Electrically controlled dielectric band gap engineering in a two-dimensional semiconductor

Riis-Jensen, Anders C.; Lu, Jiong; Thygesen, Kristian Sommer

Published in:
Physical Review B

Link to article, DOI:
10.1103/physrevb.101.121110

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Electrically controlled dielectric band gap engineering in a two-dimensional semiconductor

Anders C. Riis-Jensen,1 Jiong Lu,2 and Kristian S. Thygesen1,3,*

1CAMD, Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark
2Department of Chemistry, National University of Singapore, 3 Science Drive 3, 117543 Singapore
3Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

(Received 15 January 2020; accepted 28 February 2020; published 16 March 2020)

The emergent class of atomically thin two-dimensional (2D) materials has opened up completely new opportunities for manipulating electronic quantum states at the nanoscale. Here we explore the concept of dielectric band gap engineering, i.e., the controlled manipulation of the band gap of a semiconductor via its dielectric environment. Using first-principles calculations based on the GW self-energy approximation we show that the band gap of a two-dimensional (2D) semiconductor, such as the transition metal dichalcogenides, can be tuned over several hundreds of meV by varying the doping concentration in a nearby graphene sheet. Importantly, these significant band gap renormalizations are achieved via nonlocal Coulomb interactions and do not affect the structural or electronic integrity of the 2D semiconductor. We investigate various heterostructure designs, and show that, depending on the size of the intrinsic dielectric function of the 2D semiconductor, the band gap can be tuned by up to 1 eV for graphene carrier concentrations reachable by electrostatic doping. Our work provides opportunities for electrically controllable band gap engineering in 2D semiconductors.

DOI: 10.1103/PhysRevB.101.121110

I. INTRODUCTION

Two-dimensional (2D) materials such as graphene, hexagonal boron nitride (hBN), and the semiconducting transition metal dichalcogenides (TMDCs), represent an emergent class of materials with unique electronic and optical properties making them interesting candidates for applications in a number of areas including (opto)electronics, photonics, energy, and quantum technology [1–5]. A unique merit of these materials is that their properties can be tuned much more easily than is possible with conventional bulk materials. This feature stems from the extreme thinness of 2D materials, which makes the electronic states living inside them highly susceptible to their environment [6]. In particular, 2D semiconductors encapsulated in van der Waals heterostructures constitute an ideal platform for practicing and exploiting the art of band gap engineering.

In addition to the conventional band gap engineering schemes based on doping and alloying, the band gap of 2D semiconductors can be controlled via mechanical strain [7,8], layer stacking [9,10], and external electric fields [11,12]. In all these cases, the fundamental physical mechanism underlying the band gap modification can be explained within a single-particle picture. A fundamentally different type of band gap engineering can be achieved in 2D materials by exploiting the dependence of the electronic quasiparticle (QP) energies on the dielectric environment outside the 2D layer. This dependence stems from the electron’s self-energy, which describes the interaction between the electron (or hole) and its self-induced screening cloud. Because the shape and size of the screening cloud depends on the dielectric environment, this effect can be exploited to change the QP energies without changing the spatial shape or hybridization pattern of the 2D semiconductor wave functions. Because of its nonlocal and dynamical nature, dielectric screening effects on QP energies are not captured by traditional mean-field approximations, but require more rigorous QP theories such as the many-body GW method [13,14].

Quasiparticle screening effects manifest themselves most clearly in molecules deposited on metallic or dielectric substrates. In such systems, the dielectric screening from the substrate can shift the molecule’s orbital energies by several eV relative to the gas-phase energies. These effects, which have been demonstrated both experimentally [15–17] and theoretically [18–20], are essential to take into account for a proper description of electron transport in molecular junctions. Here the dielectric screening from the metallic electrodes can strongly influence the position of the frontier molecular orbitals relative to the electrode Fermi level and thereby affect the tunneling rate [21,22]. The shift in molecular energy levels can often be explained semiquantitatively by a classical image charge model [18,19], although dynamical corrections and wave function modifications may also play a role [23,24]. Compared to molecules, 2D semiconductors have better intrinsic screening ability making them less susceptible to their dielectric environment. Still, the internal screening is weak enough that the QP band gap can be reduced by up to 1 eV by substrate screening, as first shown by first-principles GW calculations for hBN on graphene [13]. Experimentally, it was demonstrated using nonlinear optical spectroscopy that the QP band gap of monolayer WS2 can be varied by 0.3 eV depending on the choice of substrate [25], and by GW calculation that the QP band gap is reduced by about 0.5 eV for MoS2 on a gold substrate [26]. Subsequently, the concept of dielectric band gap engineering in
2D materials has been explored for both lateral [27,28] and vertical [29–31] heterostructure designs. In a recent work, Qiu et al. [32] investigated the effect of dielectric screening on the quasiparticle band gap of a single layer of ReS$_2$ placed on top of a graphene sheet separated by $N$ layers of hBN. The concentration of charge carriers in the graphene sheet can be adjusted via a bottom gate. An increased doping level enhances the dielectric function of the graphene sheet, which in turn weakens the screened electron-electron interaction inside the 2D semiconductor leading to a reduction of its QP band gap.

In this work we present a theoretical analysis of a similar setup providing both fundamental insight into the physical mechanisms governing the idea of electrically controllable dielectric band gap engineering in a 2D semiconductor and numerical relations between key system parameters that can be used to guide the future design of van der Waals (vdW) heterostructures with tailored band gaps. To this end we consider a van der Waals heterostructure consisting of the 2D semiconductor placed on a graphene sheet (see Fig. 1). For various types of 2D semiconductors, we explore how the QP band gap depends on the doping concentration in the graphene sheet. In practice, the latter can be controlled by an electrostatic bottom gate. This setup allows the strength of the screened Coulomb interaction inside the 2D semiconductor and thus the electron self-energy, to be controlled externally via the concentration of free charge carriers in the graphene sheet. The calculations are made possible by the recently developed $G\Delta W$ method [29], which calculates the change in the GW self-energy due to the additional screening provided by the heterostructure environment, which in turn is obtained using the quantum electrostatic heterostructure (QEH) model [33].

II. RESULTS AND DISCUSSION

The QP energies are obtained from the linearized QP equation

$$\epsilon_{n,\text{QP}} = \epsilon_n + Z_n \langle \psi_n | \Sigma (\epsilon_n) | \psi_n \rangle,$$

where $\psi_n$ and $\epsilon_n$ represent approximate single-particle wave functions and energies that are typically obtained from a density functional theory calculation. In this work, we use the Perdew-Burke-Ernzerhof (PBE) exchange-correlation ($xc$) functional to determine $\psi_n$ and $\epsilon_n$. In the QP equation, the self-energy operator, $\Sigma$, which describes the interaction of the electron/hole with its self-induced screening cloud, replaces the local mean-field exchange-correlation potential, $v_{xc}$. The renormalization factor $Z_n$ accounts for the energy variation of $\Sigma (\epsilon)$ around the energy $\epsilon_n$.

Within the so-called $G_0W_0$ method, the electron self-energy is approximated by the product of the Green’s function ($G$) and dynamically screened electron-electron interaction ($W$). The $0$ subscript indicates that $G$ and $W$ are evaluated from the noninteracting $\psi_n$ and $\epsilon_n$, i.e., the non-self-consistent $GW$ approximation. The screened interaction, which is the main object of interest in the present work, is defined as

$$W(\mathbf{r}, \mathbf{r'}, \omega) = \int \frac{1}{|\mathbf{r''} - \mathbf{r'}|} d\mathbf{r''},$$

where $\epsilon(\mathbf{r}, \mathbf{r'}, \omega)$ is the microscopic dielectric function of the material. We shall not go deeper into the theory of the GW method here but refer the interested reader to one of the many excellent reviews on the topic [34,35]. For the purpose of the present work it suffices to note that due to the strong nonlocality of the Coulomb kernel and dielectric function, it is possible to affect the screened interaction $W$ between electrons in one region of space (here the 2D semiconductor) by modifying the dielectric response in another region (here the graphene sheet).

Conventional implementations of the GW method do not readily allow for the study of systems like the one shown in Fig. 1. This is due to the shear size of the supercells needed to describe vdW heterostructures comprising 2D materials with different in-plane lattice constants. For example, the smallest supercells required to model a WS$_2$/graphene bilayer when the strain on each layer is not allowed to exceed 1%, contains more than 300 atoms, which exceeds the system sizes that can be treated by conventional GW codes. Furthermore, a proper treatment of the dielectric function of doped graphene requires extremely dense $k$-point sampling around the Dirac cones making standard calculations intractable. In this work we rely on the $G\Delta W$ method [29] in combination with the QEH model [33]. The QEH model is used to compute the dielectric function of the entire vdW heterostructure from the dielectric functions of the individual layers. The method circumvents the problem of large supercells due to lattice mismatched 2D layers and allows for very fine $k$-point meshes for the individual layers. We use the QEH model to calculate the change in the screened interaction of the 2D semiconductor, $\Delta W$, due to the presence of the other layers of the heterostructure (hBN and doped graphene). In a next step, we calculate the change in the self-energy as $\Delta \Sigma = G\Delta W$ (integral over frequency not shown), which in turn yields the change in the QP gap of the semiconductor. For more details on the $G\Delta W$ methods we refer to Ref. [29]. We note in passing that other approximation schemes for dealing with dielectric screening in GW calculations for large systems have been proposed [36,37].

In the following we consider the setup shown in Fig. 1 consisting of a monolayer transition metal dichalcogenide (TMD) placed on top of a graphene sheet, possibly with a number of hBN layers inserted in between. In Fig. 2 we plot the QP band gap of four different TMDs as a function of...
FIG. 2. The quasiparticle band gap of four different TMD monolayers placed on top of a graphene sheet as a function of the position of the Fermi level (i.e., carrier concentration) in graphene. The configuration corresponds to Fig. 1 without any hBN layers between the TMD and graphene sheet. The band gap of the freestanding TMD is indicated by dashed lines.

FIG. 3. Quasiparticle band gap of MoS$_2$ in the geometry depicted in Fig. 1 as a function of the number of hBN layers between MoS$_2$ and graphene. Results are shown for three different doping concentrations in the graphene layer. The dashed black, green, and blue lines are fits of the form $a(E_F)/(N + b) + E^\infty_{gap}$, where the asymptotic value $E^\infty_{gap}$ is the calculated QP band gap of MoS$_2$ on bulk hBN.

In the configuration considered in Fig. 1 (no hBN layers between the TMD and graphene), the screening from undoped graphene reduces the TMD band gap by around 0.30 eV, as compared to the freestanding TMD. The band gaps are then further reduced by 0.25 eV as the graphene Fermi level is increased from 0 to 0.8 eV. We stress that these very substantial band gap reductions are achieved without modifying the atomic structure, the shape of the wave functions, or their hybridization patterns, and highlights the unique opportunities offered by atomically thin materials for tuning electronic properties.

In Fig. 3 we show the variation of the TMD band gap (here exemplified by MoS$_2$) as a function of the number of hBN layers separating the TMD and the graphene layer. Results are shown for the case of intrinsic graphene and for doping concentrations corresponding to Fermi level shifts of 0.4 and 0.8 eV, respectively. The band gap shows an approximate $1/N$ dependence on the number of hBN layers. The dashed lines indicate best fits to the functional form $a(E_F)/(N + b) + E^\infty_{gap}$, where the asymptotic value $E^\infty_{gap}$ is the band gap of MoS$_2$ on bulk hBN (which we calculate explicitly), $b$ is an image plane position, and $a$ is the strength of the image charge. In these fits only $a$ was taken to depend on the doping concentration in the graphene layer. The good quality of the fit confirms the interpretation of the band gap reduction as an "image charge" effect, i.e., the band gap reduction is due to the interaction between the electron/hole in the TMD and its self-induced screening cloud in the graphene layer. Furthermore we find $b$ to be very close to two times the MoS$_2$/graphene interlayer.
FIG. 4. The reduction of the QP band gap of the semiconductor monolayers listed in Table I when placed on intrinsic and doped graphene (corresponding to $\Delta E_F = 0.4$), respectively. The band gap reduction is plotted versus $1/\alpha$, where $\alpha$ is the static in-plane polarizability of the freestanding 2D semiconductor. It is clear that the band gap reduction is stronger for 2D materials with weaker intrinsic screening suggesting that the band gap renormalization is determined by the relative change in screening provided by the graphene layer. The dashed lines are added as guides to the eye and the orange dot corresponds to previous experimental work [32] (see main text).

distance, and we therefore believe this accurately reproduces the distance to the image charge plane.

In the simplest image charge model, the band gap reduction is given by the electrostatic interaction between a point charge in the TMD and its image charge in the graphene layer. Indeed, this picture has been found to hold to a very good approximation for molecules on metallic or insulating surfaces [18,19]. In Fig. 4 we show the calculated band gap renormalization for a number of different 2D semiconductors when placed on top of the graphene layer. The considered 2D semiconductors are listed in Table I together with their QP band gaps and static in-plane polarizability taken from the Computational 2D Materials Database (C2DB) [38]. The band gap reductions are plotted versus the size of the static in-plane polarizability of the 2D semiconductor. It is clear that, independent of the doping concentration in the graphene sheet, the band gap reduction is larger for 2D materials with lower polarizability. This behavior can be understood by realizing that the magnitude of the screening cloud in the graphene layer, i.e., the image charge, depends on the total charge of the quasiparticle in the 2D semiconductor, i.e., the bare electron/hole plus its screening cloud in the 2D semiconductor. The latter is obviously larger in 2D semiconductors with a high intrinsic polarizability. This shows that the amount by which the band gap can be tuned via the dielectric environment depends on the degree of internal screening in the 2D semiconductor itself. We note the band gap reduction found by Jiong et al. [32], where a reduction of about 0.6 eV of the band gap was found when monolayer ReSe$_2$ was placed on doped graphene. We have calculated the averaged in-plane static polarizability of ReSe$_2$ to be 6.58 Å, corresponding to the orange dot in Fig. 4 and we find an excellent agreement with the calculated values for the band gap reduction of other 2D monolayers.

In this work we have focused on the effect of dielectric screening on the QP band gap. In general, optical excitations, in particular bound excitons, are less affected by dielectric screening because of the neutral nature of such excitations. On the other hand, more loosely bound excitons or excitons with charge transfer character [39,40], present larger electron-hole separations and consequently experience stronger renormalization by the dielectric environment. While the effect of dielectric screening on optical excitations, in particular the

<table>
<thead>
<tr>
<th>Monolayer</th>
<th>Band gap (eV)</th>
<th>Static polarizability (Å$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BN</td>
<td>7.17</td>
<td>0.96</td>
</tr>
<tr>
<td>HfO$_2$</td>
<td>4.70</td>
<td>1.36</td>
</tr>
<tr>
<td>SnO$_2$</td>
<td>1.84</td>
<td>1.84</td>
</tr>
<tr>
<td>TiO$_2$</td>
<td>3.95</td>
<td>1.99</td>
</tr>
<tr>
<td>HfS$_2$</td>
<td>2.94</td>
<td>3.16</td>
</tr>
<tr>
<td>SnS$_2$</td>
<td>2.29</td>
<td>3.56</td>
</tr>
<tr>
<td>HfSe$_2$</td>
<td>2.12</td>
<td>4.38</td>
</tr>
<tr>
<td>WS$_2$</td>
<td>2.52</td>
<td>5.58</td>
</tr>
<tr>
<td>MoS$_2$</td>
<td>2.53</td>
<td>6.19</td>
</tr>
<tr>
<td>WSe$_2$</td>
<td>2.10</td>
<td>6.59</td>
</tr>
<tr>
<td>MoSe$_2$</td>
<td>2.12</td>
<td>7.28</td>
</tr>
<tr>
<td>MoTe$_2$</td>
<td>1.56</td>
<td>9.47</td>
</tr>
<tr>
<td>CrSe$_2$</td>
<td>1.17</td>
<td>10.31</td>
</tr>
</tbody>
</table>

FIG. 5. Contour plots: real part of the dielectric function of MoS$_2$/graphene evaluated in the MoS$_2$ layer for different doping concentrations along the in-plane wave vector. Bottom right plot: The real part of the dielectric function evaluated along $q_\parallel = 0.05$ Å$^{-1}$ for intrinsic MoS$_2$ (gray) and for three different doping concentrations for MoS$_2$/graphene (black, green, and blue).

TABLE I. Band gap and static in-plane polarizability for the 13 monolayers shown in Fig. 4.
lowest bound excitons in 2D materials, are easier to probe experimentally as compared to the QP gap, the theoretical treatment is limited by the assumption of static screening. This assumption is valid when the dielectric function shows weak variation with frequency up to a characteristic frequency given by the exciton binding energy. This condition is usually satisfied in intrinsic semiconductors and insulators. However, when the dielectric function used to screen the electron-hole interaction has a metallic component (like the intraband screening in doped graphene; cf. Fig. 5) the assumption of static screening cannot be justified and more elaborate methods are required in order to deal with the frequency dependent electron-hole interaction [41].

In conclusion, our ab initio GW calculations show that it is possible to control the quasiparticle band gap of an atomically thin 2D semiconductor by varying the doping concentration in a nearby graphene sheet. The physical mechanism underlying the band gap renormalization is that the size of the screening cloud dressing an electron/hole of the 2D semiconductor, which determines its self-energy, is controlled by the screening ability of the graphene. The effect was shown to be well described by a simple image charge model with the caveat that the magnitude of the image charge depends on the intrinsic screening of the 2D semiconductor. Thus, in general, the effectiveness of dielectric band gap engineering is determined by the intrinsic screening properties of the 2D material itself. Our work highlights the unique opportunities for shaping electron energy landscapes in 2D materials by dielectric engineering without altering the atomic or electronic integrity of the 2D material itself.

III. METHOD

All calculations have been performed with the electronic structure software code GPAW [42]. The atomic structures were relaxed using PBE with a plane-wave cut-off energy of 800 eV, on a k-point grid with a density of 6.0 Å\(^{-1}\), and with a Fermi smearing of 0.05 eV. All structures were relaxed until the maximum force on any atom was 0.01 eV/Å and the maximum stress on the unit cell was 0.002 eV/Å\(^3\). The ground state was calculated using the same parameters as for the relaxation, but with a k-point grid of 12.0 Å\(^{-1}\). The GW calculations were performed starting from the PBE wave functions. Three GW calculations were done with a cut-off energy for the self-energy of 170, 185, and 200 eV from which the band gap was determined by extrapolating to an infinite plane-wave cut-off. The electronic band gaps for the van der Waals heterostructures were determined by calculating the screening correction to the valence and conduction bands for the freestanding monolayer semiconductor in the QEH model [33] which have previously been shown to yield accurate results for heterostructures with a low degree of hybridization at the band edges. The static polarizabilities are calculated within the random phase approximation (RPA), which were based on a PBE ground state calculation with a 20 Å\(^{-1}\) k-point density grid and a cut-off energy of 50 eV was used for the RPA calculation. In both the GW and the RPA calculations a truncated Coulomb interaction was used. All band structure calculations are calculated including spin-orbit interaction.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (Grant Agreement No. 773122, LIMA). The Center for Nanostructured Graphene is sponsored by the Danish National Research Foundation, Project No. DNRF103.