and should be much smaller than diagonal term. Moreover, in a semiconductor with a sizeable band gap, these off-diagonal terms will contribute minor spectral feature far away from the energy range of interest and are thus less relevant for solving the Dyson equation. Therefore, only the diagonal term will be preserved hereafter.
For the n-doped case, the original valence bands are fully filled and the SX and CH term can be combined into a simple expression, which reads

\[
\langle v k | \Sigma_3(E) | v k \rangle \\
\simeq \int_{q < q_c} \frac{d^2q}{(2\pi)^2} \left| M_{vv}(k,-q,0) \right|^2 \frac{\Omega_d^2(q)}{2\omega_d(q)\left[ E - \epsilon_{v,k} - \omega_d(q) \right]} v_{2D}(q)
\]

(6.14)

The treatment of SX and CH term in the conduction band need to be separated because the integration limits involved are different. The CH term of a conduction state \(|c_k\rangle\), regardless whether it is occupied or not after doping, is given by

\[
\langle c k | \Sigma_{3,CH}(E) | c k \rangle \\
\simeq \int_{q < q_c} \frac{d^2q}{(2\pi)^2} \left| M_{cc}(k,-q,0) \right|^2 \frac{\Omega_d^2(q)}{2\omega_d(q)\left[ E - \epsilon_{c,k} - \omega_d(q) \right]} v_{2D}(q)
\]

(6.15)

The above terms are solely from the extra screening introduced by doping without changing the occupation number, corresponding to the contribution from \(iG_{\text{int}}\delta W\).

For the bottom conduction band that is partially occupied, the SX term is evaluated with a different momentum cutoff at the fermi wavevector \(k_F\)

\[
\langle c k | \Sigma_{3,SX}(E) | c k \rangle \\
\simeq - \int_{q < k_F} \frac{d^2q}{(2\pi)^2} \left| M_{cc}(k,-q,0) \right|^2 \frac{\Omega_d^2(q)}{(E - \epsilon_{c,k} - \omega_d(q))^2} v_{2D}(q)
\]

(6.16)

Since \(k_F\) and \(\delta\epsilon\) vanishes in the light doping limits, the above SX term corresponds to the second-order contribution \(i\delta G\delta W\) to the self energy.
The two parameters in the model, $\Omega_d(q)$ and $\tilde{\omega}_d(q)(>0)$, evolves smoothly with $q$ and are easy to interpolate and extend to the above numerical integration. The self-energy contributions from $\Sigma_3$ given by Eq. (6.14)-(6.16) are evaluated for VBM and CBM at all the doping levels being investigated, and are displayed in Fig. 6.11, where $\omega$ is the relative energy to band edges in a mean-field theory ($E - \epsilon_{\text{VBM}}$ or $E - \epsilon_{\text{CBM}}$ respectively). Thus $\omega = 0$ defines the on-shell condition and is set to be the center of every self-energy plot. For VBM (Fig. 6.11 (a)), the fluctuation at negative $\omega$ results from the singularities occurring at $E = \epsilon_{\text{k-q}} - \tilde{\omega}_d(q) < E_{\text{VBM}} - \tilde{\omega}_d(q)$ in the integrand of Eq. (6.14). This fluctuation signals the fano resonance of a hole quasiparticle with the dispersive acoustic plasmon branch, which occurs below the single-particle (on-shell) energy. On the contrary, the fluctuation of the CH term in CBM emerges at positive $\omega$ (Fig. 6.11 (b)) due to the singularities occurring at $E = \epsilon_{\text{k+q}} + \tilde{\omega}_d(q) > E_{\text{CBM}} + \tilde{\omega}_d(q)$. In contrast to the previous situation, this fluctuation reflects the fano resonance of a electron quasiparticle with the acoustic plasmon branch, which occurs above the single-particle energy, as is what happens in the p-doped case where the conduction band is totally unoccupied.

The above two self-energy contributions arising from extra screening $\delta W$ can contribute to band gap narrowing, as the self energy in the valence (conduction) band is positive (negative) definite above (below) the single-particle energies. On the other hand, it should be noted that the resonance with the plasmon excitations also causes the self energy in the valence (conduction) band to change sign below (above) the single-particle energies, leading to reverse contribution to the band gap renormaliza-
tion; As screening is strengthened by doping, the renormalized band gap actually becomes even larger. Such an energy region is termed as “anomalous” region, versus the aforementioned “normal” region on the other end of the energy axis, both of which are marked in Fig. 6.11(a) and (b).

In an n-doped case, significant contribution to the self energy of CBM also comes from the SX term, which is a second-order residue but becomes important at a decent doping level. As is shown in Fig. 6.11(c), energies are largely raised near \( \omega = 0 \), resulting in a slightly asymmetric plateau region bounded by a pair of dips on each side. The fluctuations at the dips are particularly strong because the fermi level \( k_F \) coincides with the singularities at \( E = \epsilon_{k-q} \pm \tilde{\omega}_d(q) \). To some extent, the SX term resembles the pure exchange energy in a electron gas as it exhibits energy-independent plateau near the single-particle energy \( (\omega = 0) \). Combining the SX term with the CH term in CBM gives the full contribution of \( \Sigma_3 \) to CBM, which has an energy-dependence resembling to VBM except for a difference in positive \( \omega \) featured by a partially cancelled dip. This resemblance stems from the fact that both the VBM and CBM are now more close to a hole quasiparticle state after considerable n-doping.

With all the three terms summed up, one can solve the Dyson equation to find the QP energies in a doped 2D system

\[
E = \epsilon_{nk} - V_{nk}^{xc} + \Sigma_{nk}(E)
\]  

(6.17)
Figure 6.11. (Color online) Real part of the self-energy contributions from the term $\Sigma_3 = iG\delta W$ for VBM (a) and CBM (b)-(d) in monolayer MoS$_2$. Results at five doping levels are compared. $iG\delta W$ for VBM in (a) is roughly inverse to the CH part of $iG\delta W$ for CBM in (b). Near the plasmon resonance region, the signs of the two self-energy contributions in (a) and (b) flip, which divides whole energy range into the so-called normal region and anomalous region. (c) shows the second-order residue $i\delta G\delta W$ of the contributions to the self-energy upon doping. (d) shows the total self-energy contributions from $iG\delta W$, which is a summation of (b) and (c).
where $V_{nk}^{xc} = \langle nk|V^{xc}|nk \rangle$ is the exchange correlation potential from a mean-field theory and $\Sigma_{nk}(E) = \langle nk|\Sigma(E)|nk \rangle = \langle nk|\Sigma_1 + \Sigma_2 + \Sigma_3|(E)|nk \rangle$ is the total self-energy. The corresponding spectral function $A(\omega)$ is

$$A_{nk}(E) = \frac{1}{\pi} \frac{|\text{Im}\Sigma_{nk}(E)|}{(E - \epsilon_{nk} + V_{nk}^{xc} - \text{Re}\Sigma_{nk}(E))^2 + (\text{Im}\Sigma_{nk}(E))^2} \quad (6.18)$$

Fig. 6.12 demonstrates the graphical solutions to Dyson equation and the spectral functions for VBM and CBM. While the solution in VBM is unique, there are three solutions in CBM due to the large fluctuation in its self energy. The situation that there are multiple solutions to the Dyson equation indicates the strong interaction of a QP state with a plasmon, which results in a satellite structure rather than a single QP peak in the spectral function. Both the solution in VBM and the highest-energy solution in CBM gain the majority of the spectral weight, which represents the standard sharply defined QPs at the corresponding energies. The solution in the middle for CBM is obviously spurious because the large positive slope there gives rise to a negative QP renormalization factor that is not physical. Neither is the solution reflected as a peak in the spectral function. The lowest-energy solution in CBM has a relatively weak spectral weight and is seen as a smaller and broader peak in the spectral function compared with the QP solutions. This solution locates in the fluctuating region where the coupling of the quasiparticle to the acoustic plasmon is prominent. This might suggest the emergence of possible plasmaron excitations. However, the GW approximation currently employed can not accurately account for
the electron correlations near the plasmon satellites. In principle, a more accurate theory, GW plus cumulant approximation [156–160], should be employed to yield a correct description to the hole-plasmon coupling in this situation. Thus the plasmaron peak at lower energy could be averaged out from the spectral function; nevertheless, the QP solutions remain at nearly the same energy, although with a diminished spectral weight. Therefore, the QP peak positions from the GW approximation will be used to determine the size of band gap.

With increasing doping level, the band gap of a 2D structure exhibits an unusual non-monotonic scaling behavior arising from the interplay of band filling effect and the carrier-plasmon coupling. As is marked by black solid squares in Fig. 6.13, the band gap shrinks dramatically at small doping density but surprisingly it begins to open up again slowly after a certain doping level is reached. The leading reason that is responsible for the band gap narrowing for at all doping levels is a change in the exchange energy due to filling extra charge carriers on the doped band, that corresponds to the \( \Sigma_2 = i\delta GW_{\text{int}} \) in the self-energy operator. For the n-doped case, the extra exchange energy in the conduction band edge is negative and its magnitude is proportional to the Fermi wavevector \( k_F \), which is analogous to the case of a electron gas. If only this contribution were taken into account for the QP renormalization, the band gap would decrease monotonically with increasing doping level, as illustrated by the dashed blue line in Fig. 6.13. For the p-doped case, removing electrons from the top of valence band results in a less amount of negative exchange energy and hence a similar trend of contribution to band gap narrowing.
Figure 6.12. ((Color online) (a)(b) Graphic solutions to the Dyson equation $E = \epsilon_{nk} - V_{xc}^{nk} + \Sigma_{nk}(E)$ for VBM and CBM for $n_d = 3.43 \times 10^{13} \text{cm}^{-2}$. The red, dashed blue, and green curves denote the right-hand side of the equation with $\Sigma = \Sigma_{\text{int}}$, $\Sigma_{\text{int}} + i\delta GW_{\text{int}}$, and $\Sigma_{\text{int}} + i\delta GW_{\text{int}} + iG\delta W$. While $i\delta GW_{\text{int}}$ contribute little in VBM, it dramatically lowers the self energy in CBM. The black thick circles mark the QP solutions where the blue circle in (b) marks the potential plasmaron solution in CBM. The red cross marks the spurious solution at the “wrong” slope region. (c)(d) Corresponding spectral functions for VBM and CBM. The solutions are linked to the spectral peak positions by long dashed lines (QP) and dashed dot lines (plasmaron). The DFT-LDA energies are also indicated by grey squares for references.
Figure 6.13. (Color online) Band gap renormalization at various doping densities. The band gaps from the VBM to the original CBM derived from the QP peak in the spectral function are marked with solid black squares. For references, the band gaps derived from the VBM QP solution to the CBM plasmaron solutions are marked with empty squares. The virtual band gaps solved from the self-energy without extra screening $iG\delta W$ are indicated by the dashed blue curve, which drops monotonically with increasing $n_d$. 

\[ n_d \left(10^{13} \text{cm}^2\right) \]

\[ \text{Doping Density } n_d \text{ (electron/unit cell)} \]

\[ \begin{align*} 
2.7 & \quad \bullet \quad \text{VBM QP-CBM QP} \\
2.6 & \quad \square \quad \text{VBM QP-CBM Plasmaron} \\
2.5 & \quad \cdots \quad \text{w/o } iG\delta W \\
2.4 & \\
2.3 & \\
2.2 & \\
2.1 & \\
0.00 & 0.01 \quad \square \quad 0.02 \quad \square \quad 0.03 \quad \square \end{align*} \]
Countering to one’s intuition, the extra screening effect due to doping does not result in a smaller band gap and has a complex influence on the band gap renormalization dependent on doping level. In an undoped semiconductor, the QP band edges typically lie outside the single-particle band gap with similar magnitude of self-energy corrections for the valence band and conduction band, but in opposite directions, as is true in monolayer MoS$_2$. At a light doping density, the QP valence (conduction) band edge remains well below (above) its single-particle counterpart. Meanwhile, the fluctuation in the self-energy is mainly contributed by purely extra screening $iG_{\text{int}}\delta W$ whereas the second-order residue $i\delta G\delta W$ is quite insignificant (see Fig. 6.14). Therefore, the QP peak will emerge in the anomalous region of the self energy, where significant energy exchange between the QP state and the acoustic plasmon occurs, resulting in an unusual enlargement of the QP band gap. In Fig. 6.13, it can be seen that the actual QP band gap (solid squares) drop slower than the extra exchange energy (the blue dashed line) and exhibit a roughly linear scaling rule, indicating the anomalous contribution from the extra screening due to doping.

The neat physics behind is that the doping-induced intraband transitions introduces two distinct types of excitations, the single-particle transitions and the collective excitations of plasmons; While the former strengthens the screening and weakens the effective Coulomb interactions, the latter actually makes the electron system more correlated, leading to larger self-energy corrections. The many-body interactions in an intrinsic system renormalizes the band energies to where the emerging low-energy
Figure 6.14. (Color online) Comparison of solution to Dyson equation with (solid curve) or without (dashed curve) the extra SX term $i\delta G \delta W$ included in the self energy for $n_d = 2.28 \times 10^{12} \text{cm}^{-2}$. $i\delta G \delta W$ scarcely influences the QP solution at this doping level.
plasmons play an important role, resulting in the anomalous enlargement of the QP band gap.

As the doping level keeps increasing in the conduction band, the band gap ceases to decrease and takes on a new scaling rule. On the one hand, the negative exchange energy continues to contribute to lowering conduction band, pulling the QP solution out of the anomalous plasmon-resonant region back to the single-particle energy. For example, in Fig. 6.12 (b), the final QP energy $E_{\text{QP}}^{\text{CBM}}$ has almost reached the single-particle energy $E_{\text{LDA}}^{\text{CBM}}$. On the other hand, the extra screened exchange energy from $i\delta G\delta W$ becomes dominant in the self energy and reduces the negativity of the total exchange energy by virtue of the extra screening. When the doping level is sufficiently high, the total exchange energy is largely screened by the doping-induced transitions such that the QP band gap even begins to rise, although at a much slower rate than the previous band gap narrowing. For monolayer MoS$_2$, this band gap reopening occurs around a doping density of $5 \times 10^{12}$cm$^2$, as is shown in Fig. 6.13. For references, the band gap defined by the virtual plasmaron solutions are also shown in the same figure, marked by the empty squares.

In this section, we have performed a first-principles calculation within the $G0W0$ approximation. We carefully model of doping effects in 2D systems to demonstrate a abnormal scaling rule of band gap renormalization. Surprisingly, the primary contribution responsible for band gap narrowing is not the extra screening due to doping but the negative exchange energy arising from filling (withdrawing) electrons to the doped band; the intraband transitions introduce a branch of low-energy plasmon that
eventually strengthens the electron correlations in the system, leading to slower, linear band gap narrowing at light doping density. After a critical doping level, the negative exchange energy of the doped band is largely screened by the intraband transitions and the QP band gap begins to increase at a slow rate.

6.4 Quasiparticle Properties: \( GW_0 \) Approximation

It is important to point out that self-consistency has to be included in the \( GW \) method for producing the correct QP band gap for the doped reduced-dimensional structures. There could be multiple solutions to the Dyson equation because of the highly nonlinear \( \Sigma_3 \). On the \( G_0W_0 \) level, we take the solution with the largest spectral weight to be the band edge energy. For the VBM, the \( G_0W_0 \) solution (black thin circle in Fig. 6.15 (a)) lies far below the mean-field energy \( E_{\text{LDA VBM}} \), where the contribution of \( \Sigma_3 \) is negative. This suggests a unphysical situation where \( \Sigma_3 \), that stems from the extra screening introduced by doping, is responsible for enlarging the size of the QP band gap rather than reducing it. The dilemma can be resolved by imposing self-consistency on the propagator \( G \), which always requires the QP solution to coincide with the on-shell energy of \( \Sigma_3 \) (black thick circle in 6.15 (a) and (b)). As is shown in Fig. 6.15 (e), the final band gaps predicted by the \( G_0W_0 \) are considerably larger than by the self-consistent \( GW \) method, which is more reliable for solids [161–163].

After inclusion of the self-consistency, there are also two distinct solutions. The other solution below the QP solution (blue circles in Fig. 6.15(a) and (b)) are known
Figure 6.15. (Color online) (a)-(b) Graphic solutions to the Dyson equation with different levels of self-energies included. The solid dark green curves represent the right hand side of the Dyson equation with self-consistency ($GW_0$) taken into account while the dashed dark green curves represent the one without self-consistency ($G_0W_0$). The QP solutions are marked by thick black circles and the “plasmaron” solutions are marked by thin blue circles. The mean-field energies (LDA level) are also marked by gray squares for references. (e) Band gap evolution with doping density using different levels of self-energies and self-consistency.
as a “plasmaron” excitation [164], which is believed to be spurious in a 2D electron gas because further correlations need to be considered [160].

Our *ab initio* calculations reveal an unusual QP band gap scaling rule versus the doping density \( n_{2D} \): at light doping, the band gap drops dramatically with \( n_{2D} \) as in other doped 2D structures [160, 165–169]; however, as \( n_{2D} \) is sufficiently high (\( 10^{13}\text{cm}^2 \)), the band gap reaches a plateau, which is not found by previous calculations. We attribute the plateau region to two reasons. First, the contribution from the band filling with intrinsic screening \( i\delta GW_{\text{int}} \) and of the extra screening on intrinsic system \( iG_{\text{int}}\delta W \) are responsible for narrowing the band gap but the two term scale no faster than \( \sqrt{n_{2D}} (k_F) \) (see the blue dashed and red dashed curves in Fig. 6.15 (e)), due to the 2D nature of the system. By contrast, the scaling law would roughly be proportional to the doping density \( n_{1D} (k_F) \) in a 1D system [143, 144]. In addition, the amount of exchange energy due to band filling is further diminished by the extra screening from doping. The corresponding contribution, \( i\delta G\delta W \), is responsible for enlarging the band gap and it becomes important and should not be overlooked at high \( n_{2D} \).

It should be noted that the calculated band gap within the self-consistent \( GW \) method is noticeably higher than the prediction of the GPP model [62]. Because the model only includes one-single pole in the dielectric function, which can not be used to describe the normal interband plasmon and the highly dispersive carrier plasmon at low energies simultaneously. The doping-induced screening is overestimated by the GPP model due to the large difference of \( \epsilon^{-1}_{00}(q, \omega) \) in the static limit.
6.5 Conclusions

In conclusion, we have demonstrated a highly effective generic computational scheme for modelling the dielectric function of the doped reduced-dimensional structures. Using this computational scheme, we have obtained the change of the QP band gap over a wide range of doping density and find that the band gap drops to a certain limit at sufficiently high doping level because of a delicate competition between the exchange and correlation energy.
7. Conclusions

In this thesis, we have performed first-principles calculations to study novel two-dimensional structures, such as bilayer graphene and transition metal dichalcogenides. Many-body perturbation theories, including the $GW$ approximation and the Bethe-Salpeter equation (BSE), have been employed to investigate the quasiparticle and excitonic properties of these structures. Significant effort has been made to understand how many-body interactions are impacted by electrical gating, doping, and stacking order, which are commonly used experimental techniques of modifying the physical properties of materials. The results obtained within this work can be roughly categorized into two main themes.

One main theme is the discoveries of several unusual excitonic states that cannot be explained by the conventional hydrogen model from semiconductor physics. For example, according to the hydrogen model, the lowest exciton is an s-state, which is optically active. However, we find that in the gated bilayer graphene, due to the special Mexican-hat dispersion relation, the lowest exciton is optically inactive whereas the second-lowest exciton is very bright. The relative energy levels of the two excitons can be controlled by the gate voltage, suggesting that the gated bilayer graphene can serve as the gate-tunable optoelectronic devices. More interestingly, we have shown that this tunability stems from a sheer separation of the electron and
the hole in the dark exciton state, making the exciton’s energy level highly sensitive
to external electric field. This gate-tunability may also lead to intriguing behaviors
of Bose-Einstein condensation of excitons that need to be further demonstrated by
experiments.

Another example of unusual excitonic states occurs in twisted bilayer graphene,
which is a gapless (semi-metallic) system. It is generally believed that bound excitons
cannot exist in such systems because excited electrons and holes can efficiently relax
to the fermi level and recombine there easily. Our first-principles study shows that
this is not true in twisted bilayer graphene. The stacking fault in the system results
in the emergence of two sets of parallel band transitions. Each set of transitions may
give rise to a branch of subband excitons, which is strongly bound. Unfortunately, the
binding feature of these excitons could be severely deteriorated by the hybridization
with the other band transitions at lower energies. By solving the BSE, we have
demonstrated that bound exciton can indeed survive in the gapless twisted bilayer
graphene. This bound exciton is free from hybridization with the e-h continuum as a
consequence of anti-symmetric superposition of the aforementioned subband excitons.
Our calculation reveals a novel mechanism for bound exciton in gapless systems by
virtue of quantum interference effects.

The other main themes in the thesis is about the many-body effects in doped
reduced-dimensional materials. Their screening is featured by a carrier plasmon whose
frequency approaches zero in the long-wavelength limit. Therefore, the corresponding
dielectric function cannot be captured by the traditional plasmon-pole approach for
semiconductors. In principle, a full-frequency treatment, which is computationally expensive, have to employ for the doped reduced-dimensional materials. However, based on the analysis from our first-principles, we have shown that the dielectric function in these systems can be modelled just by slightly modifying the plasmon-pole approach, making the many-body calculation much more efficient. Using our modified plasmon-pole approach, we have studied the band gap renormalization of the doped monolayer MoS$_2$ and revealed an unusual band gap scaling law versus the doping level; as the doping level is sufficiently high, the band gap ceases to narrow down and plateau due to a delicate cancellation of the exchange energy and correlation energy. This result provides an important clue for band gap engineering of 2D materials via doping.
REFERENCES


[10] DC Elias, RR Nair, TMG Mohiuddin, SV Morozov, P Blake, MP Halsall, AC Ferrari, DW Boukhvalov, MI Katsnelson, AK Geim, et al. Control of


[40] Frantisek Karlicky and Michal Otyepka. Band gaps and optical spectra of chlorographene, fluorographene and graphane from g0w0, gw0 and gw calculations on top of pbe and hse06 orbitals. *Journal of Chemical Theory and Computation*, 9(9):4155–4164, 2013.


VITA

Yufeng Liang was born in Guangzhou, a big city no more than two-hour drive from Hong Kong. He received his B. S. degree from the University of Science and Technology of China, where he was specialized in theoretical physics. There he also briefly worked with Prof. Jiangfeng Du on quantum computing based on nuclear resonance systems. After he came to the Washington University in fall 2009, he turned to study physics of reduced-dimensional structures under the supervision of Dr. Li Yang, with a passion of searching for exotic phenomena in the presence of many-body interactions. He will work as postdoctoral fellow at the Molecular Foundry located at the Lawrence Berkeley National Laboratory, Berkeley, California.

Publications and Preprints:


