

***Ab initio*-based Microscopic Modeling of the Optoelectronic Properties of Semiconductors**

Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Physik
der Philipps-Universität Marburg
vorgelegt

von

Sven Christian Liebscher

aus Einbeck

Marburg, 2022

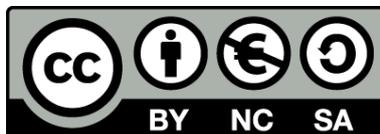
Vom Fachbereich Physik der Philipps-Universität Marburg
als Dissertation angenommen am: 24.11.2022

Hochschulkenziffer: 1180

Erstgutachter: Prof. Dr. Ermin Malic
Zweitgutachter: Prof. Dr. Ralf Tonner-Zech

Tag der mündlichen Prüfung: 02.12.2022

Originaldokument gespeichert auf dem Publikationsserver der
Philipps-Universität Marburg
<http://archiv.ub.uni-marburg.de>



Dieses Werk bzw. Inhalt steht unter einer
Creative Commons
Namensnennung
Keine kommerzielle Nutzung
Weitergabe unter gleichen Bedingungen
3.0 Deutschland Lizenz.

Die vollständige Lizenz finden Sie unter:
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Inhaltsverzeichnis

Summary	VII
Zusammenfassung	XI
Danksagung	XV
Author's contributions	XVII
1 Introduction	1
2 Theoretical Background	5
2.1 Density Functional Theory	5
2.1.1 Exchange-correlation functionals	8
2.1.2 LDA-1/2 method	9
2.2 Microscopic Theory for Optical Properties	11
2.2.1 Semiconductor Bloch Equations	11
2.2.2 Semiconductor Luminescence Equations	13
2.2.3 Auger Losses	13
2.2.4 High Harmonic Generation	14
3 Application of the LDA-1/2 method to Te and GaSe	17
3.1 Te	17
3.2 GaSe	20
4 Electro-Optical Properties of Dilute Bismides	21
4.1 LDA-1/2 method for Bismuth compounds	21
4.2 Dilute materials	23
4.3 Optical Properties	25

5	Influence of TDM Phases on High Harmonic Generation	29
5.1	Triple Dipole Products	29
5.2	Berry Phase	32
6	Coulomb Interaction in Monolayer TMDCs	35
7	Conclusion	37
	Bibliography	39
	Publications	45
	Paper I	47
	Paper II	61
	Paper III	69
	Paper IV	81

Summary

In order to find new semiconductor materials for applications, it is desirable to have a theoretical approach that can predict material properties without having the need to test these properties experimentally. This way, when looking for new materials or material combinations it is easier to find promising candidates. In this thesis, DFT is used as a method that can predict electronic and structural properties of periodic semiconductor crystals from *ab initio* calculations without the need for experimental input. It is combined with the semiconductor Bloch approach, which is a quantum theory including many-body effects used to calculate optical properties.

One material class of interest for optoelectronic applications are monolayer TMDCs. We derived an expression for the quasi-2D Coulomb potential, which includes a form factor that contains the difference between the interaction in the ideally 2D case and the quasi-2D interaction in the TMDCs. DFT calculations were performed to obtain the wavefunctions necessary to evaluate the Coulomb interaction in various TMDCs. A MDF Hamiltonian was fitted to the parameters obtained from the DFT calculation. Then, the many-body theory was used to study excitonic properties and band gap renormalization effects of the TMDCs. A density-dependent band gap renormalization occurs due to the dynamic screening of the Coulomb interaction by excited carriers. It was shown that an electron-doped state does not have as strong an effect on the band gap as a symmetric electron-hole population which is usually created from optical excitation. Furthermore, the effect of the dielectric environment on the band gap renormalization was studied for freely suspended monolayers, monolayers on a fused silicon substrate and h-Bn encapsulated monolayers. By solving the Dirac-Wannier equations, exciton resonances and binding energies were calculated. It was found that an increased dielectric constant leads to an enhanced screening of the Coulomb interaction, which decreases the exciton binding energy.

The incorporation of dilute amounts of bismuth into III-V semiconductors decreases the band gap, which allows for band gap engineering to design materials for specific applications. In order to predict the material properties of these dilute bismides with DFT, large supercells are needed, which increases the computational cost immensely.

Therefore, we applied the LDA-1/2 method in DFT, which improves the band gaps obtained from DFT at a low computational cost. First, we calculated the band gaps of several III-V semiconductors using this method and found good agreement with experimental values. Then, we had to extend the method to be able to perform calculations on III-Bi compounds, which are not semiconductors in a purely binary compound. This allowed us to calculate the properties of Ga(SbBi) as a representative example of the dilute bismides. We performed DFT calculations on different supercell geometries and Bi concentrations. We studied the band gap narrowing and SO splitting increase with increasing Bi concentration and calculated effective band structures of the supercells showing defect states from the localized interaction with the Bi atoms. Absorption and PL spectra were calculated for the different concentrations and showed an earlier onset of absorption and a red-shift of the luminescence peak for increasing Bi concentrations because of the decreasing band gap. Furthermore, we studied the decreasing Auger losses in these materials due to the increasing detuning of the band gap and the SO splitting energy when incorporating higher amounts of Bi.

Another interesting candidate for optoelectronic applications is tellurium due to its chiral structure and strong nonlinear properties. We used the shLDA-1/2 method within DFT to obtain accurate electronic and structural parameters which compare well to experiments. We analyzed the TDMs of the highest four valence bands and the lowest two conduction bands for the two light polarization directions $E \parallel c$ and $E \perp c$. This revealed that at the direct band gap at the H-point, the two highest valence bands do not couple significantly to the two lowest conduction bands in the $E \parallel c$ polarization direction, making transitions between them virtually forbidden. We used our microscopic approach to evaluate the optical properties of this material and show their strong dependence on the polarization direction. The absorption is stronger for the $E \perp c$ direction, especially near the band gap, due to the weaker coupling of the valence bands in the $E \parallel c$ direction. Optical gain for incoherent carrier populations was calculated, revealing that at higher densities, the peak shifts from the transition to the lowest conduction band to the transition to the second-lowest conduction band in the $E \perp c$ direction, while no significant gain was found in the $E \parallel c$ direction. Furthermore, we calculated PL spectra and found that the emission is weaker for $E \parallel c$ and the peak shifts more than for $E \perp c$, because the coupling increases in the energy regions that are farther away from the band gap. Another interesting property of semiconductors that is studied recently is their ability to generate high harmonic spectra after excitation with non-resonant strong-field THz radiation. We studied this nonlinear optical effect in Te by solving the coupled dynamics of interband polarizations and intraband currents. We illustrated the effect of the phase of the TDMs on the HHG spectra and performed a gauging of the TDMs obtained from DFT by using the phase of triple dipole products in order to obtain smooth and periodic phases across the whole BZ. We found that for the $E \parallel c$ direction, the two topmost valence bands do not contribute significantly to the HHG emission. Even harmonics are suppressed in the $E \parallel c$ direction due to quantum interference effects. Furthermore, we studied the effect of sample thickness on the spectra using a unidirectional pulse propagation solver. For the $E \perp c$ direction, longer sample thicknesses quickly weaken

the amplitude of higher harmonics, while for the $E \parallel c$ direction, this effect is not as strong because of the lower absorption.

Since loss processes are crucial when determining the suitability of a material for applications, we performed another study on the losses in Te through radiative recombination and Auger losses. For this, the input from the DFT calculations with the shLDA-1/2 method is used to calculate radiative lifetimes and Auger rates. In the literature, different values for the splitting between the valence bands H_4 and H_6 have been reported. Since the splitting energy influences the Auger losses significantly, two different values have been considered in the study. In order to look at losses at different temperatures, the DFT band structure was shifted according to an experimentally found formula. The carrier lifetimes due to radiative and Auger recombination were calculated as a function of the carrier density at different temperatures for both splitting energies. While for low temperature and low density, the lower splitting energy causes the Auger lifetimes to be significantly shorter, at higher temperature and higher densities, the difference between the splitting energies becomes smaller, since holes far away from the band gap can reach the valence band H_6 . The radiative lifetimes are unaffected by the different splitting energies. The Auger loss coefficient and the bimolecular recombination coefficient for both splitting energies are calculated as a function of the density for different temperatures and analytical fit functions are determined for the temperature-dependent coefficients at the low-density limit.

Zusammenfassung

Um neuartige Halbleitermaterialien für Anwendungen zu finden, ist es wünschenswert, einen theoretischen Ansatz zu haben, mit dem man Materialeigenschaften vorhersagen kann, ohne diese im Experiment bestimmen zu müssen. Auf diese Art können leichter vielversprechende Kandidaten ausgemacht werden, wenn man neue Materialien oder Materialkombinationen sucht. In dieser Arbeit wird DFT als eine Methode genutzt, die elektronische und strukturelle Eigenschaften von periodischen Halbleiterkristallen durch *ab initio* Rechnungen bestimmen kann, ohne dafür experimentelle Werte zu brauchen. Sie wird verbunden mit dem Halbleiter-Bloch-Ansatz, einer quantenmechanischen Theorie, die Vielteilcheneffekte miteinbezieht und benutzt wird, um optische Eigenschaften zu berechnen.

Eine Materialklasse, die interessant für opto-elektronische Anwendungen ist, sind Monolagen-TMDCs. Wir haben einen Ausdruck für das quasi-zweidimensionale Coulombpotenzial hergeleitet, welcher einen Formfaktor enthält, der den Unterschied zwischen der Wechselwirkung in einem perfekt zweidimensionalen Fall und der quasi-zweidimensionalen Wechselwirkung in den TMDCs beschreibt. Die Wellenfunktionen, die nötig sind, um die Coulombwechselwirkung zu berechnen, wurden durch DFT-Rechnungen für diverse TMDCs bestimmt. Ein MDF-Hamiltonian wurde an die Parameter aus den DFT-Rechnungen angefügt. Danach wurde die Vielteilchentheorie verwendet, um die exzitonischen Eigenschaften und Bandlücken-Renormierungs-Effekte in den TMDCs zu untersuchen. Eine dichteabhängige Bandlücken-Renormierung findet durch die dynamische Abschirmung der Coulombwechselwirkung von den angeregten Ladungsträgern statt. Es konnte gezeigt werden, dass in einem elektronendotierten Zustand dieser Effekt weniger stark ausfällt als für eine symmetrische Elektronen-Loch-Besetzung, welche üblicherweise bei einer optischen Anregung erzeugt wird. Desweiteren wurde der Einfluss der dielektrischen Umgebung auf die Bandlücken-Renormierung für frei schwebende Monolagen, für Monolagen auf einem Substrat aus Quarzglas und für in hBN eingekapselte Monolagen untersucht. Exziton-Resonanzen und Bindungsenergien wurden durch das Lösen der Dirac-Wannier-Gleichungen bestimmt. Es zeigte sich, dass eine höhere dielektrische Konstante zu einer stärkeren Abschirmung der

Coulombwechselwirkung führt, was die Exziton-Bindungsenergie verringert.

Der Einbau von geringen Mengen an Bismuth in III-V-Halbleiter verringert deren Bandlücke, was ein gezieltes Anpassen der Bandlücke (*engl. band gap engineering*) für spezifische Anwendungen erlaubt. Um die Materialeigenschaften dieser sogenannten verdünnten Bismide (*engl. dilute bismides*) mit DFT vorherzusagen, werden große Superzellen benötigt, welche den Rechenaufwand erheblich steigern. Aus diesem Grund haben wir die LDA-1/2-Methode in DFT angewandt, welche die damit erhaltenen Bandlücken bei niedrigem Rechenaufwand verbessert. Zunächst haben wir die Bandlücken einiger III-V-Halbleiter mit dieser Methode bestimmt und eine gute Übereinstimmung mit experimentellen Werten festgestellt. Anschließend mussten wir die Methode erweitern, damit sie auch für III-Bi-Verbindungen angewandt werden kann, die als binäres Material keine Halbleiter sind. Dadurch war es uns möglich, die Eigenschaften von Ga(SbBi) als ein Vertreter der verdünnten Bismide zu berechnen. Wir haben DFT-Rechnungen mit verschiedenen Superzellgeometrien und Bismuth-Konzentrationen durchgeführt. Wir haben die Verkleinerung der Bandlücke und die Vergrößerung der Spin-Bahn-Aufspaltung mit steigendem Bi-Gehalt untersucht und effektive Bandstrukturen der Superzellen bestimmt, in denen Defektzustände durch die lokalisierte Wechselwirkung mit den Bismuth-Atomen zu sehen sind. Absorptions- und Photolumineszenzspektren wurden für verschiedene Bismuth-Konzentrationen berechnet und zeigten ein Einsetzen der Absorption bei niedrigeren Energien und eine Rotverschiebung des Maximums der Photolumineszenz für höhere Bismuth-Konzentrationen aufgrund der Verkleinerung der Bandlücke. Außerdem haben wir die Abnahme der Augerverlustrate untersucht, die dadurch entsteht, dass die Bandlücke und die Spin-Bahn-Aufspaltung weiter auseinander gehen, wenn höhere Mengen von Bismuth eingebaut werden.

Ein weiterer spannender Kandidat für optoelektronische Anwendungen ist Tellur aufgrund seiner chiralen Struktur und starken nicht-linearen optischen Eigenschaften. Wir haben die shLDA-1/2-Methode in DFT verwendet, um genaue elektronische und strukturelle Parameter zu bestimmen, die gut mit dem Experiment übereinstimmen. Wir haben die Übergangsdipolmomente der obersten vier Valenzbänder und untersten zwei Leitungsbänder für die beiden Polarisationsrichtungen des Lichts $E \parallel c$ and $E \perp c$ analysiert. Dies hat gezeigt, dass an der direkten Bandlücke am H-Punkt die beiden obersten Valenzbänder nicht nennenswert an die untersten Leitungsbänder in der Polarisationsrichtungen $E \parallel c$ koppeln, was Übergänge zwischen ihnen quasi unmöglich macht. Wir haben unseren mikroskopischen Ansatz verwendet, um die optischen Eigenschaften dieses Materials zu bestimmen und ihre starke Abhängigkeit von der Polarisationsrichtung des Lichts zu zeigen. Die Absorption ist in der $E \perp c$ -Richtung besonders in der Nähe der Bandlücke stärker, was an der schwächeren Kopplung der Valenzbänder in der $E \parallel c$ -Richtung liegt. Die optische Verstärkung für inkohärente Ladungsträgerbesetzungen wurde berechnet, wobei in der $E \perp c$ -Richtung bei höheren Ladungsträgerdichten das Maximum des Spektrums nicht mehr durch den Übergang in das unterste, sondern in das zweitunterste Leitungsband erzeugt wird, während in der $E \parallel c$ -Richtung keine nennenswerte optische Verstärkung auftritt.

Desweiteren haben wir Photolumineszenz-Spektren berechnet, welche zeigen, dass die

Emission in der $E \parallel c$ -Richtung schwächer ist und die Maxima sich stärker verschieben als in der $E \perp c$ -Richtung, da die Kopplung in den Energieregionen weiter weg von der Bandlücke zunimmt. Eine weitere spannende Eigenschaft von Halbleitern, die in letzter Zeit untersucht wird, ist die Erzeugung von Spektren Hoher Harmonischer nach einer nicht-resonanten Anregung mit THz-Strahlung mit einem starken elektrischen Feld. Wir haben diesen nichtlinearen optischen Effekt in Tellur untersucht, indem wir die gekoppelte Dynamik von Interband-Polarisationen und Intraband-Strömen lösten. Wir haben den Einfluss der Phasen der Übergangsdipolmomente auf die Hohen-Harmonischen-Spektren illustriert und die Dipole aus DFT-Rechnungen mit Dreifachprodukten aus Übergangsdipolmomenten geeicht, um glatte und periodische Phasen über die gesamte Brillouinzone zu erhalten. Wir haben festgestellt, dass in der $E \parallel c$ -Richtung die obersten zwei Valenzbänder nicht wesentlich zur Erzeugung der Hohen-Harmonischen-Spektren beitragen. Gerade Harmonische werden in der $E \parallel c$ -Richtung durch Quanteninterferenzeffekte unterdrückt. Außerdem haben wir den Einfluss der Probendicke auf die Spektren untersucht, indem wir die Gleichungen für eine unidirektionale Puls-Propagation gelöst haben. In der $E \perp c$ -Richtung werden Harmonische im oberen Teil des Spektrums schnell mit zunehmender Probendicke abgeschwächt, während in der $E \parallel c$ -Richtung dieser Effekt wegen der niedrigeren Absorption nicht so stark ist.

Da Verlustprozesse entscheidend sind, wenn es um die Eignung eines Materials für Anwendungen geht, haben wir die Verlustprozesse durch strahlende Rekombinationen und Auger-Verluste in Tellur untersucht. In der Literatur findet man verschiedene Werte für die Aufspaltung zwischen den Valenzbändern H_4 und H_6 in Tellur. Da diese Aufspaltungsenergie die Auger-Verluste erheblich beeinflusst, haben wir zwei verschiedene Werte für die Untersuchungen angenommen. Um Verluste bei verschiedenen Temperaturen zu betrachten, wurde die Bandstruktur aus DFT anhand einer experimentell bestimmten Formel verschoben. Die durch strahlende und Auger-Rekombination bedingten Ladungsträger-Lebensdauern wurden als Funktion der Ladungsträgerdichte bei verschiedenen Temperaturen für beide Aufspaltungsenergien bestimmt. Während bei niedriger Temperatur und niedriger Ladungsträgerdichte die kleine Aufspaltungsenergie dafür sorgt, dass die Auger-Lebensdauern wesentlich kürzer sind, verkleinert sich dieser Effekt für höhere Temperaturen und höhere Ladungsträgerdichten, da weit von der Bandlücke entfernte Löcher dann das Valenzband H_6 erreichen können. Die strahlenden Lebensdauern werden von der unterschiedlichen Aufspaltungsenergie nicht beeinflusst. Der Auger-Verlustkoeffizient und der bimolekulare Rekombinationskoeffizient für beide Aufspaltungsenergien wurden als Funktion der Ladungsträgerdichte für verschiedene Temperaturen bestimmt und analytische Fit-Funktionen für die temperaturabhängigen Koeffizienten im Grenzfall niedriger Ladungsträgerdichten bestimmt.

Danksagung

Ich möchte mich ganz herzlich bei Prof. Dr. Stephan Koch bedanken für die gute Betreuung während meiner Zeit als Doktorand in seiner AG. Außerdem danke ich Prof. Dr. Ralf Tonner-Zech für die Übernahme des Zweitgutachtens. Für die gute Zusammenarbeit im GRK bedanke ich mich auch bei Prof. Dr. Kerstin Volz und gleich mehrfach bei Isabelle Kimmel, sowie auch bei den anderen GRK-Mitgliedern für den guten Austausch.

Ein besonderer Dank geht an Christian Berger, durch den ich in die AG gekommen bin, da er mich motiviert hat, dort ein Projektpraktikum zu machen. Vielen Dank auch an Lars Bannow dafür, dass er mich so gut in die Forschungsarbeit und die IT-Aufgaben in der AG eingearbeitet hat. Den beiden und auch den restlichen Mitgliedern der AG danke ich auch für die gute Zusammenarbeit und das nette Miteinander: Josefine, Maria H., Ada, Tineke, Hanno, Lars M., Philipp, Osmo, Gaurav. Sehr viel Dank geht auch an Prof. Dr. Jerome Moloney dafür, dass er mir ermöglicht hat, eine Zeit lang in seiner AG mitzuarbeiten, und an die Mitglieder seiner Arbeitsgruppe, besonders Jörg H., Anton R., Paris P. und Jalen C. für den herzlichen Empfang. Jörg gebührt auch zusätzlich mein tiefer Dank für die vielen hilfreichen, wenn auch zeitverschobenen Email-Korrespondenzen.

Außerhalb des akademischen Bereichs danke ich zunächst einmal meinen (Stief-)Eltern, Jens, Katrin und Martina dafür, dass sie mich immer unterstützt haben. Desweiteren danke ich auch Dr. Henning Vogt und Dr. Oliver Böhnke, meinen Physik- und Chemielehrern, die in mir die Lust geweckt haben, ein naturwissenschaftliches Fach zu studieren. Zu guter letzt danke ich Emma dafür, dass sie meine beste Freundin ist.

Author's contributions

Publications in peer-reviewed journals

- [I] J. Neuhaus, S. C. Liebscher, L. Meckbach, T. Stroucken, and S. W. Koch, “Microscopic coulomb interaction in transition-metal dichalcogenides”, *Journal of Physics: Condensed Matter* **33**, 035301 (2020).
- [II] S. C. Liebscher, L. C. Bannow, J. Hader, J. V. Moloney, and S. W. Koch, “Extension of the LDA-1/2 method to the material class of bismuth containing III–V semiconductors”, *AIP Advances* **10**, 115003 (2020).
- [III] S. C. Liebscher, M. K. Hagen, J. artificial, J. V. Moloney, and S. W. Koch, “Microscopic theory for the incoherent resonant and coherent off-resonant optical response of tellurium”, *Phys. Rev. B* **104**, 165201 (2021).

Current Submissions to Peer-Reviewed Journals

- [IV] J. Hader, S. C. Liebscher, J. V. Moloney, and S. W. Koch, “Intrinsic carrier losses in tellurium due to radiative and auger recombinations”, *10.48550/ARXIV.2208.05532* (2022).

Talks and posters

- **S. C. Liebscher**, J. Hader, S. W. Koch, *High-Harmonic Generation in Tellurium*, Contributed Talk, GRK 1782 Seminar in Herborn, Germany (2021)
- **S. C. Liebscher**, L. C. Bannow, J. Hader, J. V. Moloney, S. W. Koch, *Extension of the LDA-1/2 method to the material class of bismuth containing III-V semiconductors*, Contributed Talk, GRK 1782 Seminar in Höchst i. Odw., Germany (2020)

Author's Contributions

- **S. C. Liebscher**, L. C. Bannow, S. W. Koch, *Ab initio-based calculation of the optical and electronic properties of dilute Ga(SbBi)*, Poster Presentation, GRK 1782 Seminar in Schwäbisch Gmünd, Germany (2019)

Original Contributions

The research that led to the results that are summarized in this thesis and published in Refs. [I–IV] have been regularly discussed in group meetings with my supervisor Stephan W. Koch and our collaborators. In the first project, whose results are published in Ref. [I] and discussed in Chapter 6 in this thesis, I derived the analytical expression for the Coulomb matrix element in a quasi-2d material system, ran DFT calculations on MoS₂ and wrote the code that uses the wave functions from the DFT calculation to calculate the Coulomb matrix element according to the derived formula.

The second project was initiated by the idea of Lars Bannow to extend the LDA-1/2 method to bismides by using an extrapolation and was done in cooperation with Jörg Hader and Jerome Moloney (both University of Arizona, Tucson, USA). The results are published in Ref. [II] and are presented in Chapter 4 in this thesis. While L. Bannow constructed the LDA-1/2 Pseudopotentials for the semiconductor compounds and extrapolated the cutoff radius to create the potentials for the Bismuth compounds, I used the potentials to perform DFT calculations. I created the SQS supercells for different Bi concentrations and geometries, performed the relaxations and the electronic calculations and unfolded the band structure. I analyzed the reduction of the band gap and the increase of the spin orbit splitting with increasing Bi content. I used codes provided by Jörg Hader to calculate the optical properties of the dilute bismides, namely the absorption, photoluminescence and Auger losses and analyzed them. I wrote most of the manuscript for the paper. I presented intermediate results of this project in the context of the research training group (RTG) 1782 as a poster presentation at the RTG seminar in 2019 in Schwäbisch Gmünd, Germany and a contributed talk at the RTG seminar in 2020 in Höchst i. Odw., Germany.

For the third project, where we again cooperated with Jörg Hader and Jerome Moloney, our interest was led towards Te as a material with interesting optical properties by Sergei Tochitsky (University of California, Los Angeles). The results are published in Ref. [III] and are summarized in Chapters 3 and 5 in this thesis. I used the shLDA-1/2 method to construct the pseudopotential for the DFT calculations on Te, performed the calculations and analyzed the band structure and dipole matrix elements. I used the output of the DFT calculations together with a code provided by J. Hader to perform calculations on the linear optical properties of Te. The idea to gauge the dipoles using a triple product came from J. Hader. I implemented the gauging procedure for the DFT data and performed local high harmonic generation calculations using a code provided by Maria Hagen. I analyzed the effect of the triple dipole product phase on the resulting spectra as well as the choice of different bands in the computational model. I also wrote parts of the manuscript. I presented the results of this study in a contributed talk at the RTG seminar 2021 in Herborn, Germany.

For the study of the losses in Te published in Ref. [IV], I provided the DFT calculations.

List of abbreviations

BZ	Brillouin zone
DFT	density functional theory
LDA	local density approximation
GGA	generalized gradient approximation
KS	Kohn-Sham
(sh)LDA-1/2	(shell) local density approximation minus one half
SIC	self interaction correction
TMDCs	transition-metal dichalcogenides
2D/3D	two-/three-dimensional
SBEs	semiconductor Bloch equations
SLEs	semiconductor luminescence equations
HHG	high harmonic generation
TDM	transition dipole moment
SO	spin-orbit
VdW	Van der Waals
SQS	special quasirandom structures
(T)PTG	(twisted) parallel transport gauge

During the ongoing global chip shortage[1], it has become painfully clear how much our technology nowadays depends on semiconductors and integrated circuits. Since the invention of the transistor in 1947 from semiconductor materials, they have become ubiquitous in all electronic devices and are now part of our daily lives in many visible and invisible ways. Their periodic crystal structure gives rise to energy bands on which the states are located which can be occupied by electrons. In the case of semiconductors there is an energy region between the occupied and unoccupied states in which no state is located, called the band gap. Many of the properties of a semiconductor are mainly influenced by the size of this energy gap. In order to absorb light, an electron has to be excited from an occupied to an unoccupied state, and therefore has to overcome the band gap, which means that only light with a specific wavelength can be efficiently absorbed. The reverse process, where excited electrons above the band gap relax back into an unoccupied state and emit energy in the form of light, is used e.g. in LEDs. Here, the size of the band gap determines the wavelength and thus the color that is emitted. The increasing scope of applications for semiconductor technology comes from our increasing understanding of their properties and the microscopic processes taking place within them. Scientists are always on the lookout for materials with certain beneficial properties for new specific applications.

This search can be made easier with a theoretical model that can predict the properties of novel materials before producing them. This is the aim of this thesis, in which an *ab initio* based methodology is used to study electronic and optical properties of different materials. The first part of this methodology is DFT, which is an established theory to calculate the electronic ground state of atoms, molecules and solids. It is an *ab initio* theory and does not rely on experimental inputs and is therefore well suited for our purpose. The second part of our methodology is the semiconductor Bloch theory. It is a quantum theory including many-body effects that can describe the dynamics of holes and electrons in a periodic crystal as well as radiative recombinations and Auger losses. Together, these two techniques allow us to perform studies on novel materials and predict their suitability for optoelectronic applications.

DFT is a powerful tool to study semiconductors, but there are well known short-

comings in getting accurate band gaps[2]. This is due to the inherent uncertainty in its formulation regarding the so called exchange-correlation energy. Different approximations to this energy have been proposed in an attempt to alleviate this shortcoming. In this thesis, we use the LDA-1/2 method, which aims to correct the band gap by including a self-energy correction. It has been shown to have an accuracy similar to the G_0W_0 method, which is computationally much more demanding[3].

For some applications, a material with a specific band gap is needed, e.g. for the efficient absorption of sun light in solar cells. Band-gap engineering is a process through which the band gap of a material can be altered by a desired amount by incorporating impurity atoms in the crystal structure. It has been found that the incorporation of Bi decreases the band gap of III-V semiconductors, e.g. the band gap of Ga(AsBi) is reduced by 88 meV / % Bi for low Bi concentrations[4]. The III-V semiconductors with a dilute amount of bismuth, called bismides, are speculated to be suitable for applications e.g. as semiconductor lasers, since the altered band gap can effectively reduce loss processes[5].

Another way of altering the properties of semiconductors is to change their dimensionality. Since the discovery of Graphene in 2004, the interest in 2D semiconductors has increased because they exhibit properties that are distinctly different from their bulk counterparts[6]. In the case of TMDCs, they change from an indirect to a direct band gap in the visible range when going from a bulk crystal to a monolayer, which makes them promising candidates for optical applications[7].

The microscopic theory used to describe the optical properties in this thesis is not limited to linear optical properties, but also applies to nonlinear regimes. One nonlinear effect of specific interest recently has been the generation of high harmonics in solids. While this effect has been known in atoms and molecules for a long time, it has been made possible in solids only recently[8] due to the realization of suitable light sources[9]. It is a nonlinear effect from strong-field non-resonant excitation with THz light, producing interband polarizations and intraband currents which cause the emission of multiples of the driving frequency across a broad spectral range. Recent studies have shown that in the calculation of HHG the phase of the TDMs cannot be neglected when the crystal does not exhibit inversion or reflection symmetry, since it leads to the emission of even harmonics in those cases[10]. This poses a challenge for our approach, since the TDMs calculated from DFT do not have a smooth phase relation across the BZ.

One material with exceptional nonlinear properties is Te. Its 2nd order nonlinear coefficient has been measured to be $d_{11} = 5 \cdot 10^{-9} \text{ m V}^{-1}$, making it one of the highest coefficients known[11]. It has an unusual crystal structure with helical chains in a trigonal symmetry perpendicular to the chain axis, making it sensitive to the light polarization direction[12]. So far, comprehensive studies of the suitability of Te for optical applications are scarce.

In the next chapter, the theoretical models used in this thesis are explained in more detail. First, the ideas behind DFT as well as some exchange-correlation functionals and the LDA-1/2 method are introduced, and then, the semiconductor Bloch approach is described. In Chapter 3, the application of DFT and the LDA-1/2 method to bulk Te

and GaSe to obtain their crystal structures and electronic properties is demonstrated. Furthermore, optical properties of Te are studied. In order to be used for calculations on bismides, the LDA-1/2 method has to be extended. This is done in Chapter 4, along with the introduction of necessary concepts to describe dilute materials with DFT. Subsequently, the electronic and optical properties of Ga(SbBi) as a representative of the dilute bismides are evaluated. In Chapter 5, we investigate ways of gauging the TDMs obtained from DFT and the effect of the phases on HHG calculations for Te and GaAs. In Chapter 6, the Coulomb interaction in monolayer TMDCs is studied by deriving an expression for the quasi-2D Coulomb matrix element. Finally, Chapter 7 concludes this thesis and sums up the findings.

Theoretical Background

The theoretical models used to obtain the results in this thesis are introduced in this chapter. First, DFT is introduced and some exchange-correlation functionals as well as the LDA-1/2 method presented, and second, the microscopic many-body theory which is used to calculate the optical properties is introduced.

2.1 Density Functional Theory

Density Functional Theory is an established method for calculating the electronic ground state of molecules and crystalline solids. The software package Vienna ab initio simulation package (VASP)[13–16] which employs the projector-augmented wave (PAW) method[17, 18] has been used to obtain the results presented in this thesis. The foundational theoretical principles of DFT are presented briefly in this section, a more thorough description can be found in numerous textbooks, e.g.[19].

In general, a many-body system like a solid is described by its wave function $\psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$, which depends on the coordinates of the electrons $\mathbf{r}_1, \dots, \mathbf{r}_N$. It can be obtained by solving the Schrodinger Equation $\mathcal{H}\psi = E\psi$ for the many-body Hamiltonian[20]

$$\mathcal{H} = \mathcal{T} + \mathcal{V} + \mathcal{U} = -\frac{\hbar^2}{2m_0} \sum_i^N \nabla_i^2 + \sum_i^N v(\mathbf{r}_i) + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.1)$$

where \mathcal{T} is the kinetic energy, \mathcal{V} is the external potential and \mathcal{U} is the Coulomb interaction, yielding

$$\left[-\frac{\hbar^2}{2m_0} \sum_i^N \nabla_i^2 + \sum_i^N v(\mathbf{r}_i) + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi = E\psi. \quad (2.2)$$

However, it is practically impossible to solve it exactly because of the many interactions between the particles. In DFT, the complexity of this problem is reduced and thus made feasible. Instead of looking at the many-body wave function, which grows

2 Theoretical Background

exponentially in complexity with increasing number of electrons, DFT considers the electron density $n(\mathbf{r})$, which only depends on the three spatial coordinates, regardless of the system size. The basis of DFT are the Hohenberg-Kohn-Theorems[21]. For a fixed number of electrons, the kinetic and Coulomb interaction part of the Hamiltonian stay fixed, and different systems differ only in the external potential. The first theorem states that the ground state density to a given external potential is unique, so that no other external potential will yield this particular ground state density. As a consequence, the potential energy $V[n]$ resulting from the external potential \mathcal{V} can be written as a functional of the ground state density,

$$V[n(\mathbf{r})] = \int d^3\mathbf{r} n(\mathbf{r})\mathcal{V}(\mathbf{r}). \quad (2.3)$$

Since the Coulomb interaction and kinetic energy part of the Hamiltonian stay fixed, the ground state energy is also defined by the ground state density and can be written as its functional. However, the kinetic energy and Coulomb interaction contain many-body contributions that cannot be written as a functional of the density explicitly. The kinetic energy consists of the correlation part $T_c[n]$ and the single-particle part $T_s[n]$,

$$T[n] = T_s[n] + T_c[n]. \quad (2.4)$$

The Coulomb interaction consists of the Hartree term $U_H[n]$ and the exchange term $U_x[n]$,

$$U[n] = U_H[n] + U_x[n]. \quad (2.5)$$

Only the single-particle part of the kinetic energy $T_s[n]$ and the Hartree term $U_H[n]$ of the Coulomb interaction can be written explicitly as a functional of the density,

$$T_s[n] = -\frac{\hbar^2}{2m_0} \sum_i^N \int d^3\mathbf{r} \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) \quad (2.6)$$

with single-particle wave functions $\phi_i(\mathbf{r})$, and

$$U_H[n] = \frac{e^2}{8\pi\epsilon_0} \int d^3\mathbf{r} \int d^3\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.7)$$

The correlation part of the kinetic energy and the exchange term of the Coulomb interaction are combined into the exchange-correlation energy, for which no explicit expression is known,

$$E_{xc}[n] = U_x[n] + T_c[n]. \quad (2.8)$$

Using this, the energy functional $E[n] = T[n] + V[n] + U[n]$ can be written as

$$E[n] = T_s[n] + V[n] + U_H[n] + E_{xc}[n]. \quad (2.9)$$

The second theorem states that the electron density that minimizes the total energy functional $E[n]$ is the ground state density $n_0(\mathbf{r})$ of the system, yielding the ground state energy $E_0[n_0]$. Therefore, it is possible to obtain the ground state density using

a variational principle with the minimization condition $0 = \frac{\delta E[n]}{\delta n(\mathbf{r})}$. Using the above expression for $E[n]$, it can be written as

$$0 = \frac{\delta E[n]}{\delta n(\mathbf{r})} = -\frac{\hbar^2}{2m_0} \nabla^2 + \mathcal{V}(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0} \int d^3\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \quad (2.10)$$

Together, the last three terms on the RHS can be interpreted as an effective potential, and the first term as the kinetic energy of a single-particle system. As a consequence, the equation effectively describes a non-interacting single-particle system. The resulting equations are called the Kohn-Sham equations[22],

$$\left(-\frac{\hbar^2}{2m_0} \nabla^2 + \mathcal{V}(\mathbf{r}) + \mathcal{U}_H(\mathbf{r}) + \mathcal{V}_{xc}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (2.11)$$

with the Hartree part of the Coulomb interaction $\mathcal{U}_H(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \int d^3\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$ and the exchange-correlation potential $\mathcal{V}_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}$. Their solutions are the Kohn-Sham orbitals $\phi_i(\mathbf{r})$ and eigenvalues ϵ_i . The electron density can be calculated as

$$n(\mathbf{r}) = \sum_i^N f_i |\phi_i(\mathbf{r})|^2, \quad (2.12)$$

with the fractional occupations f_i . Since this density defines a new Hamiltonian, the Kohn-Sham equations have to be set up with the new Hamiltonian and solved again, yielding a new density. The old and new charge density are then mixed and the process is started again. This so-called self-consistency cycle is started with a guess for the orbitals and repeated until the old and new density agree within a certain tolerance.

DFT is an *ab initio* method in the sense that only the atom species and positions of the structure have to be known for the calculation without the need for experimental parameters. After completing the calculation of the electronic ground state, also the forces acting on the atoms can be evaluated[23]. Using these forces, a structural relaxation can be performed, which moves the atoms according to the forces to get closer to the equilibrium geometry. This is repeated in several cycles of calculating the ground state of the new geometry and moving the atoms until a certain convergence criterion is met.

For periodic systems such as solids, it is reasonable to choose a plane wave basis set to represent the orbitals,

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\Omega^{1/2}} \sum_{\mathbf{G}} C_{\mathbf{G}n\mathbf{k}} e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}}, \quad (2.13)$$

with the band index n , momentum k , normalization volume Ω , reciprocal lattice vector \mathbf{G} and plane wave coefficient $C_{\mathbf{G}n\mathbf{k}}$. The expansion into plane waves is exact if an infinite amount of plane waves are included. However, since this is not feasible in a computation, the plane wave expansion is cut off at a specified cut-off energy E_{cut} including only plane wave coefficients which satisfy $\frac{1}{2}|\mathbf{G} + \mathbf{k}|^2 < E_{cut}$. The choice of

2 Theoretical Background

the convergence criterion for the self-consistency cycle and the cut-off energy for the plane wave set affect the accuracy of DFT calculations in practice. Additionally, since the exact expression of the exchange-correlation functional is unknown, the accuracy depends on a good approximation to it. Different approaches to such an approximation are a topic of ongoing and extensive research in the field of DFT[24]. The ideas behind the functionals used in this thesis are shortly presented in the following section.

2.1.1 Exchange-correlation functionals

The approximations to the exchange-correlation potential are divided into four groups with increasing sophistication and computational effort: local density approximations (LDA), generalized gradient approximations (GGA), meta-GGAs and hybrid functionals (HF). The ones used to obtain the results in this thesis will be described in the following subsections.

LDA

The Local Density Approximation depends only on the local value of the density $n(\mathbf{r})$. It is divided into an exchange E_x and a correlation E_c energy. For the exchange energy, the analytical expression for the exchange energy density of a homogenous electron gas with the given density $n(\mathbf{r})$ is used.

For the correlation part, only high and low limits are known[25], so that an interpolation between those limits is used. The particular parametrization by Perdew *et al.* [26] is used in this thesis. LDA is known to underestimate band gaps[27]. It is used in this thesis only for the relaxation of the supercells in section 4.1 and in combination with the LDA-1/2 method for consistency between structural relaxation and electronic calculation.

PBE (GGA)

In addition to the local value of the density, the Generalized Gradient Approximation uses also the gradient of the density at the same point, $\nabla n(\mathbf{r})$. In order to distinguish it from the LDA, the GGA is often referred to as semi-local. It does not contain any contributions from a point $\mathbf{r}' \neq \mathbf{r}$, however, it does contain the behavior of the density outside of the point $n(\mathbf{r})$ indirectly through its derivative. The GGA functional used in this thesis is a parametrization which is solely based on fundamental constants and no empirical input done by Perdew, Burke and Ernzerhof [28].

PBEsol (GGA)

PBEsol is a revised version of the PBE functional with empirical modifications in order to improve equilibrium properties specifically for solids [29].

HSE (HF)

Hybrid Functionals use an exact expression for the exchange energy that is derived from Hartree-Fock theory[30]. This energy is mixed with the exchange-correlation energy from a different approximation, which can be a LDA, GGA or meta-GGA. The calculation of the Hartree-Fock exchange requires the Kohn-Sham orbitals from the self-consistency cycle, so that this approximation is not a functional of the density alone. The HSE functional uses an error-function-screened Coulomb potential to calculate the exchange portion of the energy in order to improve computational efficiency, especially for metallic systems[31].

2.1.2 LDA-1/2 method

The LDA-1/2 method does not represent an exchange-correlation functional itself but is a correction that is used to improve other functionals. The name indicates that it uses a LDA functional but strips half an electron from it (“minus one half”). It can also be used with a GGA functional and will then sometimes be referred to as GGA-1/2. It has been developed by Ferreira *et al.* in 2008[32, 33]. In DFT, the energy difference between the highest occupied and lowest unoccupied KS orbital is not actually the true band gap. Instead, since DFT solves the non-interacting Kohn-Sham equations, from which the KS eigenvalues are derived, the physical meaning of this KS band gap changes. The true band gap, however, is the difference in total energy between the ground state and the excited state with one electron missing in the valence band and one added to the conduction band. The true band gap differs from the KS band gap by the derivative discontinuity Δ_{xc} [34, 35]. The interpretation of this discontinuity in the scope of the LDA-1/2 method is that it represents the non-null self-energy of the localized hole that is created when moving to the excited state[33]. In contrast to other methods of correcting for the self-energy, e.g. the GW method, the LDA-1/2 method stays within the DFT framework and uses an approach based on Slater’s half-occupation technique[36]. This technique has been used to get correct ionization energies for atoms and molecules and is based on the Janak theorem,

$$\frac{\partial E}{\partial f_\alpha} = e_\alpha(f_\alpha), \quad (2.14)$$

from which it derives that the ionization energy is the energy of the highest occupied KS orbital at half occupation,

$$E_\alpha(0) - E_\alpha(-1) = \int_{-1}^0 \frac{\partial E}{\partial f_\alpha} df_\alpha = e_\alpha(-1/2). \quad (2.15)$$

2 Theoretical Background

In the LDA-1/2 method, the same idea is applied to crystals. The difference between the ionization energy and the KS eigenvalue can be derived as a self-energy

$$\begin{aligned}
S_\alpha = & \frac{1}{2} \iint d^3r d^3r' \frac{n_\alpha(\mathbf{r})n_\alpha(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{1}{2} \iint d^3r d^3r' n_\alpha(\mathbf{r}) \frac{\delta^2 E_{xc}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} n_\alpha(\mathbf{r}') \\
& + \frac{1}{2} \iint d^3r d^3r' \frac{n_\alpha(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \sum_\beta f_\beta \frac{\partial n_\beta(\mathbf{r}')}{\partial f_\alpha} + \frac{1}{2} \iint d^3r d^3r' n_\alpha(\mathbf{r}) \frac{\delta^2 E_{xc}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} \sum_\beta f_\beta \frac{\partial n_\beta(\mathbf{r}')}{\partial f_\alpha},
\end{aligned} \tag{2.16}$$

which can be written as the quantum-mechanical average of a self-energy potential. In the ground state, the electrons of a solid are described as completely delocalized Bloch functions, so that there is no self-energy. When exciting a hole in the valence band, however, the localized hole has a finite self-energy. This self-energy potential is obtained as the difference between the KS potentials of the system at neutral and at half-ionized occupation. Approximately, the self-energy potential of an extended system can be obtained as the sum of the potentials of the atoms. The self-energy potential can be added to the KS potential of the system before the DFT calculation, and the calculation will yield the KS eigenvalue at half occupation, which is the self-energy-corrected band gap. Therefore, the LDA-1/2 method offers a computationally efficient way of implementing the self-energy correction with the same computational cost as a standard LDA calculation. Since the stripping of half an electron from the atomic KS potential leads to a long-ranged change in the potential, in an infinite system these changes would add up to infinity. Therefore, in the LDA-1/2 method, the self-energy potential is added using a step function,

$$\Theta(r) = \begin{cases} \left[1 - \left(\frac{r}{r_{\text{cut}}}\right)^n\right]^3 & r \leq r_{\text{cut}} \\ 0 & r > r_{\text{cut}} \end{cases} \tag{2.17}$$

with the cutoff radius r_{cut} , thus limiting the correction to a certain spherical region around the atom. The exponential factor n is used to smoothen the edge of the step function. As proposed by Xue *et al.*[37], we choose $n = 20$ for our calculations. The correct cutoff radius for the step function is determined by the condition that it yields the largest band gap, so that it can be found by varying the cutoff radius r_{cut} . Physically, this means that the most part of the localized hole has to be located inside the spherical step function. In some materials, the excited hole is not localized around the atom core. Then, it is more accurate to use a shell-like step function with an inner and an outer cutoff radius r_{in} and r_{out} , respectively,

$$\Theta(r) = \begin{cases} \left[1 - \left(\frac{r}{r_{\text{out}}}\right)^m\right]^3 \frac{1 + \tanh\left[\frac{n(r-r_{\text{in}})}{2}\right]}{2} & r \leq r_{\text{out}} \\ 0 & r > r_{\text{out}} \end{cases}. \tag{2.18}$$

Both radii are again found by variation until the band gap reaches a maximum. When this shell function is used, the method is called shLDA-1/2.

2.2 Microscopic Theory for Optical Properties

2.2.1 Semiconductor Bloch Equations

The optical properties of the materials studied in this thesis are obtained from a fully microscopic many-body quantum theory[38–41]. It uses the electron-hole picture in second quantization, where the creation operator $a_{\lambda,k}^\dagger$ creates an electron with the electron band index λ and momentum k , and the destruction operator $a_{\lambda,k}$ destroys it accordingly. Similarly, holes with the hole band index ν and momentum k are created and destroyed using the operators $a_{\nu,k}^\dagger$ and $a_{\nu,k}$, respectively. The starting point is the Hamiltonian of an interacting many-body electron system,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_C + \mathcal{H}_{l-m}. \quad (2.19)$$

The first term is the kinetic energy of the electrons and holes in a solid,

$$\mathcal{H}_0 = \sum_{k,\lambda} \varepsilon_k^\lambda a_{\lambda,k}^\dagger a_{\lambda,k} + \sum_{k,\nu} \varepsilon_k^\nu a_{\nu,k}^\dagger a_{\nu,k}, \quad (2.20)$$

where ε_k^λ and ε_k^ν are the single-particle energies of electrons and holes, respectively. The single-particle energy of the holes also contains the Coulomb exchange correction according to $\varepsilon_k^\nu \equiv -E_{\nu,k}^\nu + \sum_{q \neq 0} V_q$, with the valence band energies $E_{\nu,k}^\nu$ and the Coulomb matrix element V_q . The second term is the Coulomb interaction between charge carriers,

$$\begin{aligned} \mathcal{H}_C = & \frac{1}{2} \sum_{k,k',q \neq 0, \lambda, \lambda'} V_q a_{\lambda,k+q}^\dagger a_{\lambda',k'-q}^\dagger a_{\lambda',k'} a_{\lambda,k} + \frac{1}{2} \sum_{k,k',q \neq 0, \nu, \nu'} V_q a_{\nu,k+q}^\dagger a_{\nu',k'-q}^\dagger a_{\nu',k'} a_{\nu,k} \\ & - \sum_{k,k',q \neq 0, \lambda, \nu} V_q a_{\lambda,k+q}^\dagger a_{\nu,k'-q}^\dagger a_{\nu,k'} a_{\lambda,k}, \end{aligned} \quad (2.21)$$

where V_q is the Coulomb matrix element and q is the momentum that is exchanged in the process. The third term of the Hamiltonian contains the light-matter interaction,

$$\mathcal{H}_{l-m} = -E(t) \sum_{k,\lambda,\nu} \left(d_k^{\lambda\nu} a_{\lambda,k}^\dagger a_{\nu,-k}^\dagger + \text{h.c.} \right), \quad (2.22)$$

that describes the creation of electron-hole pairs by the applied electric field $E(t)$. The coupling between two bands λ and ν at momentum k is given by the transition dipole matrix element $d_k^{\lambda\nu}$. From this Hamiltonian, the Heisenberg equation of motion,

$$\frac{\partial}{\partial t} \langle \mathcal{O} \rangle = \frac{i}{\hbar} \langle [\mathcal{H}, \mathcal{O}] \rangle, \quad (2.23)$$

can be used to derive the dynamics of the microscopic quantities. The most important expectation values are the electron occupation $f_k^\lambda = \langle a_{\lambda,k}^\dagger a_{\lambda,k} \rangle$, the hole occupation $f_k^\nu = \langle a_{\nu,-k}^\dagger a_{\nu,-k} \rangle$ and the interband polarization $p_k^{\nu\lambda} = \langle a_{\lambda,k}^\dagger a_{\nu,k}^\dagger \rangle$. The EOM of one of these quantities couples to the dynamics of a higher-order term, leading to the

2 Theoretical Background

hierarchy problem, which means that no closed set of coupled equations can be found. This issue can be resolved using the cluster-expansion approach[41], in which the N -particle contributions are recursively separated into factorizations and correlations. Then, the operator dynamics can be truncated at the Hartree-Fock level, yielding the Semiconductor Bloch Equations

$$\frac{\partial}{\partial t} p_k^{\nu\lambda} = \frac{1}{i\hbar} \sum_{\lambda', \nu'} \left(\epsilon_k^{\nu, \nu'} \delta_{\lambda\lambda'} + \epsilon_k^{\lambda, \lambda'} \delta_{\nu\nu'} \right) p_k^{\nu'\lambda'} - i \left(f_k^\lambda + f_k^\nu - 1 \right) \Omega_k^{\lambda\nu} + \left. \frac{\partial}{\partial t} p_k^{\nu\lambda} \right|_{\text{scatt}}, \quad (2.24)$$

$$\frac{\partial}{\partial t} f_k^\lambda = -2 \text{Im} \left\{ \sum_{\nu} \Omega_k^{\lambda\nu} \left(p_k^{\nu\lambda} \right)^* \right\} + \left. \frac{\partial}{\partial t} f_k^\lambda \right|_{\text{scatt}}, \quad (2.25)$$

$$\frac{\partial}{\partial t} f_k^\nu = -2 \text{Im} \left\{ \sum_{\lambda} \Omega_k^{\lambda\nu} \left(p_k^{\nu\lambda} \right)^* \right\} + \left. \frac{\partial}{\partial t} f_k^\nu \right|_{\text{scatt}}. \quad (2.26)$$

They contain the renormalized energies

$$\epsilon_k^{\lambda, \lambda'} = \epsilon_k^\lambda \delta_{\lambda\lambda'} - \sum_{\lambda'', q} V_{k-q}^{\lambda\lambda''\lambda'\lambda''} f_q^{\lambda''}, \quad (2.27)$$

$$\epsilon_k^{\nu, \nu'} = \epsilon_k^\nu \delta_{\nu\nu'} - \sum_{\nu'', q} V_{k-q}^{\nu'\nu''\nu\nu''} f_q^{\nu''}, \quad (2.28)$$

and the renormalized Rabi frequency,

$$\Omega_k^{\lambda\nu} = \frac{1}{\hbar} \left(d_k^{\lambda\nu} E(t) + \sum_{\lambda', \nu', q} V_{k-q}^{\lambda\nu'\nu\lambda'} p_q^{\nu'\lambda'} \right). \quad (2.29)$$

The scattering terms represent the higher-order correlations from the cluster expansion. They contain the electron-electron and electron-phonon scattering, which are included on a fully microscopic level by solving quantum-Boltzmann type scattering equations. Their effect is the dephasing of the polarization and the homogeneous broadening of the spectrum.

After obtaining the microscopic polarizations, the macroscopic polarization is calculated according to

$$P(t) = \frac{1}{\mathcal{S}} \sum_{k, \nu, \lambda} (d_k^{\lambda\nu})^* p_k^{\nu\lambda}(t) + \text{c.c.} \quad (2.30)$$

A Fourier transformation gives the frequency dependent macroscopic polarization $P(\omega)$, from which the absorption coefficient

$$\alpha(\omega) = \frac{\omega}{n_0 c_0 \epsilon_0} \text{Im} \left(\frac{P(\omega)}{E(\omega)} \right) \quad (2.31)$$

can be calculated with the background refractive index n_0 , the speed of light in vacuum c_0 and the the vacuum permittivity ϵ_0 .

When calculating the linear absorption, the material is assumed to be unexcited and the field weak enough so that no carrier populations are created. When using these equations to calculate gain spectra, the occupations are non-zero and the carriers are assumed to be in thermal equilibrium, so that they can be described by Fermi distributions.

2.2.2 Semiconductor Luminescence Equations

In order to describe the incoherent process of spontaneous light emission from semiconductors, a fully quantized description of the light field has to be included in the many-body Hamiltonian. A photon with momentum q is created or destroyed with the bosonic operators B_q^\dagger and B_q , respectively. The free-particle energy of the photons,

$$H_L = \sum_q \hbar\omega_q \left(B_q^\dagger B_q + \frac{1}{2} \right), \quad (2.32)$$

is added to the Hamiltonian, where ω_q is the optical frequency. The semi-classical light-matter interaction is replaced by a fully quantized interaction Hamiltonian in second quantization, which for one electron band e and one hole band h reads

$$H_Q = -i \sum_{q,k} d_k^{eh} \mathcal{E}_q u_q a_{e,k+q}^\dagger a_{h,-k}^\dagger B_q + \text{h.c.}, \quad (2.33)$$

with the mode function u_q and the vacuum field amplitude \mathcal{E}_q . In incoherent conditions, the spontaneous emission is given by the correlation of the photon number, $\Delta\langle B_q^\dagger B_q \rangle$. To obtain its time dynamics, the Heisenberg equation of motion is used. Again using the cluster-expansion approach to truncate correlations beyond the second Born approximation, the time dynamics for the more general operator

$$\frac{\partial}{\partial t} \Delta\langle B_q^\dagger B_{q'} \rangle = -i(\omega_{q'} - \omega_q) \Delta\langle B_q^\dagger B_{q'} \rangle + \sum_k \left[\mathcal{F}_{k,q}^{eh} (\Pi_{k,q'})^* + (\mathcal{F}_{k,q'}^{eh})^* \Pi_{k,q} \right] \quad (2.34)$$

are obtained. This expression directly couples to the photon-assisted polarization $\Pi_{k,q} = \Delta\langle B_q^\dagger a_{h,-k+q} a_{e,k} \rangle$, whose time dynamics are derived in Ref. [41]. These coupled differential equations are called the SLEs and can be solved numerically. The photoluminescence is then calculated as the change of the number of photons in a light mode q ,

$$S(\omega_q) = \frac{\partial}{\partial t} \Delta\langle B_q^\dagger B_q \rangle = 2 \text{Re} \left[\sum_k \mathcal{F}_{k,q}^{eh} (\Pi_{k,q})^* \right]. \quad (2.35)$$

2.2.3 Auger Losses

Auger recombination refers to a process where an electron and a hole recombine non-radiatively and the energy is transferred to a third carrier, either an electron or a hole. Impact ionization is the reverse process and describes the creation of an electron-hole pair for which the energy is taken from a third carrier. The Auger losses are calculated by including the terms for Auger recombination and impact ionization[42],

$$\begin{aligned} \mathcal{H}_C^{\text{aug}} = & 2 \sum_{\substack{k,k',q \neq 0, \\ \lambda_1, \nu_1, \nu_2, \nu_3}} \left(a_{\lambda_1, k+q}^\dagger a_{\nu_1, k'-q}^\dagger a_{\nu_2, -k}^\dagger a_{\nu_3, k'} \tilde{V}_q^{\lambda_1 \nu_1 \nu_2 \nu_3} + a_{\nu_3, k+q}^\dagger a_{\nu_2, -k'+q}^\dagger a_{\nu_1, k} a_{\lambda_1, k'} \tilde{V}_q^{\nu_2 \nu_1 \nu_3 \lambda_1} \right) \\ & - 2 \sum_{\substack{k,k',q \neq 0, \\ \lambda_1, \lambda_2, \lambda_3, \nu_1}} \left(a_{\lambda_1, k+q}^\dagger a_{\nu_1, -k'+q}^\dagger a_{\lambda_2, k} a_{\lambda_3, k'} \tilde{V}_q^{\nu_1 \lambda_1 \lambda_2 \lambda_3} + a_{\lambda_1, k+q}^\dagger a_{\lambda_2, k'-q}^\dagger a_{\nu_1, -k} a_{\lambda_3, k'} \tilde{V}_q^{\lambda_1 \lambda_2 \lambda_3 \nu_1} \right), \end{aligned} \quad (2.36)$$

into the Hamiltonian. The processes are caused by the Coulomb interaction, which is screened by the dielectric function. The static Lindhard formula is solved to obtain the dielectric function. To calculate the Auger losses, the carrier dynamics are solved from the EOM. The change of the carrier occupations is the Auger loss rate,

$$J_{\text{aug}} = e \sum_{k,\lambda} \frac{\partial}{\partial t} f_k^\lambda = e \sum_{k,\nu} \frac{\partial}{\partial t} f_k^\nu. \quad (2.37)$$

The Auger coefficient is defined as

$$C = \frac{J_{\text{aug}}}{e n^3 w}, \quad (2.38)$$

with the carrier density n .

2.2.4 High Harmonic Generation

The generation of high harmonics in semiconductors is a nonlinear optical process that occurs when exciting a semiconductor with intense laser light. The emitted radiation is coherent and exhibits a characteristic pattern of multiples of the exciting frequency.

When exciting a semiconductor with a strong THz light pulse, the excitation energy is not resonant with the typical semiconductor band gap. Therefore, the amount of excited charge carriers is comparatively small. However, the non-resonant excitation can still lead to polarizations between the bands and excite carriers in a multi-photon process, and, in contrast to typical resonant excitations, between different band combinations, since there is no resonant transition that dominates. Furthermore, the strong long-wavelength excitation accelerates carriers along the bands, leading to Bloch oscillations across the whole BZ. The high harmonic emission is caused by the coupled dynamics of interband polarizations and intraband currents. The SBEs can be used to describe the HHG with a driving field $E_{\text{THz}}(t)$. The Coulomb interaction in this situation is negligible[43]. The SBEs then take the form for the interband polarization between valence band h_i and conduction band e_j

$$\begin{aligned} i\hbar \frac{d}{dt} p_{\mathbf{k}}^{h_i e_j} &= \left(\epsilon_{\mathbf{k}}^{e_j} + \epsilon_{\mathbf{k}}^{h_i} + i|e|E_{\text{THz}}(t)\nabla_{\mathbf{k}} \right) p_{\mathbf{k}}^{h_i e_j} \\ &\quad - \hbar\Omega_{\mathbf{k}}^{h_i e_j}(t) \left(1 - f_{\mathbf{k}}^{e_j} - f_{\mathbf{k}}^{h_i} \right) + \Gamma_{\mathbf{k}}^{h_i e_j} \\ &\quad + \sum_{e_\lambda \neq e_j} \left[\hbar\Omega_{\mathbf{k}}^{h_i e_\lambda}(t) p_{\mathbf{k}}^{e_\lambda e_j} - \hbar\Omega_{\mathbf{k}}^{e_\lambda e_j}(t) p_{\mathbf{k}}^{h_i e_\lambda} \right] \\ &\quad + \sum_{h_\lambda \neq h_i} \left[\hbar\Omega_{\mathbf{k}}^{h_i h_\lambda}(t) p_{\mathbf{k}}^{h_\lambda e_j} - \hbar\Omega_{\mathbf{k}}^{h_\lambda e_j}(t) p_{\mathbf{k}}^{h_i h_\lambda} \right] \\ &\quad + \left. \frac{d}{dt} p_{\mathbf{k}}^{h_i e_j} \right|_{\text{corr}}, \end{aligned} \quad (2.39)$$

and similar expressions for the polarizations between valence bands $p_{\mathbf{k}}^{h_i h_j}$ and between conduction bands $p_{\mathbf{k}}^{e_i e_j}$, where $\Omega^{h_i e_j}$ are the renormalized Rabi frequencies. For the

carrier occupations, one obtains for the electron occupation

$$\begin{aligned} \hbar \frac{d}{dt} f_{\mathbf{k}}^{e_i} = & -2\hbar \times \\ & \times \text{Im} \left[\sum_{e_\lambda \neq e_i} \Omega_{\mathbf{k}}^{e_\lambda e_i}(t) (p_{\mathbf{k}}^{e_\lambda e_i})^* + \sum_{h_\lambda} \Omega_{\mathbf{k}}^{h_\lambda e_i}(t) (p_{\mathbf{k}}^{h_\lambda e_i})^* \right] \\ & + |e| E_{\text{THz}}(t) \nabla_{\mathbf{k}} f_{\mathbf{k}}^{e_i} + \Gamma_{\mathbf{k}}^{e_i}, \end{aligned} \quad (2.40)$$

and a similar expression for the hole occupation $f_{\mathbf{k}}^{h_i}$ [44]. After solving the time dynamics, the macroscopic polarization

$$P(t) = \sum_{\lambda, \lambda', k} d_k^{\lambda \lambda'} p_k^{\lambda \lambda'} \quad (2.41)$$

and the macroscopic current

$$J(t) = \sum_{\lambda, k} j_\lambda(k) f_k^\lambda \quad (2.42)$$

can be calculated. From these quantities, it is possible to calculate the emitted electric field,

$$E_{\text{out}}(t) \propto \frac{\partial}{\partial t} P(t) + J(t), \quad (2.43)$$

as well as the emission intensity,

$$I_{\text{out}}(\omega) \propto |\omega P(\omega) + iJ(\omega)|^2. \quad (2.44)$$

Application of the LDA-1/2 method to Te and GaSe

3.1 Te

Tellurium is a semiconductor with helical chains that are Van der Waals bonded to each other, while the atoms in the chains themselves are covalently bonded[45]. Perpendicular to the chain axis, it exhibits a trigonal symmetry. Due to the helical structure, the optical properties are different for the excitation directions parallel to the helix axis, $E \parallel c$, and perpendicular to it, $E \perp c$. It is interesting to study because of its large second order nonlinearity coefficient[46]. We calculate the electronic properties of Te in this section using the shLDA-1/2 method. Afterwards, the linear optical properties are calculated to validate the approach by comparison to experiment. Later, in section 5.1, the data obtained in this section is used to study the nonlinear optical properties of Te in HHG calculations.

First, the structure was set up according to the symmetry group of Te and a structural relaxation performed. The computational details of the calculation can be found in Paper [III]. Using the PBE functional, a good agreement of the structural parameters with experimental data could be achieved, yielding the lattice constants $a = 4.51 \text{ \AA}$ and $c = 5.96 \text{ \AA}$, as compared to the experimental values $a = 4.46 \text{ \AA}$ and $c = 5.92 \text{ \AA}$ [47]. Because of the different nature of the bonds between the Te atoms, both the LDA-1/2 and the shLDA-1/2 method have been used for the electronic calculation to find the method that yields better results. For the LDA-1/2 method, only the outer cutoff radius r_{out} has to be found. After a first coarse scanning for the optimal cutoff radius from $2.6 - 4.0 \text{ \AA}$ in steps of 0.2 \AA to find the approximate location of the maximum, the region from $3.2 - 3.6 \text{ \AA}$ has been scanned with a step width of 0.05 \AA . The resulting band gaps as a function of the cutoff radius and the quadratic interpolation used to find the maximum is shown in Fig. 3.1 for the outer and the inner cutoff radius. The interpolation has been done using only the data points in the second scanning region, since the behavior farther away from the optimal cutoff radius does not follow a quadratic function. The optimized cutoff radius is found as $r_{\text{out}} = 3.395 \text{ \AA}$. For the shLDA-1/2 method, the outer cutoff radius is taken from the LDA-1/2 method and is kept fixed while varying the inner cutoff radius r_{in} . Again, first a broader region has

3 Application of the LDA-1/2 method to Te and GaSe

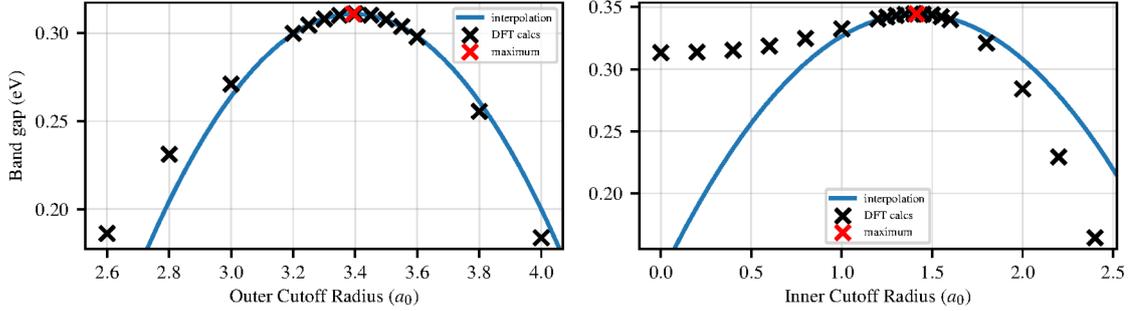


Figure 3.1: Band gaps obtained from DFT calculations on Te with different cutoff radii, (a) with the LDA-1/2 method, varying the outer cutoff radius, and (b) with the shLDA-1/2 method, varying the inner cutoff radius. The blue line is a quadratic interpolation to find the maximum.

been scanned with a larger step size, and then in the vicinity of the maximum more data points with a finer step width have been calculated and used for the interpolation. From the interpolation, we obtained the optimized cutoff radius $r_{\text{in}} = 1.328 \text{ \AA}$. The optimized band gaps obtained with these cutoff radii are $E_g = 0.29 \text{ eV}$ for the LDA-1/2 method and $E_g = 0.32 \text{ eV}$ for the shLDA-1/2 method. Using the shLDA-1/2 method, it is possible to get closer to the experimental value of $E_g = 0.33 \text{ eV}$ [48], which means that the valence hole is located mostly between the Te atoms. Therefore, the results from the shLDA-1/2 method will be used for all further calculations. However, the shLDA-1/2 method assumes in its formulation a rotational symmetry of the spherical step function, which is not necessarily accurate for the case of Te, where the intra-chain bonds are different from the inter-chain bonds. After having obtained the optimized cutoff radii, the electronic band structure and interaction matrix elements from DFT can be used for the microscopic calculations. The TDMs of Te are shown in Fig. 3.2. At the direct band gap at the H-point, the TDMs between the two highest valence bands and the two lowest conduction bands have a maximum in the $E \perp c$ direction, while the TDMs vanish directly around the H-point in the $E \parallel c$ direction.

First, the absorption is calculated according to 2.2.1 and shown in Fig. 3.3. It is found that the absorption is stronger for the $E \perp c$ direction and also starts at a lower energy. This is due to the vanishing coupling of the two topmost valence bands to the lowest conduction bands in the $E \parallel c$ direction. The experimental data points in the figure taken from Ref. [49] show a similar behaviour as our theoretical results, giving credence to our approach. Because of the good agreement, the input data from DFT has also been used to study the radiative and Auger losses in Te in Paper [IV].

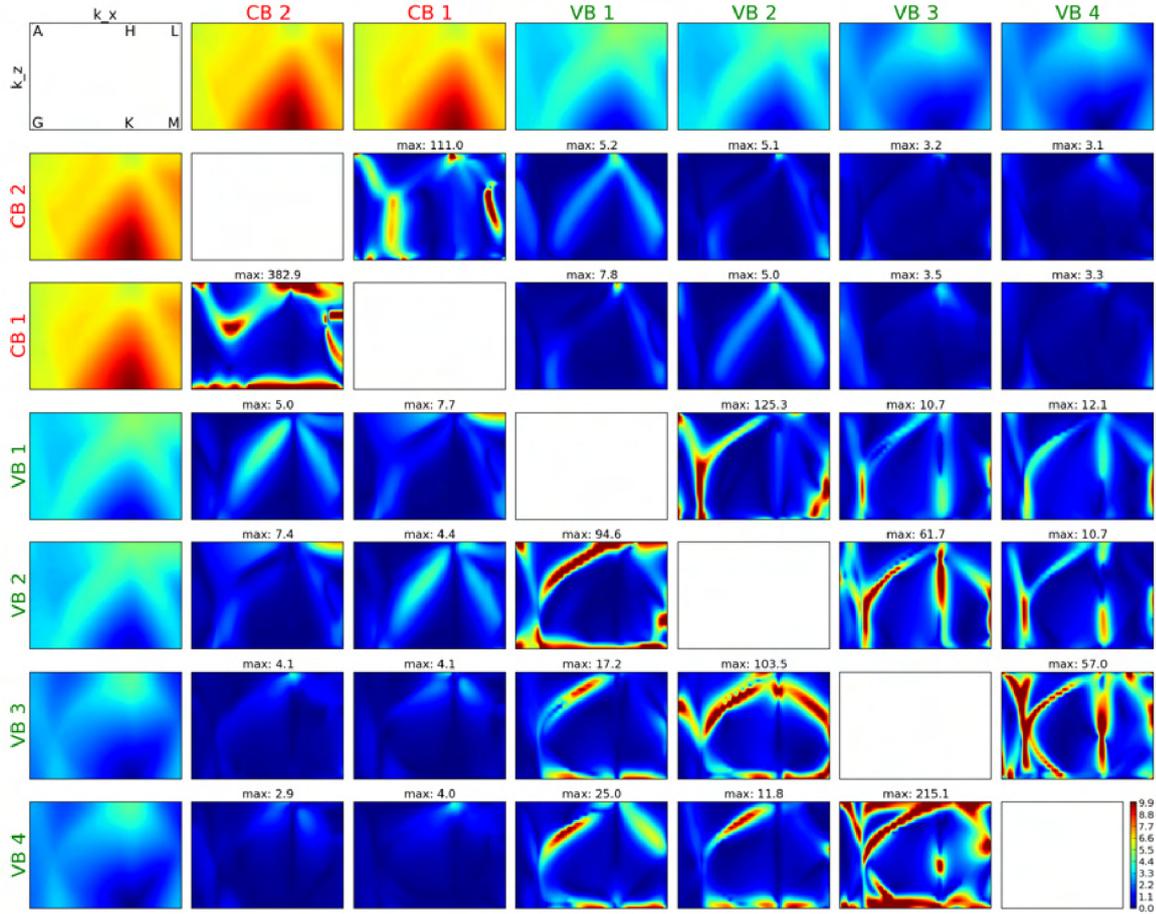


Figure 3.2: The first row and first column show the dispersions of the energy bands in a 2d plot. In the inner cells, the TDMs between the respective bands in that particular row and column are shown. In the part below the empty diagonal, the TDMs in $E \parallel c$ direction are given, while the part above the diagonal shows the TDMs in $E \perp c$ direction. The color map given in the bottom right corner applies to all TDM plots. The values of the TDMs vary over a wide range of values, so that the color map is capped at $10 e\text{\AA}$ and values higher than that, which occur mainly in the intervalence band and interconduction band TDMs, are not distinguished. The location of high symmetry points in the 2d plane is indicated in the top left corner.

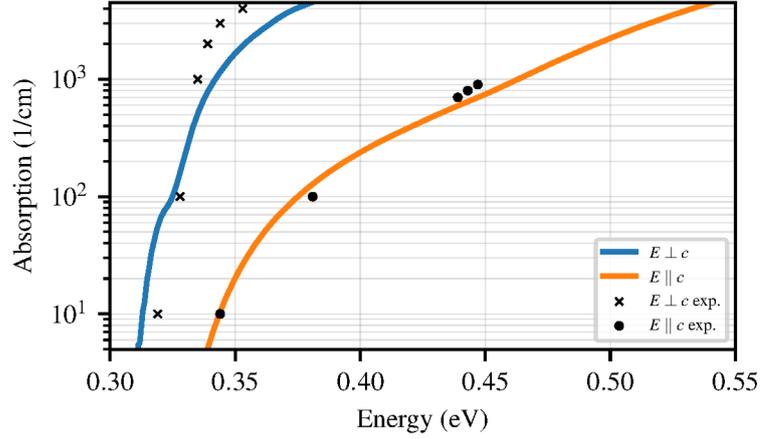


Figure 3.3: Absorption calculated from SBEs using input data from DFT calculations with the shLDA-1/2 method for Te in two polarization directions. Experimental data points taken from Ref. [49]

3.2 GaSe

Bulk ϵ -GaSe consists of Van der Waals bonded monolayers. To get a good prediction for the structural parameters, Van-der-Waals corrections have to be included in the DFT calculations. The result which is closest to experimental values was obtained with the DFT-D2 VdW correction by Grimme[50], yielding the lattice constants $a = 3.746 \text{ \AA}$ and $c = 15.94 \text{ \AA}$. For the structural relaxation and the electronic ground state calculation, a k -point grid of $12 \times 12 \times 2$ has been used. The cutoff energy was set to $E_{cut} = 500 \text{ eV}$ and the convergence criterion for electronic minimization to 10^{-8} eV . The closest agreement with experimental results for the direct band gap could be achieved using the GGA-1/2 method with a cutoff radius of $r_{cut} = 3.062 \text{ \AA}$, yielding a band gap of $E_g = 1.92 \text{ eV}$, which is still short of the experimental value $E_g = 2.1 \text{ eV}$ [51].

Electro-Optical Properties of Dilute Bismides

In this chapter, the LDA-1/2 method will be used and extended to perform calculations on semiconductors with a dilute amount of bismuth. Most of the results in this chapter are published in Paper [II].

4.1 LDA-1/2 method for Bismuth compounds

As explained in Section 2.1.2, the LDA-1/2 method requires finding the cutoff radius for the self-energy correction to the pseudopotential used in the DFT calculation. The radius is found variationally by maximizing the band gap that is obtained in a calculation with the modified potential. We found the cutoff radius and calculated the optimized band gap for zinc-blende structure compounds of Al, Ga and In with the anions N, P, As, Sb. The self-energy correction only has to be applied to the anions, since it was found by Ferreira *et al.*[33] that the effect of applying the self-energy correction to the ion on the band gap is negligible. An overview of the lattice constants and the optimal cutoff radii for the studied compounds is given in Paper [II]. A comparison of the resulting band gaps with literature values can be found in Fig. 4.1.

The III-Bi compounds are semi-metals, not semiconductors. Therefore, it is not possible to maximize the band gap and find the cutoff radius as is necessary for the LDA-1/2 method. To make the LDA-1/2 method usable for the bismuth compounds, a different method of finding the cutoff radius has to be used. Our approach is based on the observation that the cutoff radius and the covalent radius of the anions are not independent of each other. The cutoff radius of the anions in compounds with the same ion is increasing almost linearly with their covalent radius. Therefore, it is possible to fit the cutoff radius as a linear function of the covalent radius using the compounds which are semiconductors. Afterwards, using the covalent radius of Bi, the function can be extrapolated to give a value for the cutoff radius of Bi in the III-Bi compound, as shown in Fig. 4.2.

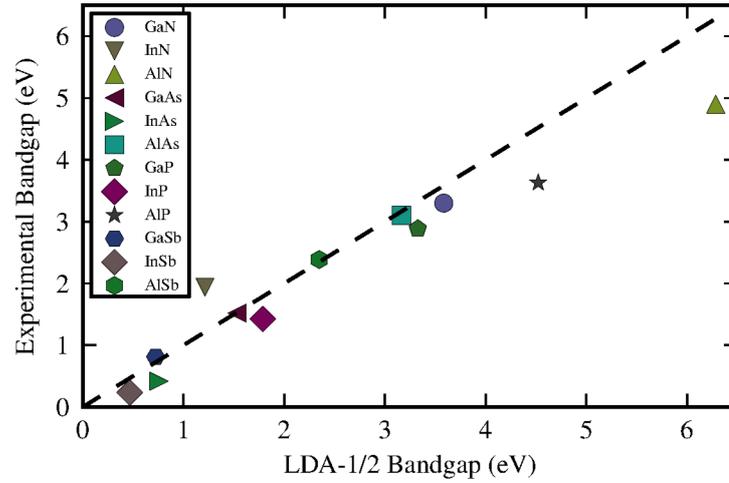


Figure 4.1: Comparison of theoretical (x -value) and experimental (y -value) values for the band gaps at the Γ -point of several III-V semiconductors. Theoretical values are calculated with the LDA-1/2 method, experimental values for $T = 0\text{ K}$ are taken from Ref. [52].

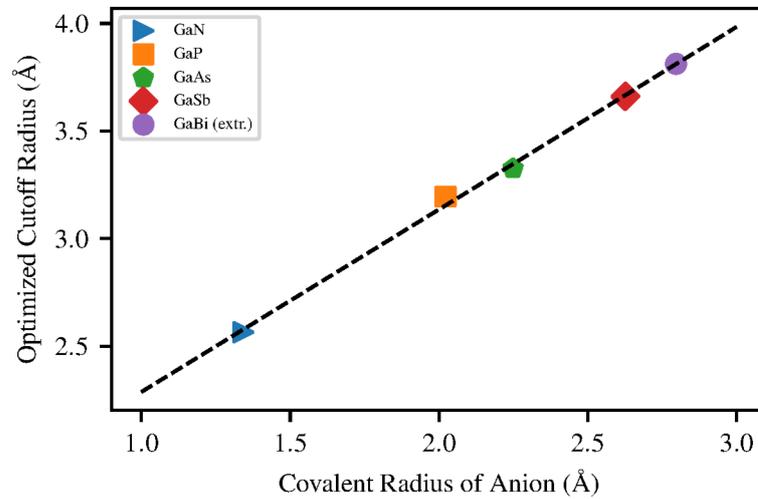


Figure 4.2: Optimized cutoff radii from the LDA-1/2 method for Ga-V compounds plotted against the covalent radii of the V elements. The dashed line is a linear fit from which the extrapolation to GaBi is obtained.

4.2 Dilute materials

Periodic boundary conditions as used in most DFT codes make it easy to describe an infinite, perfect crystal using only a small primitive cell. However, dilute materials have a small percentage of impurity atoms incorporated in their crystal structure, which usually substitute the species from the same group of the compound. These imperfections cannot be described by repeating a primitive cell. Instead, reaching a low percentage requires a high number of atoms. This can be achieved by repeating the primitive cell in space into supercells and replacing a small amount of atoms in them. Due to computational limitations, it is not possible to create a supercell that represents a real crystal. Therefore, periodic boundary conditions still have to be used for calculations of the supercells. In consequence, the repetition of the supercell can lead to spurious interactions between the impurity atoms and their images in neighboring supercells and create artificial ordering. The choice of the size of the supercell is therefore a trade-off between computational demand and limitation of the influence of these spurious interactions. It is desirable to create supercells with an optimally randomized distribution of the defect atoms, so that the least amount of artificial ordering can be achieved with small supercell sizes. For this, special quasirandom structures (SQS) that retain correlations from a random distribution even in smaller supercells are used[53]. In order to quantize the randomness of a distribution, the concept of clusters is introduced. A certain distribution of the impurity atoms on a sublattice is defined as a configuration S , consisting of the occupations σ_i for each site on the lattice, which has the value $\sigma_i = \pm 1$ depending on the atom species at that site. A cluster has a size n , which is the number of atoms it contains, and a neighbor distance k , which defines how far away the atoms in the cluster are from each other, with $k = 1$ only taking next-neighbors into account, $k = 2$ also next-next-neighbors and so on. It is evaluated as the product of all occupations σ_i of the atoms in the cluster,

$$\Pi_{(k,n)}(\mathbf{l}, S) = \prod_i \sigma_i. \quad (4.45)$$

The parameter l distinguishes the different possibilities for rotation and translation of the cluster. By evaluating all clusters of a configuration S , the correlation function can be determined as

$$\bar{\Pi}_{(k,n)}(S) = \frac{1}{NE_{(k,n)}} \sum_l \Pi_{(k,n)}(\mathbf{l}, S). \quad (4.46)$$

It is a measure for the randomness of the configuration. The goal is to achieve a value of the correlation function which is as close as possible to the completely randomized correlation function,

$$\bar{\Pi}_{(k,n)}(R) = (2x - 1)^k, \quad (4.47)$$

which is a theoretical result for a configuration R with the impurity atom concentration x , where every occupation is the expectation value for the occupation $\sigma_i = x - (1 - x)$. The difference between the correlation function of a configuration and the respective

completely randomized correlation function is called the correlation of that configuration. Since the number of atoms in the configuration only allows for a certain number of different configurations, it is not generally possible to obtain a configuration with a vanishing correlation. It also depends on the size of clusters n and the neighbor distances k that are considered when evaluating the correlation function. For our supercells, we evaluated clusters of size 4 with only next-neighbors considered, clusters of size 3 with also next-next-neighbors considered, and clusters of size 2 with also next-next-next-neighbors considered. In this thesis, the software ATAT[54, 55] was used to evaluate the clusters and create different supercell configurations. It uses a statistical Monte-Carlo algorithm to make small modifications to the configurations and find the configuration with the lowest correlation.

For the study presented in Paper [II], supercells of different geometries have been created. As shown in Fig. 4.3, it was found that the supercells with the repetitions of the primitive cell evenly distributed in the spatial directions have the most linear decrease of the band gap. The stronger band gap decrease seen in the other geometries

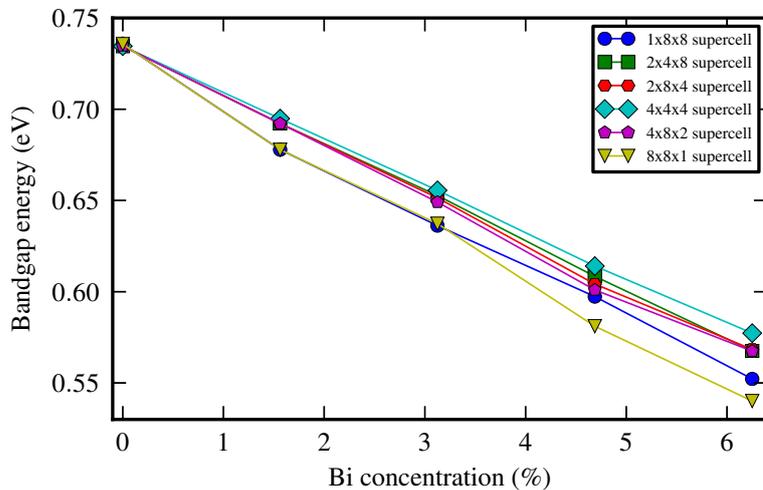


Figure 4.3: Band gaps obtained from DFT calculations with the LDA-1/2 method of $\text{Ga}(\text{Sb}_{1-x}\text{Bi}_x)$ supercells consisting of 64 primitive cells in different cuboidal geometries plotted against the Bi concentration x . The connecting lines are guides to the eye.

is attributed to the spurious interactions between defect atoms[56, 57], since in these geometries they are closer to their images in the directions where less primitive cells are repeated. Therefore, the band gap narrowing in the 4x4x4 supercells is taken as the most realistic representation. A linear fit yields a value of (25.3 ± 0.2) meV/%Bi for the band gap narrowing and an increase of the spin-orbit splitting energy of (17.9 ± 0.2) meV/%Bi.

The repetition of primitive cells in a supercell causes the band structure to change. Due to the larger extension in real space, the 1st BZ will become smaller. If no impurity

atoms are introduced, the band structure of the primitive cell will simply be folded back onto the smaller 1st BZ of the supercell. In order to be able to better compare the band structure of the supercell with that of the primitive cell to see the effect of the impurity atoms, it is necessary to unfold the supercell band structure[58, 59]. Since states from multiple primitive cells are folded into the 1st BZ of the supercell, each state in the supercell band structure unfolds into multiple states for the effective band structure. The states of the supercell band structure are projected onto the states they unfold into,

$$w_{n,\mathbf{K}}(\mathbf{k}_j) = \sum_m \left| \langle \psi_{n,\mathbf{K}}^{\text{SC}} | \psi_{m,\mathbf{k}_j}^{\text{PC}} \rangle \right|^2 = \sum_{\{\mathbf{g}\}} |C_{n,\mathbf{K}}(\mathbf{g} + \mathbf{G}_j)|^2, \quad (4.48)$$

where \mathbf{k} is the k -vector in the primitive cell, \mathbf{K} the k -vector in the supercell, $\langle \psi_{n,\mathbf{K}}^{\text{SC}} |$ is the state in the supercell, $|\psi_{m,\mathbf{k}_j}^{\text{PC}}\rangle$ is the state in the primitive cell, and n and m are the band indices. The projection can be calculated directly using the plane wave coefficients from DFT and gives the so-called spectral weight $w_{n,\mathbf{K}}(\mathbf{k}_j)$, which is a measure for how much of the Bloch character of the respective state is preserved in the supercell. If no impurity atoms are introduced into the supercell, the unfolding procedure yields the primitive cell band structure. The unfolded band structure of a supercell with impurity atoms, however, exhibits new impurity states with low spectral weight and distortions in the states of the primitive cell band structure. These changes are caused by the local interactions of the impurity atoms. In Fig. 4.4, the band structures of the 4x4x4 supercells with different concentrations of Bi atoms are shown. Every dot represents a state, the spectral weight is represented by the opacity of the dot, making states with less spectral weight less visible than those with a high spectral weight. On the left, the primitive cell band structure is regained after the unfolding procedure, since for this concentration, no Bi atoms are incorporated. With increasing Bi concentration, more and more impurity states with low spectral weight emerge. No impurity states are formed inside the band gap. However, the energy of the valence band is shifted upwards by the interaction with the impurity atoms as predicted by the valence band anticrossing model[60]. This is seen in the band structure as a downward shift of the conduction bands, since for the plotting, the edge of the valence band was set to 0 eV.

4.3 Optical Properties

The microscopic theory explained in Section 2.2 has been used to calculate the optical properties presented in this section. In the left part of Fig. 4.5, the onset of absorption for the supercells with different Bi concentrations is shown.

The arrows indicate the band gap energy of the respective concentrations. It can be seen that with increasing Bi concentration, the absorption begins to rise at a lower excitation energy, in accordance with the decrease of the band gap. In the right part of Fig. 4.5, the photoluminescence spectra of the supercells are shown. The peak of the individual curves for the different supercells moves to lower energies,

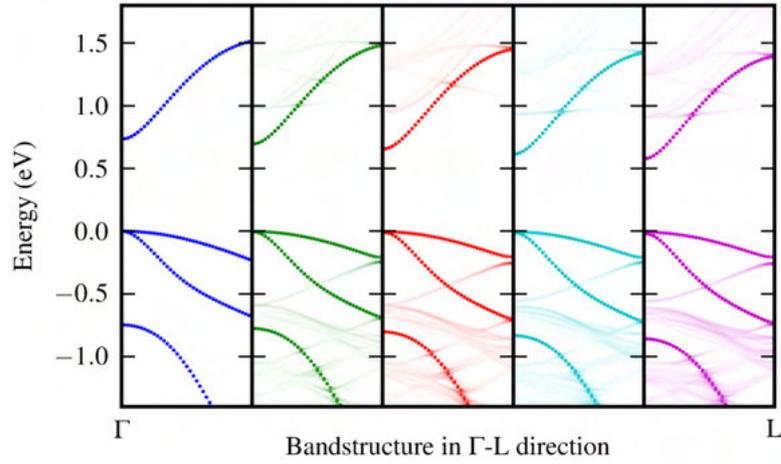


Figure 4.4: Band structures from Γ - to L-point of $\text{Ga}(\text{Sb}_{1-x}\text{Bi}_x)$ supercells for Bi concentrations $x = 0\%$, 1.56% , 3.13% , 4.69% , 6.25% from left to right. Band structures are calculated with DFT with the LDA-1/2 method and unfolded to obtain the effective primitive cell band structure. Each point represents one state, the opacity is proportional to the spectral weight of the state after the unfolding procedure. The highest energy of the valence band is normalized to 0 eV for each band structure.

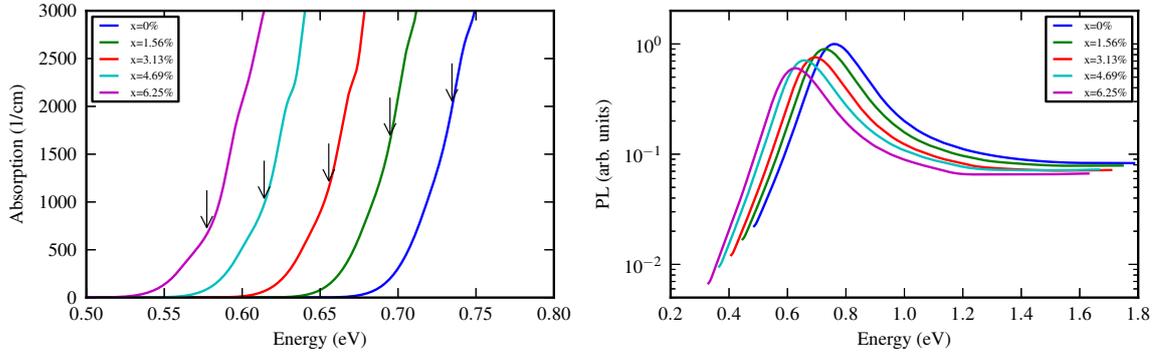


Figure 4.5: Onset of absorption (left) and Photoluminescence spectra (right) of different $\text{Ga}(\text{Sb}_{1-x}\text{Bi}_x)$ supercells, whose concentration is given in the legend, calculated with the SBEs and SLEs using input from DFT calculations with the LDA-1/2 method. Arrows indicate the band gap energy of the respective supercell.

while also decreasing in intensity. This indicates that it is possible to tune the band gap by incorporating different amounts of Bi into the crystal and thus change the wavelength at which light is absorbed or emitted. Besides changing the band gap,

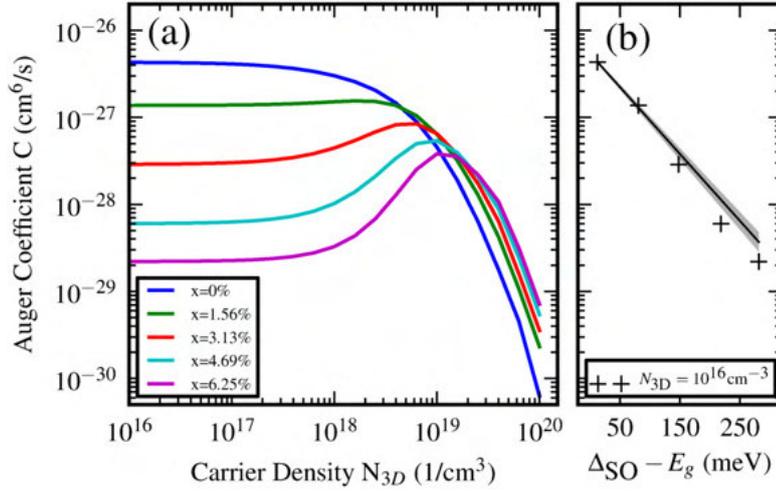


Figure 4.6: Auger loss coefficients of $\text{Ga}(\text{Sb}_{1-x}\text{Bi}_x)$ supercells with different Bi concentrations (a) as a function of the carrier density, (b) at the low density limit as a function of the detuning between band gap and SO energy.

higher Bi concentrations also influence the Auger losses as shown in Fig. 4.6. The Auger coefficient has been calculated according to Sec. 2.2.3 as a function of the carrier density. In the low density limit, the loss coefficient decreases for higher concentrations. This is due to the increasing detuning between the band gap and the SO splitting energy. The CHSH Auger process leads to recombinations of electron-hole pairs across the band gap by excitation of a hole from the SO band to the valence band. For pure $\text{Ga}(\text{SbBi})$, the band gap and SO splitting energy are very close to each other, making this loss channel resonant to the band gap. Because of the Bi-induced band gap narrowing, this loss channel is effectively blocked with higher Bi contents. For higher densities, however, the losses for all Bi concentrations increase, since more transitions become possible, making the loss coefficients roughly the same for all concentrations. For even higher densities, the losses start to decrease for all concentrations due to phase-space filling.

Influence of TDM Phases on High Harmonic Generation

In this chapter, using TDMs obtained from DFT with a random phase for HHG calculations is discussed. The influence of the phase on the HHG spectra is illustrated and two different gauging procedures to obtain smooth phases are presented.

5.1 Triple Dipole Products

The KS orbitals ϕ_{nk} obtained from DFT calculations are only defined up to an arbitrary phase factor. Since the KS equation is solved at every k -point independently, there is no fixed phase relation between orbitals at different k -points. From the orbitals, the TDM matrix element at momentum k between bands n and n' is calculated in DFT according to

$$d_{\mathbf{k}}^{nn'} = \frac{\hat{\mathbf{e}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \cdot \left\langle \phi_{n'\mathbf{k}} \left| \frac{\partial(\mathbf{H} - \epsilon_{n\mathbf{k}}\mathbf{S})}{\partial\mathbf{k}} \right| \phi_{n\mathbf{k}} \right\rangle, \quad (5.49)$$

where ϵ_{nk} are the single-particle energies, \mathbf{H} is the Hamiltonian, \mathbf{S} is the overlap operator and $\hat{\mathbf{e}}$ is the polarization direction[61]. Therefore, the TDMs also contain the random phase of the orbitals. The absolute value of the TDMs is unaffected by this, but when their derivative with regard to k is calculated, this becomes problematic. Since the phase is random for each k -point, it is not smooth along the k -axis, so that the derivative is also not smooth.

Therefore, it is necessary to perform a gauge transformation in order to get smooth and BZ periodic TDM phases.

One possibility is to only take the absolute values of the TDMs and thus make them essentially real quantities. However, for the strong fields used in HHG, different excitation paths become possible[43]. Apart from direct excitation paths from the valence to the conduction band, $h \rightarrow e$, the creation of holes in a valence band h_2 can enable indirect excitation paths $h_1 \rightarrow h_2 \rightarrow e$. The HHG emission is made up of the sum of different direct and indirect excitation paths, leading to a quantum

interference between the paths[43]. For this quantum interference, the phase of the TDM becomes important and should not be neglected. In the calculation of the macroscopic polarization from such an indirect excitation path, the product of the TDMs connecting the three bands enters[44],

$$P_{\text{indir}} |_{h_1 \rightarrow e}(t) \propto \sum_{\mathbf{k}} d_{\mathbf{k}}^{\text{eh}_1} d_{\mathbf{k}}^{\text{h}_2} d_{\mathbf{k}}^{\text{eh}_2} \left[E(t)^2 + \left(d_{\mathbf{k}}^{\text{eh}_2} \right)^2 E(t)^4 + \mathcal{O} \left(E(t)^6 \right) \right]. \quad (5.50)$$

It contains the product of the three TDMs connecting the three bands, which we will refer to as the triple dipole product. If, e.g., the phase of one of the constituting TDMs is anti-symmetric in k , while the other ones are symmetric, the sum over k causes the triple dipole product to vanish. According to the definition of the TDMs (5.49), the TDM between bands n and n' contains the KS orbital of band n as a bra, $\langle \phi_{n'k} |$, and the KS orbital of the band n' as a ket, $|\phi_{nk}\rangle$. In the triple dipole product, since the bands are connected in a circular way, the KS orbital of each constituent band will appear once as a bra and once as a ket. Therefore, if a random phase $e^{i\theta_{n_i}}$ is added to each of the orbitals $n_{1,2,3}$, the dipoles will transform as $\tilde{d}_{\mathbf{k}}^{n_i n_j} = e^{-i\theta_{n_i}} d_{\mathbf{k}}^{n_i n_j} e^{i\theta_{n_j}}$ and the phases will cancel in the triple dipole product,

$$\tilde{d}_{\mathbf{k}}^{n_1 n_2} \tilde{d}_{\mathbf{k}}^{n_2 n_3} \tilde{d}_{\mathbf{k}}^{n_3 n_1} = e^{-i\theta_{n_1} + i\theta_{n_2} - i\theta_{n_2} + i\theta_{n_3} - i\theta_{n_3} + i\theta_{n_1}} d_{\mathbf{k}}^{n_1 n_2} d_{\mathbf{k}}^{n_2 n_3} d_{\mathbf{k}}^{n_3 n_1}. \quad (5.51)$$

This way, also the random phase from DFT is canceled out. This does not allow us to obtain the phase of every TDM in the triple dipole product, but it gives us a smooth phase for the triple product across the BZ. We extract this phase and impose it on one of the constituent TDMs, while taking only the absolute value of the other TDMs. This way, the correct phase of the triple dipole product is preserved. However, since every TDM appears in multiple triple dipole products for different indirect excitation paths, not all triple dipole product phases can be conserved this way, and a choice has to be made.

The result of such a gauging procedure is shown in Fig. 5.1 for the TDMs in $E \perp c$ direction between the highest two valence bands and the lowest two conduction bands in Te. The TDMs were gauged using two triple dipole products and imposing their phases on the $d_{\mathbf{k}}^{\text{VB}_1, \text{CB}_2}$ and $d_{\mathbf{k}}^{\text{VB}_1, \text{CB}_1}$ TDMs. The plane extends to positive and negative k -values and contains the whole BZ, so that it can be continued periodically. Some parts of the TDMs exhibit anti-symmetric behavior, while others exhibit symmetric behavior. For the study in Paper [III], the imposition of the triple dipole product phase on different choices of TDMs was performed and the resulting polarization of the HHG compared. The result is shown in part (b) of Fig. 5.2. It was found that there is no significant difference, so that this method of avoiding the entering of the random phase into the HHG calculation seems to be a viable approach. Moreover, part (a) of the figure shows the result when taking different subsets of the highest four valence bands into account in the calculation. The highest two valence bands do not influence the spectrum in a meaningful way, meaning that they can be left out of the calculations.

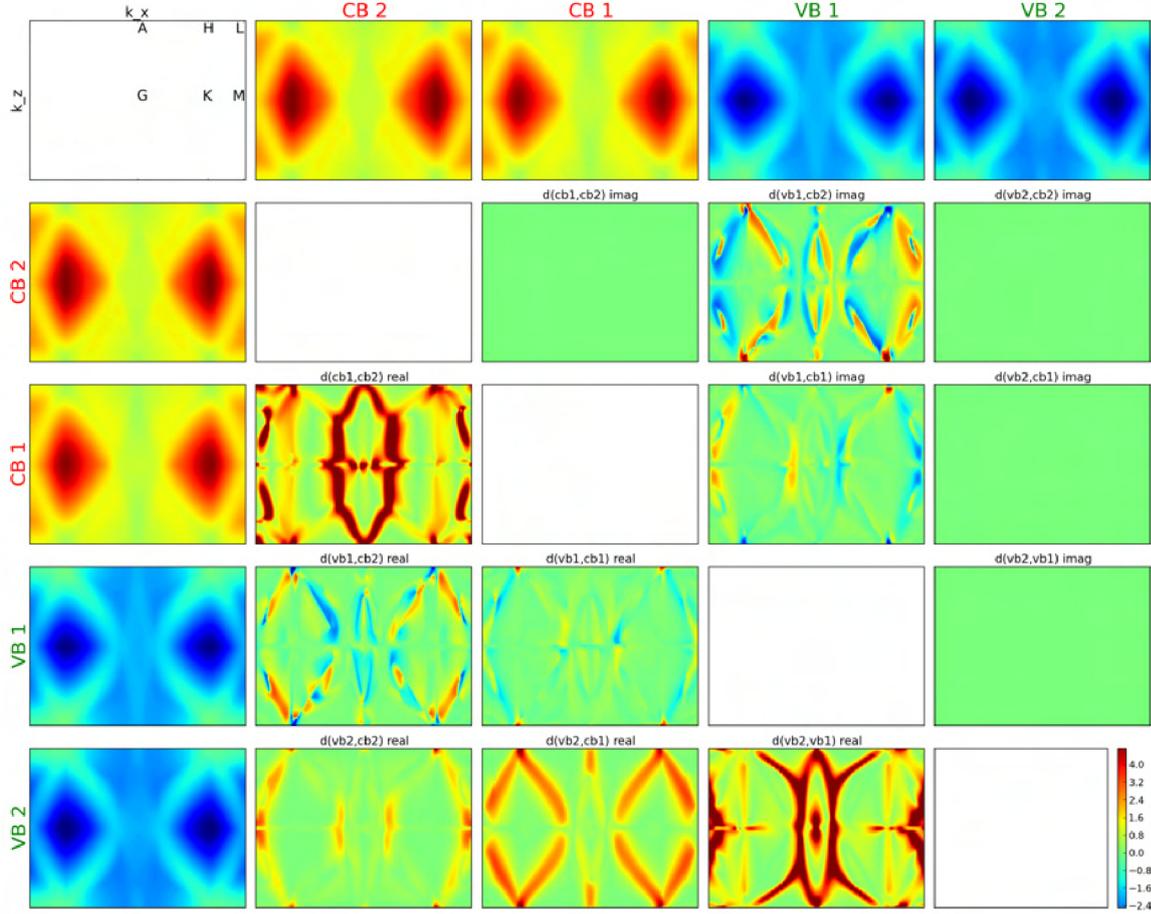


Figure 5.1: TDMs in $E \perp c$ direction of Te in a 2d plane for different band combinations after applying the parallel transport gauge. The first row and first column show the dispersions of the energy bands. In the inner cells, the TDMs between the respective bands in that particular row and column are shown. In the part below the empty diagonal, the real part of the TDMs is shown, while the part above the diagonal shows the imaginary part. The location of high symmetry points in the 2d plane is indicated in the top left corner.

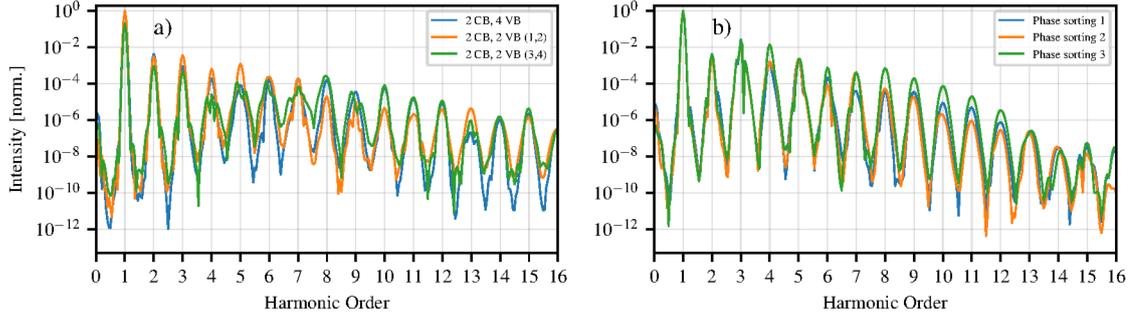


Figure 5.2: Polarization part of the high harmonic spectrum of Te in $E \parallel c$ direction plotted as a function of the driving frequency (a) for different choices of the valence bands included in the calculation (b) for triple dipole product gauges with different TDMs chosen on which to impose the phases.

5.2 Berry Phase

In order to get smooth and periodic phases for every TDM individually, we resorted to a more sophisticated approach to eliminating the random phase based on Berry phases. The theory behind the approach is explained in detail in Ref. [62]. When moving along a closed loop of states in k -space, a certain phase is accumulated, called the Berry phase, which is specific to the system. The berry phase is defined using the the cell-periodic Bloch functions $u_{nk}(\mathbf{r})$.

The Berry phase of the path is calculated as

$$\phi_n = -\text{Im} \ln \left[\langle u_{nk_0} | u_{nk_1} \rangle \langle u_{nk_1} | u_{nk_2} \rangle \dots \langle u_{nk_{N-1}} | e^{-2\pi i x/a} | u_{nk_0} \rangle \right], \quad (5.52)$$

where the phase factor $e^{-2\pi i x/a}$ with the lattice constant a is added when winding around the border of the BZ. The gauge freedom of the Bloch states can be used to rotate the phases in such a way that the overlap between states at neighboring k -points is purely real,

$$\text{Im} \ln \langle u_{n,\mathbf{k}} | u_{n,\mathbf{k}'} \rangle = 0. \quad (5.53)$$

This way, the whole Berry phase accumulated along the path is contained in the overlap between the last and the first state. This gauge is called the parallel transport gauge (PTG). The discontinuity in the phase at the end of the loop can be avoided by distributing the whole Berry phase ϕ evenly across the $N + 1$ states in the loop, so that the overlap between neighboring states always yields the same fraction of the Berry phase ϕ/N . This way, the state $|u_j\rangle$ becomes

$$|\tilde{u}_j\rangle = e^{-ij\phi/N} |u_j\rangle. \quad (5.54)$$

The gauge obtained this way is called the twisted parallel transport gauge (TPTG). Since the Berry phase is distributed evenly across the path, there is no phase jump

and the phase is smooth and periodic across the BZ if the path is chosen accordingly. This approach works for isolated bands which do not become degenerate with other bands across the BZ.

In the case of connected bands, however, a more generalized approach is needed. Bands are connected when at some point in the BZ they become degenerate. At these points, the bands cannot be distinguished by DFT and an arbitrary unitary mixing between the bands can occur. In the multiband approach, multiple bands are taken into account at once. First, one has to find a set of connected bands which are not connected to any other bands. The goal is to find a unitary mixing of the bands,

$$|\tilde{u}_{n\mathbf{k}}\rangle = \sum_{m=1}^J U_{mn}(\mathbf{k}) |u_{m\mathbf{k}}\rangle , \quad (5.55)$$

so that there is an optimal alignment of the states between neighboring k -points. For this, first the overlap matrix $M_{mn}^{(\mathbf{k}_0, \mathbf{k}_1)}$ between all band combinations m and n inside the chosen group at neighboring k -points \mathbf{k}_0 and \mathbf{k}_1 is calculated,

$$M_{mn}^{(\mathbf{k}_0, \mathbf{k}_1)} = \langle \bar{u}_{m\mathbf{k}_0} | u_{n\mathbf{k}_1} \rangle . \quad (5.56)$$

Then, a singular value decomposition is performed on the overlap matrix, yielding

$$M^{(\mathbf{k}_0, \mathbf{k}_1)} = V\Sigma W^\dagger . \quad (5.57)$$

The best unitary approximation to the overlap matrix is given by

$$\mathcal{M}^{(\mathbf{k}_0, \mathbf{k}_1)} = VW^\dagger . \quad (5.58)$$

This matrix is then used to perform the unitary mixing of the Bloch states according to

$$|\bar{u}_{n\mathbf{k}_1}\rangle = \sum_m (\mathcal{M}^\dagger)_{mn} |u_{m\mathbf{k}_1}\rangle . \quad (5.59)$$

In order to implement this procedure, first a DFT calculation is performed. Then, the algorithm goes through the k -points along the chosen path in k -space. At every k -point, the plane wave coefficients are read in and the Bloch functions $u_{n\mathbf{k}}(\mathbf{r})$ are calculated. Then, the overlap matrix between neighboring k -points is evaluated and the mixing of the states is performed according to the above detailed method. When going to the next k -point, the newly mixed states are used, so that after reaching the end of the loop, the PTG is obtained.

This method has been used to gauge the TDMs of GaAs obtained from a DFT calculation using the GGA-1/2 method. An example of the resulting TDMs, the TDM $d_{\mathbf{k}}^{\text{LH1, CB2}}$ between one spin combination of the light hole band and the conduction band, in both the PTG and the TPTG is shown in Fig. 5.3. The starting point of the loop was chosen in the region between the K- and Γ -point. Therefore, a clear phase jump by the amount of the accumulated Berry phase ϕ can be seen in part (a) with the PTG. This phase jump is smoothed out in part (b) with the TPTG.

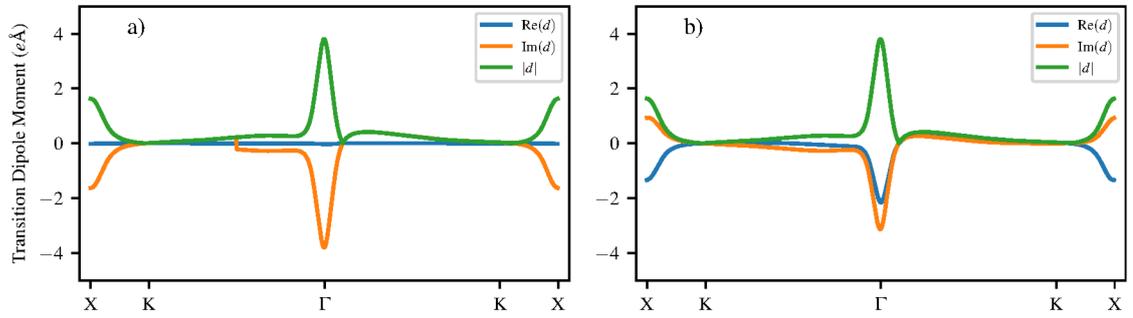


Figure 5.3: Real part, imaginary part and absolute value of TDM in $E \perp c$ direction between light hole and conduction band in Te obtained from DFT calculation with shLDA-1/2 method. The smooth phase is obtained using (a) the parallel transport gauge (b) the twisted parallel transport gauge.

Coulomb Interaction in Monolayer TMDCs

In this chapter, the introduced theoretical concepts will be used and extended to describe the Coulomb interaction in quasi two-dimensional systems. After the discovery of Graphene in 2004[63], TMDCs have recently become a prominent representative of quasi-2d materials with interesting properties. Studying monolayer materials is interesting because they exhibit very different properties from their bulk counterparts and could be useful for opto-electronic applications. The movement of charge carriers in these materials is confined to two dimensions, while the Coulomb potential can also extend in the direction perpendicular to the layer. Therefore, these materials can be regarded as quasi two-dimensional. Furthermore, they are not perfectly two-dimensional, because even a single layer has a finite extension perpendicular to the layer. Since the material is periodic only in two dimensions, the theoretical description of the crystal should be adapted.

The description of electrons as Bloch functions is only valid in the 2d plane of the layer,

$$u_{\alpha, \mathbf{k}_{\parallel}}(\mathbf{r}) = \sum_{\mathbf{G}_{\parallel}} e^{i\mathbf{G}_{\parallel} \cdot \mathbf{r}_{\parallel}} c_{\alpha, \mathbf{k}_{\parallel}}(\mathbf{G}_{\parallel}, z), \quad (6.60)$$

where \mathbf{k}_{\parallel} is the in-plane momentum, \mathbf{G}_{\parallel} are the reciprocal lattice vectors in the plane and $c_{\alpha, \mathbf{k}_{\parallel}}$ are the plane wave coefficients, leaving a dependence on the spatial coordinate z , which is chosen to be the direction perpendicular to the layer. The DFT system used for the calculations always assumes periodic boundary conditions in all directions. Therefore, to account for the lack of periodicity in z -direction, an approximation

$$u_{\alpha, \mathbf{k}_{\parallel}}(\mathbf{r}) = u_{\alpha, \mathbf{k}_{\parallel}, k_z=0}^{\text{SC}}(\mathbf{r}) \quad (6.61)$$

has to be made. The unit cell is extended into a supercell with a considerable amount of vacuum in z -direction, so that the repeated layers do not interact significantly with each other. The wave functions of this supercell $u_{\alpha, \mathbf{k}_{\parallel}, k_z=0}^{\text{SC}}(\mathbf{r})$ are assumed to be equal to the wave functions of the 2d-periodic material. Using these wave functions and their

2d periodicity, the Coulomb matrix element in second quantization can be derived as

$$V_{\mathbf{k}_{\parallel 1,2}}^{\alpha_1-4} = \sum_{\mathbf{q}_{\parallel}} \int_{\text{ec}} d^3r d^3r' u_{\alpha_1, \mathbf{k}_{\parallel 1}}^*(\mathbf{r}) u_{\alpha_4, \mathbf{k}_{\parallel 1} - \mathbf{q}_{\parallel}}(\mathbf{r}) \quad (6.62)$$

$$V_{\mathbf{q}_{\parallel}}(z, z') u_{\alpha_2, \mathbf{k}_{\parallel 2}}^*(\mathbf{r}') u_{\alpha_3, \mathbf{k}_{\parallel 2} + \mathbf{q}_{\parallel}}(\mathbf{r}'). \quad (6.63)$$

By inserting the quasi-2d Coulomb potential,

$$V_{\mathbf{q}_{\parallel}}(z) = V_{\mathbf{q}_{\parallel}}^{2\text{D}} e^{-q_{\parallel}|z|}, \quad (6.64)$$

with the ideally 2d Coulomb potential $V_{\mathbf{q}_{\parallel}}^{2\text{D}}$ into the above expression, it is found that the quasi-2d Coulomb matrix element,

$$V_{\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}}^{\text{e,h}} = V_{|\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}|}^{2\text{D}} \int_{\text{ec}} d^3r d^3r' u_{\mathbf{c}, \mathbf{k}_{\parallel}}^*(\mathbf{r}) u_{\mathbf{c}, \mathbf{k}'_{\parallel}}(\mathbf{r}) e^{-|\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}||z - z'|} u_{\mathbf{v}, \mathbf{k}'_{\parallel}}^*(\mathbf{r}') u_{\mathbf{v}, \mathbf{k}_{\parallel}}(\mathbf{r}'), \quad (6.65)$$

can be written as the 2d Coulomb matrix element $V_{|\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}|}^{2\text{D}}$ multiplied by an integral. Since this integral represents the influence of the three-dimensionality on the quasi-2d layer, it is referred to as the form factor. Its values vary between 0 and 1.

In Paper [I], a DFT calculation is performed and the resulting band structure is used to fit a MDF model Hamiltonian to it. The wave functions are used to calculate the form factor and the Coulomb matrix elements.

Conclusion

In this thesis, a combined approach of density functional theory and many-body theory was used to study the optoelectronic properties of monolayer TMDCs, dilute bismides and tellurium. The aim was to evaluate and develop a methodology for obtaining predictions on material properties from *first principles*, meaning that no experimental input is needed. This was achieved by using DFT for *ab initio* calculations of structural and electronic parameters as well as wave functions and interaction matrix elements, that are subsequently used as input to the microscopic theory to calculate optical properties. By employing the LDA-1/2 and shLDA-1/2 methods, the DFT calculations yielded accurate input data for the microscopic calculations. The microscopic theory, including many-body effects, dynamics of carrier populations and polarizations as well as the possibility of including Auger and spontaneous emission effects, is able to calculate a broad range of optical properties.

In Chapter 3, these methods were used on GaSe and Te. For Te, the lattice constants $a = 4.51 \text{ \AA}$ and $c = 5.96 \text{ \AA}$ as well as the direct band gap $E_g = 0.32 \text{ eV}$ at the H-point were calculated, showing good agreement with experimental data. For GaSe, the lattice constants $a = 3.746 \text{ \AA}$ and $c = 15.94 \text{ \AA}$ as well as the band gap $E_g = 1.92 \text{ eV}$ was calculated.

In Chapter 4, the LDA-1/2 method was extended to be able to use it for dilute bismides. This made it possible to calculate the optoelectronic properties of Ga(SbBi) as a representative example of this material class. A band gap narrowing of $(25.3 \pm 0.2) \text{ meV/\%Bi}$ and an increase of the spin-orbit splitting energy of $(17.9 \pm 0.2) \text{ meV/\%Bi}$ was found. An unfolding routine yielded effective band structures, showing impurity states and distortion of the bands due to the interaction with the Bi atoms. The calculated absorption and photoluminescence spectra show that the onset of absorption and emission peaks change in accordance with the narrowing band gap, so that the optical properties can effectively be influenced by incorporating a certain amount of bismuth. Furthermore, the loss channel through Auger recombination of an electron-hole pair with the excitation of a hole from the SO band to the valence band can be decreased by incorporating Bi, since the band gap becomes smaller than the SO splitting energy. However, the decrease of the Auger

7 Conclusion

losses occurs at low carrier densities only, where it can be described using the function $C = (5.4 \pm 0.1) \cdot 10^{-27} \text{ cm}^6 \text{ s}^{-1} \times \exp\left((17 \pm 1) \frac{\Delta_{\text{SO}} - E_g}{\text{eV}}\right)$. While these results look promising for the applications of dilute bismides, in reality, it is a challenging task to grow crystals with a significant amount of Bi incorporation, since it tends to form droplets in the growth process[64]. Furthermore, it was not tested how the results change for bigger supercells.

In Chapter 5, triple dipole products were introduced as a way of canceling out the random phase in the TDMs obtained from DFT, thus making it possible to perform HHG calculations while preserving correct phase relations. The gauging of the TDMs using the Berry phase is also discussed, making it possible to obtain TDMs with smooth and periodic phases across the whole BZ.

In Chapter 6, our microscopic theory was extended to quasi-2d materials. For these structures, the quasi-2d Coulomb matrix element was derived and the form factor in the expression was identified, which represents the deviations from the ideally 2d case.

Bibliography

- [1] X. Wu, C. Zhang, and W. Du, “An analysis on the crisis of “chips shortage” in automobile industry —based on the double influence of COVID-19 and trade friction”, *Journal of Physics: Conference Series* **1971**, 012100 (2021).
- [2] Á. Morales-García, R. Valero, and F. Illas, “An empirical, yet practical way to predict the band gap in solids by using density functional band structure calculations”, *The Journal of Physical Chemistry C* **121**, 18862–18866 (2017).
- [3] R. R. Pela, M. Marques, and L. K. Teles, “Comparing LDA-1/2, HSE03, HSE06 and G_0W_0 approaches for band gap calculations of alloys”, *Journal of Physics: Condensed Matter* **27**, 505502 (2015).
- [4] S. Francoeur, M.-J. Seong, A. Mascarenhas, S. Tixier, M. Adamczyk, and T. Tiedje, “Band gap of $\text{GaAs}_{1-x}\text{Bi}_x$, $0 < x < 3.6\%$ ”, *Applied Physics Letters* **82**, 3874–3876 (2003).
- [5] I. P. Marko and S. J. Sweeney, “Progress toward iii–v bismide alloys for near- and midinfrared laser diodes”, *IEEE Journal of Selected Topics in Quantum Electronics* **23**, 1–12 (2017).
- [6] K. S. Novoselov, A. Mishchenko, A. Carvalho, and A. H. C. Neto, “2d materials and van der waals heterostructures”, *Science* **353**, aac9439 (2016).
- [7] H. Tian, M. L. Chin, S. Najmaei, Q. Guo, F. Xia, H. Wang, and M. Dubey, “Optoelectronic devices based on two-dimensional transition metal dichalcogenides”, *Nano Research* **9**, 1543–1560 (2016).
- [8] S. Ghimire, A. D. DiChiara, E. Sistrunk, P. Agostini, L. F. DiMauro, and D. A. Reis, “Observation of high-order harmonic generation in a bulk crystal”, *Nature physics* **7**, 138–141 (2011).
- [9] J. Hebling, K.-L. Yeh, M. C. Hoffmann, B. Bartal, and K. A. Nelson, “Generation of high-power terahertz pulses by tilted-pulse-front excitation and their application possibilities”, *J. Opt. Soc. Am. B* **25**, B6–B19 (2008).

Bibliography

- [10] S. Jiang, H. Wei, J. Chen, C. Yu, R. Lu, and C. D. Lin, “Effect of transition dipole phase on high-order-harmonic generation in solid materials”, *Phys. Rev. A* **96**, 053850 (2017).
- [11] Y.-R. Shen, “Principles of nonlinear optics”, (1984).
- [12] S. S. Tsirkin, P. A. Puentes, and I. Souza, “Gyrotropic effects in trigonal tellurium studied from first principles”, *Phys. Rev. B* **97**, 035158 (2018).
- [13] G. Kresse and J. Hafner, “Ab initio molecular dynamics for liquid metals”, *Phys. Rev. B* **47**, 558–561 (1993).
- [14] G. Kresse and J. Hafner, “Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium”, *Phys. Rev. B* **49**, 14251–14269 (1994).
- [15] G. Kresse and J. Furthmüller, “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set”, *Computational Materials Science* **6**, 15–50 (1996).
- [16] G. Kresse and J. Furthmüller, “Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set”, *Phys. Rev. B* **54**, 11169–11186 (1996).
- [17] P. E. Blöchl, “Projector augmented-wave method”, *Phys. Rev. B* **50**, 17953–17979 (1994).
- [18] G. Kresse and D. Joubert, “From ultrasoft pseudopotentials to the projector augmented-wave method”, *Phys. Rev. B* **59**, 1758–1775 (1999).
- [19] D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction* (John Wiley & Sons, Inc., Hoboken, New Jersey, 2009).
- [20] L. Bannow, “Optical and electronic properties of novel semiconductor materials”, Ph.D. Thesis (Philipps-Universität Marburg, Marburg, 2019).
- [21] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas”, *Phys. Rev.* **136**, B864–B871 (1964).
- [22] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects”, *Phys. Rev.* **140**, A1133–A1138 (1965).
- [23] O. H. Nielsen and R. M. Martin, “Stresses in semiconductors: ab initio calculations on si, ge, and gaas”, *Phys. Rev. B* **32**, 3792–3805 (1985).
- [24] P. Verma and D. G. Truhlar, “Status and challenges of density functional theory”, *Trends in Chemistry* **2**, Special Issue - Laying Groundwork for the Future, 302–318 (2020).
- [25] D. M. Ceperley and B. J. Alder, “Ground state of the electron gas by a stochastic method”, *Phys. Rev. Lett.* **45**, 566–569 (1980).
- [26] J. P. Perdew and A. Zunger, “Self-interaction correction to density-functional approximations for many-electron systems”, *Phys. Rev. B* **23**, 5048–5079 (1981).

- [27] J. P. Perdew, “Density functional theory and the band gap problem”, *International Journal of Quantum Chemistry* **28**, 497–523 (1985).
- [28] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple”, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- [29] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, “Restoring the density-gradient expansion for exchange in solids and surfaces”, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [30] J. P. Perdew, M. Ernzerhof, and K. Burke, “Rationale for mixing exact exchange with density functional approximations”, *The Journal of Chemical Physics* **105**, 9982–9985 (1996).
- [31] J. Heyd, G. E. Scuseria, and M. Ernzerhof, “Hybrid functionals based on a screened coulomb potential”, *The Journal of Chemical Physics* **118**, 8207–8215 (2003).
- [32] L. G. Ferreira, M. Marques, and L. K. Teles, “Approximation to density functional theory for the calculation of band gaps of semiconductors”, *Phys. Rev. B* **78**, 125116 (2008).
- [33] L. G. Ferreira, M. Marques, and L. K. Teles, “Slater half-occupation technique revisited: the lda-1/2 and gga-1/2 approaches for atomic ionization energies and band gaps in semiconductors”, *AIP Advances* **1**, 032119 (2011).
- [34] L. J. Sham and M. Schlüter, “Density-functional theory of the energy gap”, *Phys. Rev. Lett.* **51**, 1888–1891 (1983).
- [35] R. W. Godby, M. Schlüter, and L. J. Sham, “Self-energy operators and exchange-correlation potentials in semiconductors”, *Phys. Rev. B* **37**, 10159–10175 (1988).
- [36] J. C. Slater and K. H. Johnson, “Self-consistent-field $X\alpha$ cluster method for polyatomic molecules and solids”, *Phys. Rev. B* **5**, 844–853 (1972).
- [37] K.-H. Xue, J.-H. Yuan, L. R. Fonseca, and X.-S. Miao, “Improved lda-1/2 method for band structure calculations in covalent semiconductors”, *Computational Materials Science* **153**, 493–505 (2018).
- [38] M. Kira and S. W. Koch, *Semiconductor quantum optics* (Cambridge University Press, 2011).
- [39] H. Haug and S. W. Koch, *Quantum theory of the optical and electronic properties of semiconductors* (World Scientific Publishing Company, 2009).
- [40] M. Kira, F. Jahnke, W. Hoyer, and S. Koch, “Quantum theory of spontaneous emission and coherent effects in semiconductor microstructures”, *Progress in quantum electronics* **23**, 189–279 (1999).
- [41] M. Kira and S. W. Koch, “Many-body correlations and excitonic effects in semiconductor spectroscopy”, *Progress in quantum electronics* **30**, 155–296 (2006).

- [42] J. Hader, J. V. Moloney, and S. W. Koch, “Microscopic evaluation of spontaneous emission-and auger-processes in semiconductor lasers”, *IEEE journal of quantum electronics* **41**, 1217–1226 (2005).
- [43] U. Huttner, M. Kira, and S. W. Koch, “Ultrahigh off-resonant field effects in semiconductors”, *Laser & Photonics Reviews* **11**, 1700049 (2017).
- [44] U. Huttner, “Theoretical analysis of ultrafast strong terahertz-field effects in semiconductors”, Ph.D. Thesis (Philipps-Universität Marburg, Marburg, 2016).
- [45] T. Doi, K. Nakao, and H. Kamimura, “The valence band structure of tellurium. i. the $k \cdot p$ perturbation method”, *Journal of the Physical Society of Japan* **28**, 36–43 (1970).
- [46] C. Patel, “Efficient phase-matched harmonic generation in tellurium with a CO_2 laser at 10.6μ ”, *Physical Review Letters* **15**, 1027 (1965).
- [47] C. Adenis, V. Langer, and O. Lindqvist, “Reinvestigation of the structure of tellurium”, *Acta Crystallographica Section C* **45**, 941–942 (1989).
- [48] V. B. Anzin, M. I. Eremets, Y. V. Kosichkin, A. I. Nadezhdinskii, and A. M. Shirokov, “Measurement of the energy gap in tellurium under pressure”, *physica status solidi (a)* **42**, 385–390 (1977).
- [49] S. Tutihasi, G. G. Roberts, R. C. Keezer, and R. E. Drews, “Optical properties of tellurium in the fundamental absorption region”, *Phys. Rev.* **177**, 1143–1150 (1969).
- [50] S. Grimme, “Semiempirical gga-type density functional constructed with a long-range dispersion correction”, *Journal of Computational Chemistry* **27**, 1787–1799 (2006).
- [51] E. Aulich, J. L. Brebner, and E. Mooser, “Indirect energy gap in gase and gas”, *physica status solidi (b)* **31**, 129–131 (1969).
- [52] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, “Band parameters for iii–v compound semiconductors and their alloys”, *J. Appl. Phys.* **89**, 5815 (2001).
- [53] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, “Special quasirandom structures”, *Phys. Rev. Lett.* **65**, 353–356 (1990).
- [54] A. Van De Walle, M. Asta, and G. Ceder, “The alloy theoretic automated toolkit: a user guide”, *Calphad* **26**, 539–553 (2002).
- [55] A. Van De Walle, “Multicomponent multisublattice alloys, nonconfigurational entropy and other additions to the alloy theoretic automated toolkit”, *Calphad* **33**, 266–278 (2009).
- [56] M. Usman, C. A. Broderick, A. Lindsay, and E. P. O’Reilly, “Tight-binding analysis of the electronic structure of dilute bismide alloys of gap and gaas”, *Physical Review B* **84**, 245202 (2011).

- [57] R. Kini, A. Ptak, B. Fluegel, R. France, R. Reedy, and A. Mascarenhas, “Effect of bi alloying on the hole transport in the dilute bismide alloy $1-x\text{Bi}_2\text{Te}_3$ ”, Physical Review B **83**, 075307 (2011).
- [58] V. Popescu and A. Zunger, “Effective band structure of random alloys”, Physical review letters **104**, 236403 (2010).
- [59] V. Popescu and A. Zunger, “Extracting E versus \mathbf{k} effective band structure from supercell calculations on alloys and impurities”, Phys. Rev. B **85**, 085201 (2012).
- [60] M. Rajpalke, W. Linhart, M. Birkett, K. Yu, D. Scanlon, J. Buckeridge, T. S. Jones, M. Ashwin, and T. D. Veal, “Growth and properties of gasbbi alloys”, Applied Physics Letters **103**, 142106 (2013).
- [61] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, “Linear optical properties in the projector-augmented wave methodology”, Phys. Rev. B **73**, 045112 (2006).
- [62] D. Vanderbilt, *Berry phases in electronic structure theory: electric polarization, orbital magnetization and topological insulators* (Cambridge University Press, 2018).
- [63] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, “Electric field effect in atomically thin carbon films”, Science **306**, 666–669 (2004).
- [64] L. Wang, L. Zhang, L. Yue, D. Liang, X. Chen, Y. Li, P. Lu, J. Shao, and S. Wang, “Novel dilute bismide, epitaxy, physical properties and device application”, Crystals **7**, 10.3390/cryst7030063 (2017).

Publications

This chapter contains the publications whose material was subject of this thesis. First, the abstracts of each publication are listed and afterwards the publications themselves have been included.

Paper I: *Microscopic Coulomb interaction in transition-metal dichalcogenides*

The quasi-two dimensional Coulomb interaction potential in transition metal dichalcogenides is determined using the Kohn–Sham wave functions obtained from *ab initio* calculations. An effective form factor is derived that accounts for the finite extension of the wave functions in the direction perpendicular to the material layer. The resulting Coulomb matrix elements are used in microscopic calculations based on the Dirac Bloch equations yielding an efficient method to calculate the band gap and the opto-electronic material properties in different environments and under various excitation conditions.

Paper II: *Extension of the LDA-1/2 method to the material class of bismuth containing III-V semiconductors*

The LDA-1/2 method is employed in density functional theory calculations for the electronic structure of III-V dilute bismide systems. For the representative example of Ga(SbBi) with Bi concentrations below 10%, it is shown that this method works very efficiently, especially due to its reasonably low demand on computer memory. The resulting bandstructure and wavefunctions are used to compute the interaction matrix elements that serve as input to microscopic calculations of the optical properties and intrinsic losses relevant for optoelectronic applications of dilute bismides.

Paper III: *Microscopic theory for the incoherent resonant and coherent off-resonant optical response of tellurium*

An *ab initio*-based fully microscopic approach is applied to study the nonlinear optical response of bulk tellurium. The structural and electronic properties are calculated from first principles using the shLDA-1/2 method within density

functional theory. The resulting band structure and dipole matrix elements serve as input for the quantum mechanical evaluation of the anisotropic linear optical absorption spectra, yielding results in excellent agreement with published experimental data. Assuming quasiequilibrium carrier distributions in the conduction and valence bands, absorption/gain and spontaneous emission spectra are computed from the semiconductor Bloch and luminescence equations. For ultrafast intense off-resonant excitation, the generation of high harmonics is evaluated and the emission spectra are calculated for samples of different thicknesses.

Paper IV: *Intrinsic Carrier Losses in Tellurium Due to Radiative and Auger Recombinations*

Fully microscopic many-body models based on inputs from first principle density functional theory are used to calculate the carrier losses due to radiative and Auger-recombinations in bulk tellurium. It is shown that Auger processes dominate the losses for carrier densities in the range typical for applications as lasers. The Auger loss depends crucially on the energetic position of the H_6 valence bands. At cryogenic temperatures of 50 K (100 K) the Auger coefficient, C , varies by about six (three) orders of magnitude within the range of published distances between these bands and the valence band edge. Values for C at the high and low end of these ranges are found if the distance is smaller or larger than the bandgap, respectively. At room temperature the sensitivity is reduced to about a factor of four with C values ranging between 0.4 and $1.6 \times 10^{-27} \text{cm}^6 \text{s}^{-1}$. Here, radiative losses dominate for carrier densities up to about $10^{16} / \text{cm}^3$ with a loss coefficient $B \approx 10^{-11} \text{cm}^3 \text{s}^{-1}$. The radiative losses are about two to three times lower than in typical bulk III-V materials for comparable wavelengths.

Paper I

J. Neuhaus, **S. C. Liebscher**, L. Meckbach, T. Stroucken, and S. W. Koch

“Microscopic Coulomb interaction in transition-metal dichalcogenides”

J. Phys.: Condens. Matter **33**, 035301 (2021) **DOI:** 10.1088/1361-648X/abb681

©IOP Publishing. Reproduced with permission. All rights reserved.

Microscopic Coulomb interaction in transition-metal dichalcogenides

J Neuhaus, S C Liebscher, L Meckbach, T Stroucken*  and S W Koch

Department of Physics and Material Sciences Center, Philipps University Marburg, Renthof 5, D-35032 Marburg, Germany

E-mail: tineke.stroucken@t-online.de

Received 23 June 2020, revised 26 August 2020

Accepted for publication 9 September 2020

Published 19 October 2020



CrossMark

Abstract

The quasi-two dimensional Coulomb interaction potential in transition metal dichalcogenides is determined using the Kohn–Sham wave functions obtained from *ab initio* calculations. An effective form factor is derived that accounts for the finite extension of the wave functions in the direction perpendicular to the material layer. The resulting Coulomb matrix elements are used in microscopic calculations based on the Dirac Bloch equations yielding an efficient method to calculate the band gap and the opto-electronic material properties in different environments and under various excitation conditions.

Keywords: TMDCs, Coulomb interaction, band gap renormalization, excitons, many-body-theory

(Some figures may appear in colour only in the online journal)

1. Introduction

With a thickness of just one unit cell, TMDC (transition metal dichalcogenide) monolayer materials can be viewed as the realization of effectively two-dimensional (2D) system. As revealed by density functional theory (DFT) calculations, the resulting 2D quasi-particle dispersions differ not only quantitatively, but also qualitatively from the three dimensional (3D) bandstructure of the corresponding bulk materials, making these systems interesting both for fundamental material science as well as technological applications. In particular, monolayers of TMDCs and TMDC-based heterostructures are currently investigated with regards to their application potential in opto-electronic devices. Unlike the bulk TMDC materials, the monolayers display a direct gap in the visible range with strong light matter interaction and many-body effects due to carrier confinement and weak intrinsic screening of the Coulomb interaction [1–5].

A central challenge for the predictive modeling of the TMDC opto-electronic properties is the analysis of the many-body effects and their influence on the optical spectra. Although state of the art, full *ab initio* GW–BSE calculations have been employed to calculate the linear optical spectra of

freely suspended TMDC monolayers, these are numerically extremely challenging and practically intractable for the general nonlinear response. Furthermore, numerical implementations of GW–BSE calculations for quasi-2D TMDCs are artificially 3D and require the insertion of large vacuum regions or truncations of the Coulomb interaction to avoid spurious inter-layer interactions. This not only increases the numerical effort despite decreasing the material's dimensionality, but also limits the applicability to situations where a monolayer is embedded into a more complex heterostructure or photonic crystal cavity. Hence, there is a need for *ab initio* based theoretical descriptions that are both accurate and flexible to describe the linear and nonlinear optical response under various excitation conditions and different geometries, and at the same allow for the identification and intuitive interpretation of the relevant physical processes.

A powerful tool to compute the optical response of semiconductors for a wide variety of excitation conditions is provided by the semiconductor Bloch equations (SBE), respectively the Dirac Bloch equations (DBE) for TMDCs. Based on the observation that only few bands contribute to the optical response, one derives the equation of motion (eom) for the relevant material parameters from an effective system Hamiltonian including only the relevant valence and

* Author to whom any correspondence should be addressed.

conduction bands, thus reducing the numerical complexity enormously. Due to its non-perturbative nature, this approach is particularly well suited to describe the linear and nonlinear optical response in the presence of strong many-body interactions.

As an essential input, the SBE/DBE approach needs the single-particle band structure and the interaction matrix elements. Whereas band dispersions and dipole matrix elements can be either accessed directly from *ab initio* DFT calculations or from a DFT based model Hamiltonian, the determination of the quasi-2D Coulomb potential and its matrix elements is a nontrivial task. For this purpose, we develop in this paper a parameter-free approach that allows us to efficiently determine the Coulombic input for the SBE from the Kohn–Sham wavefunctions without the need of additional approximations.

The paper is organized as follows: in the following section 2, we briefly summarize the microscopic SBE/DBE approach and introduce an orbital dependent form factor that accounts for the quasi-2D nature of Coulomb interaction potential in TMDCs. We show how this form factor can be computed from the Kohn–Sham wave functions. In section 3, we then provide details of the needed DFT calculations and present a detailed analysis of the form factor. In section 5, we develop an analytic approximation for the form factors that allows us to efficiently calculate the density dependent band-gap renormalization and exciton resonances for different material systems. We discuss and compare our results to available experimental and GW–BSE based *ab initio* results in section 5.1.

2. Preliminaries

2.1. 2D semiconductor Bloch equations

For a strictly 2D semiconductor, the basic system Hamiltonian is given by

$$H = H_0 + H_I + H_C$$

where

$$H_0 = \sum_{\alpha\mathbf{k}_\parallel} \epsilon_{\alpha\mathbf{k}_\parallel} c_{\alpha\mathbf{k}_\parallel}^\dagger c_{\alpha\mathbf{k}_\parallel}$$

describes the single-particle part,

$$H_I = \frac{e}{m_0c} \sum_{\alpha\alpha'\mathbf{k}_\parallel} \mathbf{A} \cdot \mathbf{p}_{\alpha\alpha'\mathbf{k}_\parallel} c_{\alpha\mathbf{k}_\parallel}^\dagger c_{\alpha'\mathbf{k}_\parallel},$$

contains the light–matter interaction, and

$$H_C = \frac{1}{2} \sum_{\substack{\alpha\alpha'\beta\beta' \\ \mathbf{k}_\parallel\mathbf{k}_\parallel'\mathbf{q}_\parallel \neq 0}} V_{\mathbf{q}_\parallel|\mathbf{k}_\parallel';\mathbf{k}_\parallel}^{\alpha\beta\beta'\alpha'} c_{\alpha\mathbf{k}_\parallel-\mathbf{q}_\parallel}^\dagger c_{\beta\mathbf{k}_\parallel'+\mathbf{q}_\parallel}^\dagger c_{\beta'\mathbf{k}_\parallel'} c_{\alpha'\mathbf{k}_\parallel}.$$

represents the Coulomb interaction, respectively. Here, $\hbar\mathbf{k}_\parallel$ is the in-plane crystal momentum, α combines spin and band index, $\epsilon_{\alpha\mathbf{k}_\parallel}$ contains the effective single-particle band dispersion, and $\mathbf{p}_{\alpha\alpha'\mathbf{k}_\parallel}$ and $V_{\mathbf{q}_\parallel|\mathbf{k}_\parallel';\mathbf{k}_\parallel}^{\alpha\beta\beta'\alpha'} = V_{\mathbf{q}_\parallel} \langle \alpha\mathbf{k}_\parallel - \mathbf{q}_\parallel | \alpha'\mathbf{k}_\parallel \rangle \langle \beta\mathbf{k}_\parallel' | \beta'\mathbf{k}_\parallel' \rangle$ denote the momentum and Coulomb matrix ele-

ments, respectively. Furthermore, $V_{\mathbf{q}_\parallel}$ is the Fourier transform of the 2D Coulomb potential. In the presence of a dielectric environment, the environmental screening of the Coulomb potential can be included into the definition of the ‘bare’ Coulomb potential.

The optical response is then given by $\mathbf{j} = -\frac{e}{m_0} \sum_{\alpha\alpha'\mathbf{k}_\parallel} \mathbf{p}_{\alpha\alpha'\mathbf{k}_\parallel} \langle c_{\alpha\mathbf{k}_\parallel}^\dagger c_{\alpha'\mathbf{k}_\parallel} \rangle$ and can be computed from the Heisenberg equations of motion for the transition amplitudes $P_{\mathbf{k}_\parallel}^{\alpha\alpha'} = \langle c_{\alpha\mathbf{k}_\parallel}^\dagger c_{\alpha'\mathbf{k}_\parallel} \rangle$:

$$i\hbar \frac{d}{dt} P_{\mathbf{k}_\parallel}^{\alpha\alpha'} = \left(\epsilon_{\alpha'\mathbf{k}_\parallel} - \epsilon_{\alpha\mathbf{k}_\parallel} - i\gamma \right) P_{\mathbf{k}_\parallel}^{\alpha\alpha'} + \tilde{Q}_{\mathbf{k}_\parallel}^{\alpha\alpha'} - \left(\tilde{Q}_{\mathbf{k}_\parallel}^{\alpha'\alpha} \right)^* + i\hbar \left. \frac{d}{dt} P_{\mathbf{k}_\parallel}^{\alpha\alpha'} \right|_{\text{corr}}. \quad (1)$$

Here,

$$\tilde{Q}_{\mathbf{k}_\parallel}^{\alpha\alpha'} = \sum_{\beta} P_{\mathbf{k}_\parallel}^{\alpha\beta} \tilde{\Omega}_{\mathbf{k}_\parallel}^{\beta\alpha'} \quad (2)$$

$$\tilde{\Omega}_{\mathbf{k}_\parallel}^{\beta\alpha'} = \frac{e}{m_0c} \mathbf{A} \cdot \mathbf{p}_{\alpha'\beta\mathbf{k}_\parallel} - \sum_{\substack{\gamma\gamma' \\ \mathbf{k}_\parallel' \neq \mathbf{k}_\parallel}} V_{\mathbf{k}_\parallel-\mathbf{k}_\parallel';\mathbf{k}_\parallel;\mathbf{k}_\parallel'}^{\alpha'\beta\gamma\gamma'} P_{\mathbf{k}_\parallel'}^{\gamma\gamma'} \quad (3)$$

contains the sources and the renormalizations of the single particle energies and internal field. For the interband polarization $P_{\mathbf{k}_\parallel}^{\nu c}$, we explicitly have

$$\begin{aligned} \tilde{Q}_{\mathbf{k}_\parallel}^{\nu c} - \left(\tilde{Q}_{\mathbf{k}_\parallel}^{c\nu} \right)^* &= \left(\tilde{\Omega}_{\mathbf{k}_\parallel}^{cc} - \tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu\nu} \right) P_{\mathbf{k}_\parallel}^{\nu c} + \left(f_{\mathbf{k}_\parallel}^{\nu} - f_{\mathbf{k}_\parallel}^c \right) \tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu c}, \\ \tilde{\Omega}_{\mathbf{k}_\parallel}^{cc} &= \frac{e}{m_0c} \mathbf{p}_{c c \mathbf{k}_\parallel} \cdot \mathbf{A} - \sum_{\mathbf{k}_\parallel' \neq \mathbf{k}_\parallel} \\ &\quad \times \left(V_{\mathbf{k}_\parallel-\mathbf{k}_\parallel';\mathbf{k}_\parallel;\mathbf{k}_\parallel'}^{cccc} f_{\mathbf{k}_\parallel'}^c + V_{\mathbf{k}_\parallel-\mathbf{k}_\parallel';\mathbf{k}_\parallel;\mathbf{k}_\parallel'}^{cc\nu\nu} f_{\mathbf{k}_\parallel'}^{\nu} \right) + \text{NR}, \\ \tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu c} &= \frac{e}{m_0c} \mathbf{p}_{\nu c \mathbf{k}_\parallel} \cdot \mathbf{A} \\ &\quad - \sum_{\mathbf{k}_\parallel' \neq \mathbf{k}_\parallel} V_{\mathbf{k}_\parallel-\mathbf{k}_\parallel';\mathbf{k}_\parallel;\mathbf{k}_\parallel'}^{c\nu\nu c} P_{\mathbf{k}_\parallel'}^{\nu c} + \text{NR}, \end{aligned}$$

and a similar expression for $\tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu\nu}$. Here, NR refers to non-resonant Auger and pair-creation contributions. As can be recognized, Coulomb renormalizations of the single particle energies are contained in $\tilde{\Omega}_{\mathbf{k}_\parallel}^{cc}$ and $\tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu\nu}$ whereas the attractive electron–hole interaction responsible for the formation of bound excitons is contained in $\tilde{\Omega}_{\mathbf{k}_\parallel}^{\nu c}$. Correlation effects beyond the mean field approximation are contained in

$$i\hbar \left. \frac{d}{dt} P_{\mathbf{k}_\parallel}^{\alpha\alpha'} \right|_{\text{corr}}. \quad (4)$$

It can be shown that the dominant correlation effect is the replacement of the ‘bare’ Coulomb potential in equation (3) by its dynamically screened counterpart, yielding the screened Hartree–Fock approximation [6], whereas if correlation effects are neglected, the renormalizations of

the single particle energies correspond to the Hartree–Fock approximation.

If the system is excited with an optical pulse of central frequency ω_L , the induced optical current is dominated by transition amplitudes $P_{\mathbf{k}_\parallel}^{\alpha\alpha'}$ corresponding to dipole allowed transitions with transition energies $\epsilon_{\alpha'\mathbf{k}_\parallel} - \epsilon_{\alpha\mathbf{k}_\parallel} \approx \hbar\omega_L$ that are resonant or nearly resonant with the optical frequency. If the relevant bands are sufficiently isolated, one can restrict the microscopic analysis to the resonant transition amplitudes and occupation numbers $f_{\mathbf{k}_\parallel}^\alpha = P_{\mathbf{k}_\parallel}^{\alpha\alpha}$ of those bands and the only coupling to remote bands is via screening of the Coulomb potential. Treating screening within the RPA approximation, we can make the separation

$$\Pi_{\mathbf{q}_\parallel}(\omega) = \Pi_{\mathbf{q}_\parallel}^{\text{GS}}(\omega) + \Delta\Pi_{\mathbf{q}_\parallel}(\omega) \quad (5)$$

$$W_{\mathbf{q}_\parallel}(\omega) = \frac{V_{\mathbf{q}_\parallel}}{1 - V_{\mathbf{q}_\parallel}\Pi_{\mathbf{q}_\parallel}(\omega)} = \frac{W_{\mathbf{q}_\parallel}^{\text{GS}}}{1 - W_{\mathbf{q}_\parallel}^{\text{GS}}\Delta\Pi_{\mathbf{q}_\parallel}(\omega)}. \quad (6)$$

Here, $W_{\mathbf{q}_\parallel}^{\text{GS}}$ includes all ground state and environmental screening contributions,

$$\Delta\Pi_{\mathbf{q}_\parallel}(\omega) = \sum_{\alpha\alpha'\mathbf{k}_\parallel} \frac{\Delta f_{\mathbf{k}_\parallel - \mathbf{q}_\parallel}^\alpha - \Delta f_{\mathbf{k}_\parallel}^{\alpha'}}{\hbar\omega + i\gamma_T + \epsilon_{\alpha\mathbf{k}_\parallel - \mathbf{q}_\parallel} - \epsilon_{\alpha'\mathbf{k}_\parallel}} \times |\langle \alpha\mathbf{k}_\parallel - \mathbf{q}_\parallel | \alpha'\mathbf{k}_\parallel \rangle|^2,$$

contains the optically induced deviations from the ground state polarization function only, and γ_T is the triplet dephasing rate, respectively. Hence, if the single particle dispersions, optical dipole and Coulomb matrix elements as well as the ground state screening of the Coulomb potential are known from *ab initio* calculations, the SBE provide a very efficient scheme to compute the optical response since they only need to be solved for a few bands and that part of \mathbf{k}_\parallel space in which populations and/or polarizations are optically induced.

Specifically, for TMDCs it is often sufficient to include the two spin–split valence and conduction bands near the K -points where the single particle dispersion displays a direct gap. Here, the single particle dispersion can be approximated by the relativistic dispersion

$$\epsilon_{i\mathbf{K}+\mathbf{k}}^{e/v} = E_{F,i} \pm \frac{1}{2} \sqrt{\Delta_i^2 + (2\hbar v_{F,i}k)^2}, \quad (7)$$

that results from the widely used massive Dirac–Fermion (MDF) model Hamiltonian⁷. Here, $i = s\tau$ combines the spin and valley index, Δ_i , $v_{F,i}$ and $E_{F,i}$ are the spin and valley dependent gap, Fermi velocity and Fermi level, respectively. The occurring parameters can be adjusted to reproduce the DFT bandstructure near the K -points. The overlap matrix elements resulting from the MDF Hamiltonian are

$$\begin{aligned} \langle c\mathbf{K} + \mathbf{k}_\parallel | c\mathbf{K} + \mathbf{k}_\parallel' \rangle &= \langle \nu\mathbf{K} + \mathbf{k}_\parallel' | \nu\mathbf{K} + \mathbf{k}_\parallel \rangle \\ &= u_{\mathbf{k}_\parallel} u_{\mathbf{k}_\parallel'} + v_{\mathbf{k}_\parallel} v_{\mathbf{k}_\parallel'} e^{-i\theta_{\mathbf{k}_\parallel - \mathbf{k}_\parallel'}} \\ \langle c\mathbf{K} + \mathbf{k}_\parallel | \nu\mathbf{K} + \mathbf{k}_\parallel' \rangle &= \left(u_{\mathbf{k}_\parallel} v_{\mathbf{k}_\parallel'} e^{-i\theta_{\mathbf{k}_\parallel'}} - v_{\mathbf{k}_\parallel} u_{\mathbf{k}_\parallel'} e^{-i\theta_{\mathbf{k}_\parallel}} \right) \end{aligned}$$

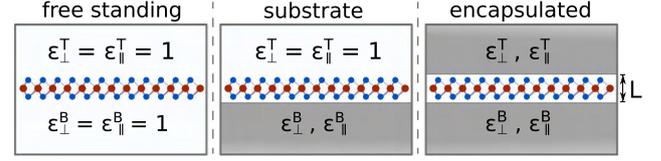


Figure 1. Illustration of the different dielectric settings considered in the calculations presented in this paper.

$$\text{with } u_{\mathbf{k}_\parallel}^2 = 1 - v_{\mathbf{k}_\parallel}^2 = \frac{\sqrt{\Delta_i^2 + (2\hbar v_{F,i}k)^2} + \Delta_i}{2\sqrt{\Delta_i^2 + (2\hbar v_{F,i}k)^2}}.$$

2.2. Coulomb potential and screening in quasi-2D materials

Although the material excitations in quasi-2D materials are confined to a region $|z| \lesssim d/2 \leq L/2$, the Coulomb potential is long ranged and the interaction among particles confined in the layer is sensitive to the dielectric environment. Furthermore, for the evaluation of the opto-electronic system response, one needs the screened Coulomb potential that not only contains screening contributions from the dielectric environment but also all ground state contributions from the material itself.

Within DFT, the macroscopic (3D) dielectric constant is obtained as

$$\epsilon_{\text{M}}^{-1}(\mathbf{q}) = \frac{\langle \tilde{V}_{\mathbf{q}}(\mathbf{r}) \rangle_{\Omega}}{V_0} = \epsilon_{00}^{-1}(\mathbf{q}),$$

where $\tilde{V}_{\mathbf{q}}(\mathbf{r})$ is a lattice periodic function and $\langle \dots \rangle_{\Omega}$ denotes a spatial average over the elementary unit cell. If applied to an artificial 3D crystal consisting of parallel monolayers separated by large vacuum regions, the averaging over a region much larger than the extension of the electron density yields a dielectric function that decreases with the size of the artificial unit cell and hence, cannot be used to construct the screened 2D Coulomb potential [7]. To avoid this complication, we developed in a previous publication [6] an electrostatic model that is based on the bulk dielectric functions which can be directly computed from 3D DFT calculations, thus avoiding any difficulties related to the artificial insertion of large vacuum layers.

Whereas we refer to reference [6] for a detailed derivation of our electrostatic model, we briefly summarize the most important steps with respect to the screened interaction for further use in this paper (figure 1). As visualized in figure 1, we considered TMDCs embedded in different dielectric surroundings, such as in vacuum, on a substrate and embedded in two different dielectric media.

To account for the environmental and ground state screening, we consider a slab geometry with

$$\epsilon_{\parallel}(z) = \begin{cases} \epsilon_{\parallel}^{\text{T}} & z < -L/2, \\ \epsilon_{\parallel} & |z| < L/2, \\ \epsilon_{\parallel}^{\text{B}} & L/2 < z \end{cases} \quad \epsilon_{\perp}(z) = \begin{cases} \epsilon_{\perp}^{\text{T}} & z < -L/2, \\ \epsilon_{\perp} & |z| < L/2, \\ \epsilon_{\perp}^{\text{B}} & L/2 < z. \end{cases}$$

and determine the interaction potential as solution of Poisson’s equation. The resulting interaction potential within the slab is given by

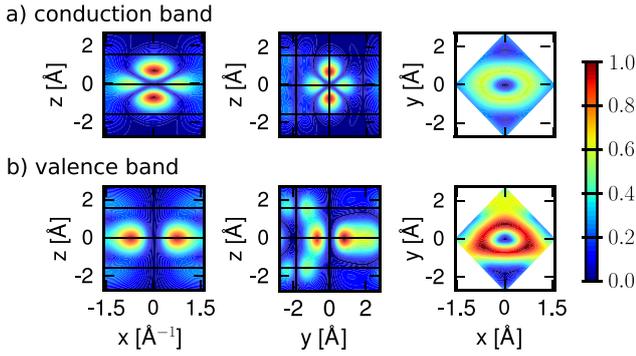


Figure 2. Probability density $\rho = |\psi(\mathbf{r})|^2$ for (a) the conduction and (b) the valence band of MoS₂ at the K -point plotted for cuts through the unit cell along the x -, y - and z -axis centered around the molybdenum atom. The black lines indicate the atomic z -position and illustrate the wave function localization in between the chalcogen layers.

$$\begin{aligned} \phi_{\mathbf{q}_{\parallel}}(z, z') &= \frac{2\pi}{\kappa q} \left(e^{-\tilde{q}|z-z'|} \right. \\ &\quad + c^{+-} e^{-\tilde{q}(L+z+z')} + c^{-+} e^{-\tilde{q}(L-z-z')} \\ &\quad \left. + c^{--} e^{-\tilde{q}(2L-z+z')} + c^{--} e^{-\tilde{q}(2L+z-z')} \right) \\ &= \frac{2\pi}{\kappa q} e^{-\tilde{q}|z-z'|} + \Delta\phi_{\mathbf{q}_{\parallel}}(z, z') \end{aligned} \quad (8)$$

with $\tilde{q} = \sqrt{\frac{\epsilon_{\parallel}}{\epsilon_{\perp}}} q$, $\kappa = \sqrt{\epsilon_{\parallel}\epsilon_{\perp}}$ and similar for $\kappa_{T/B}$ and

$$c^{m'} = \frac{(\kappa + \eta\kappa_T)(\kappa + \eta'\kappa_B)}{(\kappa + \kappa_T)(\kappa + \kappa_B) - (\kappa - \kappa_T)(\kappa - \kappa_B) e^{-2\tilde{q}L}},$$

from which the 2D Coulomb potential is obtained as $\phi_{\mathbf{q}_{\parallel}}^{2D} = \phi_{\mathbf{q}_{\parallel}}(0, 0)$. In the last line of equation (8), the first term describes the direct interaction of two electrons located at z and z' whereas the second term describes the interaction with image charges. From equation (8), the ‘bare’ Coulomb interaction $V_{\mathbf{q}_{\parallel}}(z, z')$, containing environmental screening only, is obtained by setting $\epsilon_{\parallel} = \epsilon_{\perp} = 1$ within the slab, whereas the true vacuum Coulomb interaction $V_{\mathbf{q}_{\parallel}}^{\text{vac}}(z, z')$ is obtained by additionally setting $\kappa_T = \kappa_B = 1$. Hence, equation (8) suggests that the Coulomb potential without intrinsic screening can be expressed as $V_{\mathbf{q}_{\parallel}}(z, z') = V_{\mathbf{q}_{\parallel}}^{\text{vac}}(z, z') + \Delta V_{\mathbf{q}_{\parallel}}(z, z')$, where the last term describes the interaction via image charges. Similarly, one obtains the screened Coulomb interaction in bulk, $W_{\mathbf{q}_{\parallel}}^{\text{bulk}}(z, z')$, by inserting the bulk parameters for the in- and out-of-plane dielectric constants $\epsilon_{\parallel}^{\text{B}}$ and $\epsilon_{\perp}^{\text{B}}$ both within the slab and for the top and bottom environment.

Whereas both $V_{\mathbf{q}_{\parallel}}^{\text{vac}}(z, z')$ and $W_{\mathbf{q}_{\parallel}}^{\text{bulk}}(z, z')$ are independent of the slab thickness L , the interaction with image charges in the presence of environmental screening introduces a dependence on the slab thickness L both for the ‘bare’ and screened monolayer interaction potential. To model the monolayer potential, we shall assume a slab thickness $L = D$ where $D = c/2$ is the natural layer-to-layer distance in the naturally stacked bulk crystal. In the long wavelength limit

$q_{\parallel}D \rightarrow 0$ one obtains $W_{\mathbf{q}_{\parallel}}^{\text{GS}, 2D} = \frac{2\pi}{q(\kappa_T + \kappa_B)/2 + r_0 q}$, corresponding to the widely used Rytova–Keldysh potential with a screening length $r_0 = D(2\kappa^2 - \kappa_T^2 - \kappa_B^2)/4\epsilon_{\perp}$ that depends on the dielectric contrast between the TMDC material and dielectric environment. In particular, if the dielectric environment is the bulk TMDC crystal itself, the screening length vanishes. For the short wavelengths $q_{\parallel}L/2 \rightarrow \infty$, the screened potential approaches $2\pi/\kappa q e^{-\tilde{q}|z-z'|}$, independently of the dielectric environment. Comparison with the bulk Coulomb potential allows us to identify $\epsilon_{\parallel} = \epsilon_{\parallel}^{\text{B}}$ and $\epsilon_{\perp} = \epsilon_{\perp}^{\text{B}}$ for the ground state screening even in the case of a monolayer. Hence, we can write for the Coulomb potential including environmental and ground state screening

$$W_{\mathbf{q}_{\parallel}}^{\text{GS}}(z, z') = W_{\mathbf{q}_{\parallel}}^{\text{bulk}}(z, z') + \Delta W_{\mathbf{q}_{\parallel}}(z, z'), \quad (9)$$

where $\Delta W_{\mathbf{q}_{\parallel}}(z, z')$ describes the screened interaction via image charges. Equation (9) should be compared with the division $W = W^{\text{ML}} + \Delta W$ that has been used in the $G\Delta W$ approach where ΔW contains the changes in the screened potential as compared to the suspended monolayer [7].

2.3. Form factor to quantify the impact of three-dimensionality

Even monolayer TMDCs have an intrinsic thickness since they consist of three atomic layers and their atomic orbitals have a finite extension perpendicular to the layer plane. Consequently, the Coulomb interaction in these materials differs both from the exact 2D and the three-dimensional cases. Taking the in-plane periodicity and the finite out-of-plane extension into account, a 2D ansatz gives the Bloch states,

$$\phi_{\alpha\mathbf{k}_{\parallel}}(\mathbf{r}) = \frac{1}{\sqrt{A}} e^{i\mathbf{k}_{\parallel}\cdot\mathbf{r}} u_{\alpha\mathbf{k}_{\parallel}}(\mathbf{r}) \quad (10)$$

with strictly 2D crystal momenta but 3D spatial coordinates \mathbf{r} . Hence, the Coulomb Hamiltonian contains the quasi-2D matrix elements

$$\begin{aligned} V_{\mathbf{k}_{\parallel}'\mathbf{k}_{\parallel}:\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'} &= \int_{ec} d^3r \int_{ec} d^3r' u_{\alpha\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}}^*(\mathbf{r}) u_{\beta\mathbf{k}_{\parallel}'+\mathbf{q}_{\parallel}}^*(\mathbf{r}') \\ &\quad \times V_{\mathbf{q}_{\parallel}}(z, z') u_{\beta'\mathbf{k}_{\parallel}'}(\mathbf{r}') u_{\alpha'\mathbf{k}_{\parallel}}(\mathbf{r}), \end{aligned}$$

and a similar expression holds for the matrix elements of the screened Coulomb potential.

According to equation (8), the Coulomb interaction differs from the exact two-dimensional potential only in the exponential terms. Hence, we can write for the screened Coulomb potential

$$W_{\mathbf{k}_{\parallel}'\mathbf{k}_{\parallel}:\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'} = W_{\mathbf{q}_{\parallel}}^{2D} F_{\mathbf{k}_{\parallel}'\mathbf{k}_{\parallel}:\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel}) \quad (11)$$

$$- \Delta W_{\mathbf{q}_{\parallel}}^{2D} \Delta F_{\mathbf{k}_{\parallel}'\mathbf{k}_{\parallel}:\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel}). \quad (12)$$

Here,

$$\begin{aligned} F_{\mathbf{k}_{\parallel}'\mathbf{k}_{\parallel}:\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel}) &= \int_{ec} d^3r \int_{ec} d^3r' u_{\alpha\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}}^*(\mathbf{r}) u_{\beta\mathbf{k}_{\parallel}'+\mathbf{q}_{\parallel}}^*(\mathbf{r}') \\ &\quad \times e^{-\tilde{q}_{\parallel}|z-z'|} u_{\beta'\mathbf{k}_{\parallel}'}(\mathbf{r}') u_{\alpha'\mathbf{k}_{\parallel}}(\mathbf{r}) \end{aligned} \quad (13)$$

is a form factor and

$$\Delta F_{\mathbf{k}'_{\parallel};\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel}) = F_{\mathbf{k}'_{\parallel};\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel}) - f_{\mathbf{k}_{\parallel}}^{\alpha\alpha'}(\mathbf{q}_{\parallel})f_{\mathbf{k}'_{\parallel}}^{\beta\beta'}(\mathbf{q}_{\parallel}) \quad (14)$$

$$f_{\mathbf{k}_{\parallel}}^{\alpha\alpha'}(\mathbf{q}_{\parallel}) = \int_{ec} d^3r u_{\alpha\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}}^*(\mathbf{r})e^{-\tilde{q}_{\parallel}z}u_{\alpha'\mathbf{k}_{\parallel}}(\mathbf{r}) \quad (15)$$

is its correlated part, respectively. In the 2D limit, the exponential term in the form factor approaches unity and the 2D Coulomb matrix element is recovered due to the wavefunctions orthonormality. Furthermore, $F_{\mathbf{k}'_{\parallel};\mathbf{k}_{\parallel}}^{\alpha\beta\beta'\alpha'}(\mathbf{q}_{\parallel} = 0) = \delta_{\alpha\alpha'}\delta_{\beta\beta'} = f_{\mathbf{k}_{\parallel}}^{\alpha\alpha'}(\mathbf{q}_{\parallel} = 0)f_{\mathbf{k}'_{\parallel}}^{\beta\beta'}(\mathbf{q}_{\parallel} = 0)$, while for large scattering vectors $\mathbf{q}_{\parallel} \rightarrow \infty$ the form factor approaches 0. Since the potential of the image charges is particularly important in the long wavelength limit where the correlated part of the form factor is negligible, the Coulomb matrix elements of the quasi-2D TMDC structures can be expressed in a good approximation by the exact 2D term modified by the form factor.

2.4. Combined SBE/DFT approach

In the following, we will assume that the single-particle band dispersion and wave functions can reliably be computed from DFT and use the DFT parameters as input for the SBE/DBE evaluations. From a fundamental point, this is in fact a non-trivial assumption. As is well known, DFT is based on the assumption that the two-particle interaction in any given system is a *universal* functional of the electron density while all system specific contributions to the many-body Hamiltonian are contained in the external potential. Using the dielectric model presented in section 2.2, it becomes clear that the substrate not only changes the screened Coulomb potential but also modifies the ‘vacuum’ potential that now contains interactions with image charges. Thus, in the presence of external screening, all particles (electrons and ions) interact via a modified Coulomb potential and hence, the electron–electron interaction can no longer be considered as universal.

However, analytical estimates based on the expression for the ‘bare’ Coulomb potential presented in section 2.2 show that the influence of the dielectric screening on the DFT single particle energies should be small for particles confined to a region $|z| \lesssim d/2 < D/2$. Indeed, rigorous treatments of monolayer TMDCs embedded in different dielectric environments have shown that the single-particle bandstructure in the proximity of the direct band gap at the K point remains unchanged for different environments [10]. Therefore, the single-particle bandstructure near the K -point can be obtained from an artificial 3D supercell calculation even in the presence of external screening.

3. Details of DFT computations

For our DFT calculations, we use the plane-wave based code Vienna *ab initio* Simulation Package (VASP) [11–14] including the core electrons contribution by precalculated projector augmented-wave pseudopotentials [15]. All calculations were performed using the non-empirical generalized gradient Perdew–Burke–Ernzerhof (PBE) functional [16], with

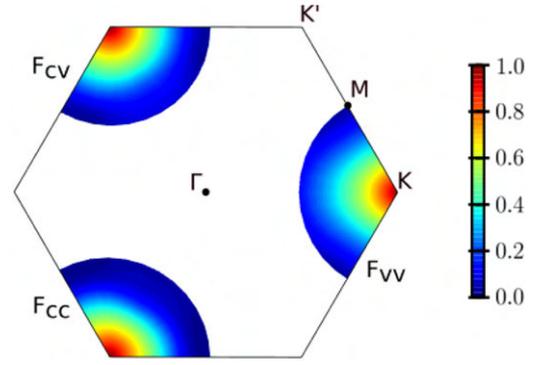


Figure 3. Form factor for inter- and intraband interaction of MoS₂ in a range of $|\overline{KM}| = |\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}| \approx 0.62 \text{ \AA}^{-1}$ around the K -point. As path dependent changes of the form factor are only small, they are approximated as isotropic around the K -point and subsequent calculations are performed using path-averaged form factors.

additionally including spin–orbit coupling and van-der-Waals interactions by the dispersion correction as proposed by Grimme (DFT-D3) [17].

In a first step, the materials structure was relaxed until remaining forces acting on the atoms were less than 1 meV \AA^{-1} . The number of plane waves was hereby restricted by a cutoff energy of 750 eV and convergence was checked with respect to the discretization of the Brillouin zone and the vacuum region that was added in z -direction to prevent interactions between the monolayers despite periodic boundary conditions. For further calculations of electronic properties, the Brillouin zone was discretized by a $12 \times 12 \times 1$, Γ centered k -point Monkhorst–Pack [18] grid and the energy cutoff was set consistently with the precalculated pseudopotentials. In practice, approximately 15 \AA of vacuum were included in z -direction.

Using the relaxed structures, the electronic properties were calculated for different paths in the Brillouin zone originating in one of the K points where the direct TMDC band gap is found. The paths were chosen along \overline{KM} in addition to paths of the same length but rotated by 30° and 60° which, due to the hexagonal symmetry, is sufficient to describe the behavior around each K point in steps of 30° .

As the Kohn–Sham wave functions are determined up to an arbitrary phase that can be chosen at each \mathbf{k}_{\parallel} -point independently, we pick the phases such that the intra- and interband matrix elements are given by $\mathbf{p}_{\alpha\alpha\mathbf{k}_{\parallel}} = \frac{m_0}{\hbar} \vec{\nabla}_{\mathbf{k}_{\parallel}} \epsilon_{\alpha\mathbf{k}_{\parallel}}$ and $\mathbf{p}_{\alpha\alpha'\mathbf{k}_{\parallel}} = \frac{m_0}{\hbar} (\epsilon_{\alpha'\mathbf{k}_{\parallel}} - \epsilon_{\alpha\mathbf{k}_{\parallel}}) \langle \alpha\mathbf{k}_{\parallel} | \vec{\nabla}_{\mathbf{k}_{\parallel}} | \alpha'\mathbf{k}_{\parallel} \rangle$ respectively, corresponding to the $\mathbf{p} \cdot \mathbf{A}$ gauge. The remaining free global phase is chosen such that bright excitons are predominantly s -type.

3.1. Ab initio wave functions

The finite extension of the electron density in the direction perpendicular to the TMDC plane is illustrated in figure 2(a) for the example of MoS₂. Here, the normalized density distribution for both valence and conduction band at the K -point is plotted for different cuts through the unit cell.

For MoS₂, the wave functions are linear combinations of the transition metal’s d -orbitals that are strongly localized in between the chalcogenite layers. While the conduction-band

wave function is dominated by the d_z -orbital, the wave function of the valence band is a linear combination of the in-plane d_{xy} - and $d_{x^2-y^2}$ -orbitals [19]. This is different in tungsten based materials where the transition metal's s -orbital and the chalcogen atom's p_z -orbital both have a finite contribution to the conduction or valence band respectively. These features result in a relatively higher density distribution in the tungsten layer for the conduction band whereas for the valence band the density distribution extends farther in the direction towards the chalcogen atoms, respectively.

4. Analysis of the form factor

4.1. Path-dependency

For the optical response, the properties of the form factor are of interest especially close to the direct band. Therefore, we compute the form factor for scattering vectors $|\mathbf{k}_\parallel - \mathbf{k}'_\parallel|$ with \mathbf{k}_\parallel fixed at the K -point and \mathbf{k}'_\parallel modified along different paths through the first Brillouin zone. The band structure is similar for all paths close to the K -point but changes significantly farther away, particularly for the conduction band where in the $K + 60^\circ$ direction a secondary minimum can be seen that forms the indirect band gap in the reference bulk structure.

In Figure 3, we show a density plot of the form factor for inter- and intraband interaction in the region around the K -points for the example of MoS_2 . Despite the anisotropy of the band structure, the MoS_2 form factor of the interband interactions F_{cv} decreases isotropically around the K -point up to a range of $|\mathbf{k}_\parallel - \mathbf{k}'_\parallel| \approx 0.45 \text{ \AA}^{-1}$. The same is true for tungsten based materials, but only up to $|\mathbf{k}_\parallel - \mathbf{k}'_\parallel| \approx 0.30 \text{ \AA}^{-1}$. Thereafter, the gradient in K -direction is smaller than in Γ -direction, whereby the absolute differences are less than 0.06. Path dependent differences are larger with regard to the intraband interactions. Here, the gradient is smaller in the direction towards the Γ -point for the valence and towards the M -points for the conduction band, respectively. However, but even here, the form factor is isotropic in a range of $|\mathbf{k}_\parallel - \mathbf{k}'_\parallel| \approx 0.15 \text{ \AA}^{-1}$ and absolute differences are always less than 0.12. Thus, the differences occurring for the intraband form factor depending on different orbital compositions of the states seem to balance out in the integration over both valence and conduction band states in the interband form factor. In the following, averages over calculations along different paths are shown and used in subsequent calculations.

4.2. Supercell dependency

In order to be useful, the introduced form factor should be insensitive to the periodic boundary conditions applied in the DFT code and it should be applicable not only to model the Coulomb interaction in monolayers but also in bulk structures. To check these properties, we calculated the form factor for different extensions of the vacuum region included in the supercell.

Our results show that as long as we properly relax the structure before calculating the form factor we obtain similar results independent of the used super cell size. The same is true even if we consider the transition to a quasi-bulk structure differing

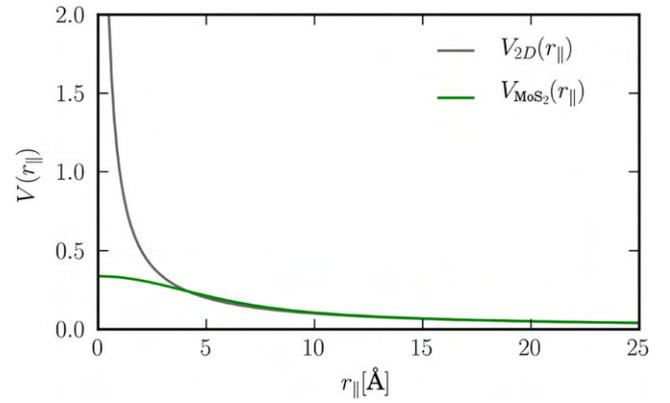


Figure 4. Comparison of the exact 2D Coulomb potential and the quasi-2D Coulomb potential in real space for the example of MoS_2 .

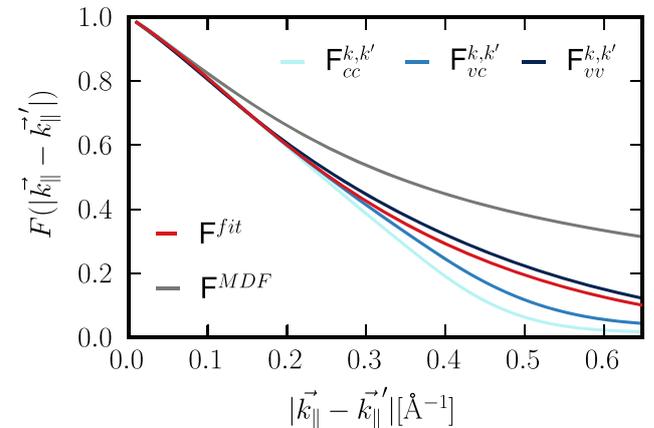


Figure 5. Comparison of the interband (subscript vc) and intraband (subscripts cc and vv) form factors together with the analytical approximation (red curve). The grey curve shows $F^{\text{MDF}} = u_{\mathbf{k}_\parallel}^2 \int_{ec} d^3 r \int_{ec} d^3 r' u_{\alpha\mathbf{K}}^*(\mathbf{r}) u_{\beta\mathbf{K}}^*(\mathbf{r}') e^{-q_\parallel |z-z'|} u_{\beta'\mathbf{K}}(\mathbf{r}') u_{\alpha'\mathbf{K}}(\mathbf{r})$.

from the common crystalline structure of MoS_2 in stacking order (usually 2H-/3R-phases) and in the interlayer distance by about 0.6 \AA . On this basis, we conclude that we can use the determined form factors to model a wide range of mono- and multi-layer systems in different dielectric environments.

4.3. Analytical description

To illustrate the effect of the form factor in real space, we compare in figure 4 the exact 2D and the quasi-2D Coulomb potentials for MoS_2 . As we can see, the form factor effectively removes the Coulomb singularity at the origin but leaves the potential unchanged for larger distances, i.e., for $r_\parallel \gtrsim 7 \text{ \AA}$. the slope of the quasi-2D potential approaches that of the original 2D potential.

In figure 5, we show examples of the momentum dependent form factors for inter- and intraband transitions that were numerically evaluated using the full wavefunctions. To simplify detailed calculations on the SBE/DBE level, it is useful to develop analytical approximations for the form factor. For this purpose, we replace the full wavefunctions by their respective values at the K -point and incorporate the orbital dependent overlap matrix elements that occur within the MDF model.

As shown by the green curve in figure 5, this procedure captures the main features of the full form factors for small scattering vectors. However, for scattering vectors larger than 0.1 \AA^{-1} or 0.2 \AA^{-1} clear deviations are seen such that additional corrections are needed.

In some of our previous works (see e.g. references [6, 20]), we adopted a semi-empirical Ohno potential to account for the finite thickness of the TMDC layers. This approximation has also been used to describe the Coulomb potential of molecules and nanotubes (see references [21, 22]) and was successfully applied in calculations of optical properties of graphene and monolayer TMDCs.

By using the Ohno potential, one regularizes the singularity of the Coulomb potential by introducing an effective thickness parameter d , which in reciprocal space occurs in an additional prefactor $e^{-q_{\parallel}d}$ of the Coulomb potential. Even though this ansatz reduces the Coulomb scattering for larger scattering vectors, the detailed analysis shows that it underestimates the form factor for small scattering vectors. Therefore, we additionally introduce an orbital independent exponential function of the form $F(\mathbf{q}_{\parallel}) = e^{-(\sigma^2 q_{\parallel}^2/2 + q_{\parallel}d)}$, while the orbital dependency is given by the MDF overlap matrix elements. With help of this function, the form factor can be described properly and physically correctly as it approximates 0 for large scattering vectors. Since the orbital dependence is contained in the MDF overlap matrix elements, the orbital independent part can be combined with the strictly 2D Coulomb potential to define the quasi-2D Coulomb potential $V_{\mathbf{q}_{\parallel}} = F(\mathbf{q}_{\parallel})V_{\mathbf{q}_{\parallel}}^{2D}$.

A physical interpretation of the quadratic contribution can be obtained by considering the real space representation of the quasi-2D Coulomb potential:

$$V(\mathbf{r}_{\parallel}) = \sum_{\mathbf{q}_{\parallel}} e^{i\mathbf{q}_{\parallel} \cdot \mathbf{r}_{\parallel}} F(\mathbf{q}_{\parallel}) V_{\mathbf{q}_{\parallel}}^{2D} = \frac{1}{2\pi\sigma^2} \int d^2r'_{\parallel} e^{-|\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}|^2/2\sigma^2} \frac{1}{\sqrt{r_{\parallel}^2 + d^2}}. \quad (16)$$

Hence, the linear term in the exponent of the form factor changes the pure 2D Coulomb potential into the Ohno potential with thickness d , whereas the quadratic term lead to an additional convolution with a Gaussian of with σ .

4.4. Material dependence of the form factor

In figure 6, we show a comparison of the interband form factor for five commonly discussed examples of semiconducting TMDCs. We notice that for tungsten based materials the form factor decreases slightly faster than for molybdenum based materials. Even more significant modifications occur for different chalcogen atoms. The comparison in figure 6 reveals a more rapid form factor decay when changing from S to Se to Te demonstrating that the gradient becomes steeper with increasing atom size. Since the microscopic distance between the chalcogenide sheets increases with the size of the involved atoms, the systems become slightly more 3D which is reflected in the form factor being a measure for the influence of the finite thickness.

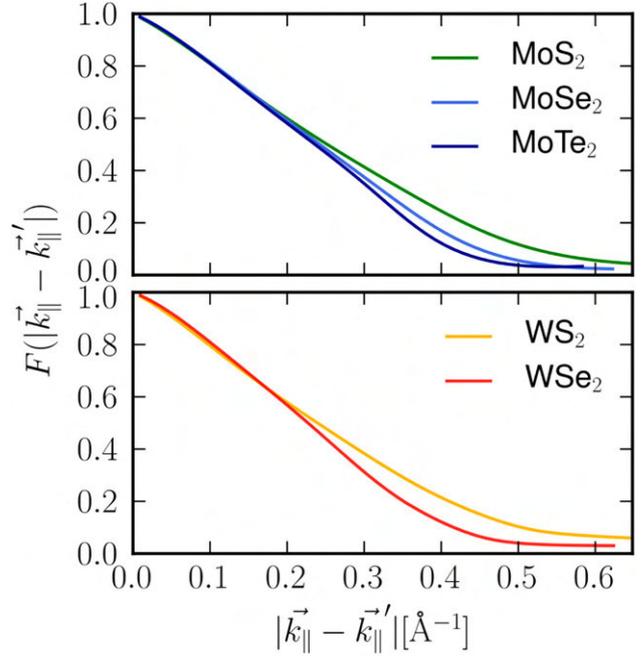


Figure 6. Interband form factors for different (a) Mo based and (b) W based TMDC monolayers. While the general slope is similar for all materials, the decay steepens with increasing size of the respective chalcogenide atoms.

5. Applications

5.1. Environmental dependent band gap renormalization

As a first application, we calculate the renormalization of the direct K point band gap for different dielectric environments. For the unexcited monolayer, the renormalized band gap is given by

$$\Delta_{\mathbf{K}} = \Delta^{\text{DFT}} + \sum_{\mathbf{q}_{\parallel}} \left(W_{\mathbf{q}_{\parallel}; \mathbf{K}; \mathbf{K} + \mathbf{q}_{\parallel}}^{vvvv} - W_{\mathbf{q}_{\parallel}; \mathbf{K}; \mathbf{K} + \mathbf{q}_{\parallel}}^{ccvv} \right) \quad (17)$$

In figure 7, we show the dependency of the gap on the environmental screening for five widely studied semiconducting monolayers (upper part). Since we plot the gap against the inverse value of the substrate dielectric constant, the figure covers the whole range of $\{1, \infty\}$ and $\{-\infty, -10\}$. For all materials investigated, we find a similar gap reduction for increased environmental screening. In the lower part of figure 7, we compare the gap shift of the different materials relative to the respective suspended monolayer with and without the form factor included. We note that the shift is almost independent of the form factor. To understand this behavior, we write the equation for the renormalized gap as

$$E_G = \Delta^{\text{DFT}} + (W_{\mathbf{K}; \mathbf{K}}^{vvvv}(\mathbf{r}_{\parallel} = 0) - W_{\mathbf{K}; \mathbf{K}}^{ccvv}(\mathbf{r}_{\parallel} = 0)) \quad (18)$$

and make use of equation (16).

Thus, for a strictly 2D system, the gap renormalization is determined by the screened Coulomb matrix elements at $\mathbf{r}_{\parallel} = 0$ and $z = z' = 0$. However, the form factor replaces $z = z' = d$ in the value of the screened Coulomb potential and additionally averages the Coulomb potential over a region

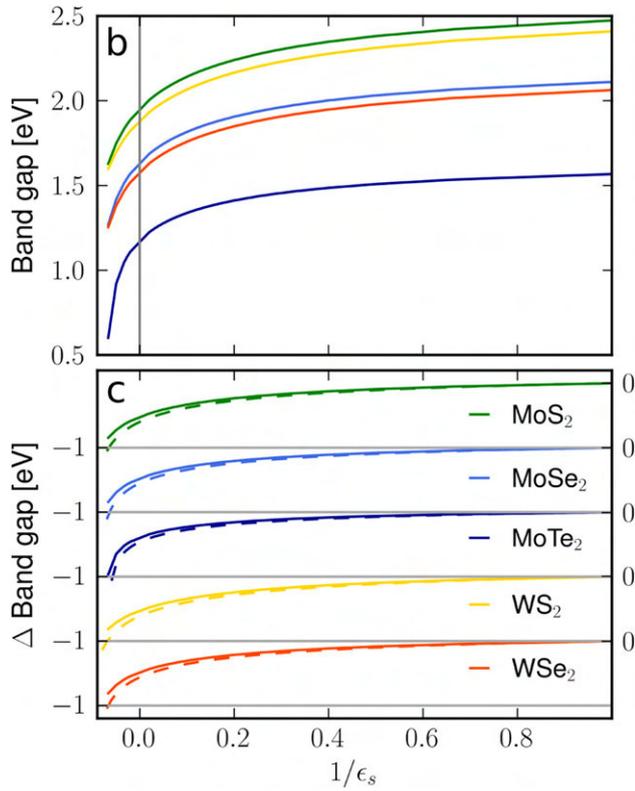


Figure 7. Dependence of the band gap on dielectric substrate screening of the five canonical TMDCs. Here, the limit $\epsilon_s \rightarrow -\infty$ represents the case of a conducting metallic substrate. The upper part (a) shows the absolute value of the gap, the lower part (b) the shift of the gap relative to the gap of a suspended monolayer with (full lines) and without (dashed lines) inclusion of the form factor.

around the origin using a Gaussian weight function. Both effects have a significant impact on the direct interaction contained in $W^{2\text{Dbulk}}(r_{\parallel}) = 1/\kappa r_{\parallel}$ that varies strongly in the region around the origin. In contrast, changes contained in $\Delta W \propto 1/\sqrt{r_{\parallel}^2 + L^2}$ vary only weakly within the region around the origin and hence are less affected by the form factor.

Furthermore, a comparison of the relative changes for the different materials in the regime of positive dielectric constants shows that these are slightly smaller for Mo rather than for W based materials with maximum differences in the range of few meV. Again, the influence of the incorporated chalcogenide atoms is more significant. Here, the maximum relative differences increase from Te to Se to S based compounds by approximately 80 meV and 45 meV, respectively.

In table 1, we give an overview of the computed band gaps for a variety of TMDCs and different dielectric environments. Comparison with available experimental data shows excellent agreement for all three Mo-based material systems. E.g. for MoS₂ we find $E_G = 2.025$ eV for the direct gap at the *K*-points in bulk, whereas we find $E_G = 2.473$ eV for the freely suspended, $E_G = 2.279$ eV on top of a fused silicon substrate, and $E_G = 2.160$ eV for an hBN encapsulated monolayer. The gap on a metal substrate (e.g. gold) can be estimated from the limit $1/\epsilon_s \rightarrow 0^-$, yielding $E_G = 1.944$ eV. These values are in excellent agreement with available experimental data, listed also in table 1. In contrast, the gaps for the W-based

materials appear to be 130–150 meV below reported experimental values. Comparison with the MoX₂ systems shows that the predicted gaps for the corresponding WX₂ and MoX₂ are very similar, whereas the experimentally reported WX₂ gaps are always above those of the corresponding MoX₂ systems under comparable conditions. Since the environmentally induced band gap changes are captured very well by our analysis, we speculate that the origin for the systematic underestimation of the gap in the W-based systems is most likely at the DFT level.

For the suspended monolayers, we can also compare our results with other *ab initio* methods. Reported quasi-particle gaps based on the GW/GdW approach vary over a wide range of up to 400 meV and we find our results at the lower end of this spectrum. Regarding the relative differences in the gaps of the W- and Mo-based monolayers, the LDA and PBE based GW results show the same trend as our results, namely almost equal gaps for the sulfides and the selenides, thus supporting the assumption that the systematic discrepancy for the WX₂ originates at the DFT level. Regarding the shifts of the band gap with increased dielectric screening, our results agree well with the available GW-based predictions.

5.2. Density dependent band gap renormalization

As a second application, we calculate the band gap renormalization resulting from finite carrier densities in different TMDCs. In the presence of excited carriers, the renormalized band gap

$$\Delta_{\mathbf{k}} = \Delta^{\text{DFT}} + \sum_{\mathbf{q}_{\parallel}} \left(W_{\mathbf{q}_{\parallel};\mathbf{K};\mathbf{K}+\mathbf{q}_{\parallel}}^{\nu\nu\nu\nu} - W_{\mathbf{q}_{\parallel};\mathbf{K};\mathbf{K}+\mathbf{q}_{\parallel}}^{\text{cc}\nu\nu} \right) \times \left(f_{\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}}^{\nu} - f_{\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}}^{\text{c}} \right). \quad (19)$$

contains the dynamically screened Coulomb matrix elements $W_{\mathbf{q}_{\parallel}}(\hbar\omega - \tilde{\epsilon}_{\mathbf{c}\mathbf{K}+\mathbf{k}_{\parallel}-\mathbf{q}_{\parallel}} + \tilde{\epsilon}_{\nu\mathbf{K}+\mathbf{k}_{\parallel}})$, that, in turn, contain the renormalized single particle dispersions $\tilde{\epsilon}_{c/\nu\mathbf{k}_{\parallel}}$. These equations are solved self consistently together with equation (6), assuming thermal carrier distributions at 300 K within the renormalized bands. Here, we distinguish between optically excited carrier densities (equal numbers of electrons and holes) and carrier densities due to doping, where we restrict the analysis to electron doping only. In the presence of excited carriers, both, screening of the Coulomb potential by the excited carriers and phase space filling contribute to the conduction and valence band renormalization. In contrast, for the case of electron doping, the valence band renormalization is solely due to screening effects corresponding to the Coulomb hole, whereas the phase space filling effects additionally contribute to the conduction band renormalization.

The interplay between phase space filling and screening contributions sensitively depends on the employed screening model. In particular, the approximation of static screening leads to an overestimation of screening effects. To show the importance of dynamical screening, we compare the band gap of a SiO₂ supported MoS₂ monolayer in figure 8 for different carrier excitation conditions. Using the static approximation for the screened Coulomb interaction, the gap decreases

Table 1. Substrate dependent quasiparticle band gap of five semiconducting TMDC systems. The band gap is given for freely suspended, a substrate supported (fused silicon and metal), and hBN encapsulated monolayers. Additionally, literature values based on experimental and/or theoretical investigations are listed for comparison. Here, we distinguish between theoretical GW/GdW calculations building on DFT calculations employing different functionals, namely a local density approach (LDA), the generalized gradient approximation parametrized as by PBE or the hybrid functional Heyd–Scuseria–Ernzerhof (HSE). Additionally, results based on the GvJ-2e approach are summarized. Furthermore, concerning experimental results, we distinguish between photoluminescence (excitation) [PL(E)] measurements yielding the optical band gap from which the quasiparticle gap is extrapolated on the basis of a model for the exciton resonance energies, and scanning tunneling spectroscopy (STS) and angle-resolved photoemission spectroscopy (ARPES) measurements.

Material	Substrate	E_g (eV)	Theor.			Exp.	
MoS ₂		2.473	LDA + GW/GdW 2.8 [10], 2.9 [23],	PBE/HSE + GW/GdW 2.55 (@300 K), 2.63 (@0 K) [24], 2.53 [26], 2.62 [27], 2.54/2.65 [28]	GvJ-2e 2.38 [25]	PL/PLE + model	STS/ARPES
	SiO ₂	2.279					2.17(4) [29]
	hBN encaps.	2.16				2.16 [30], 2.146 [29]	
	Au	1.945	2.3 [10]				1.95(5) [32], 1.90 [33]
MoSe ₂		2.111	2.6 [23], 2.26 [34],	2.12 [26], 2.24 [35]	2.03 [25]		
	SiO ₂	1.941					
	hBN encaps.	1.835				1.874 [30]	
MoTe ₂	Au	1.628					
	SiO ₂	1.57		1.72 [35], 1.72 [36]			
	hBN encaps.	1.444				1.352 [30]	
WS ₂	Au	1.356					
	SiO ₂	1.166					
	hBN encaps.	2.41	2.81 [23]	2.53 [26], 2.83 [35]	2.51 [25]	2.73 [37], 3.01 [38], 2.37 [3], 2.40 [39] 2.238 [30]	
WSe ₂	Au	2.084					
	SiO ₂	1.877					
	hBN encaps.	2.063	2.4 [23]	2.35 [40]	2.11 [25]	2.63 [38], 2.35(20) [41], 2.02 [5]	2.38(6) [29]
	SiO ₂	1.885					
	hBN encaps.	1.777		2.22 [40]		1.884 [42], 1.890 [43]	
	Au	1.573					1.75 [33]

very rapidly with increasing carrier density. Here, the decrease depends only little on the fact whether the carriers are generated by optical excitation or doping, indicating that the gap renormalization is completely dominated by the Coulomb hole. In contrast, for the case of dynamic screening, the gap decreases more slowly with increasing carrier density and is larger for the symmetrical, optically induced electron–hole densities than for electron densities alone.

Using the dynamical screening calculations, we compare in figure 9 the density dependent gap for the five investigated monolayer materials assuming optically induced electron–hole densities (solid lines) and electron densities only (dashed lines). Whereas the overall behavior is similar in all materials, the effects are somewhat more pronounced in Mo than in based systems, also increasing from S to Te. Furthermore, the renormalization depends on the distribution of carrier densities, thus the band gap changes are weaker due to doping densities than due to optically excited carriers. This effect is more pronounced in than in Mo based materials.

Finally, we compare the influence of excited carriers on the band gap of MoS₂ for different dielectric environments in

figure 10. As we can see, the environmentally induced band gap offset vanishes rapidly with increasing carrier density, showing that density dependent screening effects dominate the band gap over dielectric screening effects for moderate to large carrier densities.

5.3. Exciton resonances and binding energies

As third and final application, we calculate the resonance energies of the A exciton series for different TMDCs in various dielectric environments. Similar to the band gap renormalization, we find that the exciton binding energy decreases due to enhanced screening of the Coulomb interaction in media with increased dielectric constant. For the 1s exciton, the decrease in the band gap renormalization and binding energy nearly cancel. As a net result, the 1s exciton shows only a small red shift with increased dielectric screening. In contrast, for the higher exciton states, the band gap renormalization dominates over the decreased binding, yielding a pronounced red shift of the excited states with enhanced dielectric screening.

As representative examples, we show in figure 11 the excitonic resonances generated by solving the Dirac–Wannier

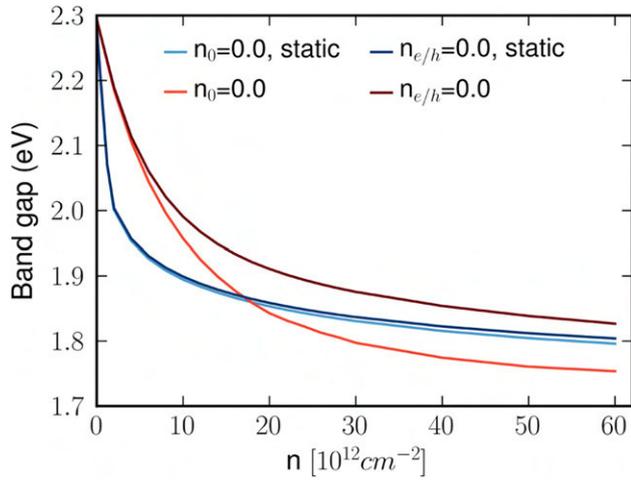


Figure 8. Comparison of the band gap renormalization in dependence of the carrier density n for static and dynamic screening for the example of a quartz supported MoS₂ monolayer. Here, $n_0 = 0$ indicates the absence of doping and $n_{e/h} = 0$ denotes the absence of optically excited carriers, respectively.

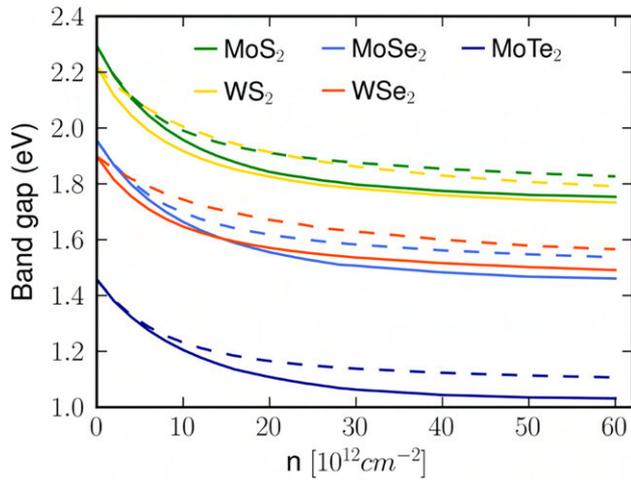


Figure 9. Dependence of the band gap renormalization on the carrier density for dynamical screening with $\gamma_T = 100$ meV for different transition metal dichalcogenides. Solid lines mark calculations with $n_0 = 0$ cm⁻², dashed lines belong to calculations with $n_{e/h} = 0$ cm⁻².

equation for MoS₂ and WS₂ on a quartz substrate and for an hBN encapsulated configuration. Here, the four energetically lowest resonances are shown together with the quasiparticle band gap. As mentioned earlier, the gap of the W-based systems is systematically underestimated such that we have to shift the DFT gap by 149 meV for WS₂ while keeping the dipole matrix elements constant BEFORE we compute the renormalizations and exciton binding energies. While the band gap decreases from the SiO₂ to the hBN encapsulated sample by about 110 meV, the $1s$ resonance energy decreases only by less than a fourth of this amount. Thus, the binding energy is reduced by about 90–100 meV for WS₂ and MoS₂ respectively. In contrast, the binding energies of the $2s$, $3s$, $4s$ excitons are reduced by 40 – 20 meV only.

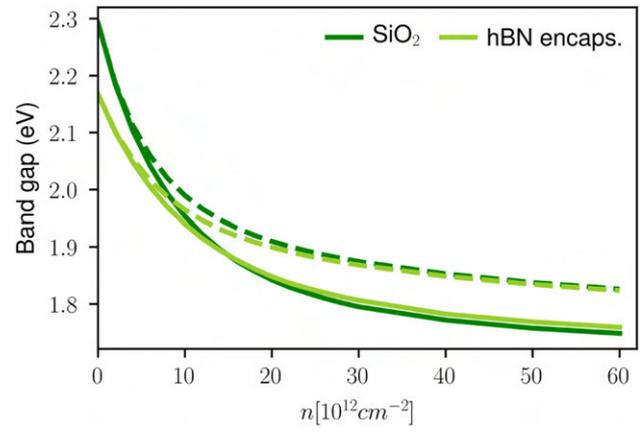


Figure 10. Dependence of the band gap renormalization on the carrier density for dynamical screening with $\gamma_T = 100$ meV for different MoS₂ in different dielectric environment, namely SiO₂ supported or hBN encapsulated.

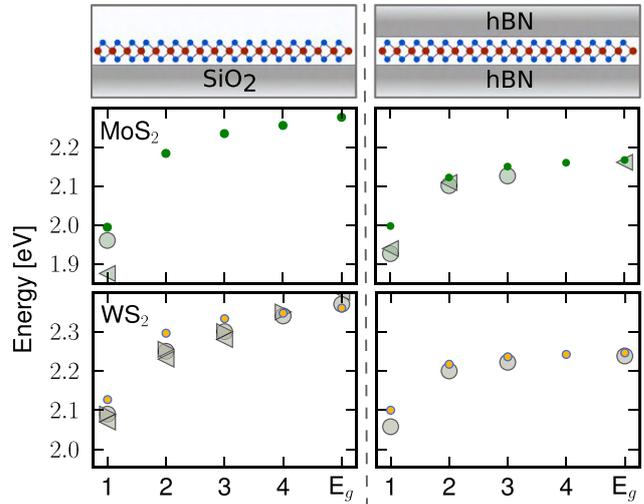


Figure 11. Resonance energies of the A exciton series and quasiparticle band gap for MoS₂ and WS₂ at a SiO₂ substrate (left) and hBN encapsulated (right) as illustrated in the above sketch. Colored symbols mark our results. Experimental values were taken from 28, 29, 34, 39–44 and are shown for comparison (grey symbols). As for WS₂ the band gap is underestimated in DFT calculations, we shifted the DFT gap for WS₂ by 149 meV while keeping the dipole matrix elements constant. For both materials, the exciton resonances are in acceptable agreement with experimentally derived values.

6. Summary and conclusions

In summary, we investigated the effectively quasi-2D nature of the Coulomb interaction potential in TMDC monolayers. We compute the matrix elements at the K -points of the Brillouin zone and introduce a form factor that effectively captures the deviations from the corresponding ideally 2D case. We apply this concept to efficiently compute fundamental properties such as excitonic resonances and density-dependent band gap renormalizations for a range of TMDC monolayers in different dielectric environments and under various excitation conditions.

Table 2. MDF model parameters for different TMDCs based on DFT calculations using the PBE functional, namely the spin- and valley dependent band gap $\Delta_{s\tau}$, the spin splitting $2\lambda_v$, the dielectric constants for bulk materials $\epsilon_{\parallel}^B, \epsilon_{\perp}$ and the bulk interlayer distance d resulting directly from the calculations. Fermi velocity $\hbar v_F$ and in-plane dielectric constants ϵ_{\parallel} are derived from the band structure and the bulk value, respectively.

Material	$s \cdot \tau$	$\Delta_{s\tau}$ (eV)	$\hbar v_F$ (eV Å)	λ_v (eV)	ϵ_{\parallel}^B	ϵ_{\parallel}	ϵ_{\perp}	d (Å)
MoS ₂	+	1.682	3.532	0.146	15.19	11.65	6.38	6.18
	−	1.831	3.467					
MoSe ₂	+	1.41	3.027	0.183	16.72	12.81	7.81	6.52
	−	1.614	3.001					
MoTe ₂	+	1.017	2.526	0.214	20.30	15.43	10.9	6.99
	−	1.266	2.574					
WS ₂	+	1.626	4.433	0.428	13.98	10.56	5.93	6.22
	−	2.022	4.208					
WSe ₂	+	1.385	3.941	0.46	15.58	11.81	7.65	6.51
	−	1.801	3.757					

Comparing our results both to current experimental data and theoretical approaches based on GW/GdW–BSE calculations, we obtain good quantitative agreement as summarized in table 1. In particular, our computed dependence of the band gap dependence on the dielectric surrounding is qualitatively in good agreement with the GdW based study of Riis-Jensen *et al* [47]. Also the comparison with experimental data yields excellent agreement for the quasiparticle band gap of all Mo based materials.

For the W based systems, the experimentally obtained values vary in a wide range with differences up to 600 meV. Using the DFT results as an input, our computed band gaps are about 100–200 meV below most experimental values. Interestingly, it can be observed that results based on the PBE functional tend to be below results using an LDA functional. Thus, the choice of the XC-functional has an impact on the absolute band gap value and absolute differences may result from an underestimation of the band gap in the DFT calculations. Fitting the underestimated DFT gap for subsequent calculations for WS₂-systems but maintaining the gap unchanged for MoS₂-systems, we find an acceptable agreement of the A exciton series with experimental data, see figure 11. Furthermore, our approach reproduces the experimentally observed influence of the dielectric environment.

To illustrate the usefulness of our approach for the excitation dependent system properties, we compute the modifications induced by charge carriers on the dynamic band gap renormalization. In experimental studies, this effect has been analyzed by employing electrical gating, examining the influence of doping to the band gap [39, 48], or by pump-probe experiments, using laser pulses to modify the excited electron–hole density in the material [44, 49]. In agreement with experimental observations, our analysis shows that smaller band-gap modifications are obtained in the configuration where additional electrons are introduced by doping relative to the case where symmetrical electron–hole populations are generated by optical excitation.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft via the Collaborative Research Center 1083 (DFG:SFB1083) and the GRK 1782 ‘Functionalization of Semiconductors’. Computing resources from the HRZ Marburg are acknowledged.

Appendix. Material parameters

In table 2 we summarize the DFT based parameters used to model the dielectric and optical material properties near the respective K -points. While the bulk dielectric constants ϵ_{\perp} , ϵ_{\parallel}^B and interlayer distance are direct results of DFT calculations, the band gap $\Delta_{s\tau}$, the valence band splitting $2\lambda_v$, and the Fermi velocity $\hbar v_F$ are determined by fitting the DFT band structure. The background dielectric constant ϵ_{\parallel} is derived in dependence of the bulk value as $\epsilon_{\parallel}^B = \epsilon_{\parallel} + \lim_{q_{\parallel} \rightarrow 0} 4\pi e^2 \chi_L(\mathbf{q}_{\parallel}, \omega)/D$ [6] where $\chi_L(\mathbf{q}_{\parallel}, \omega)$ is the contribution of upper spin-split valence and lowest conduction bands to the total longitudinal susceptibility.

ORCID iDs

T Stroucken  <https://orcid.org/0000-0002-0546-7289>

References

- [1] Mak K F, Lee C, Hone J, Shan J and Heinz T F 2010 *Phys. Rev. Lett.* **105** 136805
- [2] Splendiani A, Sun L, Zhang Y, Li T, Kim J, Chim C Y, Galli G and Wang F 2010 *Nano Lett.* **10** 1271–5
- [3] Chernikov A, Berkelbach T C, Hill H M, Rigosi A, Li Y, Aslan O B, Reichman D R, Hybertsen M S and Heinz T F 2014 *Phys. Rev. Lett.* **113** 076802
- [4] Ross J S *et al* 2013 *Nat. Commun.* **4** 1–6
- [5] He K, Kumar N, Zhao L, Wang Z, Mak K F, Zhao H and Shan J 2014 *Phys. Rev. Lett.* **113** 026803
- [6] Meckbach L, Hader J, Huttner U, Neuhaus J, Steiner J T, Stroucken T, Moloney J V and Koch S W 2020 *Phys. Rev. B* **101** 075401

- [7] Latini S, Olsen T and Thygesen K S 2015 *Phys. Rev. B* **92** 245123
- [8] Meckbach L, Stroucken T and Koch S W 2018 *Phys. Rev. B* **97** 035425
- [9] Winther K T and Thygesen K S 2017 *2D Mater.* **4** 025059
- [10] Ryou J, Kim Y S, Santosh K and Cho K 2016 *Sci. Rep.* **6** 1–8
- [11] Kresse G and Hafner J 1994 *Phys. Rev. B* **49** 14251–69
- [12] Kresse G and Hafner J 1993 *Phys. Rev. B* **47** 558–61
- [13] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 11169–86
- [14] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15–50
- [15] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758–75
- [16] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865–8
- [17] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
- [18] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188–92
- [19] Kormányos A, Burkard G, Gmitra M, Fabian J, Zólyomi V, Drummond N D and Fal'ko V 2015 *2D Mater.* **2** 022001
- [20] Stroucken T, Grönqvist J H and Koch S W 2011 *Phys. Rev. B* **84** 205445
- [21] Barford W, Boczarow I and Wharram T 2011 *J. Phys. Chem. A* **115** 9111–9
- [22] Ignatov S K, Razuvaev A G, Kokorev V N and Aleksandrov Y A 1995 *J. Struct. Chem.* **36** 538–43
- [23] Drüppel M, Deilmann T, Noky J, Marauhn P, Krüger P and Rohlfing M 2018 *Phys. Rev. B* **98** 155433
- [24] Soklaski R, Liang Y and Yang L 2014 *Appl. Phys. Lett.* **104** 193110
- [25] Gusakova J, Wang X, Shiao L L, Krivosheeva A, Shaposhnikov V, Borisenko V, Gusakov V and Tay B K 2017 *Phys. Status Solidi A* **214** 1700218
- [26] Haastrup S *et al* 2018 *2D Mater.* **5** 042002
- [27] Gerber I C *et al* 2019 *Phys. Rev. B* **99** 035443
- [28] Thygesen K S 2017 *2D Mater.* **4** 022004
- [29] Rigosi A F, Hill H M, Rim K T, Flynn G W and Heinz T F 2016 *Phys. Rev. B* **94** 075440
- [30] Goryca M *et al* 2019 *Nat. Commun.* **10** 1–12
- [31] Robert C *et al* 2018 *Phys. Rev. Mater.* **2** 011001
- [32] Bruix A *et al* 2016 *Phys. Rev. B* **93** 165422
- [33] Park S *et al* 2018 *2D Mater.* **5** 025003
- [34] Ugeda M M *et al* 2014 *Nat. Mater.* **13** 1091–5
- [35] Echeverry J P, Urbaszek B, Amand T, Marie X and Gerber I C 2016 *Phys. Rev. B* **93** 121107
- [36] Robert C *et al* 2016 *Phys. Rev. B* **94** 155425
- [37] Zhu B, Chen X and Cui X 2015 *Sci. Rep.* **5** 9218
- [38] Hanbicki A, Currie M, Kioseoglou G, Friedman A and Jonker B 2015 *Solid State Commun.* **203** 16–20
- [39] Chernikov A, van der Zande A M, Hill H M, Rigosi A F, Velauthapillai A, Hone J and Heinz T F 2015 *Phys. Rev. Lett.* **115** 126802
- [40] Gerber I C and Marie X 2018 *Phys. Rev. B* **98** 245126
- [41] Wang G, Marie X, Gerber I, Amand T, Lagarde D, Bouet L, Vidal M, Balocchi A and Urbaszek B 2015 *Phys. Rev. Lett.* **114** 097403
- [42] Liu E, van Baren J, Taniguchi T, Watanabe K, Chang Y C and Lui C H 2019 *Phys. Rev. B* **99** 205420
- [43] Stier A V, Wilson N P, Velizhanin K A, Kono J, Xu X and Crooker S A 2018 *Phys. Rev. Lett.* **120** 057405
- [44] Liu F, Ziffer M E, Hansen K R, Wang J and Zhu X 2019 *Phys. Rev. Lett.* **122** 246803
- [45] Cadiz F *et al* 2016 *2D Mater.* **3** 045008
- [46] Molas M R *et al* 2019 *Phys. Rev. Lett.* **123** 096803
- [47] Riis-Jensen A C, Lu J and Thygesen K S 2020 *Phys. Rev. B* **101** 121110
- [48] Yao K *et al* 2017 *Phys. Rev. Lett.* **119** 087401
- [49] Ulstrup S *et al* 2016 *ACS Nano* **10** 6315–22

Paper II

S. C. Liebscher, L. C. Bannow, J. Hader, J. V. Moloney, and S. W. Koch

“Extension of the LDA-1/2 method to the material class of bismuth containing III-V
semiconductors”

AIP Advances **10**, 115003 (2020) DOI: 10.1063/5.0024843

Reproduced from [II], with the permission of AIP Publishing.

Extension of the LDA-1/2 method to the material class of bismuth containing III–V semiconductors

Cite as: AIP Advances 10, 115003 (2020); doi: 10.1063/5.0024843

Submitted: 11 August 2020 • Accepted: 7 October 2020 •

Published Online: 2 November 2020



Sven C. Liebscher,^{1,a)}  Lars C. Bannow,¹  Jörg Hader,²  Jerome V. Moloney,²  and Stephan W. Koch^{1,3} 

AFFILIATIONS

¹Fachbereich Physik, Philipps-Universität Marburg, Renthof 5, 35032 Marburg, Germany

²Wyant College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA

³Material Sciences Center, Philipps-Universität Marburg, Hans-Meerwein-Straße 6, 35032 Marburg, Germany

^{a)} Author to whom correspondence should be addressed: sven.liebscher@physik.uni-marburg.de

ABSTRACT

The local density approximation-1/2 method is employed in density functional theory calculations for the electronic structure of III–V dilute bismide systems. For the representative example of Ga(SbBi) with Bi concentrations below 10%, it is shown that this method works very efficiently, especially due to its reasonably low demand on computer memory. The resulting band structure and wavefunctions are used to compute the interaction matrix elements that serve as input to the microscopic calculations of the optical properties and intrinsic losses relevant for the optoelectronic applications of dilute bismides.

© 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0024843>

I. INTRODUCTION

An ongoing interest in III–V materials that contain a dilute amount of bismuth (Bi) exists in the semiconductor research community.¹ These so-called bismides are promising candidates for opto-electronic applications such as light-emitting diodes,² semiconductor lasers,³ or solar cells.⁴ As can be seen from the material combinations listed in Ref. 1, bismuth can nowadays be incorporated into almost every III–V material. Since this allows for many different material combinations, a predictive approach for their electronic and optical properties is desirable.

For the example of In(AsBi), a scheme that combines density functional theory (DFT) and a microscopic many-body methodology has been applied and validated by detailed theory-to-experiment comparisons.^{5,6} Unfortunately, however, the DFT modeling of dilute III–V material systems requires the use of very large supercells, which are needed to approximate a random distribution of the relatively low concentration of the incorporated atoms. Previously, it has been shown that reliable bandgap calculations for dilute nitrides based on DFT require supercells with up to 432 atoms.⁷ Since commonly used functionals such as the local density approximation (LDA)⁸ or the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation⁹ severely underestimate the bandgap

of semiconductors, the meta-GGA TB09¹⁰ had to be applied in Refs. 5–7. In addition to the dependence on the electron density and its gradient, the TB09 functional also depends on the kinetic energy density. Therefore, DFT calculations utilizing this functional are in need of significantly more computational memory when compared to the LDA, which only depends on the electron density. This can become a limiting factor when not only the supercells are large but also many k -points have to be included in the calculation. For instance, this is the case when DFT is used to calculate the band structure and k -dependent dipole matrix elements as input for optical calculations.

In order to overcome the existing limitations, in the current study, the LDA-1/2 method^{11,12} is utilized, which requires only as much memory as the LDA. The LDA-1/2 method is based on Slater's half-occupation technique¹³ and improves the Kohn–Sham bandgap toward the real bandgap of the system. Recently, the strength of the LDA-1/2 method when it comes to large supercells was emphasized.¹⁴ The LDA-1/2 method depends on the variation of a cutoff radius until the bandgap of the system is maximized.¹¹ However, since III–Bi materials are semi-metals,¹⁵ this procedure cannot be used to determine the cutoff radii for the Bi atoms in AlBi, GaBi, and InBi. To overcome this problem, we use an alternative extrapolation method to obtain the cutoff radii for the Bi atoms, which allows us

to extend the LDA-1/2 method for application in dilute bismides. As a representative example, we calculate the electronic structure for Ga(SbBi) assuming different concentrations of bismuth below 10% and use the electronic structure results as input for the evaluation of the intrinsic optoelectronic properties.

II. COMPUTATIONAL METHOD

A. The LDA-1/2 method

Details of the LDA-1/2 method, especially its theoretical formulation, are presented in Refs. 11 and 12. Therefore, we restrict the discussion in this paper to the aspects important for the application of the LDA-1/2 method to dilute bismides. Ferreira *et al.*¹¹ introduced what they call a self-energy potential $V_S(r)$ whose quantum-mechanical average is the self-energy,

$$S_\alpha = \int d^3r n_\alpha(r) V_S(r), \quad (1)$$

with state α and electron density n_α . This self-energy potential $V_S(\mathbf{r})$ is obtained for the specific atom by subtracting the Kohn–Sham effective potential of the half-ionized state from that of the neutral state. Subsequently, for calculations of solids, $V_S(\mathbf{r})$ is added to the LDA pseudopotentials. However, as $V_S(\mathbf{r}) \rightarrow 1/r$ for long ranges, this would result in a divergent result. Therefore, $V_S(\mathbf{r})$ is first multiplied by a spherical step function given as

$$\Theta(r) = \begin{cases} \left[1 - \left(\frac{r}{r_{\text{cut}}} \right)^n \right]^3, & r \leq r_{\text{cut}} \\ 0, & r > r_{\text{cut}}, \end{cases} \quad (2)$$

where the usual choice is $n = 8$.¹¹ As shown by Xue *et al.*,¹⁶ for $n = 8$, the tails of the step function are long such that larger values of n are better suited for adding $V_S(\mathbf{r})$ only in the region where the valence hole is located. Following Ref. 16, we, therefore, set $n = 20$ in our calculations. The precise cutoff radius r_{cut} is determined by its variation until the bandgap of the solid in question is maximized. With this cutoff radius, the spherical step function $\Theta(r)$ covers as much as possible of the region that is occupied by the valence band hole.¹⁶ Furthermore, following Refs. 11 and 12, here, the LDA-1/2 method is only applied to the anions as the valence band states usually originate from these.

B. Supercell generation

DFT calculations were performed with the Vienna *ab initio* Simulation Package (VASP 5.4.4),^{17–20} which uses a plane-wave basis set and the projector-augmented wave method.^{21,22} In our evaluations, a 370 eV cutoff energy of the plane-wave basis was used and the criteria for electronic convergence and ionic relaxation were set to 10^{-7} eV and 10^{-6} eVÅ⁻¹, respectively. For relaxations, the LDA²³ exchange-correlation potential was used and a Γ -centered Monkhorst–Pack-grid²⁴ of $8 \times 8 \times 8$ k -points was implemented for primitive cell calculations. After the relaxation of the primitive cells of GaSb and GaBi, supercells of different geometries were created using the Alloy Theoretic Automated Toolkit (ATAT).²⁵ The supercells consist of 64 primitive cells of GaSb, arranged in different cuboidal geometries with up to eight primitive cells in one

direction. Since the x -direction and y -direction are equivalent, geometries with permutations of x and y can be omitted, leaving the geometries $1 \times 8 \times 8$, $2 \times 4 \times 8$, $2 \times 8 \times 4$, $4 \times 4 \times 4$, $4 \times 8 \times 2$, and $8 \times 8 \times 1$. The k -point grid was reduced accordingly for the supercells. Special quasirandom structures (SQSs)²⁶ were generated to statistically determine an optimal approximation to the random placement of the Bi atoms in the supercell. Up to four Bi atoms were incorporated into the supercells, yielding the concentrations 0%, 1.5625%, 3.125%, 4.6875%, and 6.25%.

The relaxations of the supercells were constrained to the z -direction in order to simulate the effect of a substrate with fixed in-plane lattice constants without explicitly including the substrate into the supercells. In order to speed-up the convergence, an initial value for the relaxed lattice constant in the z -direction was set according to the formula $\frac{(x \cdot a_{\text{GaBi}} + (1-x) \cdot a_{\text{GaSb}})^3}{a_{\text{GaSb}}^2}$, assuming that the relaxation in the z -direction will enhance the volume by the same factor as an interpolation between the lattice constants of the pure compounds in all directions would. Using the LDA-1/2 potentials obtained for Sb and Bi, energy calculations were performed on the supercells. First, a self-consistent calculation with the regular grid of k -points described above was performed to obtain the charge densities and bandgap energies. Second, non-self-consistent calculations using the charge densities were evaluated with finer k -point grids along a high-symmetry line in the first Brillouin zone from the Γ -to L-point to obtain the band structure and Coulomb and dipole matrix elements. Spin-orbit coupling was included for all energy calculations.

III. RESULTS AND DISCUSSION

In Table I, we summarize the lattice constants for the different III–V binaries as obtained from DFT calculations within the LDA. Comparing our values with those found in the literature,²⁷ we obtain good agreement with the exception of GaP. Moreover, our computed lattice constants for the III–Bi binaries compare well to the values found in the literature.^{28–30} It should be noted that bulk InBi is reported to crystallize in the PbO structure;³⁰ however, here, we are interested in incorporating Bi into zinc-blende lattices, and therefore, the lattice constant from the zinc-blende InBi structure is given in Table I.

TABLE I. Lattice constants a in Å calculated within the LDA for the zinc-blende structures of AlN, AlP, AlAs, AlSb, AlBi, GaN, GaP, GaAs, GaSb, GaBi, InN, InP, InAs, InSb, and InBi. In addition to the LDA lattice constants, the deviations Δ_{exp} to the experimental lattice constants are listed.

X	N	P	As	Sb	Bi
a (AlX)	4.3435	5.4337	5.6351	6.1155	6.2741
$\Delta_{\text{exp}}^{\text{AlX}}$ (%)	−0.8	−0.6	−0.5	−0.3	...
a (GaX)	4.4570	5.3208	5.6000	6.0529	6.2538
$\Delta_{\text{exp}}^{\text{GaX}}$ (%)	−1.0	−2.4	−0.9	−0.7	...
a (InX)	4.9412	5.8097	6.0071	6.4304	6.6081
$\Delta_{\text{exp}}^{\text{InX}}$ (%)	−0.8	−1.0	−0.9	−0.8	...

A. Determination of the cutoff radii

The cutoff radii for the anions of the zinc-blende structures of AlX, GaX, and InX for $X \in \{N, P, As, Sb\}$ were found by adapting the procedure described in Ref. 11. Thereby, several pseudopotentials with different r_{cut} for the anions were generated for each material system. The cutoff radii r_{cut} were varied in steps of $0.1a_0$, where a_0 denotes the Bohr radius, and the direct bandgap at the Γ -point was extracted from the calculations. A parabolic function $E_g(r_{\text{cut}}) = m(r_{\text{cut}} - r_{\text{cut}}^{\text{max}})^2 + E_g^{\text{max}}$ was fit to the results of the calculations with the adjustable parameters m , E_g^{max} , and $r_{\text{cut}}^{\text{max}}$. The maximum of this function is located at $r_{\text{cut}} = r_{\text{cut}}^{\text{max}}$, which is the optimized cutoff radius for the respective material.

In Table II, the cutoff radii yielding the maximum bandgaps are shown together with the covalent radii for the anions. In comparison with Ref. 11, here, the cutoff radii are smaller, which is due to the use of $n = 20$ for the spherical step function as compared to $n = 8$ in Ref. 11.

For each of the group III atoms, Al, Ga, and In, the cutoff radius for the Bi atom is obtained using a linear extrapolation. The function $r_{\text{cut}}/a_0(r_{\text{cov}}/a_0)$ for the respective ion is found by a linear fit to the cutoff radii for N, P, As, and Sb. Thereafter, the function is used to calculate the cutoff radius for Bi. The cutoff radii for the III-Bi compounds are also listed in Table II.

Figure 1 compares the experimental values²⁷ and those from the LDA-1/2 calculations for the direct bandgaps at the Γ -point for all investigated binary III-V semiconductors. Almost perfect agreement is found for GaSb, GaAs, AlSb, and AlAs, while all other bandgaps with the exception of InN are overestimated by the LDA-1/2 method. One reason for this overestimation is that the LDA was used to calculate the lattice constants and the LDA is known to yield smaller values than those reported in the experiment. Nonetheless, overall, a strong improvement as compared to the LDA and PBE bandgaps is evident.

Applying the LDA-1/2 method to dilute bismides, the bandgap for $\text{GaAs}_{1-x}\text{Bi}_x$ was calculated for five 128 atom supercells with $x = 1.5625\% \dots 7.8125\%$. The results are compared to bandgap values obtained with the TB09 functional in Ref. 31. It is found that the absolute values between both sets of calculations have a maximal difference of $\Delta E_g = 0.03$ eV, which is well within the accuracy of a specific configuration of Bi. Moreover, the LDA-1/2 bandgap decreases by 75 ± 6 meV/1% Bi, which compares extremely well to the 72 ± 4 meV/1% Bi found in Ref. 31. While the supercell size in our work is limited by the computational demand of DFT calculations, the effects of the supercell size in GaAsBi have been

TABLE II. Cutoff radii r_{cut} as obtained for the zinc-blende structures of AlN, AlP, AlAs, AlSb, GaN, GaP, GaAs, GaSb, InN, InP, InAs, and InSb together with the covalent radii for the five anions. The cutoff radii for the III-Bi materials were obtained using a linear extrapolation, as described in the text. Here, a_0 denotes the Bohr radius.

	N	P	As	Sb	Bi
r_{cov}/a_0	1.342	2.022	2.249	2.627	2.797
$r_{\text{cut}}(\text{Al})/a_0$	2.573	3.329	3.403	3.749	3.934
$r_{\text{cut}}(\text{Ga})/a_0$	2.565	3.195	3.326	3.661	3.812
$r_{\text{cut}}(\text{In})/a_0$	2.737	3.351	3.454	3.865	3.986

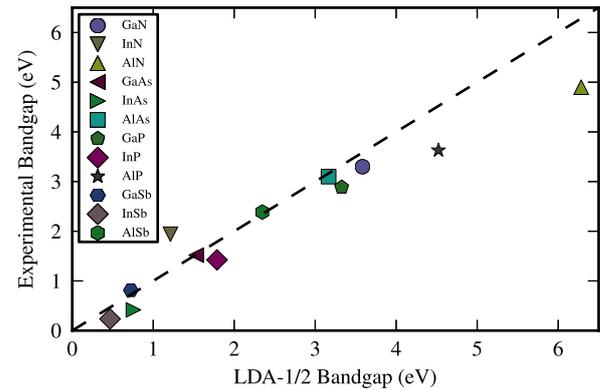


FIG. 1. Experimental bandgaps for $T = 0$ K taken from Ref. 27 compared to the bandgaps as-obtained from the LDA-1/2 method for several III-V semiconductors. Each material is represented by a data point with the experimental bandgap energy as the y-value and the calculated bandgap energy as the x-value so that data points close to the straight line correspond to an almost perfect agreement for the bandgaps. Shown are the direct bandgaps at the Γ -point, even though some materials, i.e., AlN, AlP, GaP, AlAs, and AlSb, are predicted to be indirect semiconductors.

investigated by Usman *et al.* using a tight-binding Hamiltonian, showing that supercells containing more than 1000 atoms are the optimum choice.^{32,33}

B. Bandgap narrowing in Ga(SbBi)

Using the obtained LDA-1/2 potentials for GaSb and GaBi, supercell calculations on $\text{GaSb}_{1-x}\text{Bi}_x$ were performed, as described in Sec. II B. A comparison of the bandgap energies with increasing Bi content for the different supercell geometries is shown in Fig. 2. An almost linear decrease in the bandgap can be seen for all supercell geometries. For all concentrations, the bandgap energy of the isotropic $4 \times 4 \times 4$ supercell is the highest. The supercells with two primitive cells in one direction, $2 \times 4 \times 8$ and its permutations,

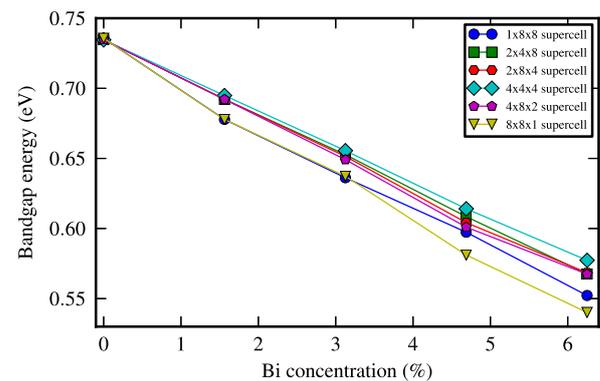


FIG. 2. Theoretical bandgap energies from DFT calculations of $\text{GaSb}_{1-x}\text{Bi}_x$ supercells against Bi concentration x using the LDA-1/2 potentials. The supercells consist of 64 primitive cells, which are arranged in different cuboidal geometries. The lines are guides to the eye.

behave similarly and have slightly lower bandgap energies than the isotropic supercell.

The strongest, non-linear decrease in bandgap energy is found for the supercells that have only one primitive cell in one of the spatial directions. Along the direction in which only one primitive cell is used, the periodic boundary conditions lead to an infinite line of neighboring Bi atoms. Their interaction causes a stronger bandgap reduction.^{33,35} While this clustering can occur in real materials, in this work, we want to model materials with more evenly distributed Bi atoms using the SQSs described in Sec. II B. Therefore, we conclude that the isotropic supercells with $4 \times 4 \times 4$ primitive cells yield the best results by minimizing the effects of Bi clustering due to the limited supercell size.

The bandgap energies and split-off energies of these supercells are shown in Fig. 3. A linear fit yields a bandgap reduction of (25.3 ± 0.2) meV/%Bi and an increase in the spin-orbit splitting energy of (17.9 ± 0.2) meV/%Bi. The spin-orbit splitting is in good agreement with 18 meV/%Bi found in a theoretical study,³⁶ and the bandgap narrowing is comparable to 30 meV/%Bi observed in an experimental study.³⁴ For the following analysis of the band structure and optical properties, only the isotropic $4 \times 4 \times 4$ supercells are considered.

C. Effective band structure of Ga(SbBi)

The repetition of primitive cells in real space leads to a backfolding of states in k -space toward the zone center. Therefore, in order to compare the band structure of a supercell with that of the primitive cell, it is convenient to perform an unfolding of the bands to obtain the effective band structure.³⁷ The spectral weight of the unfolded states represents how much of the Bloch character of the state is preserved in the supercell. The unfolded band structures of the isotropic supercells for all Bi concentrations (increasing from left to right) are shown in Fig. 4, where the opacity of the scatter points was set according to the spectral weight of the respective states. Therefore, states with lower spectral weight are less visible

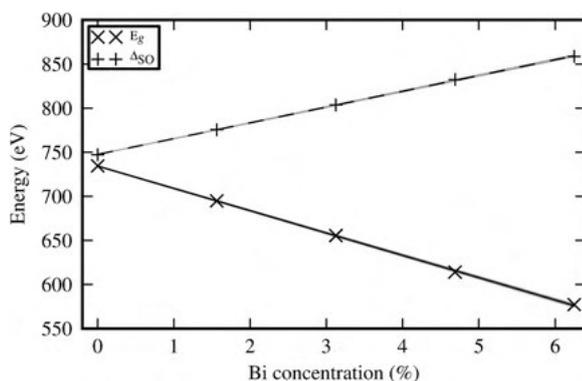


FIG. 3. Bandgap and spin-orbit splitting energy (Δ_{SO}) of the $\text{GaSb}_{1-x}\text{Bi}_x$ supercells consisting of $4 \times 4 \times 4$ primitive cells over Bi concentration x obtained from DFT calculations using the LDA-1/2 method. The lines represent linear fits with the shaded areas indicating the standard deviations. A constant shift was applied to the experimental values taken from Ref. 34 to match the theoretical value at $x = 0\%$.

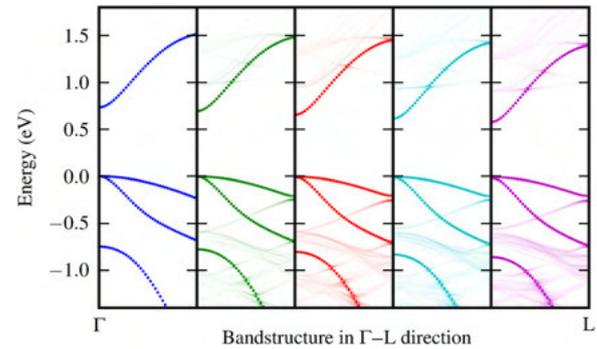


FIG. 4. Band structures from the Γ - to L -point of $\text{GaSb}_{1-x}\text{Bi}_x$ supercells for different Bi concentrations, from left to right $x = 0\%$, 1.56%, 3.13%, 4.69%, and 6.25%. The electronic structure has been calculated using DFT with the LDA-1/2 method, and the supercell band structure was unfolded to obtain the effective primitive cell band structure. The opacity of the data points indicates their spectral weight.

than those with higher spectral weight. The energies are given relative to the valence band edge so that the valence band maximum is always at 0 eV.

In the pure band structure, the light and heavy hole band, the split-off band, and the conduction band can be seen. These bands remain the ones with the highest spectral weight for all concentrations and can be clearly identified. However, we note a shift of the conduction and the split-off band to lower energies with increasing Bi concentration. As predicted by the valence band anticrossing model,³⁸ the valence band is actually shifting upwards, but due to the normalization of the valence band edge to 0 eV, it is seen as a downward shift of the conduction band in this plot. At the same time, more and more additional bands with low spectral weight emerge. In a region between the valence bands and the split-off band, a lot of defect bands are very close to one another. Additionally, the spectral weight of the defect bands increases for higher Bi concentrations.

At two points, the conduction band is crossed by defect bands with low spectral weight. Near the crossing points, the spectral weight of the defect bands increases. The spectral weight of some defect bands also increases toward the zone edges. A slight splitting of the heavy- and light-hole band can be seen for higher concentrations. This splitting can be quantified approximately using an exponential fit, yielding

$$E_{hh-lh}(x) = (3.5 \pm 0.5) \text{ meV} \cdot \exp((27 \pm 3)x), \quad (3)$$

for $\text{GaSb}_{1-x}\text{Bi}_x$. However, as shown in Ref. 39, the splitting is caused by substitutional disorder reducing the lattice symmetry. Therefore, the amount of splitting depends on the specific choice of the supercell. While the SQS supercells aim to correctly model the randomized distribution of Bi atoms in a real material, their optimization is limited by the supercell size and the splitting may vary for other configurations.

D. Optical properties of Ga(SbBi)

The DFT wavefunctions are used to compute the Coulomb and dipole matrix elements needed for our microscopic semiconductor

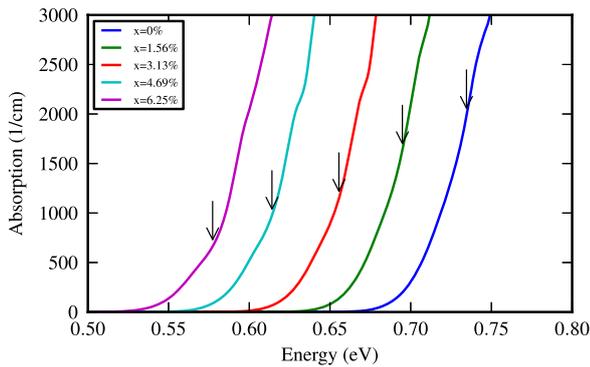


FIG. 5. Onset of interband absorption of $\text{GaSb}_{1-x}\text{Bi}_x$ supercells for different Bi concentrations as a function of photon energy at 300 K with carrier density $N_{3D} = 5 \cdot 10^{16} \text{ cm}^{-3}$. Dipole and matrix elements from DFT calculations using the LDA-1/2 method were used as input for a microscopic theory. The arrows mark the bandgap energy for the respective concentration.

Bloch equation approach to calculate the intrinsic optical properties.⁴⁰ The details of the calculation can be found in Refs. 41 and 42. In Fig. 5, the resulting onset of interband absorption for the different Bi concentrations is shown. For the calculations, 20 conduction bands and 30 valence bands were included. The temperature was set to 300 K and the carrier density was set to $N_{3D} = 5 \cdot 10^{16} \text{ cm}^{-3}$. It should be noted that the DFT band structure was used, although it was calculated at 0 K temperature. As expected, an increase in the absorption coefficient can be seen at the bandgap energy. For higher Bi concentrations, this increase shifts to lower energies, in agreement with the decreasing bandgap.

In order to calculate the emission properties, the microscopic approach is used to solve the semiconductor luminescence equations.^{43,44} Photoluminescence spectra at 300 K for the supercells have been calculated and are shown in Fig. 6. Again, 20 conduction bands and 30 valence bands were included in the calculations. An intensity peak for each concentration occurs at the respective bandgap

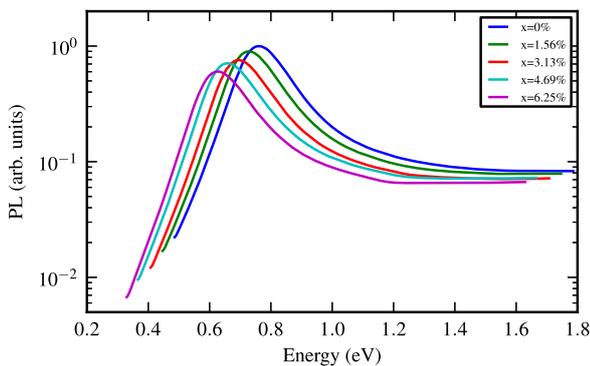


FIG. 6. Photoluminescence spectra of $\text{GaSb}_{1-x}\text{Bi}_x$ supercells for different Bi concentrations as a function of photon energy at 300 K with carrier density $N_{3D} = 5 \cdot 10^{16} \text{ cm}^{-3}$. Dipole and matrix elements from DFT calculations using the LDA-1/2 method were used as input for a microscopic theory.

energy. Therefore, a shift to lower energies can be seen for increasing Bi concentration.

E. Auger losses of $\text{Ga}(\text{SbBi})$

The losses due to Auger recombination at 300 K have been calculated within the microscopic approach, as described in Ref. 45 by solving Eqs. (9) and (10) therein and using Eq. (11) to obtain the Auger loss coefficients. The energy region from 1.1 eV below the valence band maximum up to 1.7 eV above the valence band maximum was used for the calculations. The results for different carrier densities are shown in Fig. 7.

For low densities up to $N_{3D} = 10^{17} \text{ cm}^{-3}$, we see that the Auger coefficient is virtually constant for all Bi concentrations considered. The coefficient is highest for pure GaSb with $4.3 \cdot 10^{-27} \text{ cm}^6/\text{s}$, which compares well to the experimental value of $(12 \pm 4) \cdot 10^{-27} \text{ cm}^6/\text{s}$ found in Ref. 46 and decreases for higher Bi content. Since for pure GaSb the bandgap and spin-orbit splitting energy are very close to each other (see Fig. 3), the losses are high. With increasing Bi content, the bandgap decreases, while the spin-orbit splitting increases so that the loss channel is weaker and the Auger coefficient decreases. The dependence of the Auger coefficient on the difference between the bandgap and spin-orbit splitting energy in this low density region is shown in Fig. 7(b), exhibiting a near exponential decrease, which can be approximated by

$$C = (5.4 \pm 0.1) \cdot 10^{-27} \text{ cm}^6 \text{ s}^{-1} \exp\left((17 \pm 1) \frac{\Delta_{\text{SO}} - E_g}{\text{eV}}\right). \quad (4)$$

For higher densities, from $N_{3D} = 10^{18} \text{ cm}^{-3}$ to $N_{3D} = 10^{19} \text{ cm}^{-3}$, more states are filled with electrons and more transitions become possible so that the Auger coefficient rises. The rise only occurs for the GaSbBi supercells, not for pure GaSb, and is more prominent for higher Bi concentrations. However, at the same time, from $N_{3D} = 10^{18} \text{ cm}^{-3}$ on, the filling of the phase-space leads to a general decrease in the Auger coefficient for all Bi concentrations.⁴⁷ The rise in the loss coefficient for the GaSbBi

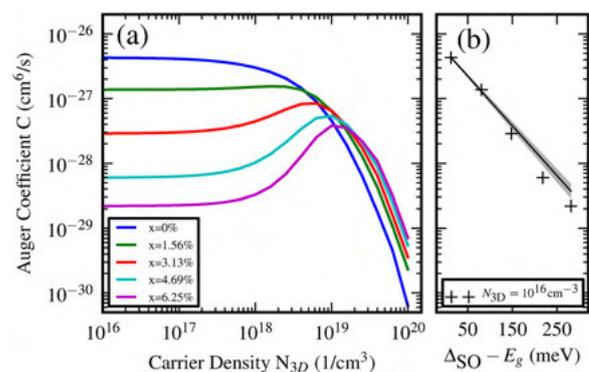


FIG. 7. (a) Auger loss coefficients of $\text{GaSb}_{1-x}\text{Bi}_x$ supercells for different Bi concentrations as a function of carrier density. Dipole and matrix elements from DFT calculations using the LDA-1/2 method were used as input for a microscopic theory to calculate the Auger losses. (b) Auger loss coefficients as a function of the difference between spin-orbit splitting and bandgap energy. The line represents an exponential fit with the shaded area indicating the standard deviation.

supercells is stronger for densities up to $N_{3D} = 10^{19} \text{ cm}^{-3}$ so that the Auger coefficient reaches a maximum. For densities higher than $N_{3D} = 10^{19} \text{ cm}^{-3}$, the phase-space filling effect dominates and the Auger coefficients for all supercells decrease exponentially. For a given density above $N_{3D} = 10^{19} \text{ cm}^{-3}$, pure GaSb has the lowest loss coefficient and higher Bi content leads to higher loss coefficients. It should be noted that the Auger coefficients have been calculated only using the Coulomb matrix elements of the Γ -L direction and results may differ in other directions.

IV. CONCLUSION

The LDA-1/2 method for DFT calculations requires less memory than other exchange-correlation functionals while yielding good results for most semiconductor compounds. To date, the LDA-1/2 method could not be used for dilute bismides because the construction of the LDA-1/2 potential involves maximizing the bandgap of the compound. In this work, a new method of finding the cutoff radius for the construction of LDA-1/2 potentials by extrapolation has been introduced. It was used to construct LDA-1/2 potentials for the semi-metal Bi in semiconductor compounds AlBi, GaBi, and InBi, allowing for DFT calculations on dilute bismides with significantly reduced computational cost. In consequence, larger supercells, which are needed for low Bi concentrations, and dense k -point grids can be calculated. The usefulness of the approach for dilute bismides was demonstrated by calculating the electronic and optical properties of Ga(SbBi) for different Bi concentrations by using the results from the DFT calculations with the LDA-1/2 method as input to a microscopic theory.

ACKNOWLEDGMENTS

The work at Marburg was supported by the DFG in the framework of the Research Training Group "Functionalization of Semiconductors" (Grant No. GRK 1782). The authors would like to thank the HRZ Marburg and CSC-Goethe-HLR Frankfurt for computational resources. The work at Tucson was supported by the Air Force Office of Scientific Research under Award No. FA9550-17-1-0246.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- 1 L. Wang, L. Zhang, L. Yue, D. Liang, X. Chen, Y. Li, P. Lu, J. Shao, and S. Wang, *Crystals* **7**, 63 (2017).
- 2 N. Hossain, I. P. Marko, S. R. Jin, K. Hild, S. J. Sweeney, R. B. Lewis, D. A. Beaton, and T. Tiedje, *Appl. Phys. Lett.* **100**, 051105 (2012).
- 3 I. P. Marko, C. A. Broderick, S. Jin, P. Ludewig, W. Stolz, K. Volz, J. M. Rorison, E. P. O'Reilly, and S. J. Sweeney, *Sci. Rep.* **6**, 28863 (2016).
- 4 R. D. Richards, A. Mellor, F. Harun, J. S. Cheong, N. P. Hylton, T. Wilson, T. Thomas, J. S. Roberts, N. J. Ekins-Daukes, and J. P. R. David, *Sol. Energy Mater. Sol. Cells* **172**, 238 (2017).
- 5 J. Hader, S. C. Badescu, L. C. Bannow, J. V. Moloney, S. R. Johnson, and S. W. Koch, *Appl. Phys. Lett.* **112**, 062103 (2018).
- 6 J. Hader, S. C. Badescu, L. C. Bannow, J. V. Moloney, S. R. Johnson, and S. W. Koch, *Appl. Phys. Lett.* **112**, 192106 (2018).
- 7 P. Rosenow, L. C. Bannow, E. W. Fischer, W. Stolz, K. Volz, S. W. Koch, and R. Tonner, *Phys. Rev. B* **97**, 075201 (2018); [arXiv:1705.10763](https://arxiv.org/abs/1705.10763).
- 8 D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- 9 J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- 10 F. Tran and P. Blaha, *Phys. Rev. Lett.* **102**, 226401 (2009).
- 11 L. G. Ferreira, M. Marques, and L. K. Teles, *Phys. Rev. B* **78**, 125116 (2008).
- 12 L. G. Ferreira, M. Marques, and L. K. Teles, *AIP Adv.* **1**, 032119 (2011).
- 13 J. C. Slater and K. H. Johnson, *Phys. Rev. B* **5**, 844 (1972).
- 14 J. Doumont, F. Tran, and P. Blaha, *Phys. Rev. B* **99**, 115101 (2019).
- 15 A. B. Shick, J. B. Ketterson, D. L. Novikov, and A. J. Freeman, *Phys. Rev. B* **60**, 15484 (1999).
- 16 K.-H. Xue, J.-H. Yuan, L. R. C. Fonseca, and X.-S. Miao, *Comput. Mater. Sci.* **153**, 493 (2018).
- 17 G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- 18 G. Kresse and J. Hafner, *Phys. Rev. B* **49**, 14251 (1994).
- 19 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- 20 G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- 21 P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- 22 G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- 23 J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- 24 H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- 25 A. van de Walle, P. Tiwary, M. de Jong, D. L. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, and Z.-K. Liu, *Calphad* **42**, 13 (2013).
- 26 A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, *Phys. Rev. Lett.* **65**, 353 (1990).
- 27 I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, *J. Appl. Phys.* **89**, 5815 (2001).
- 28 S. Q. Wang and H. Q. Ye, *Phys. Rev. B* **66**, 235111 (2002).
- 29 P. Carrier and S.-H. Wei, *Phys. Rev. B* **70**, 035212 (2004).
- 30 M. Ferhat and A. Zaoui, *Phys. Rev. B* **73**, 115107 (2006).
- 31 L. C. Bannow, S. C. Badescu, J. Hader, J. V. Moloney, and S. W. Koch, *Appl. Phys. Lett.* **111**, 182103 (2017); [arXiv:1709.09983](https://arxiv.org/abs/1709.09983).
- 32 Y. Zhang, A. Mascarenhas, and L.-W. Wang, *Phys. Rev. B* **71**, 155201 (2005).
- 33 M. Usman, C. A. Broderick, A. Lindsay, and E. P. O'Reilly, *Phys. Rev. B* **84**, 245202 (2011).
- 34 J. Kopaczek, R. Kudrawiec, W. M. Linhart, M. K. Rajpalke, K. M. Yu, T. S. Jones, M. J. Ashwin, J. Misiewicz, and T. D. Veal, *Appl. Phys. Lett.* **103**, 261907 (2013).
- 35 R. N. Kini, A. J. Ptak, B. Fluegel, R. France, R. C. Reedy, and A. Mascarenhas, *Phys. Rev. B* **83**, 075307 (2011).
- 36 M. P. Polak, P. Scharoch, and R. Kudrawiec, *Semicond. Sci. Technol.* **30**, 094001 (2015).
- 37 V. Popescu and A. Zunger, *Phys. Rev. B* **85**, 085201 (2012).
- 38 M. K. Rajpalke, W. M. Linhart, M. Birkett, K. M. Yu, D. O. Scanlon, J. Buckeridge, T. S. Jones, M. J. Ashwin, and T. D. Veal, *Appl. Phys. Lett.* **103**, 142106 (2013).
- 39 M. Usman, C. A. Broderick, Z. Batool, K. Hild, T. J. C. Hosea, S. J. Sweeney, and E. P. O'Reilly, *Phys. Rev. B* **87**, 115104 (2013).
- 40 H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 5th ed. (World Scientific, Singapore, 2009).
- 41 M. Lindberg and S. W. Koch, *Phys. Rev. B* **38**, 3342 (1988).
- 42 A. Girndt, F. Jahnke, A. Knorr, S. W. Koch, and W. W. Chow, *Phys. Status Solidi B* **202**, 725 (1997).
- 43 M. Kira and S. W. Koch, *Phys. Rev. A* **73**, 013813 (2006).
- 44 M. Kira and S. W. Koch, *Prog. Quantum Electron.* **30**, 155 (2006).
- 45 J. Hader, J. V. Moloney, and S. W. Koch, *IEEE J. Quantum Electron.* **41**, 1217 (2005).
- 46 S. Marchetti, M. Martinelli, and R. Simili, *J. Phys.: Condens. Matter* **14**, 3653 (2002).
- 47 J. Hader, J. V. Moloney, and S. W. Koch, *Appl. Phys. Lett.* **87**, 201112 (2005).

Paper III

S. C. Liebscher, M. K. Hagen, J. Hader, J. V. Moloney, and S. W. Koch

“Microscopic theory for the incoherent resonant and coherent off-resonant optical response of tellurium”

Phys. Rev. B **104**, 165201 (2021) **DOI:** 10.1103/PhysRevB.104.165201

Reproduced from [III], with the permission of APS Publishing.

Microscopic theory for the incoherent resonant and coherent off-resonant optical response of tellurium

Sven C. Liebscher ¹, Maria K. Hagen ¹, Jörg Hader ², Jerome V. Moloney ² and Stephan W. Koch ¹

¹Department of Physics and Material Sciences Center, Philipps-University Marburg, Renthof 5, 35032 Marburg, Germany

²Wyant College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA



(Received 18 August 2021; revised 30 September 2021; accepted 4 October 2021; published 11 October 2021)

An *ab initio*-based fully microscopic approach is applied to study the nonlinear optical response of bulk tellurium. The structural and electronic properties are calculated from first principles using the shLDA-1/2 method within density functional theory. The resulting band structure and dipole matrix elements serve as input for the quantum mechanical evaluation of the anisotropic linear optical absorption spectra, yielding results in excellent agreement with published experimental data. Assuming quasiequilibrium carrier distributions in the conduction and valence bands, absorption/gain and spontaneous emission spectra are computed from the semiconductor Bloch and luminescence equations. For ultrafast intense off-resonant excitation, the generation of high harmonics is evaluated and the emission spectra are calculated for samples of different thicknesses.

DOI: [10.1103/PhysRevB.104.165201](https://doi.org/10.1103/PhysRevB.104.165201)

I. INTRODUCTION

Elemental tellurium is known to have a wide variety of unusual optical characteristics. It is the only elemental semiconductor with a direct band gap in the technically interesting midinfrared wavelength range near $3.8 \mu\text{m}$. Furthermore, Te is considered to have exceptional nonlinear optical properties [1,2] due to its chiral structure where the atoms form helical chains. Relatively few studies of the optical properties of Te have been published so far. Reflectivity and absorption spectra and their temperature-dependent variations have been analyzed in Refs. [3–6] and references therein. These papers report a strong polarization dependence of the optical response, reflecting the uniaxial nature of the Te crystal. Bulk crystals of Te exhibit large refractive indices with a prominent difference between the ordinary and extraordinary directions (about 4.9 and 6.3 near the band gap) [7].

The strongly directional crystal structure also leads to prominent optical nonlinearities. An exceptionally large nonlinear coefficient was confirmed by a phase-matched harmonic generation measurement on an elemental Te crystal in Ref. [8]. Furthermore, the chiral structure has been shown to lead to gyroscopic nonlinear optical responses depending on the helicity of the light (see, e.g., Ref. [9] and references therein). Measurements of the photoluminescence (PL) from bulk Te crystals have been reported for cryogenic temperatures in Ref. [10] and for room temperature in Ref. [11]. These publications also document indications of stimulated emission and lasing as well as strong second- and third-order harmonic generation.

To complement and extend the earlier investigations, we present in this paper a comprehensive analysis of the nonlinear optical properties of bulk Te. For this purpose, we performed a systematic microscopic study of its resonant incoherent and off-resonant coherent properties. We employ an *ab initio*-

based approach where we use density functional theory (DFT) together with the shell local density approximation-1/2 (shLDA-1/2) method to obtain accurate structural and electronic parameters. We evaluate the dispersion of the energetically highest valence and the lowest conduction bands and determine the relevant dipole and Coulomb interaction matrix elements.

Using these results as input for the semiconductor Bloch equations (SBEs) [12], we first evaluate the Te absorption spectra for different excitation conditions. Our results show excellent agreement with published experimental data. Assuming quasiequilibrium carrier populations in the relevant valence and conduction bands, we compute the transition from absorption to optical gain. The corresponding luminescence spectra are evaluated using the semiconductor luminescence equations (SLEs) [13]. Both gain and luminescence exhibit strong dependence on the light polarization direction.

For strongly off-resonant excitation, we investigate the generation of high harmonics in a wide spectral range extending far above the fundamental Te band gap. Currently, high harmonic generation (HHG) in semiconductors after excitation with short high-intensity pulses is a field of active research [14–29]. Microscopically, semiconductor HHG can be related to the nonequilibrium dynamics of the induced electron-hole excitations, including interband polarizations and intraband currents probing the conduction and valence band structure in the entire Brillouin zone (BZ). To analyze these effects, we use our DFT results as structural input for the SBE and compute HHG spectra for different excitation conditions. Besides local evaluations of HHG spectra, we also study the effects of different sample thicknesses performing calculations which explicitly include field propagation effects.

This paper is organized as follows: In Sec. II, we give an overview of our DFT approach and discuss the resulting band structure and the relevant dipole matrix elements for bulk Te.

Section III summarizes our calculations for optical absorption, gain, and PL, whereas Sec. IV is devoted to the modeling of HHG in Te for different excitation conditions and sample lengths. A short summary and outlook in Sec. V concludes our presentation.

II. ELECTRONIC STRUCTURE CALCULATIONS

A. Computational details

In our approach to construct the electronic structure for bulk Te, we use the VIENNA AB INITIO SIMULATION PACKAGE [30–33] (VASP) version 5.4.4 which implements the projector-augmented wave (PAW) method [34,35]. Starting from the symmetry group of right-handed tellurium $P3_121 - D_3^4$, the crystal structure was relaxed using the generalized gradient approximation (GGA) by Perdew *et al.* [36] for the exchange-correlation energy. A Γ -centered Monkhorst-Pack [37] grid of $15 \times 15 \times 15$ k points and a plane-wave basis-set cutoff energy of 500 eV was used. The cell volume, cell shape, and ion positions were optimized using the conjugate gradient algorithm. The convergence criteria were set to 10^{-9} eV for electronic minimization and 3×10^{-4} eV/Å for the forces acting on the ions.

After relaxation, the PAW pseudopotential for Te was modified according to the shLDA-1/2 method as proposed by Xue *et al.* [38]. This method is based on the LDA-1/2 [39,40] method, which aims to avoid the underestimation of band gaps with a GGA by correcting for the self-interaction of a localized hole in the valence band by adding a so-called self-energy potential to the pseudopotential. Based on Slater's half-occupation technique [41], the self-energy potential is found by subtracting the potential of the half-ionized atom from the unionized atom. Since this self-energy potential is added to every atom, it has to be trimmed to avoid divergent contributions. In the LDA-1/2 method, this is achieved with a spherical trimming function

$$\Theta(r) = \begin{cases} \left[1 - \left(\frac{r}{r_{\text{cut}}}\right)^n\right]^3 & r \leq r_{\text{cut}} \\ 0 & r > r_{\text{cut}}, \end{cases} \quad (1)$$

in which the cutoff radius has to be determined variationally with the condition that the resulting band gap is maximized. In the shLDA-1/2 method, the trimming function is replaced by a spherical shell

$$\Theta(r) = \begin{cases} \left[1 - \left(\frac{r}{r_{\text{out}}}\right)^m\right]^3 \frac{1 + \tanh[n(r - r_{\text{in}})]}{2} & r \leq r_{\text{out}} \\ 0 & r > r_{\text{out}}, \end{cases} \quad (2)$$

which is more suitable for crystals where the charge is not centered around the atom cores, but lies between two atoms. In this case, in addition to the outer cutoff radius r_{out} , an inner cutoff radius r_{in} has to be determined by the same method as before, keeping the outer cutoff radius constant. The self-energy corrected pseudopotentials for different cutoff radii have been constructed and the optimal cutoff radius determined by fitting a quadratic function of the cutoff radius to the resulting band gaps and finding the maximum. The corresponding DFT calculations used the same computational parameters as the relaxation, however, the crystal structure was kept constant and spin-orbit coupling was included.

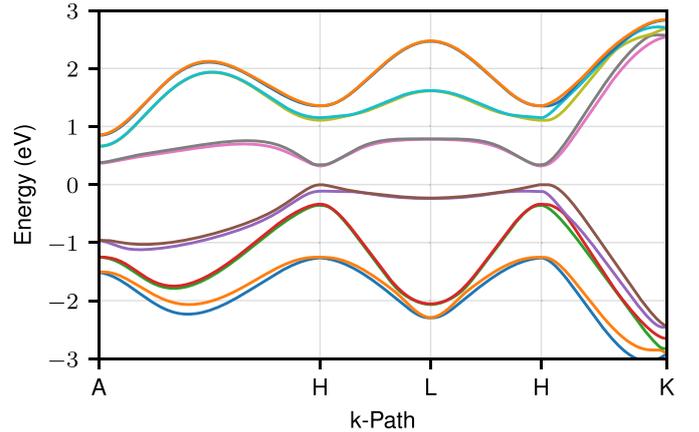


FIG. 1. Lowest six electron bands and highest six hole bands of tellurium calculated with the shLDA-1/2 method.

In a third set of calculations, the band structure and dipole matrix elements were determined. To this end, the charge-density of the self-consistently calculated ground-state obtained with the constructed pseudopotential was read in and kept constant. The k points were chosen along high symmetry lines in the BZ and the number of bands was increased, since a significant amount of empty conduction bands is needed for the optical routines of the VASP program that calculate the dielectric properties [42].

B. Band structure and dipole matrix elements

The results of the structural relaxation can be found in the first three columns of Table I, where a and c are the lattice constants and u is the parameter that determines the position of the atoms in the plane perpendicular to the helical chains. Comparison to experimental values shows that both lattice constants are slightly overestimated. For the construction of the self-energy corrected pseudopotential, the optimized inner and outer cutoff radii were determined as 1.328 Å and 3.395 Å, respectively. The resulting direct band gap at the H point, E_g , and splitting of the light-hole and heavy-hole valence band at the H point, E_{LH-HH} , are compared to the experimental values in Table I. Both the gap and the valence band splitting are in very good agreement with the experiment, underestimating the experimental values slightly by 2% and 1%, respectively. The complete band structure along high symmetry lines in the BZ is shown in Fig. 1.

From the wave functions, ϕ , obtained from DFT, the transition dipole moments (TDMs) $\mathbf{d}_{\mathbf{k}}^{n'n}$ between bands n and n' at every k -point \mathbf{k} are determined:

$$d_{\mathbf{k}}^{n'n} = \frac{\hat{\mathbf{e}}}{\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}}} \cdot \left\langle \phi_{n'\mathbf{k}} \left| \frac{\partial(\mathbf{H} - \epsilon_{n\mathbf{k}}\mathbf{S})}{\partial\mathbf{k}} \right| \phi_{n\mathbf{k}} \right\rangle. \quad (3)$$

Here ϵ denotes the single-particle energies and $\hat{\mathbf{e}}$ is the polarization direction. \mathbf{H} is the Hamilton operator for the cell periodic wave functions and \mathbf{S} is the corresponding overlap operator [42]. An overview of the TDMs projected onto the z direction for optical fields polarized parallel to the c axis ($E \parallel c$) and onto the x direction for $E \perp c$ is shown in Fig. 2. Here, the modulus of the TDMs between the two lowest conduction bands and four highest valence bands are shown

TABLE I. Comparison of structural and electronic parameters from *ab initio* DFT calculations using the shLDA-1/2 method with experimental results.

	Structural parameters			Electronic properties	
	a	c	u	E_g	E_{LH-HH}
DFT	4.51 Å	5.96 Å	0.27	0.323 eV	0.111 eV
Exp. [4,45,46]	4.46 Å	5.92 Å	0.267	0.33 eV	0.112 eV

in a momentum vector plane spanned by the Γ , A , H , L , M , and K points of the BZ.

While the scale of the color map in Fig. 2 is capped at ten, the maximum value of the dipoles between two valence bands or two conduction bands far exceeds that limit. This can be explained from Eq. (3), since bands of the same type are very close to each other up to the point of almost becoming

degenerate, so the factor $1/(\epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}})$ becomes very large. For the interband dipoles between a conduction and a valence band for $E \perp c$, the strongest dipole coupling is found around the direct band gap at the H point. For $E \parallel c$, the interband dipoles involving the two highest valence bands are vanishingly small at the H point, however, they become stronger when moving away from the H point along the H - L line. Only

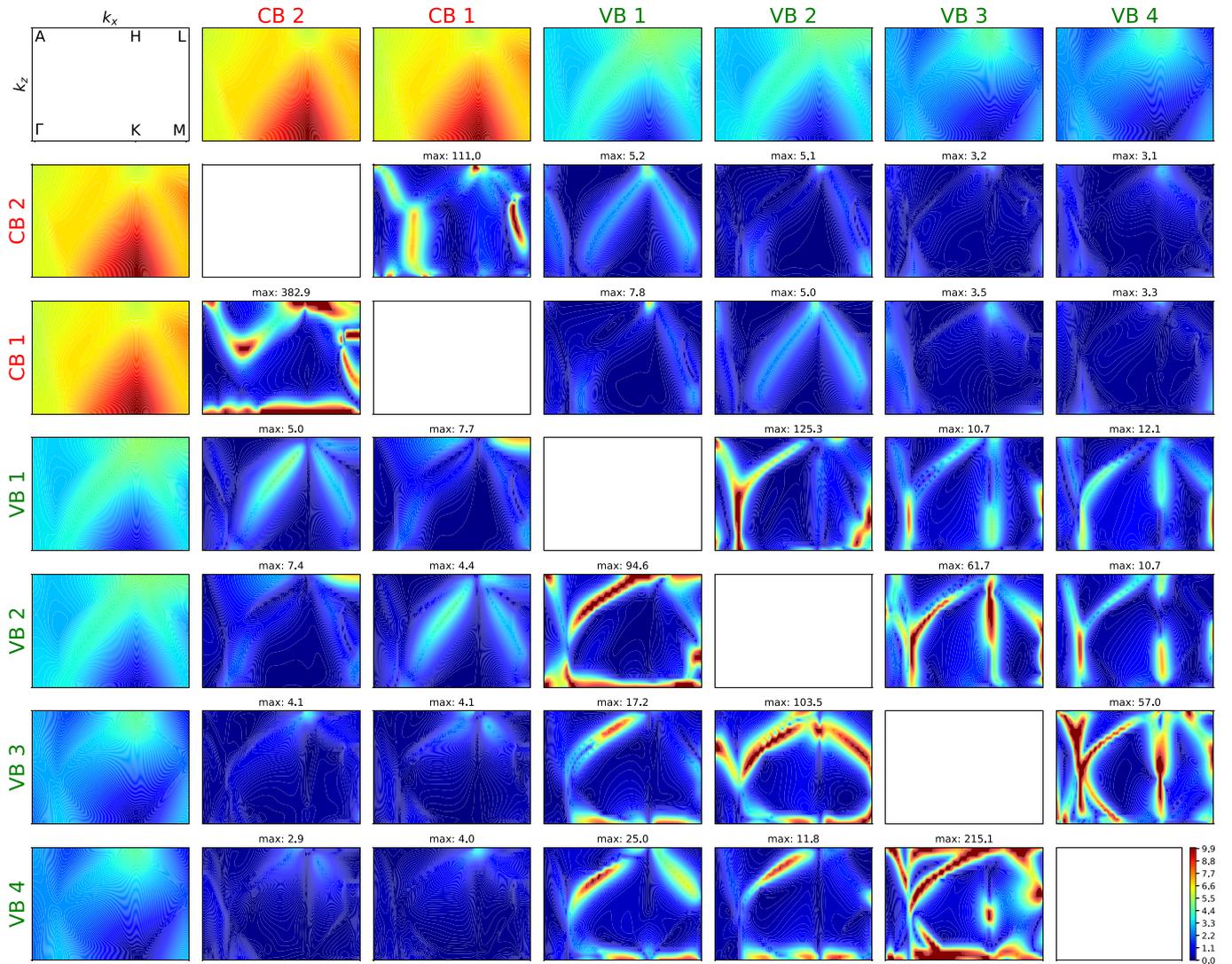


FIG. 2. Dipole transition matrix elements between bands in a 2D plane of the 1. BZ spanned by the Γ , M , L and A points. The first row and first column show the band energies, while the inner plots show the dipole strengths. The dipole in a particular cell corresponds to the band combination given by the bands in the respective row and column. In the bottom left triangle, the dipoles for $E \parallel c$ direction are given, in the top right triangle, the dipoles for $E \perp c$ direction are given. The color bar in the bottom right cell pertains to all dipoles. Values higher than ten are not distinguished in the color plot. The max value above each dipole plot indicates the maximum value of the respective dipole at any point in the plane.

the interband dipoles of the two lower-lying valence bands, VB3 and VB4, are significant at the H point. We will utilize this feature to simplify the optical response calculations for $E \parallel c$ by omitting the two upper valence bands, VB1 and VB2.

Generally, strong dipole coupling is found in parameter regions where the bands are close to each other. However, there are differences between the dipoles for $E \parallel c$ and $E \perp c$ although the band energies are the same for both polarization directions, e.g., the intraband dipoles are strong along the Γ - K - M line for $E \parallel c$, while there is no significant coupling for $E \perp c$. Conversely, the coupling along the H - K line is much stronger for $E \perp c$ than for $E \parallel c$.

III. INCOHERENT RESONANT NONLINEARITIES

To test the results of our DFT calculations, we use the band structures, wave functions, and TMDs to evaluate absorption spectra for Te and compare them to experimentally measured results. The absorption is calculated for two polarization directions of the exciting light field, $E \parallel c$ and $E \perp c$. In the BZ, these directions correspond to the H - K and H - L - H - A paths, respectively.

Linear absorption spectra are computed by applying an arbitrarily small field $E(t)$ and calculating the material response $P(t)$ by solving the equations of motion for the microscopic polarizations, $p_{\mathbf{k}}^{ji}$, i.e., the SBE [12,43],

$$\begin{aligned} \frac{d}{dt} p_{\mathbf{k}}^{j_1 i_1} &= \frac{1}{i\hbar} \left(\sum_{i_2, j_2} [\tilde{\epsilon}_{j_1 j_2, \mathbf{k}}^h \delta_{i_1 i_2} + \tilde{\epsilon}_{i_1 i_2, \mathbf{k}}^e \delta_{j_1 j_2}] p_{\mathbf{k}}^{j_2 i_2} \right. \\ &\quad \left. + [1 - f_{i_1, \mathbf{k}}^e - f_{j_1, \mathbf{k}}^h] \Omega_{\mathbf{k}}^{i_1 j_1} \right) \\ &\quad + \left. \frac{d}{dt} p_{\mathbf{k}}^{j_1 i_1} \right|_{\text{corr}}, \end{aligned} \quad (4)$$

with the renormalized electron and hole energies

$$\tilde{\epsilon}_{i_1 i_2, \mathbf{k}}^e = \epsilon_{i_1, \mathbf{k}}^e \delta_{i_1 i_2} - \sum_{i_3, q} V_{\mathbf{k}-\mathbf{q}}^{i_1 i_3 i_2 i_3} f_{i_3, \mathbf{q}}^e, \quad (5)$$

$$\tilde{\epsilon}_{j_1 j_2, \mathbf{k}}^h = \epsilon_{j_1, \mathbf{k}}^h \delta_{j_1 j_2} - \sum_{j_3, q} V_{\mathbf{k}-\mathbf{q}}^{j_2 j_3 j_1 j_3} f_{j_3, \mathbf{q}}^h, \quad (6)$$

and the renormalized generalized Rabi frequency:

$$\Omega_{\mathbf{k}}^{i_1 j_1} = -d_{\mathbf{k}}^{i_1 j_1} E(t) - \sum_{i_2, j_2, q} V_{\mathbf{k}-\mathbf{q}}^{i_1 j_2 j_1 i_2} p_{\mathbf{q}}^{j_2 i_2}. \quad (7)$$

Here, i_1, i_2, i_3 are electron band indices and j_1, j_2, j_3 are hole band indices. Like the dipole matrix elements, the Coulomb matrix elements V are evaluated using the DFT wave functions.

For the linear absorption calculations, the material is assumed to be in the unexcited ground state and the field is too weak to create carriers such that the occupations for electrons/holes $f^{e/h}$ remain zero. For gain calculations, the carriers are assumed to be in thermal equilibrium and described by Fermi distributions within the respective bands. This fully microscopic approach has been shown to yield very good quantitative agreement with the experiment for a wide variety of materials spanning the mid-IR to visible wavelength ranges (see, e.g., Ref. [44] for examples). When using proper

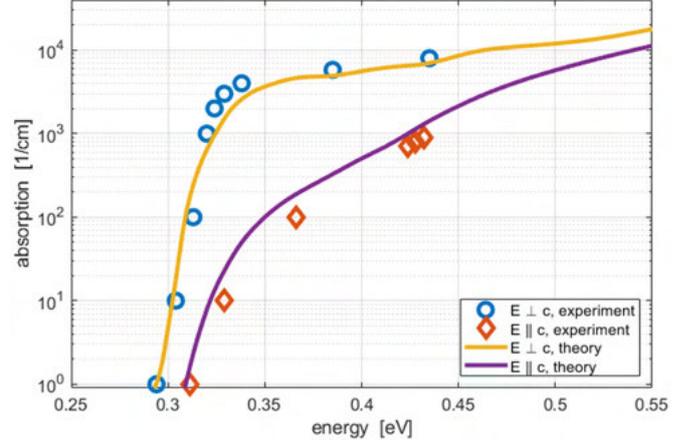


FIG. 3. Room-temperature material absorption of Te for light polarized $\parallel c$ (smaller) and $\perp c$ (larger). Solid lines: Theoretical results based on DFT. Symbols: Experimental data extracted from Ref. [6]. The experimental data was shifted by 14 meV to lower energies.

band structures and wave functions like the ones here from first-principles DFT calculations, the approach does neither require nor allow any fit or scaling parameters to improve agreement with the experiment. Thus, finding quantitative agreement with the experiment demonstrates a truly predictive quality of the model.

The term $\frac{d}{dt} p_{\mathbf{k}}^{j_1 i_1} |_{\text{corr}}$ summarizes higher order correlations that include the electron-electron and electron-phonon scattering which lead to the dephasing of the polarization and the resulting homogeneous broadening of the spectra. We include the scatterings on a fully microscopic level by solving the corresponding quantum-Boltzman type scattering equations. Standard literature parameters are used for the phonon scattering as discussed in Ref. [43]. The explicit calculation of the dephasing processes not only eliminates adjustments requiring empirical parameters but has also been shown essential to obtain correct line shapes, amplitudes, spectral positions, and density dependencies.

From the Fourier transform of the macroscopic polarization $P(t) = \sum_{i,j,\mathbf{k}} p_{\mathbf{k}}^{ji} d_{\mathbf{k}}^{ij*}$, the absorption coefficient α is calculated according to

$$\alpha(\omega) = \frac{\omega}{\epsilon_0 n_r(\omega) c E(\omega)} \text{Im}[P(\omega)]. \quad (8)$$

In Fig. 3, we plot the resulting absorption spectra for the polarizations $E \perp c$ and $E \parallel c$. Especially near the band gap, the absorption for $E \perp c$ is much larger than the one for $E \parallel c$ due to the weaker coupling between the topmost valence bands and the conduction bands near the band gap for $E \parallel c$ (see Sec. II B). The blue and red dots in Fig. 3 show the results of measurements extracted from Ref. [6]. As we can see, our computed results agree well with the experimentally measured spectra.

As noted in Ref. [6], it is difficult to determine the reason for the strong polarization dependence of the absorption from measurement alone. While the authors of Ref. [6] assumed that the absorption for $E \parallel c$ is suppressed due to an indirect gap, Refs. [47,48] concluded that the real reasons are the selection rules leading to forbidden transitions for this

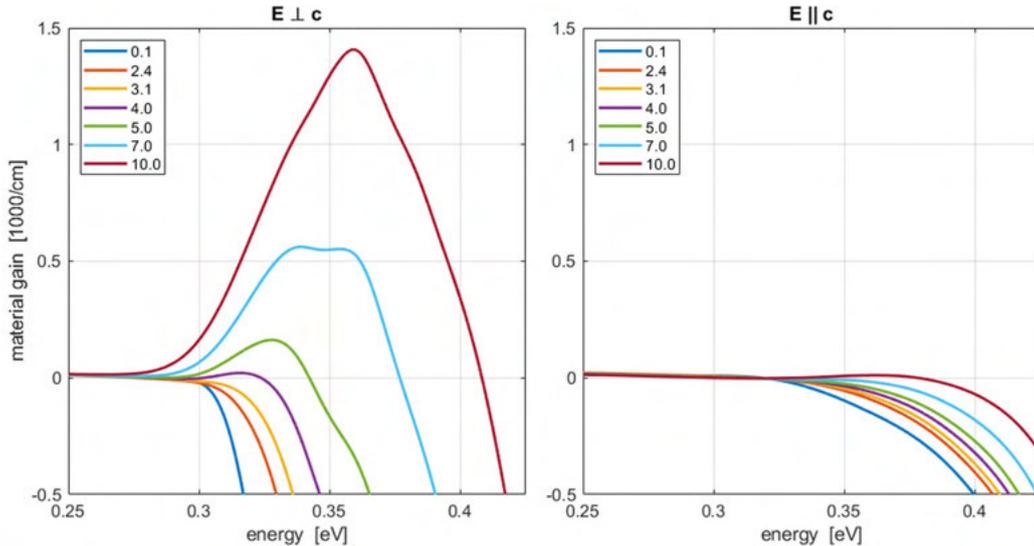


FIG. 4. Room-temperature material gain (negative absorption) spectra of Te for light polarized $\perp c$ (left) and $\parallel c$ (right) at various carrier densities. The carrier densities are given in the labels in units of $10^{19}/\text{cm}^3$.

configuration. This assumption is fully confirmed by our DFT calculations.

The authors of Ref. [6] state the band gap of their sample to be around $0.325 - 0.337$ eV compared to our value of 0.323 eV and the value of about 0.33 eV from Refs. [4,45,46]. In Ref. [11], the authors report that it was possible to shift the band-edge PL of their sample by about 24 meV through annealing. This indicates that the band gap of Te can vary due to effects like sample quality or strain by amounts that can explain the difference found between our results and those in Ref. [6]. In Fig. 3, we account for the difference in band gaps by shifting the experimental data extracted from Ref. [6] by 14 meV.

Encouraged by the good agreement of the computed and measured linear absorption spectra, we use our microscopic approach to investigate the nonlinear optical properties of bulk Te. In a first step, we assume that the material has been excited to generate significant densities of incoherent electron and hole populations in the respective bands. As an example, we show in Fig. 4 the calculated optical material gain $[-\alpha(\omega)]$ for $E \perp c$ and $E \parallel c$ and various carrier densities. We see that for $E \perp c$, gain begins to occur for carrier densities above $4 \times 10^{19}/\text{cm}^3$. For densities above about $7 \times 10^{19}/\text{cm}^3$, the peak gain shifts from the CB1-VB1 transition with a peak around $0.33 - 0.35$ eV to the second conduction band transition, CB2-VB1, with a peak near 0.37 eV.

As seen in the linear absorption spectra, the TDMs are much smaller for $E \parallel c$ than for $E \perp c$ in the spectral range where gain would occur. This leads to virtually no gain at all for this polarization direction at realistic carrier densities.

Assuming the same excitation conditions, the resulting PL is calculated by solving the SLE [13], i.e., the microscopic equations of motion for the photon-assisted polarizations. Structurally, the SLE have the same form as the SBE, Eq. (4), but include higher excitonic correlations as an additional source term. As for the SBE, we include in our SLE evaluations the electron-electron and electron-phonon scattering on a fully microscopic level.

Figure 5 shows PL spectra for $E \perp c$ and $E \parallel c$ at various levels of electron-hole-pair populations. In the low density regime, the PL scales quadratically with the carrier density. Plotting the PL divided by the square of the density as in Fig. 5 reveals deviations for higher excitation levels from this quadratic variation that are due to phase-space filling [49]. In this regime, the density dependence becomes less than quadratic and the PL peak shifts to higher transition energies. For the case of $E \parallel c$, the peak shift is stronger and the amplitude reduction is slower. These features can be attributed to the fact that the TDMs for $E \parallel c$ increase significantly with increasing energy above the gap which enhances energetically higher PL contributions. Like the gain and absorption, the PL is much weaker for $E \parallel c$ than for $E \perp c$ due to the much smaller TDMs in the energy region of interest. This agrees with the experimentally observed dominant polarization $E \perp c$ of PL in Ref. [10]. The spectral position as well as the line shape of our calculated PL agrees very well with experimentally measured data from Ref. [11] that we include in Fig. 5 for comparison. This demonstrates the high accuracy of the fully microscopic modeling approach including the explicit treatment of scattering processes that lead to an almost perfect agreement with the experimentally observed linewidth of about 80 meV.

IV. COHERENT OFF-RESONANT NONLINEAR RESPONSE

A. Microscopic approach

To model the nonlinear optical response of a crystal to a strong exciting THz field, the coupled dynamics of interband polarizations and intraband currents have to be investigated. For this purpose, we again use the SBE. However, in contrast to the quasistationary nonlinear response investigated so far, we now have to explicitly include the nonequilibrium carrier dynamics. In particular, the strong long-wavelength excitation field leads to an acceleration of carriers along the bands throughout the entire BZ. Thus, the results depend critically on the dispersion relation across the whole BZ.

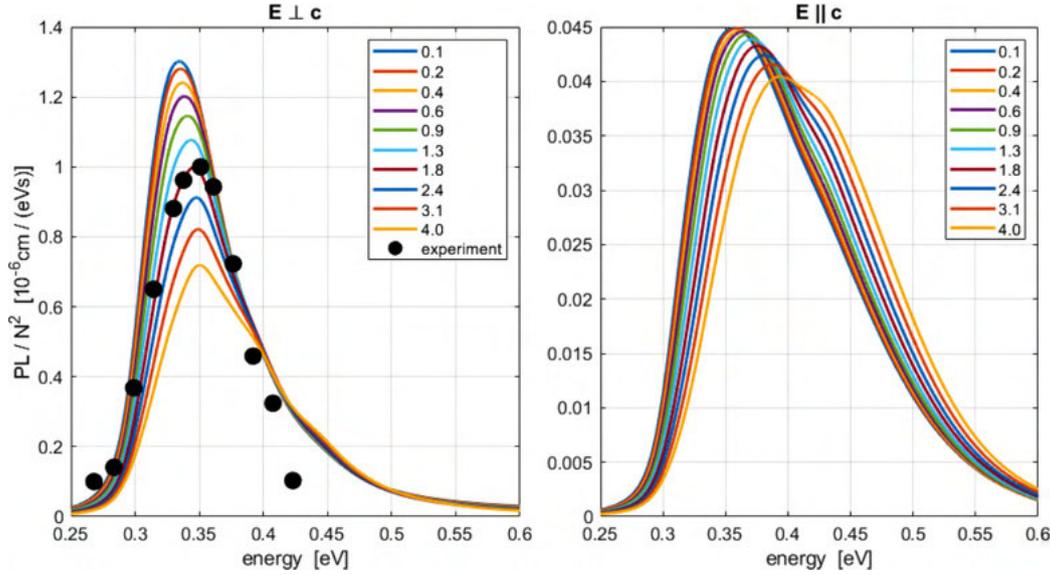


FIG. 5. Theoretical (lines) and experimental (symbols) room-temperature photoluminescence spectra of Te for light polarized $\perp c$ (left) and $\parallel c$ (right) at various carrier densities. The theoretical spectra have been divided by the respective density squared. The carrier densities are given in the labels in units of $10^{19}/\text{cm}^3$. The experimental data extracted from Ref. [11] are given in arbitrary units.

Furthermore, pulse propagation effects have to be included to study the dependence of HHG on sample length.

In earlier studies, we have shown that for the strongly off-resonant excitation assumed here, the Coulomb renormalizations have a negligible influence [50] such that the equations of motion can be simplified to

$$\begin{aligned}
 i\hbar \frac{d}{dt} p_{\mathbf{k}}^{h_i e_j} &= (\epsilon_{\mathbf{k}}^{e_j} + \epsilon_{\mathbf{k}}^{h_i} + i|e|E_{\text{THz}}(t)\nabla_{\mathbf{k}}) p_{\mathbf{k}}^{h_i e_j} \\
 &\quad - \hbar\Omega_{\mathbf{k}}^{h_i e_j}(t)(1 - f_{\mathbf{k}}^{e_j} - f_{\mathbf{k}}^{h_i}) + \Gamma_{\mathbf{k}}^{h_i e_j} \\
 &\quad + \sum_{e_{\lambda} \neq e_j} [\hbar\Omega_{\mathbf{k}}^{h_i e_{\lambda}}(t) p_{\mathbf{k}}^{e_{\lambda} e_j} - \hbar\Omega_{\mathbf{k}}^{e_{\lambda} e_j}(t) p_{\mathbf{k}}^{h_i e_{\lambda}}] \\
 &\quad + \sum_{h_{\lambda} \neq h_i} [\hbar\Omega_{\mathbf{k}}^{h_{\lambda} h_i}(t) p_{\mathbf{k}}^{h_{\lambda} e_j} - \hbar\Omega_{\mathbf{k}}^{h_{\lambda} e_j}(t) p_{\mathbf{k}}^{h_{\lambda} h_i}] \\
 &\quad + \left. \frac{d}{dt} p_{\mathbf{k}}^{h_i e_j} \right|_{\text{corr}}, \tag{9}
 \end{aligned}$$

$$\begin{aligned}
 \hbar \frac{d}{dt} f_{\mathbf{k}}^{e_i} &= -2\hbar \\
 &\quad \times \text{Im} \left[\sum_{e_{\lambda} \neq e_i} \Omega_{\mathbf{k}}^{e_{\lambda} e_i}(t) (p_{\mathbf{k}}^{e_{\lambda} e_i})^* + \sum_{h_{\lambda}} \Omega_{\mathbf{k}}^{h_{\lambda} e_i}(t) (p_{\mathbf{k}}^{h_{\lambda} e_i})^* \right] \\
 &\quad + |e|E_{\text{THz}}(t)\nabla_{\mathbf{k}} f_{\mathbf{k}}^{e_i} + \Gamma_{\mathbf{k}}^{e_i}. \tag{10}
 \end{aligned}$$

We have similar expressions for the intraband polarizations between conduction bands $p_{\mathbf{k}}^{e_i e_j}$ and between valence bands $p_{\mathbf{k}}^{h_i h_j}$ and the carrier occupations in the valence band $f_{\mathbf{k}}^{h_i}$, respectively. For HHG, we model the dephasing of the polarization as represented by the last term in Eq. (9) using a dephasing time $T_2 = 40$ fs.

The macroscopic polarization $P(t) = \sum_{\lambda, \lambda', \mathbf{k}} d_{\mathbf{k}}^{\lambda \lambda'} p_{\mathbf{k}}^{\lambda \lambda'}$ and the macroscopic current $J(t) = \sum_{\lambda, \mathbf{k}} j_{\lambda}(\mathbf{k}) f_{\mathbf{k}}^{\lambda}$ due to the acceleration of carriers along the bands contribute to the emitted

electric field $E_{\text{out}}(t) \propto \frac{\partial}{\partial t} P(t) + J(t)$ and create the characteristic local high harmonic emission spectrum which is given by the emission intensity $I_{\text{out}}(\omega) \propto |\omega P(\omega) + iJ(\omega)|^2$.

To gain some insights before doing the full propagation calculations, we performed local evaluations which need significantly less numerical effort. Here, we use a one-dimensional \mathbf{k} -space model which assumes that carriers are predominantly excited near the fundamental gap, i.e., near the H point with negligible momentum perpendicular to the field. For linearly polarized light, the carriers are then driven along a one-dimensional path through the BZ. For $E \parallel c$, the path is from K to H and back to K . For $E \perp c$, the path goes from A to H to L and back. For all HHG simulations, we assume excitation with a Gaussian pulse, $E(t) = E_0 \exp(-(t/\sigma)^2) \cos(\omega_0 t)$, with a width $\sigma = 100$ fs and a central frequency ω_0 corresponding to a wavelength of $10.6 \mu\text{m}$.

In a first step, we use this local model to identify those bands that are relevant for HHG generation under typical off-resonant excitation conditions. Clearly, the HHG signal is dominated by transitions between those bands which are energetically closest to the band gap unless these transitions are suppressed due to symmetry selection rules leading to small TDMs. As can be seen in Fig. 1, only four valence and two electron bands are in the energetically relevant region. Since the TDMs presented in Fig. 2 show that the coupling of the top two valence bands to the lowest two conduction bands vanishes at the H point for $E \parallel c$, we studied whether these bands are significant for the resulting HHG spectrum.

A comparison of the computed spectra including different valence bands is shown in Fig. 6(a). We note that by considering only the bottom two valence bands, we obtain a spectrum that agrees rather well with the full six-band calculation, allowing us to reduce the complexity of our propagation studies for $E \parallel c$ by including only this subset of bands. In contrast, for the $E \perp c$ configuration, the top two valence

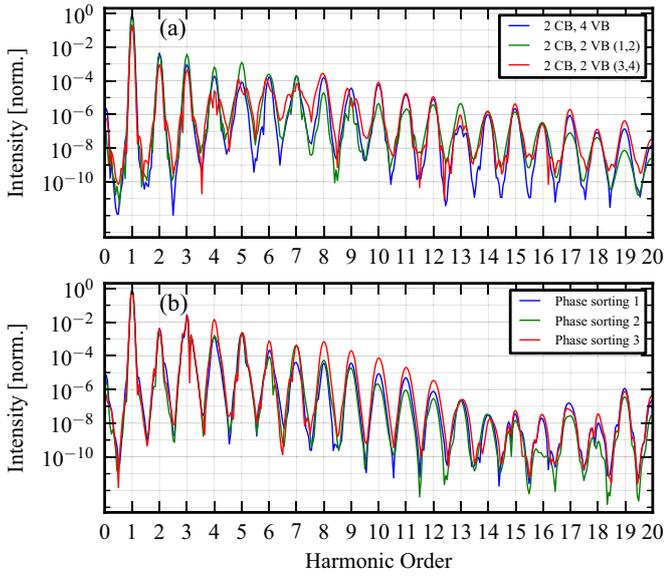


FIG. 6. Polarization part of $E \parallel c$ HHG emission in Te. (a) Influence of different choices of bands on HHG emission. (b) Influence of choice of TDM phases on HHG emission.

bands dominate the response and are thus included in the HHG simulations.

B. Phase of transition dipole matrix elements

In general, the TDMs presented in Sec. II B are complex valued. To illustrate the influence of the phases on the HHG emission, we consider a perturbative power series of the polarization response to an electric field for a situation with two valence bands h_1, h_2 , and one conduction band e . In the first order of the field, all polarizations and occupations are 0, so we obtain from Eq. (9):

$$(p_{\mathbf{k}}^{h_1e}(t))^{(1)} \propto \frac{1}{\hbar\omega} d_{\mathbf{k}}^{eh_1} E(t). \quad (11)$$

The resulting macroscopic polarization then yields

$$(P_{h_1e}(t))^{(1)} = \sum_{\mathbf{k}} d_{\mathbf{k}}^{h_1e} (p_{\mathbf{k}}^{h_1e}(t))^{(1)} \propto \frac{|d_{\mathbf{k}}^{h_1e}|^2}{\hbar\omega} E(t). \quad (12)$$

Hence, the phase of the TDMs in this first-order response is irrelevant. However, since the polarizations are nonzero in second order, the creation of a polarization between the valence bands allows for an indirect excitation into the conduction band:

$$(p_{\mathbf{k}}^{h_1e}(t))^{(2)} \propto \frac{d_{\mathbf{k}}^{eh_1}}{\hbar\omega} E(t) + \frac{d_{\mathbf{k}}^{h_2h_1} d_{\mathbf{k}}^{eh_2}}{2\hbar^2\omega^2} E^2(t) + \dots \quad (13)$$

This leads to a term in the macroscopic polarization,

$$(P_{h_1e}(t))^{(2)} \propto \sum_{\mathbf{k}} \frac{d_{\mathbf{k}}^{h_2h_1} d_{\mathbf{k}}^{h_1e} d_{\mathbf{k}}^{eh_2}}{2\hbar^2\omega^2} E^2(t) + \dots, \quad (14)$$

where the phases of the TDMs do not vanish. If, e.g., one of the TDMs in Eq. (14) is antisymmetric in \mathbf{k} and the other two are symmetric, the integration over \mathbf{k} will lead to a zero contribution to the macroscopic polarization and resulting

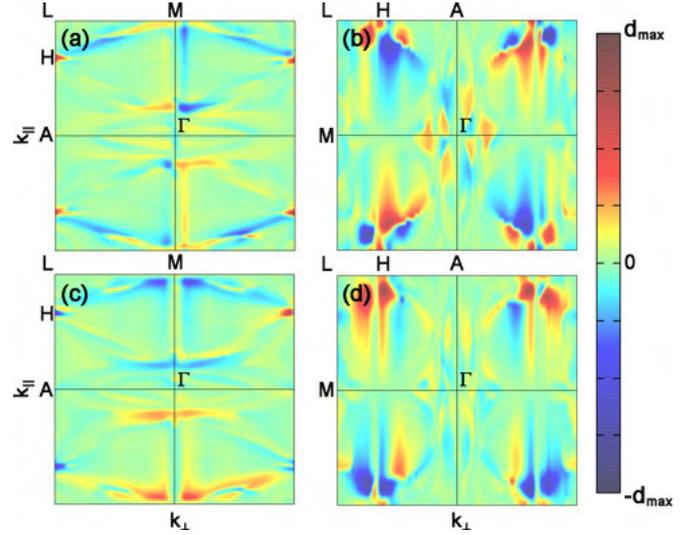


FIG. 7. Complex dipole matrix elements between the lowest conduction and highest valence band. (a) and (c) are the real and imaginary parts for $E \perp c$. (b) and (d) are the real and imaginary parts for $E \parallel c$. k_{\perp} (k_{\parallel}) is the momentum perpendicular (parallel) to the field polarization. $d_{\max} = 6, 8, 4,$ and 5 for (a)–(d), respectively.

HHG signal while a strong nonzero contribution would be obtained if the phases are neglected. Thus, the phases need to be considered correctly to obtain the correct symmetry-related selection rules and amplitudes in the HHG calculations.

It was shown in Ref. [50] that quantum interference between intervalenceband and valence-to-conductionband polarizations can lead to the appearance of even harmonics. Recently, studies showed that phase-sensitive effects contribute to even harmonic generation in solids and inclusion of the transition dipole phase is crucial for models of HHG [51–53]. Moreover, neglecting the phase of the TDMs can result in an overall change of the symmetry of the system. Since HHGs are known to be highly sensitive especially to inversion symmetry, this often leads to significant changes in the proposed HHG spectra [54]. Thus, the correct inclusion of the phases is essential to obtain the correct selection rules for HHG.

While the TDMs are complex valued, the plot in Fig. 2 only shows the absolute value. In DFT, the Schrödinger equation of every \mathbf{k} point is solved individually, so there is no phase relation between different \mathbf{k} points. Therefore, the computed TDMs contain a random phase which is not smooth across the BZ. As it turns out, this random phase can be eliminated for all \mathbf{k} points by evaluating the product of the three complex TDMs connecting the bands $n, n',$ and n'' in a circular way, e.g., $T_{\mathbf{k}}^{nn'n''} = d_{\mathbf{k}}^{nn'} d_{\mathbf{k}}^{n'n''} d_{\mathbf{k}}^{n''n}$. The random phase of each band vanishes in the product, so the phase of $T_{\mathbf{k}}^{nn'n''}$ along any direction in the BZ is smooth. Since this only gives us the phase information about the product of three TDMs, this phase is applied to one of the constituent TDMs while taking the other ones as purely real. In that way, the triple product will have the correct phase.

As an example, we show in Fig. 7 the complex TDMs for the respective transitions between the lowest electron and highest hole bands taken into account for $E \parallel c$ and $E \perp c$.

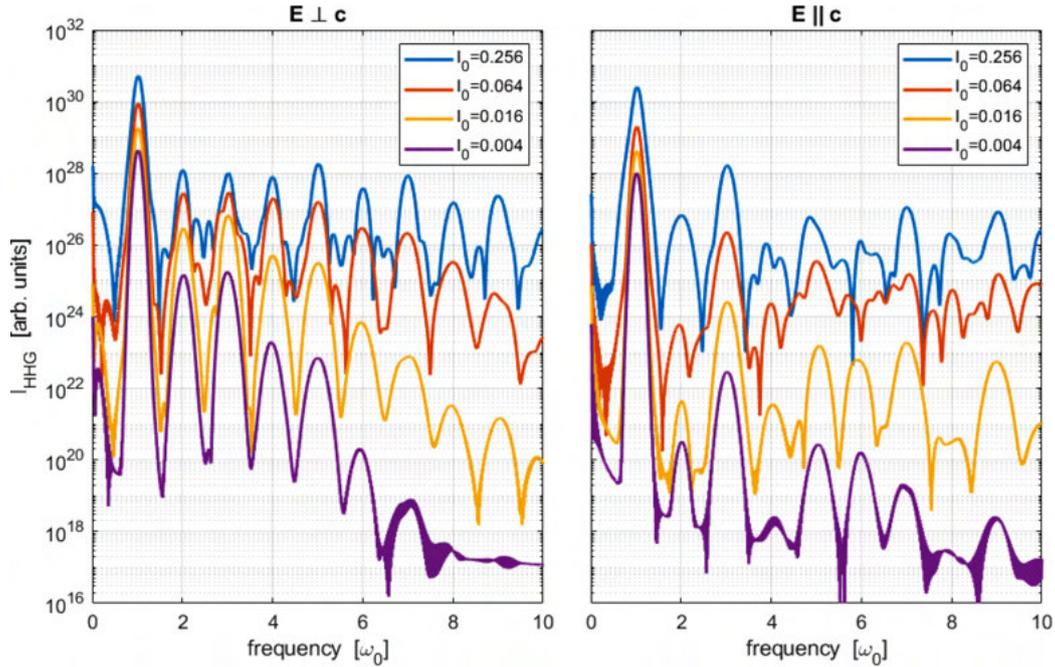


FIG. 8. HHG spectra in Te for $E \perp c$ (left) and $E \parallel c$ (right) and various peak intensities I_0 . Intensities given in the labels are in units of $10^{14}W/m^2$.

In all four plots, the momentum parallel to the field polarization is vertically aligned. Once the phases of the dipoles are taken into account, it becomes obvious that the Te system does not have pure radial or inversion symmetry. Thus, the \mathbf{k} domain has to be expanded from the positive sector Γ - A - L - M to four times the size to include also negative k_x and k_z . For the $e1 - h1$ transitions presented in Fig. 7, the real and imaginary parts of the dipoles for $E \parallel c$, shown in the two right-hand plots, appear nearly antisymmetric along the polarization direction. In contrast, the symmetry properties of the real and imaginary parts of the dipoles for $E \perp c$ are a little more ambiguous. As for the Te crystal itself, the TDMs do not have perfect (anti)symmetry. This can be seen, e.g., in the real parts of the TDMs for $E \perp c$ in Fig. 7. These are nearly symmetric near H while they appear mostly antisymmetric in most regions of small k_{\perp} . The imaginary parts for $E \perp c$ in Fig. 7(c) appear mostly antisymmetric, with slight deviations around the H-point.

In our procedure to assign the TDM phase, we arbitrarily choose the dipoles onto which we impose the smoothed phase of the triple dipole products. To check how this choice influences the HHG spectrum, we calculated the polarization part of the spectra for different phase projections. As can be seen in Fig. 6(b), our phase assignment does not influence the overall structure of the spectra, leading only to insignificant amplitude changes, so the comparisons between HHG calculations for different intensities, propagation lengths, etc. is robust against this choice for the dipole phases.

C. High harmonics in Te

To determine the dependence of HHG production in Te on the field strength, we performed calculations for the material response only, without pulse propagation. Figure 8 shows

the resulting emission for various intensities of the exciting pulse.

For both polarization configurations, a significant signal above the band gap (frequencies above the third harmonic) develops for peak intensities above about $10^{11}W/m^2$. A plateau starts to form for about 100 times higher intensities. Harmonics below the band gap emerge rather quickly for $E \perp c$ and start to already saturate at amplitudes about three orders below that of the fundamental. For $E \parallel c$, the signal below the band gap develops much slower with field intensity. In particular, the third harmonic shows less saturation for the intensities investigated here. The differences at and below the band gap are due to the fact that the interband coupling is much weaker at and near the gap, as can be seen from the absorption spectra.

Even harmonics are strongly suppressed for $E \parallel c$ while for $E \perp c$ no obvious discrimination occurs between even and odd harmonics. This behavior is a consequence of the symmetry of the dipole matrix elements. As in the case for the lowest electron-hole transition shown in Fig. 7, all dipoles that are relevant for even harmonics are nearly inversion symmetric for $E \parallel c$. This leads to a destructive quantum interference that suppresses the even harmonics. In contrast, for $E \perp c$ the relevant dipoles are dominantly symmetric which effectively enables quantum interference and allows for the even harmonics to reach similar levels as the odd harmonics.

To evaluate HHG for samples of different thicknesses, we include pulse propagation effects by coupling the SBE [Eq. (9)] to a unidirectional pulse propagation solver as described in Ref. [55] and references therein. Here we assume a perfect crystal without background absorption losses. The background refractive index for Te and $E \perp c$ and $E \parallel c$ has been taken from Ref. [56]. We calculate the contribution to the index from the optical polarization in the linear, weak field case and subtract that from the background index.

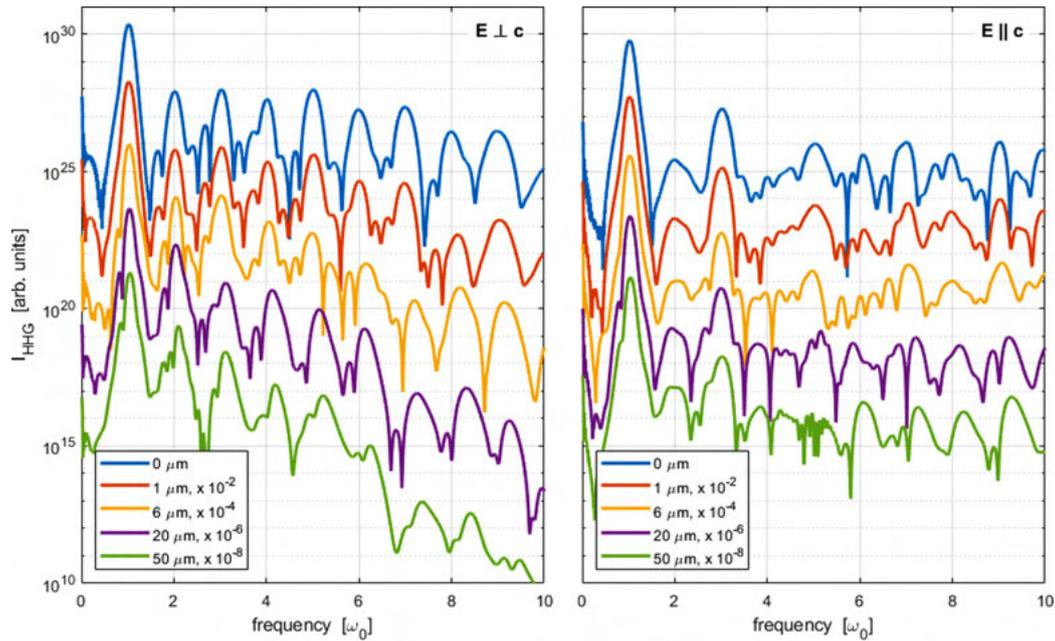


FIG. 9. HHG spectra in Te for $E \perp c$ (left) and $E \parallel c$ (right), a peak pulse intensity of $0.128 \times 10^{14} \text{W/m}^2$ and for various propagation distances. Spectra for different propagation distances have been scaled by factors of 100 for better visibility.

As an example of the results, Fig. 9 shows HHG spectra after propagation through Te for various distances. The initial pulse has a peak intensity of $0.128 \times 10^{14} \text{W/m}^2$.

For $E \perp c$, the higher harmonics quickly weaken with propagation distance. In part, this is a consequence of the gradual decreasing excitation pulse due to HHG and absorption of spectral components above the band gap. In part, this is also due to propagation-induced dephasing [55]. This weakening is less pronounced for $E \parallel c$ since the absorption is weaker and less HHG signal is produced. Over the maximum propagation distance investigated here ($50 \mu\text{m}$), the amplitude of the fundamental drops by about a factor of 10 for $E \perp c$ and only a factor of 4 for $E \parallel c$. The reduced amount of even harmonics for $E \parallel c$ likely also leads to a reduced amount of quantum interference and resulting propagation-induced dephasing within the remaining signal.

V. SUMMARY AND OUTLOOK

In summary, we present a comprehensive microscopic analysis of optical nonlinearities in bulk Te. We determine the band structure, the optical dipoles, and the Coulomb interaction matrix elements using an DFT-based approach. Investigating the near band-gap optical response for different levels of electron-hole-pair excitations, we numerically solve the stationary SBE and SLE to compute the strongly orientation dependent absorption and PL modifications. Comparing the linear absorption and PL spectra with experimental findings, we obtain excellent agreement. For elevated excitation levels, we obtain a transition from absorption to optical gain for $E \perp c$ gain with a peak in the technologically interesting mid-IR region. Since the TDMs are much smaller for $E \parallel c$ than for $E \perp c$, virtually no gain occurs for this polarization direction at realistic carrier densities.

The generation of high-harmonic emission in Te is analyzed using the fully dynamic SBE systematically treating the nonequilibrium dynamics of the optically induced polarizations and currents. Pulse propagation effects are modeled by coupling the SBE to a unidirectional propagation solver that allows us to study the sample length and field-orientation dependence of the even- and odd-order HHG for the different field polarization configurations. The importance of a correct treatment of the complex phases of dipole matrix elements for the correct description of optical selection rules is demonstrated.

As a next step, we plan to evaluate the intrinsic losses in bulk Te, in particular, the Auger losses that typically hamper the laser application potential of mid-IR emitting structures. Furthermore, we will extend our comprehensive microscopic approach to low-dimensional Te [57] to investigate its nonlinear optoelectronic properties and device application potential.

ACKNOWLEDGMENTS

The authors thank D. Matteo, S. Tochitsky, UCLA, for stimulating discussions during the early part of these investigations and I. Kilen, M. Kolesic, University of Arizona, for development of the numerical HHG propagation code. The Marburg work was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Research Training Group Functionalization of Semiconductors (GRK No. 1782) and the Collaborative Research Center No. SFB 1083. The authors thank the HRZ Marburg and CSC-Goethe-HLR Frankfurt for computational resources. The Tucson work was supported by the Air Force Office of Scientific Research under Awards No. FA9550-17-1-0246 and No. FA9550-19-1-0032.

- [1] J. H. McFee, G. D. Boyd, and P. H. Schmidt, *Appl. Phys. Lett.* **17**, 57 (1970).
- [2] V. V. Berezovskii, Y. A. Bykovskii, M. I. Goncharov, and I. S. Rez, *Sov. J. Quantum Electron.* **2**, 180 (1972).
- [3] V. V. Sobolev, *Dokl. Akad. Nauk SSSR* **151**, 1308 (1963).
- [4] J. Stuke and H. Keller, *Phys. Stat. Sol.* **7**, 189 (1964).
- [5] J. J. Loferski, *Phys. Rev.* **93**, 707 (1954).
- [6] S. Tutihasi, G. G. Roberts, R. C. Keezer, and R. E. Drews, *Phys. Rev.* **177**, 1143 (1969).
- [7] R. S. Caldwell and H. Y. Fan, *Phys. Rev.* **114**, 664 (1959).
- [8] C. K. N. Patel, *Phys. Rev. Lett.* **15**, 1027 (1965).
- [9] S. S. Tsirkin, P. A. Puente, and I. Souza, *Phys. Rev. B.* **97**, 035158 (2018).
- [10] A. Benoit, C. La Guillaume, and J. M. Debver, *Sol. Stat. Commun.* **3**, 19 (1965).
- [11] D. Choi and K. S. Jeong, *J. Phys. Chem. Lett.* **10**, 4303 (2019).
- [12] M. Lindberg and S. W. Koch, *Phys. Rev. B* **38**, 3342 (1988).
- [13] M. Kira and S. W. Koch, *Semiconductor Quantum Optics* (Cambridge University Press, Cambridge, 2012).
- [14] S. Ghimire, A. D. DiChiara, E. Sistrunk, P. Agostini, L. F. DiMauro, and D. A. Reis, *Nat. Phys.* **7**, 138 (2011).
- [15] T. T. Luu, M. Garg, S. Y. Kruchinin, A. Moulet, M. T. Hassan, and E. Goulielmakis, *Nature (London)* **521**, 498 (2015).
- [16] N. Yoshikawa, T. Tamaya, and K. Tanaka, *Science* **356**, 736 (2017).
- [17] G. Vampa, T. Hammond, N. Thiré, B. Schmidt, F. Légaré, C. McDonald, T. Brabec, and P. Corkum, *Nature (London)* **522**, 462 (2015).
- [18] M. Hohenleutner, F. Langer, O. Schubert, M. Knorr, U. Huttner, S. W. Koch, M. Kira, and R. Huber, *Nature (London)* **523**, 572 (2015).
- [19] G. Ndabashimiye, S. Ghimire, M. Wu, D. A. Browne, K. J. Schafer, M. B. Gaarde, and D. A. Reis, *Nature (London)* **534**, 520 (2016).
- [20] H. Liu, Y. Li, Y. S. You, S. Ghimire, T. F. Heinz, and D. A. Reis, *Nat. Phys.* **13**, 262 (2017).
- [21] Y. S. You, D. A. Reis, and S. Ghimire, *Nat. Phys.* **13**, 345 (2017).
- [22] P. Xia, C. Kim, F. Lu, T. Kanai, H. Akiyama, J. Itatani, and N. Ishii, *Opt. Express* **26**, 29393 (2018).
- [23] A. Kemper, B. Moritz, J. Freericks, and T. Devereaux, *New J. Phys.* **15**, 023003 (2013).
- [24] G. Vampa, C. R. McDonald, G. Orlando, D. D. Klug, P. B. Corkum, and T. Brabec, *Phys. Rev. Lett.* **113**, 073901 (2014).
- [25] P. G. Hawkins, M. Y. Ivanov, and V. S. Yakovlev, *Phys. Rev. A* **91**, 013405 (2015).
- [26] T. Higuchi, M. I. Stockman, and P. Hommelhoff, *Phys. Rev. Lett.* **113**, 213901 (2014).
- [27] T. Tamaya, A. Ishikawa, T. Ogawa, and K. Tanaka, *Phys. Rev. Lett.* **116**, 016601 (2016).
- [28] M. Wu, S. Ghimire, D. A. Reis, K. J. Schafer, and M. B. Gaarde, *Phys. Rev. A* **91**, 043839 (2015).
- [29] S. Ghimire, A. D. DiChiara, E. Sistrunk, G. Ndabashimiye, U. B. Szafruga, A. Mohammad, P. Agostini, L. F. DiMauro, and D. A. Reis, *Phys. Rev. A* **85**, 043836 (2012).
- [30] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [31] G. Kresse and J. Hafner, *Phys. Rev. B* **49**, 14251 (1994).
- [32] G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- [33] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [34] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [35] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [36] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [37] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [38] K.-H. Xue, J.-H. Yuan, L. R. Fonseca, and X.-S. Miao, *Comput. Mater. Sci.* **153**, 493 (2018).
- [39] L. G. Ferreira, M. Marques, and L. K. Teles, *Phys. Rev. B* **78**, 125116 (2008).
- [40] L. G. Ferreira, M. Marques, and L. K. Teles, *AIP Adv.* **1**, 032119 (2011).
- [41] J. C. Slater and K. H. Johnson, *Phys. Rev. B* **5**, 844 (1972).
- [42] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* **73**, 045112 (2006).
- [43] A. Girndt, F. Jahnke, A. Knorr, S. W. Koch, and W. W. Chow, *Phys. Stat. Sol. (b)* **202**, 725 (1997).
- [44] See <http://www.nlcstr.com/publications.htm> and <http://www.nlcstr.com/examples0.htm> for examples.
- [45] C. Adenis, V. Langer, and O. Lindqvist, *Acta Crystallogr. Sect. C* **45**, 941 (1989).
- [46] V. B. Anzin, M. I. Eremets, Y. V. Kosichkin, A. I. Nadezhdinskii, and A. M. Shirokov, *Phys. Status Solidi A* **42**, 385 (1977).
- [47] C. Rigaux and G. Drilhon, *J. Phys. Soc. Jpn. Suppl.* **21**, 193 (1966).
- [48] P. Grosse and K. Winzer, *Phys. Stat. Sol.* **26**, 139 (1968).
- [49] J. Hader, J. V. Moloney, and S. W. Koch, *Appl. Phys. Lett.* **87**, 201112 (2005).
- [50] U. Huttner, M. Kira, and S. W. Koch, *Laser Photonics Rev.* **11**, 1700049 (2017).
- [51] S. Jiang, H. Wei, J. Chen, C. Yu, R. Lu, and C. D. Lin, *Phys. Rev. A* **96**, 053850 (2017).
- [52] L. H. Thong, C. Ngo, H. T. Duc, X. Song, and T. Meier, *Phys. Rev. B* **103**, 085201 (2021).
- [53] A. Chacón, D. Kim, W. Zhu, S. P. Kelly, A. Dauphin, E. Pisanty, A. S. Maxwell, A. Picón, M. F. Ciappina, D. E. Kim, C. Ticknor, A. Saxena, and M. Lewenstein, *Phys. Rev. B* **102**, 134115 (2020).
- [54] S. Jiang, J. Chen, H. Wei, C. Yu, R. Lu, and C. D. Lin, *Phys. Rev. Lett.* **120**, 253201 (2018).
- [55] I. Kilen, M. Kolesik, J. Hader, J. V. Moloney, U. Huttner, M. K. Hagen, and S. W. Koch, *Phys. Rev. Lett.* **125**, 083901 (2020).
- [56] S. Adachi, *The Handbook on Optical Constants of Semiconductors* (World Scientific, Singapore, 2012).
- [57] S. Zhe, R. Cao, K. Khan, A. K. Tareen, X. Liu, W. Liang, Y. Zhang, C. Ma, Z. Guo, X. Luo, and H. Zhang, *Nano-Micro Lett.* **12**, 99 (2020).

Paper IV

J. Hader, **S. C. Liebscher**, J. V. Moloney, and S. W. Koch

“Intrinsic Carrier Losses in Tellurium Due to Radiative and Auger Recombinations”
arXiv:2208.05532 [cond-mat.mtrl-sci] (2022) **DOI:** 10.48550/ARXIV.2208.05532

submitted to Applied Physics Letters on 27-Jul-2022

Intrinsic Carrier Losses in Tellurium Due to Radiative and Auger Recombinations

Jörg Hader,^{1, a)} Sven C. Liebscher,² Jerome V. Moloney,¹ and Stephan W. Koch²

¹⁾*Wyant College of Optical Sciences, University of Arizona, 1630 E. University Blvd., Tucson, Arizona 85721, USA*

²⁾*Department of Physics and Material Sciences Center, Philipps-University Marburg, Renthof 5, 35032 Marburg, Germany*

(Dated: 27 July 2022)

Fully microscopic many-body models based on inputs from first principle density functional theory are used to calculate the carrier losses due to radiative- and Auger-recombinations in bulk tellurium. It is shown that Auger processes dominate the losses for carrier densities in the range typical for applications as lasers. The Auger loss depends crucially on the energetic position of the H_6 valence bands. At cryogenic temperatures of 50 K (100 K) the Auger coefficient, C , varies by about six (three) orders of magnitude within the range of published distances between these bands and the valence bandedge. Values for C at the high and low end of these ranges are found if the distance is smaller or larger than the bandgap, respectively. At room temperature the sensitivity is reduced to about a factor of four with C values ranging between 0.4 and $1.6 \times 10^{-27} \text{cm}^6 \text{s}^{-1}$. Here, radiative losses dominate for carrier densities up to about $10^{16} / \text{cm}^3$ with a loss coefficient $B \approx 10^{-11} \text{cm}^3 \text{s}^{-1}$. The radiative losses are about two to three times lower than in typical bulk III-V materials for comparable wavelengths.

Triagonal tellurium (t-Te) combines the simplicity of a single elemental material with a wide variety of technically interesting properties. Its distinctive structure made of weakly interacting helical chains of atoms and triagonal symmetry perpendicular to the axis of the chains, c , lead to a strong anisotropy of its characteristics like an exceptionally strong optical nonlinearity¹ or an extraordinarily high piezoelectric coefficient². Recently, the interest in tellurium has intensified as it was shown that it transforms into a topological insulator under pressure³ and it has become possible to synthesize it in the two-dimensional form tellurene and as one-dimensional wires^{4,5}.

t-Te is a semiconductor with a bandgap in the mid-infrared wavelength range near $3.7 - 3.8 \mu\text{m}$. Reaching these wavelengths using conventional III-V materials requires complex heterostructures like quantum cascade systems, or involve complex mixtures of multiple binary materials, like quinternary AlInGaAsSb, which are difficult to grow in a controlled manner⁶. The size of this bandgap combined with a high conductivity make the material ideal for applications as a (chiral) thermoelectric material^{7,8}.

While t-Te has an indirect gap due to a camel-back structure of the highest valence band, the dip in the valence band is only of the order of 1meV ^{7,9} and the material behaves like a direct semiconductor under common conditions. This makes t-Te a candidate for optoelectronic applications like photo-detectors, solar cells or light emitters¹⁰. To be successful for these applications requires good crystal quality, strong optical coupling and low carrier losses. The atomic simplicity of t-Te enables

high crystal quality and measured¹¹ and calculated¹² absorption spectra have confirmed strong optical coupling at the bandgap. The latter was found very similar to that in bulk III-V materials with a similar bandgap, like InAs¹³ or dilute InAsBi¹⁴.

The high crystal quality of t-Te allows for low defect related carrier losses¹⁵. Carrier losses in t-Te due to radiative recombinations have recently been measured in Refs.¹⁵ and Ref.¹⁶. The works determined recombination coefficients, B , of $1.1 \times 10^{-8} \text{cm}^3 / \text{s}$ and $1.9 \times 10^{-9} \text{cm}^3 / \text{s}$, respectively. Both values are significantly larger than the commonly accepted values in direct gap bulk III-V materials like GaAs¹⁷⁻²⁰, GaSb²¹, InP²¹, InSb²¹, or InAs²¹. For these, B has been determined to be in the range of $0.2-2 \times 10^{-10} \text{cm}^3 / \text{s}$. Generally, the radiative recombination due to spontaneous emission should scale like the near bandgap absorption since in good approximation the two can be directly related through the Kubo-Martin-Schwinger (KMS) relation²². Since the absorption in t-Te is of similar strength as in these other materials, the radiative lifetimes should not be expected to be significantly different either.

The large magnitude and spread of the values for B in Refs.^{15,16} are likely due to the fact that the material is inherently p-doped. The doping arises due to the fact that the chiral chains are finite and unsaturated bonds occur at their ends. The exact level of doping is generally not known and sample-dependent. The intrinsic radiative lifetime that assumes equal electron and hole carrier densities can be vastly different from the minority carrier lifetime in doped samples and the latter strongly depends on the dopant level²³.

Carrier losses due to Auger recombinations are known to increase dramatically with decreasing bandgap²⁵. For materials with bandgaps similar to t-Te, like InAs or InSb, Auger rates have been found to be about three

^{a)} Author to whom correspondence should be addressed: jhader@acms.arizona.edu

orders of magnitude higher than in materials with a bandgap around one micron, like GaAs (Auger coefficients, C around $10^{-27} \text{ cm}^6/\text{s}$ versus $10^{-30} \text{ cm}^6/\text{s}$ ^{14,26,27}). This makes Auger loss the limiting factor for mid-IR materials in applications as optoelectronic devices. To the best of our knowledge, Auger rates have neither been calculated nor measured for t-Te. Generally, measuring the losses experimentally is difficult since only the total non-radiative recombination rate is accessible, but the Auger contribution cannot be measured individually.

Knowing the intrinsic recombination rates allows to judge the prospects of t-Te for applications. Here we use fully microscopic many-body models based on inputs from first principle density functional theory (DFT) in order to calculate the intrinsic carrier losses due to radiative- and Auger-recombination. These models have been shown to yield excellent agreement with the experiment for a wide variety of materials with wavelengths ranging from the mid-IR to the ultra-violet²⁴. The *a-priori* nature of the modeling approach eliminates fit parameters that would otherwise be required to be extracted from experiments, like linewidth broadenings or dephasing times. Elimination of such parameters, combined with the demonstrated accuracy of the results makes this approach quantitatively predictive.

The carrier lifetimes due to spontaneous emission processes, also known as bimolecular recombination lifetimes or, short, as radiative carrier lifetimes, are calculated using the semiconductor luminescence equations (SLE)²⁸. These are the equations of motion for microscopic photon-assisted polarizations. They explicitly contain carrier-carrier and carrier-phonon scatterings that lead to the dephasing of the polarizations and influence photoluminescence (PL) lineshapes, spectral positions and amplitudes. The corresponding terms also enter the SLE as part of higher order correlations that are sources for the photon-assisted polarizations. These source terms are absent in the equations of motion for the optical polarizations that describe absorption/gain spectra. This absence leads to an error when using the KMS relation to derive the PL from the absorption that can easily be on the order of a factor of two or more. The PL spectrum is obtained from the total photon assisted polarization via Fourier transformation and the radiative lifetime is obtained by integrating over the spectrum²⁹.

Like the scattering terms involved in the SLE, Auger losses are calculated in second-Born and Markov approximation by solving quantum-Boltzmann type scattering equations³⁰. The equations that we solve for the Auger losses are basically the same as derived decades earlier³¹. However, unlike earlier attempts to solve these equations, we do not use any uncontrolled approximations like using Boltzmann instead of Fermi distributions, approximating the Coulomb coupling matrix elements or considering only occupations near zero momentum or at 0K. These approximations lead to errors that can easily exceed one order of magnitude and are to a large degree the reason why values for Auger losses in the literature

vary for many materials by such large margins. As mentioned, our calculations have been tested against the experiment for a wide variety of materials, temperature- and density-ranges and have shown uncertainties in the few ten-percent range^{24,28}.

We assume equal electron and hole carrier densities N and all Auger calculations are for direct inter-band processes which are dominant in materials for this wavelength range³². The bimolecular (radiative) recombination coefficients, B , and Auger coefficients, C are derived from the corresponding recombination times, τ_B and τ_C , through:

$$B = \frac{1}{\tau_B N}, \quad C = \frac{1}{\tau_C N^2}. \quad (1)$$

The band energies and wavefunctions that enter the SLE and Auger equations are calculated using the Vienna Ab Initio Simulation Package³³⁻³⁶ (VASP) with the Projector-Augmented Wave (PAW) method^{37,38}. The structure was relaxed using the Generalized Gradient Approximation (GGA) by Perdew, Burke and Ernzerhof (PBE)³⁹ for the exchange-correlation energy. The PAW pseudopotential was modified according to the sHLDA-1/2 method as proposed in Ref.⁴⁰. More details of the DFT calculation can be found in Ref.¹².

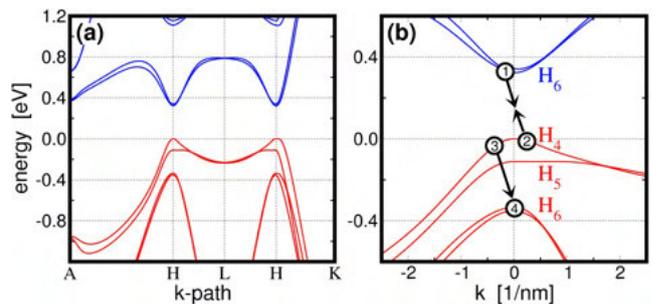


FIG. 1. (a): Lowest four electron (blue) and highest hole (red) bands along some high symmetry directions of the Brillouin zone of t-Te as calculated with DFT. (b): Zoom into the region around H along the A-H-L path with the schematic of an Auger recombination of an electron (1) with a hole (2) where the excess energy is being given to another hole (3) which is excited into a state (4) deep in the valence band.

Fig.1 shows the results of the bandstructure calculation for t-Te along some high symmetry directions of the Brillouin zone (BZ). Like Ref.⁷, we find a small camel-back shape at the top of the valence band edge with a global maximum slightly away from the H -point and about 1 meV above the energy at H . At 0K the bandgap at the H -point is $E_0 = 333 \text{ meV}$ and the splitting between the highest two valence bands $E_{H_4-H_5} = 111 \text{ meV}$. The splitting between the highest and third highest valence bands $E_{H_4-H_6} = 337 \text{ meV}$. The values for the gap and $E_{H_4-H_5}$ agree very well with experimentally measured values⁴¹⁻⁴³ and other calculations^{3,7,44}. However, there is some uncertainty in the literature for $E_{H_4-H_6}$. While bandstruc-

ture shown in Ref.⁷ agrees with our value, Ref.⁴³ extracted a value of 427 meV from experiments and Ref.⁴⁴ calculates a splitting of about 415 meV. The error margin of the measured value in Ref.⁴³ is not clear, but could be on the order of several tens of meV.

As we will show below, the exact value for $E_{H_4-H_6}$ has a significant impact on the Auger losses. At the low end of the range of published values the splitting is almost exactly equal to the bandgap. This is an ideal situation for Auger transitions since it allows holes to be excited from the H_4 to the H_6 band with minimal momentum transfer as indicated in the schematic in Fig.1 (b). On the other hand, for the larger values, the band separation is larger than the bandgap and carriers near the gap are not able to reach the H_6 bands. Here, the only possible transitions are within the H_4 and H_5 bands or within the H_6 electron band. These require large momentum transfers to reach states one bandgap away from the bandedge and are therefore unfavorable since the Coulomb coupling decreases with increasing momentum transfer.

To compare results for the two different splittings we apply for the larger splitting case a momentum-independent shift to the H_6 hole bands from our calculation. All other bands and all wavefunctions are kept the same.

The DFT calculations are performed for a temperature of 0 K. We include the temperature dependence of the bandstructure by applying a shift to the conduction bands according to the formula for the bandgap as derived from the experiment in Rev.⁴³. This leads to a gap that has a maximum of 335 meV around 75 K, is about 333 meV at 0 K and 327 meV at 300 K. For temperatures above 300 K the bandgap shrinks by about 0.052 meV/K.

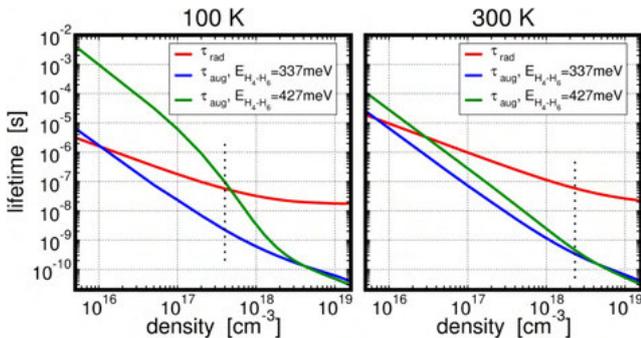


FIG. 2. Calculated carrier lifetimes due to radiative (red) and Auger recombinations as function of the carrier density at 100 K (left) and 300 K (right). Blue (green): Auger lifetimes for the case that the hole splitting $E_{H_4-H_6}$ is similar to (337 meV) (larger than (427 meV)) the bandgap. Dotted lines mark the approximate carrier density at which the absorption at the bandgap is bleached and optical gain emerges.

Fig.2 shows the calculated carrier lifetimes due to radiative and Auger recombinations as function of the carrier density. Auger losses are shown for two cases of the hole splitting $E_{H_4-H_6}$, the case where it is almost res-

onant with the bandgap (337 meV) and the case where it is about 95 meV larger than the bandgap (427 meV). At 300 K Auger losses dominate already at densities on the order of $10^{17}/\text{cm}^3$ which is more than one order of magnitude below the transparency density at that temperature. At 300 K the Auger losses at low densities are about four times smaller if $E_{H_4-H_6}$ is larger than the gap rather than when $E_{H_4-H_6}$ and the gap are almost equal. At 100 K this difference is close to three orders of magnitude.

At the higher temperature the larger high energy tail of the Fermi distributions leads a larger percentage of holes at energies far enough below the bandedge to reach the H_6 bands even for the larger splitting. This increases the Auger loss for the large splitting case and makes it more similar to the small splitting case. Similarly, the sensitivity to the splitting decreases with increasing density since an increasing fraction of carriers occupies states sufficiently far away from the gap to reach the H_6 bands.

In the limit of very high densities the Auger losses for the larger splitting start to exceed somewhat those for the smaller splitting. This is due to the fact that for the larger splitting at these densities more carriers can reach final states with less momentum transfer and larger Auger coupling.

The radiative lifetimes are virtually identical for both values for $E_{H_4-H_6}$. This is due to the fact that for both cases of the splitting virtually no carriers occupy the H_6 hole bands even at the highest temperatures and densities considered here. Non-zero occupations occur only for the H_4 and H_5 hole bands and these occupations are not affected by the position of the H_6 band. Since the spontaneous emission only involves occupied states it is unaffected by the splitting and so are the resulting lifetimes.

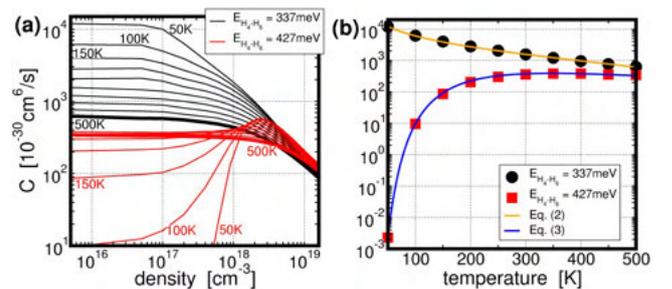


FIG. 3. (a) Auger coefficient as function of the carrier density for temperatures between 50 K and 500 K in 50 K intervals. The result for 500 K is marked with a bold line. Black (Red): for a hole splitting $E_{H_4-H_6}=337$ meV (427 meV). (b) Auger coefficients in the low density limit as function of the temperature. Circles (Squares): calculation results for $E_{H_4-H_6}=337$ meV (427 meV). Orange (blue) Line: fit according to Eq. (2) (Eq. (3)).

Fig.3 shows the Auger coefficients as function of the density for various temperatures. In the limit of low den-

sities ($N \lesssim 10^{17}/\text{cm}^3$) one finds the traditional quadratic density dependence of the lifetimes and, thus, density independent coefficients C . In the high density limit ($N \gtrsim 5 \times 10^{18}/\text{cm}^3$) the density dependence is reduced due to phase space filling⁴⁵. At low densities the occupation probabilities increase approximately linearly with the density leading to the quadratic density dependence of the lifetimes. At high densities Pauli-blocking limits a further increase of the occupations and additional carrier have to be filled instead into states at higher energies. Here, this leads to a reduction of the density dependence of the Auger lifetimes from quadratic to about linear and, correspondingly, a nearly linear decrease of $C(N)$.

The values found here for C for the case where $E_{H_4-H_6}$ is larger than the bandgap are very similar to those in InAs. Like t-Te, InAs has a bandgap in the 3-4 μm range (356 meV at 300 K). Its spin-orbit hole splitting is larger than the bandgap and of similar magnitude as the $H_4 - H_6$ -splitting here (410 meV). Its Auger coefficient at room temperature has been found to be around $10^{-27} \text{cm}^6/\text{s}^{14}$.

Fig.3 (b) shows the temperature dependence of the Auger coefficient in the low density limit. For the case where $E_{H_4-H_6}$ is about equal with the gap C decreases with increasing temperature. This is mostly due to the fact that the occupation probabilities near the bandgap decrease with increasing temperature. The dependence can be fitted using the expression:

$$C(T) = \frac{7.4 \times 10^{-25} - 8.6 \times 10^{-28} T}{T + 8} [\text{cm}^6/\text{s}], \quad (2)$$

where T is the temperature in Kelvin.

The low density $C(T)$ increases dramatically at low temperatures if $E_{H_4-H_6}$ is larger than the bandgap. In the cryogenic limit, no carriers are at energies close enough to reach the H_6 band and transitions within the near-gap bands are unfavorable since they require large momentum transfer. This leads to very low Auger rates. With rising temperature an increasing fraction of holes occupies states close enough to the H_6 -band to allow for Auger transitions to it. This leads to a strong increase of $C(T)$ for temperatures up to about room temperature. For even higher temperatures the trend seen for the case where the splitting is smaller than the gap starts to set in and $C(T)$ starts to decrease. The behavior for low temperatures can be described by an activation energy law with an energy similar to the difference between the splitting and the gap. The decrease at higher temperatures leads to a modification of the dependence that can be modeled by an additional $1/T^3$ -scaling:

$$C(T) = 3.3^{-19} \times \exp\left(\frac{-E_a}{k_B T}\right) T^{-3} [\text{cm}^6/\text{s}], \quad (3)$$

where k_B is the Boltzmann constant and $E_a=0.09 \text{eV}$.

Fig.4 shows the calculated bimolecular recombination coefficient B as function of the carrier density and for various temperatures. As mentioned above, the $H_4 - H_6$

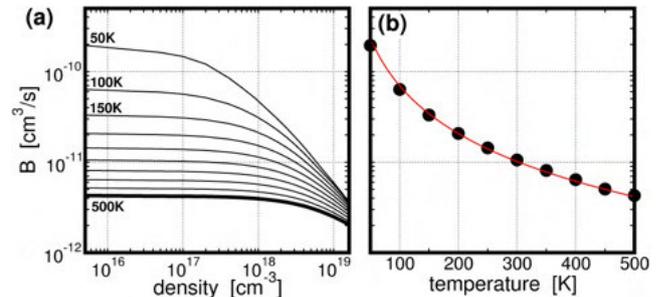


FIG. 4. (a) Bimolecular recombination coefficient, B as function of the carrier density for temperatures between 50 K and 500 K in 50 K intervals. (b) Coefficients B in the low density limit as function of the temperature. Circles show calculation results. The red line is a fit according to Eq. (4).

hole splitting has virtually no influence on the radiative recombination rate. Within the parameter range investigated here we find the differences between results for the two splittings on the order of 2×10^{-4} and indistinguishable on the scale of Fig.4.

In the limit of low carrier densities we find $B(77 \text{K}) \approx 1.1 \times 10^{-10} \text{cm}^3/\text{s}$ and $B(300 \text{K}) \approx 1.1 \times 10^{-11} \text{cm}^3/\text{s}$. The temperature dependence of the low density value of B as shown in Fig.4 can be fitted using:

$$B(T) = 2.2 \times 10^{-7} T^{-7/4} [\text{cm}^3/\text{s}], \quad (4)$$

As for Auger recombination, the bimolecular recombination rates found here are similar to those reported for III-V bulk materials at similar wavelengths. Ref.⁴⁶ measured a coefficient B of $3 \times 10^{-10} \text{cm}^3/\text{s}$ at 77 K in InAsSb. Using SimuLaseTM software⁴⁷, which implements the same level of microscopic many-body physics as used here, we calculate for bulk InAs $B(77 \text{K}) \approx 2.0 \times 10^{-10} \text{cm}^3/\text{s}$ and $B(300 \text{K}) \approx 2.5 \times 10^{-11} \text{cm}^3/\text{s}$.

The radiative lifetimes in those III-V materials are about two to three times longer than what we find for t-Te. In part this is likely due to the fact that the optical coupling for polarization parallel to the c-axis is symmetry forbidden at the bandgap of t-Te. This effect which is reflected in the suppressed near-bandgap absorption of t-Te^{11,12} also leads to a limited spontaneous emission into that direction¹².

Like the Auger coefficients, the bimolecular recombination coefficients decrease with density due to phase-space filling⁴⁵. In the high density limit the coefficients tend toward a $1/N$ density dependence.

In summary, we have used fully microscopic many-body models based on input from first-principle DFT calculations to study intrinsic carrier losses in bulk trigonal tellurium. Assuming that the splitting between the H_4 and H_6 valence bands is as in our calculations about equal to the bandgap we find Auger recombination losses of similar strength as in bulk III-V materials for similar wavelengths, like InAs. Auger coefficients of about

10^{-27} cm⁶/s at 300 K make it unlikely that the material is viable for laser applications. However, the Auger losses could be dramatically lower if the splitting is larger than the bandgap as suggested in some other publications. Independent of the $H_4 - H_6$ splitting, the radiative losses are about a factor three lower than in III-V materials for these wavelengths. This means that t-Te could yield better performance for low-density applications like photo detectors or solar cells.

The authors thank the HRZ Marburg and CSC-Goethe-HLR Frankfurt for computational resources. The Tucson work was supported by the Air Force Office of Scientific Research under award numbers FA9550-19-1-0032 and FA9550-21-1-0463.

DATA AVAILABILITY

The data that supports the findings of this study are available within the article.

ORCID IDS

Jörg Hader <https://orcid.org/0000-0003-1760-3652>

Sven C. Liebscher <https://orcid.org/0000-0003-3505-1521>

Jerome V. Moloney <https://orcid.org/0000-0001-8866-0326>

Stephan W. Koch <https://orcid.org/0000-0001-5473-0170>

- ¹C. K. N. Patel, Phys. Rev. Lett. **15**, 1027 (1965).
- ²T.I. Lee, S. Lee, E. Lee, S. Sohn, Y. Lee, S. Lee, G. Moon, D. Kim, Y.S. Kim, J.M. Myoung, and Z.L. Wang, Adv. Mater. **25**, 2920 (2013).
- ³L.A. Agapito, N. Kioussis, W.A. Goddard III, and N.P. Ong, Phys. Rev. Lett. **110**, 176401 (2013).
- ⁴J. Qiao, Y. Pan, F. Yang, C. Wang, Y. Chai, and W. Ji, Sci. Bulletin **63** 159 (2018).
- ⁵G. Qiu, A. Charnas, C. Niu, Y. Wang, W. Wu, and P.D. Ye, npj 2D Mater. Appl. **6**, 17 (2022).
- ⁶E. Tournié, and L. Cerutti, *Mid-infrared Optoelectronics; Materials, Devices, and Applications*, (Woodhead Publishing Series in Electronic and Optical Materials, Woodhead Publishing, Elsevier, Cambridge, 2019).
- ⁷H. Peng, N. Kioussis, and G.J. Snyder, Phys. Rev. B **89**, 195206 (2014).
- ⁸S. Lin, W. Li, Z. Chen, J. Shen, B. Ge, and Y. Pei, Nat. Commun. **7**, 10287 (2016).
- ⁹T. Doi, K. Nakao, and H. Kamimura, J. Phys. Soc. Jpn. **28**, 822 (1970).
- ¹⁰A. Das, and B.K. Banik, "Semiconductor characteristics of tellurium and its implementations," Phys. Sci. Rev., published online May 18, 2022.
- ¹¹S. Tutihasi, G. G. Roberts, R. C. Keezer, and R. E. Drews, Phys. Rev. **177**, 1143 (1969).
- ¹²S.C Liebscher, M.K. Hagen, J. Hader, J.V. Moloney, and S.W. Koch, Phys. Rev. B **104** 165201 (2021).
- ¹³P. T. Webster, A. J. Shalindar, N. A. Riordan, C. Gogineni, H. Liang, A.R. Sharma, and S. R. Johnson, J. Appl. Phys. **119**, 225701 (2016)
- ¹⁴J. Hader, S.C. Badescu, L.C. Bannow, J.V. Moloney, S.R. Johnson, and S.W. Koch, Appl. Phys. Lett. **112**, 192106 (2018).
- ¹⁵P. Bhaskar, A.W. Achstein, M.J.W. Vermeulen, L.D.A. Sibbeles, J. Phys. Chem. C **123**, 841 (2019).
- ¹⁶V. Iyer, M. Segovia, Y. Wang, W. Wu, P. Ye, and X. Xu, Phys. Rev. B **100**, 075436 (2019).
- ¹⁷G.W. 't Hooft, Appl. Phys. Lett. **39**, 389 (1981).
- ¹⁸S. Strauss, W.W. Rühle, and K. Köhler, Appl. Phys. Lett. **62**, 55 (1993).
- ¹⁹F.K. Reinhart, J. Appl. Phys. **97** 123536 (2005).
- ²⁰M. Niemeyer, P. Kleinschmidt, A. W. Walker, L. E. Mundt, C. Timm, R. Lang, T. Hannappel, and D. Lackner, AIP Advances **9**, 045034 (2019).
- ²¹M. Levinshtein, S. Rumyantsev, and M. Shur, *Handbook Series on Semiconductor Parameters*, (World Scientific, Singapore, 2000).
- ²²W. Chow, M. Kira, and S.W. Koch, Phys. Rev. B **60**, 1947 (1999).
- ²³P.Y. Yu, and M. Cardona *Fundamentals of Semiconductors, Physics and Material Properties*, (Fourth Ed., Springer, Heidelberg, 2010).
- ²⁴See <http://www.nlcstr.com/publications.htm> and <http://www.nlcstr.com/examples0.htm> for examples.
- ²⁵W.W. Bewley, J.R. Lindle, C.S. Kim, M. Kim, C.L. Canedy, I. Vurgaftman, and J.R. Meyer, Appl. Phys. Lett. **93**, 041118 (2008).
- ²⁶[Landolt-Börnstein - Group III Condensed Matter] *Group IV Elements, IV-IV and III-V Compounds. Part b - Electronic, Transport, Optical and Other Properties Volume b - Indium arsenide (InAs), recombination and generation, lifetimes*, ed. O. Madelung, U. Rössler, and M. Schulz (Springer, Berlin, 2002).
- ²⁷[Landolt-Börnstein - Group III Condensed Matter] *Data extract from Landolt-Börnstein III/44A: Semiconductors - New Data and Updates for I-VII, III-V, III-VI and IV-VI Compounds*, ed. U. Rössler (Springer, Berlin, 2009).
- ²⁸M. Kira, F. Jahnke, W. Hoyer, and S. W. Koch, Prog. Quantum Electron. **23**, 189 (1999).
- ²⁹W. W. Chow and S. W. Koch, *Semiconductor-Laser Fundamentals: Physics of the Gain Materials*, (Springer, Berlin 1999).
- ³⁰J. Hader, J.V. Moloney, and S.W. Koch, IEEE J. Quantum Electron. **41**, 1217 (2005).
- ³¹A. R. Beattie and P. T. Landsberg, Proc. Royal Soc. London Series A, **249**, 16 (1958).
- ³²J.-X. Shen, D. Steinauf, A. McAllister, G. Shi, E. Kioupakis, A. Janotti, and G. Van de Walle, Phys. Rev. B **100**, 155202 (2019).
- ³³G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
- ³⁴G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
- ³⁵G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
- ³⁶G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
- ³⁷G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
- ³⁸P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
- ³⁹J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
- ⁴⁰K.-H. Xue, J.-H. Yuan, L. R. Fonseca, and X.-S. Miao, Comput. Mater. Sci. **153**, 493 (2018).
- ⁴¹R.S. Caldwell and H.Y. Fan, Phys. Rev. **114**, 664 (1959).
- ⁴²V.B. Anzin, M.I. Eremets, Y.V. Koshichkin, A.I. Nadezhdinskii, and A. M. Shirokov, Phys. Status Solidi a **42**, 385 (1977).
- ⁴³D.-Q. Yang, L.-Q. Zhu, J.L. Wang, W. Xia, J.-Z. Zhang, K. Jiang, L.-Y. Shang, Y.-W. Li, and Z.-G. Hu, Phjys. Status Solidi b **2022**, 2100625 (2022).
- ⁴⁴J. Qiao, Y. Pan, F. Yang, C. Wang, and W. Ji, Sci. Bulletin **63**, 159 (2018).
- ⁴⁵J. Hader, J.V. Moloney, and S.W. Koch, Appl. Phys. Lett., **87** (20) 201112 (2005).
- ⁴⁶B. V. Olson, E. A. Shaner, J. K. Kim, J. F. Klem, S. D. Hawkins, L. M. Murray, J. P. Prineas, M. E. Flatté, and T. F. Boggess, Appl. Phys. Lett. **101** 092109 (2012).
- ⁴⁷*SimuLaseTM*, AllnGaAsSb, Version 3.0 by Nonlinear Control Strategies Inc., Tucson, AZ (2022).

