



DTU Physics
Department of Physics

Ph.D. Thesis

Light-Matter Interactions from First Principles

Mark Kamper Svendsen, September 2022



DTU Physics
Department of Physics

Ph.D. Thesis

Light-Matter Interactions from First Principles

Mark Kamper Svendsen

Kongens Lyngby, 2022

Light-Matter Interactions from First Principles

Ph.D. Thesis
Mark Kamper Svendsen
September 30th, 2022

Copyright: Reproduction of this publication in whole or in part must include the customary bibliographic citation, including author attribution, thesis title, etc.

Cover: Extinction efficiency map of the magnetic dipole resonance as a function of radius calculated for spherical nano-particles of Boron Phosphide. See Fig. 2c in Publication [II].

Technical University of Denmark
DTU Physics
Department of Physics

Fysikvej 311
2800 Kongens Lyngby, Denmark
Phone: +45 4525 3344
info@fysik.dtu.dk

www.fysik.dtu.dk

Preface

This thesis is submitted in candidacy for a Ph.D. degree in Physics from the Technical University of Denmark (DTU). The Ph.D project was carried out at the Center for Atomic-scale Materials Design (CAMD) at the DTU Department of Physics between October 2019 and September 2022. The project was supervised by Main Supervisor Professor Kristian Sommer Thygesen and Co-Supervisor Professor Jesper Mørk.

The thesis also includes work conducted during an external research stay hosted by Assistant Professor Johannes Flick and Professor Angel Rubio at the Flatiron Institute in New York from February to June 2022. I want to extend a sincere thank you to the Simons Foundations for the financial support of this external research stay

Mark Kamper Svendsen

Mark Kamper Svendsen
Kongens Lyngby, September 30th, 2022

Abstract

Our understanding of light-matter interactions enables a wide range of the technologies that define modern life. Examples include lasers, solar cells and light emitting diodes, all of which are likely to play a key role in our transition towards a more sustainable society. In this thesis, light-matter interactions are studied using first-principles quantum mechanical calculations. It is demonstrated how such calculations can be used to both interpret novel experimental results and actively drive materials discovery for opto-electronic applications. The thesis both presents new methodologies, as well as novel applications of existing methodologies.

The quantum nature of light has traditionally been neglected in the context of condensed matter physics and quantum chemistry. However, it can become important when the electromagnetic environment is structured in such a way as to host resonant or confined modes of light. A particularly interesting version of this is how resonant electromagnetic environments can be used to alter the electronic structure of materials, even in the absence of actual light. This happens as a result of the altered quantum fluctuations of the electromagnetic field, and the effect is directly related to the fundamental quantum nature of both light and matter. This thesis presents a novel methodology combining Macroscopic Quantum Electrodynamics with different flavors of Density Functional Theory. These results represent a step towards a fully first-principles description of quantum light-matter interactions in arbitrary electromagnetic environments.

Cavity-QED with color centers in 2D hexagonal Boron Nitride is also investigated. This investigation is carried out using a combination of the tensor network based TEMPO algorithm for time propagation of open quantum systems and a first-principles description of the electron-phonon coupling. This investigation shows that the fine structure of the electron-phonon coupling's spectral density manifests as clear signatures of non-Markovianity in the dynamics of the coupled cavity-defect system. Furthermore, clear signs of hybridisation between the light-matter polaritons and the phonon modes of the environment are found.

Another focus of this thesis is materials discovery for nano-photonic- and opto-electronic applications. Using a combination of high-throughput computational screening, accurate many-body perturbation theory, experimental synthesis as well as near- and far-field optical characterization, Boron-Phosphide is identified as an overlooked high-performance material for near-UV dielectric nano-photonics.

The thesis also addresses the potential of indirect band gap semiconductors for photovoltaic applications. Combining a computationally efficient approximation of phonon-assisted indirect absorption with high-throughput computational screening, 9 new com-

pounds with potential use in future thin-film solar cells are identified.

New results from the Computational 2D Materials Database (C2DB) project are also presented. Comparison with full phonon calculations validates the Center and Boundary Protocol as an efficient tool to evaluate dynamical stability of relaxed 2D atomic structures. Furthermore, it is shown that the protocol can be iteratively applied to generate dynamically stable structures from unstable initial structures via systematic dislocations of the lattice. Finally, the existing data in the C2DB is harnessed to train a machine learning model for dynamical stability prediction. This model performs very well, showing an excellent receiver operating characteristic (ROC) curve with an area under the curve (AUC) of 0.9. The results show that the integration of such classification models into the workflows can drastically reduce computational efforts in high-throughput studies.

Furthermore, the elementary electronic excitations of layered PtSe₂ were unravelled using both momentum-resolved Electron Energy Loss Spectroscopy and first-principles calculations. Additionally, it was shown how to turn theoretical considerations about the momentum-dependence of the Coulomb interaction matrix elements into actionable spectroscopic insight for 2D materials. This could potentially lead to new possibilities in spectroscopy.

Finally, the thesis presents results from a large collaboration seeking to use van der Waals (vdW) heterostructures under illumination of electrons from a Transition Electron Microscope as a source of energy tunable X-ray radiation. It is shown that the energy of the emission can be tuned both via the incoming electron energy and the configuration of the heterostructure. This opens the door to highly tunable sources of X-ray radiation with very small physical footprints compared to existing technologies.

Resumé

Vores forståelse af hvordan lys og stof vekselvirker muliggør en lang række af de teknologier, som definerer livet i den moderne verden. Som eksempler kan nævnes lasere, solceller og lysemitterende dioder, der alle sandsynligvis vil spille en nøglerolle i vores overgang til et mere bæredygtigt samfund. Denne afhandling undersøger lys-stof vekselvirkninger ved hjælp af kvantemekaniske beregninger. Det demonstreres, hvordan sådanne beregninger kan bruges til både at fortolke nye eksperimentelle resultater, og aktivt drive materialeopdagelse til optoelektroniske applikationer. Afhandlingen præsenterer både nye metoder såvel som nye anvendelser af eksisterende metoder.

Lysets kvantemekaniske natur er traditionelt set blevet negligeret i faststoffysik og kvantekemi. Den kan dog blive vigtig, når det elektromagnetiske miljø er struktureret på en sådan måde, at det er vært for resonante lystilstande. En særlig interessant version af dette er, hvordan resonante elektromagnetiske miljøer kan bruges til at ændre den elektroniske struktur af materialer, selv i fravær af faktisk lys. Dette sker som et resultat af de ændrede kvantefluktuationer i det elektromagnetiske felt, og effekten er direkte relateret til, at både lys og stof fundamental set kræver en kvantemekanisk beskrivelse. Denne afhandling præsenterer en ny metodik, der kombinerer makroskopisk kvanteelektrodynamik med forskellige varianter af tæthedsfunktionalteori. Dette er et skridt imod en fuldstændig *ab-initio* beskrivelse af kvantemekaniske lys-stof interaktioner i vilkårlige elektromagnetiske miljøer.

Kavitetskvanteelektrodynamik med farvecentre i 2D-materialer er også blevet undersøgt. Dette blev gjort ved hjælp af en kombination af den tensorsnetværksbaserede TEMPO-algoritme til tidsudbredelse af åbne kvantesystemer og en *ab-initio* beskrivelse af elektron-fonon koblingen. Denne undersøgelse viser, at strukturen i elektron-fononkoblingens spektrale tæthed manifesterer sig som klare signaturer af umarkovianitet i dynamikken i det koblede system. Ydermere findes der klare tegn på hybridisering mellem lys-stof polaritonerne og miljøets fonontilstande.

Et andet fokus i denne afhandling er materialeopdagelse til nanofotoniske og optoelektroniske applikationer. Ved at bruge en kombination af høj gennemstrømnings beregningscreening, nøjagtig mangelegeme forstyrrelsesteori, eksperimentel syntese samt nær- og fjernfelts optisk karakterisering, identificeres Bor-Fosfid som et overset højtydende materiale til dielektrisk nanofotonik i det synlige og ultraviolette spektrum.

Afhandlingen behandler også halvledere med indirekte båndgab til fotovoltaiske applikationer. Ved at kombinere en beregningsmæssigt effektiv tilnærmelse af fononasisteret indirekte absorption med høj gennemstrømnings beregningscreening, identificeres 9 nye materialer som potentielt brugbare i fremtidig tyndfilmssolceller.

Nye resultater fra projektet: Computational 2D Materials Database (C2DB) præsenterer

teres også. Sammenligning med fulde fononberegninger validerer Center og Boundary Protocol som et effektivt værktøj til at evaluere dynamisk stabilitet af optimerede 2D atomstrukturer. Endvidere vises det, at protokollen kan anvendes iterativt til at generere dynamisk stabile strukturer fra ustabile input strukturer via systematiske dislokationer af gitteret. Afslutningsvis bruges de eksisterende stabilitetsdata i C2DB til at træne en maskinlæringsmodel til at forudsige dynamisk stabilitet. Denne model præsterer godt og viser en fremragende modtagerdrifts karakteristik (ROC) kurve med et areal under kurven (AUC) på 0.9. Resultaterne viser, at integrationen af sådanne klassifikationsmodeller i beregnings workflows kan reducere beregningsindsatsen drastisk i høj gennemstrømnings beregningsscreeninger.

Ydermere blev de elementære elektroniske excitationer af det lagdelte materiale PtSe₂ optrevet ved hjælp af både momentum-opløst elektronenergita bsspektroskopi og *ab-initio* beregninger. Derudover blev det vist, hvordan man kan vende teoretiske overvejelser om momentum-afhængighed af Coulomb interaktionsmatricielementerne til brugbar spektroskopisk indsigt for 2D-materialer. Dette kan potentielt føre til nye muligheder inden for spektroskopi.

Endelig præsenterer afhandlingen resultater fra et stort samarbejde, der forsøger at bruge van der Waals (vdW) heterostrukturer som en kilde til energiafstembar røntgenstråling. Det vises, at emissionens energi kan tunes både via den indkommende elektronenergi og konfigurationen af heterostrukturen. Dette kan være med til at bane vejen for energijusterbare røntgenkilder, som er meget kompakte sammenlignet med eksisterende teknologier.

Acknowledgements

No thesis is written in a vacuum and there are a lot of people that I need to thank for their support over the past three years.

I first want to thank my supervisor Kristian. It has really been a pleasure working with you, and I am grateful that you always let me pursue even the more crazy ideas. The freedom to try, and fail at, a lot of different projects really helped me grow as a scientist.

Furthermore, I want to thank Angel Rubio and Johannes Flick for their great supervision during my external research stay and beyond. It has been a true pleasure working with you and I have learned a lot from it! I also want to thank the CCQ gang Carl, Giacomo, Florian, Kevin, Angkun and Nicolas for all the great times we shared!

Moreover, I want to thank the people of CAMD for always being up for both serious and very unserious discussions about everything between heaven and earth! I would have never thought that I would be designing computers made out of microscopic birds as part of the coffee breaks! Thanks for making the experience a lot more fun: Morten, Estefania, Asbjørn, Daniele, Anders, Alireza, Sajid, Joachim, Peder, Jens Jørgen, Ole, Ask, Sahar, Cuau, Casper, Julian, Urko, Sami, Mikael, Fredrik, Peter, Frederik, Luca and Jiban. I want to extended a particular thanks to Fabian, Nikolaj, Matthew, Thorbjørn, Mads, Hadeel and Stefano which are the CAMD people that I have shared the most of this journey with. A special thanks also goes to Bettina and Lone for their help navigating the sometimes very dense jungle of non-scientific tasks which also constitute an important part of doing a Ph.D.

I next want to extend a special thanks to Mikael, Fredrik, Thorbjørn and Julian for your help with proofreading of this thesis. It really helped me significantly improve the quality of this thesis and I will be forever grateful to you for taking time out to help me get this thesis to where I wanted it to be!

To Asbjørn, thank you for some great discussions and a lot of fun along the way! I'm still betting on you making me the duke of Næstved if this whole science thing doesn't work out in the end.

To Rasmus, thank you for everything, professional and personal! It has been absolutely great to know that I always had a friend on the outside that understood what it was about when things got tough. One might say that it was amazing, la, la, la...

To Fabian, there is a small misunderstanding in the literature that I want to clarify for you. Recently, bogus claims about zinc's superiority over iron have been made¹.

¹Bertoldo, F. (2022) *Computational Design of Quantum Defects in Low-Dimensional Semiconductors* [Ph.D. Thesis, Technical University of Denmark]

However, given that iron accounts for the vast majority of all metal production, and the main application of zinc is as corrosion protection for iron, it is safe to say that those claims were unfounded. Apart from that, thank you for all the good times during the home office days! It was a true pleasure sharing those times with you, and I hope and believe that it became the beginning of a life long friendship.

To my family in danish: En afhandling som denne tager tre år at skrive, men den er kun mulig som følge af den støtte I har givet mig siden jeg var helt lille. Tak for alt Mor, Far, Mathias og Micka!

To Anna, there is no one I need to thank more than you! These past three years had their ups and downs but you've stood by me the whole time. Thank you for always having my back when things get tough, and for reminding me that life has much more to offer than physics. I love you and I could not have done this without you!

List of Publications

Primary Publications

- [I] Mark Kamper Svendsen[†], Yaniv Kurman[†], Peter Schmidt, Frank Koppens, Ido Kaminer, Kristian S Thygesen
“Combining density functional theory with macroscopic QED for quantum light-matter interactions in 2D materials,”

[†] The authors contributed equally to the work,
[Nature Communications 12, 2778 \(2021\)](#).

Printed in copy from page 123.

- [II] Mark Kamper Svendsen, Hiroshi Sugimoto, Artyom Assadillayev, Daisuke Shima, Minoru Fujii, Kristian S Thygesen, Søren Raza,
“Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators,”
[Advanced Optical Materials \(2022\): 2200422](#).

Printed in copy from page 137.

- [III] Jinhua Hong[†], Mark Kamper Svendsen[†], Masanori Koshino, Thomas Pichler, Hua Xu, Kazu Suenaga, Kristian S Thygesen,
[†] The authors contributed equally to the work,
“Momentum-Dependent Oscillator Strength Crossover of Excitons and Plasmons in Two-Dimensional PtSe₂,”

[ACS Nano 2022, 16, 8, 12328–12337](#).

Printed in copy from page 147.

- [IV] Simone Manti, Mark Kamper Svendsen, Nikolaj R. Knøsgaard, Peder M. Lyngby, Kristian S. Thygesen,
“Predicting and machine learning structural instabilities in 2D materials,”
[Submitted to npj Computational Materials: arXiv:2201.08091](#).

Printed in copy from page 158.

- [V] Jiban Kangsabanik, Mark Kamper Svendsen, Alireza Taghizadeh, Andrea Crovetto, Kristian S. Thygesen,
“Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption,”
Submitted to Journal of the American Chemical Society: arXiv:2207.09569.

Printed in copy from page 172.

- [VI] Mark Kamper Svendsen, Sajid Ali, Nicolas Stenger, Kristian Sommer Thygesen, Jake Iles-Smith,
“Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials,”
Submitted to arXiv: arXiv:2207.10630.

Printed in copy from page 186.

Secondary Publications (Minor Contributions)

- [VII] M. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N. R. Knøsgaard, M. Kruse, A. H. Larsen, S. Manti, T. G. Pedersen, U. Petralanda, T. Skovhus, M. K. Svendsen, J. J. Mortensen, T. Olsen and K. S. Thygesen,
“Recent Progress of the Computational 2D Materials Database (C2DB),”
2D Mater. 2021, **8**, 044002.

Printed in copy from page 194.

- [VIII] Michael Shentcis, Adam K Budniak, Xihang Shi, Raphael Dahan, Yaniv Kurman, Michael Kalina, Hanan Herzig Sheinfux, Mark Blei, Mark Kamper Svendsen, Yaron Amouyal, Sefaattin Tongay, Kristian Sommer Thygesen, Frank HL Koppens, Efrat Lifshitz, F Javier García de Abajo, Liang Jie Wong, Ido Kaminer,
“Tunable free-electron X-ray radiation from van der Waals materials,”
Nat. Photonics 14, 686–692 (2020).

Printed in copy from page 223.

Other Publications not included in this thesis

- [IX] Mark Kamper Svendsen, Christian Wolff, Antti-Pekka Jauho, N Asger Mortensen, Christos Tserkezis,
“Role of diffusive surface scattering in nonlocal plasmonics,”
J. Phys.: Condens. Matter 32 395702.

-
- [X] Niels CH Hesp, Mark Kamper Svendsen, Kenji Watanabe, Takashi Taniguchi, Kristian S Thygesen, Iacopo Torre, Frank HL Koppens, “WSe₂ as Transparent Top Gate for Infrared Near-Field Microscopy,” *Nano Letters* 2022, 22, 15, 6200–6206.
- [XI] Moritz Fischer, Ali Sajid, Jake Iles-Smith, Alexander Hötger, Denys I. Miakota, Mark Kamper Svendsen, Christoph Kastl, Stela Canulescu, Sanshui Xiao, Martijn Wubs, Kristian S. Thygesen, Alexander W. Holleitner, and Nicolas Stenger “Exploring the phonon-assisted excitation mechanism of luminescent centres in hexagonal boron nitride by photoluminescence excitation spectroscopy,” *Submitted to arXiv: arXiv:2209.08910*.

Contents

Preface	i
Abstract	ii
Resumé	iv
Acknowledgements	vi
List of Publications	viii
Primary Publications	viii
Secondary Publications (Minor Contributions)	ix
Other Publications not included in this thesis	ix
Contents	xi
Introduction	1
Thesis Outline	3
1 Quantum Electrodynamics and Light Matter Interactions	4
1.1 Quantizing the Electromagnetic field: Macroscopic Quantum Electrodynamics	5
1.2 Interaction of the electromagnetic field and matter	11
1.3 Regimes of Quantum Light Matter Interactions	16
1.4 The Dyadic Green's Function	19
2 The Many-Body Problem	23
2.1 Ground state Density Functional Theory	23
2.2 Time Dependent Density Functional Theory with photons	28
3 Linear Response Theory	31
3.1 Linear response theory	32
3.2 Linear response of a non-relativistic QED system	32
3.3 Calculating linear response functions within the QEDFT framework	34
3.4 The Random Phase Approximation	35
3.5 The generalized Casida formulation of LR-QEDFT for finite systems	36
4 Towards a first-principles description of real molecules in real cavities by combining MQED and QEDFT	39
4.1 Formalism	40

4.2	Spherical Micro-Cavities	42
4.3	Adding an emitter to the cavity	45
4.4	Conclusion	48
4.5	Computational details for the QEDFT calculations	49
5	Dielectric Screening in Materials	50
5.1	The Dielectric Function	50
5.2	Many Body Perturbation Theory and Dielectric Screening	54
5.3	Spectroscopic quantities and dielectric screening	62
6	Tensor network methods for quantum light-matter interactions with open quantum systems	70
6.1	Open Quantum Systems	70
6.2	The TEMPO algorithm	73
6.3	Linear response within the TEMPO algorithm	77
7	Summary of the results related to quantum light-matter interactions	80
7.1	Publication [I]: Combining density functional theory with macroscopic QED for quantum light- matter interactions in 2D materials	80
7.2	Publication [VI]: Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials	83
8	Summary of the results related to Computational Materials Discovery	87
8.1	Publication [II]: Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators	87
8.2	Publication [V]: Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption	92
8.3	Publication [VII]: Recent progress on the computational 2D materials database	94
8.4	Publication [IV]: Predicting and machine learning structural instabilities in 2D materials	96
9	Summary of the results from experimental collaborations	100
9.1	Publication [III]: Momentum-dependent oscillator strength crossover of excitons and plasmons in two-dimensional PtSe ₂	100
9.2	Publication [VIII]: Tunable free-electron X-ray radiation from van der Waals materials	104
10	Summary and Outlook	107
	Bibliography	110
11	Papers	123
11.1	Paper [I]: Combining density functional theory with macroscopic QED for quantum light- matter interactions in 2D materials	123
11.2	Paper [II]: Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators	137
11.3	Paper [III]: Momentum-Dependent Oscillator Strength Crossover of Excitons and Plasmons in Two-Dimensional PtSe ₂	147
11.4	Paper [IV]: Predicting and machine learning structural instabilities in 2D materials,	158

11.5	Paper [V]: Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption,	172
11.6	Paper [VI]: Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials,	186
11.7	Paper [VII]: Recent Progress of the Computational Materials Database (C2DB) . . .	194
11.8	Paper [VIII]: Tunable free-electron X-ray radiation from van der Waals materials, . .	223

Introduction

This thesis is about light-matter interactions. It is hard to know exactly where to begin the story of how light and matter interacts. The reasons why it is interesting are so numerous that it is hard to do it justice in the few lines that the introduction to a Ph.D. dissertation can afford. What is certain is that light-matter interactions are an extremely fundamental part of the world around us. It is the reason that we are able to see, the reason that the sun is able to heat the earth, and the reason that plants are able to convert sunlight into the sugars that ultimately feed every single organism on earth.

Our understanding of light-matter interactions has also helped enable a wide range of the technologies that define life in the modern world. One very famous and very important example is the laser[1]. Lasers play a central role in a wide range of different technologies. Along with our understand of how materials can be engineered to guide light in optical waveguides, lasers form the backbone of data transfer on the internet[2]. Other important examples are solar cells and LED lighting, both of which are likely to play a central role in the transition to a future sustainable society. What these examples have in common is that they are made possible by our scientific understanding of how light and matter interacts on a quantum level. In addition to being fundamentally interesting, there is therefore good reason to study light-matter interactions from a societal point of view.

Light is an electromagnetic wave of a certain frequency. Matter on the other hand can take many forms. In all cases, it is made of atoms which consists of negatively charged electrons and positive charged protons. The charge of these constituents means that even though most materials are globally charge neutral, the matter can still respond to the electromagnetic field carried by the light. At the heart of our theoretical understanding of light-matter interactions lies quantum mechanics. This theory, which grew from the work of the likes of Bohr, Planck, Schrödinger, Dirac and Heisenberg, has managed to completely change our scientific understanding of the world around us. Among other things, these efforts culminated in the theory of Quantum Electrodynamics (QED) around the middle of the twentieth century. QED is the fundamental theory describing charged particles and their interaction with the electromagnetic field[3]. Since matter is made up of charged particles interacting with each other via the electromagnetic field, this meant that the fundamental theoretical framework to describe condensed matter physics and quantum chemistry was now at hand. Comparing the situation to a game of chess, it was now known what the game pieces are and how they move. This led the famous physicist Paul Dirac to playfully remark:

Quantum Theory has explained all of chemistry and most of physics.

Paul A.M. Dirac

However, as with many things in life, the devil is in the detail. As is also the case in chess, knowing what the pieces are and how they move is a far cry from understanding all the intricacies of the game. It is also nowhere near the same as having an overview of all the possible games that can exist. A good example of just how complicated condensed matter physics can be is carbon. Carbon compounds comes in a wide range of allotropes[4]. While all of the allotropes consist of the same basic building blocks, carbon atoms, their difference in structure can result in extremely different properties. In particular, carbon can be either one of the best conductors known to man (graphene), or almost impossible to pass a current through (diamond).

Collectively, the properties of materials that result from the intricate way the charged particles that make them up interact are called the electronic structure of the material. This electronic structure dictates a materials interaction with light, and it can be calculated using so-called first-principles or *ab-initio* methodology. First-principles means that the only information going into the calculation is the atoms constituting the material and their positions. This thesis presents examples of how these types of calculations can be used to both interpret novel experimental results, and actively drive materials discovery for different opto-electronic applications. Especially the latter approach is only possible because computing power has advanced to the point where it is possible to perform high-throughput calculations to determine material properties. Part of the less glamorous, but equally important, work leading to this thesis has been dedicated to how high-throughput computational workflows can be constructed and potentially improved using techniques from Machine Learning.

Light can also be used to change material properties. A particularly interesting version of this is how resonant electromagnetic environments can be used to alter the electronic structure of materials, even in the absence of actual light in the cavity. This happens as a result of the altered quantum fluctuations of the electromagnetic field in such environments, and the effect is directly related to the fundamental quantum nature of both light and matter[5, 6]. It also presents a uniquely challenging computational problem because it becomes necessary to quantum mechanically treat both the light and matter at the same high level of accuracy. The development of methods capable of this is one of the central topics of the present thesis.

This thesis focuses on crystalline materials which can be treated with standard first-principles methods. The thesis investigates both conventional 3D materials like silicon, and the more novel 2D materials. 2D materials are a class of materials consisting of atomically thin layers. Most 2D materials are naturally found as layered bulk crystals wherein the layers are bound together by weak van der Waals (vdW) interactions. In fact, the first example of a 2D material was graphene which was realized as a single layer of a graphite crystal exfoliated from the bulk crystal using simple scotch tape[7]. Other well known examples of 2D materials are the Transition Metal Dichalcogenides (TMD) monolayers MoS_2 , MoSe_2 , WS_2 and WSe_2 [8]. The weak interlayer vdW bonds means that different 2D materials can be stacked into almost arbitrary composite structures without regard to lattice matching. Such heterostructure represents a completely unique and very versatile material platform called vdW heterostructures[9]. Addition-

ally, the atomically tight confinement of the electrons in the out-of-plane direction results in unique physical properties. Such structures can for example host excitations with very confined electromagnetic fields[10, 11], strong exciton resonances[12], and thickness tunable infrared emission originating from intersubband transitions with large out-of-plane dipole moments[13]. This makes vdW heterostructures an interesting platform to study strong quantum light-matter interactions.

Thesis Outline

The introduction concludes with an outline of the thesis. The main body of the thesis is split into two parts. The first part of the thesis is the methodology part and it is structured as follows:

1. The first chapter will discuss relevant aspects of Quantum Electrodynamics (QED) which is the theory that governs light-matter interactions. Particular emphasis is put of Macroscopic QED (MQED) which is the theory governing quantized electromagnetism in the presence of lossy magnetoelectric backgrounds.
2. The second chapter discusses Density Functional Theory (DFT) as a practical solution to address the many-body problem of interacting electrons and photons.
3. The third chapter discusses the linear response of a non-relativistic QED system and its description in terms of the DFT construction.
4. The fourth chapter discusses ongoing work that seeks to combine concepts discussed in the first three chapters into a general method to describe practically relevant electronic systems in real optical cavities from first principles.
5. The fifth chapter discusses dielectric screening in solids and its relation to different kinds of spectroscopy. This discussion is build around a combination of linear response theory and many-body perturbation theory.
6. The sixth chapter discusses the use of tensornetwork based methods to describe cavity-QED for open quantum systems.

The second part of the thesis presents the main results from the papers included in this thesis. This discussion is subdivided into three parts that group the papers thematically. Copies of all the papers are provided at the end of the thesis.

7. The seventh chapter reviews Publications [I] and [VI] which can be labelled as relating to quantum light-matter interactions. It should be mentioned that the work presented in chapter 4 also falls within this same category.
8. The eighth chapter reviews Publications [II], [IV], [V] and [VII] which can be labelled as relating to computational materials discovery.
9. The ninth chapter reviews Publications [III] and [VIII] which have both resulted from collaborations with experimental colleagues.

CHAPTER 1

Quantum Electrodynamics and Light Matter Interactions

At a fundamental level, light-matter interactions are described by Quantum Electrodynamics (QED). This Chapter will discuss aspects of QED as it relates to condensed matter physics and quantum chemistry. This means that the discussion will be restricted to low-energy, non-relativistic QED.

QED is fundamentally a quantum theory of both the matter and the electromagnetic field. Nevertheless, the quantum nature of light has traditionally broadly been neglected in the context of condensed matter physics and quantum chemistry. This is often a good approximation, since many phenomena are well described within a classical treatment of the electromagnetic field. However, the quantum nature of the light can become important when the electromagnetic environment is structured in such a way as to host resonant or confined modes of light. Such situations can broadly be labelled as cavity-QED (c-QED). An example of this is illustrated in Fig. 1.1 where a $V_N C_B$ color center in a hexagonal boron nitride monolayer is embedded in an optical cavity. To provide a general description of such configurations, it is necessary to account for the quantum nature of the light. While this situation has traditionally been restricted to the domain of quantum optics[14], there has in recent years been a lot of progress and interest in utilizing the quantum nature of light with quantum chemistry and condensed matter physics. Some interesting examples include the emerging field of QED chemistry where the chemical landscape is altered by the coupling to quantized light[5, 15–18], alteration of photophysical processes such as e.g. singlet fission[19–21], cavity control of excitons[22], suggestions to realize novel topological states of matter with specially engineered cavities[6], and alterations of the quantized Hall conductance by the interaction with a cavity[23, 24].

The discussion in this chapter is structured as follows,

1. First the quantum theory of the electromagnetic field in the presence of a general, lossy material background is discussed.
2. Next, the interaction between the quantized electromagnetic fields and matter is discussed.

3. Finally, the concept of the Dyadic Green's Function is introduced.

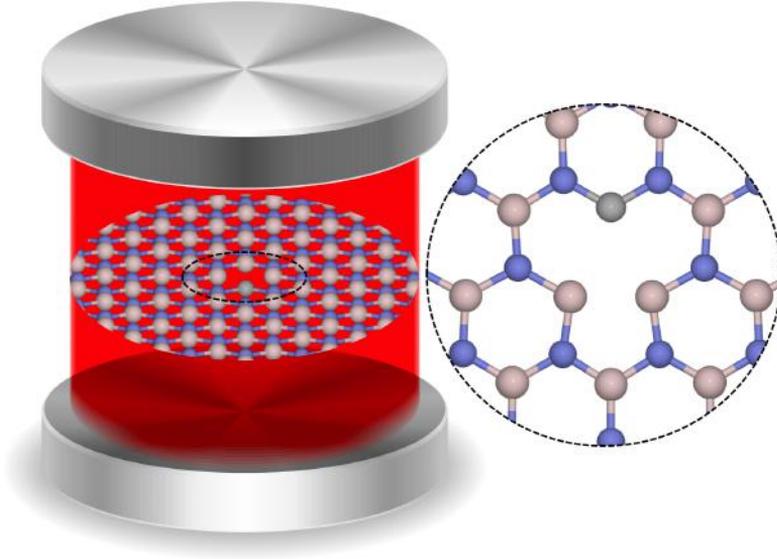


Figure 1.1: Illustration of a $V_N C_B$ color center in a hexagonal boron nitride monolayer embedded in an optical cavity. Figure reproduced from Publication [VI].

1.1 Quantizing the Electromagnetic field: Macroscopic Quantum Electrodynamics

This section discusses the quantum theory of the electromagnetic field. First, the case of electromagnetic fields in a lossless background is reviewed to get the general feel for what a quantum theory of electromagnetism looks like. Subsequently, the difficulties relating to the quantization in a general magnetoelectric background are outlined and a solution in the form of Macroscopic Quantum Electrodynamics (MQED) is discussed.

Before diving into the details of quantization, it makes sense to briefly discuss the representation of the electromagnetic field. The electromagnetic field is fundamentally a gauge field[3]. Formally, this gauge field is represented by a four-vector $A^\mu(s)$, where $\mu = t, x, y, z$ and s denotes a spacetime point $s = (\mathbf{r}, t)$. It is customary to denote the temporal component of the electromagnetic field as the scalar potential $\phi(\mathbf{r}, t)$, and the spatial components as the vector potential $\mathbf{A}(\mathbf{r}, t)$, such that $A^\mu(s) = (\phi(s), \mathbf{A}(s))$. To ensure gauge invariance, all scalar and vector potentials that are related by a transfor-

mation of the form,

$$\mathbf{A}(\mathbf{r}, t) \rightarrow \mathbf{A}(\mathbf{r}, t) + \nabla\chi(\mathbf{r}, t) \quad (1.1)$$

$$\phi(\mathbf{r}, t) \rightarrow \phi(\mathbf{r}, t) - \partial_t\chi(\mathbf{r}, t), \quad (1.2)$$

have to be physically equivalent. Here $\chi(\mathbf{r}, t)$ is any twice differentiable function of space and time. $\phi(\mathbf{r}, t)$ and $\mathbf{A}(\mathbf{r}, t)$ are therefore not physical fields, but merely mathematical constructions to ensure the gauge invariance of QED. They are however generally very useful constructions and directly related to the physical electric- and magnetic fields,

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t), \quad \mathbf{E}(\mathbf{r}, t) = -\partial_t\mathbf{A}(\mathbf{r}, t) - \nabla\phi, \quad (1.3)$$

which in turn obey Maxwell's equations. As will be seen in this chapter, the gauge freedom of electromagnetism is a very important concept in QED because the choice of gauge ultimately dictates the form of the Hamiltonian describing the coupled light-matter system.

1.1.1 Quantization in a Lossless Medium

There are essentially two ways of approaching the quantization of the electromagnetic field in the lossless case. The first is the formal field theoretical approach, starting from the Lagrangian density of the $U(1)$ -gauge field, and then proceeding to quantize the theory via the five step canonical quantization procedure[3, 25]. While this approach is very elegant, it obscures the connection between the modes of the electromagnetic field and any structure that might exist in the environment such as e.g. the mirrors of a cavity etc.¹ Alternatively, the electromagnetic field can also be quantized starting from Maxwell's equations[26, 27]. This approach has the advantage of being conceptually simpler and more directly related to the boundary conditions of the problem. Regardless of the approach, the end result is that the fields are promoted to operators, which in turn are expanded in terms of a set of orthonormal modes. In the following discussion, the Schrödinger picture of quantum mechanics is adopted wherein the time dependence is carried by the wave functions.

As a first step, it makes sense to fix the gauge freedom. A popular choice in quantum optics is the Coulomb gauge which fixes $\nabla \cdot \mathbf{A} = 0$. In the absence of external charges and currents, this further translates into $\phi = 0$. Choosing the Coulomb gauge, the quantized field operators in the lossless medium can be written as,

$$\hat{\mathbf{E}}(\mathbf{r}) = i \sum_m \omega_m [\mathbf{A}_m(\mathbf{r})\hat{a}_m - \mathbf{A}_m^*(\mathbf{r})\hat{a}_m^\dagger], \quad (1.4)$$

$$\hat{\mathbf{B}}(\mathbf{r}) = \nabla \times \sum_m [\mathbf{A}_m(\mathbf{r})\hat{a}_m + \mathbf{A}_m^*(\mathbf{r})\hat{a}_m^\dagger], \quad (1.5)$$

¹Note that this section considers *lossless* setups and therefore any such mirrors would have to be lossless mirrors.

where $\hat{a}_m, \hat{a}_m^\dagger$ are bosonic creation/annihilation operators that create/destroy photons in mode m and satisfy $[\hat{a}_m, \hat{a}_n^\dagger] = \delta_{mn}$. ω_m is the frequency of mode m , and $\mathbf{A}_m(\mathbf{r})$ are the solutions to the spatial Helmholtz equation,

$$\nabla^2 \mathbf{A}_m(\mathbf{r}) = -k_m^2 \mathbf{A}_m(\mathbf{r}). \quad (1.6)$$

Importantly, in the lossless case, the Helmholtz equation in Hermitian and the mode functions form a complete, orthogonal basis.

The energy density of the electromagnetic field can be expressed with Poynting's theorem,

$$\mathcal{H} = \frac{1}{2} \left(\epsilon \hat{\mathbf{E}}(\mathbf{r})^2 + \frac{1}{\mu} \hat{\mathbf{B}}(\mathbf{r})^2 \right). \quad (1.7)$$

Here ϵ and μ are the permittivity and permeability of the medium respectively.² The total energy of the system, the Hamiltonian, is thus,

$$H = \int d\mathbf{r} \mathcal{H} = \sum_m \hbar \omega_m \left(\hat{a}_m^\dagger \hat{a}_m + \frac{1}{2} \right), \quad (1.8)$$

where the last equality follows from the orthogonality of the mode functions. The Hamiltonian is often alternatively represented in terms of a set of canonical operators, $\hat{p}_m = \left(\frac{\hbar}{2\omega}\right)^{1/2} (\hat{a}_m + \hat{a}_m^\dagger)$ and $\hat{q}_m = i \left(\frac{\hbar}{2\omega}\right)^{1/2} (\hat{a}_m^\dagger - \hat{a}_m)$,

$$\hat{H} = \frac{1}{2} \sum_m [\hat{p}_m^2 + \omega_m^2 \hat{q}_m^2]. \quad (1.9)$$

which are interpreted as the momentum and coordinate of mode m respectively. In the absence of losses, the quantized theory of the electromagnetism can therefore be written in terms of a discrete set of uncoupled Harmonic oscillators representing the eigenmodes of the Helmholtz equation. The excitations of the system are bosonic and represent the photons of the electromagnetic field, created and annihilated by \hat{a}_m and \hat{a}_m^\dagger respectively.

Finally the orthogonality of the Helmholtz equation's mode functions ensures that the Equal Time Commutation Relation (ETCR) of the field operators becomes,

$$\left[\hat{\mathbf{E}}(\mathbf{r}), \hat{\mathbf{B}}(\mathbf{r}') \right] = -\frac{i\hbar}{\epsilon_0} \nabla \times \delta^\perp(\mathbf{r} - \mathbf{r}'), \quad (1.10)$$

where $\delta^\perp(\mathbf{r} - \mathbf{r}')$ is the transverse Dirac delta function, and ϵ_0 is the vacuum permittivity. Eq. 1.10 is important because it agrees with the classical Poisson bracket of the electromagnetic fields and thus obeys the correspondence principle of quantum mechanics.

²Importantly, in the lossless medium both quantities have to be real and non-dispersive.

1.1.2 Quantization in the presence of losses

Quantizing the electromagnetic field becomes more complicated in the presence of losses. In classical optics, the effect of the matter background on light is contained in the optical parameters such as for example the refractive index, $n(\omega)$ [2]. The refractive index enters the Helmholtz equation as a modification of the wave number, $k = n(\omega)k_0$, where k_0 would be the wave number in free space. For strictly real $n(\omega)$, the quantization procedure follows directly from the previous section by simple substitution of k_0 with k . However, in any real material, the refractive index will be frequency dependent/dispersive. To comply with the Kramers-Kronig relations, this means that the refractive index must also have a non-zero imaginary part, at least in some frequency range[2]. The non-zero imaginary part of k renders the Helmholtz equation non-Hermitian and its eigenfunctions will no longer form a complete, orthogonal basis. This is problematic because the orthogonality of the mode functions was what ensured the commutation relation of the electric and magnetic fields in Eq. 1.10. One can insist on the commutator of the fields being correct, but this will come at the cost of losing the bosonic commutation relations of the photonic creation and annihilation operators[27].

The straightforward approach to the quantization in the presence of losses fails because there is an underlying matter theory encoded in the frequency dependence of the refractive index which is not properly accounted for[28, 29]. The solution to the problem is therefore conceptually simple but complicated in practice: A proper description of the quantized fields must be one that accounts for the underlying matter explicitly.

1.1.3 The Huttner-Barnett model

The first example of a quantum theory of the electromagnetic field in the presence of a lossy background was the microscopic theory developed by Huttner and Barnett in 1992[28, 29]. Here a Hopfield model was considered, with the electromagnetic vector potential coupled to a harmonic oscillator mode representing the matter polarization, and the matter polarization further coupled to a continuum of harmonic oscillators representing a bath. The latter coupling results in unidirectional energy transfer from the polarization mode of the matter into the bath which represents the losses. This coupled theory can be diagonalized in terms of a continuous set of polaritonic operators, $\hat{\mathbf{f}}(\mathbf{r}, \omega)$, that have both light and matter character. The resulting Hamiltonian of the coupled system reads,

$$\hat{H} = \int_0^\infty d\omega \hbar \omega \int d\mathbf{r} \hat{\mathbf{f}}^\dagger(\mathbf{r}, \omega) \cdot \hat{\mathbf{f}}(\mathbf{r}, \omega), \quad (1.11)$$

where the polaritonic operators obey bosonic commutation relations, $[\hat{\mathbf{f}}(\mathbf{r}, \omega), \hat{\mathbf{f}}^\dagger(\mathbf{r}', \omega')] = \delta(\omega - \omega')\delta(\mathbf{r} - \mathbf{r}')$. Using the inverse of the transformation leading to the diagonalization, it is possible to express the electromagnetic fields in terms of these polaritonic operators such that the ETCR of the electric and magnetic fields equals their classical Poisson bracket. The full diagonalization of the problem thus results in the desired properties for a quantized theory of the electromagnetic field but comes at the cost

of having a continuous set of modes instead of the discrete set seen in section 1.1.1. Importantly, the expansion coefficients of the vector potential turn out to be Dyadic Green's Function (DGF) of a homogeneous, isotropic material with a dielectric function related to the properties of the matter polarization and the bath (see section 1.4 for details on the DGF). The DGF is a completely classical object but it encodes both the material properties of the constituents and the boundary conditions of the problem. A generalization of the Huttner-Barnett approach to arbitrary geometries would therefore represent a very powerful description of the quantized electromagnetic field.

1.1.4 The Langevin noise approach

The Huttner-Barnett model can be generalized to arbitrary media and geometries as an effective Langevin noise model[27, 30]. Whereas the Huttner-Barnett model is a microscopic approach, the Langevin noise model starts from the macroscopic Maxwell's equations in matter and from them seeks to identify the proper dynamical variable around which to build the quantized theory.

The starting point of the theory is the coupling between the electromagnetic field and matter. Assuming that the material responds linearly to the field, the induced current in response to an electric field is given by Ohm's law,

$$\mathbf{j}_{\text{in}}(\mathbf{r}, \omega) = \int d\mathbf{r}' \mathbf{Q}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{E}(\mathbf{r}', \omega) + \mathbf{j}_{\text{N}}(\mathbf{r}, \omega). \quad (1.12)$$

Here $\mathbf{Q}(\mathbf{r}, \mathbf{r}', \omega)$ is the complex conductivity tensor of the medium, and $\mathbf{j}_{\text{N}}(\mathbf{r}, \omega)$ is the random noise current necessary for the linear constitutive relations to satisfy the Fluctuation-Dissipation theorem[31]. The real part of the complex conductivity is related to energy dissipation $\sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{2} (\mathbf{Q}(\mathbf{r}, \mathbf{r}', \omega) + \mathbf{Q}^\dagger(\mathbf{r}, \mathbf{r}', \omega))$ and for absorbing media it represents a positive-definite integral kernel.

The fields will also respond to the current. This will be generally be governed by the inhomogeneous Helmholtz equation, which in the presence of the current in Eq. 1.12 reads,

$$\left[\nabla \times \nabla \times - \frac{\omega^2}{c^2} \right] \mathbf{E}(\mathbf{r}, \omega) - i\mu_0\omega \int ds \mathbf{Q}(\mathbf{r}, \mathbf{s}, \omega) \cdot \mathbf{E}(\mathbf{s}, \omega) = i\mu_0\omega \mathbf{j}_{\text{N}}(\mathbf{r}, \omega), \quad (1.13)$$

with the noise current acting as a source of the electric field. The presence of the source terms renders a solution in terms of eigenmodes impossible and this emphasizes that a theory of quantized electromagnetism in lossy media that neglects the underlying matter degrees of freedom is impossible to construct. Eq. 1.13 be formally solved in terms of the DGF,

$$\mathbf{E}(\mathbf{r}, \omega) = i\mu_0\omega \int ds \mathbf{G}(\mathbf{r}, \mathbf{s}, \omega) \cdot \mathbf{j}_{\text{N}}(\mathbf{s}, \omega). \quad (1.14)$$

Similarly, a Helmholtz equation describing the evolution of the magnetic field can also be derived, the solution to which can also be written in terms of the DGF,

$$\mathbf{B}(\mathbf{r}, \omega) = \mu_0 \nabla \times \int ds \mathbf{G}(\mathbf{r}, \mathbf{s}, \omega) \cdot \mathbf{j}_{\text{N}}(\mathbf{s}, \omega). \quad (1.15)$$

The DGF is a completely classical object, but it is of central importance to the quantized theory of electromagnetic fields in lossy environments because it simultaneously carries the information about the electromagnetic boundary conditions, and, as discussed below, it serves as a projector from the coupled light-matter system onto the electromagnetic degrees of freedom. The DGF is a central topic for the present thesis and section 1.4 is dedicated to a thorough discussion of it.

Eqs. 1.14 and 1.15 indicate that the noise current, $\mathbf{j}_N(\mathbf{s}, \omega)$, is the fundamental dynamical variable of the theory, and the quantized theory should therefore be build around it. The theory is quantized by promoting $\mathbf{j}_N(\mathbf{s}, \omega)$ to an operator and imposing equal time commutation relations,

$$\left[\hat{\mathbf{j}}_N(\mathbf{r}, \omega), \hat{\mathbf{j}}_N(\mathbf{r}', \omega') \right] = \frac{\hbar\omega}{\pi} \delta(\omega - \omega') \boldsymbol{\sigma}(\mathbf{r}, \mathbf{r}', \omega), \quad (1.16)$$

in accordance with the fluctuation-dissipation theorem[30]. The commutation relation in Eq. 1.16 makes it clear that the noise current operator is not a bosonic vector field. However, in the absence of gain, the fact that the real part of the conductivity tensor is a positive definite operator allows the MQED theory to be recast in terms of a new set of simple bosonic vector fields. In that case, the matrix square-root of the conductivity tensor exists, and one can define a set of bosonic operators $\hat{\mathbf{f}}(\mathbf{r}, \omega)$ as from the noise current,

$$\hat{\mathbf{j}}_N(\mathbf{r}, \omega) = \left(\frac{\hbar\omega}{\pi} \right)^{1/2} \int d\mathbf{s} \mathbf{K}(\mathbf{r}, \mathbf{s}, \omega) \cdot \hat{\mathbf{f}}(\mathbf{s}, \omega), \quad (1.17)$$

where $\boldsymbol{\sigma}(\mathbf{r}, \mathbf{r}', \omega) = \int d\mathbf{s} \mathbf{K}(\mathbf{r}, \mathbf{s}, \omega) \cdot \mathbf{K}^\dagger(\mathbf{s}, \mathbf{r}', \omega)$, and the bosonic operators obey standard commutation relations,

$$\left[\hat{\mathbf{f}}(\mathbf{r}, \omega), \hat{\mathbf{f}}^\dagger(\mathbf{r}', \omega') \right] = \delta(\omega - \omega') \delta(\mathbf{r} - \mathbf{r}') \quad (1.18)$$

The interpretation of these new bosonic operators is that they represent the fundamental dipolar, polaritonic excitations of the coupled light-matter system.

The Hamiltonian of the coupled light-matter system can also be found. The time dynamics of the operators is governed by the Heisenberg equation of motion. In the frequency domain the Heisenberg equation of motion reads,

$$\hbar\omega \hat{\mathbf{j}}_N(\mathbf{r}, \omega) = \left[\hat{\mathbf{j}}_N(\mathbf{r}, \omega), \hat{H} \right]. \quad (1.19)$$

For equation 1.19 to hold the Hamiltonian must have the following form:

$$\hat{H} = \pi \int_0^\infty d\omega \int d\mathbf{r} d\mathbf{s} \hat{\mathbf{j}}_N^\dagger(\mathbf{r}, \omega) \cdot [\boldsymbol{\sigma}(\mathbf{r}, \mathbf{s}, \omega)]^{-1} \cdot \hat{\mathbf{j}}_N(\mathbf{s}, \omega) \quad (1.20)$$

which in terms of the renormalized bosonic operators gives,

$$\hat{H} = \int_0^\infty d\omega \hbar\omega \int d\mathbf{r} \hat{\mathbf{f}}^\dagger(\mathbf{r}, \omega) \cdot \hat{\mathbf{f}}(\mathbf{r}, \omega), \quad (1.21)$$

Equation 1.21 represents a very important result, namely the combined matter-light Hamiltonian written in terms of an infinite set of harmonic oscillators. Comparing to Eq. 1.8, it is seen that the expansion in terms of a discrete set of harmonic oscillators is replaced by a continuous one when losses are present.

Finally, the electromagnetic field operators can be recast in terms of the fundamental polaritonic operators,

$$\hat{\mathbf{E}}(\mathbf{r}, \omega) = i\mu_0\omega\sqrt{\frac{\hbar\omega}{\pi}} \int d\mathbf{r}' \int d\mathbf{s} \mathbf{G}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{K}(\mathbf{r}', \mathbf{s}, \omega) \cdot \hat{\mathbf{f}}(\mathbf{s}, \omega), \quad (1.22)$$

$$\hat{\mathbf{B}}(\mathbf{r}, \omega) = \mu_0\sqrt{\frac{\hbar\omega}{\pi}} \nabla \times \int d\mathbf{r}' \int d\mathbf{s} \mathbf{G}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{K}(\mathbf{r}', \mathbf{s}, \omega) \cdot \hat{\mathbf{f}}(\mathbf{s}, \omega). \quad (1.23)$$

The Schrödinger picture operators can be defined as,

$$\hat{\mathbf{E}}(\mathbf{r}) = \int_0^\infty d\omega \hat{\mathbf{E}}(\mathbf{r}, \omega) + \text{h.c.}, \quad (1.24)$$

$$\hat{\mathbf{B}}(\mathbf{r}) = \int_0^\infty d\omega \hat{\mathbf{B}}(\mathbf{r}, \omega) + \text{h.c.}, \quad (1.25)$$

and they can be shown to obey a commutator in compliance with the correspondence principle,

$$\left[\hat{\mathbf{E}}(\mathbf{r}), \hat{\mathbf{B}}(\mathbf{r}') \right] = -\frac{i\hbar}{\epsilon_0} \nabla \times \delta^\perp(\mathbf{r} - \mathbf{r}'). \quad (1.26)$$

A final note is that the electromagnetic fields $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$ only represent the degrees of freedom of the electromagnetic field, which is further confirmed by their commutation relation. This seems to somehow contradict the fact that they are now expanded in terms of the noise current or polaritonic operators which are clearly not solely operators associated with the field. To resolve this apparent paradox, the DGF is to be interpreted as the integral kernel of a projection operator projecting from the Hilbert space of the coupled system onto the part of the Hilbert space associated with the electromagnetic field[27].

1.2 Interaction of the electromagnetic field and matter

With the general theory of the quantized electromagnetic field in lossy media at hand, it is now possible to discuss the interaction between the quantized electromagnetic field and emitters. It turns out that the exact form of the system Hamiltonian depends strongly on the choice of gauge. While all gauges are formally equivalent, different representations can be advantageous in different situations. As one example, it has been shown that different gauges show different resilience to the truncation of the electronic Hilbert space that always accompany calculations for practically relevant systems[32–34].

The following discusses two relevant choices of gauge: the Coulomb gauge and the Multipolar or Power-Zienau-Woolley (PZW) gauge.

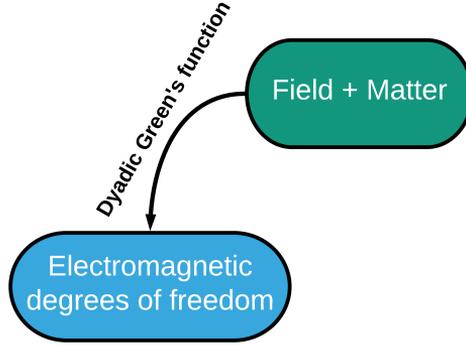


Figure 1.2: Illustration of the Dyadic Green's Function as the projector from the combined light matter system of MQED onto the electromagnetic degrees of freedom.

The Minimal Coupling Hamiltonian in the Coulomb gauge: The first example is the Coulomb gauge that was used in section 1.1.1. Since $\nabla \cdot \mathbf{A} = 0$, the vector potential describes the transverse components of the field, while the scalar potential, which will now generally be nonzero, describes the longitudinal components of the field. This construction results in the following Hamiltonian for the coupled system,

$$\hat{H} = \frac{1}{2} \int d\mathbf{r} \left(\epsilon_0 \hat{\mathbf{E}}(\mathbf{r})^2 + \frac{1}{\mu_0} \hat{\mathbf{B}}(\mathbf{r})^2 \right) + \sum_{\alpha} \frac{1}{2m_{\alpha}} \left[\hat{\mathbf{p}}_{\alpha} - q_{\alpha} \hat{\mathbf{A}}(\mathbf{r}_{\alpha}) \right]^2 + \sum_{\alpha \neq \beta} \frac{q_{\alpha} q_{\beta}}{8\pi \epsilon_0 |\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}|}. \quad (1.27)$$

Here the sums run over the charges in the system, \mathbf{r}_{α} denotes the position of particle α and q_{α} denotes the charge of particle α . In Eq. 1.27, all the microscopic degrees of freedom of the electronic systems are explicit and for this reason it is often called the minimal coupling scheme. The fact that the microscopic degrees of freedom are explicit in the Hamiltonian means that this representation can sometimes be a bit cumbersome to deal with in practice. However, the explicit microscopic nature of the Hamiltonian also means that it is well suited for rapidly varying fields.

The Multipolar representation in the Power-Zienau-Wolley gauge: Another popular choice of gauge is the Power-Zienau-Wolley (PZW) gauge, also sometimes known as the multipolar gauge[35–37]. It expresses the light-matter interactions in terms of the physical fields, \mathbf{E} and \mathbf{B} , instead of the potentials, and it allows for a systematic expansion of the fields in terms of multipole moments. Furthermore, if the matter can be thought of as made up of multiple, separate ensembles of charge (e.g. molecules), the Coulomb interaction between these molecules is removed from the microscopic Hamiltonians of the ensembles and instead carried by the field. As an added benefit, the representation has proven to be very robust to the effects of truncation of

the electronic Hilbert space[32]. For these reasons, the PZW gauge is particularly well suited when studying the coupling between the quantized electromagnetic field and molecules.

The PZW transformation is a formally a canonical transformation that is carried out by applying the following unitary transformation to Eq. 1.27,

$$\hat{U} = \exp \left[\frac{i}{\hbar} \int d\mathbf{r} \sum_i \hat{\mathbf{P}}_i(\mathbf{r}) \cdot \hat{\mathbf{A}}(\mathbf{r}) \right]. \quad (1.28)$$

Here $\hat{\mathbf{P}}_i(\mathbf{r})$ is the polarization operator of the i 'th ensemble of charge. Neglecting magnetic interactions, this leads to the following Hamiltonian,

$$\begin{aligned} \mathcal{H} = & \sum_i \left(\sum_{\alpha \in i} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \frac{1}{2\epsilon_0} \int d\mathbf{r} \hat{\mathbf{P}}_i(\mathbf{r})^2 \right) + \\ & \frac{1}{2} \int d\mathbf{r} \left(\epsilon_0 \hat{\mathbf{E}}(\mathbf{r})^2 + \frac{1}{\mu_0} \hat{\mathbf{B}}(\mathbf{r})^2 \right) - \\ & \sum_i \int d\mathbf{r} \hat{\mathbf{P}}_i(\mathbf{r}) \cdot \hat{\mathbf{D}}(\mathbf{r}_i). \end{aligned} \quad (1.29)$$

Importantly, the coupling between the electromagnetic field and the matter is in terms of the displacement field $\hat{\mathbf{D}}$ and *not* the electric field $\hat{\mathbf{E}}$ [36]. However, in the following it will nonetheless be denoted by $\hat{\mathbf{E}}$ to follow standard conventions. The PZW gauge is particularly convenient when the coupling to the electromagnetic field can be described within the dipole approximation. In this case, the interaction Hamiltonian can be written as,

$$H_{\text{int}} \approx - \sum_i \hat{\mathbf{d}}_i \cdot \hat{\mathbf{E}}(\mathbf{r}_i). \quad (1.30)$$

Specifying to electrons, it often makes sense to write the dipole moment as $\hat{\mathbf{d}}_i = -e\hat{\mathbf{R}}_i$ where $\hat{\mathbf{R}} = \sum_j \hat{\mathbf{r}}_j$ is the center of mass position of the charges.

The representation of the matter is also affected by the PZW transformation. Specifically, the matter Hamiltonian reads,

$$\mathcal{H}_{\text{Mat}} = \sum_{\alpha} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \frac{1}{2\epsilon_0} \int d\mathbf{r} \hat{\mathbf{P}}(\mathbf{r})^2. \quad (1.31)$$

The polarization field can be split into a longitudinal part and a transverse part. The longitudinal contribution to the \mathbf{P}^2 term corresponds to the usual Coulomb interaction between the charged particles of each ensemble, while the transverse part of the polarization is an additional term. The matter Hamiltonian is thus modified relative to its form in the Coulomb gauge by the transverse polarization term,

$$\mathcal{H}_{\text{Mat}} = \mathcal{H}_{\text{Mat}}^{\text{C}} + \frac{1}{2\epsilon_0} \int d\mathbf{r} \left[\hat{\mathbf{P}}_{\perp}(\mathbf{r}) \right]^2, \quad (1.32)$$

where $\mathcal{H}_{\text{Mat}}^{\text{C}}$ denotes the usual Coulomb gauge matter Hamiltonian. Generally, the transverse polarization term should be derived from the full polarization term via a transverse projection[36], $\hat{\mathbf{P}}_{\perp}(\mathbf{r}) = \frac{2}{3}\hat{\mathbf{P}}(\mathbf{r}) + \int d\mathbf{r}' \hat{\mathbf{T}}_{\perp}(\mathbf{r}, \mathbf{r}') \cdot \hat{\mathbf{P}}(\mathbf{r}')$. This projection is, however, quite computationally cumbersome. Alternatively, for a single ensemble coupled to a lossless magnetoelectric background within the dipole approximation, the transverse polarization term can be expanded in terms of the mode functions of the electromagnetic field expansion in Eq. 1.5 and the resulting Hamiltonian of the coupled system is[38],

$$\mathcal{H} = \mathcal{H}_{\text{Mat}}^{\text{C}} + \frac{1}{2} \sum_n \left(\hat{p}_n^2 + \omega_n^2 \left[\hat{q}_n - \frac{\boldsymbol{\lambda}_n \cdot \hat{\mathbf{R}}}{\omega_n} \right]^2 \right), \quad (1.33)$$

where $\boldsymbol{\lambda}_n$ is a scaled version of the mode functions of the field expansion in Eq. 1.5, $\boldsymbol{\lambda}_n = -\sqrt{\frac{2}{\hbar\omega_n}} e\mathbf{E}_n(\mathbf{R})$. The Hamiltonian of the coupled system can therefore be represented as the Coulomb gauge Hamiltonian of the matter coupling the harmonic oscillator modes of the quantized electromagnetic field. The coupling manifests as a displacement of the equilibrium of the harmonic oscillator mode proportional to the product of the dipole moment and the modes field strength at the emitter position. This form of the Hamiltonian is the foundation of most of the literature with Quantum Electrodynamical Density Functional Theory (QEDFT)[38, 39] which will be discussed extensively in the following Chapters.

1.2.1 The emitter centered representation of Macroscopic Quantum Electrodynamics in the dipole approximation

This section discusses a particularly convenient representation of the MQED electric field expansion which is appropriate when the coupling to the field can be described within the dipole approximation[40–42]. The starting point for the discussion is the PZW gauge Hamiltonian for a single emitter coupled to the cavity field within the dipole approximation. Specializing to a spatially local magnetoelectric medium and neglecting magnetic interactions, the Hamiltonian reads,

$$\begin{aligned} \mathcal{H} = & \sum_{\alpha} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \frac{1}{2\epsilon_0} \int d\mathbf{r} \hat{\mathbf{P}}(\mathbf{r})^2 \\ & + \sum_{\lambda} \int_0^{\infty} d\omega \hbar\omega \int d\mathbf{r} \hat{\mathbf{f}}_{\lambda}^{\dagger}(\mathbf{r}, \omega) \cdot \hat{\mathbf{f}}_{\lambda}(\mathbf{r}, \omega) - \hat{\mathbf{d}} \cdot \hat{\mathbf{E}}(\mathbf{r}_0), \end{aligned} \quad (1.34)$$

where the λ index denotes the electric and magnetic contributions respectively. The MQED expansion of the electric field can be written as[27],

$$\hat{\mathbf{E}}(\mathbf{r}) = \int d\omega \hat{\mathbf{E}}(\mathbf{r}, \omega) + \text{h.c.}, \quad (1.35)$$

$$\hat{\mathbf{E}}(\mathbf{r}, \omega) = \sum_{\lambda=e,m} \int d\mathbf{r}' \mathbf{G}_{\lambda}(\mathbf{r}, \mathbf{r}', \omega) \cdot \hat{\mathbf{f}}_{\lambda}(\mathbf{r}', \omega). \quad (1.36)$$

In the current form, the Hamiltonian is written in terms of the spatially resolved polaritonic operators from MQED, $\hat{\mathbf{f}}_\lambda(\mathbf{r}', \omega)$. Because the coupling is described within the dipole approximation, the field is only sampled at the emitter position, \mathbf{r}_0 , and the current formulation contains more information than necessary. Therefore, it makes sense to seek a simplified representation by constructing new modes such that at each frequency a minimum numbers of modes couple to the emitter at position \mathbf{r}_0 with dipole orientation $\hat{\mathbf{n}}$.

For a given dipole moment orientation and frequency, there will be a set of "bright modes" which couple to dipoles with that orientation, and a set of "dark modes" which do not. Which modes couple to a given transition is determined by the DGF. In particular, the bright modes can be represented by a new set of operators which are written as linear superpositions of the fundamental bosonic modes from MQED,

$$\hat{B}_i(\omega) = \sum_\lambda \int d\mathbf{r} \beta_{i,\lambda}(\mathbf{r}, \omega) \cdot \hat{\mathbf{f}}_\lambda(\mathbf{r}, \omega), \quad (1.37)$$

$$\beta_{i,\lambda}(\mathbf{r}, \omega) = \frac{\mathbf{n}_i \cdot \mathbf{G}_\lambda(\mathbf{r}_i, \mathbf{r}, \omega)}{G_i(\omega)}, \quad (1.38)$$

where $i = x, y, z$ denotes spatial orientation. The normalisation constant, $G_i(\omega)$, is chosen such that $[\hat{B}_i(\omega), \hat{B}_j^\dagger(\omega')] = S_{ij} \delta(\omega - \omega')$ with $S_{ii} = 1$ and this leads to,

$$G_i(\omega) = \left(\frac{\hbar \omega^2}{\pi \epsilon_0 c^2} \mathbf{n}_i \cdot \text{Im} \mathbf{G}(\mathbf{r}_i, \mathbf{r}_i, \omega) \cdot \mathbf{n}_i \right)^{1/2}. \quad (1.39)$$

Note that $G_i(\omega)$ is almost the local spectral density of a dipole coupled to the electromagnetic environment (see section 1.4). The bright modes associated with different dipole orientations are not generally orthogonal because their overlap is related to the symmetry of the DGF,

$$S_{ij}(\omega) = \sum_\lambda \int d\mathbf{r} \beta_{i,\lambda}^*(\mathbf{r}, \omega) \cdot \beta_{j,\lambda}(\mathbf{r}, \omega) = \frac{\hbar \omega^2}{\pi \epsilon_0 c^2} \frac{\mathbf{n}_i \cdot \text{Im} \mathbf{G}(\mathbf{r}_i, \mathbf{r}_j, \omega) \cdot \mathbf{n}_j}{G_i(\omega) G_j(\omega)}. \quad (1.40)$$

It is seen that $S_{ii} = 1$, but that in general the off-diagonal elements can be non-zero. To arrive at a set of bosonic modes with appropriate commutation relations, the modes need to be explicitly orthogonalized. A new set of modes, $\hat{C}_i(\omega)$ can be defined as linear superpositions of the bright modes,

$$\hat{C}_i(\omega) = \sum_j V_{ij}(\omega) \hat{B}_j(\omega), \quad (1.41)$$

where the transformation matrix is chosen such that $[\hat{C}_i(\omega), \hat{C}_j^\dagger(\omega')] = \delta_{ij} \delta(\omega - \omega')$. There are several ways to perform such a orthogonalization such as e.g. Gram-Schmidt. The original MQED bosonic modes can then be expanded in terms of the active modes

and a potentially infinite modes of "dark" modes, $D_j(\omega)$,

$$\hat{\mathbf{f}}_\lambda(\mathbf{r}, \omega) = \sum_i^N \chi_{i,\lambda}^*(\mathbf{r}, \omega) \hat{C}_i(\omega) + \sum_j \mathbf{d}_{j,\lambda}^*(\mathbf{r}, \omega) \hat{D}_j(\omega), \quad (1.42)$$

$$\chi_{i,\lambda}(\mathbf{r}, \omega) = \sum_j V_{ij}(\omega) \beta_{j,\lambda}(\mathbf{r}, \omega), \quad (1.43)$$

$$\int d\mathbf{r} \mathbf{d}_{j,\lambda}^*(\mathbf{r}, \omega) \cdot \chi_{i,\lambda}(\mathbf{r}, \omega) = 0. \quad (1.44)$$

Importantly, it was recently noticed that the electric field operator could also be expanded in terms of the bright modes[40],

$$\hat{\mathbf{E}}(\mathbf{r}) = \sum_i \int_0^\infty d\omega \mathbf{E}_i(\mathbf{r}, \omega) \hat{C}_i(\omega) + \text{h.c.}, \quad (1.45)$$

$$\mathbf{E}_i(\mathbf{r}, \omega) = \frac{\hbar\omega^2}{\pi\epsilon_0 c^2} \sum_j V_{ij}^*(\omega) \frac{\text{Im}\mathbf{G}(\mathbf{r}, \mathbf{r}_0, \omega) \cdot \hat{n}_i}{G_i(\omega)}. \quad (1.46)$$

Note that here the dark modes have been removed from the expansion of the electric field since these do not play a role for the light matter coupling. Since i runs over the different dipole orientations, this provides the most compact representation of the MQED electric field that contains all the information necessary to describe light-matter interactions within the dipole approximation.

1.3 Regimes of Quantum Light Matter Interactions

This section will discuss the different regimes of light-matter interactions that can be realized in c-QED. In principle, the starting point for computational non-relativistic QED should be the Hamiltonians outlined above. However, to discuss the potential effects that can arise in c-QED, it makes sense to consider the simpler case of a two-level electronic system coupled to a single cavity mode with the dipole approximation. In this limit, the problem reduces to the Rabi Hamiltonian which in the multipolar gauge can be cast as[5],

$$\hat{H} = \frac{1}{2} \hbar\omega_e \hat{\sigma}_z + \hbar\omega_c \hat{a}^\dagger \hat{a} + \hbar g (\hat{a}^\dagger + \hat{a}) \hat{\sigma}_x. \quad (1.47)$$

Here ω_e is the electronic transition frequency, and ω_c is the cavity frequency. The strength of the light matter interaction is characterized by of the so-called Rabi frequency,

$$g = \frac{1}{\hbar} \mathbf{d} \cdot \mathbf{E}, \quad (1.48)$$

where \mathbf{d} is the transition dipole moment of the electronic transition, and \mathbf{E} is the electric field strength at the position of the emitter. Due to the vacuum fluctuations,

the interaction between light and matter will persist even in the absence of actual photons in the cavity. At the vacuum level, the strength of the field fluctuations in a general electromagnetic environment can be calculated using MQED³[27],

$$\langle 0 | \hat{E}_k(\mathbf{r}, \omega) \hat{E}_l^\dagger(\mathbf{r}', \omega') | 0 \rangle = \frac{\hbar \omega^2}{\pi \epsilon_0 c^2} \text{Im} G_{kl}(\mathbf{r}, \mathbf{r}', \omega) \delta(\omega - \omega'). \quad (1.49)$$

In the anti-node of a simple cavity, the fluctuation strength can be recast as[5],

$$|\langle \hat{E}^2 \rangle_{\text{vac}}|^{1/2} = \sqrt{\frac{\hbar \omega_c}{2 \epsilon_0 V}}, \quad (1.50)$$

where V is the cavity mode volume. This simple example shows that there are essentially two ways to increase the coupling strength: either by choosing a transition with a large transition dipole moment, or by decreasing the cavity mode volume. This reflects the fact that the coupling is dependent on both the properties of the emitter and of the field, which explains why properties such as spontaneous emission can be altered by engineering the electromagnetic environment[43]. The inverse relationship between the mode volume and the coupling strength further explains the wide use of plasmonic resonators to achieve strong light matter interactions[44]. Other important parameters are the cavity linewidth, κ , and the linewidth of the emitter Γ . Depending on the relative sizes of these parameters, the nature of the light matter coupling can be very different. The coupling in c-QED can be broadly separated into two categories, the weak coupling regime and the strong coupling regime.

Weak coupling regime: In the weak coupling regime, the losses dominate over the coupling, $g < \frac{1}{\sqrt{2}}(\kappa^2 + \Gamma^2)^{\frac{1}{2}}$. For this reason this is sometimes referred to as the bad

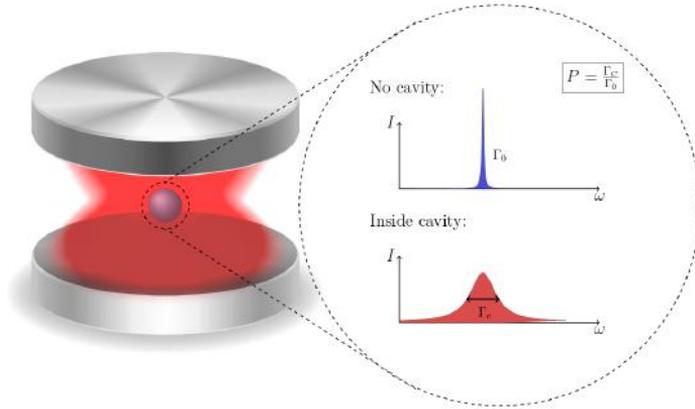


Figure 1.3: Illustration of an emitter weakly coupled to an optical cavity.

³Note the similarities with the coefficients of the field expansion in Eq. 1.46.

cavity regime. In this regime, the light and matter are unable to coherently interact, and consequently no hybridization of the light and matter takes place. Instead, the interaction between the light and matter leads to alterations of the emissions rates of the electronic system via the Purcell effect[43, 45]. To understand this, consider the case where the cavity losses are dominant and the cavity width is larger than the coupling strength, $g \ll \kappa$. In this case, the emitter sees the cavity mode as a broad continuum and the energy transfer from the emitter to the cavity mode is thus irreversible. However, because the local density of states of the electromagnetic field is enhanced due to the presence of the cavity, and the emission rate is proportional to the local density of states of the electromagnetic field[44], the emission rate of the emitter is enhanced. This is illustrated in Fig. 1.3. Similarly, if the emitter is not resonant with the cavity, the emission rate is suppressed due to the lower density of states as compared to vacuum. For completeness, it is noted that the case where the intrinsic linewidth dominates, $\Gamma \gg g, \kappa$, would also be considered weak coupling. However, this case is not very interesting because the emission into the far field would dominate over the interaction with the cavity and no significant effects of the cavity would be observable.

c-QED in the weak coupling regime is widely used in applications, for example for low threshold solid state lasers [46] and improving the coherence properties of single photon emitters[47].

Strong coupling regime: When increasing the coupling strength, the system starts to enter the so-called strong coupling regime which is characterized by coherent interactions between the matter and cavity photons, This hybridization manifests as an avoided crossing leading to the appearance of two new hybrid light matter states called polaritons[5, 46]. As illustrated in Fig. 1.4, these new polariton states are separated in energy by $2\hbar g$. The onset of strong coupling is not entirely well defined, but it

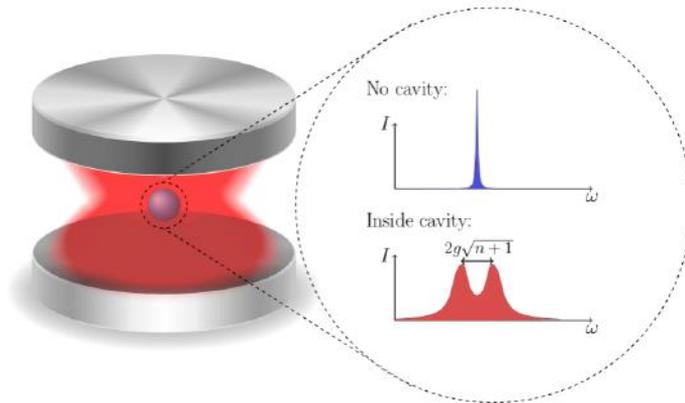


Figure 1.4: Illustration of an emitter strongly coupled to an optical cavity. Here n denotes the photon number, highlighting that these effects persist even in the limit where the expectation value of the photon number vanishes.

is customarily taken as the coupling strength at which the splitting can be resolved experimentally, corresponding roughly to the criterion that $g > \frac{1}{\sqrt{2}}(\kappa^2 + \Gamma^2)^{\frac{1}{2}}$ [5]. The proper theoretical description of the strong coupling regime depends on the ratio of the coupling strength and the transition frequency, $\eta = g/\omega_e$. In fact, it makes sense to further subdivide the strong coupling regime into the Strong Coupling (SC) regime ($\eta < 0.1$), the Ultra Strong Coupling regime (USC) ($0.1 < \eta < 1$), and the Deep Strong Coupling (DSC) regime ($\eta > 1$) [46].

In the SC regime, the so-called Rotating Wave Approximation (RWA) can be performed and the system can be described by the Jaynes-Cummings Model (JCM) [48],

$$\hat{H} = \frac{1}{2}\hbar\omega_e\hat{\sigma}_z + \hbar\omega_c\hat{a}^\dagger\hat{a} + \hbar g(\hat{a}^\dagger\hat{\sigma}_- + \hat{a}\hat{\sigma}_+). \quad (1.51)$$

The big difference between the JCM and the Rabi model is the interaction term. After the RWA, the electron-photon interaction becomes number conserving in terms of the combined number of electronic and photonic excitations. This has profound consequences because it means that while alteration of the excited states is possible within the JCM, it is impossible for the photon field to alter the ground state properties of the system [5, 46]. While the SC regime can be reached with single molecules [49], it is more commonly reached by increasing the number of emitters. In this case, the coupling scales as $\sqrt{N}g$ and it is therefore easier to reach SC with an ensemble of emitters [5, 46].

The simple JCM picture begins to change when the system enters the USC, at which point the RWA breaks down and it becomes necessary to include multi-photon states in the description. The USC also signals the regime where the ground state of the coupled systems gets significantly dressed by photons and where virtual photon states are needed for the description of the system [46]. As the coupling is increased further towards the DSC regime, these effects are enhanced and the description is complicated further. Importantly, in this regime it is also generally not sufficient to describe the emitter by simple few level approximations, and it becomes necessary to employ more sophisticated models of the electronic system that can account for many electronic states. For completeness, it makes sense to mention that the coupling strength at which the breakdown of the few level approximations happens is gauge dependent again highlighting the intricacies of the gauge freedom in computational QED [32, 34]. Platforms to realize the USC and DSC regimes include intersubband polaritons [50], superconducting circuits [51] and Landau polaritons [52, 53].

1.4 The Dyadic Green's Function

The Dyadic Green's function (DGF) is the Green's function of the electromagnetic field [54]. It is defined as the Green's function of the Helmholtz equation,

$$\left[\nabla \times \nabla \times - \frac{\omega^2}{c^2} \right] \mathbf{G}(\mathbf{r}, \mathbf{r}', \omega) - i\mu_0\omega \int d\mathbf{s} \mathbf{Q}(\mathbf{r}, \mathbf{s}, \omega) \mathbf{G}(\mathbf{s}, \mathbf{r}, \omega) = \delta(\mathbf{r} - \mathbf{r}'). \quad (1.52)$$

Physically, the DGF represents the electric field at point \mathbf{r} due to a point source at point \mathbf{r}' , and given the DGF the field from an arbitrary source current, $\mathbf{j}_{\text{ext}}(\mathbf{r}')$, can be

written as,

$$\mathbf{E}(\mathbf{r}) = i\omega\mu_0 \int \mathbf{G}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{j}_{\text{ext}}(\mathbf{r}') d\mathbf{r}'. \quad (1.53)$$

The DGF is a very important object in electromagnetism in general and it is the basis for many numerical methods in computational electromagnetism[43]. For the purposes of this thesis, the main interest in the DGF stems from its central role in MQED. In this case, the central quantity is really the imaginary part of the DGF which holds the information about the local density of states for the electromagnetic field and consequently it describes the coupling of an emitter to the vacuum field. Specifically, the local density of states for a dipole orientation $\hat{\mathbf{n}}$ is given as,

$$\rho(\mathbf{r}, \omega) = \frac{6\omega}{\pi c^2} \{ \hat{\mathbf{n}} \cdot \text{Im} [\mathbf{G}(\mathbf{r}, \mathbf{r}, \omega)] \cdot \hat{\mathbf{n}} \}. \quad (1.54)$$

This can be better understood by noting that what drives light-matter interactions at the zero photon level are the vacuum fluctuations of the field. This is for example the ultimate origin of spontaneous emission. To understand how this is related to the DGF, it makes sense to consider the vacuum field fluctuations of the electric field which can be calculated in terms of the MQED expansion of the electric field[27],

$$\langle 0 | \hat{E}_k(\mathbf{r}, \omega) \hat{E}_l^\dagger(\mathbf{r}', \omega') | 0 \rangle = \frac{\hbar\omega^2}{\pi\epsilon_0 c^2} \text{Im} G_{kl}(\mathbf{r}, \mathbf{r}', \omega) \delta(\omega - \omega'). \quad (1.55)$$

Eq. 1.55 tell us that the fluctuations of the electromagnetic field are given directly by the DGF. Since the ultimate origin of coupling between the electromagnetic environment and an electronic system at the vacuum level is the vacuum fluctuations of the field, this explains the central role of the DGF in the MQED formulation in particular, and the description of quantum light-matter interactions in general.

1.4.1 Calculating the Dyadic Green's function

The central question in practice is how to calculate the DGF for the cavity setup of interest. In general, the trick is to find suitable basis functions to in terms of which to expand the DGF.

The simplest possible case is vacuum or a homogeneous medium. In this case, the DGF is expanded onto spherical waves centered at the dipole position and it is usually written in the following way[43, 54],

$$\mathbf{G}^{\text{vac}}(\mathbf{r}, \mathbf{r}', \omega) = \left[\frac{\nabla \otimes \nabla}{k_0^2} - \mathbb{1} \right] \frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}. \quad (1.56)$$

At this point, one notices that Eq. 1.56 looks frighteningly singular at $\mathbf{r} = \mathbf{r}'$. This would be very problematic, because it would mean that the local density of states in Eq. 1.54 would blow up everywhere in free space and this would mean that coupling between emitters and the electromagnetic field would always be infinite which is clearly unphysical. However, with some algebra, it turns out that the singularity only affects

the real part of the DGF and that the imaginary part is well behaved at $\mathbf{r} = \mathbf{r}'$ [55]. In fact, it can be shown that the imaginary part of the vacuum DGF at $\mathbf{r} = \mathbf{r}'$ is[55],

$$\text{Im}\mathbf{G}^{\text{vac}}(\mathbf{r}, \mathbf{r}, \omega) = \frac{\omega}{6\pi c} \mathbf{1}. \quad (1.57)$$

Plugging this into Eq. 1.54 gives the spatially uniform free space density of states of the electromagnetic field,

$$\rho^{\text{vac}}(\mathbf{r}, \omega) = \frac{\omega^2}{\pi^2 c^3}. \quad (1.58)$$

Note that the singular nature of the real part of the DGF is a general feature of DGFs and great care has to be taken to deal with this singularity when using the DGF for numeric electromagnetism[43, 54]. However, the statement that the singularity does not affect the imaginary part is generally true. Since this thesis only ever explicitly deals with the imaginary part of the DGF we will not discuss this further. For the interested reader it is noted that exhaustive discussion of the singularity can be found in Refs. [43] and [54].

The problem is significantly more challenging if one considers structured media. In general, no closed solution solution for the DGF exist and numerical methods have to be employed. However, in some high-symmetry geometries it is possible to write down concrete expressions for the DGF, specifically layered media[54]. In these cases, the DGF can be expanded in terms of appropriate vector wave functions that capture the symmetry of the problem. Two such layered examples are relevant for the work in this thesis: a planarly layered configuration (Publication [I]) and a spherically layered configuration (see Chapter 4).

For planarly layered media, the DGF is expanded onto a set of in-plane plane waves with wave vector \mathbf{q} , modulated by a z -dependent function that accounts for the reflection between the layers,

$$\mathbf{G}(\mathbf{r}, \mathbf{r}') = \frac{i}{8\pi^2} \int \frac{d^2q}{k_z q^2} [\mathbf{M}(\mathbf{q}, \mathbf{r}, \mathbf{r}') + \mathbf{N}(\mathbf{q}, \mathbf{r}, \mathbf{r}')] - \frac{\hat{\mathbf{z}} \otimes \hat{\mathbf{z}}}{k^2} \delta(\mathbf{r} - \mathbf{r}'). \quad (1.59)$$

The DGF is split into transverse electric (TE), $\mathbf{M}(\mathbf{q}, \mathbf{r}, \mathbf{r}')$, and transverse magnetic (TM), $\mathbf{N}(\mathbf{q}, \mathbf{r}, \mathbf{r}')$, components, as these are the natural basis for reflection of planar interfaces. There is no closed general expression for the \mathbf{M} and \mathbf{N} tensors, and for a given problem these have to be determined by applying the electromagnetic boundary conditions at each interface in the problem. For an example of this, see Publication [I] and its Supplementary Information.

The second relevant example for the purposes of this thesis is that of a spherically layered medium. In a spherically layered medium, the DGF is expanded onto an infinite number of vector spherical harmonics[54],

$$\mathbf{G}(\mathbf{r}, \mathbf{r}', \omega) = ik \sum_n^{\infty} \frac{1}{n(n+1)} [\mathbf{E}_n(\mathbf{r}, \mathbf{r}', \omega) + \mathbf{M}_n(\mathbf{r}, \mathbf{r}', \omega)] - \frac{\hat{\mathbf{r}} \otimes \hat{\mathbf{r}}}{k^2} \delta(\mathbf{r} - \mathbf{r}'). \quad (1.60)$$

It is again natural to split into electric and magnetic type modes, \mathbf{E}_n and \mathbf{M}_n respectively. Again, for a given configuration, the electromagnetic boundary conditions are used to determine the exact form of \mathbf{E}_n and \mathbf{M}_n . For an example, see Chapter 4. Note that both Eq. 1.60 and Eq. 1.59 contain singularities at $\mathbf{r} = \mathbf{r}'$ which have been explicitly extracted.

For more complex geometries, it is not straightforward to write down the proper mode functions. In such cases, it is necessary to determine the mode functions numerically. In general for lossy backgrounds, the Helmholtz equation is a non-Hermitian operator and special care has to be taken when determining the appropriate mode functions. One example of how this can be done is with the biorthonormal mode construction discussed in Ref. [56].

CHAPTER 2

The Many-Body Problem

The previous chapter introduced the interacting light-matter problem. However, writing down the equations is a far easier task than actually solving them for practically relevant systems. This chapter introduces Density Functional Theory (DFT) which is a practical approach for solving the interacting many-body problem for realistic systems. The chapter starts with a discussion of ground state DFT and how it can be used to accurately describe the ground state properties of an interacting many electron system. Subsequently, the generalization of DFT to time dependent DFT (TDDFT) is introduced. Because a major focus of this thesis is the interaction between quantized light and matter, the generalization of TDDFT to the presence of quantized light, the so-called Quantum Electrodynamics Density Functional Theory (QEDFT), is the focus of this discussion.

2.1 Ground state Density Functional Theory

2.1.1 The many electron problem

The starting point for the discussion of ground state DFT is the stationary many-body Schrödinger equation for the interacting electron-nuclear problem,

$$\hat{H}\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M). \quad (2.1)$$

Here Ψ is the many-body wave function, E is its energy, and $\{\mathbf{r}_n\}$ and $\{\mathbf{R}_m\}$ are the sets of electronic- and nuclear positions respectively. The interactions in the problem are given by the Coulomb interaction between the charged particles. The electron-nuclear wave function, Ψ , is a function of $3(N + M)$ variables and it very quickly becomes unfeasible to solve it directly even for simple systems. Therefore, to solve Eq. 2.1 in practice simplifications are needed.

The first commonly applied simplification is the Born-Oppenheimer approximation[57]. It assumes that due to the significantly smaller mass of the electrons, their response to changes in the nuclear positions will be adiabatic such that the two are uncorrelated. An uncorrelated state can be written as a product state, and within this approximation it becomes possible to decouple the electronic problem from the nuclear one, $\Psi(\{\mathbf{r}_n\}, \{\mathbf{R}_m\}) \approx \chi(\{\mathbf{R}_m\})\psi(\{\mathbf{r}_n\}; \{\mathbf{R}_m\})$. Here χ is the nuclear wave function which now solely depends on the nuclear coordinates, and ψ is the electronic wave function which depends on the electronic coordinates and parametrically on the nuclear coordi-

notes. In the end, the electronic Hamiltonian can be written as,

$$\hat{H}_{\text{el}} = \sum_i^N \frac{\hat{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} - \sum_i^N \sum_I^M \frac{Z_I e^2}{|\hat{\mathbf{r}}_i - \mathbf{R}_I|.} \quad (2.2)$$

The first term represents the kinetic energy of the electrons, the second represents the electron-electron interaction, and the final term represents the potential felt by the electrons due to the now static nuclei. Within the Born-Oppenheimer approximation, the role of the nuclei is therefore to set up a potential felt by the electrons which can be considered equivalent to an external potential, \hat{V}_{ext} .

Even within the Born-Oppenheimer approximation, solving for the N -electron wave function is a daunting task, as the size of the wave function still grows exponentially with the number of electrons in the system. It therefore makes sense to ask if there is a simpler quantity than the wave function that contains the information needed to compute the ground state observables of the system. One answer to this question is Density Functional Theory (DFT) which replaces the many-body wave function with the electronic density,

$$n(\mathbf{r}) = N \int \prod_{i=1}^N d\mathbf{r}_i |\psi(\mathbf{r}_1, \dots, \mathbf{r}, \dots, \mathbf{r}_N)|^2. \quad (2.3)$$

Unlike the many-body wave function, the electronic density only depends on one spatial position \mathbf{r} and is therefore a much simpler object to work with.

2.1.2 The Hohenberg Kohn Theorems

The theoretical foundations of DFT are the two Hohenberg Kohn Theorems[58]. These theorems are the formalization of the Thomas-Fermi conjecture[59, 60] that the ground state energy of a many-body system can be written as a functional of the electronic density.

The first Hohenberg Kohn theorem states that the many-body ground state wave function and the external potential can be written as a unique functional of the electronic density, and any ground state observable can therefore be written as a functional of the electronic density. In particular, the ground state energy of the interacting system can be expressed as a unique functional of the electronic density,

$$E[n_{\text{GS}}] = T[n_{\text{GS}}] + V_{\text{el-el}}[n_{\text{GS}}] + \int d\mathbf{r} v_{\text{ext}}(\mathbf{r}) n_{\text{GS}}(\mathbf{r}). \quad (2.4)$$

The two first terms in Eq. 2.4 are known as the universal part of the functional because they are the same for all systems. The reformulation of the problem in terms of the electronic density is a massive simplification, because unlike the many-body wave function which was a function of $3N$ variables, the electronic density only depends on three variables *regardless* of system size.

The second theorem states that the ground state energy functional satisfies the variational principle with respect to the ground state density,

$$E[n] \geq E[n_{\text{GS}}], \quad \left. \frac{\delta E[n_{\text{GS}}]}{\delta n} \right|_{n=n_{\text{GS}}} = 0. \quad (2.5)$$

This allows for variational optimization of the density which is very important because even though it can be proved that the mapping from the external potential to the ground state density exists, its form is unknown. Importantly, no approximations are needed to arrive at the above results, making DFT an exact ground state theory.

2.1.3 The Kohn-Sham equations

In practice, direct minimization of the energy functional is extremely challenging. This is further complicated by the fact that the universal part of the functional is unknown. An alternative to the direct minimization route is provided by the Kohn-Sham construction, first introduced in 1965[61]. Because the ground state energy is a unique functional of the density, any system with the same density will produce the same energy. This means that if a auxiliary non-interacting system can be constructed to have the same density as the fully interacting system, the ground state energy will be the same for the two systems.

To understand how the Kohn-Sham system is constructed, it is instructive to rewrite Eq. 2.4 as,

$$E[n] = T^{\text{NI}}[n] + V^H[n] + V^{\text{ext}}[n] + E^{\text{xc}}[n] \quad (2.6)$$

The terms in Eq. 2.6 are as follows:

- $T^{\text{NI}}[n]$ is the kinetic energy functional of a non-interacting system of density n . Using this functional is a significant simplification, as the exact kinetic energy functional for an interacting system is unknown.
- $V^H[n]$ is the Hartree potential. The Hartree potential describes the electrostatic repulsion between the charged particles in the system.
- $V^{\text{ext}}[n]$ is the external potential containing all the information about the position and charges of the nuclei, as well as any externally applied perturbations.
- $E^{\text{xc}}[n]$ is the exchange-correlation term which will be discussed further in the following.

For Eqs. 2.4 and 2.6 to be equivalent, the definition of the exchange-correlation terms has to be,

$$E^{\text{xc}}[n] = T[n] - T^{\text{NI}}[n] - V^H[n] + V^{\text{el-el}}[n]. \quad (2.7)$$

While formally no different from Eq. 2.4, Eq. 2.6 has the great conceptual advantage that all of the difficult and unknown terms are grouped into $E^{\text{xc}}[n]$. There is now thus only a single term that needs to be approximated.

For a non-interacting system, the energy functional should take the following form,

$$E^{\text{NI}}[n] = T^{\text{NI}}[n] + V^{\text{KS}}[n]. \quad (2.8)$$

Here $V^{\text{KS}}[n]$ is the potential energy functional of the non-interacting system and it can be written as,

$$V^{\text{KS}}[n] = \int d\mathbf{r} v^{\text{KS}}(\mathbf{r})n(\mathbf{r}). \quad (2.9)$$

The form of the potential is tuned such that the non-interacting system has the same density as the interacting system. Because the Kohn-Sham system is non-interacting, one ends up with N separate equations for the N non-interacting electrons,

$$\left(-\frac{\hbar^2 \nabla^2}{2m_e} + v^{\text{KS}}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (2.10)$$

where $\psi_i(\mathbf{r})$ and ϵ_i are the i 'th the Kohn-Sham orbital and the Kohn-Sham energy respectively. The effective single-particle Kohn-Sham potential takes the following form,

$$v^{\text{KS}}(\mathbf{r}) = v^{\text{KS}}(\mathbf{r}) + v^{\text{ext}}(\mathbf{r}) + v^{\text{xc}}(\mathbf{r}), \quad (2.11)$$

$$v^{\text{xc}}(\mathbf{r}) = \frac{\delta E^{\text{xc}}[n]}{\delta n(\mathbf{r})}. \quad (2.12)$$

Finally, the density of the non-interacting can be calculated from the Kohn-Sham orbitals,

$$n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2, \quad (2.13)$$

where f_i is the occupation function of orbital i .

Since the effective potential is a functional of the electronic density, and the electronic density depends on the Kohn-Sham wave functions, the Kohn-Sham equations have to be solved self-consistently. Furthermore, since the general form of $E^{\text{xc}}[n]$ is still unknown, suitable approximations for this term are needed in practice. Some common approximations are discussed in the next section.

2.1.4 Exchange-correlation functionals for ground state DFT

There are many different types xc functionals for ground state DFT. In the following a few relevant to the work in this thesis are discussed.

The Local Density Approximation (LDA) The first and simplest example of a xc-functional is the local density approximation[61]. The LDA is constructed to be exact for the homogeneous electron gas and calculates the exchange-correlation energy according to,

$$E_{\text{xc}}^{\text{LDA}}[n] = \int d\mathbf{r} e_{\text{xc}}^{\text{hom}}(n(\mathbf{r})), \quad (2.14)$$

where $e_{xc}^{\text{hom}}(n(\mathbf{r}))$ is the exchange-correlation energy per volume in the homogeneous electron gas. $e_{xc}^{\text{hom}}(n(\mathbf{r}))$ can be split into an exchange contribution and a correlation contribution, the former of which has an analytic form[62] and the latter of which has to be approximated using high-level Monte-Carlo calculations[63–65]. The LDA is thus exact for the homogeneous electron gas, and as Eq. 2.14 would suggest, it works quite well for systems with slowly varying densities and metals. In contrast, it performs poorly on systems with localized states because it is unable to properly account for the effect of variations in the electron density. Additionally, the LDA suffers from the band gap problem inherent to local Kohn-Sham constructions and systematically underestimates electronic band gaps.

The Generalized Gradient Approximation (GGA) The first step of Jacob’s ladder beyond LDA is a functional that includes the gradient of the electron density[66],

$$E_{xc}^{\text{GGA}}[n] = \int d\mathbf{r} f(n(\mathbf{r}), \nabla n(\mathbf{r})). \quad (2.15)$$

Conceptually, the idea behind the GGA construction is that including the gradient in the integrand, $f(n(\mathbf{r}), \nabla n(\mathbf{r}))$, will allow the functional to better describe the effect of spatial variations in the density. While many different GGAs exist, distinguished by their choice of $f(n(\mathbf{r}), \nabla n(\mathbf{r}))$, by far the most commonly used is the Perdew, Burke and Ernzerhof (PBE) functional[67]. Compared to LDA, GGAs generally improve the description of bond lengths, binding energies and band gaps, although the GGAs still suffer from the band gap problem.

The bandgap problem in DFT and approaches to overcome it: The bandgap problem refers to the fact that the DFT band gap underestimates the true band gap because it fails to account for material dependent the derivative discontinuity (DD), Δ_{xc} . The DD is a system dependent quantity and it arises from the fact that in real systems the derivatives of certain quantities depends on the number of electrons in a discontinuous way and this is not captured by standard formulations of DFT[68]. This can be expressed as,

$$E_G^{\text{True}} = E_G^{\text{DFT}} + \Delta_{xc}, \quad \Delta_{xc} \geq 0. \quad (2.16)$$

Δ_{xc} is in practice always positive and the neglect of it in the standard (semi)local formulations of DFT is the reason for the systematic underestimation of the band gap of solids in DFT calculations. One approach to overcome this shortcoming of semi-local functionals is to attempt to evaluate the derivative discontinuity, as done by e.g. the GLLBSC exchange correlation functional[69, 70].

Another approach is to use so-called Hybrid functionals. In the hybrid approach, the exchange part of the functional is mixed or hybridized with some fraction of the exchange from Hartree Fock[71],

$$E_{xc}^{\text{Hybrid}}[n] = \alpha E_x^{\text{HF}}[n] + (1 - \alpha) E_x^{\text{DFT}}[n] + E_c^{\text{DFT}}[n]. \quad (2.17)$$

The motivation behind the hybrid functionals is that since DFT tends to underestimate the bandgap due to the bandgap problem and Hartree Fock tends to overestimate the band gap, a combination of the two should perform better than the two individually. The fraction of exchange taken from Hartree Fock, α , is a fitting parameter and it is determined by fitting towards a large dataset. Some famous examples of hybrid functionals are HSE06[72] and PBE0[73].

2.2 Time Dependent Density Functional Theory with photons

While DFT is a powerful tool to describe the ground state of an electronic system, it is not sufficient to describe the full range of properties of a material. Imagine for example that the system is perturbed by a laser pulse, and the quantity of interest is the optical response of the material. This response will be related to the system's time evolution under the influence of the laser pulse, and it will therefore be necessary to go beyond the ground state properties. The time dependent many-body problem is even more complicated than its ground state equivalent, and directly solving the time dependent Schrödinger equation is therefore infeasible for practically relevant systems.

Fortunately, DFT can be generalized to the time dependent case under quite general conditions, in which case it is called Time Dependent DFT (TDDFT). The foundation of TDDFT is the Runge-Gross (RG) theorem. It states that given the initial state of the electronic system, there is a one to one correspondence between the electronic density and the external potential up to a purely time dependent constant, if the external potential is Fourier expandable[74]. Any observable of the system can thus be expressed in terms of the time dependent density and the initial state of the system instead of the full time dependent wave function. Similar to the ground state case, it can be proven that the time dependent density can be reproduced by a non-interacting Kohn-Sham system, now formulated with a time-dependent Kohn-Sham potential[75].

Standard formulations of TDDFT do not consider the photon field, but only deals with the perturbation of electronic systems by external fields. For the purposes of this thesis, it makes sense to discuss a generalized version of TDDFT that can account for the quantized electromagnetic field, the so-called Quantum Electrodynamical Density Functional Theory (QEDFT). QEDFT is similar to TDDFT in that it is a reformulation of the time-dependent many-body problem, but instead of only considering the quantum nature of the electrons, QEDFT is a DFT type reformulation of the full interacting QED problem[38, 39]. QEDFT differs from TDDFT in terms of the number of internal variables needed to describe the observables of the system. In normal DFT, only the matter part is considered quantized and consequently the dynamics can be mapped onto one internal variable, either the electronic density or the electronic current depending on the form on the external potential[75]. In QEDFT, both the light and matter are quantized and as a result two internal variables are needed to reproduce the dynamics: the electronic density or current, and the expectation value

of the vector potential[38, 39]. By mapping the dynamics of the problem onto these internal variables, the problem is reformulated in terms of a nonlinear fluid equation for the density coupled non-linearly to the modes of the electromagnetic field and their corresponding Maxwell's equations[38, 39]. As is also the case with ground state DFT, in this new set of coupled equations there are unknown terms and they therefore cannot be solved directly. One therefore again constructs a non-interacting Kohn-Sham system to reproduce the dynamics and approximate the exchange-correlation terms that appear.

While the QEDFT reformulation of the full many-body QED problem can be proved under very general conditions[39], most practical formulations of QEDFT use the non-relativistic PZW gauge Hamiltonian of an electronic system coupled to a mode resolved electromagnetic field within the dipole approximation (See section 1.2),

$$\hat{H}_0 = \hat{H}_e + \frac{1}{2} \sum_{\alpha}^N \left(\hat{p}_{\alpha}^2 + \omega_{\alpha}^2 \left[\hat{q}_{\alpha} - \frac{\boldsymbol{\lambda}_{\alpha}}{\omega_{\alpha}} \cdot \hat{\mathbf{R}} \right]^2 \right). \quad (2.18)$$

This Hamiltonian is also the one employed by the Octopus code[76] which was used for the QEDFT calculations in Chapter 4. If the external perturbation can be represented by a scalar potential, the internal variables are the electronic density, and the expectation values of the photon mode coordinates, $(n(\mathbf{r}, t), \{q_{\alpha}(t)\})$ [38, 39]. The Kohn-Sham construction of QEDFT defines the following exchange-correlation potentials,

$$v_{\text{Mxc}}[n, q_{\alpha}](\mathbf{r}, t) = v_s[n](\mathbf{r}, t) - v[n, q_{\alpha}](\mathbf{r}, t), \quad (2.19)$$

$$j_{\alpha, \text{M}}[n, q_{\alpha}](\mathbf{r}, t) = j_{\alpha, s}[q_{\alpha}](\mathbf{r}, t) - j_{\alpha}[n, q_{\alpha}](\mathbf{r}, t), \quad (2.20)$$

where $v_s[n](\mathbf{r}, t)$ is the potential in the non interacting system and $v[n, \hat{q}_{\alpha}](\mathbf{r}, t)$ is the potential in the interacting system, and similarly for the currents. The Mxc subscript denotes mixed exchange-correlation because it contains both electronic and photonic contributions. Importantly, since the electrons and photons are decoupled in the non interaction system, the potential and currents in the non-interacting system have only functional dependence on the electronic density and displacement operators respectively. In the time dependent case, the photon xc -current only has the mean field contribution[77]. This means that one can write, $j_{\alpha, \text{M}}[n, \hat{q}_{\alpha}](\mathbf{r}, t) = -\omega_{\alpha}^2 \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{R}(t)$. With these definitions, the problem can be recast in terms of the following set of coupled equations for a set of non-interacting electrons, coupled to a set of photonic modes,

$$i\hbar\partial_t\phi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + v_{\text{KS}}[v, n, q_{\alpha}](\mathbf{r}, t) \right] \phi(\mathbf{r}, t), \quad (2.21)$$

$$(\partial_t^2 + \omega_{\alpha}^2) q_{\alpha}(t) = -\frac{j_{\alpha}(t)}{\omega_{\alpha}} + \omega_{\alpha} \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{R}(t). \quad (2.22)$$

Here the KS potential is defined as, $v_{\text{KS}}[v, n, q_{\alpha}](\mathbf{r}, t) = v_{\text{ext}}[v](\mathbf{r}, t) + v_{\text{Mxc}}[n, q_{\alpha}](\mathbf{r}, t)$. Instead of solving the coupled problem, one can thus solve the set of uncoupled equations, provided that appropriate approximations of the exchange correlation terms can be found. The exchange correlation naturally divides into an electronic part,

which is customarily denoted as the Hartree exchange-correlation (Hxc), and a photonic part, which is denoted as the photonic exchange correlation (pxc)[38],

$$v_{\text{Mxc}}[n, q_\alpha](\mathbf{r}, t) = v_{\text{Hxc}}[n](\mathbf{r}, t) + v_{\text{pxc}}[n, q_\alpha](\mathbf{r}, t). \quad (2.23)$$

This allows the well known xc-functionals from DFT and Time Dependent DFT (TDDFT) to be reused in the case of the electronic part of QEDFT. Because the xc-potentials depend on the density and photon coordinates, the equations have to be solved self-consistently. This thesis mainly employs TDDFT and QEDFT within the linear response regime, as discussed in the next chapter. Exchange and correlation in TDDFT and QEDFT will therefore be discussed in the next chapter only in the ways that it relates to the linear response formalism.

Besides quantum light, the QEDFT formulation can also be thought of as a way to perform TDDFT calculations for an open quantum system,[38] which in principle allows for the e.g. intrinsic linewidth of spectral peaks to be calculated directly instead of being included as a simulation parameter[77]. However, a major question for both applications related to quantum light, and for TDDFT for open quantum systems, is how the set of field parameters $\{\lambda_\alpha\}$ should be related to a real electromagnetic environment. It turns out, as shown partially in Publication [I] and elaborated significantly in 4, that the MQED expansion of the electromagnetic fields in terms of the DGF can be used to directly relate the field parameters to a general electromagnetic environment. The work presented in Chapter 4 in particular paves the way for both all-modes simulation of quantum optics with large electronic systems, and TDDFT for a general open quantum system in a general electromagnetic environment.

CHAPTER 3

Linear Response Theory

Having discussed methods to efficiently determine the electronic structure of matter, this chapter focuses on how this knowledge can be translated into spectroscopic information, focusing on how electronic systems respond to weak external perturbations. For weak external perturbations, the response of the material will be proportional to the applied perturbation, and it is governed by so-called linear response theory. A simple example of linear response is Hooke's law which states that the restoring force of a spring, F , is proportional to the displacement from equilibrium, Δx ,

$$F = -k\Delta x, \tag{3.1}$$

where k is the spring constant. The larger k is, the larger the restoring force will be and thus the harder the spring is to displace from equilibrium. Therefore, k is a function, albeit a trivial one, which encodes the first order response of the system. This makes it an example of a *linear response function*. Of course, if you're very strong, it is possible to pull the spring so far that it starts plastic deformation, and eventually it will break. When this point is reached, the linear approximation breaks and Eq. 3.1 no longer applies. At this point, the response of the system to the perturbation has gone beyond the realm of linear response and more complicated descriptions are needed. However, within the regime that a spring would normally be operated, Eq. 3.1 remains valid and even with its limitations it is therefore a very powerful formula.

This thesis is not about springs, but in some ways it is concerned with questions that are conceptually similar to the question of how strong a spring is. However, instead of asking what the response of the material will be to a mechanical perturbation, this thesis is focused on what the response of a material to an electromagnetic perturbation will be, and how knowing this can help us derive spectroscopic information about the materials from *ab-initio* calculations. As is also the case for springs, considerations of a material's response to electromagnetic fields within linear response are only valid if the perturbing field is weak. While linear response goes a long way towards describing the response of matter to electromagnetic fields, it also makes sense to mention that there are many properties that require a beyond linear response description, including nonlinear optics phenomena such as high harmonic generation[78] and supercontinuum generation in optical fibers[79].

The structure of this chapter is as follows: First the formal quantum theory behind the linear response is outlined and the Kubo formula is introduced. With this theory at hand, the linear response of a non-relativistic, interacting QED system is discussed and the relevant linear response functions are introduced. Following this, the practical

formulation of linear response in terms of QEDFT is introduced. Finally, some approximations of the exchange-correlation kernels are discussed and the Random Phase Approximation (RPA) is introduced.

3.1 Linear response theory

Consider a system described by the Hamiltonian \hat{H}_0 . Suppose that at some time t_0 , a weak perturbation, $\hat{H}_{\text{ext}}(t)$, is applied, resulting in the system being driven from equilibrium,

$$\hat{H}(t) = \hat{H}_0 + \theta(t - t_0)\hat{H}_{\text{ext}}(t). \quad (3.2)$$

To linear order, the response of the system is proportional to the applied perturbation and given by the famous Kubo formula[31]. For the purposes of this thesis, it is sufficient to consider perturbations whose time dependence can be written as,

$$\hat{H}_{\text{ext}}(t) = \int d\mathbf{r} \hat{A}(\mathbf{r}) f(\mathbf{r}, t), \quad (3.3)$$

where $\hat{A}(\mathbf{r})$ is some one-body electronic operator, and $f(\mathbf{r}, t)$ is a function that generally depends on both spatial position and time. According to the Kubo formula, the response of an observable described by $\hat{B}(\mathbf{r}, t)$ is given as,

$$\langle \delta \hat{B}(\mathbf{r}, t) \rangle = \int d\mathbf{r}' \int_{t_0}^{\infty} dt' \chi_A^B(\mathbf{r}, \mathbf{r}', t - t') f(\mathbf{r}', t'), \quad (3.4)$$

$$\chi_A^B(\mathbf{r}, \mathbf{r}', t - t') = -\frac{i}{\hbar} \theta(t - t') \langle [\hat{B}_I(\mathbf{r}, t), \hat{A}_I(\mathbf{r}', t')] \rangle_0. \quad (3.5)$$

Here $\chi_A^B(\mathbf{r}, \mathbf{r}', t - t')$ is the retarded response function, $\langle \cdot \rangle_0$ denotes the expectation value with respect to the equilibrium state of the system, and the subscript I denotes an interaction picture operator. Eq. 3.5 embodies the remarkable result that a non-equilibrium property, the linear response to an external perturbation, can be entirely described in terms of a retarded correlation function of the system in equilibrium. This result is a manifestation of the fluctuation-dissipation theorem, which states that the ability of a system to dissipate energy is directly related to the system's fluctuations in equilibrium[80].

3.2 Linear response of a non-relativistic QED system

Based on the theory outlined in the previous section, the linear response of a coupled non-relativistic QED system can be studied. Following Ref. [77], an electronic system coupled within the dipole approximation to a discrete set of photon modes, $\{\hat{q}_\alpha\}$, is considered. Adopting the PZW gauge and neglecting magnetic interactions, the Hamiltonian becomes,

$$\hat{H}_0 = \hat{H}_e + \frac{1}{2} \sum_{\alpha=1}^N \left(\hat{p}_\alpha^2 + \omega_\alpha^2 \left[\hat{q}_\alpha - \frac{\boldsymbol{\lambda}_\alpha}{\omega_\alpha} \cdot \hat{\mathbf{R}} \right]^2 \right). \quad (3.6)$$

There are two ways to probe the coupled system light-matter system: Either with an external field coupling to the charged particles, or with an external current coupling to the photon modes. A general external perturbation can thus be written as,

$$\hat{H}_{\text{ext}}(t) = \hat{V}_{\text{ext}}(t) + \hat{J}_{\text{ext}}(t), \quad (3.7)$$

$$\hat{V}_{\text{ext}}(t) = \int d\mathbf{r}' \hat{n}(\mathbf{r}') v_{\text{ext}}(\mathbf{r}', t), \quad (3.8)$$

$$\hat{J}_{\text{ext}}(t) = \sum_{\alpha} \frac{j_{\alpha}(t)}{\omega_{\alpha}} \hat{q}_{\alpha}. \quad (3.9)$$

Here $v_{\text{ext}}(\mathbf{r}, t)$ is the externally applied potential and $j_{\alpha}(t)$ is the external current pumping mode α . Importantly, the perturbation in Eq. 3.8 is only sufficient to describe perturbation of the electronic system by longitudinal fields. Consequently, Eq. 3.8 is generally not sufficient to study the interaction between matter and electromagnetic radiation, because electromagnetic radiation is described by a transverse field. However, in the long wavelength limit, the longitudinal and transverse response functions coincide.[81] Because photons in the optical and near UV regimes carry negligible momentum, the optical response can thus be studied from the longitudinal response function.

Because the light and matter are coupled via \hat{H}_0 , the density will respond to the external current and the photon modes will respond to the external potential respectively, even though the perturbations do not directly couple to these degrees of freedom. For this reason, there are fundamentally four response functions to consider,

$$\delta n(\mathbf{r}, t) = \int_{t_0}^{\infty} dt' \int d\mathbf{r}' \chi_n^n(\mathbf{r}t, \mathbf{r}'t') \delta v(\mathbf{r}', t') + \sum_{\alpha=1}^N \int_{t_0}^{\infty} dt' \chi_{q_{\alpha}}^n(\mathbf{r}t, t') \delta j_{\alpha}(t'), \quad (3.10)$$

$$\delta q_{\alpha}(t) = \int_{t_0}^{\infty} dt' \int d\mathbf{r}' \chi_n^{q_{\alpha}}(t, \mathbf{r}'t') \delta v(\mathbf{r}', t') + \sum_{\beta=1}^N \int_{t_0}^{\infty} dt' \chi_{q_{\beta}}^{q_{\alpha}}(t, t') \delta j_{\beta}(t'), \quad (3.11)$$

where the response functions can be derived from the Kubo formula in Eq. 3.5,

$$\chi_n^n(\mathbf{r}t, \mathbf{r}'t') = -i\hbar^{-1} \theta(t-t') \langle [\hat{n}_I(\mathbf{r}, t), \hat{n}_I(\mathbf{r}', t')] \rangle_0, \quad (3.12)$$

$$\chi_{q_{\alpha}}^n(\mathbf{r}t, t') = -i\hbar^{-1} \theta(t-t') \omega_{\alpha}^{-1} \langle [\hat{n}_I(\mathbf{r}, t), \hat{q}_{I,\alpha}(t')] \rangle_0, \quad (3.13)$$

$$\chi_n^{q_{\alpha}}(t, \mathbf{r}'t') = -i\hbar^{-1} \theta(t-t') \langle [\hat{q}_{I,\alpha}(t), \hat{n}_I(\mathbf{r}', t')] \rangle_0, \quad (3.14)$$

$$\chi_{q_{\beta}}^{q_{\alpha}}(t, t') = -i\hbar^{-1} \theta(t-t') \omega_{\beta}^{-1} \langle [\hat{q}_{I,\alpha}(t), \hat{q}_{I,\beta}(t')] \rangle_0. \quad (3.15)$$

These response functions describe the density-density, the density-photon, the photon-density, and the photon-photon responses respectively. Of particular importance in condensed matter physics is the density-density response function in Eq. 3.12. In the absence of quantized light modes, this is the only remaining response function, and it is the response function of interest when considering most types of spectroscopy and the dielectric screening in materials. For this reason, Chapter 5 will be dedicated exclusively to it.

Because of the form of perturbations considered, the linear response functions can only depend on the time difference, $t - t'$, making it favourable to adopt a frequency space representation. In the Lehmann representation, the frequency space response functions are represented in terms of the eigenstates and energies of the coupled electron-photon system,

$$\chi_n^n(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{\hbar} \lim_{\eta \rightarrow 0^+} \sum_n \left[\frac{f_n(\mathbf{r})f_n^*(\mathbf{r}')}{\omega - (E_n - E_0)/\hbar + i\eta} - \frac{f_n(\mathbf{r}')f_n^*(\mathbf{r})}{\omega + (E_n - E_0)/\hbar + i\eta} \right], \quad (3.16)$$

$$\chi_{q\alpha}^n(\mathbf{r}, \omega) = \frac{1}{\hbar} \lim_{\eta \rightarrow 0^+} \sum_n \frac{1}{\omega_\alpha} \left[\frac{f_n(\mathbf{r})g_{n,\alpha}^*}{\omega - (E_n - E_0)/\hbar + i\eta} - \frac{g_{n,\alpha}f_n^*(\mathbf{r})}{\omega + (E_n - E_0)/\hbar + i\eta} \right], \quad (3.17)$$

$$\chi_n^{q\alpha}(\mathbf{r}, \omega) = \frac{1}{\hbar} \lim_{\eta \rightarrow 0^+} \sum_n \left[\frac{f_n^*(\mathbf{r})g_{n,\alpha}}{\omega - (E_n - E_0)/\hbar + i\eta} - \frac{g_{n,\alpha}^*f_n(\mathbf{r})}{\omega + (E_n - E_0)/\hbar + i\eta} \right], \quad (3.18)$$

$$\chi_{q\beta}^{q\alpha}(\omega) = \frac{1}{\hbar} \lim_{\eta \rightarrow 0^+} \sum_n \frac{1}{\omega_\beta} \left[\frac{g_{n,\alpha}g_{n,\beta}^*}{\omega - (E_n - E_0)/\hbar + i\eta} - \frac{g_{n,\alpha}^*g_{n,\beta}}{\omega + (E_n - E_0)/\hbar + i\eta} \right], \quad (3.19)$$

where $f_n(\mathbf{r}) = \langle \Psi_0 | \hat{h}(\mathbf{r}) | \Psi_n \rangle$ and $g_{n,\alpha} = \langle \Psi_0 | \hat{q}_\alpha | \Psi_n \rangle$ are the transition matrix elements, $|\Psi_0\rangle$ is the correlated electron-photon ground state, and $|\Psi_n\rangle$ is the n 'th excited state of the correlated electron-photon system. The infinitesimal η ensures causality.

The response functions introduced in the previous section are often referred to as polarizabilities because they encode the systems' ability of polarize in response to an external perturbation, and they provide access to a range of spectroscopic information. χ_n^n and χ_q^q encode the direct response of the matter or photonic modes to an external field or current respectively, and thus describe the absorption coefficients of the electronic system and cavity modes respectively. Another way to phrase the same is that they represent the normal ways of doing spectroscopy of the systems individually, although they of course include the coupling in the general case. The cross polarizations, χ_q^n and χ_n^q , allows for the probing of one of the subsystems indirectly and these functions encode the cross talk between the systems[77].

The problem in practice is that the wave function of the correlated electron-photon system is unknown and as discussed in section 2.2 very hard to solve for directly. It therefore makes sense to formulate the problem in terms of QEDFT, as discussed below in section 3.3.

3.3 Calculating linear response functions within the QEDFT framework

In contrast to the interacting case, for the non-interacting Kohn-Sham system of QEDFT it is comparatively easy to obtain the eigenstates and energies of the system. From the set of energies and wave functions, it is possible to construct the response functions of the non-interacting system via the Lehmann representation. The response functions of the non-interacting system is denoted with a "0"-subscript. Because the photons and electrons are decoupled in the non-interacting system it only has two

response functions, $\chi_{n,0}^n$ and $\chi_{q\beta,0}^{q\alpha}$. The response functions of the interacting system can then be calculated in terms of the non-interacting response function by solving the following coupled set of Dyson equations[75, 77],

$$\begin{aligned} \chi_n^n(\mathbf{r}t, \mathbf{r}'t') &= \chi_{n,0}^n(\mathbf{r}t, \mathbf{r}'t') + \\ &\int \int \int \int d\tau d\tau' d\mathbf{r}_1 d\mathbf{r}_2 \chi_{n,0}^n(\mathbf{r}t, \mathbf{r}_1\tau) f_{\text{Mxc}}^n(\mathbf{r}_1\tau, \mathbf{r}_2\tau') \chi_n^n(\mathbf{r}_2\tau', \mathbf{r}'t') + \\ &\sum_{\alpha} \int \int \int \int d\tau d\tau' d\mathbf{r}_1 \chi_{n,0}^n(\mathbf{r}t, \mathbf{r}_1\tau) f_{\text{Mxc}}^{q\alpha}(\mathbf{r}_1\tau, \tau') \chi_n^{q\alpha}(\tau', \mathbf{r}'t') \end{aligned} \quad (3.20)$$

$$\begin{aligned} \chi_{q\alpha}^n(\mathbf{r}t, t') &= \int \int \int \int d\tau d\tau' d\mathbf{r}_1 d\mathbf{r}_2 \chi_{n,0}^n(\mathbf{r}t, \mathbf{r}_1\tau) f_{\text{Mxc}}^n(\mathbf{r}_1\tau, \mathbf{r}_2\tau') \chi_{q\alpha}^n(\mathbf{r}_2\tau', t') + \\ &\sum_{\beta} \int \int \int \int d\tau d\tau' d\mathbf{r}_1 \chi_{n,0}^n(\mathbf{r}t, \mathbf{r}_1\tau) f_{\text{Mxc}}^{q\alpha}(\mathbf{r}_1\tau, \tau') \chi_{q\alpha}^{q\beta}(\tau', t') \end{aligned} \quad (3.21)$$

$$\chi_n^{q\alpha}(t, \mathbf{r}'t') = \sum_{\beta} \int \int \int \int d\tau d\tau' d\mathbf{r}_1 \chi_{q\beta,0}^{q\alpha}(t, \tau) g_{\text{Mxc}}^{n\beta}(\tau, \mathbf{r}_1\tau') \chi_n^n(\mathbf{r}_1\tau', \mathbf{r}'t') \quad (3.22)$$

$$\begin{aligned} \chi_{q\beta}^{q\alpha}(t, t') &= \chi_{q\beta,0}^{q\alpha}(t, t') + \\ &\sum_{\beta} \int \int \int \int d\tau d\tau' d\mathbf{r}_1 \chi_{q\beta,0}^{q\alpha}(t, \tau) g_{\text{Mxc}}^{n\beta}(\tau, \mathbf{r}_1\tau') \chi_{q\beta}^n(\mathbf{r}_1\tau', t') \end{aligned} \quad (3.23)$$

where the exchange-correlation kernels are defined as,

$$f_{\text{Mxc}}^n(\mathbf{r}t, \mathbf{r}'t') = \frac{\delta v_{\text{Hxc}}[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t')} + \frac{\delta v_{\text{pxc}}[n, q_{\alpha}](\mathbf{r}, t)}{\delta n(\mathbf{r}', t')} \equiv f_{\text{Hxc}}^n(\mathbf{r}t, \mathbf{r}'t') + f_{\text{pxc}}^n(\mathbf{r}t, \mathbf{r}'t') \quad (3.24)$$

$$f_{\text{Mxc}}^{q\alpha}(\mathbf{r}t, t') = \frac{\delta v_{\text{pxc}}[n, q_{\alpha}](\mathbf{r}, t)}{\delta q_{\alpha}(t')} \equiv f_{\text{pxc}}^{q\alpha}(\mathbf{r}t, t') \quad (3.25)$$

$$g_M^{n\alpha}(t, \mathbf{r}'t') = \frac{\delta j_{\alpha, M}(t)}{\delta n(\mathbf{r}'t')} = -\delta(t-t') \omega_{\alpha}^2 \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{e}\mathbf{r}' \quad (3.26)$$

Note that the last constant, $g_M^{n\alpha}(t, \mathbf{r}'t')$, can be derived exactly since $j_{\alpha, M}(t)$ is known exactly, see section 2.2. The observant reader will notice that there is seemingly a kernel missing, namely $g_M^{q\alpha}(t, t) = \frac{\delta j_{\alpha, M}(t)}{\delta q_{\alpha}(t')}$. However, since the xc-current has no functional dependence on the displacement field it vanishes.

The approach of reformulating the linear response problem in terms of QEDFT provides a practical way to calculate the linear response of even large systems with many electrons, provided that suitable approximations for the exchange-correlation kernels can be found.

3.4 The Random Phase Approximation

In practice, one needs to approximate the exchange-correlation kernels to perform actual calculations. The most widely applied approximation is the so-called Random Phase Approximation(RPA). The RPA was initially developed for the case without

photons, in which case the only exchange-correlation kernel is $f_{\text{Hxc}}^n(\mathbf{r}_1\tau, \mathbf{r}_2\tau')$. For the electronic part, the RPA is equivalent to only keeping the Hartree (or mean field) contribution to the exchange-correlation kernel. In the electronic case, the RPA is therefore equivalent to time-dependent Hartree theory,

$$f_{\text{Hxc}}^n(\mathbf{r}_1t, \mathbf{r}_2t') = \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \delta(t - t'). \quad (3.27)$$

In practice, RPA is often sufficient to accurately describe the plasmonic excitations of a system, but completely fails to capture excitonic effects because the direct interactions are neglected. In the context of materials, the RPA is therefore most appropriate for the description of metallic systems, but often falls short in the description of semiconductors. The shortcomings of RPA for the Hartree-exchange are discussed further in section 5.2.3, and the consequences of going beyond it are discussed in 5.2.5 on the Bethe-Salpeter Equation.

Inspired by the standard RPA, Ref. [77] defined the photonic RPA (pRPA) as taking the mean-field contributions to the photonic kernels,

$$f_{\text{pxc}}^n(\mathbf{r}_1t, \mathbf{r}_2t') = \delta(t - t') \sum_{\alpha} (\boldsymbol{\lambda}_{\alpha} \cdot \mathbf{e}\mathbf{r}_1) \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{e}\mathbf{r}_2, \quad (3.28)$$

$$f_{\text{pxc}}^{g\alpha}(\mathbf{r}_1t, t') = -\delta(t - t') \omega_{\alpha} \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{e}\mathbf{r}_1, \quad (3.29)$$

$$g_M^{n\alpha}(t, \mathbf{r}_1t') = -\delta(t - t') \omega_{\alpha}^2 \boldsymbol{\lambda}_{\alpha} \cdot \mathbf{e}\mathbf{r}_1. \quad (3.30)$$

Note that at this level of approximation $g_M^{n\alpha}(t, \mathbf{r}'t')$ is exact. As discussed in Ref. [77], the main short-coming of the pRPA is that it neglects multi-photon processes, and the pRPA is therefore appropriate in the weak and strong coupling regimes where multi-photon processes are limited, but fails in the ultra strong coupling regime where multi-photon processes begin to significantly affect the dynamics of the system[46].

3.5 The generalized Casida formulation of LR-QEDFT for finite systems

It is possible to extract spectroscopic information from the polarizabilities of the system. Assume for example that a molecular system is probed by an external field. This will induce a change in the dipole moment of the molecule,

$$\delta\mathbf{D}(t) = - \int d\mathbf{r} \mathbf{e}\mathbf{r} \delta n(\mathbf{r}, t). \quad (3.31)$$

To characterize this coupling, the dipole polarizability tensor can be defined as,

$$\mathbf{D}(\omega) = \boldsymbol{\alpha}(\omega) \mathbf{E}(\omega), \quad (3.32)$$

which leads to the expression,

$$\alpha_{\mu\nu}(\omega) = - \int d\mathbf{r} \frac{\delta n(\mathbf{r}, \omega)}{\delta E^{\nu}(\omega)} \mathbf{e}r_{\mu} = e^2 \int \int d\mathbf{r} d\mathbf{r}' r_{\mu} \chi_n^n(\mathbf{r}, \mathbf{r}', \omega) r'_{\nu}. \quad (3.33)$$

In practice, one often defines the dipole polarizability as the trace of the tensor in 3.33[82],

$$\alpha(\omega) = \text{Tr} [\alpha^{\mu\nu}(\omega)] / 3, \quad (3.34)$$

which is directly related to the photoabsorption cross section of the electronic system,

$$\sigma(\omega) = \frac{4\pi\omega}{3c} \text{Im}\alpha(\omega). \quad (3.35)$$

For finite systems, the dipole polarizability can be written in terms of so-called oscillator strengths of the transitions between the many-body states of the system, f_I [82],

$$\alpha(\omega) = \sum_I \frac{f_I}{\omega^2 - \Omega_I^2}, \quad (3.36)$$

$$f_I = \frac{2}{3} \Omega_I \sum_{\mu=1}^3 |\langle \Psi_0 | e r_\mu | \Psi_I \rangle|^2. \quad (3.37)$$

where $|\Psi_0\rangle$ is the many-body ground state of the system, $|\Psi_I\rangle$ is the I 'th many-body excited state, and Ω_I is the energy of the transition between the two states. Note that these states are the many-body states of the coupled electron-photon system and they will generally have both electronic and photonic character. Importantly the oscillator strengths obey the so-called f-sum rule,

$$\sum_I f_I = N_{\text{el}} \quad (3.38)$$

where N_{el} is the number of electrons in the system.

In the absence of quantized light, the oscillator strengths can be determined by solving the Casida equation[75, 83]. As proven by Flick *et. al.* in 2019, the Casida equation can be generalized to situations with quantized light. This generalized Casida equation reads,

$$\begin{bmatrix} U & V \\ V^T & \omega_\alpha \end{bmatrix} \begin{bmatrix} \mathbf{F} \\ \mathbf{P} \end{bmatrix} = \Omega_S^2 \begin{bmatrix} \mathbf{F} \\ \mathbf{P} \end{bmatrix}. \quad (3.39)$$

Here ω_α is a diagonal matrix with the frequency of photon modes in the diagonal. U accounts for the coupling between the electrons. Introducing the pair orbital index $S = (ia)$, corresponding to the pair orbital $\Phi_S(\mathbf{r}) = \phi_i(\mathbf{r})\phi_a^*(\mathbf{r})$ with energy $\epsilon_S = \epsilon_a - \epsilon_i$, U can be expanded as,

$$U_{SS'} = \epsilon_S^2 \delta_{SS'} + 2\epsilon_S^{1/2} K_{SS'}(\Omega_S) \epsilon_{S'}^{1/2}, \quad (3.40)$$

$$K_{SS'}(\Omega_S) = \int \int d\mathbf{r} d\mathbf{r}' \Phi_S(\mathbf{r}) f_{\text{Mxc}}^n(\mathbf{r}, \mathbf{r}', \Omega_S) \Phi_{S'}(\mathbf{r}'). \quad (3.41)$$

If N_{pair} pair-orbitals are included in the calculation, U is therefore an $N_{\text{pair}} \times N_{\text{pair}}$ matrix. $V_{\alpha S}$ is the matrix accounting for the coupling between the electrons and the photon modes,

$$V_{\alpha S} = 2\sqrt{\epsilon_S M_{\alpha S}(\Omega_S) N_{\alpha S} \omega_\alpha}, \quad (3.42)$$

$$M_{\alpha, S}(\Omega_S) = \int d\mathbf{r} \Phi_S(\mathbf{r}) f_{\text{Mxc}}^{q_\alpha}(\mathbf{r}, \Omega_S), \quad (3.43)$$

$$N_{\alpha, S} = \frac{1}{2\omega_\alpha^2} \int d\mathbf{r} \Phi_S(\mathbf{r}) g_{\text{M}}^{q_\alpha}(\mathbf{r}). \quad (3.44)$$

The size of these matrices will be $N_{\text{photon}} \times N_{\text{pair}}$. The \mathbf{F} part of the eigenvector in Eq. 3.39 can be interpreted as the matter part and the \mathbf{P} can be interpreted as the photon part. The square norm of the two gives the electronic- and photonic fraction of the excitation respectively[77]. In terms of the eigenvectors it is possible to calculate the transition matrix elements that enter Eq. 3.37[77]. Finally, note that when neglecting the quantized light, only the top left block in Eq. 3.39 survives and one recovers the standard Casida formulation of linear response TDDFT for finite systems[83].

CHAPTER 4

Towards a first-principles description of real molecules in real cavities with MQED and QEDFT

The work presented in this section is still ongoing and therefore not included as a publication. This work is a result of my external research stay at the Flatiron Institute and it was supervised by Professor Angel Rubio and Assistant Professor Johannes Flick. It builds on top of the ideas presented in Publication [I] and seeks to combine MQED with QEDFT, providing a truly first principles description of many electron systems coupled to complex electromagnetic environments.

As discussed in section 2.2, QEDFT is a generalization of TDDFT to situations with quantized light modes. While QEDFT is a very powerful tool allowing for an *ab-initio* treatment of systems with many electrons coupled to many photon modes, a major limitation is that most practical formulations are based on a discrete mode expansion of the electromagnetic field. As discussed in Chapter 1, this implies a lossless medium. It has recently been suggested that the losses can be modelled by densely sampling the photonic environment within the discrete approximation[77, 84]. However, so far the formal connection between the standard Hamiltonian used by practical formulations of QEDFT, Eq. 3.6, and the PZW gauge Hamiltonian in the presence of losses has not been established. Furthermore, the connection between the electromagnetic environment and the cavity field parameters, $\{\lambda_\alpha\}$ has not been investigated in the QEDFT literature. This means that current studies limit them limit themselves to very simple cavity loss profiles, and even in these cases the cavity parameters are not linked to a real cavity setup.

In this project we want to study how the QEDFT framework can be used for real molecules in real cavity setups. Specifically, we will investigate how QEDFT can be formulated in terms of the field expansion from MQED within the dipole approximation. This will allow us to express the coupling constants in terms of the DGF, which can in turn be related to a real cavity setup.

4.1 Formalism

The starting point for this project is the PZW gauge Hamiltonian describing an electronic system coupled to the MQED field within the dipole approximation. Neglecting magnetic interactions, the Hamiltonian reads,

$$\mathcal{H} = \sum_{\alpha} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \frac{1}{2\epsilon_0} \int d\mathbf{r} \hat{\mathbf{P}}(\mathbf{r})^2 + \sum_{\lambda} \int_0^{\infty} d\omega \hbar\omega \int d\mathbf{r} \hat{\mathbf{f}}_{\lambda}^{\dagger}(\mathbf{r}, \omega) \cdot \hat{\mathbf{f}}_{\lambda}(\mathbf{r}, \omega) - \hat{\mathbf{d}} \cdot \hat{\mathbf{E}}(\mathbf{r}_0). \quad (4.1)$$

As discussed in section 1.2.1, this Hamiltonian can be reformulated in terms of a set of emitter centered bright modes. Within this formulation, the MQED expansion of the electric field can be written as[85],

$$\hat{\mathbf{E}}(\mathbf{r}) = \sum_i \int_0^{\infty} d\omega \mathbf{E}_i(\mathbf{r}, \omega) \hat{C}_i(\omega) + \text{h.c.}, \quad (4.2)$$

$$\mathbf{E}_i(\mathbf{r}, \omega) = \frac{\hbar\omega^2}{\pi\epsilon_0 c^2} \sum_j V_{ij}^*(\omega) \frac{\text{Im}\mathbf{G}(\mathbf{r}, \mathbf{r}_0, \omega) \cdot \hat{n}_j}{G_j(\omega)}. \quad (4.3)$$

To connect this to the Hamiltonian used in the Octopus code LR-QEDFT implementation, we follow Ref. [84] in defining the auxiliary variables, $\boldsymbol{\lambda}_i(\omega) = -\left(\frac{2}{\hbar\omega}\right)^{1/2} e\mathbf{E}_i(\mathbf{r}_0, \omega)$, and rewrite the field expansion at the emitter position as,

$$\hat{\mathbf{E}}(\mathbf{r}_0) = -\sum_i \int_0^{\infty} d\omega \left(\frac{\hbar\omega}{2e^2}\right)^{1/2} \boldsymbol{\lambda}_i(\omega) \hat{C}_i(\omega) + \text{h.c.}. \quad (4.4)$$

We want to change from the ladder operators to the canonical operators, $\hat{q}_i(\omega) = \left(\frac{\hbar}{2\omega}\right)^{1/2} (\hat{C}_i(\omega) + \hat{C}_i^{\dagger}(\omega))$ and $\hat{p}_i(\omega) = \left(\frac{\hbar\omega}{2}\right)^{1/2} (\hat{C}_i(\omega) - \hat{C}_i^{\dagger}(\omega))$, again to be consistent with the Octopus implementation of QEDFT. This is possible as long as the orthogonalization scheme used in the construction of $\hat{C}_i(\omega)$ results in real valued \mathbf{V} matrices. With this change, the electric field operator becomes,

$$\hat{\mathbf{E}}(\mathbf{r}_0) = -\sum_i \int_0^{\infty} d\omega \frac{\omega}{e} \boldsymbol{\lambda}_i(\omega) \hat{q}_i(\omega). \quad (4.5)$$

The idea is then to use Eq. 4.5 to put Eq. 4.1 in a form similar to the Hamiltonian in Eq. 3.6. The intuitive generalization of the Hamiltonian to the continuous case suggests that,

$$\mathcal{H} = \mathcal{H}_{\text{Mat}}^{\text{C}} + \frac{1}{2} \sum_i \int_0^{\infty} d\omega \left\{ \hat{p}_i(\omega)^2 + \omega^2 \left[\hat{q}_i(\omega) - \frac{\boldsymbol{\lambda}_i(\omega)}{\omega} \cdot \hat{\mathbf{R}} \right]^2 \right\}. \quad (4.6)$$

Here $\hat{\mathbf{R}} = \sum_i \hat{\mathbf{r}}_i$ is the center of mass position of the charges making up the emitter which is related to the dipole moment as $\hat{\mathbf{d}} = -e\hat{\mathbf{R}}$. The question is then under what conditions this holds. Inserting the field expansion into the emitter-field coupling term in the Hamiltonian results in,

$$\hat{H}_{\text{int}} = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}}(\mathbf{r}_0) = -\sum_i \int_0^\infty d\omega \omega \left[\boldsymbol{\lambda}_i(\omega) \cdot \hat{\mathbf{R}} \right] \hat{q}_i(\omega). \quad (4.7)$$

which is the straightforward generalization of the lossless case. The other term in which the field parameters appear is the dipole self energy term. As discussed in Chapter 1, the origin of the dipole self-energy terms is the matter Hamiltonian, and generally it should be derived in terms of the transverse projection of the full polarization term[36],

$$\hat{\mathbf{P}}_\perp(\mathbf{r}) = \frac{2}{3}\hat{\mathbf{P}}(\mathbf{r}) + \int d\mathbf{r}' \hat{\mathbf{T}}_\perp(\mathbf{r}, \mathbf{r}') \cdot \hat{\mathbf{P}}(\mathbf{r}'). \quad (4.8)$$

However, the direct application of this projection is computationally cumbersome. In the lossless case, the Helmholtz equation is Hermitian and its transverse solutions form a orthogonal and complete basis for the transverse space. Consequently, the transverse polarization can be expanded in terms of these modes and the projection can be simplified significantly. This procedure leads to the standard form of the dipole self-energy term,

$$\frac{1}{2\epsilon_0} \int d\mathbf{r} \left[\hat{\mathbf{P}}_\perp(\mathbf{r}) \right]^2 = \frac{1}{2} \sum_\alpha \left[\boldsymbol{\lambda}_\alpha \cdot \hat{\mathbf{R}} \right]^2. \quad (4.9)$$

The fact that the transverse modes can be used to span the transverse projector indicates that any truncation of the photon Hilbert space needs to be performed carefully as it indirectly results in a truncation of the transverse basis and therefore also the transverse polarization. The two should therefore generally always be truncated consistently[5]. In the limit of a complete basis for the photonic Hilbert space, there are some arguments that the term can be neglected[85]. This argument however seem to suppose that the coupling to the field at arbitrarily large frequencies can be described within the dipole approximation which would not be the case unless the system is a true point emitter. Furthermore, in any numerical implementation, a finite basis for the photonic Hilbert space would be employed, and the term is therefore generally important.

The expansion of the transverse polarization field in the presence of losses is more complicated. If the cavity losses are not true absorptive losses, but simply a result of e.g. finite reflectivity of the cavity mirrors, the modes will globally obey,

$$\int d\mathbf{r} \mathbf{E}_i(\mathbf{r}, \omega_1) \cdot \mathbf{E}_j(\mathbf{r}, \omega_2) = \delta_{ij} \delta(\omega_1 - \omega_2), \quad (4.10)$$

and thus remain orthogonal and complete. In this case the transverse polarization can be straightforwardly generalized,

$$\frac{1}{2\epsilon_0} \int d\mathbf{r} \left[\hat{\mathbf{P}}_\perp(\mathbf{r}) \right]^2 = \frac{1}{2} \sum_i \int_0^\infty d\omega \left[\boldsymbol{\lambda}_i(\omega) \cdot \hat{\mathbf{R}} \right]^2. \quad (4.11)$$

In the presence of true absorption losses it is not possible to define the transverse projector in terms of the field modes because these no longer form a complete, orthonormal set. We speculate that the transverse projection should instead be defined in terms of a biorthonormal construction as discussed in Ref. [56], but this is still a matter of ongoing research. From the point of the emitter it should not be possible to distinguish these two cases because it only samples the field at its position. We thus speculate that, at least for a single emitter position, the true form of the dipole self energy term in the presence of losses will functionally be no different than Eq. 4.11. As a first approximation it might also be generally prudent to approximate the dipole self energy via Eq. 4.11. We note that in the regimes of coupling strength considered in the following, the dipole self-energy terms is expected to have little effect[84] and this is therefore not important for the conclusions that follow.

4.2 Spherical Micro-Cavities

Next we turn to a discussion of the mode structure of a real cavity. As an example, we consider the spherically layered micro-cavity from Ref. [86]. As shown in Fig. 4.2a, the cavity consists of three concentric spherical layers, each characterized by a frequency dependent dielectric function, $\varepsilon_n(\omega)$.

To use the emitter centered framework, the DGF must be derived. In general, the source and field points (\mathbf{r} and \mathbf{r}' respectively) can be in either the same layer or different layers. Consequently, the DGF for the reflected field is a 9 component object, where each component is a 3D dyad. Labelling the possible combinations of source and field points by two extra indices $m, n = 1, 2, 3$, the full DGF of the system can be written as,

$$\mathbf{G}_{mn}(\mathbf{r}, \mathbf{r}', \omega) = \mathbf{G}^{\text{vac}}(\mathbf{r}, \mathbf{r}', \omega)\delta_{nm} + \mathbf{G}_{mn}^{\text{ref}}(\mathbf{r}, \mathbf{r}', \omega). \quad (4.12)$$

Due to the spherical symmetry of the problem, the DGF of the system is most efficiently expanded onto vector spherical harmonics[54]. The 9 different components of the reflection contribution can be worked out by invoking the electromagnetic boundary conditions at the material interfaces. In this work we consider the situation where the emitter is placed in the inner region of the cavity, and it is therefore only necessary to consider the $n = m = 3$ component of the DGF.

For a general emitter position inside the cavity, it is necessary to carefully converge the number of vector spherical harmonics used in the calculation of the DGF. However, if the emitter is placed exactly in the center of the cavity things simplify significantly. In this case, only the lowest order transverse magnetic mode contributes and the reflection DGF can be written as[86],

$$\mathbf{G}_{33}^{\text{ref}}(\mathbf{r}, \mathbf{r}', \omega)|_{r, r' \rightarrow 0} = \frac{i\omega}{6\pi c} r_{n=1}^N(\omega) \mathbb{1}. \quad (4.13)$$

where $r_{n=1}^N(\omega) \mathbb{1}$ is the reflection coefficient for the lowest order transverse magnetic

mode. If we assume that region 2 is very thick we can write[86],

$$r_{n=1}^N(\omega) = \frac{\left[i + \rho(n(\omega) + 1) - i\rho^2 n(\omega) - \frac{\rho^3 n(\omega)^2}{n+1} \right] e^{i\rho}}{\sin\rho - \rho(\cos\rho + in(\omega)\sin\rho) + i\rho^2 n(\omega)\cos\rho - \rho^3(\cos\rho - in(\omega)\sin\rho) \frac{n(\omega)^2}{n(\omega)^2 - 1}}, \quad (4.14)$$

where $\rho = R\omega/c$ is the standard size parameter from Mie theory[43].

The contribution from the vacuum DGF is $\text{Im}\mathbf{G}^{\text{vac}}(\mathbf{r}, \mathbf{r}, \omega) = \frac{\omega}{6\pi c} \mathbb{1}$. The imaginary part of the full DGF, evaluated in the center of the cavity, thus becomes,

$$\text{Im}\mathbf{G}(\mathbf{0}, \mathbf{0}, \omega) = \frac{\omega}{6\pi c} \left[1 + \text{Re}r_{n=1}^N(\omega) \right] \mathbb{1}. \quad (4.15)$$

Denoting the unit vector in the i 'th direction \mathbf{n}_i it can be noticed that, $\mathbf{n}_i \cdot \text{Im}\mathbf{G}(\mathbf{0}, \mathbf{0}, \omega) \cdot \mathbf{n}_j \propto \delta_{ij}$ which means that the mode orthogonalization needed in the emitter centered formulation is trivial in this case, (see section 1.2.1). The cavity field parameters can therefore be derived directly,

$$\boldsymbol{\lambda}_i(\omega) = -e \left(\frac{\omega^2}{6\pi^2 \epsilon_0 c^3} \left[1 + \text{Re}r_{n=1}^N(\omega) \right] \right)^{1/2} \mathbf{n}_i. \quad (4.16)$$

Notice that the spherical symmetry of the cavity means that at each frequency there are three orthogonal modes with identical field strengths, differing only by their spatial orientation i . Any dipole orientation therefore couples identically to the cavity modes if the emitter is placed in the center of the cavity.

4.2.1 Discretization of the cavity field parameters

In order to solve the problem using the linear response QEDFT formalism outlined in Chapter 3, the approach will be to densely sample the field parameters in Eq. 4.16. Assuming that we have uniformly spaced modes we can perform the simple discretization,

$$\int_0^\infty d\omega \rightarrow \sum_k \Delta\omega. \quad (4.17)$$

In the continuous formulation, the operators $\hat{C}_i(\omega)$ have units of $\text{s}^{1/2}$ and the field, $\mathbf{E}(\mathbf{r}, \omega)$ has units of $\text{Vs}^{1/2}/\text{m}$. However, in a discrete expansion the creation and annihilation operators should generally be unitless and the electric field should be in units of V/m . It therefore makes sense to define the a new discretized field and creating/annihilation operators as,

$$\mathbf{E}_k(\omega_k) \equiv \sqrt{\Delta\omega} \mathbf{E}_i(\mathbf{0}, \omega_k), \quad (4.18)$$

$$\hat{a}_k^{(\dagger)} \equiv \sqrt{\Delta\omega} \hat{C}_i^{(\dagger)}(\omega). \quad (4.19)$$

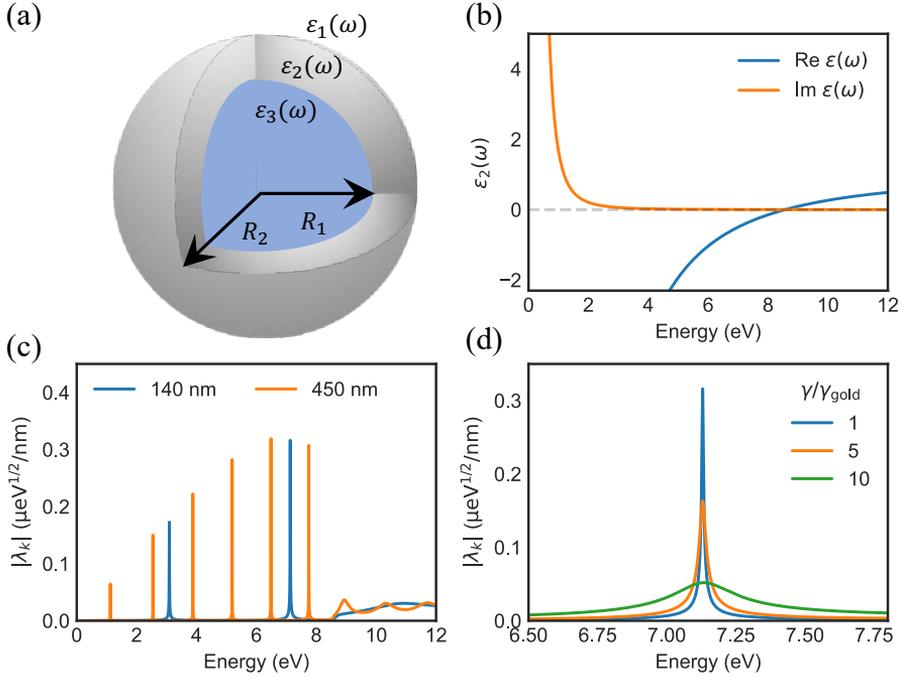


Figure 4.1: Spherical micro-cavity setup: (a) Illustration of the three layer spherical micro-cavity setup. (b) The Drude model dielectric function of gold. (c) The modes of the spherical micro-cavity plotted for two different radii, $R = 140\text{nm}$ in blue and $R = 450\text{nm}$ in orange. (d) The impact from the Drude damping parameter on the cavity resonances.

Transforming to canonical operators, this leads to the following discretized Hamiltonian,

$$\hat{H}_0 = \hat{H}_e + \frac{1}{2} \sum_k^N \left(\hat{p}_k^2 + \left[\hat{q}_k - \frac{\lambda_k}{\omega_k} \cdot \hat{\mathbf{R}} \right]^2 \right). \quad (4.20)$$

where the cavity field parameters are,

$$\lambda_k = -e \left(\frac{\Delta\omega\omega_k^2}{6\pi^2\epsilon_0c^3} [1 + \text{Re}r_{n=1}^N(\omega_k)] \right)^{1/2} \mathbf{n}_k. \quad (4.21)$$

Notice again that the spherical symmetry of this specific cavity considered here means that any dipole orientation couples identically to the cavity.

4.2.2 The modes of a gold shell cavity

A simple way to make a mirror is with a metal surface and the dielectric function of a metal can be approximated by the Drude model[81],

$$\epsilon_2(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}. \quad (4.22)$$

Here ω_p is the metal plasma frequency and γ is the Drude dampening rate. Below the plasma frequency, the real part of the dielectric function will be negative and the material surface will consequently be strongly reflective[2]. Above the plasma frequency, the real part of the dielectric function becomes positive and the material will lose its metallic characteristics resulting in a significant loss of surface reflectivity. As a concrete example of a metal, Fig. 4.1b shows the frequency dependent Drude dielectric function for gold[87]. Within the Drude mode, the plasma frequency of gold is around 8.5 eV, and the surface will therefore be highly reflective below this frequency, and significantly less reflective above this frequency.¹

To illustrate the effect of the cavity geometry and the material properties on the electromagnetic environment, Fig. 4.1c shows the the cavity modes of a gold cavity with two different R_1 . It is clearly observed that the number of modes, as well as their spectral position is directly linked to the radius. Furthermore, above the plasma frequency of around 8.5 eV where the mirrors lose their reflectivity, the sharp mode structure is replaced by a broad continuum of modes. Finally, Fig. 4.1d zooms in on the mode around 7.1 eV in the cavity with $R = 140\text{nm}$ and shows the effect of changing the Drude dampening parameter. Here the width of the peak increases with increasing dampening in the gold which highlights the connection between the width of the cavity modes and the losses in the gold.

This example highlights how the use of the emitter centered representation of MQED allows us to directly relate the cavity field parameters to a real electromagnetic environment. While this is a relatively simple example, the approach is general and should work for an arbitrary electromagnetic environment provided that the DGF can be determined.

4.3 Adding an emitter to the cavity

We next add a benzene molecule to the cavity. Benzene is chosen mainly because of its prevalence as a test system in the existing TDDFT and QEDFT literature on strong coupling[77, 84, 88], but we emphasize that the method can treat systems with more electrons if needed. We focus on finding a cavity configurations with a mode resonant with the $\Pi \rightarrow \Pi^*$ transition of the benzene molecule. The first step is thus to determine the spectral position of this transition. Using the Casida framework presented in section 3.5 without photons, we find that this transition occurs at 6.808 eV (182 nm) in free space, consistent with previous TDDFT calculations for benzene[77, 84].

¹It should be noted that the Drude model is not a good model for the dielectric properties of gold due to the neglect of d-band absorption. In the future, better models for the metal will be explored.

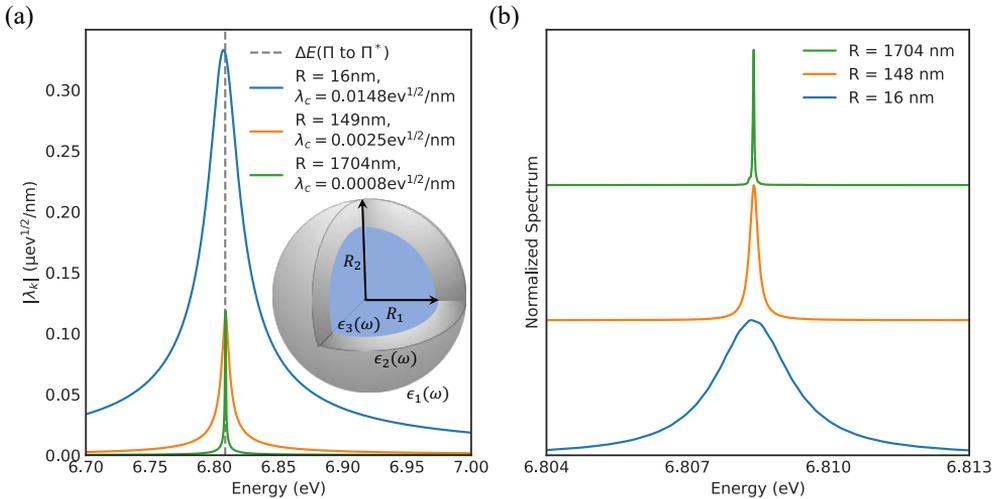


Figure 4.2: (a) Different radii of the spherical micro-cavity that host a resonance that aligns with the $\Pi \rightarrow \Pi^*$ of the benzene molecule. The total light-matter coupling strength of the modes is also shown in the legend, and it can be clearly observed that the cavity-coupling strength grows with decreasing cavity radius. (b) The linear absorption spectrum of benzene as a function of cavity radius showing a clear radius dependent Purcell enhancement.

As shown in Fig. 4.2a, it is possible to find different radii of the gold micro-cavity for which there is a cavity mode resonant with the benzene $\Pi \rightarrow \Pi^*$ transition. As expected, the cavity coupling strength increases as the cavity is made smaller. All but the smallest cavity are optical cavities in the sense that the characteristic dimension of the cavity, the radius, is larger than the wavelength of the transition. For the smallest cavity of radius 16 nm, a significant increase in the coupling strength relative to the other sizes is observed. This happens exactly because this cavity is sub-wavelength sized and therefore significantly near-field coupling to the surface plasmon mode of the gold starts to occur. Fig. 4.2b shows the linear absorption spectra of the coupled emitter-cavity system calculated for the different cavity radii using the linear response QEDFT method. We emphasize that all linewidths in the figure are *true* linewidths in the sense that they are not related to any broadening parameters in the QEDFT calculation and only reflect the density of states in the optical environment. A radius-dependent Purcell enhancement with decreasing radius is clearly observed, reflecting the reduction in radiative lifetime resulting from the altered optical environment. However, it is not possible to achieve strong coupling with a single benzene molecule using the gold-shell cavity. We attribute this to the fact that as the coupling strength gets larger with decreasing radius, the optical losses also increase, resulting in a broader cavity resonance. We note in passing that the Purcell enhancement for the smallest 16 nm cavity is around 3500 which means that the local field enhancement at the center of the spherical micro-cavity is comparable to what is found in experiments with

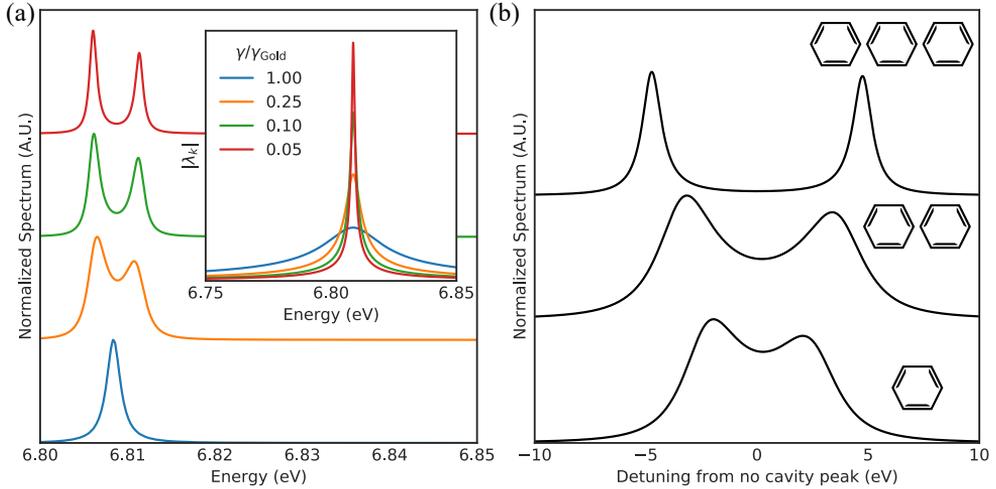


Figure 4.3: (a) The linear absorption spectrum of benzene as a function in the gold cavity as a function of the Drude damping parameter. (b) The linear absorption spectrum in a cavity with $\gamma = \gamma_{\text{gold}}/4$ as a function of the number of benzene molecules.

plasmonic microcavities.[89]

In an attempt to reach the SC regime, we next seek to reduce the losses in the cavity. As already discussed above, the width of the cavity resonance is reduced for smaller Drude damping parameters, γ . For this reason, Fig. 4.3a shows the absorption function for the case with the true gold damping, as well as 25%, 10% and 5% of the damping respectively. As shown in the inset of Fig. 4.3a, we find that reducing the losses results in a narrower cavity mode without a significant reduction in the cavity coupling strength. At around 25% of the true damping, we begin to observe indications of the two polariton peaks in the linear absorption spectrum. Further reducing the damping we see clear strong coupling with a Rabi splitting of around $2g \sim 10$ eV. This emphasizes the importance of the optical losses in reaching the strong coupling regime, and further highlights the significant strength of our method that we are able to study the effect of different cavity parameters from first principles via our combination of QEDFT and MQED.

Another way to engineer the coupling strength is by changing the emitter. There are two ways to do this, either by changing the emitter or by changing the number of emitters. As discussed in section 1.3, the effective coupling strength should scale with the square root of the number of emitters. To investigate this, we take the cavity with 25% of the true gold losses, and compute the absorption spectrum for 1, 2 and 3 benzene molecules. As shown in Fig. 4.3b, we see a clear evolution of the Rabi splitting with the number of benzene molecules. To perform this analysis, one needs to solve a coupled many electron, many photon problem and it again highlights the strength of the method.

Finally, it is possible to increase the coupling strength by exchanging the benzene molecule with another emitter with a larger transition dipole moment. To investigate this, we consider if we consider molecules with more benzene rings anthracene, naphthalene etc., and compute the spectrum and oscillator strength of the $\Pi \rightarrow \Pi^*$ -transition as a function of the number of rings using the Casida method without photons, as shown in Fig. 4.4. It can be seen that the transition dipole moment of the $\Pi \rightarrow \Pi^*$ -transition scales linearly with the length of the molecule. However, the position of the peak also changes significantly meaning that the cavity has to be re-optimized to be resonant with the transition. For the size systems where the calculations were feasible, we find that it is not possible to reach the strong coupling regime this way. However, it again highlights the strength of the approach that we can treat electronic systems of this size.

4.4 Conclusion

In conclusion, we have shown how a combination of MQED and QEDFT can be used to eliminate the arbitrariness of the cavity field parameters in QEDFT calculations.

We have used this approach to treat the case of benzene in a spherical micro-cavity, and showed that the effect of changes in the electromagnetic environment can be captured directly in our first principles calculations. To highlight the strength of the QEDFT method, we further showed that it is possible to treat many molecules coupled to a multi-mode photonic environment while accounting for the full electronic structure of

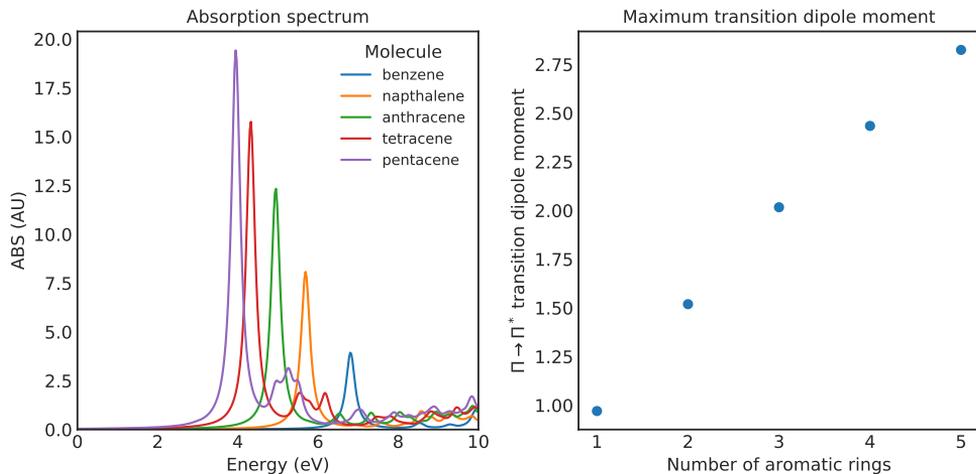


Figure 4.4: (a) The linear absorption spectrum of different aromatic compounds without any photon modes. (b) The transition dipole moment of the $\Pi \rightarrow \Pi^*$ -transition as a function of the number of aromatic rings.

the emitters.

This work sets is a step towards parameter-free QEDFT calculations and also opens open the possibility for proper treatment of real experimental setups. We emphasize that the connection between the optical environment and the DGF is not limited to cavity like setups but instead provides a general way to determine the electromagnetic spectral density of an arbitrary environment. In addition to the QED setup, our method therefore also provides a way to perform TDDFT in a lossy optical environment, potentially removing the need for artificial spectral broadening in such calculations.

4.5 Computational details for the QEDFT calculations

We use the publicly available real space psuedo-potential DFT code OCTOPUS[76]. Molecular geometries are optimized with the LDA exchange-correlation functional, on a real space grid consisting of spheres of radius 6\AA around each atom with a grid spacing of 0.16\AA .

We calculate the ground state on a real space grid consisting of spheres of radius 6\AA around each atom with a grid spacing of 0.08\AA . This grid configuration ensures that the ground state energy is converged to within 1 meV/atom .

For the Casida calculation, we need to converge the excited states. We perform a fixed-density calculation on top of the ground state calculation to determine the excited states. For the systems considered here, we find that the relevant transitions are well converged with 300 excited states included. We find that the coupling to the optical environment is well converged when sampling each cavity peak with around 1000 photon modes.

CHAPTER 5

Dielectric Screening in Materials

In much of condensed matter physics and chemistry, the quantum nature of the electromagnetic field can be neglected and replaced with a classical description. In this limit, the field is treated solely as an external perturbation, and of the four response functions of the full non-relativistic QED linear response theory only the density-density response function ($\chi_n^n(\mathbf{r}, \mathbf{r}', \omega)$) retains relevance. This limiting case is very important as it describes many different types of spectroscopy, as well as the dielectric response of materials. Since the density-density response function is the only relevant response function in this regime, the super- and subscripts are suppressed in the following.

This chapter explores the relation between dielectric screening and spectroscopic observables, with a discussion structured in the following way: Initially, the dielectric function is introduced and some general considerations about response functions in discretely periodic crystalline systems are outlined and the concept of *local field effects* is introduced. Secondly, a convenient mode decomposition of the dielectric function is introduced which allows for more detailed investigation of excitation spectra.

Subsequently, another way to describe the role of the dielectric functions is that it describes the *neutral* excitations of the system. The dielectric function is therefore intricately connected to the electronic structure of the material in question. Because the excitation spectrum of the system depends on both the occupied and unoccupied states, any proper description of the dielectric properties requires an accurate description of both the ground and excited states. Furthermore, electronic screening is also strongly dimensionality dependent and electronic screening therefore differs fundamentally in 2D and 3D materials. These relations are explored using Many Body Perturbation theory, and in this discussion proper approximations are also outlined.

Finally, the relation between the dielectric function and spectroscopic properties are discussed as it relates to the work in this thesis.

5.1 The Dielectric Function

When dealing with screening, the central response function is in fact not the density response function but instead the dielectric function, ϵ . Formally, the inverse dielectric function is defined as the linear response function relating the total potential to the

externally applied one,

$$v_{\text{tot}}(\mathbf{r}, \omega) = \int d\mathbf{r}' \epsilon^{-1}(\mathbf{r}, \mathbf{r}', \omega) v_{\text{ext}}(\mathbf{r}', \omega). \quad (5.1)$$

Because the total potential is the sum of the external potential and the potential induced by the material in response to the perturbation, ϵ encodes the material response, as is also the case for χ . In principle, spectroscopic observables such as e.g. optical absorption, electron energy loss spectra and refractive indices can be therefore formulated in terms of both χ and ϵ . The two response functions are in fact directly related to each other via,

$$\epsilon^{-1}(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r} - \mathbf{r}') + \int d\mathbf{r}'' v(\mathbf{r} - \mathbf{r}'') \chi(\mathbf{r}'', \mathbf{r}', \omega), \quad (5.2)$$

where $v(\mathbf{r})$ is the real space Coulomb interaction. However, in any spectroscopic experiment, the external potential is the quantity that is controlled while the total potential is the measured quantity. The dielectric function is therefore the more natural choice making it the central object in computational spectroscopy.

5.1.1 The macroscopic dielectric function and response functions in periodic systems

In most real experiments, the potential applied to the material varies on a length scale that is much larger than the size of the unit cell. (The exception to this rule of thumb is spatially resolved Electron Energy Loss Spectroscopy). However, because the material structure has variations on a microscopic scale, the potential induced in response to the external potential will nonetheless include microscopic variations on the size of the unit cell. These rapid spatial variations cannot be resolved experimentally, and the thing that is actually measured in an experiment is the average of the microscopic total potential over a unit cell,

$$V_{\text{tot}}(\mathbf{r}, \omega) = \frac{1}{\Omega_{\text{uc}}} \int_{\Omega(\mathbf{r})} d\mathbf{r}' v_{\text{tot}}(\mathbf{r}', \omega). \quad (5.3)$$

Even though these induced rapid spatial variations cannot be resolved in the experiment directly, they will impact the long wavelength component of the total potential and therefore affect the measurement. The impact of the microscopic variations on the macroscopic response is called *local field effects* and it is a very important concept in condensed matter physics.

Because the measured part of the total potential is an average over the unit cell, it is translationally invariant and it makes sense to define a macroscopic dielectric function, ϵ_M , which relates the macroscopic component of the total potential to the external potential,

$$V_{\text{tot}}(\mathbf{r}, \omega) = \int d\mathbf{r}' \epsilon_M^{-1}(\mathbf{r} - \mathbf{r}', \omega) v_{\text{ext}}(\mathbf{r}', \omega). \quad (5.4)$$

It is again important to emphasize that even though the microscopic variations of the field are not directly resolved in the experiment, they can have a profound impact on ϵ_M via the local field effects.

To develop a better understanding of this, consider the fact that the response functions of a periodic system has to obey the underlying translational invariance. For a crystal, this means that the displacement of any response function, f , by any lattice vector \mathbf{R} , has to leave it unchanged, $f(\mathbf{r}, \mathbf{r}') = f(\mathbf{r} + \mathbf{R}, \mathbf{r}' + \mathbf{R})$. Consequently, the inverse Fourier transform can be expressed as,

$$f(\mathbf{r}, \mathbf{r}') = \frac{1}{\Omega} \sum_{\mathbf{q}}^{\text{BZ}} \sum_{\mathbf{G}, \mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} f_{\mathbf{G}, \mathbf{G}'}(\mathbf{q}) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}, \quad (5.5)$$

where \mathbf{q} is the momentum transfer, \mathbf{G} and \mathbf{G}' are reciprocal lattice vectors, and Ω is the unit cell volume. The $\mathbf{G}, \mathbf{G}' = 0$ component represents the spatially slowly varying component induced directly by the perturbation with wave vector \mathbf{q} , and the $\mathbf{G}, \mathbf{G}' \neq 0$ components represents the components with rapid spatial variation induced by the periodicity of the crystal. The potentials can be similarly expanded in the plane wave basis as,

$$v_{(\text{ext}, \text{tot})}(\mathbf{r}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{q}}^{\text{BZ}} \sum_{\mathbf{G}} v_{\mathbf{G}}^{(\text{ext}, \text{tot})}(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}, \quad (5.6)$$

and Eq. 5.1 can be rewritten as,

$$v_{\mathbf{G}}^{\text{tot}}(\mathbf{q}, \omega) = \sum_{\mathbf{G}'} \epsilon_{\mathbf{G}, \mathbf{G}'}^{-1}(\mathbf{q}, \omega) v_{\mathbf{G}'}^{\text{ext}}(\mathbf{q}, \omega). \quad (5.7)$$

While the microscopic total potential will contain $\mathbf{G} \neq 0$ components, any external potential applied in a realistic spectroscopic experiment will not as it will vary on a length scale much larger than the crystal unit cell. Eq. 5.7 thus reads,

$$v_{\mathbf{G}}^{\text{tot}}(\mathbf{q}, \omega) = \epsilon_{\mathbf{G}, 0}^{-1}(\mathbf{q}, \omega) v_0^{\text{ext}}(\mathbf{q}, \omega), \quad (5.8)$$

in practice. Using Eq. 5.8, the expression for the macroscopic component of the total potential in Eq. 5.3 can be written as,

$$V_{\text{tot}}(\mathbf{r}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{q}}^{\text{BZ}} v_0^{\text{tot}}(\mathbf{q}, \omega) e^{-i\mathbf{q}\cdot\mathbf{r}}. \quad (5.9)$$

The reciprocal space representation of the macroscopic total potential is thus the $\mathbf{G} = 0$ component of $v_{\mathbf{G}}^{\text{tot}}$. Since the external potential only contains a long wavelength component, the reciprocal space representation of Eq. 5.3 becomes,

$$V_{\text{tot}}(\mathbf{q}, \omega) = \epsilon_{00}^{-1}(\mathbf{q}, \omega) v_{\text{ext}}(\mathbf{q}, \omega). \quad (5.10)$$

Comparing with Eq. 5.4, it is clear that the macroscopic dielectric function should be defined as,

$$\epsilon_M(\mathbf{q}, \omega) = \frac{1}{\epsilon_{00}^{-1}(\mathbf{q}, \omega)}. \quad (5.11)$$

Because $\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$ is a matrix it is important to stress that $\frac{1}{\epsilon_{00}^{-1}(\mathbf{q},\omega)} \neq \epsilon_{00}(\mathbf{q},\omega)$ and the macroscopic dielectric function should always be defined via Eq. 5.11. This reflects the fact that even though the rapid variations in the field induced in the material cannot be directly resolved in the experiment, they have real experimental consequences because they affect the macroscopic component of the total potential.

It is very important to note that the *macroscopic* dielectric function is not well defined for a 2D material in the optical limit ($q \rightarrow 0$)[90]. For a 2D material, the properties in the optical limit should be formulated in terms of the polarizabilities.

5.1.2 Spectral decomposition of the dielectric function

While the macroscopic dielectric function is the key object describing the spectroscopic properties of materials, it is not sufficient to fully characterize the excitations in a material. As an example, modes with odd spatial symmetry will *not* show up in the macroscopic EELS spectrum. In some cases it can therefore be beneficial to investigate the microscopic dielectric function in more detail.

A powerful tool to investigate the microscopic dielectric function is its spectral representation[91]. Because the dielectric function is generally non-Hermitian, its spectral representation takes the following form,

$$\epsilon(\mathbf{r},\mathbf{r}',\omega) = \sum_n \epsilon_n(\omega) \phi_n(\mathbf{r},\omega) \rho_n(\mathbf{r}',\omega). \quad (5.12)$$

Here $\epsilon_n(\omega)$ is n 'th complex eigenvalue, and $\phi_n(\mathbf{r},\omega)$, $\rho_n(\mathbf{r},\omega)$ the right and left eigenvectors respectively. The latter pair can be shown to satisfy Poisson's equation, $\nabla^2 \phi_n(\mathbf{r},\omega) = 4\pi \rho_n(\mathbf{r},\omega)$, and therefore represent the density and induced potentials associated with the n 'th mode of the dielectric function respectively. This leads to the interpretation of the complex eigenvalue $\epsilon_n(\omega)$ as the effective dielectric function of mode n .

For periodic systems, the eigenmodes and right eigenvectors of the dielectric function are determined by diagonalizing the dielectric matrix,

$$\sum_{\mathbf{G}'} \epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) \phi_{n,\mathbf{G}'}(\mathbf{q},\omega) = \epsilon_n(\mathbf{q},\omega) \phi_{n,\mathbf{G}}(\mathbf{q},\omega). \quad (5.13)$$

The dual vectors (the induced densities), $\rho_{n,\mathbf{G}}(\mathbf{q},\omega)$, can be found by inverting the matrix $\Phi_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$, the columns of which are made up of $\phi_{\mathbf{G}}(\mathbf{q},\omega)$. The real space representation of the induced potentials and densities associated with the modes can be found by means of a Fourier transformation, and this allows for visualization of the modes. Furthermore by studying the evolution of $\epsilon_n(\mathbf{q},\omega)$ as a function of \mathbf{q} it becomes possible to directly probe the momentum transfer dependence of the excitations in the system. In Publication [III], the spectral representation was used to study the unique way in which the reduced screening in 2D impacts plasmon formation.

5.2 Many Body Perturbation Theory and Dielectric Screening

The dielectric function is a linear response function, and as already discussed in section 3.1 on linear response and the Kubo formula, any linear response function can be written in terms of the retarded correlation functions of the system in equilibrium. Since these correlation functions have a Lehmann representation, it means that a solid understanding of these can be developed if they can be discussed in terms of a theory that properly represents the electronic structure of the system. Importantly, any such theory would have to simultaneously account for the ground and excited states of the system. Ground state DFT is constructed to be an exact *ground state* theory, and it is known that even with the exact functional the bandstructure is not supposed to be exact because DFT does not correctly account for the derivative discontinuity. This is the origin of the so-called bandgap problem which haunts most DFT functionals, and in practice it also haunts linear response TDDFT because it builds on top of the ground state calculation. The problem can be remedied somewhat with the hybrid functionals, or by explicitly evaluating the derivative discontinuity as done in GLLBSC, but in practice it is nice to have a method that correctly describes both the ground- and excited states by construction.

In order to properly describe the single particle energies of a system, one should seek a method that can properly account for both the ground and excited state of the system. To this end, this section discusses Many Body Perturbation Theory (MBPT) in an attempt to develop a better understanding of the screening processes in a material. This discussion allows for the natural discussion of important concepts such as the GW-approximation, the Random Phase approximation, and the Bethe-Salpeter equation.

5.2.1 The single particle Green's Function and Quasi Particles

The central object in MBPT is the single particle GF,

$$G(\mathbf{r}, t; \mathbf{r}' t') = \langle N | \hat{\mathcal{T}} \left\{ \hat{\psi}(\mathbf{r}, t) \hat{\psi}^\dagger(\mathbf{r}', t') \right\} | N \rangle. \quad (5.14)$$

Here $\hat{\psi}(\mathbf{r})$ is a fermionic field operator, $|N\rangle$ is the N -particle ground state of the interacting system, and $\hat{\mathcal{T}}\{\circ\}$ denotes time ordering of the operators. Physically, the single particle GF propagates a single particle added to the system from one spacetime point to another. Depending on the time ordering in Eq. 5.14, the particle in questions is either an electron or a hole. For $(t' < t)$, Eq. 5.14 corresponds to the amplitude that an electron added at (\mathbf{r}', t') propagates to (\mathbf{r}, t) where it is annihilated. Contrastingly, for $(t' > t)$, Eq. 5.14 corresponds to the amplitude that a hole added at (\mathbf{r}, t) propagates to (\mathbf{r}', t') where it is annihilated. The definition of the GF in Eq. 5.14 is therefore able to describe all, single particle charged excitations in the system.

In frequency space, the single particle GF has a Lehmann representation in terms of

the charged single particle excitations of the system,

$$G(\mathbf{r}, \mathbf{r}', \omega) = \sum_j \frac{\psi_j(\mathbf{r})^* \psi_j(\mathbf{r}')}{\omega - \epsilon_j + i\eta \text{sgn}(e_j - E_F)}, \quad (5.15)$$

where,

$$\epsilon_j = E(N) - E(N - 1, j), \quad \psi_j(\mathbf{r}) = \langle N - 1, j | \hat{\psi}(\mathbf{r}) | N \rangle, \quad \text{For } \epsilon_j < E_F, \quad (5.16)$$

$$\epsilon_j = E(N + 1, j) - E(N), \quad \psi_j(\mathbf{r}) = \langle N | \hat{\psi}(\mathbf{r}) | N + 1, j \rangle, \quad \text{For } \epsilon_j > E_F. \quad (5.17)$$

Here $E(N)$ is the energy of the N -particle ground state, $|N - 1, j\rangle$ and $E(N - 1, j)$ are the state and energy of the system with one electron removed from state j , and $|N + 1, j\rangle$ and $E(N + 1, j)$ are the state and energy of the system with one electron added to state j . It can be seen that the single particle GF has poles exactly at the addition and removal energies of the system, and it is related to the intrinsic spectral function of the system,

$$A(\mathbf{r}, \mathbf{r}', \omega) = \pi^{-1} \text{Im} G(\mathbf{r}, \mathbf{r}', \omega) \text{sgn}(E_F - \omega). \quad (5.18)$$

Within the sudden approximation which disregards complications related to the propagation from sample to detector, the intrinsic spectral function is directly related to what would be measured in for example Photoemission Spectroscopy (PES) and Inverse PES (IPES) respectively [92, 93]. The thing that is measured with these techniques is really the energies that it takes to remove or add an extra electron to the system, and with a reliable way to calculate the GF of the interacting many-body system it is therefore possible to determine the true addition and removal energies of the system.

For a non-interacting system, the ground state of the system is a single Slater determinant, and consequently the GF is made up of a series of separated delta function poles representing true single particle excitations in the system. Another way to think of this is that in the non-interacting system, the single particle excitations do not feel their surroundings and they have an infinite lifetime. In an interacting system, the ground state cannot be represented by a single Slater determinant, and as a result the poles can be grouped into structures that resemble broadened peaks. This means that the excitations are no longer true single particle excitations but instead represent many-body excitations of the interacting system. In some case, the peaks look like simple broadened Lorentzian peaks and in this case they can be represented as first order poles in the complex plane. Such structures are functionally similar to charged single particle excitations in non-interacting systems, and they represent single-particle like excitations of the interacting system. For this reason, they are called almost particles or *quasi-particles*. A physical explanation of such peaks is that sometimes the effect of the interactions is merely to *dress* the single particles. For example, a negatively charged electron introduced into the system can repel the electrons around it. This results in a buildup of positive charge around the negatively charged electron, effectively reducing its charge. The single particle is thus "dressed" by the interactions, resulting in something that acts like a single particle would do, but is something more complicated. Near such quasi-particle peaks, the GF is often expanded into a quasi-particle

part and a multi-particle part. The quasi-particle part reads,

$$G^{\text{QP}}(\mathbf{r}, \mathbf{r}', \omega) = \sum_j \frac{\psi_j^{\text{QP}}(\mathbf{r})^* \psi_j^{\text{QP}}(\mathbf{r}')}{\omega - \epsilon_j^{\text{QP}}} \quad (5.19)$$

where the quasi-particle energies are generally complex ϵ_j^{QP} to reflect the lifetime broadening resulting from the interactions in the system. The interactions can thus both shift the energy of the excitation, and result in a finite lifetime reflected by the width of the peak. The norm of the quasi-particle states is not unity but instead given by the so-called quasi-particle weight Z_j . If Z_j is close to one, the systems excitations are well described as single particle like and the quasi-particle picture applies. If Z_j is small, it means that the excitations of the system are more multi-particle like. In such cases, the multi-particle part of the GF carries most of the information and the weight is flushed into satellites. This reflects stronger correlation in the system and the quasi-particle picture becomes in appropriate.

5.2.2 The Self-Energy and Hedin's equations

Actual calculation of the single particle GF relies on the equations of motion technique. In this approach, the one particle Green's function of an interacting system is related to the two-particle Green's function, which is related to the three particle Green's function and so on to the infinite particle GF in what is called the Martin-Swinger Hierarchy[94]. This connection between the N and $N - 1$ particle GFs reflects the presence of interactions in the system.

While formally interesting, the infinite series is not practical for actual calculations, and as a formal way of truncating the series the so-called *self-energy*, Σ , is defined such that the equation of motion for the single particle GF can be written as a Dyson equation,

$$G(1, 2) = G_0(1, 2) + \int d3d4 G_0(1, 3) \Sigma(3, 4) G(4, 2). \quad (5.20)$$

In the above, the standard notation $(N) = (\mathbf{r}_N, t_N)$ has been introduced. Here, G_0 is defined so as to contain the kinetic energy, the external potential and the Hartree potential. These terms are the easy ones that can be treated without approximations, as discussed in section 2.1. In that way, the self-energy is constructed to capture the effect of exchange and correlation, and its interpretation is therefore that it is the correction to the single-particle Hamiltonian due to the presence of inter-particle interactions in the system. Importantly, the self-energy is a frequency dependent quantity reflecting its relation to many-particle effects[93].

The GF and the self-energy are themselves related to a set of other 2 and 3 point

GFs through the so-called Hedin's equations[92, 95],

$$\Sigma(1, 2) = i \int d(34) G(1, 3) \Gamma(3, 2, 4) W(4, 1^+), \quad (5.21)$$

$$W(1, 2) = V(1, 2) + \int d(34) V(1, 3) P(3, 4) W(4, 2), \quad (5.22)$$

$$P(1, 2) = -i \int d(34) G(1, 3) G(4, 1^+) \Gamma(3, 4, 2), \quad (5.23)$$

$$\Gamma(1, 2, 3) = \delta(1, 2) \delta(1, 3) + \int d(4567) \frac{\delta \Sigma(1, 2)}{\delta G(4, 5)} G(4, 6) G(7, 5) \Gamma(6, 7, 3). \quad (5.24)$$

Here W is the screened Coulomb interaction which in terms of the dielectric function can be expressed as $W(1, 2) = \int d(3) \epsilon^{-1}(1, 3) V(3, 2)$. P is the irreducible polarizability representing all the possible screening processes in the material. It can be thought of as the set of fundamental process by which the system is able to polarize. The actual polarization of the system can be made up of any number of repetitions of these processes and is described by the full or reducible polarization, χ . The reducible and irreducible polarizations are related via a Dyson equation,

$$\chi(1, 2) = P(1, 2) + \int d(34) P(1, 3) V(3, 4) \chi(4, 2). \quad (5.25)$$

Finally, Γ is the interaction vertex of the system, and it describes a renormalization of the electron-photon interaction. In principle, Hedin's equations capture the full range of interactions in the system, and if iteratively solved they very accurately describe the ground and excited states of the system. However, in practice it is impossible to solve the full set of equations and suitable approximations have to be made. Fortunately, discussing these approximations also brings insight into the different screening processes that occur in a material.

5.2.3 The GW approximation

The most famous, and to the best of the authors knowledge, first approximation to Hedin's equations is the so-called *GW* approximation[95]. Within the *GW* approximation, the vertex corrections are neglected completely, $\Gamma(1, 2, 3) \approx \delta(1, 2) \delta(1, 3)$, in which case the rest of Hedin's equations reduce to,

$$\Sigma(1, 2) = iG(1, 2)W(2, 1^+), \quad (5.26)$$

$$W(1, 2) = V(1, 2) + \int d(34) V(1, 3) P(3, 4) W(4, 2), \quad (5.27)$$

$$P(1, 2) = -iG(1, 2)G(2, 1^+). \quad (5.28)$$

It is seen that the self energy reduces to the product of the Green's function (G) and the screened interaction (W), which gives the approximation its name. From these equations, it is seen that the *GW*-approximation can be understood a dynamically screened version of Hartree-Fock (HF). Indeed, replacing W with V in the expression

for Σ on recovers the standard HF expression. HF can there be thought of as the lowest order GW-approximation in the interaction. The inclusion of the dynamic screening effects in the full GW approximation accounts for the fact that the screening in the material reduces the direct Coulomb interaction between the particles, and the dynamical nature accounts for the fact that the system is interacting. In practice, the GW approximation greatly improves of HF and gives a good description of things like band gaps of solids[93].

The simplified Hedin's equations can now be solved iteratively in principle. While possible, in practice one often relies on the further so-called G_0W_0 approximation, where the equations are only iterated a once. The G_0W_0 approximation has the dual benefit of being computationally cheaper, and also often giving better results than the full self-consistent GW calculation. This happens because it turns out that the vertex corrections are important to include in the self-consistent scheme[92, 93]. Within the simplest G_0W_0 approximation, Hedin's equation are not iterated and the correction to the non-interacting single particle energies are evaluated using first order perturbation theory. In practice, the starting point for the G_0W_0 calculation is often the set of Kohn-Sham states and energies from a DFT calculation. For example, in a periodic system where the states can be labeled by a band index, n , and a wave vector from the first Brillouin Zone, \mathbf{k} , the energy correction is evaluated as,

$$\epsilon_{n\mathbf{k}}^{\text{QP}} = \epsilon_{n\mathbf{k}}^{\text{KS}} + Z_{n\mathbf{k}} \langle \psi_{n\mathbf{k}}^{\text{KS}} | \hat{\Sigma}(\epsilon_{n\mathbf{k}}^{\text{KS}}) - \hat{V}_{\text{xc}} | \psi_{n\mathbf{k}}^{\text{KS}} \rangle \quad (5.29)$$

Note importantly that the exchange correlation energy used in the DFT calculation is subtracted from the self-energy in Eq. 5.29. As discussed in section 5.2.2, the self-energy is there to account for exactly the exchange and correlation and this is therefore necessary to avoid double counting.

The main convergence parameters in G_0W_0 for solids are the k-point grid used in the calculation, as well as the plane wave cut-off energy and number of bands used for the calculation of the screened interaction. In practice, the number of bands is usually converged directly, while the cut-off energy and the k-point density are usually extrapolated to infinity.

5.2.4 The Random Phase approximation

It is also interesting to take a extra look at the irreducible polarization in these simplified equations to understand which screening processes are included at this level of approximation. Within the GW approximation, the $P(1,2)$ reduces to the product of two Green's functions, essentially representing an electron-hole pair created at 1 and annihilated at 2 without interacting with each other. This means that all multi-particle processes that could contribute to the screening are neglected within the GW-approximation, and it is a consequence of neglecting $\delta\Sigma/\delta G$. Throwing this term away means that the response of the polarization to the excitations in the screening process is neglected[92]. Another way of arriving at the same result is to state that because all multi-particle contributions to the screening have rapidly varying phase these will contribute less than the single particle ones, giving the approximation its name: The Random Phase Approximation (RPA). As discussed in section 3.4, the RPA is also

obtained when neglecting the exchange correlation kernels in linear response TDDFT and linear response QEDFT. The above arguments now explain why the RPA only captures single particle contributions to the screening process. It is important to note that the form of the RPA calculated within the linear response DFT formalisms is actually slightly more approximate, because, in the language of Hedin's equations, it constructs the polarization in terms of the GF of the Kohn-Sham system, which would be G_0 instead of the full G .

The remaining discussion point is the validity of the RPA. In metals, the screening is very large, and an electron and a hole quasi particle in the material will therefore only interact weakly. For this reason, the RPA expression for P is often sufficient to describe metallic systems and it can account for the plasmonic excitations that dominate in those systems. In contrast, in semi-conductors and insulators, the screening is much weaker than in metals, and in general an electron and a hole will interact significantly. This mutual interaction gives rise to excitons, and properly accounting for these is critical to obtain a proper description of the dielectric properties of semi-conductors and insulators. This is especially true in 2D due to the reduced screening, as discussed further in section 5.2.6. However, the applicability of the RPA depends strongly on the questions asked. Near the band-edge of the semiconductors where the excitonic effects manifest directly it is not good approximation, but it does perform relatively well for example for static screening in 3D semiconductors. These points are elaborated more extensively in Publication [II] with respect to 3D semi-conductors, and in Publication [III] for 2D materials.

5.2.5 Vertex corrections and the Bethe-Salpeter Equation

As mentioned above, within the GW approximation the screening is evaluated within the RPA where multi particle processes contributing to the irreducible polarization are neglected. This was a consequence of neglecting the vertex corrections by assuming $\frac{\delta \Sigma}{\delta G} = 0$, and it can be understood physically as neglecting the polarization response to the virtual excitations involved in the screening process. To give an example of why this is approximate: Within the RPA, the polarization does not respond to the electron-hole pairs that constitute the fundamental building block of the screening process. In reality, this would generally not be true, as the negatively charged electron and the positively charged hole would attract and form a two-particle bound state called an exciton. As mentioned above, this two-particle state would have a lower energy than the two quasi-particles separately, and it can have a strong impact on physical properties and result in e.g. optical absorption onset below the quasi-particle bandgap.

To take the vertex corrections into account, the vertex equation in Hedin's equation can be iterated once with Σ evaluated at the GW level. It is commonly assumed that the alteration of the screening by the excitations is negligible and that the functional derivative of the self energy can be approximated as [92],

$$\frac{\delta \Sigma}{\delta G} = i \frac{\delta}{\delta G} (GW) \approx iW. \quad (5.30)$$

Within this approximation, the Hedin's equation for the polarizability (Eq. 5.23) can be reformulated with this lowest order vertex taken into account. It turns out that the kernel in this new Dyson equation is a four point function, and it therefore makes most sense to formulate the Dyson equation in terms of a four point polarizability function ${}^4P(1, 2, 3, 4)$. Since the physically relevant quantity to understand experiments is not the irreducible polarizability P , but the reducible polarizability χ , it makes sense to formulate the so-called Bethe-Salpeter Equation (BSE) for χ ,

$${}^4\chi(1, 2, 3, 4) = {}^4P_{\text{QP}}(1, 2, 3, 4) + \int d(5678) {}^4P_{\text{QP}}(1, 2, 5, 6) K(5, 6, 7, 8) {}^4\chi(7, 8, 3, 4). \quad (5.31)$$

where ${}^4P_{\text{QP}}(1, 2, 3, 4) = \delta(1, 2)\delta(3, 4)P(1, 3)$ is the four-point version of the RPA polarizability, and the two point reducible polarization can be obtained from ${}^4\chi$ as ${}^4\chi(1, 1, 2, 2)$. The interaction kernel in Eq. 5.31 describes the Coulomb interaction,

$$K(1, 2, 3, 4) = \delta(1, 2)\delta(3, 4)V(1, 3) - \delta(1, 3)\delta(2, 4)W(1, 2). \quad (5.32)$$

Unlike the RPA, there are now two contributions to this kernel which can be seen to connect differently. The first term in K can be understood as an exchange interaction, whereas the second can be understood as a direct interaction connecting the electron and hole Green's Functions in the screening bubble. By dropping the second term, the interaction between the quasi-particles is lost and one recovers the RPA. Interestingly, the direct interaction term is screened whereas the exchange interaction is not. This can be understood from the origin of the two terms. One way to think about it is that the exchange is the lowest order polarization response, and it comes from the approximation that the polarization of the system does not react to the polarization. For this reason it is not screened. Contrastingly, the direct interaction comes from the polarization response to the virtual excitation processes and it is therefore screened.

Solving the BSE directly is very expensive and often the further approximation that the screened interaction can be approximated by its static limit has to be made. In that case, the problem can be reformulated in terms of an effective two particle Hamiltonian. In particular, for extended systems, the excitation energies at crystal momentum \mathbf{q} can be found by solving a Casida-like eigenvalue problem,

$$\sum_{S'} \mathcal{H}_{SS'}(\mathbf{q}) F_{S'}^\lambda(\mathbf{q}) = E^\lambda(\mathbf{q}) F_S^\lambda(\mathbf{q}), \quad (5.33)$$

$$\mathcal{H}_{SS'}(\mathbf{q}) = \left(\epsilon_{m\mathbf{k}+\mathbf{q}}^{\text{QP}} - \epsilon_{n\mathbf{k}}^{\text{QP}} \right) \delta_{SS'} - (f_{m\mathbf{k}+\mathbf{q}} - f_{n\mathbf{k}}) K_{SS'}(\mathbf{q}), \quad (5.34)$$

where $f_{n\mathbf{k}}$ is the occupation function of band n and wavevector \mathbf{k} , and the Hamiltonian is formed in a basis of two particle eigenstates made from the quasi-particle states of the system, $\psi_S(\mathbf{r}_e, \mathbf{r}_h) = \psi_{n\mathbf{k}}^{\text{QP}}(\mathbf{r}_h) \psi_{m\mathbf{k}+\mathbf{q}}^{\text{QP}}(\mathbf{r}_e)$, labelled by $S = (n\mathbf{k}, m\mathbf{k} + \mathbf{q})$. The interaction kernel in Eq. 5.34 again inherits the structure from the full frequency dependent kernel and contains an exchange term and a direct interaction term,

$$K_{SS'}(\mathbf{q}) = 2V_{SS'}(\mathbf{q}) - W_{SS'}(\mathbf{q}), \quad (5.35)$$

where the factor of two accounts for spin. With this formulation it is possible to investigate the internal structure of these two terms, and it turns out that the q -dependence of these two terms very different and strongly dimensionality dependent. The result is fundamental differences in how excitation formation works in 2D and 3D, as discussed in Publication [III]. This thesis returns to this point in section 5.2.6 below.

In terms of the eigenvalues and eigenvectors of the two-particle Hamiltonian, the polarizability can be expanded as,

$$\chi(\mathbf{q}, \omega) = \sum_{\lambda\lambda'} \frac{N_{\lambda\lambda'}^{-1}(\mathbf{q}) F_{\lambda}^S(\mathbf{q}) [F_{\lambda'}^{S'}(\mathbf{q})]^*}{\omega - E_{\lambda}(\mathbf{q}) + i\eta}. \quad (5.36)$$

In practice, it is favourable to restrict the transitions to the positive frequency transitions. This approximation is referred to as the Tamm-Dancoff approximation, and it has the clear advantage of making the two-particle Hamiltonian Hermitian thus rendering the normalisation trivial, $N_{\lambda\lambda'} = \delta_{\lambda\lambda'}$. While the effect of including the excitonic effects are usually negligible in metals, it can have significant effect on the response of semiconductors in both 2D and 3D. The effect is particularly striking in 2D materials where the reduced screening can result in very large exciton binding energies. An example from this work is the EELS spectrum of monolayer PtSe₂ presented in Publication [III]. As shown in Fig. 9.2a, the excitonic effects result in a renormalization of the lowest peak position by over -0.5 eV.

Finally, the eigenstates of the two-particle Hamiltonian are also a source of physical information about the system. In terms of the eigenfunctions, the real space exciton wave function can be written as, $\Psi_{\lambda}^{\text{ex}}(\mathbf{r}_e, \mathbf{r}_h; \mathbf{q}) = \sum_S A_{\lambda}^S(\mathbf{q}) \psi_S(\mathbf{r}_e, \mathbf{r}_h)$. Therefore, by plotting $A_{\lambda}^S(\mathbf{q})$ on the quasi particle bandstructure it is possible to determine which transitions contribute most to a given excitonic state, λ . Fig. 5.1 shows this for monolayer MoS₂ and monolayer PtSe₂.

5.2.6 The direct and exchange interaction in 2D vs 3D and its impact of the formation of excitations

To understand the difference between the Coulomb interaction in 2D and 3D, it is instructive to consider the quasi-2D Coulomb interaction for two line charges in a slab of thickness, d . In Fourier space, it takes the following form[90],

$$V_{\text{Q2D}}(\mathbf{q}_{\parallel}) = \frac{4\pi}{q_{\parallel}^2 d} \left[1 - \frac{2}{q_{\parallel} d} e^{-q_{\parallel} d/2} \sinh\left(\frac{q_{\parallel} d}{2}\right) \right]. \quad (5.37)$$

In the 3D limit ($q_{\parallel} d \gg 1$), the Fourier space Coulomb interaction is the well-known $V(\mathbf{q}_{\parallel}) = \frac{4\pi}{q_{\parallel}^2}$. However, in the 2D limit ($q_{\parallel} d \gg 1$), the the Fourier space Coulomb interaction is $V(\mathbf{q}_{\parallel}) = \frac{2\pi}{q_{\parallel}}$. This difference reflect the reduced screening in 2D, and it has very significant consequences for the formation of plasmonic- and excitonic excitations in materials, especially in the optical ($q \rightarrow 0$) limit. In particular, it impacts the direct interaction and exchange interactions quite differently.

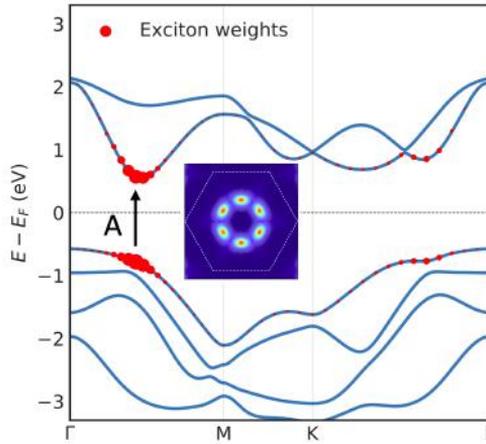


Figure 5.1: Figuring showing the exciton wave functions from BSE for the lowest energy exciton in a monolayer of PtSe₂ projected onto the electronic band structure. Figure reproduced from Publication [III].

As discussed in Publication [III] and its Supplementary information, the exchange interaction term is responsible for plasmon formation, and the direct screened interaction term is responsible for exciton formation. While W is relatively independent of \mathbf{q} in the optical limit for both 2D and 3D, the \mathbf{q} -dependence of V is strongly dimensionality dependent near the optical limit. Specifically, in 3D, the long range component of V is relatively \mathbf{q} -independent in the optical limit, while it scales as $\sim \mathbf{q}$ in 2D. Effectively therefore, the long-range Coulomb interaction is turned off in this limit for 2D systems, see Fig. 5.2. This results in the absence of plasmons formation in the optical limit for 2D materials. This is sharp contrast to the 3D case, where plasmon formation persists, and often even dominates, in the optical limit. Consequently, the excitation spectrum of 2D materials in the optical limit is dominated by excitons, the effect of which is already significantly enhanced due to the reduced screening of the direct interaction between the electron and the hole in 2D systems.

5.3 Spectroscopic quantities and dielectric screening

As alluded to in the previous sections, several important spectroscopic properties can be calculated from the dielectric function. In this section, the properties relevant to the work performed in this thesis are discussed.

5.3.1 Electron Energy Loss Spectroscopy

One particularly important type of spectroscopy for the purposes of this thesis is Electron Energy Loss Spectroscopy (EELS). In EELS, a beam of electrons with a known, narrow range of energies is shined onto the sample. By measuring the change in the

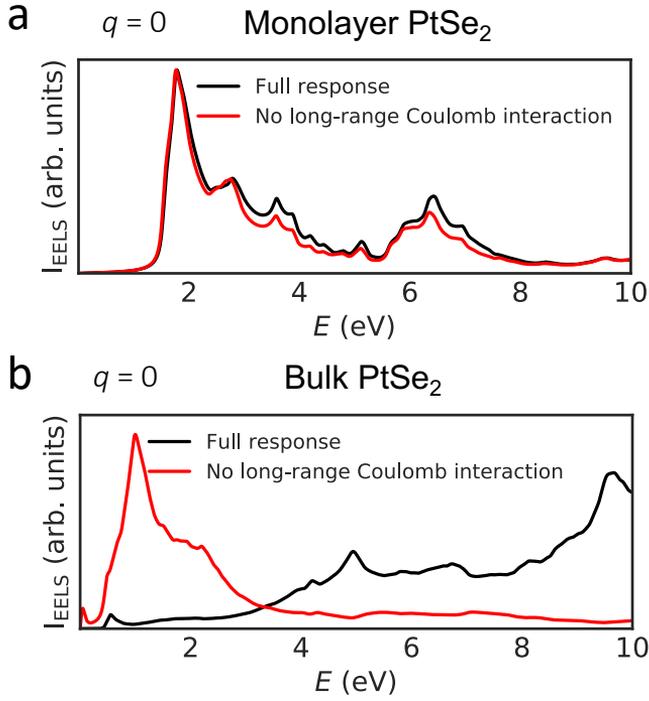


Figure 5.2: Figuring showing the importance of the long range Coulomb interaction in the optical limit in 2D vs 3D.

electrons energy and momentum after they interact with the sample, it is possible to deduce a lot of information about the electronic excitations in the sample. The electrons in EELS usually have energies in the keV range, and it is generally possible to probe both low energy excitations such as plasmons and excitons, as well as high energy core transitions. The former provides information about the chemistry and dielectric excitations and is called low energy EELS, and the latter provides information about the elemental composition and is called high energy EELS[96]. This thesis is only concerned with low energy EELS up to excitation energies of around 10 eV. In the work leading to this thesis, two different types of EELS were considered: Momentum resolved EELS (q -EELS) and spatially resolved EELS (s -EELS). The difference between these two modes of operation correspond roughly to the difference between operating the electron microscope in TEM or STEM mode[97].

q -resolved EELS: In q -EELS the incoming electron can be thought of as a plane wave and the EELS can be calculated from the dielectric function,

$$\text{EELS}(\mathbf{q}, \omega) = -\text{Im} [\epsilon_M^{-1}(\mathbf{q}, \omega)]. \quad (5.38)$$

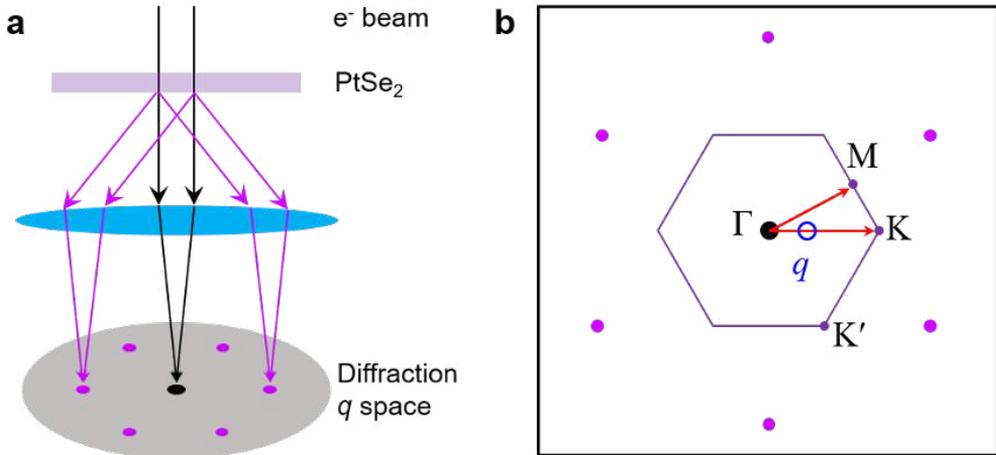


Figure 5.3: Setup for the q -EELS measurements: Depending on the deflection angle of the electrons after the interaction with the sample, it is possible to determine the momentum transferred to the sample. By collecting the transmitted with a magnetic lens, as shown in (a), one obtains the q -resolved signal in the form of a set of diffraction points, as shown in (b). The spectra at different q can therefore be obtained by looking at the energy distribution in the different diffraction peaks.

q -EELS is a very powerful tool to probe the dispersion of a material's excitation spectrum. However, the high q -resolution means that it has limited spatial resolution, and it is therefore not suited for probing e.g. morphology dependent resonances in nano-structures or edge effects.

This technique was used in Publication [III] to probe the q -dependent excitation spectrum of layered PtSe₂ both theoretically and experimentally. In the experiment, a relatively wide, collimated beam of electrons were shined onto the sample. Depending on the deflection angle of the electrons after the interaction with the sample, it is possible to determine the momentum transferred to the sample. By collecting the transmitted with a magnetic lens as illustrated in Fig. 5.3a, one obtains the q -resolved signal in the form of a set of diffraction points, see Fig. 5.3b. The spectra at different q can therefore be obtained by looking at the energy distribution in the different diffraction peaks.

Spatially resolved EELS: In s-EELS the electron beam is focused onto the sample at it is therefore best described as a moving point charge. In this case, the measured EEL spectrum is a convolution of many q and this method thus cannot be used to probe the dispersion of excitations in a system like q -EELS. However, because the electric field of the moving point charge is evanescent it provides a very good spatial resolution [96, 97]. This the excellent spatial resolution means that the technique can be used to for example spatially map out the optical resonances of nano-structures. Interestingly, in

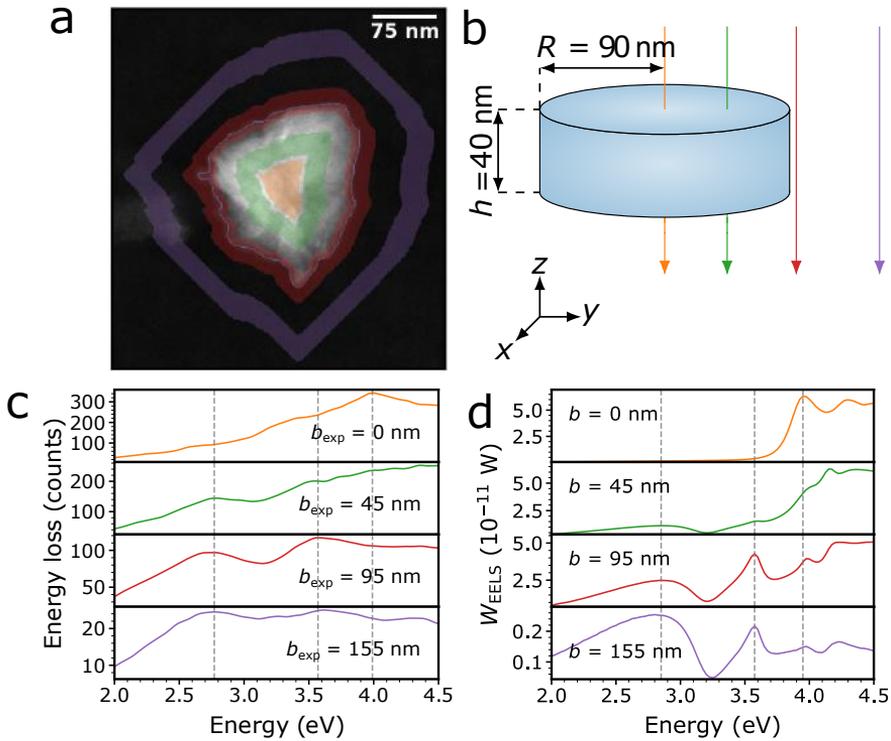


Figure 5.4: Figure showing how different beam positions can be probed via different beam positions in spatially resolved EELS. Reproduced from Publication [II].

addition to the material composition and geometry of the nano-structure, the loss rate observed in the measurement will also depend strongly on the beam position[96]. Thus by probing the structure a different beam positions different modes can be excited. In Publication [II], s-EELS was used in exactly this fashion to probe the size and morphology dependent resonances in dielectric nano-structures of Boron Phosphide (BP). This is shown in Fig. 5.4 reproduced from the publication.

For completeness, it makes sense to mention that for structures significantly smaller than the wavelength of light, the non-retarded formulation in terms of the incident scalar potential is sufficient. However, if the size of the structures are comparable to the wavelength of light, which was the case in Publication [II], retardation effects must be included. In practice, one way to deal with this is to reformulate the expression in terms of the DGF, see Ref.[96].

5.3.2 Optical absorption

Another important spectroscopic quantity that can be calculated from the dielectric function is optical absorption. The absorption is related to the energy dissipated in

the material by the current that the external field induces. In 3D, the absorption experienced by a plane wave perturbation with momentum \mathbf{q} and energy $\hbar\omega$ is directly related to the macroscopic dielectric function,

$$\text{ABS}^{3\text{D}}(\mathbf{q}, \omega) = \text{Im} \{ \epsilon_{\text{M}}(\mathbf{q}, \omega) \} \quad (5.39)$$

Since optical fields carry negligible momentum, the optical absorption can then be defined as,

$$\text{ABS}^{3\text{D}}(\omega) = \lim_{q \rightarrow 0} \text{Im} \{ \epsilon_{\text{M}}(\mathbf{q}, \omega) \}. \quad (5.40)$$

As already discussed in Chapter 3, it is important to reiterate that the theory outlined here is not able to describe a general transverse perturbation since such a perturbation cannot be expressed by a scalar potential. However, it turns out that in the long wavelength limit, the longitudinal and transverse response functions coincide[81]. Because photons in the optical and near UV regimes carry negligible momentum, the optical response can thus be studied from the longitudinal response function.

As mention in section 5.1.1, the macroscopic dielectric function is not well defined for 2D systems in the optical limit. Instead it is necessary to formulate the optical absorption in terms of the polarizability. Specifically, the absorption can be calculated from the macroscopic polarizability of the 2D system,

$$\text{ABS}^{2\text{D}}(\omega) = - \lim_{q \rightarrow 0} \left\{ \frac{4\pi}{q^2} \text{Im} [\chi_{\text{M}}^{2\text{D}}(\mathbf{q}, \omega)] \right\}. \quad (5.41)$$

In practice, the GPAW code operates with 3D periodic boundary conditions even when dealing with 2D materials. The 2D version of the macroscopic polarizability therefore has to be calculated from the macroscopic polarization of the full cell, and this is done by integrating the macroscopic component of the cell polarizability over the width of the 2D layer.

5.3.3 Refractive index

The refractive index is the ratio between the speed of light in vacuum and the effective speed at which light moves inside of a material, $n(\omega) = c/v(\omega)$. For a bulk materials, the complex refractive index can be calculated as,

$$\eta(\omega) = \sqrt{\lim_{q \rightarrow 0} \{ \epsilon_{\text{M}}(\mathbf{q}, \omega) \}}. \quad (5.42)$$

The real part of $\eta(\omega)$ is the refractive index, and the imaginary part, $k(\omega)$, is the extinction coefficient which accounts for absorption. Because the complex refractive index is given by the square root of the dielectric function, both the real and imaginary part of the dielectric function has to be well captured in order to accurately describe the refractive index. This is contrast to the optical absorption which only needs the imaginary part to be converged up to the energy of interest.

Similar to optical absorption, it is necessary to include excitonic effects in the calculation to accurately determine the frequency dependent refractive index of a semiconductor. This is illustrated in Fig. 5.5 where both RPA@G₀W₀ and BSE@G₀W₀

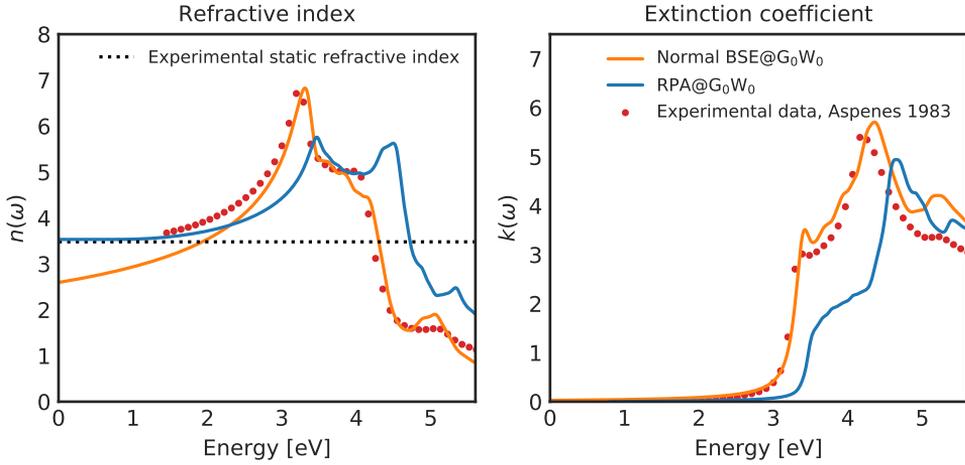


Figure 5.5: The refractive index and extinction coefficient for bulk Silicon calculated using the RPA@G₀W₀ and BSE@G₀W₀ methods compared to the experimental values from Ref. [98].

is compared to experimental data for bulk Silicon[98]. It is clearly seen that both the detailed features near the band edge, as well as the position of the absorption onset, are only well described by the BSE@G₀W₀ method. However, because the refractive index also depends on the real part of the dielectric function the most accurate BSE@G₀W₀ method severely underestimates the static refractive index. This happens because it's impossible in practice to converge the BSE calculation with respect to the number of excited states and there is therefore a significant amount of missing spectral weight. As an example of why this can be important, in Publication [II] it was necessary to perform an accurate computational evaluation of Boron-Phosphide (BP) as a material candidate for all-dielectric nano-photonics. The key material parameters determining a dielectric nano-structures ability to host geometric Mie resonances is its refractive index and extinction coefficient[99, 100]. The underestimation of the static index therefore means that the refractive index calculated using the BSE method cannot be directly used to evaluate the potential of a material for such applications. An alternative approach is to neglect the excitonic effects and use the RPA@G₀W₀ method. While this method provides a decent description of the refractive index well below the bandgap, it overestimates the onset of absorption and misrepresents the spectral features near the gap. This method is therefore also not suitable to accurate evaluation of dielectric materials for nano-optics. A method that combines the RPA@G₀W₀ method's ability to include the full spectral weight with the BSE@G₀W₀ method's ability to include excitonic effects is therefore needed to arrive at a quantitatively correct description. An effective approach to do this was developed for Publication [II] based on the f-sum rule.

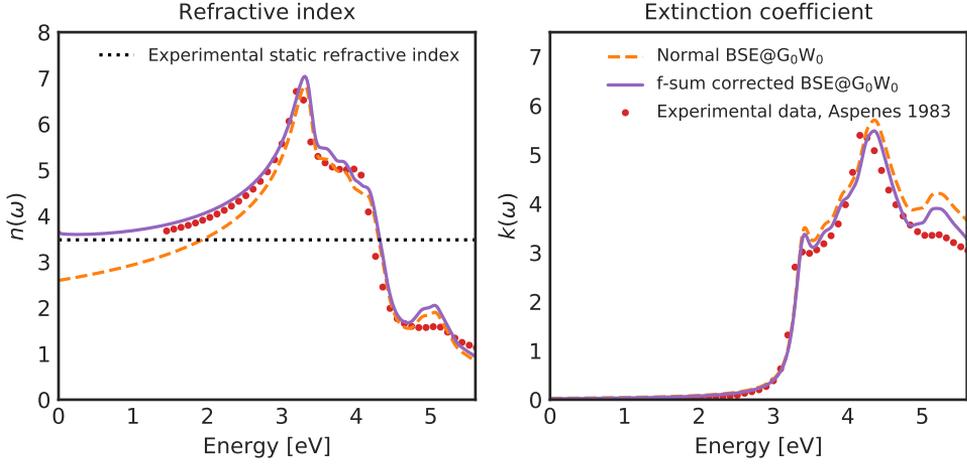


Figure 5.6: The refractive index and extinction coefficient for bulk Silicon calculated using the BSE@G₀W₀ and f-sum corrected BSE@G₀W₀ methods compared to the experimental values from Ref. [98].

As discussed in Chapter 3, the f-sum rule is the statement in Eq. 3.38 that for a finite system the sum of the oscillator strengths has to equal the number of electrons in the system. For extended systems, the f-sum rule can be generalized as an expression for the polarizability,

$$\int_0^{\infty} d\omega \omega \text{Im}\chi(\omega) = \frac{\pi n_e e^2}{2m}, \quad (5.43)$$

where n_e is the electron density, e is the elementary charge and m is the electron mass. Eq. (5.43) is a general statement and it therefore has to hold for polarizabilities independent of the method used to calculate it. Assuming that it was possible to converge both an RPA and a BSE calculation with respect to the number of excited states, the two would necessarily have to obey,

$$\int_0^{\infty} d\omega \omega \text{Im}\chi^{(\text{RPA})}(\omega) = \int_0^{\infty} d\omega \omega \text{Im}\chi^{(\text{BSE})}(\omega). \quad (5.44)$$

Eq. (5.44) can be enforced if the imaginary part of the BSE polarizability is artificially extended to ensure the proper spectral weight.

When performing the BSE calculation, all electronic bands within a certain energy range of the band extrema are included. Denoting this energy range as ΔE , one can be sure that all transitions with an energy smaller than $E_{\text{direct gap}} + \Delta E$ are included in the calculation. Note that the calculation can potentially include transitions with transition energies up to $E_{\text{direct gap}} + 2\Delta E$, but one can only be sure that *all* transitions

with energies below $E_{\text{direct gap}} + \Delta E$ are accounted for. Denoting the maximum frequency at which the BSE calculation includes all transitions, $\omega_c \equiv (E_{\text{direct gap}} + \Delta E)/\hbar$, the following extension is employed,

$$\text{Im}\chi^{(\text{BSE})}(\omega) = \begin{cases} \text{Im}\chi^{(\text{BSE})}(\omega), & \text{if } \omega \leq \omega_c, \\ C_0 e^{-\gamma\omega}, & \text{if } \omega > \omega_c. \end{cases} \quad (5.45)$$

The two parameters C_0 and γ can be chosen to ensure continuity and that Eq. (5.44) is fulfilled. With the correction, the calculations now has the full spectral weight in the imaginary part of the dielectric function and the real part can then be calculated via the Hilbert transformation. As can be seen in Fig. 5.6 for silicon, the f-sum correction fixes the underestimation of the static refractive index and also improves the description of the extinction coefficient with both showing excellent agreement with the experimental data.

CHAPTER 6

Tensor network methods for quantum light-matter interactions with open quantum systems

Within the methods discussed so far there was no distinction between the system of interest and the surroundings. The discussion has thus so far focused on *closed* quantum systems. However, any microscopic or mesoscopic system will inevitably be *open*, meaning that it will interact with its surrounding environment which results in the exchange of energy and information. One example is a color center in a material that one wants to use for e.g. quantum optical applications. While the color center is the interesting part, the presence of both lattice phonons and nuclear spins can drastically affect the system's optical- and coherence properties[101, 102]. In such situations, it makes sense to treat the system as an *open quantum system* (OQS), meaning a system that evolves in the presence of an environment[36]. Fig. 6.1 shows an illustration of a generic OQS.

This chapter discusses the Time Evolving Matrix Product Operator (TEMPO) algorithm[103]. TEMPO is a tensor network based method that can efficiently deal with OQSs coupled to Gaussian environments. Unlike the common master equations approach, the approach presented in this chapter makes no assumptions regarding memory effects in the system-environment interaction.

6.1 Open Quantum Systems

To start the discussion, this section reviews the concept of an OQS. Generically, the Hamiltonian of an OQS can be expressed as,

$$\hat{H} = \hat{H}_S + \hat{H}_E + \hat{H}_{\text{int}} \quad (6.1)$$

where \hat{H}_S is the Hamiltonian describing the system, \hat{H}_E is the Hamiltonian describing the environment, and \hat{H}_{int} is the Hamiltonian describing the coupling between the two. For the combined system, the time-evolution can in principle be described by a wave

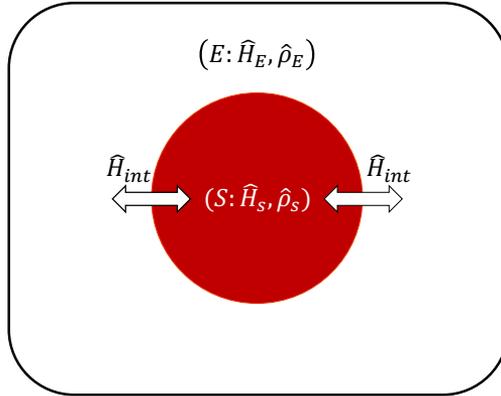


Figure 6.1: Illustration of an Open Quantum System consisting of the system of interest, an environment, and a system-environment boundary across which energy and information can flow.

function. However, because the system degrees of freedom are the only thing that can be probed, it is necessary to turn to a density-matrix description. Specifically, the time dependent state of the system is described by the density matrix that is obtained by tracing over the environment in the full density matrix. Denoting the time dependent state of the combined system and environment as $\chi(t)$, the state of the system at time t is given as,

$$\rho_S(t) = \text{Tr}_E [\chi(t)]. \quad (6.2)$$

The question then becomes how to calculate the time dependent state of the system in practice. Directly using Eq. 6.2 is generally impractical because the environment can potentially be so large that explicitly treating it is unfeasible; and even if this was possible it would likely be inefficient. The goal should therefore instead be to find a method that can propagate the state of the system in the presence of the environment without explicitly propagating the state of the environment.

6.1.1 Time evolution of an open quantum system

The time evolution of the full density matrix describing the system and the environment follows the standard Heisenberg equation of motion with respect to the full Hamiltonian,

$$\partial_t \chi(t) = -i [H, \chi(t)] \equiv \mathcal{L}[\chi(t)]. \quad (6.3)$$

The above defines the Liouvillian superoperator, \mathcal{L} and for simplicity \hbar is set to 1. Eq. 6.3 can be formally solved by defining the time evolution superoperator $\mathcal{U}(t) = \exp(\mathcal{L}t)$,

$$\chi(t) = \mathcal{U}(t)[\chi_0], \quad (6.4)$$

where χ_0 is the initial state of the combined system and environment. As mentioned above what should really be sought is the time dependent state of the system,

$$\rho_S(t) = \text{tr}_E \{ \mathcal{U}(t) [\chi_0] \}. \quad (6.5)$$

If the system and environment start out uncorrelated, i.e. $\chi_0 = \rho_0 \otimes \tau_0$, Eq. 6.5 can be understood as a quantum channel, $\mathcal{E}(t : 0)$, which propagates the initial state of the system to time t ,

$$\rho_S(t) = \text{tr}_E \{ \mathcal{U}_{t:0} [\chi_0] \} \equiv \mathcal{E}(t : 0) [\rho_0]. \quad (6.6)$$

$\mathcal{E}(t : 0)$ encodes the evolution of the systems under the influence of the interactions with the environment. It will generally represent non-unitary time evolution, reflecting the fact that the system is interacting with the environment. Generally speaking, the map will depend on the full history of the evolution and not just a single time, because the interactions between the system and the bath can have some memory. For example, imagine that the interactions result in some energy being transferred from the system to the environment at some time t . This energy can propagate around the environment and potentially return at a later time t' . Processes like this are called non-Markovian because they reflect that the full history, and not just the current time step, is important to describe general system-environment interactions, and they can only be described by propagating the memory of the system.

In some special cases, the interaction with the environment is completely memory-less, and the state of the system at some time t is fully determined by the state of the system at the previous time step. This is usually a good description when the interaction between the system and the environment is weak, and the environment is relatively unstructured. Furthermore, the environment should be large, such as to remain approximately unaltered by the interaction with the system. In such cases, the system environment interaction is Markovian and the description can be simplified significantly via the Born-Markov approximation[104]. Within this approximation, Eq. 6.6 can be recast as a Lindblad equation which, similar to the Heisenberg equation of motion, is a linear differential equation but with extra terms that lead to non-unitary time evolution of the system.

An example of a master equation is a system coupled to an empty environment under the assumption that the spectral densities of the system environment interactions are flat within the spectrum of interest,

$$\partial_t \rho_S(t) = -\frac{i}{\hbar} [H_S, \rho_S(t)] + \sum_i \Gamma_i \left[\hat{L}_i \rho_S(t) \hat{L}_i^\dagger - \frac{1}{2} \left\{ \hat{L}_i^\dagger \hat{L}_i, \rho_S(t) \right\} \right]. \quad (6.7)$$

Here the sum runs over the different system-environment interaction channels. Γ_i is the relaxation rate associated with interaction channel i , and \hat{L}_i is the so-called jump operator characterizing the interaction channel.

6.2 The TEMPO algorithm

While the master equation construction is very powerful and has been successfully applied to many systems, it necessitates the identification of a memory-less system-environment boundary. While it is sometimes possible to transform the system in such a way as to favourably redefine the system-bath boundary via a suitable unitary transformation[105, 106], this is not always the case. Furthermore, such effective theories are often only valid in limited parameter regimes[106] and should consequently be checked against more accurate methods. Consequently it makes sense to seek a method that can integrate Eq. 6.6 without making approximations as to the nature of the system-environment interactions.

The first step towards a solution is often a discretization of Eq. 6.6. In the discretized version, its solution can be conveniently represented by the so-called Augmented Density Tensor (ADT)[107, 108], $A(t_1, t_2, \dots, t_N)$. The ADT can be thought of a generalized version of the density matrix which depends on the present time and all previous times, thus encoding the memory of the system. It can be represented as a tensor of rank N , with each index representing a time step. The local size of each index is the same as the density matrix of the system, that is d^2 , and the ADT thus grows exponentially as d^{2N} with the number of time steps taken. While a lot of effort has been made to propagate the full ADT as far as possible, the exponential wall is very steep and one is therefore limited to around 20 steps[109]. To go beyond this, it is necessary to seek methods that efficiently compress the ADT. One method that achieves this is the Time Evolving Matrix Product Operator (TEMPO) algorithm[103]. In TEMPO, the ADT is represented as a matrix product state (MPS), and for system-environment interactions with finite autocorrelation time it can be efficiently compressed using standard MPS techniques.

6.2.1 Matrix product states

Before discussing TEMPO, it makes sense to briefly discuss the most important aspects of the MPS construction. MPSs are a clever way of truncating the size of a many-body wave function, motivated by the fact that ground states of local gapped Hamiltonians obey the so-called area law.[110] The area law states that if the ground state of a gapped, local Hamiltonian is cut into two halves, the entanglement entropy of the cut scales with the area of the cut and not with the volume, as would be the case for a random state in the Hilbert space. To understand how this relates to compression, consider a 1D chain of N sites with local Hilbert space dimension d , the state of which can be written as,

$$|\Psi\rangle = \sum_{j_1, \dots, j_N} C(j_1, \dots, j_N) |j_1, \dots, j_N\rangle. \quad (6.8)$$

Here j_n labels the state at site n , and $C(j_1, \dots, j_N)$ holds the coefficients of the wave function. Generally speaking, the tensor holding the coefficients $C(j_1, \dots, j_N)$ will contain d^N elements. $C(j_1, \dots, j_N)$ thus grows rapidly with system size, quickly becoming very unwieldy for large systems.

Now imagine that the chain is split into a left- and a right half. The state of the system can be represented by the Schmidt decomposition of the cut,

$$|\Psi\rangle = \sum_{\alpha} \Lambda_{\alpha} |\alpha\rangle_L \otimes |\alpha\rangle_R, \quad (6.9)$$

where Λ_{α} are the Schmidt coefficients and $|\alpha\rangle_{L,R}$ are the Schmidt states of the chain to the left and right of the cut respectively. The entanglement entropy of the cut can be cast in terms of the Schmidt coefficients,

$$S = - \sum_{\alpha} \Lambda_{\alpha}^2 \log \Lambda_{\alpha}^2. \quad (6.10)$$

It is clear that if a few Schmidt states dominate, S will be close to zero, and if there are a lot of different relevant Schmidt states, S will be large. Another way to state the same is that if the full state can be represented by only a few Schmidt states then S is small. In this case, the memory footprint of the wave function can be significantly reduced by throwing away small Schmidt states with only minimal loss of information. The ultimate limit of this is the case where the states of the left and right halves are separable, in which case only a single Schmidt state is necessary for each of the two halves. Contrastingly, if a lot of Schmidt states are needed then S will be large and the state cannot be significantly compressed without loss of information. The entanglement entropy is thus a direct measure of the compressibility of the state. Because of the area law, ground states of local gapped Hamiltonians will fall in the former category.

The insight from above directly motivates the MPS wave function ansatz, which seeks to decompose $C(j_1, \dots, j_N)$ into a product of two dimensional tensors or matrices each representing a site in the chain, which is where the “*Matrix Product*” part of the name comes from[111–113]. The MPS representation is non-unique, but the most intuitive, and in many ways also the most useful, representation is the so-called canonical representation where each bond in the chain is represented by its Schmidt decomposition,

$$|\Psi\rangle = \sum_{j_1, \dots, j_N} \Lambda^{[1]} \Gamma^{[1]} \Lambda^{[1]} \Gamma^{[1]} \dots \Lambda^{[N]} \Gamma^{[N]} \Lambda^{[N]} |j_1, \dots, j_N\rangle. \quad (6.11)$$

The full state of the system can then be compressed by sweeping left to right and back, throwing away all Schmidt values that are below the some desired cut-off[113]. One might ask why this representation is advantageous, since it still requires storing the initial version of Eq. 6.11 to compress it. However, in practice one does not start with the actual entangled state, but instead with some trial state which is evolved into the ground state via algorithms such as e.g. imaginary time Time Evolving Block Decimation[114] (TEBD) or Density Matrix Renormalization Group[115] (DMRG). In such cases, the canonical form of the MPS is constructed after each step of the algorithm and the truncation is subsequently performed. In that way, the Many Body state can be efficiently compressed, without ever having to store the full state.

Similar to the MPS, one can also define so-called matrix product operators (MPO), which are operators that act locally on the bonds of the MPS. An MPO acting on an

MPS with N indices will generally be a $2N$ index object. For the purposes of this thesis, it is only important to note that the MPO is the natural representation for operators acting on an MPS, and consequently algorithms for MPS are formulated in terms of successive applications of MPOs. For a more thorough discussion of MPS's and MPO's in particular, and tensor networks in general, the interested reader is referred to Refs. [111–113].

Interestingly, the MPS construction has also recently attracted attention for machine learning applications because the format provides a memory efficient representation of large tensors that simultaneously allows mathematical operations to be performed efficiently [116–118].

6.2.2 TEMPO

The TEMPO algorithm does not seek to find the ground state of a system, but instead to perform time propagation of an OQS while accounting for memory effects in the system-environment interaction. While this memory is crucial to account for, it is also important to realize that there will generally be some finite autocorrelation time characterizing the interaction. The correlations of the ADT will thus be limited to some finite characteristic time scale, similar to the entanglement entropy of a gapped ground state, meaning that it can be compressed analogously.

Restricting the discussion to cases where the system and environment are initially uncorrelated and the environment starts in thermal equilibrium, the state of the system at some time t can be written in terms of the time evolution operator,

$$\rho_S(t) = \text{tr}_E(\mathcal{U}(t)\rho_0 \otimes \tau_B). \quad (6.12)$$

Here ρ_0 is the initial state of the system, and τ_B is the thermal state of the environment. A discretization in time is performed, $t = k\delta t$, such that the time evolution operator is represented as $\mathcal{U}_{\delta t}^N$. As shown in the Supplementary Information of Publication [VI], the discretized propagator can be symmetrically Trotter decomposed, $\mathcal{U}_{\delta t} = \mathcal{V}_{\delta t}\mathcal{W}_{\delta t}\mathcal{V}_{\delta t} + \mathcal{O}(\delta t^3)$, by introducing the two new superoperators $\mathcal{V}_{\delta t}[\rho] = \exp(\mathcal{L}_0\delta t/2)[\rho]$, and $\mathcal{W}_{\delta t}[\rho] = U_{\delta t}\rho U_{\delta t}^\dagger$ describing the evolution of the system and environment degrees of freedom respectively. By successive application of the Trotter decomposed time evolution superoperator, the state of the system can be expressed as

$$\rho_S^{\alpha k} = \sum_{\vec{\alpha}, \vec{\beta}} \rho_S^{\alpha_0} \prod_{j=1}^k \mathcal{V}_{\beta_j}^{\alpha_j} \mathcal{V}_{\alpha_{j-1}}^{\beta_j} \mathcal{F}_{\beta_k \dots \beta_1}, \quad (6.13)$$

where $\mathcal{F}_{\beta_k \dots \beta_1} = \text{tr}_E(\mathcal{W}^{\beta_k} \dots \mathcal{W}^{\beta_1}[\tau_B])$ is the so-called influence functional describing the effect of the environment on the evolution of the system. For Gaussian environments, the trace in the expression for the influence functional can be done analytically [107, 108] and written in product form as,

$$\mathcal{F}_{\beta_N \dots \beta_1} = \prod_{j=1}^N \prod_{k=0}^{j-1} [b_k]_{\beta_j \beta_{j-k}}, \quad (6.14)$$

where the influence tensors are given as,

$$[b_{(i-j)}]_{\beta_i \beta_j} = e^{-(\lambda_{s_i} - \lambda_{r_i})(\eta_{i-j} \lambda_{s_j} - \eta_{i-j}^* \lambda_{r_j})}. \quad (6.15)$$

Here λ_s are the diagonal elements of the coupling operator describing the system-environment interaction, with discretized memory kernel,

$$\eta_{i-j} = \begin{cases} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} dt' dt'' C(t' - t''), & \text{for } i \neq j, \\ \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} dt' dt'' C(t' - t''), & \text{for } i = j. \end{cases}, \quad (6.16)$$

and environment autocorrelation function,

$$C(t) = \int_0^\infty J(\omega) \left[\coth\left(\frac{\beta\omega}{2}\right) \cos(\omega t) + i \sin(\omega t) \right], \quad (6.17)$$

where $J(\omega)$ is the spectral function of the system-environment interaction.

From Eqs. 6.13 and 6.14 it can be deduced that calculating the state of the system at some time step k requires the construction of the $(2k + 1)$ -index object:

$$\mathcal{R}_{\beta_k \cdots \beta_1}^{\alpha_k \cdots \alpha_0} = \rho^{\alpha_0} \prod_{j=1}^k \mathcal{V}_{\beta_j}^{\alpha_j} \prod_{l=1}^j [b_{(j-l)}]_{\beta_j \beta_l} \mathcal{V}_{\alpha_{j-1}}^{\beta_j}, \quad (6.18)$$

which is exactly the ADT. For arbitrarily small timesteps δt , it is an exact representation of the system-environment interaction for the OQS of interest. At this point, the size of the ADT still scales exponentially with number of timesteps taken and it is crucial to find a way to compress it effectively.

The ADT can be constructed recursively. To achieve this, the internal product in Eq. 6.18,

$$B^{\beta_k \cdots \beta_1} = \prod_{l=1}^j [\tilde{b}_{(j-l)}]_{\beta_j \beta_l} \quad (6.19)$$

is expanded with $k - 1$ dummy indices to define the propagation tensor,

$$\mathcal{B}_{\mu_k \cdots \mu_1}^{\beta_k \cdots \beta_1} = \prod_{j=1}^{k-1} \delta_{\mu_{k-j}}^{\beta_{k-j}} B^{\beta_k \cdots \beta_1}. \quad (6.20)$$

The first propagation step is,

$$\mathcal{A}^{\beta_1} = \mathcal{B}^{\beta_1} \tilde{\rho}^{\beta_1}, \quad (6.21)$$

where $\rho^{\beta_1} = \mathcal{V}_{\alpha_0}^{\beta_1} \rho^{\alpha_0}$, is the state propagated by half a timestep. Note that to extract the state from the above ADT, the tensor must be propagated by a final half timestep, this is a consequence of using a symmetric Trotter splitting. The next few timesteps can be constructed iteratively, such that:

$$\begin{aligned} \mathcal{A}^{\beta_2 \beta_1} &= \mathcal{B}_{\mu_1}^{\beta_2 \beta_1} \mathcal{A}^{\mu_1}, \\ \mathcal{A}^{\beta_3 \beta_2 \beta_1} &= \mathcal{B}_{\mu_2 \mu_1}^{\beta_3 \beta_2 \beta_1} \mathcal{A}^{\mu_2 \mu_1}, \\ &\vdots \end{aligned} \quad (6.22)$$

It can be seen that the rank of the tensors grows with each application of the propagation tensor which suggests exponential growth of the ADT. However, crucially, Ref. [103] realized that the ADT has an MPS representation and the propagation tensor has an MPO representation in terms of the influence tensors. Similar to the case of normal MPS, this representation is not unique and different representations have different advantages[119]. However, here the original MPO decomposition of the propagation tensor is used,

$$\mathcal{B}_{\mu_{k-1}\dots\mu_1}^{\beta_k\dots\beta_1} = [b_0]_{\gamma_1}^{\beta_k} \left(\prod_{j=1}^{n-2} [b_j]_{\gamma_{j+1}, \mu_{k-j}}^{\gamma_j, \beta_{k-j}} \right) [b_{k-1}]_{\mu_1}^{\gamma_{k-1}, \beta_1}. \quad (6.23)$$

The time propagation of the system is thus performed by successive applications of the MPO defined by Eq. 6.23. After each application of the MPO, the state is compressed. To compress at a given site k , a Singular Value Decomposition(SVD) of the local tensor is performed,

$$A_{\beta_{k-1}, \beta_{k-2} \dots \beta_1}^{\beta_n, \beta_{n-1} \dots \beta_k} = U_{\gamma}^{\beta_n, \beta_{n-1} \dots \beta_k} \Lambda_{\gamma}^{\beta_{k-1}, \beta_{k-2} \dots \beta_1} V_{\beta_{k-1}, \beta_{k-2} \dots \beta_1}^{\gamma, \dagger} \quad (6.24)$$

where the diagonal matrix Λ_{γ}^{β} contains the singular values and U and V are the left and right matrices from SVD respectively. Singular values less than a threshold value λ_C are discarded, reducing the internal bond dimension of the MPS, providing a controllable way to reduce the size of the tensors. The SVD is applied successively to each site, swept from left to right, and then back from right to left in each round of compression. The principal convergence parameters in the TEMPO formalism are thus the time step, dt , and the SVD cut-off λ_C . The whole process of evolving the ADT is shown visually in Fig. 6.2.

6.3 Linear response within the TEMPO algorithm

Publication [VI] uses the TEMPO formalism to investigate how the properties of color centers in hexagonal boron nitride (hBN) can be engineered using c-QED. Specifically, the lattice phonons were considered the environment, and the cavity mode and electronic states were considered the system. As part of this work, a method for calculating the linear absorption spectrum of such a system within TEMPO was devised.

As shown by Mukamel, a system's linear heterodyne absorption spectrum can be calculated as[36],

$$A(\omega) = 2\text{Re} \left[\int_0^{\infty} dt e^{i\omega t} \text{Tr}(\mu(t)\mu(0)\chi(-\infty)) \right]. \quad (6.25)$$

Here $\chi(-\infty)$ denotes the initial state of the OQS, which is taken to be its equilibrium state, and μ is the system transition operator describing the type of driving. To calculate the different linear response spectra, it is therefore necessary to calculate $\text{Tr}(\mu(t)\mu(0)\chi(-\infty))$ as a function of time.

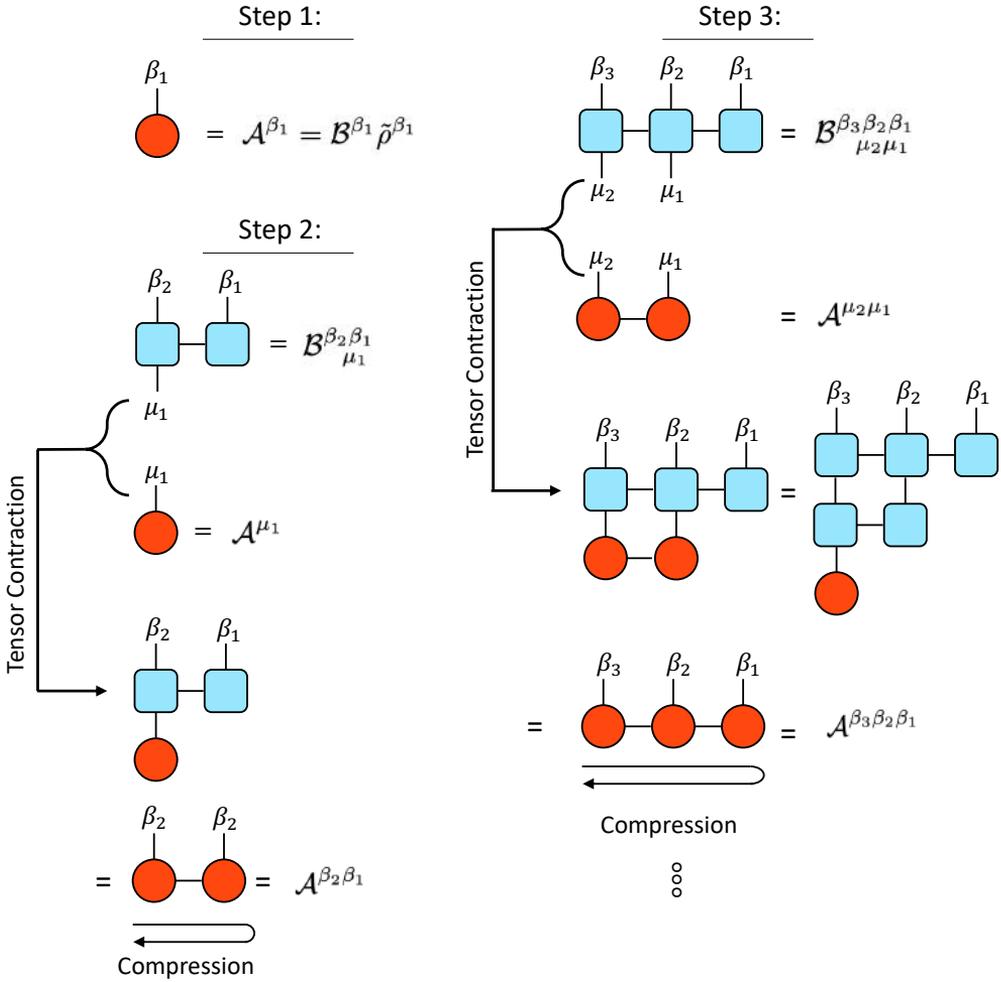


Figure 6.2: Illustration of the growth of the augmented density tensor within the TEMPO algorithm. The figure illustrates the three growth steps in Eqs. 6.21 and 6.22, including the intermediate compression steps from Eq. 6.24.

As already mentioned in section 3.2, there are two different ways of probing a QED system: Either via an external current probing the cavity mode, or via an external field probing the electronic system. The transition operator can therefore be either the dipole operator of the electronic system, or the quadrature operator of the cavity field. In either case, the operator acts solely on the state of the system. The linear absorption spectra can therefore be calculated by propagating the system from the unphysical initial state $\varrho(0) = \mu(0) |g, 0\rangle \langle g, 0|$ under the influence of the environment, and tracing with the relevant transition operator μ at time t . Here $|g, 0\rangle$ denotes

the state of system where the electron system is in its ground state and there are no photons in the cavity mode. Formally, one would have to propagate the system to $t = \infty$. However, in practice the important thing to ensure is that the system is propagated to equilibrium.

CHAPTER 7

Summary of the results related to quantum light-matter interactions

This first of the three result chapters will summarize the papers related to quantum light-matter interactions. It should be mentioned that the work presented in chapter 4 also falls within this same category. For each paper, a motivation of the study is given followed by a short summary highlighting the most important points and conclusions. Copies of all the papers are provided at the end of the thesis and the interested reader is encouraged to refer to them for additional details.

7.1 Publication [I]: Combining density functional theory with macroscopic QED for quantum light-matter interactions in 2D materials

Publication [I] explores a novel combination of MQED and DFT as a step towards a fully first-principles description of quantum light-matter interactions in arbitrary electromagnetic environments.

Van der Waals (vdW) heterostructures are a class of materials composed of stacked 2D materials[9]. Because the layers in the stack are only bounded together by the vdW-interaction, the possible combinations are not limited by lattice matching constraints like for conventional material interfaces. The freedom to stack different 2D materials in almost arbitrary order makes the vdW heterostructure a very versatile material platform. Additionally, the atomically tight confinement of the electrons in the out-of-plane direction results in unique physical properties. Such structures can for example host excitations with very confined electromagnetic fields[10, 11], strong exciton resonances[12], and thickness tunable infrared emission originating from intersubband transitions with large out of plane dipole moments[13]. This makes vdW heterostructures an interesting platform to study strong quantum light-matter interactions.

The theoretical treatment of quantum light-matter interactions in vdW heterostructures is challenging for several reasons. Within the dipole approximation for the electron-photon coupling, the coupling strength is the product of the transition dipole

moment and the local electric field. However, the ultra confined nature of the electromagnetic field associated with excitations in such structures means that it can become necessary to go beyond the dipole approximation. When this happens, the detailed shape of the electronic wave function also becomes important[120]. Furthermore, any quantum description of the electromagnetic fields in such structures must also simultaneously account for losses, as well as the very nonlocal nature of the excitations in 2D materials. Finally, because of the potential that the ultra confined fields lead to very strong light-matter interactions, it is also paramount to account for any potential effects that go beyond first order perturbation theory.

Motivated by the potential that vdW heterostructures hold as a platform to explore strong light-matter interactions, Publication [I] set out to address the theoretical challenges outlined above. The paper introduces a Wigner-Weisskopf model based method to describe quantum light-matter interactions in complex van der Waals heterostructures. The presented method should be valid under quite general conditions up until the USC regime. The methodology combines the quantization of the electromagnetic field in arbitrary optical environments from MQED, with the single particle Kohn-Sham wave functions and energies from DFT for the description of the electronic system. The combination of the two enables quantitative description of the light-matter interactions in the regime of highly confined, lossy light modes.

The methodology is used to determine the radiative transition rates of a multi-

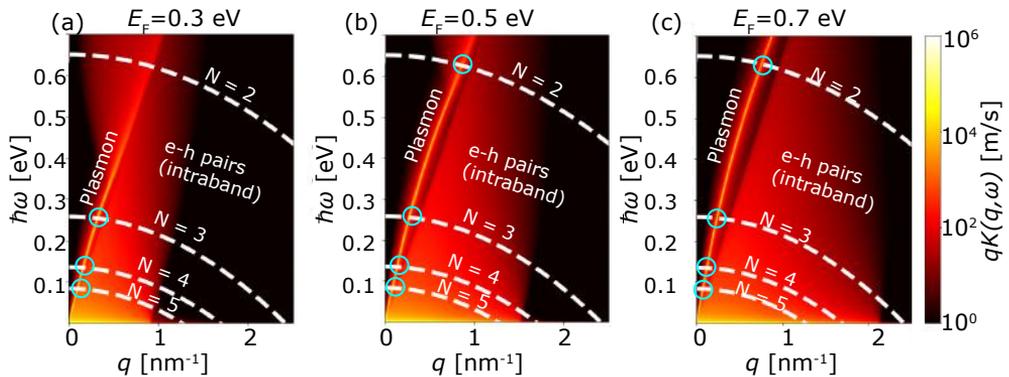


Figure 7.1: Interaction kernel: The coupling kernel for the transition between of holes between the lower and upper subband in a two-layer WSe₂ quantum well placed between a perfect conductor and a doped graphene sheet is shown for three different values of the graphene Fermi level, (a) 0.3 eV, (b) 0.5 eV and (c) 0.7 eV. The kernel describes the coupling into all optical excitations, including the plasmon and the electron-hole continuum as described by graphene's full nonlocal conductivity[121]. The dashed lines represent the non-vertical emission energies derived from the DFT band structure for WSe₂ quantum wells of $N = 1, 2, 3, 4$ and 5 layers and the light blue circles denote the intersection between the dispersion of the intersubband transition energy and the graphene plasmon. Figure reproduced from Publication [I].

layer TMD Quantum Well (QW) sandwiched between graphene and a metal surface. Such setups host strongly confined acoustic plasmon modes, which should enable very strong light-matter interactions[122–124]. Furthermore, the strong dependence of the graphene plasmon on the graphene Fermi levels means that the setup is highly tunable[121].

The interaction between the electronic system and the electromagnetic field is represented in the Weyl gauge ($\phi = 0$),

$$\hat{H}_{\text{int}} = \frac{e}{m} \hat{\mathbf{A}} \cdot \hat{\mathbf{p}} - \frac{i e \hbar}{2m} \nabla \cdot \hat{\mathbf{A}}. \quad (7.1)$$

Note that this expression neglects the diamagnetic term, the consequences of which are discussed in the paper. The Weyl gauge is an incomplete gauge and should therefore be accompanied by an additional condition to completely fix the gauge[125]. In the setup considered here, this condition is provided because $\nabla \cdot (\epsilon \cdot \mathbf{A}) = 0$ inside of the anisotropic TMD. Note that this is the case because the modes of the MQED field expansion should be defined in the absence of the intersubband transition. In an isotropic dielectric material, this choice of gauge reduces to the normal Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$) discussed in section 1.2. In such cases, the $\nabla \cdot \hat{\mathbf{A}}$ term disappears from Eq. 7.1, leaving only the $\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}$ term. The $\nabla \cdot \hat{\mathbf{A}}$ term can thus be thought of as an anisotropic correction and we have generally found it to be small relative to the $\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}$ term.

From the emitter side, the hole subbands of the TMD QW are considered, and the specific case where the electron starts in the upper subband and then decays to the lower one is chosen. As shown in the publication, the time dependence of the excited state coefficient can be expressed as,

$$\dot{C}_i(t) = - \int_0^t dt' \int d\omega \int d\mathbf{q} K(\omega, \mathbf{q}) e^{-i(\omega - \omega_{if\mathbf{q}})(t-t')} C_i(t'), \quad (7.2)$$

where $\dot{C}_i(t)$ is the excited state coefficient. The integral kernel $K(\omega, \mathbf{q})$ holds all the information about the electromagnetic environment. This includes both the acoustic plasmon mode, and the Landau continuum of the graphene, as shown in Fig. 7.1 for different values of the graphene Fermi level. The actual coupling to the intersubband transition is determined not only by the kernel, but also by the dispersion of the transition energy $\omega_{if\mathbf{q}}$. Because the intersubband splitting and the dispersion of the transition energy are dependent on the number of layers and the material, this means that different TMD stacks will couple differently. This is illustrated by the white lines in Fig. 7.1 indicating the dispersions of the transition energy for 2-5 layers of WSe₂. We explore this by calculating the radiative transition rates for the TMD quantum wells, varying both the thickness (2-5 layers), the TMD material (WSe₂, MoSe₂, WS₂ and MoS₂), and the graphene Fermi level (0.1-0.7 eV). Purcell enhancements up to 10⁷ and radiative rates approach 30 THz are found. The lifetimes corresponding to these rates are significantly lower than the electron-phonon limited lifetimes of electrons and holes in TMDs[126], and it is thus concluded that the coupling to the graphene plasmons should be dominant for the intersubband transitions at low temperatures.

Furthermore, the DFT-MQED method is used to analyze the validity of several common approximations such as the dipole approximation and the use of 1D QW wave function models. It is found that the dipole approximation can be violated, and that even Fermi's golden rule can be violated making it essential to use the full method to get accurate results. Most significantly, it is found that the rates depend strongly on the detailed shape of the wave functions, with the rates calculated using different approximations for the wave functions showing both significant qualitative variations as well as varying quantitatively by up to an order of magnitude. These results emphasize the need for a non-perturbative description, that can account for both the electron wave functions as well as the spatial profile of the electromagnetic field.

In summary, this paper presents a general method to describe light-matter interactions in the weak and SC regime for van der Waals heterostructures hosting very confined electromagnetic modes. Using this framework, the coupled acoustic graphene plasmon TMD QW setup is identified as an interesting platform for strong light-matter interactions, with Purcell factors up to 10^7 and radiative rates of up to 30 THz that dwarf the phonon limited lifetime. Furthermore, it is shown how most of the common approximations including the dipole approximation, approximate electronic wave functions and first order perturbation theory can be violated in vdW heterostructures emphasizing the need for a general method to treat quantum light-matter interactions in these setups.

7.2 Publication [VI]: Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials

Publication [VI] explores non-Markovian effects in the time evolution of color centers in 2D materials coupled to an optical cavity.

Single photon emissions from color centers in materials is a widely used platform for quantum technologies, the most notable example being the NV-center in diamond[101]. In recent years there has been a growing interest in the use of color centers in 2D materials as a source of single photons[127]. Everything else being equal, the defects in 2D materials have the distinct advantage of being intrinsically near the surface of the material and therefore lend themselves easily to for example near field manipulation[128]. In particular, hexagonal boron nitride (hBN) has attracted considerable attention owing to the material's wide band gap and the relative stability of single photon emission from its color centers[129, 130].

The emission from color centers is always accompanied by optical sidebands. These are called phonon sidebands, and they are a result of coupling between the color center and the host lattice[131]. The phonon sidebands result in significant broadening of the emission spectrum and have a negative impact on the coherence properties of the emitted photons. Specifically, the indistinguishability of the emitted photons, which is the key property for linear optical quantum computing[102], is reduced by the existence of the sidebands[47]. Strategies aimed at mitigating this usually focus on enhancing the

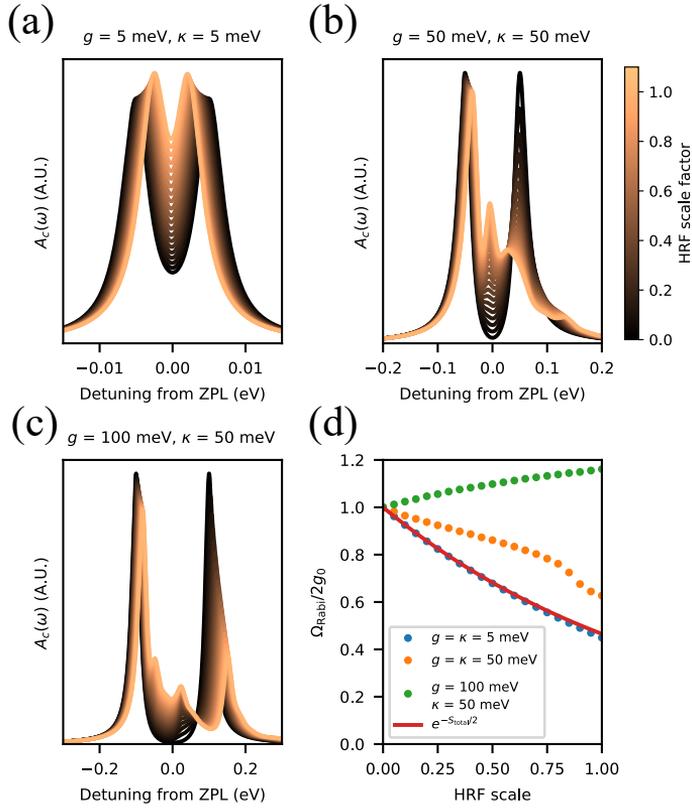


Figure 7.2: (a-c) the linear absorption spectra when probing the cavity mode, for various values of the scaling parameter α_{HRF} . Here structure in the absorption spectra becomes increasingly important for large coupling, and broad cavity widths. (d) shows the change in peak position as a function of α_{HRF} (points) for the cavity parameters studied in (a-c). Figure reproduced from Publication [VI].

emission from the zero phonon line (ZPL) via Purcell enhancement[44]. Intuitively, in the limit of infinite Purcell enhancement, the sideband emission would be completely suppressed leading to unit efficiency and ideal coherence properties of the emitted photons. However, previous studies have shown that this intuitive picture is oversimplified and that more detailed considerations are needed to properly describe the effect of the cavity on the emission[47].

Inspired by this, we set out to study how the optical properties of color centers in hBN could be engineered with c-QED. We specifically considered the $V_N C_B$ and $C_2 C_N$ defect complexes because they are the complexes that are thought responsible for the single photon emission in the visible[132, 133]. As discussed in the paper, by assuming that the inter-systems crossing time is relatively large, we restrict our treatment to the triplet manifolds which we model as two level systems. To get an

accurate description of the phonon coupling to the transition between the ground and excited state, we calculate the electron-phonon coupling from first-principles within the Duschinsky method[134]. This calculation gives us the partial Huang-Rhys factors from which the spectral density of the electron-phonon coupling can be constructed. Similar to previous studies, we employ the HSE06 exchange correlation functional for the DFT calculations which has been shown to agree well with experiments[131, 132].

The next question is how to describe the dynamics of the coupled system. If the phonons can be assumed to react adiabatically to changes in the photon field, the coupled emitter-cavity-host lattice system can be treated within the master equation formalism using polaron theory[47, 105]. However, as discussed in the paper, we quickly discovered that the polaron formation time in the hBN color centers, which can be thought of as the time it takes for the lattice to react to a transition from the ground to excited state or vice versa, is relatively long. This means that the inverse polaron formation time is comparable to even modest light matter coupling strengths, and this suggests that the interaction can be significantly impacted by non-Markovian effects. To properly describe both the coupling to the light and the coupling to the phonons, we therefore turned to the TEMPO algorithm described in section 6.2.

To study the non-markovian effects, we first study the population dynamics of the system when is initialized in the excited state and allowed to relax. We observe clear signatures of non-Markovian time evolution which we attribute to the fact that for even relatively weak cavity coupling strengths, the polaron formation time is on the same time scale as the cavity coupling strength. Via comparison with a single mode effective model of the phonon spectral density, we confirm that the origin of these non-Markovian effects is the fine structure of the electron-phonon coupling spectral densities. We also compared full TEMPO calculations with polaron theory using the proper spectral densities to confirm that the differences between the single mode and full spectral densities are true non-Markovian effects. The non-Markovian features are especially pronounced for $V_N C_B$ because it has a strong phonon peak at low detuning from the ZPL.

To further study the effect that the electron-phonon coupling has on the interaction of the cavity mode and the electronic transition, we have developed linear response within the TEMPO formalism (section 6.3). Using this, we calculated the linear absorption spectrum of the defects, while artificially scaling the electron-phonon coupling, $\mathcal{J}_{\text{Ph}}(\omega) \rightarrow \alpha_{\text{HRF}} \mathcal{J}_{\text{Ph}}(\omega)$, where $\alpha_{\text{HRF}} \in [0, 1]$ and $\mathcal{J}_{\text{Ph}}(\omega)$ is the spectral density of the electron-phonon coupling. The results are shown in Fig. 7.2. For small cavity coupling strengths and narrow cavity widths, we find a simple normalisation of the cavity coupling strength, which we attribute to a renormalization of the transition dipole moment by the electron-phonon coupling. This renormalization can be easily explained using simple Frank-Condon arguments. However, for stronger coupling and wider cavities, we observe far more complicated evolution of the spectra, clearly indicating actual hybridization of the phonons and the polaritonic states of the emitter-cavity system similar to Ref. [135].

In summary, we have shown that for color centers in hBN, the fine structure of the electron-phonon coupling spectral density plays a significant role in determining the dynamics of the cavity-defect system, even in the weak light-matter coupling regime, ne-

cessitating a non-Markovian description. The interplay between the light-matter interaction and the coupling to the host lattice was further investigated using the newly developed linear response TEMPO formalism. Specifically, the absorption spectra of the defects was computed at varying cavity coupling strengths and cavity widths. At weak cavity coupling strengths and narrow cavity widths, the coupling to the phonon environment leads to a simple renormalization of the cavity coupling strength in agreement with standard Frank Condon physics. However, at larger cavity coupling strengths the absorption spectrum show clear signatures of hybridisation between the light-matter polaritons and phonon modes.

In future work, we will seek to extend the approach to directly calculate coherence properties such as the indistinguishability of the emitted photons.

CHAPTER 8

Summary of the results related to Computational Materials Discovery

This second of the three result chapters will summarize the papers related to computational materials discovery. For each paper, a motivation of the study is given followed by a short summary highlighting the most important points and conclusions. Copies of all the papers are provided at the end of the thesis and the interested reader is encouraged to refer to them for additional details.

8.1 Publication [II]: Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators

Publication [II] employs a combination of high-throughput materials discovery, accurate many-body perturbation theory, and experimental fabrication and characterization to uncover new materials for all-dielectric nano-photonics.

Nano-photonics is the study of light confined to near or beyond the diffraction limit[43]. The diffraction limit is the statement that for *free-space* electromagnetic waves, the smallest feature that can be spatially resolved is $\sim \lambda_0/4\pi$. This is because the magnitude of the electromagnetic wave's wave number has to obey,

$$|k| = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi n(\lambda_0)}{\lambda_0}. \quad (8.1)$$

However, sub-diffraction limit confinement of light can be realized in the presence of spatial inhomogeneity owing to the fact that these can give rise to evanescent fields. For example, if k_z is imaginary then $k_z^2 < 0$ and one can have e.g. $|k_x| > |k|$, which means that at the cost of having the field decay out of plane, the light can be confined beyond the diffraction limit in the in-plane direction. This local confinement of the light results in an enhancement of the electromagnetic field strengths, and it can for example be used to boost the sensitivity of optical biosensors[136]. For this reason, the local field enhancement, measured as local field over the incident field, is a commonly

used figure of merit in nano-photonics for spectroscopic applications. As discussed in Ref. [100], it can be expressed as,

$$\left(\frac{E_{\text{local}}}{E_{\text{inc.}}}\right)^2 \propto \frac{Q^2}{V}, \quad (8.2)$$

where E_{local} is the local electric field, $E_{\text{inc.}}$ is the field incident on the nano structure, Q is the quality factor of the resonance in question and V is the mode volume of the resonance. Eq. 8.2 shows that strong local field enhancement can be realized with small mode volumes or large quality factors.

The surface plasmon polariton resonances of metallic nano-structures are a widely used platform in nano-optics due to the very confined nature of their associated electric fields[137]. However, while plasmonic resonances exhibit very large local field enhancement due to the strong confinement of the light, they are also inherently limited by the Landau dampening resulting from the scattering of the plasmons into the electron-hole pair continuum of the metals. This problem is in fact further enhanced in metallic nano-structures due to the surface enhanced Landau dampening[138, 139]. The resonances of plasmonic nano-structures are therefore often very broad which can be problematic for loss sensitive applications. Furthermore, plasmonic nano-structures often only host electric type resonances limiting their range of application. For these reasons, there has recently been a growing interest replacing the metals with dielectrics[99, 100]. While dielectric do not host plasmonic resonances, dielectric structures of sizes comparable to the wavelength of light can still host geometric Mie resonances[140]. The Mie resonances of dielectrics are far less confined than their plasmonic equivalents, but they do not suffer from Landau dampening and thus have significantly larger Q -factors. As per Eq. 8.2, these large Q factors can compensate for the weaker confinement of the mode and still lead to significant field enhancement. Furthermore, dielectric nano-structures are also able to host magnetic type resonances. This is both interesting because it can lead to local enhancement of the magnetic fields, as well as interesting directional scattering effects such as the Kerker effect via controlled interference with electric type resonances[141]. These reasons, along with the compatibility of many dielectrics with the CMOS fabrication techniques used in the semiconductor industry[100, 136], means that dielectric nano-photonics is an interesting path to pursue.

A dielectric's ability to host high-quality Mie resonances is directly linked to its refractive index. Generally speaking, a larger refractive index will lead to a larger quality factor of the resonances[100]. To be more precise, the important metric is the refractive index contrast between the nano-structure and its surroundings. This can be understood intuitively by noting that the critical angle of an interface also increases with the refractive index contrast[2], meaning that it can be thought of as a measure of the materials ability to trap the light. To maximize the quality factor of the resonances, it is critical that the absorption losses in the material are minimal. It is therefore critical to operate the material under the electronic bandgap to avoid the interband absorption. However, the larger the band gap of a material, the lower refractive index of the material usually is. This inverse relation between the refractive index and the electronic band gap is embodied in the qualitatively correct Moss formula[142],

$$n^4 E_G = C. \quad (8.3)$$

This means that while it is relatively easy to find materials that are suitable for applications in the IR and red part of the visible spectrum, it becomes a lot harder when going towards the blue end of the visible spectrum and into the near-UV. This is also reflected in the fact that most of the materials used in the near UV usually have modest refractive indices between 2 and 2.3[143–145]. It is therefore critical to expand the space of candidate materials as much as possible.

In Publication [II], we used high-throughput computational screening to search for candidate dielectrics among the 2743 thermodynamically stable unary and binary materials from the Open Quantum Materials Database (OQMD) project[146]. Stable here is defined as the materials on the OQMD convex hull. To perform this screening, a custom workflow based on the Atomic Simulations Recipe[147] (ASR) and MyQueue[148] frameworks was developed to determine the electronic properties of the materials. Due to computational constraints in the calculations of the optical response, we removed all materials with more than 11 atoms in the unit cell. To identify the semiconductors, a DFT ground state calculation using the PBE functional was performed for all materials. Removing the metals left us with 338 semiconductors as potential material candidates. Next, we calculated the frequency dependent refractive index and extinction coefficient within the RPA. Based on the fractional anisotropy of the refractive index ($FA < 0.05$), we sorted the materials into isotropic (207 materials) and anisotropic (131 materials) materials of which we focused on the 207 isotropic materials. Finally, we employed a Mie theory based evaluation to screen the scattering properties of the materials. The screening process and the results of it are summarized in Fig. 8.1. Our screening identified several different materials that have been overlooked in the context of nanophotonics, most notably boron-phosphide (BP). Based on the screening, BP shows a static refractive index comparable to the commonly used materials in the visible and near UV, while showing a larger band gap.

To more accurately characterize the optoelectronic properties of BP, we used the G_0W_0 approximation to correct the band gap from PBE. We confirmed that the G_0W_0 calculation provides a good description of the electronic structure by comparing the direct and indirect band gaps with experimental values[149, 150]. We subsequently applied the f-sum corrected BSE method (see section 5.3.3) to accurately calculate the frequency dependent refractive index including excitonic effects. As shown in Fig. 8.2(a-b), we observed that BP retains its favourable comparison with the commonly used materials at this high level of accuracy. We also used Mie theory simulations to confirm that spherical particles of BP can host clear Mie resonances across the range of energies below the direct band gap, see Fig. 8.2c. For these reasons, we decided to proceed with experimental fabrication and characterization of BP nano-structures.

BP nano particles were prepared by grinding commercially available BP powder in a mortar. The particles were then dispersing in methanol and dropcasted onto a glass substrate to isolate individual nano particles. This process resulted in quite rough nano particle shapes but they nonetheless showed clear signs of Mie resonances, which we prove by performing far-field characterization of the nano particles using dark field microscopy, see Fig. 3 in Publication [II]. We further checked the crystallinity of the nano particles using micro-Raman spectroscopy.

We then proceed with near-field characterization of the nano-particles using s-EELS

(see section 5.3.1). The combination of high spectral- and spatial resolution in s-EELS, as well as the its ability to probe resonances in the UV, allowed us to both extend our measurements into the UV, and to determine the nature of the resonances. In combination with simulations on cylindrical nano particles, we are able to spatially resolve the modes and determine the near field profiles of the different resonances. Using this process, we identify several characteristic modes such as the in- and out of plane electric dipoles, as well as the electric quadropole, despite the rough shape of the particles. This further confirmed potential of BP as a candidate material for dielectric nano-photonics.

As a final step, we investigated the possibility of reshaping the BP nano particles into more regular shapes, as this would be necessary to properly harness its potential as a candidate material for nano-photonics. We specifically investigated a laser reshaping procedure and show that it is indeed possible to process the BP nano-particles into spheres. We confirm that both the chemical composition and crystallinity is unaltered after reshaping using micro-Raman and EDS. We also characterized the scattering properties of the reshaped particles in the far field and show that it agrees very well with simulations for spheres of BP using the refractive index calculated using the f-sum corrected BSE method.

In summary, in this publication we used high-throughput computations to screen the stable binary materials from the OQMD database for interesting materials for dielectric nano optics. In the process, we developed both the computational workflow to perform the screening, and the f-sum corrected BSE method that, while not suited for high-throughput due to a large computational cost, can be used to accurately characterize interesting candidate materials. In particular, we identify BP and we show both theoretically and experimentally that the material is able to host clear resonances all the way into the UV part of the spectrum. Finally, we showed that it is possible to process the nano particles into regular shapes using a laser reshaping method. Along with other recent progress in BP fabrication[151], Publication [II] therefore opens the door to the wide spread adoption of BP for nano-photonics in the visible and near UV spectral ranges.

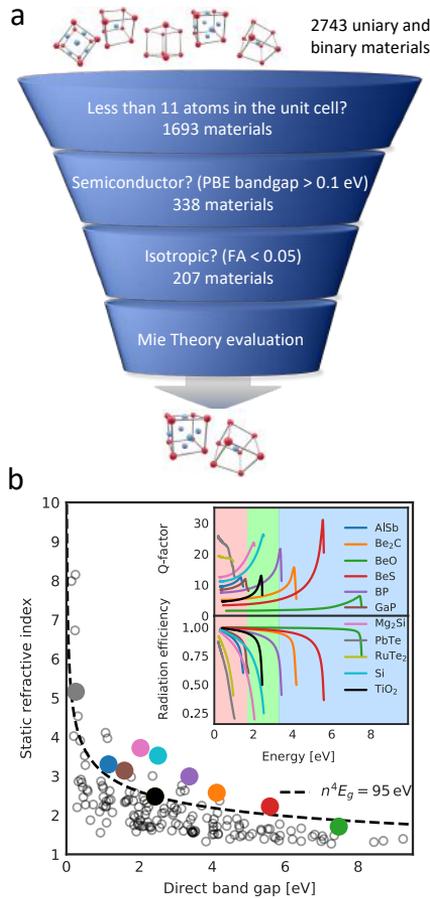


Figure 8.1: High-throughput materials screening: Schematic of the screening steps and the number of materials that survive each of them. b) The static refractive index of the isotropic materials plotted as a function of their direct band gap energies along with the Moss formula. The colored dots highlight interesting materials and the inset shows the energy-dependent Q -factor and radiation efficiency, η , of the magnetic dipole resonances in those materials. The background colors represent the infrared (red), visible (green), and ultraviolet (blue) spectral regions. Figure reproduced from Publication [II].

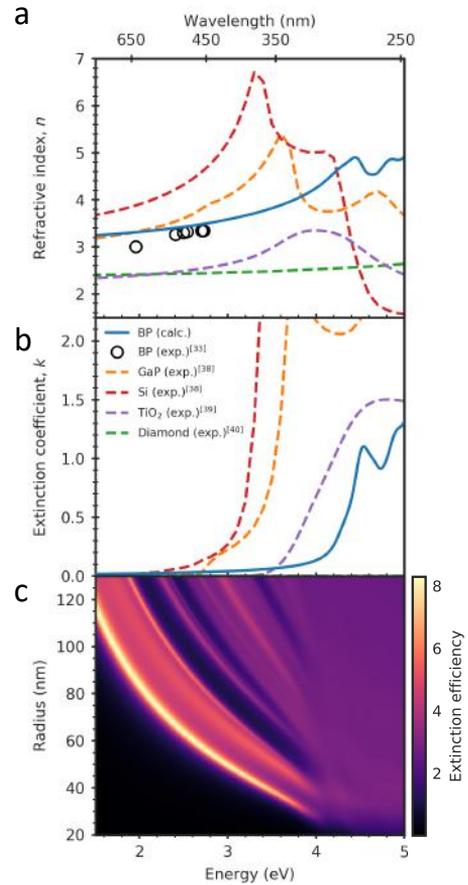


Figure 8.2: Optical response and Mie resonances of BP: a,b) Refractive index n and extinction coefficient k of BP calculated using the f -sum corrected BSE method compared with that of some of the commonly used dielectrics in the visible and ultraviolet spectral regions. c) Extinction efficiency map of BP spheres of varying radii calculated using Mie theory, which demonstrates that Mie resonances can be sustained in the visible and ultraviolet. Figure reproduced from Publication [II]

8.2 Publication [V]: Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption

Publication [V] investigates the potential of indirect bandgap materials in photovoltaic applications via high-throughput computational screening.

Photovoltaics (PV) is likely to play a central role in the transition towards sustainable energy production. While the material of choice has historically been silicon, recent years have seen growing interest in thin-film solar cells due to their superior mechanical properties, as well as their inherently lower material use. Despite of the interest, most materials currently used for thin-film cells are plagued by low abundance, toxicity concerns, and/or bad long term stability in the atmosphere. As such, the search for new materials for thin film PV cells has been the focus of many high-throughput computational screening studies[152–158]. However, due to the large computational cost of evaluating the phonon-assisted indirect absorption inherent to indirect band gap materials, these studies have primarily been focused on direct band gap materials. Because the indirect band gap materials form the largest part of semiconductor chemical space, this means that a significant part of the potential materials have been largely neglected. In this study we seek to remedy this by developing a computationally efficient approximation to evaluate indirect absorption and use it to screen the 338 semiconductors from Publication [II].

Phonon assisted absorption is a second order process involving a photon-phonon pair which collectively satisfy crystal momentum conservation. Because the phonon can carry significant crystal momentum, this allows non-vertical absorption processes to occur. For materials with an indirect electronic band gap, this results in absorption below the direct band gap[81]. The phonon assisted contribution to the absorption can be written in terms of the electron-phonon coupling matrix elements via second order perturbation theory,

$$\alpha(\omega) = \frac{4\pi e^2}{m^2 \epsilon_0 c} \frac{1}{n(\omega)\omega} \frac{1}{N_k^2} \sum_{\mathbf{k}_1 \mathbf{k}_2, c, v, l} |t_{c\mathbf{k}_2 \leftarrow v\mathbf{k}_1}^{l, \pm}|^2 \delta(E_{c\mathbf{k}_2} - E_{v\mathbf{k}_1} - \hbar\omega \pm \hbar\omega_{\mathbf{q}}^l). \quad (8.4)$$

Here $\hbar\omega$ is the energy of the incident photon and $n(\omega)$ is the refractive index of the semiconductor. To evaluate the indirect scattering matrix elements, $t_{c\mathbf{k}_2 \leftarrow v\mathbf{k}_1}^{l, \pm}$, the electron-phonon coupling, $V^l(\mathbf{q})$, has to be evaluated at $\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1$,

$$t_{c\mathbf{k}_2 \leftarrow v\mathbf{k}_1}^{l, \pm} = \sqrt{n_{\mathbf{q}}^l} \left[\sum_{\alpha} \frac{\langle \psi_{c\mathbf{k}_2} | V^l(\mathbf{q}) | \psi_{\alpha\mathbf{k}_1} \rangle \langle \psi_{\alpha\mathbf{k}_1} | \hat{\mathbf{e}} \cdot \mathbf{p} | \psi_{v\mathbf{k}_1} \rangle}{E_{v\mathbf{k}_1} - E_{\alpha\mathbf{k}_1} + \hbar\omega} + \sum_{\beta} \frac{\langle \psi_{c\mathbf{k}_2} | \hat{\mathbf{e}} \cdot \mathbf{p} | \psi_{\beta\mathbf{k}_2} \rangle \langle \psi_{\beta\mathbf{k}_2} | V^l(\mathbf{q}) | \psi_{v\mathbf{k}_1} \rangle}{E_{v\mathbf{k}_1} - E_{\beta\mathbf{k}_2} \pm \hbar\omega_{\mathbf{q}}^l} \right]. \quad (8.5)$$

Because the sum in Eq. 8.4 runs over all conduction- and valence bands, all momentum transfers, and all phonon modes, l , evaluating it requires a full phonon analysis for each

material which would be prohibitively expensive in a high-throughput setting. Instead, Publication [V] employs the approximation that the electron-phonon coupling matrix element can be approximated by the Γ -point phonons,

$$\omega_{\mathbf{q}}^l \rightarrow \omega_{\mathbf{0}}^l \quad (8.6)$$

$$\langle \psi_{c\mathbf{k}_2} | V^l(\mathbf{k}_2 - \mathbf{k}_1) | \psi_{v\mathbf{k}_1} \rangle \rightarrow \langle u_{c\mathbf{k}_2} | V^l(\mathbf{0}) | u_{v\mathbf{k}_1} \rangle. \quad (8.7)$$

This way, only a single phonon calculation in the primitive unit cell is needed for each material. Using crystalline silicon as a test, we confirm that the approximation provides a decent description of the indirect absorption. It is important to note that the approximation performs better for materials with larger unit cells. Furthermore, it can in principle be systematically improved by using larger supercells, becoming exact in the limit of an infinite supercell.

Combining this approximate method for the indirect absorption with the normal interband absorption, which we calculated from the frequency dependent RPA level dielectric functions from Publication [II], we are in a position to evaluate the PV performance of the stable binaries from the OQMD convex hull. We specifically consider the subset of the 338 semiconductors from Publication [II] which has a band gap between 0.7 and 3 eV when evaluated with the GLLBSC functional. In addition to the absorption, the materials are evaluated using screening descriptors for carrier transport, and non-radiative recombination. Generally speaking, the carrier mobility is characterized by the relaxation time due to scattering, τ , and the effective mass, m^* . While accurately calculating the relaxation time is a difficult task due to the multitude of different process that contribute to it, the effective mass of the carriers can be easily determined from the electronic band structure. The effective mass is therefore used as the measure of mobility in this study. To account for non-radiative recombination, we use the model proposed by Blank which treats the non-radiative recombination in terms of an effective internal quantum efficiency, Q_i [159].

After manually screening out materials containing toxic and rare elements, as well as materials that are known to be unstable in air, the screening leaves us with 28 potential candidate materials. The list, which contains 20 indirect band gap semiconductors, consists of both well established (3), emerging (16), and previously unexplored (9) materials. Most of the new compounds are anion rich chalcogenides (TiS3, Ga2Te5) and phosphides (PdP2, CdP4, MgP4, BaP3) containing homoelemental bonds.

Accounting for the indirect absorption is also important because it allows us to distinguish between materials with the same fundamental band gap. Consequently, in addition to the new candidate materials, our work also helps to address a discussion in literature about the potential benefits of having a small indirect band gap. It is found that an indirect band gap can be beneficial for photovoltaics, as it can result in a larger power conversion efficiency in the radiative limit compared to a direct band gap material with the same band gap. This is however only true provided that their gap is below the Shockley-Queisser optimal gap of 1.1 eV. As a concrete example, we show that CdP₄ ($E_g = 0.89$ eV, $E_g^{\text{dir}} = 1.01$ eV) has a larger power conversion efficiency than InSe ($E_g = E_g^{\text{dir}} = 0.86$ eV).

My primary contribution to this publication lies in the calculation of the electronic part to the absorption, as well as the construction of the workflow that led to the discovery of the 338 semiconductors that served as seed for this study.

In summary, Publication [V] presents a computationally efficient approximation of the phonon-assisted indirect absorption in semi-conductors. This approximation is used to screen a range of indirect band gap semiconductors, addressing a class of materials that have previously been largely neglected in high-throughput screening of PV materials. We identify 9 previously unexplored materials that could be interesting for PV applications. Interestingly, most of the new compounds contain homoelemental bonds, which is unlike most materials commonly used for PV applications. Our work highlights the previously underexplored potential of indirect band gap materials for photovoltaic applications.

8.3 Publication [VII]: Recent progress on the computational 2D materials database

Publication [VII] presents new developments for the Computational 2D Materials Database (C2DB) project.

To properly harness the potential of 2D materials in general, and vdW heterostructures in particular, it is important to expand the space of known 2D materials as much as possible. To this end, the C2DB project seeks to discover novel 2D materials using high throughput computational screening. The project was originally published back in 2018, consisting of relaxed atomic structures, electronic structures, thermodynamic properties, optical properties, elastic properties and stability for around 1500 2D materials[160]. Since its original publication, the C2DB has grown to around 4000 structures, several features have been added to the database, old features have been updated and the data provenance has been improved. These developments were a result of a collective effort of all of CAMD, and it led to the publication of [VII].

Figure 8.3 shows the C2DB workflow that all of the materials in the database have to go through. The workflow can be thought of a split into two parts. The former part deals with structure optimization and filtering out materials that do not relax into a 2D structure, whereas the latter part is property workflow that calculates a range of properties for the materials that pass the first step. The first step in the first part is a structure relaxation followed by a dimensionality analysis to determine if the material has relaxed into a single 2D material. If the material has either disintegrated into 1D or 0D clusters, or has relaxed into multiple 2D clusters, it is discarded. This is illustrated in Fig. 2 of the manuscript. Subsequently, the material is classified according to symmetry and magnetic structure. A stability analysis is further performed classifying the thermodynamic-, elastic-, and dynamical stability of the material. After these classification steps, the stable materials are fed to the property workflow. It includes the properties from the original C2DB workflow[160], as well as new properties including exfoliation energies, Bader charges, Born effective charges, optical- and infrared polarizabilities, the piezoelectric tensor, Raman spectra, second harmonic generation

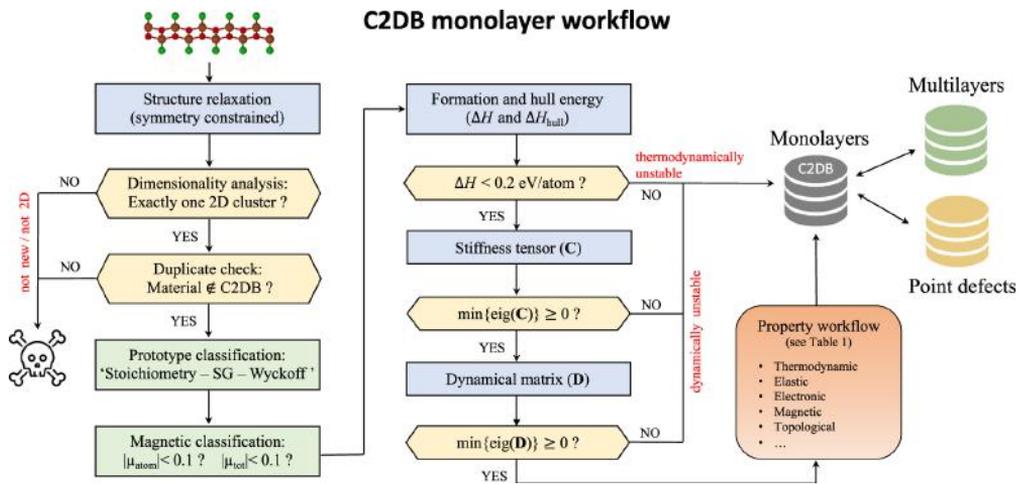


Figure 8.3: The C2DB workflow: A new material is first relaxed, and then classified according to its dimensionality. It also checked that the material does not already exist in the C2DB. If the materials passes these tests, its structure and magnetic properties are classified, before it's evaluated according to its thermodynamic-, elastic-, and dynamical stability. For the stable materials, the property workflow is then run before adding them to the C2DB. The unstable materials are also added to the database but without having the property workflow run for them.

and spontaneous polarization. The roughly 2500 new materials added to the database were identified by adding MXY Janus monolayers, as well exfoliable monolayers from experimentally known layered bulk crystals.

My personal contribution to the C2DB project lies primarily in the stability- and dimensionality analysis of the candidate structures. For the dimensionality analysis part of the workflow, a new recipe based on a dimensionality scoring parameter outlined by Ref. [161] was implemented in the ASR framework[147]. The work for the stability part of the workflow is based around the evaluation of the CBP protocol used to determine dynamic stability, as well as potential acceleration of the process using machine learning. These contributions to the C2DB project are further discussed in Publication [IV].

In summary, the C2DB project has undergone significant development in terms of the number of materials, quality of the workflow and new properties, meaning that it now contains a wide range of properties for around 4000 2D materials. The database provides a high quality data resource that can be used as a direct reference/starting point for new projects, as well as a valuable data resource for machine learning[162–164].

8.4 Publication [IV]: Predicting and machine learning structural instabilities in 2D materials

Publication [IV] is a continuation of the work on the C2DB project focusing on the dynamical stability evaluation of candidate materials (step 3 in the middle row of Fig. 8.3). Specifically, the paper assesses the Center and Boundary Phonon (CBP) protocol used by C2DB as an alternative to full phonon calculations for dynamical stability evaluation.

The relaxation of a candidate material for C2DB is indirectly constrained because the number of atoms in the unit cell is fixed. There is therefore a risk that the relaxation ends up in a saddle point on the potential energy surface. Any such material would, given the chance, make a distortion enlarging the unit cell, allowing it to relax into the lowest energy configuration. This is analogous to the case well known Peierls distortion[165], and the material would therefore be dynamically unstable. The dynamical stability of a structure can be evaluated via a phonon calculation. If any imaginary phonon frequencies are identified, it means that the material can gain energy by distorting along such a mode and the structure would therefore be dynamically unstable. In principle, imaginary phonon frequencies can occur anywhere in the phonon bandstructure and a full phonon calculation is necessary to evaluate the dynamical stability of a candidate material. Such a calculation is however very expensive, and this way of evaluating the dynamical stability is therefore not suited for high-throughput computations.

As an alternative to full phonon calculations, the C2DB project employs the CBP protocol. After the material has been relaxed, the CBP protocol evaluates the stiffness tensor of the material, and the Hessian matrix of a 2×2 supercell. More details on the CBP protocol are given in the paper but, importantly, the way the Hessian is calculated means that it only accounts for the phonons in the center and at high symmetry points along the boundary of the first Brillouin zone. This approach saves a significant amount of computational resources. However, the procedure will miss dynamical instability arising from imaginary phonon frequencies in the interior of the Brillouin zone, and the fidelity of this protocol for 2D materials remained an open question before this work. As compared to full phonon calculations, there are three possible outcomes of the CBP protocol: A true positive result in the cases where the unstable mode are at the center or the boundary of the cell, a true negative when the material is truly stable, and a false positive if the unstable mode is not at one of the high symmetry points evaluated by the CBP. The false positive result cannot exist because any material with an unstable boundary mode is de facto unstable. The protocol is tested on 20 randomly selected materials from the C2DB. The 20 materials represent 7 different prototypical crystal structures, with 10 metals and 10 semiconductors. For these 20 materials, full phonon calculations are performed and the classification is compared to the result of the CBP protocol. Of these 20 materials, 17 are correctly labeled and 3 are falsely labelled as stable by the CBP protocol. To improve the test, the CBP is also evaluated for the 258 monolayers for which full phonon calculations can be found in the literature[166]. Of these only 14 come out as false positives, resulting in a total success rate of 261

out of 278, which validates the CBP protocol as an efficient tool for the evaluation of dynamical stability. Note further that the majority of the false positives contain Nb, Co or Ta, indicating that these metals are associated with particularly complicated charge density wave reconstructions.

The effect of the structural distortions on the crystals can be investigated by pushing along the unstable phonon modes to generate dynamically stable structures. Such controlled dislocations were performed for 137 dynamically unstable materials from C2DB, which are chosen to have a single unstable mode, as predicted by CBP, and to be close (within 0.2 eV/atom) of the convex hull. The unstable structures can be split into two categories, unstable with isotropic Hessians and unstable with anisotropic Hessians. For the former, there is a clear way to displace the materials, and 43 out of 91 become stable after displacement. For the latter, it is not clear how to perform the displacement in general and the success rate is significantly lower at 6 out of 41, though it is shown that this can be improved by iterative application of the scheme. For the materials that become dynamically stable, it was found that the thermodynamic stability of the new, lower symmetry structures improves, with several of them moving near, or onto, the convex hull. As would be expected from the picture of Peirls distortion[81], there is a trend between band openings and distortions. For the new 49 stable materials the C2DB workflow is run and the materials were added to the C2DB. Iterative application of the CBP protocol could in principle be included in the C2DB stability workflow to find the dynamically stable configurations of the initial dynamically unstable materials, as shown in Fig. 8.4a.

Finally, a machine learning based approach to identify the unstable materials without a phonon calculation is evaluated. This is inspired by the fact that clear correlations between the dynamical stability and simple electronic properties of the material can be identified. In this work, the XGBoost model was employed[167]. To evaluate the model, the Receiver Operating Characteristic (ROC) curve[168] is used. The advantage of this metric that a clear overall measure of the performance is provided by the area under the curve (AUC), and it is further possible to evaluate its efficiency at different levels of willingness to sacrifice stable materials in the screening process. A simple fingerprint consisting of only simple electronic properties yields an respectable area under the roc curve (AUC) of 0.82. Improving the fingerprint using the RAD-PDOS fingerprint[163] significantly improves performance, raising the ROC from 0.82 to 0.9. The model performs slightly worse on the set on 137 materials selected for displacement, the reasons are understandable because these are chosen specifically to be close to the convex hull which is likely a metric that the model picks up on.

Due to the strong performance of the machine learning model, it is hypothesized that it can be used to speed up the screening for dynamically unstable materials. As shown in Fig. 8.4b, is envisioned that the machine learning model can be integrated into the workflow as a pre-screening step before the CBP protocol. In this way, it can be used to filter out the majority of the dynamically unstable materials without the need to evaluate even the center and boundary phonons. Such a combination would be able to further reduce the number of phonon calculations need to remove dynamically unstable materials in the workflow. For some projects, it will be advantageous to sacrifice a few dynamically stable materials if it means that more materials can be

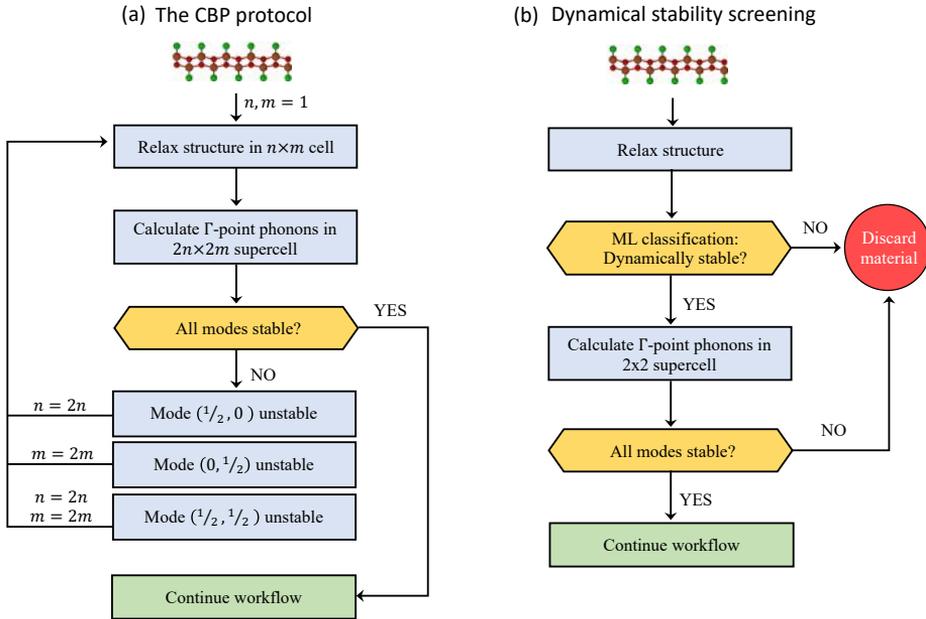


Figure 8.4: Two potential augmentations to the C2DB dynamical stability workflow: (a) After relaxation the material is evaluated using the CBP protocol. If it passes, the C2DB workflow is run for the material. If it fails, the material is pushed along its unstable mode and subsequently relaxed. The latter step is applied iteratively until a dynamically stable configuration is reached. (b) After the relaxation of a material its dynamical stability is first screened by the machine learning model. If it is labelled as unstable the material is discarded. If it passes, the material is then evaluated via the CBP protocol to confirm the machine learning prediction. The pre-screening step with the machine learning model thus serves as a filter that reduces the overall number of phonon calculations.

screened. Depending on the willingness to sacrifice stable materials, it is possible to save a significant amount of phonon-calculations by pre-screening with the ML model. The willingness to sacrifice materials is directly mapped out by the ROC curve, and the different thresholds can be read off the curve. Fig. 8.5 shows the classification thresholds where 5%, 10% and 20% of the stable materials are discarded are indicated. It is observed that $56 \pm 9\%$, $70 \pm 6\%$ and $85 \pm 3\%$ of the phonon calculations for the three thresholds, respectively.

In summary, the paper performs a systematic evaluation of the CBP protocol for the detection of dynamical instabilities in 2D materials. In total, the success rate of the CBP protocol is 261 out of 278, indicating high fidelity. A method for generating dynamically stable materials from unstable ones by pushing along the unstable phonon modes was explored, and several new materials were added to the C2DB as a result.

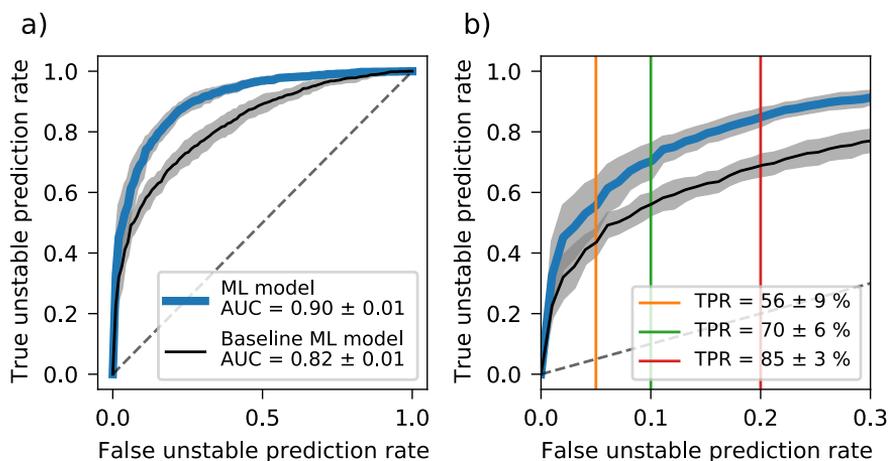


Figure 8.5: Machine learning results. a) shows the ROC curves for a machine learning model trained on RAD-PDOS fingerprints and a baseline model trained on electronic features from the C2DB. The AUC scores are 0.91 and 0.82, respectively. b) shows the ROC curves zoomed on low false prediction rates with different classification thresholds highlighted with vertical lines. By accepting 5% of the stable materials being falsely characterised as unstable, we can correctly label $56 \pm 9\%$ of the unstable materials. For 10% the true prediction rate is $70 \pm 6\%$ and for 20% it is $85 \pm 3\%$. Figure reproduced from Publication [IV].

Finally, the use of machine learning to speed up the evaluation of dynamical stability was investigated. It was found that the model performs well, achieving an excellent ROC-AUC score of 0.90 and lends itself to high-throughput assessment of dynamical stability.

CHAPTER 9

Summary of the results from experimental collaborations

This third of the three result chapters will summarize the papers related to experimental collaborations. For each paper, a motivation of the study is given followed by a short summary highlighting the most important points and conclusions. Copies of all the papers are provided at the end of the thesis and the interested reader is encouraged to refer to them for additional details.

9.1 Publication [III]: Momentum-dependent oscillator strength crossover of excitons and plasmons in two-dimensional PtSe₂

In Publication [III], we used *q*-EELS (see section 5.3.1) along with first principles calculations to probe the excitation spectrum of layered PtSe₂ as a function of thickness. This publication was a collaboration with experimental colleagues, and my contribution to the publication lies in the theoretical calculations and data analysis.

PtSe₂ is a Transition Metal Dichalcogenide (TMD) which can be isolated in monolayer or few layer forms[169], similar to other TMDs such as MoS₂ and WSe₂. However, unlike the more well-studied TMDs such as MoS₂ and WSe₂ which shows a direct to indirect bandgap transition when going from monolayer to bulk[170], the stronger interlayer coupling in layered PtSe₂ means that the material goes from a semi-conductor in the monolayer to a metal from three layers and onward[171–173]. The properties of PtSe₂ are therefore strongly layer dependent for few layers.

While the transport and catalytic properties of PtSe₂ have been extensively studied[174–176], its strong thickness dependent properties could also make it a very interesting material for optoelectronic applications[177]. However, until now its excitation spectrum remained understudied. In Publication [III] we therefore set out to perform a detailed investigation of the excitation spectrum using *q*-EELS.

We initially fabricated layered slabs of PtSe₂ using chemical vapour deposition. This resulted in samples of different thicknesses ranging from monolayer to bulk samples.

The quality of these samples was checked using ADF-STEM and it was found that high-quality regions of many different thicknesses could be identified. Next, q -EELS measurements were performed using the approach described in section 5.3.1. We specifically performed measurements on a monolayer, a 4 layer and a bulk sample, see Fig. 9.1(a-c). It was found that in all cases the excitation spectrum was dominated by three distinct peaks, a low energy A peak, an intermediate energy B peak, and a high-energy C peak. It was generally found that the A-peak dominated in the low q -region, whereas the B and C peak become dominant as q was increased.

To understand the dispersion of the different peaks, we turned to theoretical considerations regarding the q -dependence of electronic screening in 2D and 3D materials respectively. As discussed in section 5.2.5, at the BSE level the interaction between electrons in a solid is made up of two contributions: an exchange contribution, $V(\mathbf{q})$, and a contribution from the direct interaction, $W(\mathbf{q})$. In Publication [III] and its Supplementary Information we discuss how the electronic screening behaves very differently in 2D and 3D. In 3D, both contributions are relatively constant in the small q limit, with the exchange interaction usually dominating. This is in stark contrast to the 2D case where, for small q , the long-range components of $V(\mathbf{q})$ scales approximately linearly with q whereas the $W(\mathbf{q})$ remains relatively constant. In effect, the long-range contribution to $V(\mathbf{q})$ is thus turned off in the optical limit in 2D, leading to a spectrum dominated by the direct interaction. The result is that in 2D the low q excitation spectrum is dominated by excitonic effects, whereas at larger q the spectrum is dominated by the exchange interaction[90]. We employ this different scaling of the excitonic and plasmonic excitations to assign the A peak as an excitonic peak in the 2D systems, and the B and C peak as plasmonic peaks.

To support the theoretical reasoning, we calculate the EEL spectrum at different q -values with the RPA, see Fig. 9.1(d-f). We find that for all thicknesses, the RPA reproduces the B and C peaks relatively well, emphasizing their plasmonic origin. Because of the good agreement between the RPA and the experiments, we can employ the mode decomposition of the dielectric function described in section 5.1.2 to further study the excitations. Specifically, we can study the long-range exchange interaction "turn on" as we increase q to directly observe the strong q -dependence of the plasmon formation in 2D. With this method, we find that the B and C peak go from interband peaks to interband plasmons in 2D as we increase q , firstly proving their plasmonic nature, and secondly showing that plasmon formation is strongly q -dependent in 2D and that this can be used as a spectroscopic tool. The results in 2D stand in clear contrast to the 3D case where we observe basically no q -dependence of the plasmon formation in the low q -limit and observe strong plasmon formation generally. However, the fact that the RPA reproduces the B and C peak well again indicates their interband plasmonic nature in the bulk system. We finally discuss the orbital composition of the interband transitions contributing to the B and C peaks, and we find that for all thicknesses the B peak is formed primarily of transitions between Pt- d like orbitals to Se- p like orbitals, and the C-peak is made primarily of transitions between Pt- p to Pt- s orbitals.

Finally, we turned to a discussion of the A-peak which we reasoned was excitonic, at least in the 2D systems. Interestingly, the A-peak is actually relatively well represented

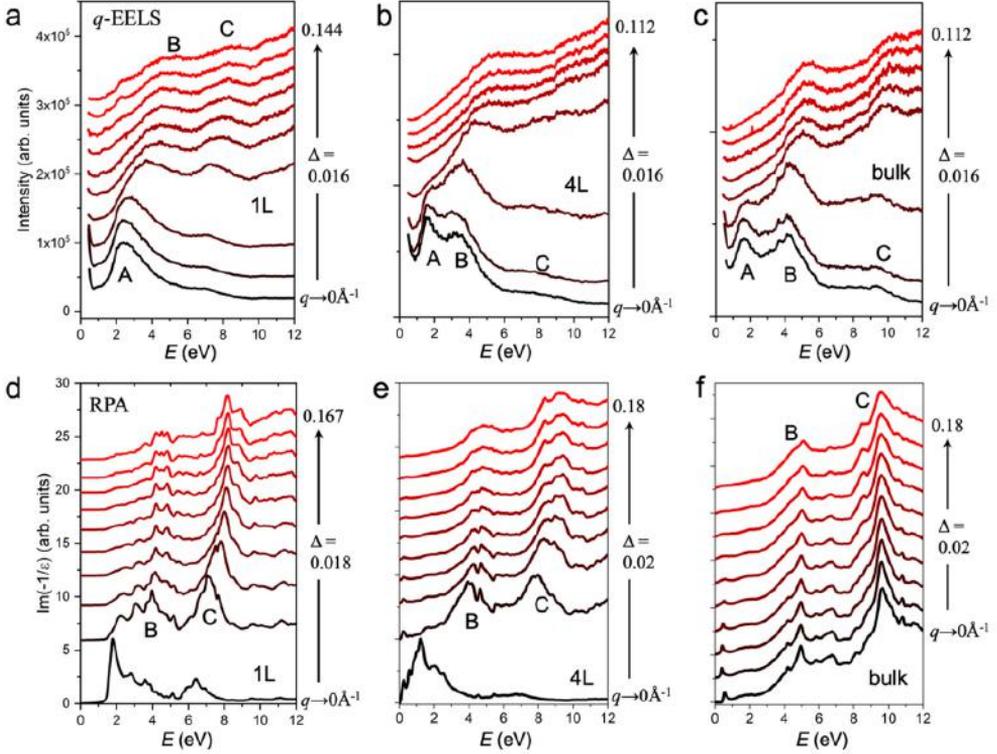


Figure 9.1: (a–c) Experimental q-EELS of 1L, 4L and bulk PtSe₂, respectively. (d–f) Calculated q-dependent loss functions in the random phase approximation (RPA) of 1L, 4L, and bulk PtSe₂, respectively. The low energy excitonic peak A, and the intermediate and high energy plasmonic peaks B and C are indicated. Figure reproduced from Publication [III].

by RPA in the 2D systems at $q = 0$, despite the lack of excitonic effects. As discussed in the paper, this turns out to happen due to a combination of different errors in the RPA@PBE calculations. For the monolayer, we can perform BSE calculations based on energies from the G_0W_0 approximation (as described in section 5.2.5) to confirm the excitonic nature of the A-peak, see Fig. 9.2a. It turns out that the opening of the bandgap in G_0W_0 is counteracted almost exactly by the exciton binding energy in the monolayer. Combined with the vanishing exchange interaction in the optical limit, this means that for the 2D systems RPA based on PBE will produce an interband peak almost exactly where the true excitonic peak lies. However, as can be seen in Fig. 9.2a, the reality is that there is significant renormalization of the A-peak due to excitonic effects and these are ultimately the true origin of the peak. In addition to calculating the excitonic correction to the EELS spectrum, we further map out the excitonic wave function in reciprocal space and show that while the electron is located in a minimum in k -space, the hole is localized in a relatively flat part of the Brillouin

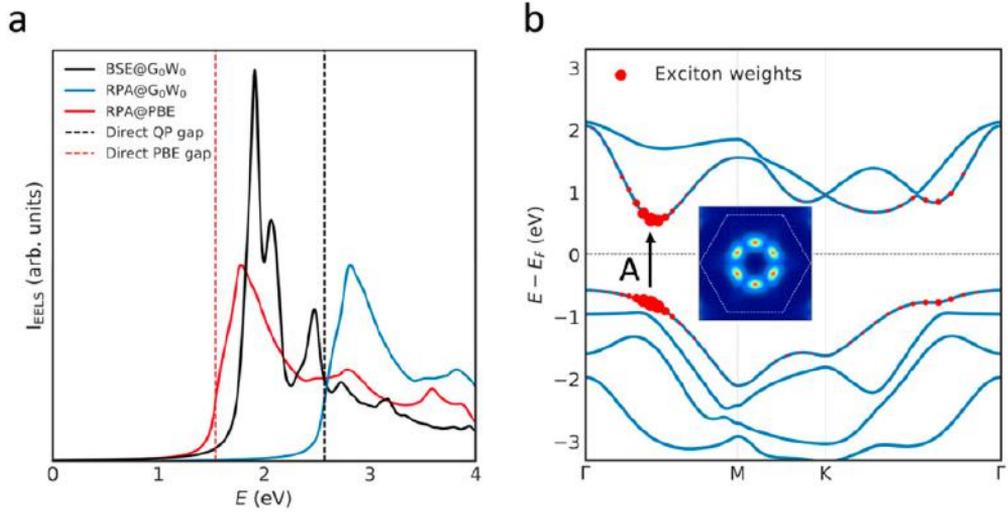


Figure 9.2: Excitonic origin of the A-peak in monolayers. (a) Comparison of BSE and RPA calculated loss functions of monolayer PtSe₂ in the $q \rightarrow 0$ limit. The BSE spectrum is calculated on top of the G₀W₀ band structure. For comparison, the RPA spectrum is calculated on top of the PBE band structure and the PBE band structure corrected to match the G₀W₀ band gap, respectively. (b) The BSE calculated excitonic weights of the A peak. The colored inset shows the 2D projected exciton wave function distribution in k space. Both the excitonic weights and k -space wave function show that the A exciton originates from direct transitions at the midpoint of $\Gamma - M$ (highlighted by the black arrow). Figure reproduced from Publication [III].

zone and consequently it is significantly more delocalized in k -space, see Fig. 9.2b. We believe that this explains why the excitonic peak in photoluminescence of PtSe₂ is weaker than expected. For the 4 layer system it is not possible to perform a BSE calculation due to the larger number of atoms in the unit cell. However, because the scaling argument for the different contributions to the Coulomb interaction also holds in the 4 layer case we believe that the A peak also represents an exciton in the 4 layer slab.

It is quite clear that the A-peak cannot represent an excitonic peak in the metallic bulk system due to the large screening in the system, and the BSE calculation for the bulk system also confirms this. Our hypothesis is that the A peak is instead an intraband plasmon in the bulk system. If one calculates the plasma frequency of the bulk system based solely on the density of states at the Fermi level, a plasma frequency of 1.6 eV is found which agrees well with the experimental position of the A peak. However, in the full RPA spectrum the intraband plasmon peak is renormalized to 0.56 eV due to the screening by interband transitions. It should, however, be noted that the PBE functional is known to underestimate interband transition energies, which generally leads to an overestimation of interband screening and thus a red-shifting of plasmon

energies[178]. For conventional metals, the description of the intraband plasmon can be improved using the GLLBSC functional[178]. We also tried this for the layered PtSe₂ system but we did not see any significant change in the bulk A-peak. We were therefore not able to confirm the nature of the A-peak and its origin consequently remains an open question.

In conclusion, this paper unravels the elementary electronic excitations of PtSe₂ using both q -resolved EELS and first principles calculations. In addition, we show how to turn theoretical considerations about the q -dependence of the different Coulomb interaction matrix elements into actionable spectroscopic insight for 2D materials, potentially opening the door to new possibilities in spectroscopy of 2D materials.

9.2 Publication [VIII]: Tunable free-electron X-ray radiation from van der Waals materials

In Publication [VIII], we investigated how parametric coherent Bremsstrahlung from van der Waals heterostructures can be used as a tunable source of monochromatic X-ray radiation.

Access to energy tunable X-ray radiation is important for core-level spectroscopy and different X-ray imaging techniques[179]. However, most laboratory scale sources can only provide emission with broad spectral width and limited tunability. Instead, energy tunable X-ray radiation is most often generated using large magnetic undulators at synchrotron- or free electron laser facilities.[180, 181]. A magnetic undulator consists of a series of magnetic dipoles with alternating directions. When a high-energy electron is passed through the undulator, the alternating magnetic field will cause the electron trajectory to oscillate back and forth resulting in emission of radiation. The wavelength of the emitted radiation depends on both the electron velocity and the undulator period, λ_u . Focusing on the first harmonic in the forward direction, the emitted radiation has wavelength[182],

$$\lambda_e(\theta = 0) = \frac{\lambda_u}{2\gamma^2} \left(1 + \frac{K^2}{2} \right). \quad (9.1)$$

Here $\gamma = \left(1 - \frac{v^2}{c^2} \right)^{-1/2}$ is the relativistic length contraction parameter, and K is the undulator strength parameter, $K = \frac{eB\lambda_u}{2\pi m_e c}$, where B is the magnetic field strength of the magnetic dipoles. K quantifies the magnitude of the oscillations the electron undergoes when traveling through the undulator. In the regime where $K \ll 1$ the perturbation to the electron trajectory is small. This means that the radiation from the same electron in different periods is able to interfere which causes the emission to have a narrow spectral width. Eq. 9.1 dictates that the wavelength of the emitted radiation in the undulator regime is $\lambda_e \approx \frac{\lambda_u}{2} \left(1 - \frac{v^2}{c^2} \right)$. Thus if either the undulator period is small enough, or the electron velocity is large enough, the emitted radiation will be X-ray radiation. Conventional undulators usually have periods on the millimeter or centimeter scales, and to achieve X-ray emission it is therefore necessary to accelerate the electrons to

very large energies. This is the reason why the operation of such sources is restricted to synchrotron- and free electron laser facilities, and it severely limits the practical use of this type of radiation. For this reason, there has recently been interest in developing more compact undulator designs. The most notable example of this is laser-driven Compton based X-ray sources that use the magnetic field of high-intensity lasers pulses as the undulating field. This results in significant shrinking of the undulator period to micrometer scales which means that lower electron energies are required[183].

In this work, we explore the ultimate limit of undulator wavelength using the parametric coherent bremsstrahlung generated by electrons traveling through layered vdW crystals. This allows us to reduce the electron acceleration voltage to levels found in a standard Transmission Electron Microscope (TEM) ($\sim 60\text{--}300$ keV). The idea is illustrated in Fig. 9.3a.

As a proof of concept, we place a stack of layered WSe₂ in the TEM oriented such that electrons travel perpendicular to the layers. The emitted X-ray is then measured as a function of the electron acceleration voltage, as shown in Fig. 9.3b. We find that the emitted radiation has a narrow spectral width, and its energy can be tuned by varying the incoming electron energy. This proves that the vdW-heterostructure can provide the necessary undulation of the electrons. We further confirm the effect for stacks of MnSP₃ and CrPS₄ respectively. Because tuning the energy of the electrons can require cumbersome realignment, we also explore how the X-ray can be tuned by changing the vdW material. Specifically, we consider the paradigmatic XPS₃ material with Fe, Co, Ni metal ions respectively. The interlayer distance differs slightly between these materials, 6.52 Å, 6.46 Å and 6.30 Å, and we show a systematic blue shift of the emitted X-ray radiation with decreasing interlayer distance.

As illustrated in Fig. 9.3b and c, there are two mechanisms responsible for the X-ray generation in the vdW heterostructures: Parametric X-ray Radiation (PXR) and Coherent Bremsstrahlung (CBS). The former can be understood as the radiation that results from the materials response to the incoming electron. Contrastingly, the latter can be understood as the radiation resulting from the perturbation of the electrons path as it interacts with the structure of the material. Both mechanisms result in energy tunable X-ray being emitted, and the emission from both mechanisms follow identical angular dispersion relations. PXR can be modelled by treating the periodic structure of the atomic lattice as a discrete array of emitting dipoles. CBS can be modelled via a two step process. The first step is solving the Newton-Lorentz equations for the electron trajectory in the presence of the electrostatic potential of the crystal lattice. We obtain this potential from a DFT calculation. The radiation can then be calculated from the electron trajectory using the Lienard-Wiechert potentials which describe the electromagnetic field of a moving point charge[184]. Our analysis shows that PXR dominates as the source of radiation in vdW materials for the electron energies available in the TEM. We also find that the model of the PXR can be used to accurately predict the spectral properties of the X-ray radiation observed in the experiments. Finally, we theoretically propose that designer undulators can be constructed by engineering the period of the vdW heterostructures. Specifically, we propose that these could be made by periodically stacking n_1 layers of one vdW material, followed by n_2 layers of another vdW material. Using the model of the dominant PXR radiation, we theoretically prove

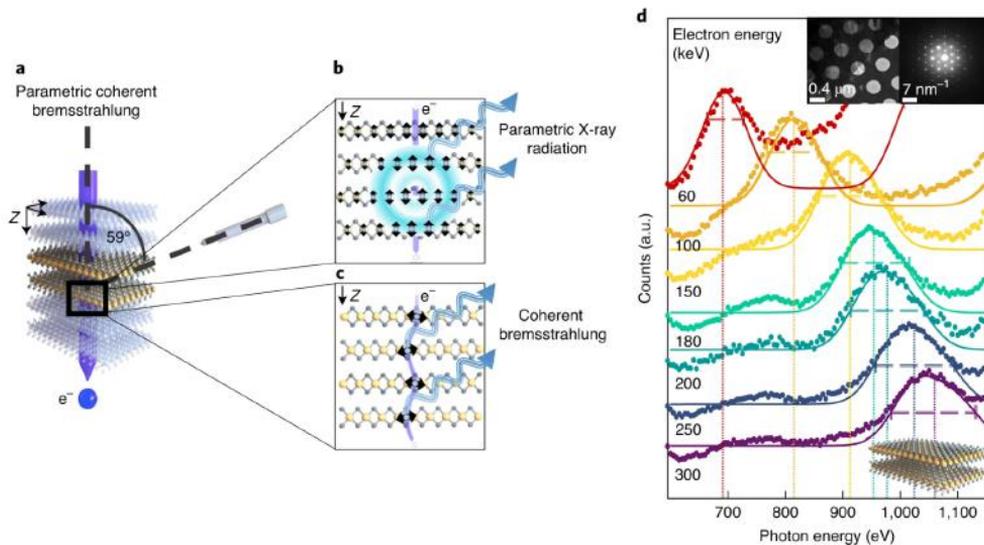


Figure 9.3: a) The radiation is produced from two combined mechanisms, PXR and CBS from an electron (e^-) propagating through periodic crystal structures. The angle between the direction of the detector and the negative direction of electron propagation is 59 degrees. b) Illustration of the PXR mechanism. c) Illustration of the CBS mechanism. d) Photon energy spectrum of X-ray radiation created by an electron moving along the [001] axis of WSe_2 . The emitted radiation peaks at different photon energies, depending on the electron kinetic energy (60–300 keV). Insets: sample image and diffraction pattern (top right). Figure reproduced from Publication [VIII].

that changing the period of the heterostructure leads to changes in the dispersion of the emitted X-ray as a function of electron energy.

Publication [VIII] was a large collective effort. My contributions were limited to assisting in the development of the CBS simulations, and carrying out the necessary DFT calculations to extract the electronic potentials.

In summary, vdW heterostructures are identified as a potential sources of energy tunable X-ray radiation when placed in a TEM. It is shown that the emission can be tuned both via the incoming electron energy and the configuration of the heterostructure. This opens the door to highly tunable sources of X-ray with very small physical footprints as compared to current technologies.

CHAPTER 10

Summary and Outlook

This thesis has studied light-matter interactions using first-principles quantum mechanical calculations. These calculations have been used to both interpret novel experimental results, and actively drive materials discovery for different opto-electronic applications. The thesis has both presented new methodologies, as well as novel applications of existing methodologies.

A major focus of the present work was quantum light-matter interactions. Macroscopic Quantum Electrodynamics (MQED) was introduced as a general approach to the quantization of the electromagnetic field in lossy magnetoelectric backgrounds. The thesis also reviews Density Functional Theory (DFT) as a practical approach to solving the interacting many-electron ground state problem, and Quantum Electrodynamical DFT (QEDFT) as a version of Time Dependent DFT (TDDFT) that can account for quantized modes of light. Furthermore, the linear response of a non-relativistic QED system was reviewed, and its formulation in terms of QEDFT was introduced.

Building on these existing ideas, a method combining MQED with DFT to provide an accurate description of light-matter interactions in van der Waals (vdW) heterostructures was presented. This is a challenging problem because such structures host ultra-confined electromagnetic modes. In particular, it was shown how most of the common approximations including the dipole approximation, approximate electronic wave functions and first order perturbation theory can be violated in vdW heterostructures. This emphasizes the need for a general method to treat quantum light-matter interactions in these setups. As a follow up to this work, an approach combining MQED with QEDFT was presented in Chapter 4. The work addresses a clear hole in the existing QEDFT literature, which has so far not connected the cavity coupling strengths to real cavity setups, and instead mostly treated these as free parameters. While the framework is limited to finite systems and coupling within the dipole approximation, it overcomes some of the other limitations of the MQED-DFT framework such as the neglect of the diamagnetic term and the approximate model of the electronic energy levels. The work therefore represents a step towards fully first principles description of quantum light-matter interactions. The framework is illustrated for a spherical micro-cavity with multiple benzene molecules embedded in it. It is found that the properties of the electromagnetic resonances can be linked directly to the geometry and material composition of the cavity. It was also shown that QEDFT's ability to treat many electrons allowed for treatment of multiple molecules coupled to the electromagnetic environment without having to resort to simplified models of the electronic structure. This work is a step towards parameter-free QEDFT calculations in general electromagnetic environments. In addition to the cavity-QED setting, the method could also provide a

way to perform TDDFT in a lossy optical environment, potentially removing the need for artificial spectral broadening in such calculations. Future work will be dedicated to combining ideas from both approaches into a truly general framework for quantum light-matter interactions capable of describing both finite and extended systems.

Furthermore, the impact of non-Markovian effects for cavity-QED with color centers in 2D materials was investigated. The color centers were described as open quantum systems interacting with their phonon environment. The study combined a first-principles description of the electron-phonon coupling with the tensor network based TEMPO algorithm. The latter is capable of describing the time evolution of an open quantum system under the influence of non-Markovian interactions with a general Gaussian environment. It was shown how the fine structure of the electron-phonon coupling's spectral density manifested as clear signatures of non-Markovianity in the dynamics of the coupled cavity-defect system. As a part of this project, a new methodology utilizing TEMPO to calculate the linear absorption spectrum of an open quantum system was introduced. Using this new approach, it was shown that at larger cavity coupling strengths the absorption spectrum shows clear signatures of hybridisation between the light-matter polaritons and the phonon modes of the environment. Future work will seek to extend the approach to directly calculate relevant coherence properties such as the indistinguishability of the emitted photons.

The thesis has also extensively discussed dielectric screening in materials. General aspects of dielectric screening in materials was discussed, and relevant methodology such as the G_0W_0 approximation, the Random Phase Approximation and the Bethe-Salpeter Equation (BSE) was introduced. Furthermore, the connection between the dielectric screening and spectroscopic quantities such as EELS, absorption and the refractive index was discussed. A new methodology to accurately calculate refractive indices of solids was introduced in the form of the f-sum corrected BSE.

The understanding of dielectric screening was used to guide the search for new materials. Using a combination of high-throughput computational screening and accurate many-body perturbation theory, Boron-Phosphide (BP) was computationally identified as an overlooked high-performance material for near-UV dielectric nano-photonics. This was subsequently confirmed via a combination of near- and far-field optical characterization on experimentally synthesized nano-particles. Apart from the identification of BP, this work resulted in a database containing optical and electronic properties for 338 semiconductors. This database was subsequently used to examine indirect band gap materials for photovoltaic applications. Because the calculation of the indirect absorption is computationally expensive, this group of semi-conductors have historically been largely neglected in high-throughput computational screening studies. To overcome this, a computationally efficient approximation for the indirect absorption was presented. The resulting screening led to the identification of several new interesting compounds. Most of these compounds are anion rich chalcogenides (TiS_3 , Ga_2Te_5) and phosphides (PdP_2 , CdP_4 , MgP_4 , BaP_4) containing homoelemental bonds. Future work will seek to build on these discoveries with new uses of the existing data, as well as extend the existing database. It is envisioned that this will result in a large database of high quality first principles calculations for the binary materials which can serve as a guide for future materials discovery efforts.

The thesis further presented new results for the C2DB project. In particular, a systematic evaluation of the Center and Boundary Protocol (CBP) as an approximate method for dynamical stability evaluation was performed. This showed that the method shows high fidelity, validating its use in the C2DB project. It was also shown how dynamically stable materials can be generated from unstable ones by iterative application of the CBP. Finally, a machine learning approach to harness the extensive amount of dynamical stability data already in the C2DB was presented. The model was evaluated using the ROC curve, showing an excellent AUC of 0.9. In the end, two modified versions of the C2DB stability workflow are presented. One uses the iterative application of the CBP to identify stable material candidates from dynamically unstable seed materials. The second employs the machine learning model as a pre-screening tool. It is discussed how the tradeoff between the speed up and the number of materials lost due to being falsely labelled as unstable by the model is directly mapped out by the ROC curve. This gives flexibility for the potential future deployment of the model in the C2DB workflow.

Furthermore, the elementary electronic excitations of layered PtSe_2 was unravelled using both q -resolved EELS measurements and first principles calculations. The material was shown to host both plasmonic- and excitonic excitations. Using the Bethe-Salpeter Equation as the starting point, it was discussed how the Coulomb interaction can be divided into a direct- and an exchange contribution. The former is responsible for exciton formation, and the latter is responsible for plasmon formation. As a result of the reduced screening, it was shown that the exchange contribution disappears in the optical limit for 2D materials. This stands in clear contrast to the 3D case, and results in the unique situation that plasmon formation is effectively turned off in 2D materials for $q \rightarrow 0$. It is envisioned that this unique feature of the electronic screening in 2D materials could lead to new spectroscopic possibilities. This was illustrated by tracking the q -evolution of the characteristic plasmon peaks in the EELS spectrum for the monolayer using the mode decomposition of the dielectric function. The analysis clearly showed that these excitations transform from interband peaks to plasmons as q was increased.

Finally, the thesis presents results from a large collaboration seeking to use vdW heterostructures as a source of energy tunable X-ray radiation under illumination of electrons from a Transition Electron Microscope. Specifically, vdW heterostructures are identified as potential sources of energy tunable X-ray, where the emission can be tuned both via the incoming electron energy and the configuration of the heterostructure. This opens the door to highly tunable sources of X-ray with very small physical footprints as compared to current technologies.

Bibliography

- [1] A. L. Schawlow and C. H. Townes, “Infrared and optical masers,” *Physical Review* **112**, 1940 (1958).
- [2] B. E. Saleh and M. C. Teich, *Fundamentals of photonics* (John Wiley & Sons, 2019).
- [3] M. E. Peskin, *An introduction to quantum field theory* (CRC Press, 2018).
- [4] A. Hirsch, “The era of carbon allotropes,” *Nature materials* **9**, 868 (2010).
- [5] J. Flick, N. Rivera, and P. Narang, “Strong light-matter coupling in quantum chemistry and quantum photonics,” *Nanophotonics* **7**, 1479 (2018).
- [6] H. Hübener, U. De Giovannini, C. Schäfer, J. Andberger, M. Ruggenthaler, J. Faist, and A. Rubio, “Engineering quantum materials with chiral optical cavities,” *Nature materials* **20**, 438 (2021).
- [7] K. S. Novoselov, A. K. Geim, S. V. Morozov, D.-e. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, “Electric field effect in atomically thin carbon films,” *Science* **306**, 666 (2004).
- [8] K. Novoselov, o. A. Mishchenko, o. A. Carvalho, and A. Castro Neto, “2D materials and van der Waals heterostructures,” *Science* **353**, aac9439 (2016).
- [9] A. K. Geim and I. V. Grigorieva, “Van der Waals heterostructures,” *Nature* **499**, 419 (2013).
- [10] F. H. Koppens, D. E. Chang, and F. J. García de Abajo, “Graphene plasmonics: a platform for strong light-matter interactions,” *Nano letters* **11**, 3370 (2011).
- [11] D. Basov, M. Fogler, and F. García de Abajo, “Polaritons in van der Waals materials,” *Science* **354**, aag1992 (2016).
- [12] G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B. Urbaszek, “Colloquium: excitons in atomically thin transition metal dichalcogenides,” *Reviews of Modern Physics* **90**, 021001 (2018).
- [13] P. Schmidt, F. Vialla, S. Latini, M. Massicotte, K.-J. Tielrooij, S. Mastel, G. Navickaite, M. Danovich, D. A. Ruiz-Tijerina, C. Yelgel, et al., “Nano-imaging of intersubband transitions in van der Waals quantum wells,” *Nature nanotechnology* **13**, 1035 (2018).
- [14] C. Gerry, P. Knight, and P. L. Knight, *Introductory quantum optics* (Cambridge University Press, 2005).

- [15] J. A. Hutchison, T. Schwartz, C. Genet, E. Devaux, and T. W. Ebbesen, “Modifying chemical landscapes by coupling to vacuum fields,” *Angewandte Chemie International Edition* **51**, 1592 (2012).
- [16] A. Thomas, J. George, A. Shalabney, M. Dryzhakov, S. J. Varma, J. Moran, T. Chervy, X. Zhong, E. Devaux, C. Genet, et al., “Ground-state chemical reactivity under vibrational coupling to the vacuum electromagnetic field,” *Angewandte Chemie* **128**, 11634 (2016).
- [17] J. Flick, M. Ruggenthaler, H. Appel, and A. Rubio, “Atoms and molecules in cavities, from weak to strong coupling in quantum-electrodynamics (qed) chemistry,” *Proceedings of the National Academy of Sciences* **114**, 3026 (2017).
- [18] M. Hertzog, M. Wang, J. Mony, and K. Börjesson, “Strong light–matter interactions: a new direction within chemistry,” *Chemical Society Reviews* **48**, 937 (2019).
- [19] L. A. Martínez-Martínez, M. Du, R. F. Ribeiro, S. Kéna-Cohen, and J. Yuen-Zhou, “Polariton-assisted singlet fission in acene aggregates,” *The Journal of Physical Chemistry Letters* **9**, 1951 (2018).
- [20] B. Liu, V. M. Menon, and M. Y. Sfeir, “The role of long-lived excitons in the dynamics of strongly coupled molecular polaritons,” *ACS Photonics* **7**, 2292 (2020).
- [21] C. Climent, D. Casanova, J. Feist, and F. J. Garcia-Vidal, “Not dark yet for strong light-matter coupling to accelerate singlet fission dynamics,” *Cell Reports Physical Science* **3**, 100841 (2022).
- [22] S. Latini, E. Ronca, U. De Giovannini, H. Hübener, and A. Rubio, “Cavity control of excitons in two-dimensional materials,” *Nano letters* **19**, 3473 (2019).
- [23] F. Appugliese, J. Enkner, G. L. Paravicini-Bagliani, M. Beck, C. Reichl, W. Wegscheider, G. Scalari, C. Ciuti, and J. Faist, “Breakdown of topological protection by cavity vacuum fields in the integer quantum hall effect,” *Science* **375**, 1030 (2022).
- [24] V. Rokaj, M. Penz, M. A. Sentef, M. Ruggenthaler, and A. Rubio, “Polaritonic hofstadter butterfly and cavity control of the quantized hall conductance,” *Physical Review B* **105**, 205424 (2022).
- [25] T. Lancaster and S. J. Blundell, *Quantum field theory for the gifted amateur* (OUP Oxford, 2014).
- [26] L. E. Ballentine, *Quantum mechanics: a modern development* (World Scientific Publishing Company, 2014).
- [27] S. Scheel and S. Y. Buhmann, “Macroscopic qed-concepts and applications,” arXiv preprint arXiv:0902.3586 (2009).
- [28] B. Huttner and S. Barnett, “Dispersion and loss in a hopfield dielectric,” *EPL (Europhysics Letters)* **18**, 487 (1992).
- [29] B. Huttner and S. M. Barnett, “Quantization of the electromagnetic field in dielectrics,” *Physical Review A* **46**, 4306 (1992).

- [30] S. Y. Buhmann, D. T. Butcher, and S. Scheel, “Macroscopic quantum electrodynamics in nonlocal and nonreciprocal media,” *New Journal of Physics* **14**, 083034 (2012).
- [31] R. Kubo, “Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems,” *Journal of the Physical Society of Japan* **12**, 570 (1957).
- [32] D. De Bernardis, P. Pilar, T. Jaako, S. De Liberato, and P. Rabl, “Breakdown of gauge invariance in ultrastrong-coupling cavity qed,” *Physical Review A* **98**, 053819 (2018).
- [33] A. Stokes and A. Nazir, “Gauge ambiguities imply jaynes-cummings physics remains valid in ultrastrong coupling qed,” *Nature communications* **10**, 499 (2019).
- [34] M. A. Taylor, A. Mandal, W. Zhou, and P. Huo, “Resolution of gauge ambiguities in molecular cavity quantum electrodynamics,” *Physical Review Letters* **125**, 123602 (2020).
- [35] E. A. Power and S. Zienau, “Coulomb gauge in non-relativistic quantum electrodynamics and the shape of spectral lines,” *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **251**, 427 (1959).
- [36] S. Mukamel, *Principles of nonlinear optical spectroscopy*, 6 (Oxford University Press on Demand, 1999).
- [37] R. G. Woolley, “Power-zienau-woolley representations of nonrelativistic qed for atoms and molecules,” *Physical Review Research* **2**, 013206 (2020).
- [38] I. V. Tokatly, “Time-dependent density functional theory for many-electron systems interacting with cavity photons,” *Physical review letters* **110**, 233001 (2013).
- [39] M. Ruggenthaler, J. Flick, C. Pellegrini, H. Appel, I. V. Tokatly, and A. Rubio, “Quantum-electrodynamical density-functional theory: bridging quantum optics and electronic-structure theory,” *Physical Review A* **90**, 012508 (2014).
- [40] J. Feist, A. I. Fernández-Domínguez, and F. J. García-Vidal, “Macroscopic qed for quantum nanophotonics: emitter-centered modes as a minimal basis for multiemitter problems,” *Nanophotonics* **10**, 477 (2021).
- [41] S. Y. Buhmann and D.-G. Welsch, “Casimir-polder forces on excited atoms in the strong atom-field coupling regime,” *Physical Review A* **77**, 012110 (2008).
- [42] T. Hümmer, F. García-Vidal, L. Martín-Moreno, and D. Zueco, “Weak and strong coupling regimes in plasmonic qed,” *Physical Review B* **87**, 115419 (2013).
- [43] L. Novotny and B. Hecht, *Principles of nano-optics* (Cambridge university press, 2012).
- [44] S. I. Bozhevolnyi, L. Martin-Moreno, and F. Garcia-Vidal, *Quantum plasmonics* (Springer, 2017).

- [45] E. M. Purcell, “Spontaneous emission probabilities at radio frequencies,” in *Confined electrons and photons* (Springer, 1995), pages 839–839.
- [46] A. Frisk Kockum, A. Miranowicz, S. De Liberato, S. Savasta, and F. Nori, “Ultrastrong coupling between light and matter,” *Nature Reviews Physics* **1**, 19 (2019).
- [47] J. Iles-Smith, D. P. McCutcheon, A. Nazir, and J. Mørk, “Phonon scattering inhibits simultaneous near-unity efficiency and indistinguishability in semiconductor single-photon sources,” *Nature Photonics* **11**, 521 (2017).
- [48] J. Keeling, “Light-matter interactions and quantum optics,” Graduate lectures (2014).
- [49] G. Zengin, G. Johansson, P. Johansson, T. J. Antosiewicz, M. Käll, and T. Shegai, “Approaching the strong coupling limit in single plasmonic nanorods interacting with j-aggregates,” *Scientific reports* **3**, 3074 (2013).
- [50] A. A. Anappara, S. De Liberato, A. Tredicucci, C. Ciuti, G. Biasiol, L. Sorba, and F. Beltram, “Signatures of the ultrastrong light-matter coupling regime,” *Physical Review B* **79**, 201303 (2009).
- [51] T. Niemczyk, F. Deppe, H. Huebl, E. Menzel, F. Hocke, M. Schwarz, J. Garcia-Ripoll, D. Zueco, T. Hümmer, E. Solano, et al., “Circuit quantum electrodynamics in the ultrastrong-coupling regime,” *Nature Physics* **6**, 772 (2010).
- [52] D. Hagenmüller, S. De Liberato, and C. Ciuti, “Ultrastrong coupling between a cavity resonator and the cyclotron transition of a two-dimensional electron gas in the case of an integer filling factor,” *Physical Review B* **81**, 235303 (2010).
- [53] A. Bayer, M. Pozimski, S. Schambeck, D. Schuh, R. Huber, D. Bougeard, and C. Lange, “Terahertz light-matter interaction beyond unity coupling strength,” *Nano Letters* **17**, 6340 (2017).
- [54] W. C. Chew, *Waves and fields in inhomogeneous media*, Vol. 16 (Inst of Electrical &, 1995).
- [55] S. M. Barnett, B. Huttner, R. Loudon, and R. Matloob, “Decay of excited atoms in absorbing dielectrics,” *Journal of Physics B: Atomic, Molecular and Optical Physics* **29**, 3763 (1996).
- [56] Y. Chen, T. R. Nielsen, N. Gregersen, P. Lodahl, and J. Mørk, “Finite-element modeling of spontaneous emission of a quantum emitter at nanoscale proximity to plasmonic waveguides,” *Physical Review B* **81**, 125431 (2010).
- [57] M. Born and R. Oppenheimer, “Zur quantentheorie der molekeln,” *Ann. Phys.* **389**, 457 (1927).
- [58] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” *Phys. Rev.* **136**, B864 (1964).
- [59] L. H. Thomas, “The calculation of atomic fields,” in *Mathematical proceedings of the cambridge philosophical society*, Vol. 23, 5 (Cambridge University Press, 1927), pages 542–548.

- [60] E. Fermi, "Un metodo statistico per la determinazione di alcune priorieta dell'atome," *Rend. Accad. Naz. Lincei* **6**, 32 (1927).
- [61] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.* **140**, A1133 (1965).
- [62] W. Koch and M. C. Holthausen, *A chemist's guide to density functional theory* (John Wiley & Sons, 2015).
- [63] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Can. J. Phys.* **58**, 1200 (1980).
- [64] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," *Phys. Rev. Lett.* **45**, 566 (1980).
- [65] J. Perdew, E. McMullen, and A. Zunger, "Density-functional theory of the correlation energy in atoms and ions: a simple analytic model and a challenge," *Phys. Rev. A* **23**, 2785 (1981).
- [66] J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas," *Phys. Rev. B* **33**, 8822 (1986).
- [67] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865 (1996).
- [68] J. P. Perdew, "Density functional theory and the band gap problem," *Int. J. Quantum Chem.* **28**, 497 (1985).
- [69] O. Gritsenko, R. van Leeuwen, E. van Lenthe, and E. J. Baerends, "Self-consistent approximation to the kohn-sham exchange potential," *Physical Review A* **51**, 1944 (1995).
- [70] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, "Kohn-sham potential with discontinuity for band gap materials," *Physical Review B* **82**, 115106 (2010).
- [71] J. Kohanoff, *Electronic structure calculations for solids and molecules: theory and computational methods* (Cambridge university press, 2006).
- [72] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened coulomb potential," *J. Chem. Phys.* **118**, 8207 (2003).
- [73] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: the pbe0 model," *The Journal of chemical physics* **110**, 6158 (1999).
- [74] E. Runge and E. K. Gross, "Density-functional theory for time-dependent systems," *Physical review letters* **52**, 997 (1984).
- [75] M. A. Marques and E. K. Gross, "Time-dependent density functional theory," *Annual review of physical chemistry* **55**, 427 (2004).
- [76] N. Tancogne-Dejean, M. J. Oliveira, X. Andrade, H. Appel, C. H. Borca, G. Le Breton, F. Buchholz, A. Castro, S. Corni, A. A. Correa, et al., "Octopus, a computational framework for exploring light-driven phenomena and quantum dynamics in extended and finite systems," *The Journal of chemical physics* **152**, 124119 (2020).

- [77] J. Flick, D. M. Welakuh, M. Ruggenthaler, H. Appel, and A. Rubio, "Light-matter response in nonrelativistic quantum electrodynamics," *ACS photonics* **6**, 2757 (2019).
- [78] S. Ghimire and D. A. Reis, "High-harmonic generation from solids," *Nature physics* **15**, 10 (2019).
- [79] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear science at the dawn of the 21st century* (Springer, 2000), pages 195–211.
- [80] R. Kubo, "The fluctuation-dissipation theorem," *Reports on progress in physics* **29**, 255 (1966).
- [81] G. Grosso and G. P. Parravicini, *Solid state physics* (Academic press, 2013).
- [82] M. E. Casida, "Time-dependent density-functional theory for molecules and molecular solids," *Journal of Molecular Structure: THEOCHEM* **914**, 3 (2009).
- [83] M. E. Casida, "Time-dependent density functional response theory for molecules," in *Recent advances in density functional methods: (part i)* (World Scientific, 1995), pages 155–192.
- [84] D. S. Wang, T. Neuman, J. Flick, and P. Narang, "Light-matter interaction of a molecule in a dissipative cavity from first principles," *The Journal of Chemical Physics* **154**, 104109 (2021).
- [85] J. Feist, A. I. Fernández-Domínguez, and F. J. García-Vidal, "Macroscopic qed for quantum nanophotonics: emitter-centered modes as a minimal basis for multi-emitter problems," *Nanophotonics* **10**, 477 (2021).
- [86] H. T. Dung, L. Knöll, and D.-G. Welsch, "Spontaneous decay in the presence of dispersing and absorbing bodies: general theory and application to a spherical cavity," *Physical Review A* **62**, 053804 (2000).
- [87] M. G. Blaber, M. D. Arnold, and M. J. Ford, "Search for the ideal plasmonic nanoshell: the effects of surface scattering and alternatives to gold and silver," *The Journal of Physical Chemistry C* **113**, 3041 (2009).
- [88] T. P. Rossi, T. Shegai, P. Erhart, and T. J. Antosiewicz, "Strong plasmon-molecule coupling at the nanoscale revealed by first-principles modeling," *Nature communications* **10**, 3336 (2019).
- [89] G. M. Akselrod, C. Argyropoulos, T. B. Hoang, C. Ciracì, C. Fang, J. Huang, D. R. Smith, and M. H. Mikkelsen, "Probing the mechanisms of large Purcell enhancement in plasmonic nanoantennas," *Nature Photonics* **8**, 835 (2014).
- [90] K. S. Thygesen, "Calculating excitons, plasmons, and quasiparticles in 2d materials and van der Waals heterostructures," *2D Materials* **4**, 022004 (2017).
- [91] K. Andersen, K. W. Jacobsen, and K. S. Thygesen, "Spatially resolved quantum plasmon modes in metallic nano-films from first-principles," *Physical Review B* **86**, 245129 (2012).
- [92] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body Green's-function approaches," *Reviews of Modern Physics* **74**, 601 (2002).

- [93] D. Golze, M. Dvorak, and P. Rinke, “The gw compendium: a practical guide to theoretical photoemission spectroscopy,” *Frontiers in chemistry* **7**, 377 (2019).
- [94] P. C. Martin and J. Schwinger, “Theory of many-particle systems. i,” *Physical Review* **115**, 1342 (1959).
- [95] L. Hedin, “New method for calculating the one-particle Green’s function with application to the electron-gas problem,” *Phys. Rev.* **139**, A796 (1965).
- [96] F. G. De Abajo, “Optical excitations in electron microscopy,” *Reviews of modern physics* **82**, 209 (2010).
- [97] D. B. Williams and C. B. Carter, “The transmission electron microscope,” in *Transmission electron microscopy*, Second edition (Springer, 2009).
- [98] D. E. Aspnes and A. Studna, “Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV,” *Phys. Rev. B* **27**, 985 (1983).
- [99] A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk’yanchuk, “Optically resonant dielectric nanostructures,” *Science* **354**, aag2472 (2016).
- [100] D. G. Baranov, D. A. Zuev, S. I. Lepeshov, O. V. Kotov, A. E. Krasnok, A. B. Evlyukhin, and B. N. Chichkov, “All-dielectric nanophotonics: the quest for better materials and fabrication techniques,” *Optica* **4**, 814 (2017).
- [101] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, “The nitrogen-vacancy colour centre in diamond,” *Phys. Rep.* **528**, 1 (2013).
- [102] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, “Linear optical quantum computing with photonic qubits,” *Reviews of modern physics* **79**, 135 (2007).
- [103] A. Strathearn, P. Kirton, D. Kilda, J. Keeling, and B. W. Lovett, “Efficient non-markovian quantum dynamics using time-evolving matrix product operators,” *Nature communications* **9**, 3322 (2018).
- [104] H.-P. Breuer, F. Petruccione, et al., *The theory of open quantum systems* (Oxford University Press on Demand, 2002).
- [105] A. Nazir and D. P. McCutcheon, “Modelling exciton–phonon interactions in optically driven quantum dots,” *Journal of Physics: Condensed Matter* **28**, 103002 (2016).
- [106] M. Bundgaard-Nielsen, J. Mørk, and E. V. Denning, “Non-markovian perturbation theories for phonon effects in strong-coupling cavity quantum electrodynamics,” *Physical Review B* **103**, 235309 (2021).
- [107] N. Makri and D. E. Makarov, “Tensor propagator for iterative quantum time evolution of reduced density matrices. i. theory,” *The Journal of chemical physics* **102**, 4600 (1995).
- [108] N. Makri and D. E. Makarov, “Tensor propagator for iterative quantum time evolution of reduced density matrices. ii. numerical methodology,” *The Journal of chemical physics* **102**, 4611 (1995).

- [109] P. Nalbach, A. Ishizaki, G. R. Fleming, and M. Thorwart, “Iterative path-integral algorithm versus cumulant time-nonlocal master equation approach for dissipative biomolecular exciton transport,” *New Journal of Physics* **13**, 063040 (2011).
- [110] M. B. Hastings, “An area law for one-dimensional quantum systems,” *Journal of statistical mechanics: theory and experiment* **2007**, P08024 (2007).
- [111] U. Schollwöck, “The density-matrix renormalization group in the age of matrix product states,” *Annals of physics* **326**, 96 (2011).
- [112] R. Orús, “A practical introduction to tensor networks: matrix product states and projected entangled pair states,” *Annals of physics* **349**, 117 (2014).
- [113] J. Hauschild and F. Pollmann, “Efficient numerical simulations with tensor networks: tensor network python (tenpy),” *SciPost Physics Lecture Notes*, 005 (2018).
- [114] G. Vidal, “Efficient simulation of one-dimensional quantum many-body systems,” *Physical review letters* **93**, 040502 (2004).
- [115] S. R. White, “Density matrix formulation for quantum renormalization groups,” *Physical review letters* **69**, 2863 (1992).
- [116] A. Novikov, D. Podoprikin, A. Osokin, and D. P. Vetrov, “Tensorizing neural networks,” *Advances in neural information processing systems* **28** (2015).
- [117] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, D. P. Mandic, et al., “Tensor networks for dimensionality reduction and large-scale optimization: part 1 low-rank tensor decompositions,” *Foundations and Trends in Machine Learning* **9**, 249 (2016).
- [118] Z.-Y. Han, J. Wang, H. Fan, L. Wang, and P. Zhang, “Unsupervised generative modeling using matrix product states,” *Physical Review X* **8**, 031012 (2018).
- [119] M. R. Jørgensen and F. A. Pollock, “Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-markovian path integrals,” *Physical review letters* **123**, 240602 (2019).
- [120] N. Rivera, I. Kaminer, B. Zhen, J. D. Joannopoulos, and M. Soljačić, “Shrinking light to allow forbidden transitions on the atomic scale,” *Science* **353**, 263 (2016).
- [121] P. A. D. Gonçalves and N. M. Peres, *An introduction to graphene plasmonics* (World Scientific, 2016).
- [122] E. J. Dias, D. A. Iranzo, P. Gonçalves, Y. Hajati, Y. V. Bludov, A.-P. Jauho, N. A. Mortensen, F. H. Koppens, and N. Peres, “Probing nonlocal effects in metals with graphene plasmons,” *Physical Review B* **97**, 245405 (2018).
- [123] D. Alcaraz Iranzo, S. Nanot, E. J. Dias, I. Epstein, C. Peng, D. K. Efetov, M. B. Lundeberg, R. Parret, J. Osmond, J.-Y. Hong, et al., “Probing the ultimate plasmon confinement limits with a van der waals heterostructure,” *Science* **360**, 291 (2018).

- [124] M. B. Lundeberg, Y. Gao, R. Asgari, C. Tan, B. Van Duppen, M. Autore, P. Alonso-González, A. Woessner, K. Watanabe, T. Taniguchi, et al., “Tuning quantum nonlocal effects in graphene plasmonics,” *Science* **357**, 187 (2017).
- [125] B. Hatfield, *Quantum field theory of point particles and strings* (CRC Press, 2018).
- [126] K. Kaasbjerg, K. S. Thygesen, and A.-P. Jauho, “Acoustic phonon limited mobility in two-dimensional semiconductors: deformation potential and piezoelectric scattering in monolayer mos2 from first principles,” *Physical Review B* **87**, 235312 (2013).
- [127] S. Michaelis de Vasconcellos, D. Wigger, U. Wurstbauer, A. W. Holleitner, R. Bratschitsch, and T. Kuhn, “Single-photon emitters in layered van der waals materials,” *physica status solidi (b)* **259**, 2100566 (2022).
- [128] A. Reserbat-Plantey, I. Epstein, I. Torre, A. T. Costa, P. Gonçalves, N. A. Mortensen, M. Polini, J. C. Song, N. M. Peres, and F. H. Koppens, “Quantum nanophotonics in two-dimensional materials,” *ACS Photonics* **8**, 85 (2021).
- [129] T. T. Tran, K. Bray, M. J. Ford, M. Toth, and I. Aharonovich, “Quantum emission from hexagonal boron nitride monolayers,” *Nature nanotechnology* **11**, 37 (2016).
- [130] T. T. Tran, C. Elbadawi, D. Totonjian, C. J. Lobo, G. Grosso, H. Moon, D. R. Englund, M. J. Ford, I. Aharonovich, and M. Toth, “Robust multicolor single photon emission from point defects in hexagonal boron nitride,” *ACS Nano* **10**, 7331 (2016).
- [131] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van de Walle, “First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres,” *New J. Phys.* **16**, 073026 (2014).
- [132] A. Sajid and K. S. Thygesen, “Vncb defect as source of single photon emission from hexagonal boron nitride,” *2D Materials* **7**, 031007 (2020).
- [133] K. Li, T. J. Smart, and Y. Ping, “Carbon trimer as a 2 ev single-photon emitter candidate in hexagonal boron nitride: a first-principles study,” *Physical Review Materials* **6**, L042201 (2022).
- [134] R. Borrelli, A. Capobianco, and A. Peluso, “Franck–condon factors—computational approaches and recent developments,” *Canadian Journal of Chemistry* **91**, 495 (2013).
- [135] S. Latini, U. De Giovannini, E. J. Sie, N. Gedik, H. Hübener, and A. Rubio, “Phonoritons as hybridized exciton-photon-phonon excitations in a monolayer h-bn optical cavity,” *Physical Review Letters* **126**, 227401 (2021).
- [136] M. L. Tseng, Y. Jahani, A. Leitis, and H. Altug, “Dielectric metasurfaces enabling advanced optical biosensors,” *ACS Photonics* **8**, 47 (2020).
- [137] J. A. Schuller, E. S. Barnard, W. Cai, Y. C. Jun, J. S. White, and M. L. Brongersma, “Plasmonics for extreme light concentration and manipulation,” *Nature materials* **9**, 193 (2010).

- [138] U. Kreibig and L. Genzel, "Optical absorption of small metallic particles," *Surface Science* **156**, 678 (1985).
- [139] J. B. Khurgin, "Ultimate limit of field confinement by surface plasmon polaritons," *Faraday discussions* **178**, 109 (2015).
- [140] G. Mie, "Beiträge zur optik trüber medien, speziell kolloidaler metallösungen," *Annalen der physik* **330**, 377 (1908).
- [141] M. Kerker, D.-S. Wang, and C. Giles, "Electromagnetic scattering by magnetic spheres," *JOSA* **73**, 765 (1983).
- [142] T. Moss, "Relations between the refractive index and energy gap of semiconductors," *physica status solidi (b)* **131**, 415 (1985).
- [143] W. T. Chen, A. Y. Zhu, V. Sanjeev, M. Khorasaninejad, Z. Shi, E. Lee, and F. Capasso, "A broadband achromatic metalens for focusing and imaging in the visible," *Nature nanotechnology* **13**, 220 (2018).
- [144] K. Huang, J. Deng, H. S. Leong, S. L. K. Yap, R. B. Yang, J. Teng, and H. Liu, "Ultraviolet metasurfaces of 80% efficiency with antiferromagnetic resonances for optical vectorial anti-counterfeiting," *Laser & Photonics Reviews* **13**, 1800289 (2019).
- [145] C. Zhang, S. Divitt, Q. Fan, W. Zhu, A. Agrawal, Y. Lu, T. Xu, and H. J. Lezec, "Low-loss metasurface optics down to the deep ultraviolet region," *Light: Science & Applications* **9**, 1 (2020).
- [146] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, "The open quantum materials database (oqmd): assessing the accuracy of dft formation energies," *npj Computational Materials* **65**, 1501 (2015).
- [147] M. Gjerding, T. Skovhus, A. Rasmussen, F. Bertoldo, A. H. Larsen, J. J. Mortensen, and K. S. Thygesen, "Atomic simulation recipes: a python framework and library for automated workflows," *Computational Materials Science* **199**, 110731 (2021).
- [148] J. J. Mortensen, M. Gjerding, and K. S. Thygesen, "MyQueue: task and workflow scheduling system," *J. Open Source Softw.* **5**, 1844 (2020).
- [149] R. Archer, R. Koyama, E. Loebner, and R. Lucas, "Optical absorption, electroluminescence, and the band gap of bp," *Physical Review Letters* **12**, 538 (1964).
- [150] E. Schrotten, A. Goossens, and J. Schoonman, "Photo- and electroreflectance of cubic boron phosphide," *J. Appl. Phys.* **83**, 1660 (1998).
- [151] A. Crovetto, J. M. Adamczyk, R. R. Schnepf, C. L. Perkins, H. Hempel, S. R. Bauers, E. S. Toberer, A. C. Tamboli, T. Unold, and A. Zakutayev, "Boron phosphide films by reactive sputtering: searching for a p-type transparent conductor," *Advanced Materials Interfaces* **9**, 2200031 (2022).
- [152] X.-G. Zhao, J.-H. Yang, Y. Fu, D. Yang, Q. Xu, L. Yu, S.-H. Wei, and L. Zhang, "Design of lead-free inorganic halide perovskites for solar cells via cation-transmutation," *Journal of the American Chemical Society* **139**, 2630 (2017).

- [153] K. Kuhar, M. Pandey, K. S. Thygesen, and K. W. Jacobsen, “High-throughput computational assessment of previously synthesized semiconductors for photovoltaic and photoelectrochemical devices,” *ACS Energy Letters* **3**, 436 (2018).
- [154] K. Choudhary, M. Bercx, J. Jiang, R. Pachter, D. Lamoen, and F. Tavazza, “Accelerated discovery of efficient solar cell materials using quantum and machine-learning methods,” *Chemistry of Materials* **31**, 5900 (2019).
- [155] Y. Li and K. Yang, “High-throughput computational design of organic–inorganic hybrid halide semiconductors beyond perovskites for optoelectronics,” *Energy & Environmental Science* **12**, 2233 (2019).
- [156] S. Luo, T. Li, X. Wang, M. Faizan, and L. Zhang, “High-throughput computational materials screening and discovery of optoelectronic semiconductors,” *Wiley Interdisciplinary Reviews: Computational Molecular Science* **11**, e1489 (2021).
- [157] D. Dahliah, G. Brunin, J. George, V.-A. Ha, G.-M. Rignanese, and G. Hautier, “High-throughput computational search for high carrier lifetime, defect-tolerant solar absorbers,” *Energy & Environmental Science* **14**, 5057 (2021).
- [158] Y. Gan, N. Miao, P. Lan, J. Zhou, S. R. Elliott, and Z. Sun, “Robust design of high-performance optoelectronic chalcogenide crystals from high-throughput computation,” *Journal of the American Chemical Society* **144**, 5878 (2022).
- [159] B. Blank, T. Kirchartz, S. Lany, and U. Rau, “Selection metric for photovoltaic materials screening based on detailed-balance analysis,” *Physical review applied* **8**, 024032 (2017).
- [160] S. Hastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, et al., “The computational 2d materials database: high-throughput modeling and discovery of atomically thin crystals,” *2D Mater.* **5**, 042002 (2018).
- [161] P. M. Larsen, M. Pandey, M. Strange, and K. W. Jacobsen, “Definition of a scoring parameter to identify low-dimensional materials components,” *Phys. Rev. Mater.* **3**, 034003 (2019).
- [162] F. Bertoldo, S. Ali, S. Manti, and K. S. Thygesen, “Quantum point defects in 2d materials—the qpod database,” *npj Computational Materials* **8**, 56 (2022).
- [163] N. R. Knøsgaard and K. S. Thygesen, “Representing individual electronic states for machine learning GW band structures of 2D materials,” *Nat. Commun.* **13**, 468 (2022).
- [164] G. M. Nascimento, E. Ogoshi, A. Fazzio, C. M. Acosta, and G. M. Dalpian, “High-throughput inverse design and bayesian optimization of functionalities: spin splitting in two-dimensional compounds,” *Scientific data* **9**, 195 (2022).
- [165] R. Peierls, *Quantum theory of solids*, International Series of Monographs on Physics (Clarendon Press, 1996).

- [166] N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, and N. Marzari, “Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds,” *Nature Nanotechnology* **13**, 246 (2018).
- [167] T. Chen and C. Guestrin, “XGBoost: a scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, KDD '16* (2016), pages 785–794.
- [168] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, Vol. 2 (Springer, 2009).
- [169] Y. Wang, L. Li, W. Yao, S. Song, J. Sun, J. Pan, X. Ren, C. Li, E. Okunishi, Y.-Q. Wang, et al., “Monolayer ptse2, a new semiconducting transition-metal-dichalcogenide, epitaxially grown by direct selenization of pt,” *Nano letters* **15**, 4013 (2015).
- [170] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, “Electronics and optoelectronics of two-dimensional transition metal dichalcogenides,” *Nature nanotechnology* **7**, 699 (2012).
- [171] Y. Zhao, J. Qiao, P. Yu, Z. Hu, Z. Lin, S. P. Lau, Z. Liu, W. Ji, and Y. Chai, “Extraordinarily strong interlayer interaction in 2d layered pts2,” *Advanced materials* **28**, 2399 (2016).
- [172] Y. Zhao, J. Qiao, Z. Yu, P. Yu, K. Xu, S. P. Lau, W. Zhou, Z. Liu, X. Wang, W. Ji, et al., “High-electron-mobility and air-stable 2d layered ptse2 fets,” *Advanced Materials* **29**, 1604230 (2017).
- [173] A. Ciarrocchi, A. Avsar, D. Ovchinnikov, and A. Kis, “Thickness-modulated metal-to-semiconductor transformation in a transition metal dichalcogenide,” *Nature communications* **9**, 919 (2018).
- [174] X. Chia, A. Adriano, P. Lazar, Z. Sofer, J. Luxa, and M. Pumera, “Layered platinum dichalcogenides (pts2, ptse2, and ptte2) electrocatalysis: monotonic dependence on the chalcogen size,” *Advanced Functional Materials* **26**, 4306 (2016).
- [175] S. Lin, Y. Liu, Z. Hu, W. Lu, C. H. Mak, L. Zeng, J. Zhao, Y. Li, F. Yan, Y. H. Tsang, et al., “Tunable active edge sites in ptse2 films towards hydrogen evolution reaction,” *Nano Energy* **42**, 26 (2017).
- [176] H. Xu, H. Zhang, Y. Liu, S. Zhang, Y. Sun, Z. Guo, Y. Sheng, X. Wang, C. Luo, X. Wu, et al., “Controlled doping of wafer-scale ptse2 films for device application,” *Advanced Functional Materials* **29**, 1805614 (2019).
- [177] X. Yu, P. Yu, D. Wu, B. Singh, Q. Zeng, H. Lin, W. Zhou, J. Lin, K. Suenaga, Z. Liu, et al., “Atomically thin noble metal dichalcogenide: a broadband mid-infrared semiconductor,” *Nature communications* **9**, 1545 (2018).
- [178] J. Yan, K. W. Jacobsen, and K. S. Thygesen, “Conventional and acoustic surface plasmons on noble metal surfaces: a time-dependent density functional theory study,” *Physical Review B* **86**, 241404 (2012).

-
- [179] B. K. Agarwal, *X-ray spectroscopy: an introduction*, Vol. 15 (Springer, 2013).
- [180] W. a. Ackermann, G. Asova, V. Ayvazyan, A. Azima, N. Baboi, J. Bähr, V. Balandin, B. Beutner, A. Brandt, A. Bolzmann, et al., “Operation of a free-electron laser from the extreme ultraviolet to the water window,” *Nature photonics* **1**, 336 (2007).
- [181] H. Winick and S. Doniach, *Synchrotron radiation research* (Springer Science & Business Media, 2012).
- [182] D. Attwood, *Soft x-rays and extreme ultraviolet radiation: principles and applications* (Cambridge university press, 2000).
- [183] N. D. Powers, I. Ghebregziabher, G. Golovin, C. Liu, S. Chen, S. Banerjee, J. Zhang, and D. P. Umstadter, “Quasi-monoenergetic and tunable x-rays from a laser-driven compton light source,” *Nature Photonics* **8**, 28 (2014).
- [184] D. J. Griffiths, *Introduction to electrodynamics*, 2005.

CHAPTER 11

Papers

- 11.1 Paper [I]: Combining density functional theory with macroscopic QED for quantum light-matter interactions in 2D materials

Mark Kamper Svendsen[†], Yaniv Kurman[†], Peter Schmidt, Frank Koppens, Ido Kaminer, Kristian S Thygesen

“Combining density functional theory with macroscopic QED for quantum light-matter interactions in 2D materials,”

[†] The authors contributed equally to the work,
[Nature Communications 12, 2778 \(2021\)](#).

Combining density functional theory with macroscopic QED for quantum light-matter interactions in 2D materials

Mark Kamper Svendsen^{1,6}[✉], Yaniv Kurman^{2,6}, Peter Schmidt³, Frank Koppens^{3,4}, Ido Kaminer² & Kristian S. Thygesen⁵

A quantitative and predictive theory of quantum light-matter interactions in ultra thin materials involves several fundamental challenges. Any realistic model must simultaneously account for the ultra-confined plasmonic modes and their quantization in the presence of losses, while describing the electronic states from first principles. Herein we develop such a framework by combining density functional theory (DFT) with macroscopic quantum electrodynamics, which we use to show Purcell enhancements reaching 10^7 for intersubband transitions in few-layer transition metal dichalcogenides sandwiched between graphene and a perfect conductor. The general validity of our methodology allows us to put several common approximation paradigms to quantitative test, namely the dipole-approximation, the use of 1D quantum well model wave functions, and the Fermi's Golden rule. The analysis shows that the choice of wave functions is of particular importance. Our work lays the foundation for practical ab initio-based quantum treatments of light-matter interactions in realistic nanostructured materials.

¹CAMD, Department of Physics, Technical University of Denmark, Kgs. Lyngby, Denmark. ²Department of Electrical Engineering, Technion, Israel Institute of Technology, Haifa, Israel. ³ICFO-Institut de Ciències Fòniques, The Barcelona Institute of Science and Technology, Castelldefels (Barcelona), Spain. ⁴ICREA – Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. ⁵CAMD and Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, Kgs. Lyngby, Denmark. ⁶These authors contributed equally: Mark Kamper Svendsen, Yaniv Kurman. [✉]email: markas@dtu.dk

The development of van der Waals (vdW) heterostructures composed of atomically thin 2D layers has opened new opportunities for manipulating and enhancing light-matter interactions at the nanoscale. The very tight confinement of the electrons in these materials endows them with unique physical properties that can be altered and controlled by varying the number, type, and sequence of the atomic layers^{1,2}. For example, the experiment, which served as motivation for the current theoretical work, demonstrated thickness-tunable infrared emission via intersubband transitions in vdW quantum wells (QWs) comprised of few-layer transition metal dichalcogenide (TMD) structures³. At the same time, vdW heterostructures can support light-matter hybrid modes such as plasmon-, phonon-, and exciton-polaritons^{4,5} with ultra-confined electromagnetic (EM) fields in the direction normal to the 2D plane. The coupling between such polaritonic light modes and electron systems in nearby layers of the heterostructure provides an interesting avenue for controlling the optical properties of the material or even creating new types of quantum states—if sufficiently strong coupling can be realized. However, the theoretical description and, in particular, the practical computation, of such coupled quantum light-matter systems present significant challenges that call for new developments^{6–8}.

To set the stage for the following discussions, we briefly outline the most central notions and approximations relevant for the description of light-matter interactions in nanoscale materials, see Fig. 1. First, the light-matter interaction is often described within the dipole approximation, which is valid whenever the EM field can be assumed to be constant over the extent of the matter system, or more precisely, the extent of the relevant electronic wave functions. The dipole approximation works well in many

cases but becomes problematic for confined EM modes, such as the plasmon polaritons considered in the present work^{9,10}. In such cases, the detailed shape of the wave function (and the light field) must be taken into account when evaluating the light-matter interaction. The question of how detailed the description of the electron wave function must be remains largely unexplored, and obviously depends on the specifics of the considered system and the targeted accuracy. For the TMD/graphene heterostructures studied in this work, we find that the difference in the spontaneous emission rates obtained with model wave functions and full ab-initio wave functions, can differ by an order of magnitude and show clear qualitative differences as well.

When cavities or plasmonic structures are used to select and enhance distinct light modes, it is possible to reach the strong-coupling regime where the pure light and matter states become entangled. For this to happen, the light-matter coupling strength must exceed the damping rate of the light mode; or equivalently, the radiative lifetime of the matter state should be shorter than the lifetime of the light mode, such that emission-reabsorption processes can occur. We note in passing that the lifetime of the light mode is determined by the optical losses occurring in the relevant materials (see below). Under strong-coupling conditions it is necessary to diagonalize the light-matter Hamiltonian, or alternatively evolve the state dynamics, using non-perturbative methods (in practice within a truncated Hilbert space)^{11–13}. However, in many cases of interest the coupling will be sufficiently weak that first-order perturbation theory, i.e. Fermi's Golden rule, applies. The combination of the dipole approximation and Fermi's Golden rule yields an electronic transition rate given by the product of the dipole strength and the local photonic density of states (DOS). In this case, owing to the equivalence of

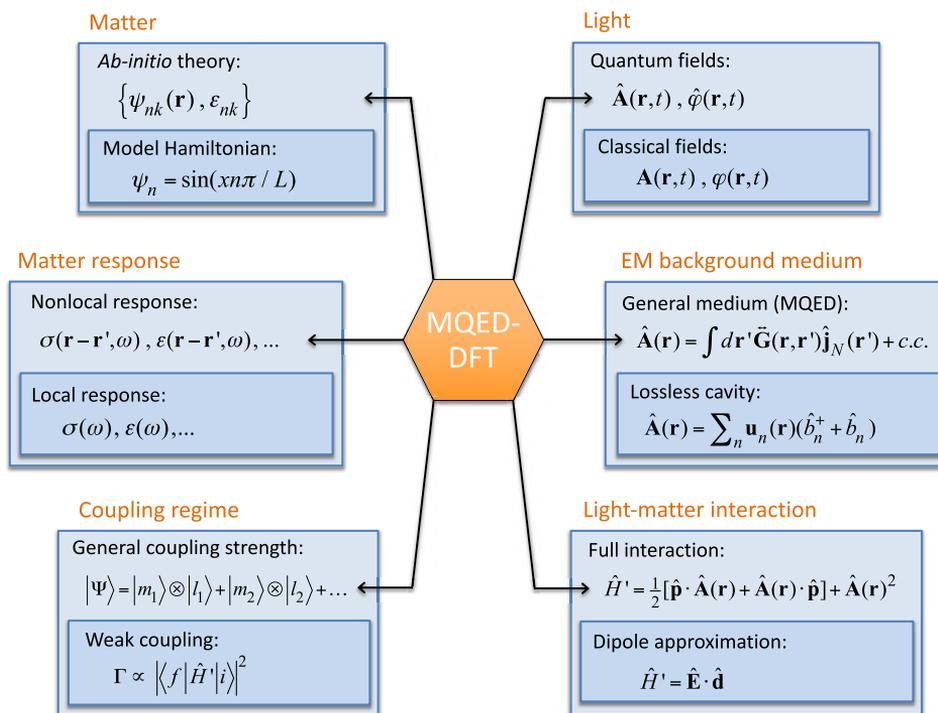


Fig. 1 Schematic illustration of the key quantities (blue boxes) entering the description of light-matter interacting systems. For each quantity, the most general/accurate level of the theory is shown at the top while commonly used approximations are shown below.

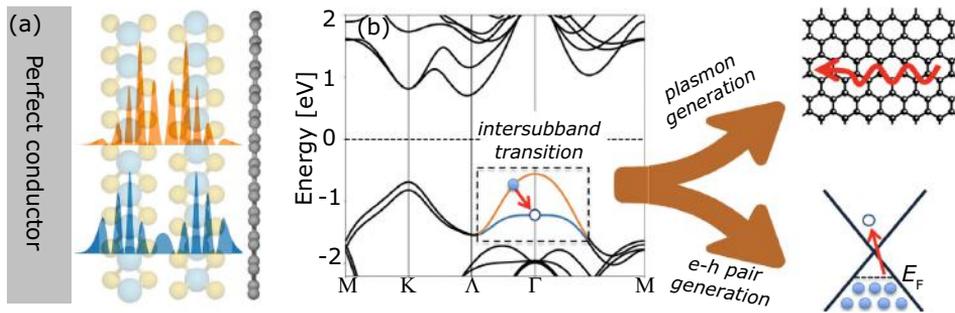


Fig. 2 Combining DFT and MQED in van der Waals materials. **a** Illustration of a 2-layer TMD quantum well sandwiched between a perfect conductor and a doped graphene sheet. The z -profile of two subband wave functions obtained from DFT are also shown. **b** The DFT band structure of bilayer WSe₂ with the two subbands at the Γ -point indicated in orange and blue. An electron from the filled upper subband can combine with a hole at the top of the lower subband by emitting either an electron-hole pair (loss) or a propagating acoustic plasmon-polariton.

the local photon DOS and the EM self-field of a point dipole, the spontaneous emission can be perceived as a semi-classical effect that depends on the electron wave functions only through the effective dipole strength. On the other hand, when the EM field varies on the scale of the electron wave function, the dipole-approximation breaks down⁹, selection rules change^{7,10} and the local semi-classical picture must be replaced by a full quantum treatment that explicitly incorporates the nonlocal nature of light-matter interactions.

The conventional approach to the quantization of the EM field is based on an expansion into orthogonal modes. However, in the presence of optical losses, i.e. when the conductivity of the EM background medium is finite in the relevant frequency range, the expansion into orthogonal modes is ill-defined and the quantization must be performed using alternative approaches such as macroscopic quantum electrodynamics (MQED)¹⁴ or expansion in quasi-normal modes¹⁵. Finally, it is often necessary to invoke the optical response functions of the involved materials. This is for example the case for the MQED formalism to be discussed below. While spatially local, i.e. q -independent, response functions are often employed, and indeed work well in extended systems, this approximation becomes questionable in nanoscale structures such as 2D materials^{16,17}.

As will be shown, all of the common approximations used to describe quantum light-matter systems (see Fig. 1) can be violated in vdW QW devices. Consequently, these systems are ideal as testbeds for the testing and development of theoretical methods in this exciting field.

In recent years, there has been significant developments toward a full ab-initio description of coupled light-matter systems based on the fundamental Hamiltonian of non-relativistic QED, also known as the Pauli-Fierz Hamiltonian¹⁸. In particular, the quantum-electrodynamical density functional theory (QEDFT), which is a density functional type reformulation of the Pauli-Fierz Hamiltonian¹⁹, has been introduced as a generalization of time-dependent density functional theory (DFT) to cases where the quantum nature of the light is important^{18,20}. The fundamental nature of the QEDFT formalism has, for example, enabled studies of atomic emitters in cavities where the back-action of the EM field leads to “dressing” of the electronic states^{21,22}. Although general, the QEDFT approach still has its limitations. An important one for making quantitative predictions of light-matter interactions in real materials, is the assumption of a lossless²³ or low-loss²⁴ EM background medium. Although this problem can in principle be alleviated by including the lossy environment as part of the active matter system¹⁸ it is unlikely that this would be practical for complex nano-optical

setups, such as those involving vdW-heterostructures^{1,6}. We further note that most applications of QEDFT have been restricted to the dipole approximation, although steps towards a more general description have recently been taken^{24,25}. Finally, most QEDFT studies have been restricted to finite systems, such as atoms and small molecules, while the application to extended systems has been limited.

In this paper, we introduce an approach based on the Wigner-Weisskopf model to describe quantum light-matter interactions in complex 2D material setups under quite general conditions. Our methodology, combines MQED for quantizing the optical modes of an arbitrary optical environment with the single particle Kohn-Sham wave functions and energies from DFT for the electrons. This combination enables a quantitative description of the physics in a regime where quantized light modes, confined to a length scale of the electron wave function, interact strongly with the electrons while being subject to losses in the EM background medium. We use the method to calculate the radiative transition rates of a multilayer TMD QW sandwiched between graphene and a metal surface, see Fig. 2a. By varying the type of TMD material, the number of layers, and the graphene doping level, we determine the maximal Purcell enhancement achievable in such a setup to be 10^7 , corresponding to intersubband transition rates of 4.5 THz with high efficiency of emission into the propagating plasmons. For the thinnest 2-layer stacks we observe somewhat lower Purcell factors of around 10^6 , but higher absolute rates approaching 30 THz. These radiative lifetimes are at least an order of magnitude lower than the intrinsic (electron-phonon limited) lifetimes of electrons and holes in TMDs at low temperatures^{3,26,27}. Consequently, the coupling to the graphene plasmons should indeed be the dominant mechanism governing the dynamics of the intersubband transitions at low temperatures. We compare our DFT-MQED results with common approximation schemes such as the dipole approximation and the use of 1D QW wave function models. We find quantitative differences in the obtained rates of up to an order of magnitude. In particular, we find that the rates are highly sensitive to the detailed spatial shape of the wave function, emphasizing the need for a proper description of the electronic states. Lastly, we explore cases where the first-order QED breaks down and higher-order processes must be considered to obtain a correct description of the Purcell effect.

Results

The MQED-DFT framework. The framework described in the following section is motivated by the new regimes of light-matter

interactions that are now accessible in nano-materials such as vdW heterostructures. Although the presented framework is general, we shall focus on the case of 2D layered geometries and apply the formalism to a TMD QW sandwiched between a conductor and a graphene sheet, see Fig. 2a. Such heterostructures feature rapid emission from the QW into acoustic graphene plasmons^{28–30} that require a beyond-dipole MQED framework with ab-initio QW wave functions. We note that our work is the first to combine MQED with DFT. Moreover, the use of the Wigner–Weisskopf model enables us to analyze cases within the strong-coupling regime. A similar use of the Wigner–Weisskopf model can be found in^{24,31}, which also go beyond the Fermi’s golden rule. We present here the main steps of our framework, while the details are described in Supplementary note 1.

We represent the light-matter interaction Hamiltonian in the Weyl gauge (which is defined by the requirement of a vanishing scalar potential),

$$\hat{H}_{\text{int}} = \frac{e}{m} \hat{\mathbf{A}} \cdot \hat{\mathbf{p}} - \frac{ie\hbar}{2m} \nabla \cdot \hat{\mathbf{A}}, \quad (1)$$

where e is the electron charge, m is the electron mass, and $\hat{\mathbf{p}}$ represents the electron momentum operator. In writing Eq. (1) we neglected the diamagnetic $\hat{\mathbf{A}}^2$ term. The justification of this assumption and its implications are discussed in the “Limitations of MQED-DFT” subsection below. The Weyl gauge is an incomplete gauge and should therefore be accompanied by an additional condition to completely fix the gauge to avoid unphysical excitations arising in the quantization scheme³². We mention in passing that this condition is met because $\nabla \cdot (\hat{\epsilon} \mathbf{A}) = 0$ inside of the anisotropic TMD layer, the justification of which is discussed later. We further note that this means that in an isotropic dielectric material, our gauge reduces to the normal Coulomb gauge ($\nabla \cdot \mathbf{A} = 0$) common in quantum optics. In such cases, the $\nabla \cdot \hat{\mathbf{A}}$ term disappears from Eq. (1), leaving only the $\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}$ term. The $\nabla \cdot \hat{\mathbf{A}}$ term can thus be thought of as an anisotropic correction and we have generally found it to be small relative to the $\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}$ term.

Within MQED in the Weyl gauge, the vector potential is expressed as^{33–35},

$$\hat{\mathbf{A}}_i(\mathbf{r}) = \sqrt{\frac{\hbar}{\pi\epsilon_0}} \int d\omega \frac{\omega}{c^2} \int d^3s \hat{\mathbf{G}}_{ii}(\mathbf{r}, \mathbf{s}, \omega) \left[\sqrt{\text{Im } \epsilon(\mathbf{s}, \omega)} \right]_{ij} \hat{f}_j(\mathbf{s}, \omega) + \text{h.c.}, \quad (2)$$

where \hbar is the reduced Planck’s constant, ϵ_0 is the vacuum permittivity, c is the speed of light, and the subscripts obey Einstein’s summation convention. The quantization scheme is build around the Langevin noise current annihilation and creation operators, $\hat{f}_i(\mathbf{s}, \omega) = \left[\sqrt{\text{Im } \epsilon(\mathbf{s}, \omega)} \right]_{ij} \hat{f}_j(\mathbf{s}, \omega)$ (and similar for \hat{f}_i^\dagger), with $\epsilon_{ij}(\mathbf{s}, \omega)$ being the relative permittivity. The classical Dyadic Green’s function (DGF), $\hat{\mathbf{G}}(\mathbf{r}, \mathbf{s}, \omega)$, which satisfies

$$\left[\nabla \times \nabla \times - \left(\frac{\omega^2}{c^2} \right) \hat{\epsilon}(\mathbf{r}, \omega) \right] \hat{\mathbf{G}}(\mathbf{r}, \mathbf{s}, \omega) = \hat{\mathbf{I}} \delta(\mathbf{r} - \mathbf{s}), \quad (3)$$

gives the electric field at position \mathbf{r} generated by a point dipole of frequency ω at position \mathbf{s} in the presence of an EM background medium with permittivity $\hat{\epsilon}$. In our case, the EM background permittivity is defined by the perfect conductor, the graphene sheet, and the TMD without contributions from the subband transitions (these define the active matter system, see below, and should be excluded from $\hat{\epsilon}$ to avoid double counting). In general, determining the DGF is a nontrivial task, which is often

accomplished using numerical methods³⁶. However, in our case, the layered geometry of the vdW-heterostructure and the accompanying in-plane translation symmetry makes it possible to obtain a closed analytical solution for the DGF (see the Methods section).

The active matter system is defined by the electrons in the valence subbands near the Γ -point of the TMD QW, see Fig. 2b. We determine the electronic structure of the subbands of the isolated TMD structure by means of DFT calculations. DFT allows us to access the band structure and the single particle wave functions of the subbands³⁷, which in turn are used to evaluate the coupling matrix elements of the light-matter interaction Hamiltonian as described below. The details of the DFT calculation are provided in the Methods section.

Due to the in-plane translational invariance of the system, the electronic states of the subbands can be labeled by a subband index and a wave vector. In the following we consider the process where an electron from the upper subband (orange band in Fig. 2b) recombines with a hole at the top of the lower subband (blue) with the EM field initially in its ground state, i.e. a spontaneous emission process. The rate of this transition can thus be seen as the inverse lifetime of a hole in the lower subband due to the coupling to the EM field. Denoting the ground state of the TMD by $|\Psi_0\rangle$, the initial state representing a hole at the top of the lower subband (l) is written $|i\rangle = \hat{c}_{l,0}|\Psi_0\rangle$, while the final state with a hole in the upper subband (u) at momentum \mathbf{q} is given by $|f_{\mathbf{q}}\rangle = \hat{c}_{u,\mathbf{q}}|\Psi_0\rangle$.

The time-dependent combined state of the QW and EM field takes the form

$$|\Psi(t)\rangle = C_l(t)e^{-i\omega_l t}|i\rangle \otimes |0\rangle + \sum_{\mathbf{q}} \sum_{\nu_{\mathbf{q}}} C_{f,\nu_{\mathbf{q}}}(t)e^{-i(\omega_{\mathbf{q}} + \omega_{\nu_{\mathbf{q}}})t} |f_{\mathbf{q}}\rangle \otimes |\nu_{\mathbf{q}}\rangle, \quad (4)$$

where $|\nu_{\mathbf{q}}\rangle$ labels a single photon state of the EM field, which due to the in-plane translational invariance of the layered structure considered in this work is described by a continuum of energies and in-plane momenta. We note that with the form of the many-body state in Eq. (4) we are limited to one-photon final states and thus we neglect multi-photon processes. Using the Wigner–Weisskopf approach, it can be shown that the time evolution of the initial state coefficient satisfies the equation of motion

$$\dot{C}_l(t) = - \int_0^t dt' \int d\omega \int d\mathbf{q} K(\omega, \mathbf{q}) e^{-i(\omega - \omega_{lq})(t-t')} C_l(t'), \quad (5)$$

where $\hbar\omega_{lq} = \epsilon_{lq} - \epsilon_i$ is the electronic transition energy and the coupling kernel is given by

$$K(\omega, \mathbf{q}) = - \frac{1}{4\pi^2 4m^2} \sum_{\nu_{\mathbf{q}}} \left| \langle f_{\mathbf{q}} \otimes \nu_{\mathbf{q}} | \hat{H}_{\text{int}} | i \otimes 0 \rangle \right|^2 \delta(\omega - \omega_{\nu_{\mathbf{q}}}). \quad (6)$$

The Wigner–Weisskopf approach defined by Eqs. (5) and (6) describes the decay of the excited state into the EM environment and includes all optical one-photon decoherence channels. We note that via proper generalization other loss and decoherence sources, such as electron–phonon coupling, not considered in the current work, could in principle be accounted for by appropriate generalization of the many-body state and the coupling kernel. To account for pure dephasing a density matrix representation of the state is required and the Hilbert space would have to be extended to include the sets of states, $\{|i\rangle \otimes |\nu_{\mathbf{q}}\rangle\}$ and $\{|f_{\mathbf{q}}\rangle \otimes |0\rangle\}$.

The kernel, $K(\omega, \mathbf{q})$, in Eq. (6) describes the coupling strength between the electronic subband transitions and the optical excitations with frequency ω and in-plane momentum \mathbf{q} . To properly account for the nonlocality of the graphene, we employ the full nonlocal conductivity of the graphene³⁸ as described in

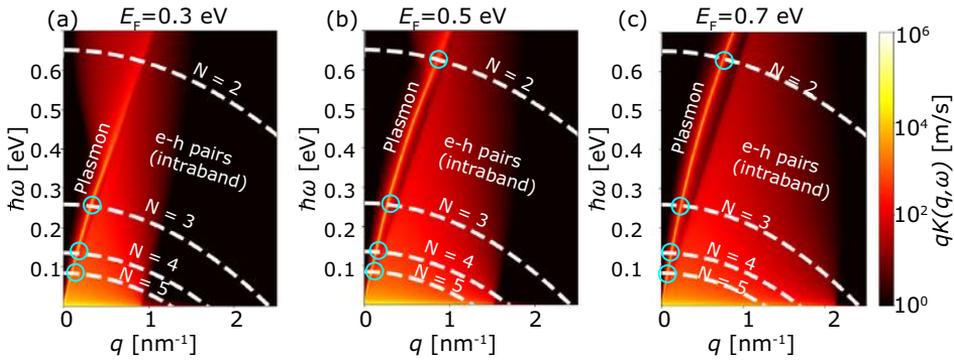


Fig. 3 Overlapping the electronic bandstructure and the plasmon dispersion to calculate their light-matter interactions. The coupling kernel for the transition between of holes between the lower and upper subband in a two-layer WSe₂ quantum well placed between a perfect conductor and a doped graphene sheet is shown for three different values of the graphene Fermi level, **a** 0.3 eV, **b** 0.5 eV and **c** 0.7 eV. The kernel describes the coupling into all optical excitations, including the plasmon and the electron-hole continuum as described by graphene’s full nonlocal conductivity³⁸. The dashed lines represent the non-vertical emission energies, $\hbar\omega_{i,q}$, derived from the DFT band structure for WSe₂ quantum wells of $N = 2, 3, 4,$ and 5 layers and the light blue circles denote the intersection between the dispersion of the intersubband transition energy and the graphene plasmon.

Methods. Importantly, we model the graphene via a nonlocal surface conductivity, neglecting the thickness of the graphene sheet. We note that treating the graphene as a 3D material with a finite thickness (with a bulk conductivity instead of a surface conductivity) should not affect the graphene plasmonic properties and thus not affect the results. The graphene conductivity depends strongly on the graphene Fermi level. The kernel therefore has to be calculated independently for each value of the graphene Fermi level. Exploiting the in-plane isotropy of the subband band structures, in Fig. 3 we show a color map of the coupling kernel integrated over the polar angle, $qK(\omega, q)$, for a 2-layer WSe₂ QW at three different values of the graphene Fermi level, alongside the transition energies, $\omega_{i,q}$, for the different numbers of layers. The acoustic graphene plasmon and the dissipative electron-hole (e-h) continuum are clearly visible.

From Eq. (6) it follows that the magnitude of the kernel at a given (ω, q) depends on two factors: the spatial overlap of the subband wave functions and the electric field of the optical mode (the matrix element), and the number of available optical modes at (ω, q) (the photonic DOS). The graphene plasmons affect the kernel via large matrix elements originating from their strong EM field. On the other hand, the e-h excitations in graphene individually produce weak EM fields but due to their large number they still can have a significant influence on the kernel, especially at small frequencies compared to the Fermi level.

There are several aspects of the formalism worth highlighting. First, we emphasize that the normal modes of the EM field, $|\nu_q\rangle$, are used only symbolically in the above equations. In fact, as mentioned previously, the normal modes are not well defined in the presence of lossy media like the graphene sheet. Instead, one has to resort to so-called quasi-normal mode formalisms, of which MQED is one example, but others exist³⁹. In practice, the MQED formalism allows us to obtain the kernel in terms of the DGF. Focusing on the dominant $\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}$ term from the interaction Hamiltonian for simplicity, it can be shown that (see Supplementary note 1),

$$\sum_{\nu_q} | \langle f_q \otimes \nu_q | \hat{H}_{int} | i \otimes 0 \rangle |^2 \delta(\omega - \omega_{\nu_q}) = \iint d\mathbf{r} d\mathbf{r}' p_i^*(\mathbf{r}) \text{Im} \hat{G}_{ii}(\mathbf{r}, \mathbf{r}', \omega) p_f(\mathbf{r}'), \tag{7}$$

where we have defined the transition current $p_f(\mathbf{r}) = [\psi_f^*(\mathbf{r}) \hat{\mathbf{p}} \psi_i(\mathbf{r})]$. By virtue of the DGF, an expansion in normal modes

becomes superfluous, and this resolves the problem of ill-defined normal modes. However, the replacement of the normal modes by the DGF is not only a technicality; it also leads to a somewhat different picture of the quantum light-matter interactions itself. The left hand side of Eq. (7) describes coupling of electronic states $|i\rangle$ and $|f_q\rangle$ via the manifold of degenerate EM modes at frequency ω . In this picture, the EM field driving the electronic transition is simply the field of the EM vacuum, i.e., $|\hat{\mathbf{A}}|_{0EM}\rangle$. In particular, it is independent of the state of the electrons. A quite different picture is suggested by the right-hand side. Recalling that $i\vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega)$ is the electric field at \mathbf{r} created by a point current source at \mathbf{r}' , we can interpret the double integral as the EM self-energy of the current distribution $\langle f_q | \hat{\mathbf{p}}(\mathbf{r}') | i \rangle$ (we must use the imaginary part of the DGF to obtain the real part of the energy). We can further notice that $\langle f_q | \hat{\mathbf{p}}(\mathbf{r}') | i \rangle$ is a measure of the quantum fluctuations of $\hat{\mathbf{p}}(\mathbf{r}')$ in the initial and final states (the matrix element vanishes if one of the states is an eigenstate of $\hat{\mathbf{p}}(\mathbf{r}')$). Therefore, in this picture, the electronic transition is driven by an EM field whose source is the current density of the transition itself.

Limitations of MQED-DFT. Although the presented framework is quite general it still has its limitations. Most importantly, the framework is limited to the weak and strong coupling regimes, but not suited for treating cases of ultra-strong coupling. The first reason is that the framework is built around the Wigner–Weisskopf model which means that we are restricted to a one-photon Hilbert space and that we neglect the counter-rotating terms. Both of these assumptions are expected to be violated in the ultra-strong coupling regime⁴⁰. Secondly, the fact that the wave functions and energies of the active matter system (the TMD subbands in the present case) are calculated independently of the vacuum fluctuations means that the states of the system might be misrepresented in situations where there is significant dressing of the electronic states by the vacuum field. In such cases, one should resort to a fully self-consistent method such as e.g., QEDFT. We stress, however, that such situations are only expected to arise in the case of ultra-strong coupling, which is very difficult to reach without external pumping of the EM field. Within the typical regimes of weak to strong coupling, in particular this includes systems in equilibrium, i.e., non-driven systems, we thus expect the MQED-DFT method to be valid and

provide an efficient alternative to more full-fledged ab-initio treatments.

To verify that the dressing of the subband states considered in this work is indeed small, we calculated the Lamb-shifts of the electronic energies and found that they never exceed 0.4% of the transition energy, (see Supplementary note 5). We thus conclude that the dressing of the electronic states by the EM vacuum field is negligible, and that a non-self consistent treatment of the light-matter interaction is justified for the systems studied in the present work.

Next we discuss the gauge fixing condition $\nabla \cdot (\vec{\epsilon} \vec{A}) = 0$. This condition is met since the MQED vector potential is defined in the absence of the intersubband transition and we assume that no other longitudinal sources exist inside of the TMD layer. The condition also follows directly from the form of the DGF presented in the methods section. This extra condition completely fixes the gauge and leaves us with a generalized version of the Coulomb gauge common in quantum optics. We note that the assumption that no other longitudinal sources exist is justified inside the TMD. At the TMD/graphene and TMD/conductor interfaces this condition is violated and the field develops a longitudinal component. However, the subband wave functions are well confined within the TMD region and have only negligible weights at the interfaces.

Next, we consider limitations arising from the chosen form of the interaction Hamiltonian used in this work. As mentioned in the above, in writing Eq. (1) we neglected the diamagnetic \hat{A}^2 term. The diamagnetic current density, the \hat{A}^2 term, has been neglected since it only describes two-photon emission/absorption and emission-reabsorption processes. The two-photon processes are not captured by our one-photon Hilbert space and neglecting them is consistent with the use of the Wigner–Weisskopf model. The emission-reabsorption processes leads to an energy renormalization of the states similar to the Lamb shift. We have calculated this shift using second-order perturbation theory for a few of the considered structures and found it to be below 0.4% of the undressed transition energies and thus negligible, see Supplementary note 4 for details on the calculation. While we believe that for the situation and method considered in this work the values calculated above give a good indication of the importance of the diamagnetic term, such an approach is only valid in a perturbative or finite level approach such as the restricted Hilbert space considered in this work⁴¹. More generally, the \hat{A}^2 term becomes dominant when reaching the ultra-strong coupling regime⁴⁰ and any attempt at treating this regime would require an extension of the model to include the diamagnetic term. We further note that in any self-consistent treatment of both light and matter, e.g. QEDFT, the diamagnetic term is essential to ensure the stability of the coupled system even at weak coupling strengths⁴¹ and any such approach must therefore naturally include the \hat{A}^2 term.

Another limitation of our approach concerns the description of the EM field distribution via the DGF, which is based on a continuum model of the underlying materials, e.g. a (nonlocal) sheet conductivity for graphene and infinitely sharp interfaces between the materials. As a result, while the modes of the field resolve the mesoscopic structure of the vdW-heterostructure, the local fields due to the atomic structure are not included in the modes of the EM field.

Finally, we note that the assumption of a perfectly conducting metal substrate is an approximation. In practice, a finite conductivity would allow the field to penetrate into the metal and cause additional optical losses. While we expect these effects to be small in the spectral region considered here²⁹, they could be significant in e.g., the optical regime and can be included in the

formalism at the price of a more elaborate (numerical rather than analytical) determination of the DGF.

Decay rates from MQED-DFT. In this section, we use the MQED-DFT framework to analyse the coupling of intersubband transitions to propagating graphene plasmons in the vdW QW devices depicted in Fig. 2. Throughout the discussion we shall distinguish the *radiative* decay rate due to emission of propagating plasmons from the *total* decay rate, which will also include the incoherent coupling to the e-h continuum in the graphene. The radiative rate is obtained by setting the coupling to the e-h continuum to zero (i.e. setting $K=0$ everywhere in the (q, ω) -parameter space except in the region where the plasmon crosses the subband transition, see Fig. 3). Although the electron is emitting to different modes (for example different propagation directions), the interaction is expressed in terms of a single decay rate that incorporates all the emission channels.

In Fig. 4 we report the calculated radiative emission rates and Purcell factors for QWs of MoS₂, MoSe₂, WS₂, and WSe₂ as a function of the number of layers ($N=2, \dots, 5$) and the graphene doping level ($E_F=0.1-0.7$ eV). In general, the four TMDs show very similar behavior with the (small) differences mainly originating from variations in the subband energy gaps. Figure 4a, b shows the maximum radiative emission rate and corresponding Purcell factor obtainable by tuning of E_F . The decay rates range from 2 to 28 THz, with the largest rates obtained for the thinnest structures. The corresponding lifetimes are at least an order of magnitude shorter than the intrinsic (electron-phonon limited) lifetimes of electrons and holes in TMDs at low temperatures^{3,26,27}. Consequently, the coupling to the graphene plasmons should indeed be the dominant mechanism governing the dynamics of the intersubband transitions, at least in the low temperatures regime. The higher rates obtained for thinner structures is somewhat counter-intuitive as the thinner structures have smaller transition dipole moments. However, this effect is dwarfed by the larger number of plasmons available at higher energies (the 2D acoustic plasmon DOS scales as $\sim\omega$) where we find the transition energies of the thinner structures, see Fig. 3. Another way to think of this is that the plasmon is more confined at the larger transition frequencies of the thinner structures leading to stronger field enhancement²⁹.

While the absolute decay rates decrease with the number of layers, N , the Purcell factors show the opposite trend. This goes against the common Q/V -scaling rule for the Purcell enhancement and can be explained as follows: The (vertical) intersubband transition energy scales as $\omega_{if}(N) \sim 1/N^2$. Within the point dipole approximation, the Purcell factor is essentially a measure of how much the photon DOS at the position of the emitter is enhanced relative to the free space value. While the 2D acoustic plasmon DOS scales as $\sim\omega$, the free space DOS scales as ω^2 . This means that even though the plasmonic DOS increases with energy, its magnitude relative to free space DOS actually decreases with energy. This argument therefore explains why the Purcell enhancement decreases with decreasing thickness. Going beyond the dipole approximation, another argument for the increase in Purcell factor with N , is the change in the matrix element due to the spatial variation of the plasmon along the QW. As we will show in Fig. 5, variations from the dipole approximation do occur in the thicker structures. However, since the deviations from the dipole approximation are relatively small, we deduce that the dominant effect leading to the larger Purcell factor in the thicker structures is the relatively larger plasmonic DOS at the intersubband transition energy as compared to its photonic counterpart.

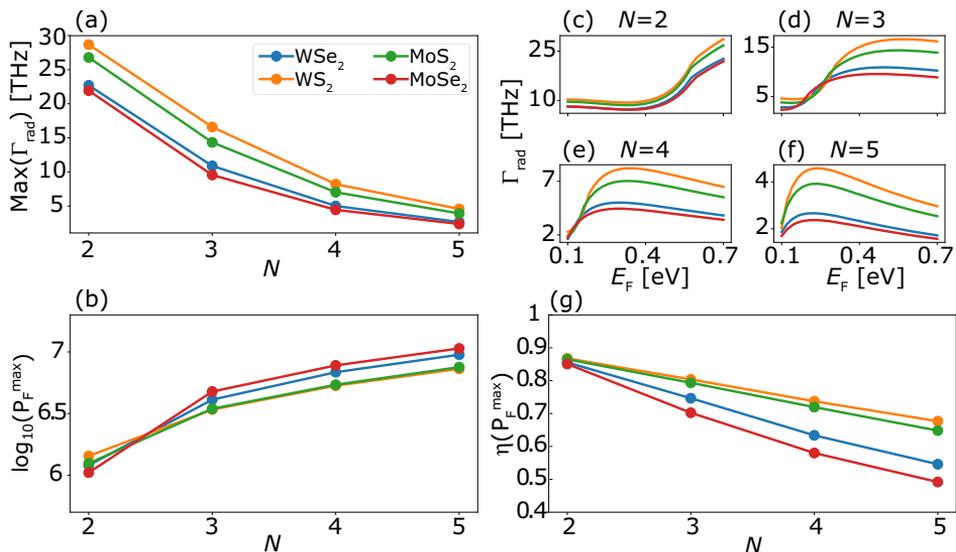


Fig. 4 The enhancement of transition rates as a function of number of layers and their composition. **a** The maximum radiative rates obtainable by adjusting the graphene Fermi level, for the different TMD materials and number of layers in the quantum well (N). **b** The Purcell factors corresponding to the maximal radiative rates in (a). Radiative emission rates as a function of graphene Fermi level for 2 (c), 3 (d), 4 (e), and 5 (f) layers. **g** Coupling efficiency, $\eta = \Gamma_{\text{rad}}/\Gamma_{\text{tot}}$, evaluated for E_F giving the maximum radiative rate. All rates were calculated using the generalized Weisskopf-Wigner MQED-DFT framework.

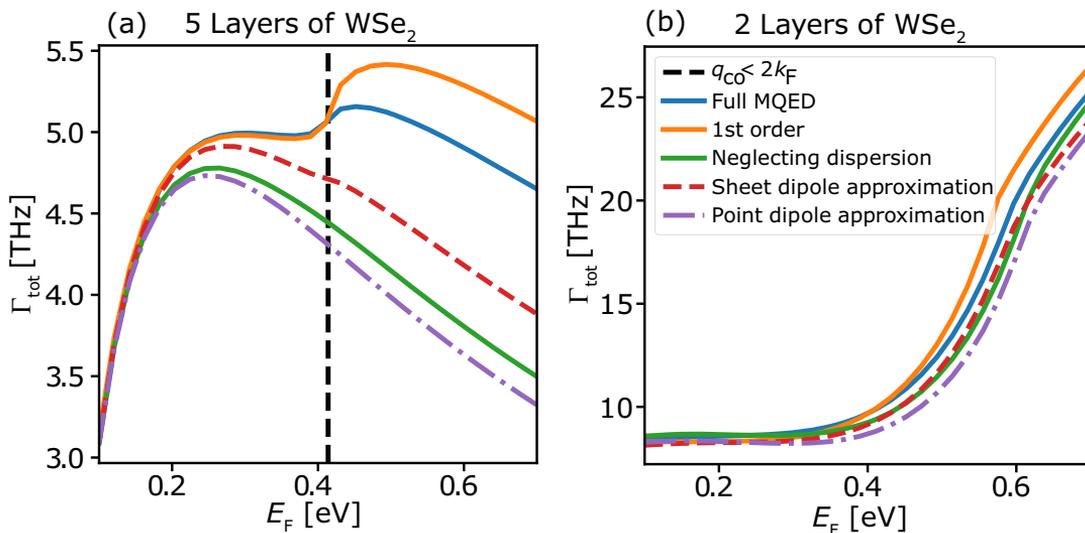


Fig. 5 Comparison of the total rates calculated using the five different approximations. The comparison is shown for 2 (a) and 5 (b) monolayers of WSe₂: The full time dependent MQED framework (blue), 1st order MQED perturbation theory (orange), the full MQED framework while neglecting the dispersion of the intersubband transition energy (green), approximating the TMD as a dipole sheet at $z_0 = d/2$ (red), and the point dipole approximation at $z_0 = d/2$ (purple). The dashed black line marks the graphene Fermi level at which horizontal transitions are opened up.

As can be seen from Fig. 4c–f, the rates show a strong dependence on the graphene doping level. This behavior is not surprising as the magnitude and dispersion of the plasmon EM field increases with the density of free carriers. This explains the initial rise of Γ_{rad} with E_F (except for the two-layer wells to which we return below). After the initial growth, the rates level-off and

even decrease as E_F is raised further. This happens because the dispersion curve of the graphene plasmon steepens as E_F is increased. This steepening of the plasmon dispersion at larger E_F results in a smaller transition momentum and thus smaller confinement of the plasmonic mode, leading to a reduction of the plasmonic DOS at the intersubband transition energy (blue circles

in Fig. 3). Within the range of E_F considered, the larger transition energies of the 2-layer stacks implies that the reduction of the plasmonic DOS due to the aforementioned dispersion steepening does not affect these thinnest structures, which explains why we do not observe a drop in the radiative rate for these systems.

Finally, we consider the radiative efficiency, $\eta = \Gamma_{\text{rad}}/\Gamma_{\text{tot}}$, evaluated for the graphene Fermi level giving the maximum Purcell enhancement, see in Fig. 4g. The efficiencies range from 0.5 to 0.9 with the highest values attained for the thinner structures when the Fermi level is chosen to maximize the radiative emission rate. This trend can again be traced to the higher subband transition energies in the thinner structures, which implies that the dispersion curve $\omega_{if}(q)$ avoids the e-h continuum, see Fig. 3. We note that since the metallic substrate is modeled as a perfect conductor there might be some additional losses arising from a real metallic substrate. However, we expect that these will be small in the infra-red region²⁹ and indeed we have found the corrections from a real model of the metal to be small relative to the rates found here. Therefore, to sum up this chapter, using our formalism we can deduce that the most effective and the fastest light-emitting TMD device should be formed by only a 2-layer TMD.

Assessment of common approximations. Having discussed the main results of the full MQED-DFT calculations, we now turn to a detailed analysis of the role played by the different elements of the method and a critical assessment of commonly used approximations. In this chapter we will focus on Fermi's golden rule, the dipole approximation, and the possible choices for modeling the electronic wave functions when going beyond the dipole approximation. The discussion will be centered around Fig. 5, which shows a comparison of the total emission rate calculated with various approximations for a 5-layer (a) and 2-layer (b) WSe₂ QW.

First-order perturbation theory. We first consider the applicability of first-order perturbation theory, i.e. Fermi's Golden rule, for calculating emission rates. In terms of the coupling kernel, the expression for the spontaneous emission rate takes the form

$$\Gamma = 2\pi \int_0^{2\pi} d\theta \int_0^{q_{\text{hz}}(\theta)} dq q K(\omega_{ifq}, q), \quad (8)$$

which follows by making the Markov approximation in Eq. (5). In the momentum integral, the $q = 0$ limit corresponds to a vertical transition in the band structure while the limit $q = q_{\text{hz}}$ denotes the wave vector at which the transition becomes horizontal. The argument of the kernel expresses the conservation of energy and in-plane momentum in all transitions. Consequently, within the 1st order picture, the emission rate is thus assumed to be the sum of the kernel values along the arc (q, ω_{ifq}) (dashed white curves in Fig. 3).

The blue curve in Fig. 5a shows the reference result of the full MQED-DFT method while the orange curve shows the 1st order result from Eq. (8). As expected, the 1st order approximation works well for low E_F where the plasmon is weak and the density of e-h pairs is low. As E_F is raised, the light-matter coupling becomes stronger, and beyond a certain threshold, marked by the vertical dashed black line, there is a clear deviation from the full-coupling reference curve. In this case, the back-action of the all the emitted light modes on the TMD electron actually prohibits a faster emission and as a result the 1st order perturbation theory overestimates the rate. The unperturbed decay rate is found from the exponential decay fit of the excited state probability decay trend (See Supplementary Figs. 4 and 5). The breakdown of the 1st order perturbation theory occurs as a result of the coupling to

low-energy horizontal e-h transitions across the graphene Fermi circle. Thus Fermi's golden rule is expected to fail when the horizontal QW transitions fall inside the e-h continuum, that is when $2k_F > q_{\text{hz}}$. For the 5-layer QW in Fig. 5a this criterion occurs at approximately $E_F = 0.4$ eV, while for the 2-layer structure in Fig. 5b it occurs outside the plotted scale.

Dipole approximations. The next approximation we investigate is the local dipole approximation. Within the dipole approximation, the graphene assisted EM vacuum field driving the emission process is assumed to vary little over the extent of the intersubband wave functions, and Eq. (7) reduces to

$$K(\omega, \mathbf{q}) \approx \hat{\mathbf{p}}_{if\mathbf{q}}^\dagger \cdot \text{Im} \vec{\mathbf{G}}(\mathbf{r}_0, \mathbf{r}_0, \omega) \cdot \hat{\mathbf{p}}_{if\mathbf{q}}, \quad (9)$$

where \mathbf{r}_0 is a point in the center of the QW and $\hat{\mathbf{p}}_{if\mathbf{q}} = \langle f_{\mathbf{q}} | \hat{\mathbf{p}} | i \rangle$ is the dipole matrix element evaluated from the DFT wave functions. Equation (9) is the well known and widely used formula for spontaneous dipole emission³⁵. Below we test this strictly local approximation against the fully nonlocal expression Eq. (7). We consider two different types of dipole approximations, namely the conventional *point-dipole* approximation and a *sheet-dipole* approximation. To model a point-dipole emitter we neglect the \mathbf{q} -dependence of the dipole matrix elements and transition frequencies and replace them on the right-hand-side of Eq. (9) by their values at the vertical ($q = 0$) transition. Unlike the point-dipole approximation, the sheet-dipole approximation takes the full \mathbf{q} -dependence of both dipole matrix elements (that is the QW wave functions) and transition frequencies into account, yet still evaluates the DGF at a single space point \mathbf{r}_0 in the center of the QW (See Supplementary S1 for more details). Thus, the sheet-dipole approximation becomes exact for EM fields with q -dependent energy dispersion and spatial variation along the in-plane direction but not in the normal direction.

The total decay rates obtained using the two dipole approximations are shown in Fig. 5 by the purple (point-dipole) and red (sheet-dipole) curves, respectively. Focusing on the 5-layer structure, we observe that both dipole approximations work well for small E_F . In this regime, the decay is dominated by the coupling to the plasmons whose electric field is long range and varies little over the extent of the QW, thus fulfilling the assumption of the dipole approximation. However, for larger values of E_F , the point-dipole approximation deviates from the full-coupling reference and underestimates the emission rate significantly. This difference arises mainly due to the neglect of the q -dispersion of the transition energies, which implies that the coupling to the low-energy e-h pairs is missed. Indeed, we obtain very similar results by setting $\omega_{if}(\mathbf{q}) = \omega_{if}(\mathbf{0})$ in the full MQED calculation (green curve). However, while the sheet dipole approximation indeed represents an improvement upon the point dipole approximation, it still falls short of the full MQED treatment. The q -dispersion of the transition energies alone is thus not enough to explain the shortcoming of the local treatment. Specifically, the coupling to the large q excitations in the e-h continuum of graphene is also underestimated in the sheet dipole approximation, due to the rapid spatial variation of their associated electric fields across the QW width, see Fig. 6d. In conclusion, any accurate description must therefore simultaneously account for the q -dispersion of the transition energies and describe the light-matter interaction beyond the dipole approximation. We note that both the neglect of q -dispersion and the dipole approximation have little influence on the rate obtained for the 2-layer QW, because the transition energy is much higher and the emission is dominated by the plasmon, see Fig. 3.

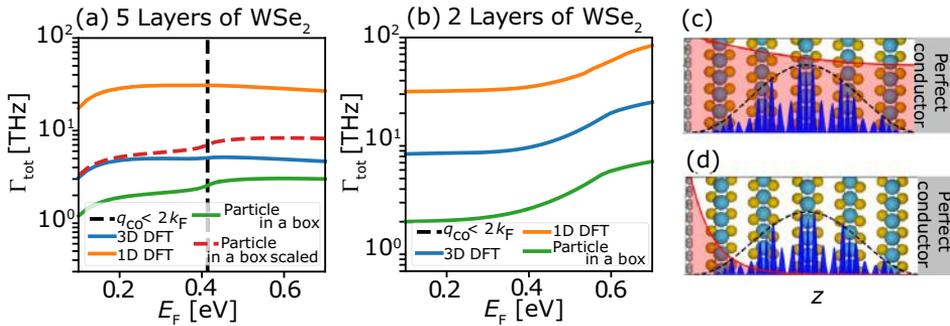


Fig. 6 Comparison of the total emission rates calculated using three different wave function approximations: particle in a box states, in-plane averaged DFT wave functions (1D DFT), and the full 3D DFT wave functions. **a, b** Comparison of the total emission rates calculated with the different wave function approximations for a 2-layer and 5-layer WSe₂ quantum well, respectively. Additionally, the dashed red line in **(a)** shows the rates calculated using the particle in a box wave functions scaled by the square of the ratio between the transition dipole moments calculated with the full 3D DFT wave functions and the particle in a box wave functions. The field enhancement profiles for the plasmon with in-plane wave vector $q = 0.2 \text{ nm}^{-1}$ **(c)** and $q = 1.21 \text{ nm}^{-1}$ **(d)**. The latter corresponds to the wave vector for horizontal transitions, q_{hz} , of the 5-layer WSe₂ QW. The simple particle in a box wave function is shown by the dashed black line and the 1D DFT wave function is shown in blue.

Electronic wave functions. The final approximation that we investigate concerns the QW wave functions. As will be shown, different reasonable wave function choices can affect the rate by an order of magnitude. In Fig. 6a, b, we compare the total rates obtained using the full MQED framework with the full 3D DFT wave function (blue) to the results obtained using model wave functions of the form, $\psi(\mathbf{r}) = \chi(z)e^{i\mathbf{k}\cdot\mathbf{r}_{\parallel}}$, where $\chi(z)$ is either the in-plane averaged DFT wave function (orange) or a “particle in a box” approximation (green). To understand the role of the wave function, it is instructive to consider the explicit form of the light-matter coupling valid for our 2D geometry,

$$K(\mathbf{q}, \omega) \approx C(\mathbf{q}, \omega) \times \left| \int_{\text{u.c.}} d\mathbf{r}_{\parallel} \int dz u_{\mathbf{f}\mathbf{q}}(\mathbf{r}) \cosh\left(\sqrt{\frac{\epsilon_{\parallel}}{\epsilon_{\perp}}} qz\right) \partial_z u_{\mathbf{i}}(\mathbf{r}) \right|^2 \quad (10)$$

In this expression u denotes the (in-plane) periodic part of the QW wave functions and $C(\mathbf{q}, \omega)$ is independent of the wave functions (it contains the density and field strength of the graphene assisted EM vacuum modes, see the Supplementary note 1).

The different wave function approximations produce two types of errors. The first type of error is a simple error in the effective dipole. This error occurs when the q of all of the relevant excitations fulfill the condition $q \ll 1/d$ (d is the QW width) such that the field is approximately constant over the width of the QW. This condition is fulfilled for the 2-layer well in Fig. 6a. In this case, the different wave functions produce different effective dipoles but with no qualitative effect on the rate as function of E_F . As shown in the Supplementary note 2, the three curves in Fig. 6a are to a very good approximation related by a simple rescaling by the square of their respective transition dipole moments.

The second type of error occurs when the wave functions are coupling to large q excitations. In this case, the coupling becomes sensitive to the spatial shape of the wave functions. This is for example the case for the 5-layer structure, where the rate from the particle-in-a-box wave function deviates qualitatively from the DFT results for higher values of E_F . This can be clearly seen if we compare the rates calculated using the full 3D DFT wave functions to the rates calculated using the particle in a box states rescaled by the ratio of the effective dipole moments, blue and dashed red lines in Fig. 6a respectively. Specifically, the qualitative deviation happens when $2k_F$ becomes comparable to q_{hz} as

indicated by the vertical dashed line. At this point, strong quenching due to the coupling of the horizontal transition to the low-energy e-h pairs becomes possible. Importantly, however, the strong quenching is not observed for the DFT wave functions because they are localized further inside the TMD structure (further away from the interface with the graphene) as compared to the standing waves, and therefore their overlap with the vacuum field associated with the $2k_F$ e-h pairs is significantly suppressed, see Fig. 6c.

Discussion

In this paper, we have combined MQED with DFT to obtain a computationally efficient framework for quantitatively describing quantum light-matter interactions in solid-state systems. The active quantum matter is represented by its DFT band structure and wave functions while the EM background media are described by their appropriate macroscopic linear response functions, which can be spatially nonlocal and dissipative. While the Kohn–Sham energies and wave functions from DFT are expected to provide a good approximation to the true subband states, it should be mentioned that they are subject to the limitations of the specific exchange-correlation functional employed and the Kohn–Sham scheme in general. In the future it would thus be interesting to explore how the use of other wave function schemes, for example hybrid functionals or GW self-energies, would affect emission rates and other quantum light-matter properties. The approach taken in the present work is in many ways complementary to the QEDFT formalism. While being less fundamental in nature than the QEDFT, our method avoids many of the approximations commonly made in the description of coupled light-matter systems, c.f. Fig. 1, while it remains applicable to relatively large/complex systems. Instead of taking the ab-initio route directly, we take advantage of the fact that the quantization of the light degrees of freedom in interacting, potentially lossy light-matter systems can be achieved within the MQED formalism^{14,33}. A powerful feature of the MQED framework is that it treats all light modes on an equal footing. Specifically, it achieves the simultaneous quantization of lossy—as well as undamped propagating modes. While the coupling to propagating modes is almost always the main objective, in practice the optical losses are often the performance-limiting factor. Since losses are present and often of significant practical importance in real materials, they are essential to include in a

quantitative theory. Even within the MQED formalism, and similar quantization schemes of lossy quasi-normal modes¹⁵, the light-matter interaction is usually described through the dipole approximation, i.e. the relevant electronic transition matrix elements are replaced by point dipole emitters¹⁴. However, the vdW heterostructures studied in the present work, comprise a clear example where the dipole approximation breaks down when emitting lossy optical modes. Furthermore, as shown in this work, the quantum description is essential to capture the deviations from first-order perturbation theory, which we observe when the carrier concentration in the plasmon-carrying graphene sheet becomes sufficiently high.

We have applied the combination of MQED and DFT to study the coupling of intersubband transitions in atomically thin TMD QWs to the optical modes of a doped graphene sheet. Our calculations predict radiative lifetimes below 1 ps and Purcell factors approaching 10⁷. Importantly, these lifetimes are an order of magnitude shorter than the intrinsic low temperature phonon-limited carrier lifetimes in TMDs and thus the spontaneous emission of graphene plasmons from intersubband transitions should indeed be efficient and observable. We note that Purcell factors as high as the ones reported here are unusual for extended 2D structures, and were previously only reported for geometries confined in all three dimensions^{42–44}. In contrast to the case of 3D confined cavities where the enhanced emission is due to the coupling to a distinct EM cavity mode, the strong enhancement of the 2D subband transitions considered here is enabled by the significant spatial overlap of the subband wave functions with the highly confined acoustic graphene plasmons.

Although the calculated Purcell factors are indeed very high, they are at least two orders of magnitude smaller than the conventional Q/V enhancement rule highlighting the limitations of the latter. The deviations from the Q/V rule are in part connected to the failure of both the first-order perturbation theory, and the dipole approximation in the examined structures. Our formalism allowed us to benchmark these approximations, which fail in our structure when coupling to high-momentum lossy e-h pair excitations in the graphene that increase the total coupling and produce a significant field variation across the TMD QW. Other common approximations, such as simpler wave function models can lead to decay rates differing by more than a factor of 10 depending on the details of the model. Apart from the difference in dipole moments, the finer details in the spatial shape of the wave function can affect the coupling to the high- q lossy modes in graphene resulting in a wrong qualitative behavior of e.g. the rate versus graphene doping level.

The developed MQED-DFT method is computationally compact, methodologically simple, and is not limited to a specific system geometry or material type. In addition, the MQED quantization scheme can deal with situations where the emitter is interacting with lossy media, e.g. plasmonic metal structures, doped semiconductors, or insulators with polar phonons, which can be included in macroscopic optical parameters of the EM background media. On the other hand, the non-self consistent treatment of the light-matter interaction limits the method to the weak to strong coupling regimes where the dressing of the electronic states by the EM field can be neglected. In the ultra-strong coupling regime, the EM field dresses the electronic states, and one must resort to fully self-consistent treatments such as QEDFT. Alternatively, our current theory could be extended using the Power-Zienau-Woolley gauge that overcome several of our approximations when using an electronic multi-polar term⁴⁵. However, in the most typical cases of weak-strong coupling, the method developed here provides an efficient and practical alternative to methods like QEDFT, which is applicable to realistic setups in quantum nano-photonics.

Methods

DFT calculations. The wave functions of the TMD stacks used to evaluate the light-matter coupling kernel Eq. (5) were obtained from DFT employing the PBE exchange-correlation (xc) functional⁴⁶ and a plane wave cut-off of 800 eV. To resolve the subband structure a Γ -point centered 60×60 k -point grid was employed. To avoid spurious interaction between periodically repeated supercells, a 20 Å vacuum was included perpendicular to the TMD stack. The relaxed atomic structures of the four TMD monolayers were obtained from the C2DB database⁴⁷. Next, N -copies of the monolayer were stacked in the 2H-configuration and the interlayer distance was relaxed using the C09-vdW xc-functional⁴⁸ while keeping the intralayer structure fixed. All calculations were performed with the GPAW code⁴⁹. The graphene sheet and the perfect mirror were both placed at a distance of $(d_{\text{Gra}} + d_{\text{TMD}})/2$ from the outer most atomic layer of the TMD stack, with d_{Gra} and d_{TMD} being the interlayer distances in graphene and the TMD respectively.

1D wave function approximations. In addition to the full 3D DFT generated wave functions, this work considers two different wave function approximations: Particle in a box states, and the in-plane average of the DFT states. The particle in a box states are given as,

$$\phi_n(\mathbf{r}) = \left(\frac{2}{d}\right)^{1/2} \sin\left(\frac{n\pi z}{d}\right), \tag{11}$$

where n is the subband index and d is the width of the TMD layer. The 1D DFT states are found when approximating $\phi_n(\mathbf{r})$ with the in-plane average of the full 3D wave functions.

The Dyadic Green's function. In this section we will find the Dyadic Green's function(DGF) for the structure discussed in the main text. The DGF, $\vec{\vec{G}}(\mathbf{r}, \mathbf{r}', \omega)$, provides a general description of the EM environment of a structure, and it plays a fundamental role in determining the interaction between matter and the modes of the EM field. $\vec{\vec{G}}(\mathbf{r}, \mathbf{r}', \omega)$ is defined as propagating a current density, $\mathbf{j}(\mathbf{r}', \omega)$, at a source points \mathbf{r}' , onto the corresponding electric field at point \mathbf{r} according to:

$$\mathbf{E}(\mathbf{r}) = i\omega\mu_0\mathbf{j}(\mathbf{r}) \int_V \vec{\vec{G}}(\mathbf{r}, \mathbf{r}')\mathbf{j}(\mathbf{r}')dV' \tag{12}$$

This translates into the DGF itself being defined as the electric field \mathbf{E} at point \mathbf{r} , generated by a radiating dipole, \mathbf{p} , located at the source point \mathbf{r}' . Therefore, the field from the radiating dipole can be expressed as, $\mathbf{E}(\mathbf{r}, \omega) = \omega^2\mu_0\mathbf{j}(\mathbf{r}', \omega) \cdot \mathbf{p}$ ³⁵. Specifying to non-magnetic materials, $\mu = 1$, the equation of motion for the DGF is the Helmholtz equation with a point source term:

$$\left[\nabla \times \nabla \times - \left(\frac{\omega^2}{c^2}\right)\vec{\vec{\epsilon}}(\mathbf{r}, \omega)\right]\vec{\vec{G}}(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r} - \mathbf{r}') \tag{13}$$

where ω is the angular frequency and c is the speed of light. In this work, we consider the electric permittivity tensor, $\vec{\vec{\epsilon}} = \text{diag}(\epsilon_{\parallel}, \epsilon_{\perp}, \epsilon_{\perp})$ to account for anisotropic dielectric properties of the TMD stacks.

Because of the layered nature of the geometry, the in-plane translational invariance of the system enables the DGF to be expressed as a mode expansion over the in-plane momentum, $\mathbf{q} = (q_x, q_y)$ in the following way^{35,50,51}:

$$\vec{\vec{G}}(\mathbf{r}, \mathbf{r}', \omega) = \int d\mathbf{q} e^{i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} [\vec{\vec{N}}(\mathbf{q}, z, z') + \vec{\vec{M}}(\mathbf{q}, z, z')] \tag{14}$$

where $\vec{\vec{N}}(\mathbf{q}, z, z')$ and $\vec{\vec{M}}(\mathbf{q}, z, z')$ describes the spectral response of the TE and TM polarizations respectively. However, since the cavity plasmon mode is TM polarized, we retain only the TM polarized part of the DGF.

To further derive the explicit form of the Green's function in the multilayered structure, we use express $\vec{\vec{M}}(\mathbf{q}, z, z')$ with the reflection coefficients from each side of the TMD^{50,51}. It can be shown that the components of the TM response are

$$\text{Im} \left[M_{\hat{q}\hat{q}}(\mathbf{q}, z, z') \right] = \frac{k_z^2 c^2}{4\pi^2 \omega^2 \epsilon_{\parallel}} \sinh(k_z z) \sinh(k_z z') \text{Im} \left[\frac{r_g e^{-2k_z d} - 1}{r_g e^{-2k_z d} + 1} \right] \tag{15}$$

$$\text{Im} \left[M_{z\hat{q}}(\mathbf{q}, z, z') \right] = \frac{i q c^2}{4\pi^2 \omega^2 \epsilon_{\perp}} \cosh(k_z z) \sinh(k_z z') \text{Im} \left[\frac{r_g e^{-2k_z d} - 1}{r_g e^{-2k_z d} + 1} \right] \tag{16}$$

$$\text{Im} \left[M_{\hat{q}z}(\mathbf{q}, z, z') \right] = \frac{-i q c^2}{4\pi^2 \omega^2 \epsilon_{\perp}} \sinh(k_z z) \cosh(k_z z') \text{Im} \left[\frac{r_g e^{-2k_z d} - 1}{r_g e^{-2k_z d} + 1} \right] \tag{17}$$

$$\text{Im} \left[M_{zz}(\mathbf{q}, z, z') \right] = \frac{q^2 c^2 \epsilon_{\parallel}}{4\pi^2 \omega^2 k_z^2 \epsilon_{\perp}^2} \cosh(k_z z) \cosh(k_z z') \text{Im} \left[\frac{r_g e^{-2k_z d} - 1}{r_g e^{-2k_z d} + 1} \right] \tag{18}$$

where $k_z = \sqrt{\frac{\epsilon_{\perp}}{\epsilon_{\parallel}} q^2 - \epsilon_{\parallel} \frac{\omega^2}{c^2}}$ is the wave vector in the out of plane direction, and $\hat{\mathbf{q}}$ is an in plane spatial unit vector pointing in the direction of \mathbf{q} . In addition, we have assumed perfect reflection from the TMD to the metal and used the quasistatic reflection coefficient $r_g = \frac{\sqrt{\epsilon_{\parallel}} \epsilon_{\perp} - 1 - \frac{\beta \epsilon_{\parallel}(\omega)}{\omega}}{\sqrt{\epsilon_{\parallel}} \epsilon_{\perp} + 1 + \frac{\beta \epsilon_{\parallel}(\omega)}{\omega}}$ as the TMD to Graphene reflection coefficient (see

Supplementary note 3 for derivation) with $\sigma(q, \omega)$ as the nonlocal graphene surface conductivity which includes plasmonic losses and Landau damping³⁸. We note that the Green's function expressions can be derived directly from the electric fields from a dipole oriented in the \hat{q} and \hat{z} directions. Finally, for our further calculations we will also use the identity, $\text{Im}[\vec{G}(\mathbf{r}, \mathbf{r}', \omega)] = \int d\mathbf{q} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} \text{Im}[\vec{M}(\mathbf{q}, z, z', \omega)]$, which describes the entire TM optical response that the TMD experiences, whether we want to consider the TMD emitter as local, nonlocal or even a dipole emitter.

Background permittivity

TMD permittivity. The permittivities of the TMD's were calculated both in-plane and out-of-plane using the GPAW code⁴⁹ and they are consistent with the values found in the literature⁵². The permittivities are chosen not to include the inter-subband transitions as to avoid double counting and are thus frequency independent. The values used are shown in Table 1.

Nonlocal graphene conductivity. In order for the light-matter interactions to take into account all possible optical excitations in the graphene, including the high-momentum e-h continuum, there is a necessity for a nonlocal graphene conductivity as opposed to a simple Drude model. This approach include any possible optical response in the graphene, within the random phase approximation. In this chapter, we will explain how to retrieve the nonlocal graphene surface conductivity in a similar manner as found in Peres and Goncalves³⁸. We note that treating the graphene as a 3D material with a finite thickness (with a bulk conductivity instead of a surface conductivity) should not affect the graphene plasmonic properties. First, the graphene conductivity, $\sigma(q, \omega)$, is defined in the following way:

$$\sigma(q, \omega) = \frac{4i\sigma_0\hbar\omega}{q^2} \chi_I\left(\frac{q}{k_F}, \frac{\hbar\omega}{E_F}\right) \quad (19)$$

with $\sigma_0 = \frac{e^2}{4\pi}$, and χ_I being the susceptibility of normalized variables, defined as:

$$\chi_I(x, y) = \frac{\left(1 + i\frac{\Gamma}{yE_F}\right)\chi_g(x, y + i\frac{\Gamma}{E_F})}{1 + i\frac{\Gamma}{yE_F}\chi_g(x, y + i\frac{\Gamma}{E_F})/\chi_g(x, 0)} \quad (20)$$

with Γ as the graphene damping parameter. This work uses the value of Γ to be 16 meV which is consistent with choice by Dias et al.⁵³. Finally $\chi_g(x, y)$ is defined according to the regions in Fig. 7, that describe the regions of different optical responses in the graphene. For more readable expressions, we define the following auxiliary functions, $F(\xi) = \xi\sqrt{\xi^2 - 1} - \text{arc cos h}(\xi)$, and, $C(\xi) = x\sqrt{1 - \xi^2} - \text{arc cos}(\xi)$. Furthermore, $x = q/k_F$ is the normalized momentum and $y = \hbar\omega/E_F$ is the normalized energy.

The different regions of the responses are:

Region 1A. Region 1A is defined as the part of the (q, ω) plane where $x > \text{Re}[y]$ and $\text{Re}[y] + x < 2$. In this region the graphene electrons stay in the same band while changing largely their momentum. The expression for the nonlocal susceptibility is:

$$\chi_g^{1A}(x, y) = \frac{-2k_F}{\pi\hbar v_F} + i\frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{x^2 - y^2}} \left[F\left(\frac{2-y}{x}\right) - F\left(\frac{2+y}{x}\right) \right] \quad (21)$$

Region 2A. Region 2A is defined by $x - \text{Re}[y] < 2$, $x > \text{Re}[y]$ and $x + \text{Re}[y] > 2$. In this region the graphene electrons change their band while also changing largely their momentum. The expression for the susceptibility is:

$$\chi_g^{2A}(x, y) = \frac{-2k_F}{\pi\hbar v_F} - \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{x^2 - y^2}} C\left(\frac{y-2}{x}\right) - i\frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{x^2 - y^2}} F\left(\frac{2+y}{x}\right) \quad (22)$$

Region 3A. Region 3A is defined by $x - 2 > \text{Re}[y]$. In this region there is no direct transition for the graphene and thus the optical response is weak. The expression for the susceptibility is:

$$\chi_g^{3A}(x, y) = \frac{-2k_F}{\pi\hbar v_F} + \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{x^2 - y^2}} \left[C\left(\frac{2+y}{x}\right) - C\left(\frac{y-2}{x}\right) \right] \quad (23)$$

Region 1B. Region 1B is defined by $\text{Re}[y] > x$ and $\text{Re}[y] + x < 2$. This region enables the propagating graphene plasmons while inhibiting other e-h excitations. The expression for the susceptibility is:

$$\chi_g^{1B}(x, y) = \frac{-2k_F}{\pi\hbar v_F} + \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{y^2 - x^2}} \left[F\left(\frac{2+y}{x}\right) - F\left(\frac{2-y}{x}\right) \right] \quad (24)$$

Region 2B. Region 2B is defined by $\text{Re}[y] > x$, $\text{Re}[y] + x > 2$ and $\text{Re}[y] - x < 2$. This region also enables the propagating graphene plasmons but includes other e-h excitations. The expression for the susceptibility is:

$$\chi_g^{2B}(x, y) = \frac{-2k_F}{\pi\hbar v_F} + \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{y^2 - x^2}} F\left(\frac{2+y}{x}\right) + \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{y^2 - x^2}} C\left(\frac{2-y}{x}\right) \quad (25)$$

Table 1 The TMD material permittivities that were used in our simulations.

Material	In-plane static dielectric constant	Out-of-plane static dielectric constant
MoS ₂	16.1	7.2
WS ₂	14.5	6.6
MoSe ₂	17.3	8.3
WSe ₂	15.7	7.8

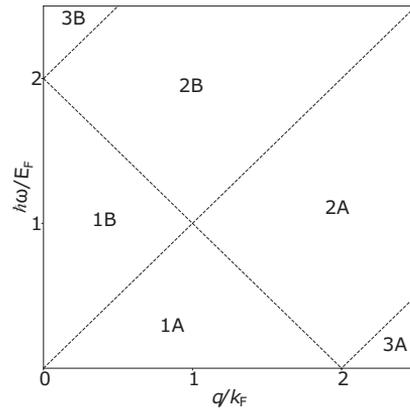


Fig. 7 Regions for the calculation of the graphene susceptibility. Regions in $(q/k_F, \omega/E_F)$ -space in which the graphene susceptibility is to be calculated.

Region 3B. Region 3B is defined as $\text{Re}[y] - x > 2$ and it captures all of the high energy interband transitions. The expression for the susceptibility is:

$$\chi_g^{3B}(x, y) = \frac{-2k_F}{\pi\hbar v_F} + \frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{y^2 - x^2}} \left[F\left(\frac{2+y}{x}\right) - F\left(\frac{2-y}{x}\right) \right] - i\frac{k_F}{\hbar v_F} \frac{x^2}{4\pi\sqrt{y^2 - x^2}} \quad (26)$$

Data availability

The data that support the plots and findings of this paper are available from M.K.S. (markas@dtu.dk) upon reasonable request.

Received: 9 August 2020; Accepted: 8 April 2021; Published online: 13 May 2021

References

- Geim, A. K. & Grigorieva, I. Van der Waals heterostructures. *Nature* **499**, 419–425 (2013).
- Novoselov, K. S., Mishchenko, A., Carvalho, A. & Castro Neto, A. H. 2D materials and van der Waals heterostructures. *Science* **353**, 6298 (2016)
- Schmidt, P., Vialla, F. & Latini, S. Nano-imaging of intersubband transitions in van der Waals quantum wells. *Nature Nanotech* **13**, 1035–1045 (2018).
- Basov, D. N., Fogler, M. M. & García de Abajo, F. J. Polaritons in van der Waals materials. *Science* **354**, 6309 (2016)
- Jablun, M., Buljan, H. & Soljačić, M. Plasmonics in graphene at infrared frequencies. *Phys. Rev. B* **80**, 245435 (2009).
- Koppens, F. H., Chang, D. E. & García de Abajo, F. J. Graphene plasmonics: a platform for strong light-matter interactions. *Nano Lett.* **11**, 3370–3377 (2011).
- Rivera, N., Kaminer, I., Zhen, B., Joannopoulos, J. D. & Soljačić, M. Shrinking light to allow forbidden transitions on the atomic scale. *Science* **353**, 263–269 (2016).
- Kurman, Y. & Kaminer, I. Tunable bandgap renormalization by nonlocal ultra-strong coupling in nanophotonics. *Nature Physics* **16**, 868–874 (2020).

9. Andersen, M. L., Stobbe, S., Sørensen, A. S. & Lodahl, P. Strongly modified plasmon–matter interaction with mesoscopic quantum emitters. *Nat. Phys.* **7**, 215–218 (2011).
10. Takase, M. et al. Selection-rule breakdown in plasmon-induced electronic excitation of an isolated single-walled carbon nanotube. *Nat. Photonics* **7**, 550–554 (2013).
11. Latini, S., Ronca, E., De Giovannini, U., HÄbener, H. & Rubio, A. Cavity control of excitons in two-dimensional materials. *Nano Lett.* **19**, 3473–3479 (2019).
12. Törmä, P. & Barnes, W. L. Strong coupling between surface plasmon polaritons and emitters: a review. *Rep. Prog. Phys.* **78**, 013901 (2014).
13. Reithmaier, J. P. et al. Strong coupling in a single quantum dot–semiconductor microcavity system. *Nature* **432**, 197–200 (2004).
14. Scheel, S. & Buhmann, S. Y. Macroscopic QED-concepts and applications. *arXiv preprint arXiv:0902.3586* (2009).
15. Van Vlack, C., Kristensen, P. T. & Hughes, S. Spontaneous emission spectra and quantum light-matter interactions from a strongly coupled quantum dot metal-nanoparticle system. *Phys. Rev. B* **85**, 075303 (2012).
16. Raza, S., Bozhevolnyi, S. I., Wubs, M. & Mortensen, N. A. Nonlocal optical response in metallic nanostructures. *J. Phys. Condens. Matter* **27**, 183204 (2015).
17. Andersen, K., Latini, S. & Thygesen, K. S. Dielectric genome of van der Waals heterostructures. *Nano Lett.* **15**, 4616–4621 (2015).
18. Ruggenthaler, M., Tancogne-Dejean, N., Flick, J., Appel, H. & Rubio, A. From a quantum-electrodynamical light-matter description to novel spectroscopies. *Nat. Rev. Chem.* **2**, 1–16 (2018).
19. Flick, J., Ruggenthaler, M., Appel, H. & Rubio, A. Kohn–Sham approach to quantum electrodynamical density-functional theory: Exact time-dependent effective potentials in real space. *Proc. Natl Acad. Sci.* **112**, 15285–15290 (2015).
20. Ruggenthaler, M. et al. Quantum-electrodynamical density-functional theory: Bridging quantum optics and electronic-structure theory. *Phys. Rev. A* **90**, 012508 (2014).
21. Flick, J., Ruggenthaler, M., Appel, H. & Rubio, A. Atoms and molecules in cavities, from weak to strong coupling in quantum-electrodynamics (QED) chemistry. *Proc. Natl Acad. Sci.* **114**, 3026–3034 (2017).
22. Wang, D. S., Neuman, T., Flick, J. & Narang, P. Weak-to-strong light-matter coupling and dissipative dynamics from first principles. *arXiv preprint arXiv:2002.10461*. (2020).
23. Tokatly, I. V. Time-dependent density functional theory for many-electron systems interacting with cavity photons. *Phys. Rev. Lett.* **110**, 233001 (2013).
24. Neuman, T., Esteban, R., Casanova, D., García-Vidal, F. J. & Aizpurua, J. Coupling of molecular emitters and plasmonic cavities beyond the point-dipole approximation. *Nano Lett.* **18**, 2358–2364 (2018).
25. Jöstäd, R., Ruggenthaler, M., Oliveira, M. J., Rubio, A. & Appel, H. Light-matter interactions within the Ehrenfest–Maxwell–Pauli–Kohn–Sham framework: fundamentals, implementation, and nano-optical applications. *Adv. Phys.* **68**, 225–333 (2019).
26. Kaasbjerg, K., Thygesen, K. S. & Jauho, A.-P. Acoustic phonon limited mobility in two-dimensional semiconductors: deformation potential and piezoelectric scattering in monolayer MoS₂ from first principles. *Phys. Rev. B* **87**, 235312 (2013).
27. Hinsche, N. F. et al. Spin-dependent electron-phonon coupling in the valence band of single-layer WS₂. *Phys. Rev. B* **96**, 121402 (2017).
28. Dias, E. J. C. et al. Probing nonlocal effects in metals with graphene plasmons. *Phys. Rev. B* **97**, 245405 (2018).
29. Iranzo, D. A. Probing the ultimate plasmon confinement limits with a van der Waals heterostructure. *Science* **360**, 291–295 (2018).
30. Lundeberg, M. B. Tuning quantum nonlocal effects in graphene plasmonics. *Science* **357**, 187–191 (2017).
31. Vogel, W. & Welsch, D.-G. *Quantum optics* (John Wiley, Sons, 2006).
32. Hatfield, B. *Quantum field theory of point particles and strings* (CRC Press, 2018).
33. Glauber, R. J. & Lewenstein, M. Quantum optics of dielectric media. *Phys. Rev. A* **43**, 467–491 (1991).
34. Raabe, C., Scheel, S. & Welsch, D.-G. Unified approach to QED in arbitrary linear media. *Phys. Rev. A* **75**, 053813 (2007).
35. Novotny, L. & Hecht, B. *Principles of nano-optics* (Cambridge university press, 2012).
36. Lalanne, P. et al. Quasinormal mode solvers for resonators with dispersive materials. *JOSA A* **36**, 686–704 (2019).
37. Dreizler, R. M. & Gross, E. K. *Density functional theory: an approach to the quantum many-body problem* (Springer Science, Business Media, 2012).
38. Gonçalves, P. A. D. & Peres, N. M. *An introduction to graphene plasmonics* (World Scientific, 2016).
39. Ren, J., Franke, S., Knorr, A., Richter, M. & Hughes, S. Near-field to far-field transformations of optical quasinormal modes and efficient calculation of quantized quasinormal modes for open cavities and plasmonic resonators. *Phys. Rev. B* **101**, 205402 (2020).
40. Kockum, A. F., Miranowicz, A., De Liberato, S., Savasta, S. & Nori, F. Ultrastrong coupling between light and matter. *Nat. Rev. Phys.* **1**, 19–40 (2019).
41. Schafër, C., Ruggenthaler, M., Rokaj, V. & Rubio, A. Relevance of the quadratic diamagnetic and self-polarization terms in cavity quantum electrodynamics. *ACS Photonics* **7**, 975–990 (2020).
42. Srinivasan, K. & Painter, O. Linear and nonlinear optical spectroscopy of a strongly coupled microdisk quantum dot system. *Nature* **450**, 862–865 (2007).
43. Hennessy, K. et al. Quantum nature of a strongly coupled single quantum dot–cavity system. *Nature* **445**, 896–899 (2007).
44. Loo, V. et al. Optical nonlinearity for few-photon pulses on a quantum dot-pillar cavity device. *Phys. Rev. Lett.* **109**, 166806 (2012).
45. Feist, J., Fernández-Domínguez, A. I. & García-Vidal, F. J. Macroscopic QED for quantum nanophotonics: emitter-centered modes as a minimal basis for multi-emitter problems. *arXiv preprint arXiv:2008.02106* (2020).
46. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
47. Haastrup, S. et al. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Materials* **5**, 042002 (2018).
48. Cooper, V. R. Van der Waals density functional: An appropriate exchange functional. *Phys. Rev. B* **81**, 161104 (2010).
49. Enkovaara, J. E. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys. Condens. Matter* **22**, 253202 (2010).
50. Chew, W. C. *Waves and fields in inhomogeneous media* (IEEE press, 1995).
51. Wasey, J. & Barnes, W. Efficiency of spontaneous emission from planar microcavities. *J. Modern Optics* **47**, 725–741 (2000).
52. Laturia, A., Van de Put, M. L. & Vandenbergh, W. G. Dielectric properties of hexagonal boron nitride and transition metal dichalcogenides: from monolayer to bulk. *npj 2D Mater. Appl.* **2**, 1–7 (2018).
53. Dias, E. J. C. et al. Probing nonlocal effects in metals with graphene plasmons. *Phys. Rev. B* **97**, 245405 (2018).

Acknowledgements

All authors thank Nicholas Rivera for the fruitful discussions. K.S.T. acknowledges support from the Center for Nanostructured Graphene (CNG) under the Danish National Research Foundation (project DNRF103) and from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant No. 773122, LIMA). I.K. acknowledges the support by the Azrieli Faculty Fellowship, by the GIF Young Scientists’ Program and from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant Grant No. 851780, and from the Israel Science Foundation grants 3334/19 and 831/19. F.H.L.K. acknowledges financial support from the Government of Catalonia through the SGR grant, and from the Spanish Ministry of Economy and Competitiveness, through the “Severo Ochoa” Programme for Centres of Excellence in R&D (SEV-2015-0522), and the Explora Ciencia FIS2017-91599-EXP. F.H. L.K. also acknowledges support by Fundació Cellex Barcelona, Generalitat de Catalunya through the CERCA program, and the Mineco grants Plan Nacional (FIS2016-81044-P) and the Agency for Management of University and Research Grants (AGAUR) 2017 SGR 1656. Furthermore, the research leading to these results has received funding from the European Union’s Horizon 2020 under grant agreement no. 881603 (Core3) Graphene Flagship, and no. 820378 (Quantum Flagship). This work was supported by the ERC TOPONANOP under grant agreement no. 726001.

Author contributions

F.K., P.S., M.K.S., Y.K., I.K., and K.S.T. were part of the process to conceive the idea for this project. M.K.S. and Y.K. derived and implemented the MQED-DFT framework and K.S.T. and I.K. supervised the project. All authors took part in drafting of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-23012-3>.

Correspondence and requests for materials should be addressed to M.K.S.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

11.2 Paper [II]: Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators

Mark Kamper Svendsen, Hiroshi Sugimoto, Artyom Assadillayev, Daisuke Shima, Minoru Fujii, Kristian S Thygesen, Søren Raza,
"Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators,"
[Advanced Optical Materials \(2022\): 2200422.](#)

RESEARCH ARTICLE

Computational Discovery and Experimental Demonstration of Boron Phosphide Ultraviolet Nanoresonators

Mark K. Svendsen, Hiroshi Sugimoto, Artyom Assadillayev, Daisuke Shima, Minoru Fujii, Kristian S. Thygesen,* and Søren Raza*

Controlling ultraviolet light at the nanoscale using optical Mie resonances holds great promise for a diverse set of applications, such as lithography, sterilization, and biospectroscopy. Access to the ultraviolet requires materials with a high refractive index and wide band gap energy. Here, the authors systematically search for such materials by computing the frequency-dependent optical permittivity of 338 binary semiconductors and insulators from first principles, and evaluate their scattering properties using Mie theory. This analysis reveals several interesting candidate materials among which boron phosphide (BP) appears most promising. Then BP nanoparticles are prepared and it is demonstrated that they support Mie resonances at visible and ultraviolet wavelengths using both far-field optical measurements and near-field electron energy-loss spectroscopy. A laser reshaping method is also presented to realize spherical Mie-resonant BP nanoparticles. With a refractive index over three and low absorption losses in a broad spectral range spanning from the infrared to the near ultraviolet, BP is an appealing material for a broad range of applications in dielectric nanophotonics.

resonances have been realized at visible and infrared wavelengths thanks to the mature lithographic processing of suitable materials,^[4] such as silicon (Si),^[5] gallium phosphide (GaP),^[6] and titanium dioxide (TiO₂).^[7] It would be desirable to extend the operation of these materials to the ultraviolet, but their small direct band gap energies ($\lesssim 3$ eV) lead to significant absorption losses in the ultraviolet. Wide band gap materials, such as niobium pentoxide^[8] and hafnium oxide,^[9] offer transparency in the ultraviolet but at the cost of a moderate refractive index ($n \approx 2.1$ – 2.3). Diamond has been theoretically suggested as a potential material,^[10,11] but comes with significant nanofabrication challenges.^[12] The scarcity of available high-index materials with wide band gap energies calls for the identification of new materials which can advance the rich optical properties of Mie resonances observed in the visible to the ultraviolet.

1. Introduction

Achieving control over ultraviolet light with nanoscale materials is essential for improving surface-enhanced spectroscopies of biological molecules and enabling new ultraviolet optical components.^[1] Geometric Mie resonances supported by resonant nanoantennas made from materials that combine a high refractive index with low absorption losses offer efficient and tunable manipulation of the near- and far-field of optical waves.^[2,3] Mie

Concurrent advances in first-principles methodology and computing power have recently made it possible to design and discover new materials via high-throughput computations.^[13–17] The approach has been successfully applied in several domains, including photovoltaics, transparent conductors, and photocatalysis.^[18–20] However, to the best of our knowledge, computational discovery of new high-index materials remains largely unexplored. Relevant previous work in this direction has been limited to the static response regime^[21,22] reflecting the fact that the major materials databases so far has focused on ground state properties.

Here we use high-throughput linear response density functional theory (DFT) to screen an initial set of 2743 elementary and binary materials with the aim to identify isotropic high-index, low loss, and broad band optical materials. For the most promising materials, the computed frequency-dependent complex refractive indices are used as input for Mie scattering calculations to evaluate their optical performance. In addition to the already known high-index materials we identify several new compounds. In particular, boron phosphide (BP) offers a refractive index above three with very low absorption losses in a spectral range spanning from the infrared to the ultraviolet. We then prepare BP nanoparticles and show, by means of dark-field optical measurements and electron energy-loss spectroscopy, that they support size-dependent Mie resonances in the visible and ultraviolet. Finally, we demonstrate a laser reshaping

M. K. Svendsen, A. Assadillayev, K. S. Thygesen, S. Raza
Department of Physics
Technical University of Denmark, Fysikvej
Kongens Lyngby DK-2800, Denmark
E-mail: thygesen@fysik.dtu.dk; sraz@dtu.dk

H. Sugimoto, D. Shima, M. Fujii
Department of Electrical and Electronic Engineering
Kobe University
Rokkodai, Nada, Kobe 657-8501, Japan

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adom.202200422>.

© 2022 The Authors. Advanced Optical Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adom.202200422

method to realize spherical BP nanoparticles, which host multiple Mie resonances in agreement with full-field optical simulations. Our experimental measurements demonstrate the potential of high-index BP across a broad spectral range as well as validate the refractive index obtained from DFT calculations. Besides the discovery of BP, we believe that our high-throughput screening provide an overview of existing materials as well as a pathway for realizing new high-index materials.

2. Results

2.1. High-Throughput Screening

Our high-throughput screening procedure is illustrated in Figure 1a. We build the workflow using the Python-based Atomic Simulation Recipes^[23] framework and the MyQueue scheduling software^[24] (see detailed workflow in Note S1 and Figure S1, Supporting Information). Starting from 2743 thermodynamically stable elementary and binary materials from the Open Quantum Materials Database^[15] we extract the 1693 materials with up to ten atoms in the unit cell and relax the atomic structure using DFT with the Perdew–Burke–Ernzerhof (PBE) exchange correlation functional^[25] and the D3 correction to account for the van der Waals forces.^[26] We perform DFT ground state calculations for all of the materials to determine their electronic band gaps. After discarding the metals, we are left with 338 semiconductors for which we calculate the optical dielectric function, $\epsilon(\omega)$, within the random phase approximation (RPA) and extract the refractive index and extinction coefficient as $n(\omega) = \text{Re}\{\sqrt{\epsilon(\omega)}\}$ and $k(\omega) = \text{Im}\{\sqrt{\epsilon(\omega)}\}$, respectively. All calculations are performed with the GPAW electronic structure code.^[27,28]

Next, we classify the materials according to the anisotropy of their refractive index tensor using a cut-off of 0.05 for the fractional anisotropy (see Figure S2, Supporting Information). This leaves us with 207 isotropic material candidates, for which we show the static refractive index as a function of the direct band gap in Figure 1b. In the Supporting Information, we provide refractive indices and band gap energies for all of the 207 material candidates. The data points qualitatively follow the Moss formula^[29] (dashed line in Figure 1b); however, there are significant deviations from the general trend, which we ascribe to variations in oscillator strength and density of the transitions across the direct gap. It is well known that (semi)local functionals like the PBE employed in the present work systematically underestimates band gaps.^[30] This effect is, however, to some extent compensated by the fact that the RPA neglects the attractive electron–hole interactions and consequently underestimates the spectral weight near the band edge. As a result, refractive indices obtained with RPA@PBE are typically in good agreement with experiments in the static limit^[28] while deviations occur at higher frequencies near the band edge region (see Note S2 and Figures S3–S5, Supporting Information). In Section 2.2, we will employ more accurate and computationally expensive many-body perturbation methods for the materials that are identified as interesting based on this screening.

We now turn to a more in-depth evaluation of the performance of the discovered materials. Specifically, we use Mie

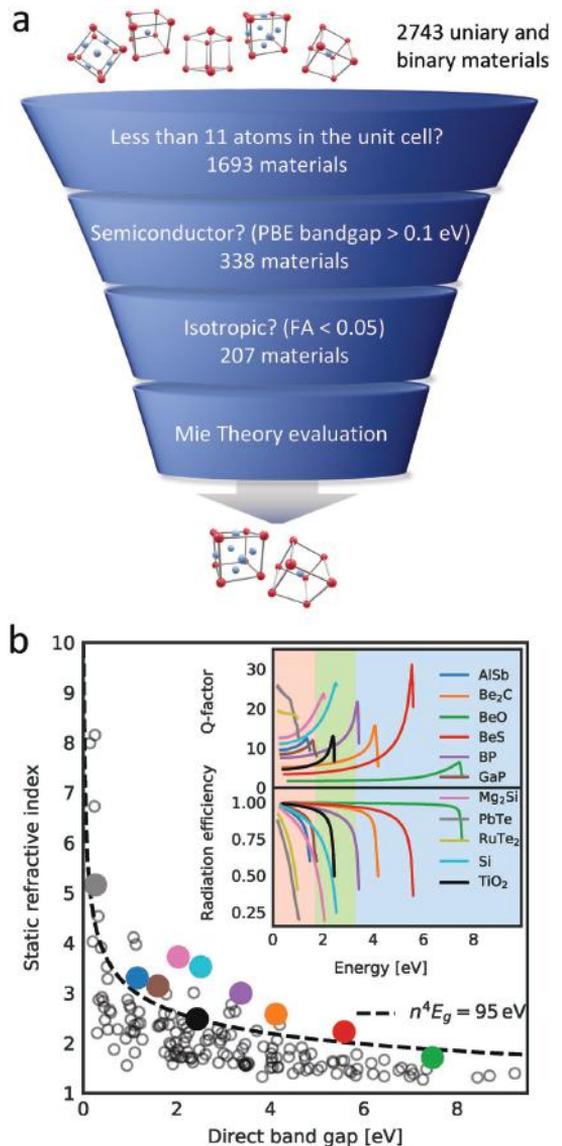


Figure 1. High-throughput materials screening. a) Schematic of the screening steps and the number of materials that survive each of them. b) The static refractive index of the isotropic materials plotted as a function of their direct band gap energies along with the Moss formula. The colored dots highlight interesting materials and the inset shows the energy-dependent Q -factor and radiation efficiency, η , of the MD resonances in those materials. The background colors represent the infrared (red), visible (green), and ultraviolet (blue) spectral regions.

theory to calculate the scattering properties of a spherical nanoparticle made from the subset of isotropic materials. We focus on the lowest-order magnetic dipole (MD) resonance of the spheres and calculate energy-dependent quality factors, Q , and

radiation efficiencies, η (see Note S1, Supporting Information).^[4] This is achieved by continuously adjusting the size of the sphere to tune the energy of the MD resonance across the infrared, visible, and ultraviolet regions. A high Q -factor is beneficial for boosting the local field enhancement of the nanoparticle, while radiation efficiency close to unity points to low absorption losses. This analysis identifies all of the commonly used materials, such as Si, TiO₂, and GaP. However, we also find a number of other highly promising materials, some of which are highlighted in the inset of Figure 1b. In particular, we identify BP, which has a refractive index exceeding that of TiO₂ and a radiation efficiency higher than both Si and TiO₂ across the entire visible part of the spectrum. For these reasons, we believe BP stands out as an overlooked material with highly desirable optical properties and we will focus on BP in the rest of the paper.

2.2. Optical Response of BP

BP crystals were successfully synthesized as early as 1957,^[31] yet experimental measurements of its refractive index are limited to a couple of data points in the visible.^[32,33] Refractive index measurements of BP thin films have also been conducted but with varying results.^[34,35]

The RPA@PBE permittivities used for the initial screening are qualitatively accurate but suffer from underestimated band gaps and missing excitonic effects. To determine the refractive index of BP with quantitative accuracy, we solve the Bethe–Salpeter equation (BSE) to obtain the permittivity using single-particle transition energies obtained from a G_0W_0 band structure calculation. The band structure calculation reveals an indirect band gap of 2.1 eV and a direct band gap of 4.41 eV, which matches experimental measurements of the band gap energies for BP.^[36,37] The square root dependence of the refractive index on the permittivity makes it crucial to converge both the real and imaginary parts of the latter. Unfortunately, the real part converges slowly with the number of bands making it impractical to obtain well converged results directly from the BSE. The problem can be alleviated by extending the imaginary part of the permittivity by an exponentially decaying tail whose weight is fixed by the f -sum rule (see Section 4), and subsequently obtain the real part via the Kramers–Kronig relation. We benchmark this approach against experimental data for the refractive index of crystalline Si and find excellent agreement (see Note S3 and Figure S6, Supporting Information).

With the f -sum rule fulfilling BSE- G_0W_0 method at hand, we are in a position to make a quantitative comparison of the refractive index of BP with some of the commonly used materials^[38,39] in the visible as well as diamond,^[40] which has been theoretically suggested for operation in the ultraviolet (Figure 2a,b). We observe that the absorption edge of BP lies significantly higher than Si, GaP, and TiO₂, while it retains a refractive index comparable to that of GaP. This suggests that BP provides low-loss operation across the entire visible spectrum and in the infrared (see Figure S7, Supporting Information). While this is also the case for TiO₂, its refractive index is significantly lower than that of BP. The higher refractive index of BP means that nanostructures can be made more compact^[41] and packed more densely for enhanced metasurface performance.^[42] Importantly, we

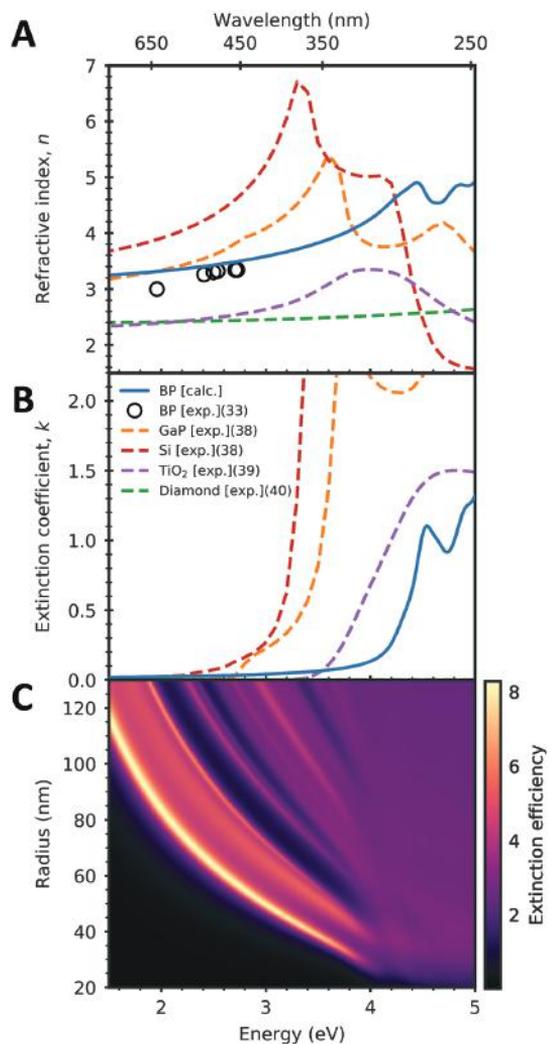


Figure 2. Optical response and Mie resonances of BP. a,b) Refractive index n and extinction coefficient k of BP calculated using the f -sum corrected BSE method compared with that of some of the commonly used dielectrics in the visible and ultraviolet spectral regions. c) Extinction efficiency map of BP spheres of varying radii calculated using Mie theory, which demonstrates that Mie resonances can be sustained in the visible and ultraviolet.

note that BP offers a high refractive index with a low extinction coefficient not only in the visible but also in the ultraviolet—a spectral region which is difficult to reach with the commonly-used materials. Diamond is transparent in the ultraviolet as well, but has a significantly lower refractive index. We illustrate the broadband performance of BP by performing extinction efficiency calculations using Mie theory for a BP sphere with varying radii, confirming that Mie resonances can be sustained across the visible and ultraviolet (Figure 2c).

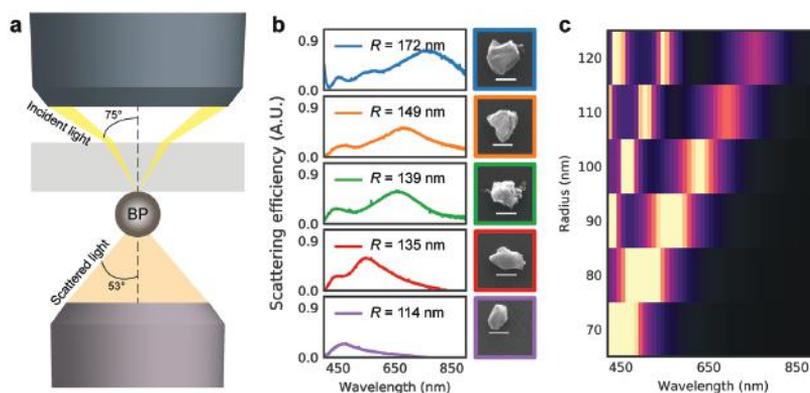


Figure 3. Dark-field scattering of BP nanoparticles. a) Schematic of the dark-field scattering spectroscopy. The nanoparticles are excited by unpolarized, white light incident at an oblique angle and the transmitted scattered light is collected. b) Scattering efficiency spectra recorded from individual BP nanoparticles of different sizes showing clear scattering peaks due to resonant interaction. The effective radius R of the nanoparticle is determined from the area of the particles, assuming a spherical shape. Scale bars: 200 nm. c) Simulated scattering efficiency map of BP nanospheres with varying particle radii.

2.3. Far- and Near-Field Characterization of BP Nanoparticles

We now turn to an experimental demonstration of the potential of BP as a Mie-resonant nanostructure. BP has been synthesized in various forms such as crystals,^[43,44] films,^[35,45–47] and nanoparticles.^[48–50] However, BP nanoparticles in the size range suitable for sustaining Mie resonances have not been reported. We prepare BP nanoparticles in the size range of a few hundred nanometers by grinding BP powder (Kojundo Chemicals) in a mortar and then dispersing it in methanol. The BP solution is subsequently dropcasted on a glass substrate.

We measure the far-field scattering efficiency of individual nanoparticles by illuminating the nanoparticles through a high-numerical-aperture dark-field objective and collecting the transmitted scattered light (see **Figure 3a**). A similar measurement setup has been used to detect Mie resonances in Si nanoparticles.^[51] The scattering spectra recorded from a series of BP nanoparticles clearly show resonant peaks, which red shift with increasing particle size (**Figure 3b**). Despite the irregular particle shapes, the scattering resonances are quite prominent and follow the trend observed in other Mie-resonant nanostructures, namely, that the smallest particle size support only the lowest-order Mie resonance while larger particle sizes also support higher-order Mie resonances.^[52] We additionally confirm the crystallinity of the nanoparticles using micro-Raman spectroscopy from individual nanoparticles (see **Figure S8**, Supporting Information). The particle radii are extracted from the scanning electron microscopy images under the assumption of a spherical shape and used as input for full-field simulations of the scattering efficiency of a BP nanosphere. The simulations account for the measurement setup as well as the glass substrate (see **Section 4**). We find that the simulations accurately reproduce both the shift in resonance wavelengths with particle size as well as the number of resonant peaks (**Figure 3c**). However, the particle sizes need to be adjusted to match the resonance wavelengths observed in the experiments. This suggests that the particle shape is better characterized as flakes

with a thickness significantly smaller than the in-plane size. Nonetheless, the distinctive, multiple scattering peaks provide strong evidence for the interpretation that these are related to geometric Mie resonances.

To gain more insight into the nature of these resonances and to access the ultraviolet spectral region, we also perform near-field characterization on similar BP nanoparticles using electron energy-loss spectroscopy (EELS). EELS is performed in a transmission electron microscope and has been employed to access near-field properties of both metallic^[53,54] and dielectric nanostructures^[55–57] as well as optical devices.^[58] The combined high spatial and spectral resolution of EELS provides unique nanoscale information on optical modes over a broad spectral range. For EELS measurements, the BP nanoparticles are deposited on a thin silicon dioxide membrane. The EELS signal recorded from a triangular-shaped BP nanoparticle (**Figure 4a**) at different beam positions is presented in **Figure 4c**. The position of the beam is directly related to the excitation efficiency of the optical modes,^[59] and thus, by judicious positioning of the beam we can selectively excite different Mie modes.^[55] When the beam is positioned in the center of nanoparticle, we observe a distinct resonance in the ultraviolet at 3.99 eV. As the beam is moved closer to the surface of the nanoparticle, two additional resonances are observed at the energies 3.57 and 2.77 eV. To identify the nature of these resonances, we simulate the EELS signal of a BP nanodisk with the same effective radius as the measured BP nanoparticle (**Figure 4b**). The thickness of the nanodisk is varied to achieve correspondence to the measured EELS resonance energies. The simulated EELS spectra for a nanodisk thickness of $h = 40$ nm at the same beam positions as in the experiments are shown in **Figure 4d**. Here, we observe that the three lowest-energy EELS peaks have the same dependence on the impact parameter of the electron beam as seen experimentally. The simulated resonance energies of all three EELS peaks are also in quantitative agreement with the experiments, albeit the lowest-energy EELS peak is slightly shifted to higher energies in the simulations. The measured EELS peaks

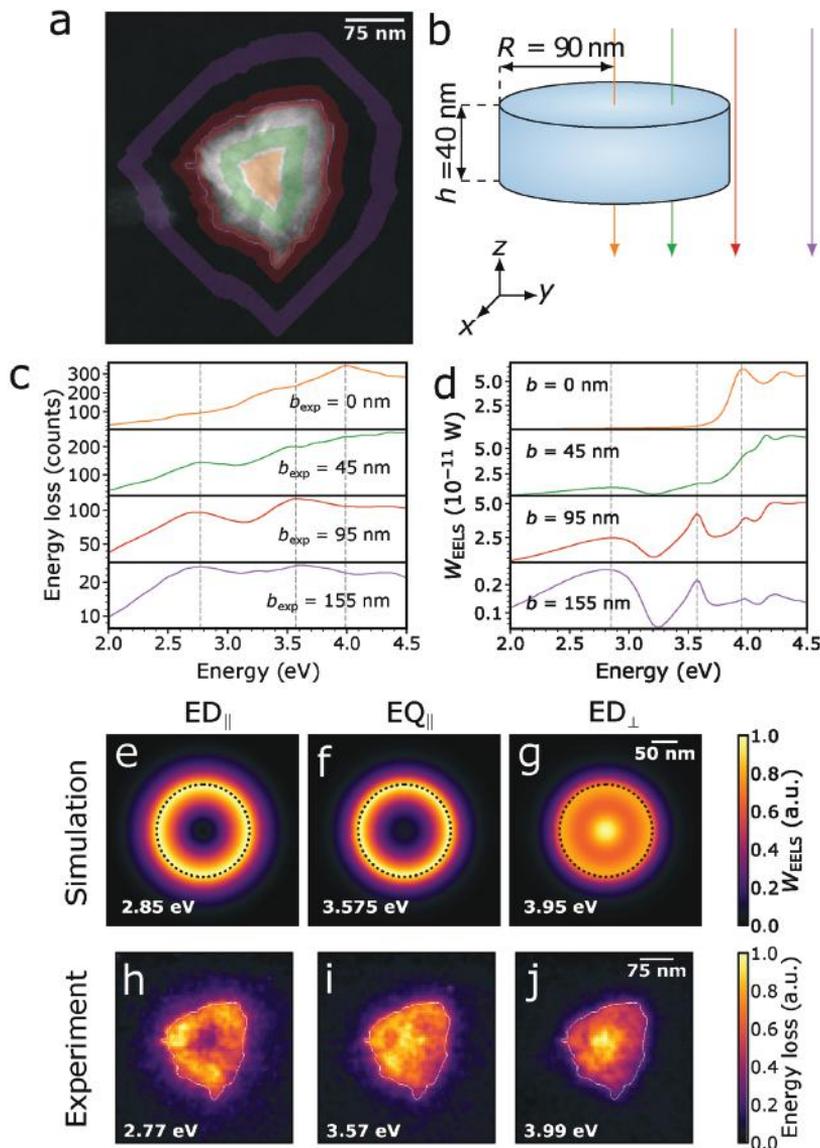


Figure 4. Electron energy-loss spectroscopy of BP nanoparticles. a,b) Experimental transmission electron microscopy image and theoretical set-up depicting the electron beam positions b used to acquire the measured and simulated EELS spectra of (c,d), respectively. The triangular-shaped BP nanoparticle is characterized by an effective radius of $R = 90$ nm. Several EELS peaks due to Mie resonances are observed. e–g) Simulated and h–j) experimental EELS intensity maps at the Mie resonance energies. The Mie modes are identified as the in-plane electric dipole (ED_{\parallel}), in-plane electric quadrupole (EQ_{\parallel}), and out-of-plane electric dipole (ED_{\perp}) using a multipole decomposition.

are broadened by the finite energy resolution of our EELS setup (see Section 4). By performing a multipole decomposition of the induced field produced by the electron beam,^[60] we identify the two lowest energy EELS peaks to be due to the Mie modes of the in-plane electric dipole ED_{\parallel} and in-plane electric quadrupole EQ_{\parallel} . The highest energy EELS peak has contributions

from both the in-plane magnetic dipole and out-of-plane electric dipole ED_{\perp} , where the latter dominates in the center of the particle (see Figure S9, Supporting Information for full decomposition). Simulated EELS intensity maps show that the electron beam couples efficiently to the in-plane modes, ED_{\parallel} and EQ_{\parallel} , for beam positions near the surface of the particle, while the

out-of-plane ED_{\perp} is excited also for beam positions in the center of the particle (Figure 4e–g). The experimental EELS intensity maps (Figure 4h–j) of these three Mie modes are in good agreement with the simulations as well as previous EELS measurements performed on Mie-resonant Si nanoparticles.^[55,56]

The dark-field scattering measurements (Figure 3) combined with the EELS measurements (Figure 4) show that the optically resonant nature of the BP nanoparticles can be identified both in their far-field and near-field responses, respectively. These resonances are present both in the visible and ultraviolet and their resonance energies can be tuned with particle size. The insight from both of these measurement techniques reveals that, despite their irregular shape, BP nanoparticles host a variety of multipolar Mie resonances, which are key attributes of low-loss high-index nanostructures.

2.4. Laser Reshaping

To alleviate the irregular shape of the as-prepared BP nanoparticles, we generate spherical BP nanoparticles by irradiating the unprocessed BP nanoparticles with a pulsed laser (see Figure 5a and Section 4). An example of a laser-processed BP nanoparticle is presented in Figure 5b, confirming that the laser processing can be used for realizing spherical-shaped BP nanoparticles without affecting the nanoparticle composition. In addition, we perform micro-Raman spectroscopy on the same particle and find that the crystallinity from the unprocessed BP nanoparticles is retained after laser reshaping (see Figure S8, Supporting Information). The dark-field scattering spectrum from the BP nanoparticle in Figure 5b is recorded and we observe multiple scattering peaks (Figure 5c). The relatively large particle radius ($R = 184$ nm) places the lowest-order Mie resonances at wavelengths longer than our measurement range, while the scattering peaks observed can be attributed to higher-order Mie resonances. The full-field simulation of a BP nanosphere with a slightly smaller radius $R = 165$ nm, where we account for the substrate and the measurement setup, shows very good agreement with the measurement. We attribute the deviation in particle radius to shape imperfections and a slight variation between the calculated and experimental refractive index of BP. Multipole decomposition reveals that the scattering peaks are due to the excitation of the magnetic quadrupole (MQ), the electric quadrupole (EQ), and radial higher-order magnetic dipole (MD^2), thereby confirming the Mie-resonant nature of BP nanoparticles. We also performed EELS measurements on a smaller laser-reshaped BP nanoparticle, where we observe Mie resonances in the ultraviolet (see Figure S10, Supporting Information).

3. Conclusion

Using a DFT-based high-throughput screening method combined with optical Mie theory of 338 dielectrics, we identify BP as a promising high-refractive-index material for ultraviolet nanooptics. We develop a new, quantitatively accurate many-body perturbation theory based methodology for calculating refractive indices and use it to reveal that BP has a high refractive index ($n > 3$) and low extinction coefficient ($k < 0.1$) up to

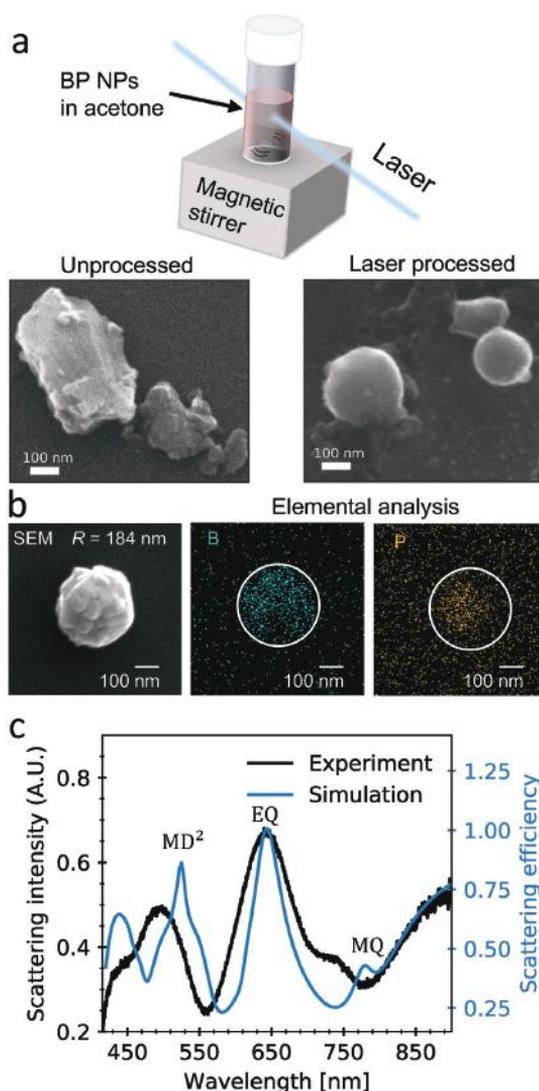


Figure 5. Laser reshaping of BP nanoparticles. a) Reshaping of the BP nanoparticles is achieved by pulsed laser irradiation of unprocessed BP nanoparticles in acetone. After the irradiation, the BP nanoparticles obtain a spherical shape. b) Scanning electron microscopy image and energy-dispersive X-ray analysis of reshaped BP nanoparticle, showing a spherical shape without change in composition. c) Scattering spectrum recorded from the BP nanoparticle in (b) and simulated spectrum for a BP nanosphere with radius $R = 165$ nm with multipole decomposition revealing the excitation of the magnetic quadrupole (MQ), electric quadrupole (EQ), and radial higher-order magnetic dipole (MD^2).

ultraviolet photon energies of 4 eV. We present an approach to fabricate BP nanoparticles as well as a laser reshaping method to generate spherical BP nanoparticles. Through dark-field optical spectroscopy and EELS measurements, we confirm the

presence of Mie resonances in BP nanoparticles across the visible and ultraviolet. Our work advances nanoscale Mie optics to the ultraviolet and may find applications in metasurface-enhanced spectroscopy of biological molecules and, more generally, in realizing metasurface optical components operating in the ultraviolet.

4. Experimental Section

Structural Relaxation: All ground- and excited state calculations were performed with the GPAW electronic structure code.^[27] The atomic structure and the unit cell of the materials were relaxed until the maximum force (stress) is below 10^{-4} eV Å⁻¹ (0.002 eV Å⁻³). The PBE functional for exchange and correlation effects, a Γ -point centered k -point grid with a density of 6.0\AA^{-3} , a 800 eV plane wave cutoff, and a Fermi-Dirac smearing of 50 meV was used. van der Waals interactions were taken into account by the D3 correction scheme.^[26]

RPA Calculations: The optical permittivity, $\epsilon(\omega)$, was calculated within the RPA using the dielectric function module in GPAW. From $\epsilon(\omega)$ the refractive index and extinction coefficient were calculated as $n(\omega) = \text{Re}\{\sqrt{\epsilon(\omega)}\}$ and $k(\omega) = \text{Im}\{\sqrt{\epsilon(\omega)}\}$, respectively. To ensure convergence across all materials a k -point grid with a high density of 20.0\AA^{-3} was employed and conduction bands up to five times the number of valence bands were included. The calculations were performed on a nonlinear frequency grid with an initial frequency spacing of 0.5 meV, a broadening of 50 meV, and a local field cutoff of 50 eV.

G_0W_0 Calculations: The G_0W_0 calculations were performed on top of the ground state calculations. To ensure converged quasi-particle gaps extrapolation of both the plane wave cut-off and the k -point resolution to infinity were performed.

BSE Calculations: The BSE calculations were performed within the Tamm–Dancoff approximation on a k -point grid with a density of 20\AA^{-3} . The calculation included all valence and conduction bands within 2.3 eV of the valence band maximum and conduction band minimum, respectively. The calculation of the screened interaction included all occupied bands and unoccupied bands up to five times the number of occupied bands, and local field effects were accounted for up to a plane wave cut-off of 50 eV. The calculation was performed on a linear 10 001 point frequency grid spanning 0–8 eV.

BSE f -Sum Rule Correction: The optical polarizability has to obey the f -sum rule

$$\int_0^\infty d\omega \omega \text{Im}\chi(\omega) = \frac{\pi}{2} \frac{n_e e^2}{m} \quad (1)$$

where n_e is the electron density, e is the elementary charge, and m is the electron mass. Since it is not obvious what to use for n_e in the PAW calculations, a different strategy was used to fix the left hand side, namely, it was obtained from an RPA calculation (which can be readily converged)

$$\int_0^\infty d\omega \omega \text{Im}\chi^{(\text{RPA})}(\omega) = \int_0^\infty d\omega \omega \text{Im}\chi^{(\text{BSE})}(\omega) \quad (2)$$

Equation (2) can be enforced if the imaginary part of the BSE polarizability was artificially extended. Denoting the largest transition energy included in the BSE calculation as ω_c , the following extension was performed

$$\text{Im}\chi^{(\text{BSE})}(\omega) = \begin{cases} \text{Im}\chi^{(\text{BSE})}(\omega), & \text{if } \omega \leq \omega_c \\ C_0 e^{-\gamma(\omega-\omega_c)}, & \text{if } \omega > \omega_c \end{cases} \quad (3)$$

The constant C_0 is used to ensure continuity and the constant γ is fixed to give the correct spectral weight as fixed by Equation (2). For benchmarks and additional details see Supporting Information.

Optical Simulations: The scattering efficiency and EELS simulations were both performed in COMSOL Multiphysics (version 5.6), which solves Maxwell's equations using finite-element modeling. For the scattering efficiency simulations in Figure 3c, a BP sphere is placed on a semi-infinite glass substrate ($n_{\text{sub}} = 1.45$) and excited by a plane wave incident from the substrate side at an oblique angle of $\theta_{\text{inc}} = 42^\circ$. The f -sum corrected refractive index was used for BP shown in Figure 2. Then a near-to-far field transformation was performed to extract the scattered far-field.^[61] The Poynting flux of the scattered far-field in the air region was integrated over a solid angle spanning an azimuthal angle of 2π and a maximum polar angle of $\theta_{\text{col}} = \arcsin(\text{NA}) = 53^\circ$ to retrieve the total scattered power collected by the collection objective (NA = 0.8). The total scattered power was normalized to the incident power and the geometrical cross sectional area of the particle to determine the scattering efficiency. To account for the unpolarized incident light in the experiments, this simulation procedure was performed for both transverse-electric and transverse-magnetic polarization of the incident wave. Finally, the scattering efficiency from both polarization states was averaged. The scattering efficiency simulation in Figure 5c followed the same steps with the only change being that the plane wave was incident from the air side at an oblique angle of $\theta_{\text{inc}} = 75^\circ$.

For the EELS calculations, the electron beam was simulated as an edge current with an amplitude of $1\text{ }\mu\text{A}$. The induced electromagnetic field was obtained by calculating the fields with and without the BP nanodisk in the simulation domain, and subsequently subtracting them. The energy loss was then calculated as the work rate done on the electron beam by the induced electromagnetic field.^[55]

Dark-Field Scattering Measurements: A custom-built inverted optical microscope was used for dark-field scattering spectroscopy of single nanoparticles (Figure 3a). For the measurements presented in Figure 3, the sample is illuminated from the top by a halogen lamp through a dark field condenser and the scattered light was collected by an objective (50 \times , NA = 0.8). For the spectrum in Figure 5c, the sample was illuminated from the bottom through a dark-field objective (50 \times , NA = 0.8) and the scattered light was collected by the same objective. To measure the spectra, scattered light was transferred to the entrance slit of a monochromator (SpectraPro-300i, Princeton Instruments) and detected by a liquid-N₂ cooled CCD (Princeton Instruments). For Raman scattering measurements, the nanoparticles were excited by a 488 nm laser (Coherent Sapphire 488–50).

EELS Measurements and Analysis: The EELS measurements were performed in a monochromated and aberration-corrected FEI Titan operated in STEM mode at an acceleration voltage of 300 kV, providing a probe size of ≈ 0.5 nm and an energy resolution of 0.08 eV (as measured by the full-width-at-half-maximum of the zero-loss peak). Richardson–Lucy deconvolution was performed to remove the zero-loss peak. An EELS spectrum recorded in vacuum was used as an input for the point-spread function. Due to a small asymmetry in the zero-loss peak, the deconvolution algorithm produced an artificial EELS peak in the energy range below 0.5 eV. However, the artificial peak did not overlap with any of the observed resonances and could be safely removed using a first-order logarithmic polynomial.

The depicted EELS spectra were obtained by integrating the deconvoluted EELS data around the experimental impact parameter b_{exp} . This parameter was directly related to the effective radius of the particle R which was extracted as a radius of the circle with the effective area of the particle found from the STEM image. The effective area of the particle is the area which was enclosed by the boundaries obtained by Otsu's thresholding of the STEM image. The integration parameter itself was calculated as a radius of the circle with the adjusted effective area of expanded/reduced initial boundaries. For the integration region centered at the nanoparticle, the experimental impact parameter changes from 0 to $0.3R$ and denotes the nanoparticle center. For the annulus-shaped regions, the experimental impact parameter denoted the mean of the inner and outer radii with a typical radius spread of 0.3 (for example, the green region in Figure 4a encloses the regions from 0.35 to $0.65R$). The depicted EELS spectra were smoothed with a Gaussian function ($\sigma = 0.03$ eV).

The EELS maps were obtained by summing the deconvoluted EELS data in a spectral window of 0.02 eV width centered at the resonance energies. The signal-to-noise ratio was improved by spatially binning the map, reducing the total number of pixels by a factor of four. A Gaussian filter with $\sigma = 0.8$ pixels was applied to smooth the image.

Laser Reshaping: The laser reshaping was performed using a laser melting in liquid process.^[62,63] The process yields spherical nanoparticles from irregular shaped nanoparticles by selective heating due to the light absorption of solid particles. Unprocessed BP nanoparticles in acetone were irradiated with the third harmonic of a Nd:YAG laser (355 nm wavelength, 5 nm pulse width, 20 Hz repetition rate) with a fluence of 50 mJ cm⁻² per pulse for 10 min to obtain spherical BP nanoparticles.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

S.R. and A.A. acknowledge support from the Independent Research Funding Denmark (7026-00117B). K.S.T. acknowledges support from the Center for Nanostructured Graphene (CNG) under the Danish National Research Foundation (project DNR103) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant No. 773122, LIMA). K.S.T. is a Villum Investigator supported by VILLUM FONDEN (Grant No. 37789). H.S. acknowledges support by JSPS KAKENHI Grant Number 21K14496.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

M.K.S. constructed the computational workflow, performed the screening, developed the f-sum approach, and prepared the figures. H.S. performed the structural and optical characterizations of nanoparticles. A.A. performed the EELS measurements, EELS analysis, EELS simulations, and prepared a figure. D.S. fabricated the nanoparticles and conducted laser shaping processes. M.F. contributed to analyses and interpretation of the data and supervised the fabrication and characterizations of nanoparticles. S.R. and K.S.T. conceived the idea and supervised the work. S.R. performed the optical simulations. All authors discussed the results and contributed to the preparation of the manuscript.

Data Availability Statement

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

Keywords

boron phosphide, electron energy-loss spectroscopy, high-refractive-index nanostructures, high-throughput screening, laser reshaping, Mie resonances

Received: February 22, 2022

Revised: April 1, 2022

Published online: May 29, 2022

- [1] D. Zhao, Z. Lin, W. Zhu, H. J. Lezec, T. Xu, A. Agrawal, C. Zhang, K. Huang, *Nanophotonics* **2021**, *10*, 2283.
- [2] A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, B. Luk'yanchuk, *Science* **2016**, *354*, aag2472.
- [3] S. Kruk, Y. Kivshar, *ACS Photonics* **2017**, *4*, 2638.
- [4] D. G. Baranov, D. A. Zuev, S. I. Lepeshov, O. V. Kotov, A. E. Krasnok, A. B. Evlyukhin, B. N. Chichkov, *Optica* **2017**, *4*, 814.
- [5] I. Staude, J. Schilling, *Nat. Photonics* **2017**, *11*, 274.
- [6] D. J. Wilson, K. Schneider, S. Hönl, M. Anderson, Y. Baumgartner, L. Czornomaz, T. J. Kippenberg, P. Seidler, *Nat. Photonics* **2020**, *14*, 57.
- [7] W. T. Chen, A. Y. Zhu, V. Sarjeev, M. Khorasaninejad, Z. Shi, E. Lee, F. Capasso, *Nat. Nanotechnol.* **2018**, *13*, 220.
- [8] K. Huang, J. Deng, H. S. Leong, S. L. K. Yap, R. B. Yang, J. Teng, H. Liu, *Laser Photon. Rev.* **2019**, *13*, 1800289.
- [9] C. Zhang, S. Divitt, Q. Fan, W. Zhu, A. Agrawal, Y. Lu, T. Xu, H. J. Lezec, *Light Sci. Appl.* **2020**, *9*, 55.
- [10] Y. Gutiérrez, D. Ortiz, J. M. Saiz, F. González, P. Albella, F. Moreno, *Appl. Sci.* **2018**, *8*, 2065.
- [11] J. Hu, M. Lawrence, J. A. Dionne, *ACS Photonics* **2020**, *7*, 36.
- [12] I. Aharonovich, A. D. Greentree, S. Prawer, *Nat. Photonics* **2011**, *5*, 397.
- [13] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, O. Levy, *Nat. Mater.* **2013**, *12*, 191.
- [14] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
- [15] S. Kirkin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, C. Wolverton, *NPJ Comput. Mater.* **2015**, *1*, 15010.
- [16] S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, J. Gath, K. W. Jacobsen, J. J. Mortensen, T. Olsen, K. S. Thygesen, *2D Mater.* **2018**, *5*, 042002.
- [17] M. N. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N. R. Knøsgaard, M. Kruse, A. H. Larsen, S. Manti, T. G. Pedersen, U. Petralanda, T. Skovhus, M. K. Svendsen, J. J. Mortensen, T. Olsen, K. S. Thygesen, *2D Mater.* **2021**, *8*, 044002.
- [18] J. B. Varley, A. Miglio, V.-A. Ha, M. J. van Setten, G.-M. Rignanese, G. Hautier, *Chem. Mater.* **2017**, *29*, 2568.
- [19] L. Yu, A. Zunger, *Phys. Rev. Lett.* **2012**, *108*, 068701.
- [20] I. E. Castelli, T. Olsen, S. Datta, D. D. Landis, S. Dahl, K. S. Thygesen, K. W. Jacobsen, *Environ. Sci. Technol.* **2012**, *5*, 5814.
- [21] I. Petousis, D. Mrdjenovich, E. Ballouz, M. Liu, D. Winston, W. Chen, T. Graf, T. D. Schladt, K. A. Persson, F. B. Prinz, *Sci. Data* **2017**, *4*, 160134.
- [22] F. Naccarato, F. Ricci, J. Suntivich, G. Hautier, L. Wirtz, G.-M. Rignanese, *Phys. Rev. Mater.* **2019**, *3*, 044602.
- [23] M. Gjerding, T. Skovhus, A. Rasmussen, F. Bertoldo, A. H. Larsen, J. J. Mortensen, K. S. Thygesen, *Comput. Mater. Sci.* **2021**, *199*, 110731.
- [24] J. J. Mortensen, M. Gjerding, K. S. Thygesen, *J. Open Source Software* **2020**, *5*, 1844.
- [25] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865.
- [26] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [27] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Duřak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, *J. Phys.: Condens. Matter* **2010**, *22*, 253202.
- [28] J. Yan, J. J. Mortensen, K. W. Jacobsen, K. S. Thygesen, *Phys. Rev. B* **2011**, *83*, 245122.
- [29] T. Moss, *Phys. Status Solidi B* **1985**, *131*, 415.
- [30] J. P. Perdew, *Int. J. Quantum Chem.* **1985**, *28*, 497.

- [31] P. Popper, T. A. Ingles, *Nature* **1957**, 179, 1075.
- [32] T. Takenaka, M. Takigawa, K. Shohno, *Jpn. J. Appl. Phys.* **1976**, 15, 2021.
- [33] W. Wetzling, J. Windscheif, *Solid State Commun.* **1984**, 50, 33.
- [34] M. Odawara, T. Udagawa, G. Shimaoka, *Jpn. J. Appl. Phys.* **2005**, 44, 681.
- [35] S. Dalui, S. Hussain, S. Varma, D. Paramanik, A. K. Pal, *Thin Solid Films* **2008**, 516, 4958.
- [36] R. J. Archer, R. Y. Koyama, E. E. Loebner, R. C. Lucas, *Phys. Rev. Lett.* **1964**, 12, 538.
- [37] E. Schroten, A. Goossens, J. Schoonman, *J. Appl. Phys.* **1998**, 83, 1660.
- [38] D. E. Aspnes, A. Studna, *Phys. Rev. B* **1983**, 27, 985.
- [39] T. Siefke, S. Kroker, K. Pfeiffer, O. Puffky, K. Dietrich, D. Franta, I. Ohlídal, A. Szeghalmi, E.-B. Kley, A. Tünnermann, *Adv. Opt. Mater.* **2016**, 4, 1780.
- [40] H. Phillip, E. Taft, *Phys. Rev.* **1964**, 136, A1445.
- [41] J. Yang, J. A. Fan, *Opt. Express* **2017**, 25, 23899.
- [42] P. Lalanne, P. Chavel, *Laser Photonics Rev.* **2017**, 11, 1600295.
- [43] Y. Kumashiro, *J. Mater. Res.* **1990**, 5, 2933.
- [44] K. Woo, K. Lee, K. Kovnir, *Mater. Res. Express* **2016**, 3, 074003.
- [45] Y. Kumashiro, K. Nakamura, T. Enomoto, M. Tanaka, *J. Mater. Sci. Mater. Electron.* **2011**, 22, 966.
- [46] B. Padavala, C. D. Frye, X. Wang, Z. Ding, R. Chen, M. Dudley, B. Raghoothamachar, P. Lu, B. N. Flanders, J. H. Edgar, *Cryst. Growth Des.* **2016**, 16, 981.
- [47] A. Crovetto, J. M. Adamczyk, R. R. Schnepf, C. L. Perkins, H. Hempel, S. R. Bauers, E. S. Toberer, A. C. Tamboli, T. Unold, A. Zakutayev, *Adv. Mater. Interfaces* **2022**, 9, 2200031.
- [48] H. Sugimoto, M. Fujii, K. Imakita, *RSC Adv.* **2015**, 5, 8427.
- [49] M. Feng, J. Zhang, H. Zhou, P. Jiang, D. Zhang, L. Wang, S. Chen, *Mater. Res. Express* **2019**, 6, 125922.
- [50] H. Sugimoto, B. Somogyi, T. Nakamura, H. Zhou, Y. Ichihashi, S. Nishiyama, A. Gali, M. Fujii, *J. Phys. Chem. C* **2019**, 123, 23226.
- [51] H. Sugimoto, M. Fujii, *Adv. Opt. Mater.* **2017**, 5, 1700332.
- [52] H. Sugimoto, T. Okazaki, M. Fujii, *Adv. Opt. Mater.* **2020**, 8, 2000033.
- [53] F. J. García de Abajo, *Rev. Mod. Phys.* **2010**, 82, 209.
- [54] C. Colliex, M. Kociak, O. Stéphan, *Ultramicroscopy* **2016**, 162, A1.
- [55] A. Assadillayev, T. Hinamoto, M. Fujii, H. Sugimoto, M. L. Brongersma, S. Raza, *ACS Photonics* **2021**, 8, 1582.
- [56] A. Assadillayev, T. Hinamoto, M. Fujii, H. Sugimoto, S. Raza, *Nanophotonics* **2021**, 10, 4161.
- [57] D. T. Alexander, V. Flauraud, F. Demming-Janssen, *ACS Nano* **2021**, 15, 16501.
- [58] J. H. Song, S. Raza, J. van de Groep, J. H. Kang, Q. Li, P. G. Kik, M. L. Brongersma, *Nat. Commun.* **2021**, 12, 48.
- [59] S. Raza, M. Esfandyarpour, A. L. Koh, N. A. Mortensen, M. L. Brongersma, S. I. Bozhevolnyi, *Nat. Commun.* **2016**, 7, 13790.
- [60] R. Alaee, C. Rockstuhl, I. Fernandez-Corbaton, *Opt. Commun.* **2018**, 407, 17.
- [61] J. Yang, J. P. Hugonin, P. Lalanne, *ACS Photonics* **2016**, 3, 395.
- [62] X. Li, A. Pyatenko, Y. Shimizu, H. Wang, K. Koga, N. Koshizaki, *Langmuir* **2011**, 27, 5076.
- [63] D. Zhang, B. Gökce, S. Barcikowski, *Chem. Rev.* **2017**, 117, 3990.

11.3 Paper [III]: Momentum-Dependent Oscillator Strength Crossover of Excitons and Plasmons in Two-Dimensional PtSe₂

Jinhua Hong[†], Mark Kamper Svendsen[†], Masanori Koshino, Thomas Pichler, Hua Xu, Kazu Suenaga, Kristian S Thygesen,

[†] The authors contributed equally to the work,

“Momentum-Dependent Oscillator Strength Crossover of Excitons and Plasmons in Two-Dimensional PtSe₂,”

[ACS Nano 2022, 16, 8, 12328–12337.](#)

Momentum-Dependent Oscillator Strength Crossover of Excitons and Plasmons in Two-Dimensional PtSe₂

Jinhua Hong,[¶] Mark Kamper Svendsen,^{*,¶} Masanori Koshino,[¶] Thomas Pichler, Hua Xu, Kazu Suenaga,^{*} and Kristian S. Thygesen



Cite This: *ACS Nano* 2022, 16, 12328–12337



Read Online

ACCESS |

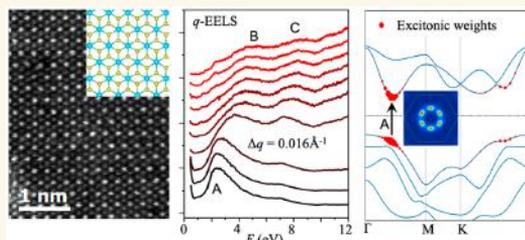
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The 1T-phase layered PtX₂ chalcogenide has attracted widespread interest due to its thickness dependent metal–semiconductor transition driven by strong interlayer coupling. While the ground state properties of this paradigmatic material system have been widely explored, its fundamental excitation spectrum remains poorly understood. Here we combine first-principles calculations with momentum (*q*) resolved electron energy loss spectroscopy (*q*-EELS) to study the collective excitations in 1T-PtSe₂ from the monolayer limit to the bulk. At finite momentum transfer, all the spectra are dominated by two distinct interband plasmons that disperse to higher energy with increasing *q*. Interestingly, the absence of long-range screening in the two-dimensional (2D) limit inhibits the formation of long wavelength plasmons. Consequently, in the small-*q* limit, excitations in monolayer PtSe₂ are exclusively of excitonic nature, and the loss spectrum coincides with the optical spectrum. The qualitatively different momentum dependence of excitons and plasmons enables us to unambiguously disentangle their spectral fingerprints in the excited state spectrum of layered 1T-PtSe₂. This will help to discern the charge carrier plasmon and locally map the optical conductivity and trace the layer-dependent semiconductor to metal transition in 1T-PtSe₂ and other 2D materials.

KEYWORDS: monolayer PtSe₂, momentum resolved electron energy loss spectroscopy, plasmons, excitons, 2D materials, electronic screening in 2D, *ab initio* calculations



Platinum dichalcogenides (PtX₂, X = S, Se, Te) are emerging as a class of two-dimensional (2D) layered materials^{1,2} with thickness- and anion-dependent electronic properties resulting from the strong interlayer coupling.^{3–5} In this octahedral-coordinated 1T-phase system, PtSe₂ is an indirect semiconductor in monolayer and bilayer forms and evolves into a metal as the thickness increases to trilayer and toward the bulk.^{4,6} This semiconductor-to-metal transition is caused by the giant coupling of *p*_z–*p*_z states of two adjacent interlayer chalcogen atoms.^{4,5} This is in contrast to the well-known 2H-MoS₂ system where the weaker interlayer coupling reduces the gap as more layers are stacked, but is insufficient to drive a semiconductor-to-metal transition. When the chalcogen changes from S over Se to Te, the bulk PtX₂ turns from a semiconductor (PtS₂) to a metal (PtSe₂ and PtTe₂). Chia et al.⁷ showed that the enhanced conductivity across the chalcogen group gives rise to a monotonic increasing catalytic performance in the hydrogen evolution reaction.⁸ Vacancy defect induced magnetism and a layer-modulated ferromagnetic-to-antiferromagnetic crossover have

been observed in both metallic PtSe₂ few-layer⁹ and semi-conducting bi- and monolayer structures.¹⁰ In the electronic transport characteristics, an *n*- to *p*-type conductivity modulation has been realized in vacancy doped monolayer PtSe₂ via controlled chemical vapor deposition.¹¹ The monolayer PtX₂ semiconductors present high electron mobilities and are stable in air,⁴ demonstrating their potential for mid-infrared optoelectronics,¹² magnetism, and catalysis applications.

While the transport and catalytic properties of PtX₂ have been thoroughly investigated, its electronic excitation spectrum, including the collective excitations (excitons and

Received: April 5, 2022

Accepted: July 26, 2022

Published: August 1, 2022



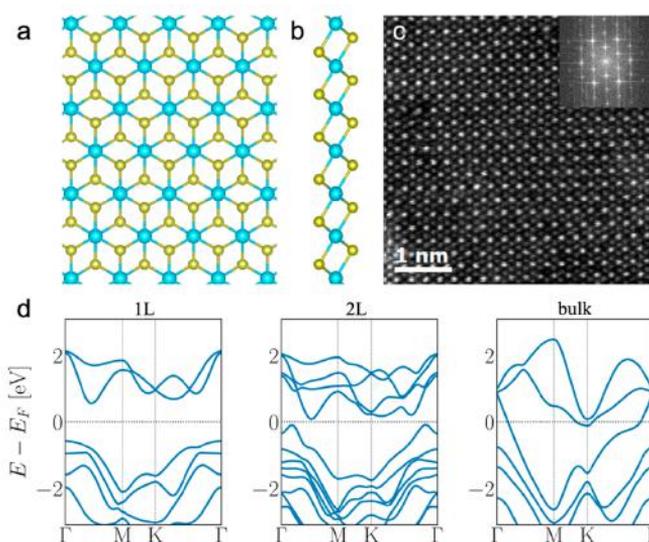


Figure 1. Atomic and electronic structures of PtSe₂. (a–b) Top- and side-view structure models of monolayer 1TPtSe₂. Cyan balls are Pt, and yellow balls are Se atoms. (c) ADF-STEM image of monolayer PtSe₂. (d) Thickness-dependent electronic band structure of PtSe₂ in the 2D Brillouin zone.

plasmons), remains largely unexplored. Understanding these excitations^{13–15} is the key to controlling the optical properties of the material and for future photonic or optoelectronic applications.^{12,16–19} The optical excitations in the MoX₂ and WX₂ semiconducting TMDs have been intensely studied during the past years because of the unique exciton physics found in their mono- and few-layer structures.^{20,21} In comparison, the optical properties of few-layer PtX₂ have only recently gained attention, and only the most basic properties of their excitons have so far been addressed.^{22–25} The plasmon excitations of bulk and thin layers of PtTe₂ were reported by Ghosh et al.²⁶ However, in these measurements the samples were not atomically thin and substrate–sample interactions could not be ruled out. To date, the electronic excitation spectrum of freestanding few-layer PtX₂ remains unexplored, and it is unclear how it depends on the number of layers and the metal–semiconductor transition taking place as the thickness decreases to the monolayer limit.

In this work, we employ momentum (q) resolved electron energy loss spectroscopy (q -EELS)^{27–32} in a transmission electron microscope (TEM) to probe the valence excitations of PtSe₂ samples of different thicknesses. In the high energy range, we identify the existence of in-plane polarized plasmons at 7–10 eV across various film thicknesses. By tracking their q -dispersion we find that this plasmon excitation is enhanced at large nonzero q , similar to the π plasmon of graphene. At lower energies, a peak at ~ 2 eV appears and dominates the EEL spectrum, particularly for the thinner samples. We ascribe this peak to excitonic effects in the two-dimensional samples and show that it is highly sensitive to q . Meanwhile, a plasmon-like feature emerges between the exciton and the high-energy plasmon, which has not been observed previously in other TMDs. The loss functions calculated from first principles within the random phase approximation (RPA) confirm the plasmonic nature of the high- and intermediate-energy peaks. We furthermore employ the Bethe Salpeter Equation (BSE) to

prove the excitonic nature of the low energy peak in the monolayer spectrum and map out the k -space wave function of the lowest exciton. Overall, we find good agreement between theory and experiment, which allows us to perform a detailed study of the collective excitations in layered PtSe₂ and elucidate the intricate connection between sample dimensionality (2D vs 3D), dielectric screening, and the relative importance of excitons and plasmons as a function of momentum transfer, q . The q -dependent crossover in the spectral weights of excitons and plasmons also provides a handle to study the thickness-dependent band gaps, metal-to-semiconductor transition, and many-body effects of PtSe₂ and PtTe₂.³³

RESULTS AND DISCUSSION

High quality monolayer and few-layer PtSe₂ are grown on mica substrates by direct selenization of PtCl₄ in a chemical vapor deposition (CVD) furnace at 700 °C. Single-crystal domains are easily formed with a typical size of tens of micrometers. The as-grown samples are transferred onto TEM grids using PMMA spin coating and KI etching, and then washed in acetone and isopropanol to remove the surface residues. Before the q -EELS experiments, the PtSe₂ samples are annealed in the TEM chamber at 200 °C for 3 h (Supporting Information, Figures S1–S2) to reduce the surface residue or contamination. All the q -EEL spectra are collected in a diffraction mode using parallel beam illumination, with an energy resolution of 40 meV and a momentum resolution of 0.03 Å⁻¹ (see more details in Figures S3–S5). In our q -EELS setup, the momentum q always refers to in-plane polarized excitation, since the out-of-plane component is estimated to be $\sim q_E = k\theta_E = (2\pi/\lambda) \cdot (\Delta E/2E_0) = 0.004$ Å⁻¹ (for a signal at 4 eV) which can be ignored in the normal incident case. The out-of-plane excitation may exist only in the $q \rightarrow 0$ EEL spectra, but get buried by the dominant in-plane components. To extract the

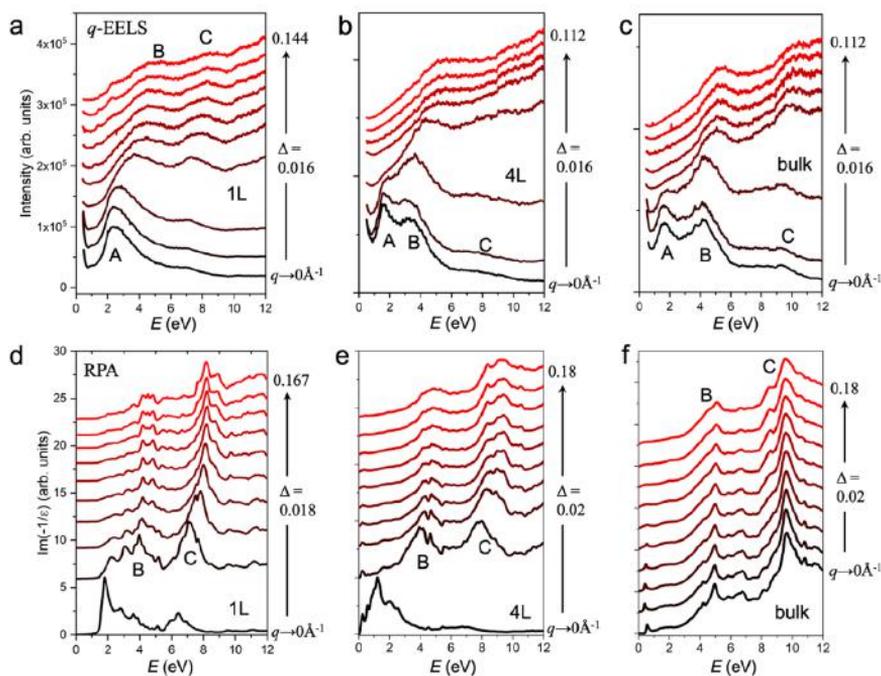


Figure 2. Momentum-dependent loss function of PtSe₂. (a–c) Experimental *q*-EELS of 1L, 4L and bulk PtSe₂, respectively. (d–f) Calculated *q*-dependent loss functions in the random phase approximation (RPA) of 1L, 4L, and bulk PtSe₂, respectively. The low energy excitonic peak A, and the intermediate and high energy plasmonic peaks B and C are indicated.

out-of-plane excitation, tilting the sample to high angles may give rise to different characteristic modes³⁴ and uncover the in-plane/out-of-plane anisotropy of the collective excitations. However, this is not allowed by the narrow pole-piece gap of our microscopes and thus beyond the scope of the current work.

Atomic and Electronic Structures of PtSe₂. Figure 1 shows the atomic and electronic band structures of layered 1T-PtSe₂. In the structural models in Figure 1a, 1b, the Pt atoms (in cyan) are coordinated with 6 Se atoms (in yellow), forming an octahedral crystal field with *d*²*sp*³ hybridization. This is the known 1T phase structure, and each layer is stacked AA' in order to form the bulk system. In Figure 1c, the atomically resolved annular dark-field scanning transmission electron microscopy (ADF-STEM) image clearly shows the atomic coordination of monolayer PtSe₂ in the 1T phase. The brighter Pt atom is surrounded by 6 darker Se atoms. Based on this atomic structure, our density functional theory (DFT) calculations provide the single-particle electronic band structure of PtSe₂ as a function of film thickness (Figure 1d). In the monolayer case, PtSe₂ is an indirect semiconductor (*E*_g = 1.1 eV) with the valence band maximum (VBM) at the Γ point and the conduction band minimum (CBM) between Γ and M. As the thickness increases to bilayer, the band gap decreases drastically to only 0.2 eV. When further increasing to trilayer or bulk, the electronic structure evolves into a metal. Figure 1d shows the metal-to-semiconductor transition as the thickness decreases from bulk to monolayer (also see Supporting Information, Figure S6). This transition originates from the strong electronic interlayer hybridization, which

occurs because the wave functions of the valence and conduction bands have a strong component of Se-*p*_z orbitals (see Figure S6). This is in contrast to the well-known indirect-to-direct gap transition in 2H-MoS₂ where the hybridization is significantly weaker due to the wave functions of the valence and conduction bands being mainly localized on the inner Mo atoms.

Loss Functions of PtSe₂ for Different Thicknesses. To explore the electronic excitations of PtSe₂, we employ both *q*-EELS measurements and first-principles calculations. Figure 2a–c show the low loss fine structure of single-layer (1L), 4-layer (4L), and bulk PtSe₂ acquired at different *q* along the Γ M direction (assuming that in-plane anisotropy can be ignored within the limited *q* range <0.2 Å⁻¹). Three clear peaks denoted A (~2 eV), B (~4 eV), and C (~8 eV) dominate the electronic excitation spectrum. Interestingly, the relative intensities of the three peaks vary strongly with both *q* and layer thickness. Understanding the origin of these variations and the nature of the electronic excitations behind the peaks (whether excitonic or plasmonic) is a main focus of this work.

For all thicknesses the B and C peaks dominate the EEL spectra for large *q*. For smaller *q*, the C peak disappears almost entirely for all thicknesses while the B peak remains in the bulk spectrum, becomes weaker in 4L, and disappears in the monolayer limit. The fact that the B and C spectral features are well reproduced by our random phase approximation (RPA) calculations (see Figure 2d–f) suggests that the underlying excitations are of plasmonic nature. This is because the RPA neglects the attractive electron–hole (e–h) interaction

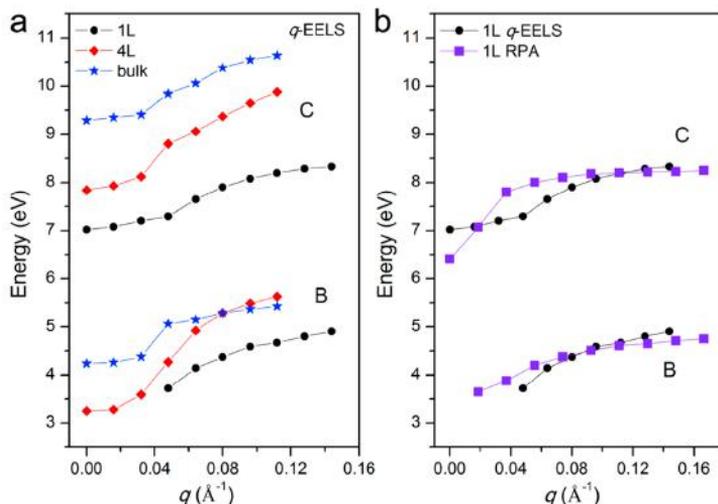


Figure 3. Dispersion of the loss peaks of PtSe₂. (a) Experimental dispersion of the B and C peaks of PtSe₂ for different thicknesses. (b) Comparison of the experimental and RPA calculated dispersion of the B and C peaks for monolayer PtSe₂.

responsible for exciton formation and only include e – h exchange that drives plasmon formation.

For momentum-transfer below 0.05 \AA^{-1} a low-lying peak A at $\sim 2 \text{ eV}$ appears in the EEL spectrum of all samples. Interestingly, the A peak is only present at small q and becomes dwarfed by the B and C plasmonic features for larger q —a trend that is particularly pronounced in the monolayer limit. In the following we discuss why this observation is indicative of the peak A being of excitonic nature in the two-dimensional samples.^{23,35}

q Dependence of the Electronic Screening in 2D vs 3D. To understand why the small- q EEL spectrum of the monolayer is dominated by the excitonic A peak while the high-energy B and C plasmon peaks dominate for larger q , we consider the Bethe Salpeter equation (BSE) which can account for both single-particle excitations, plasmons and excitons.^{36–38} In the BSE, the excitations are found as solutions to the eigenvalue problem,

$$\sum_S H_{SS'}(\mathbf{q}) F_S^\lambda(\mathbf{q}) = E^\lambda F_S^\lambda(\mathbf{q}) \quad (1)$$

where $H_{SS'}(\mathbf{q})$ is the two-particle BSE Hamiltonian in the e – h pair orbital basis labeled by the pair-orbital index $S = (n, m, \mathbf{k}, \mathbf{q})$, $F_S^\lambda(\mathbf{q})$ is the exciton wave function of the λ^{th} solution, and E^λ is the corresponding energy. The BSE Hamiltonian reads

$$H_{SS'}(\mathbf{q}) = (\epsilon_{m, \mathbf{k}+\mathbf{q}}^{\text{QP}} - \epsilon_{n, \mathbf{k}}^{\text{QP}}) \delta_{SS'} - (f_{m, \mathbf{k}+\mathbf{q}} - f_{n, \mathbf{k}}) K_{SS'}(\mathbf{q}) \quad (2)$$

The diagonal holds the quasiparticle (QP) energies of the electron and hole state. The exchange-correlation-kernel, $K_{SS'}(\mathbf{q})$, that renormalizes and mixes the single-particle transitions can be divided into two constituent parts: the e – h exchange interaction (V) and the direct screened interaction (W)

$$K_{SS'}(\mathbf{q}) = 2V_{SS'}(\mathbf{q}) - W_{SS'}(\mathbf{q}) \quad (3)$$

where the factor 2 appears because we have specialized to singlet excitations. Setting $K = 0$ yields the single-particle

spectrum of bare e – h transitions while setting $W = 0$ yields the RPA. In general, both V and W are positive, and thus the former increases the energy of the collective excitation while the latter decreases its energy, relative to the bare QP transition energies of the constituent e – h pairs. This leads to the observation (which may also be taken as a definition) that V is responsible for plasmon formation and W is responsible for exciton formation. The relative importance of these two terms therefore dictates what type of excitations will dominate at a given q . We note that in this discussion we do not distinguish between interband and intraband plasmons.

The exchange interaction takes the form

$$V_{SS'}(\mathbf{q}) = \int d\mathbf{r} d\mathbf{r}' \frac{\psi_{n\mathbf{k}}(\mathbf{r}) \psi_{m\mathbf{k}+\mathbf{q}}^*(\mathbf{r}) \psi_{n'\mathbf{k}'}^*(\mathbf{r}') \psi_{m'\mathbf{k}'+\mathbf{q}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (4)$$

It can be shown that, for small momentum transfers, the long wavelength (small q) Fourier component of $\psi_{n\mathbf{k}}(\mathbf{r}) \psi_{m\mathbf{k}+\mathbf{q}}^*(\mathbf{r})$ is proportional to q regardless of the dimensionality of the system. In two dimensions, the Fourier transformed Coulomb interaction goes as $v_q \sim 1/q$ for small q , which means that the exchange interaction goes as $V_{SS'}(\mathbf{q}) \approx q$ and therefore it becomes negligible in the optical limit ($q \rightarrow 0$). For this reason, we expect very weak plasmon formation at small q in atomically thin 2D materials. This is in stark contrast to the 3D case where the Coulomb interaction goes as $v_q \sim 1/q^2$ meaning that the exchange interaction in 3D can be significant also at small q . In the 3D case, plasmons can therefore exist even in the optical limit, and this represents a fundamental qualitative difference between plasmons in 2D and 3D. (See [Supplementary Note 1](#).)

Unlike the exchange interaction matrix elements, the direct screened interaction matrix elements $W_{SS'}(\mathbf{q})$ remain finite and more or less independent of q in the optical limit. This means that, in contrast to plasmons, excitons can still form in 2D materials for small q . In fact, in the limit of small q , excitonic effects are extremely pronounced in 2D semiconductors due to the weak screening and can greatly reconstruct the non-interacting spectrum (see [Figure 5](#) and related discussion).

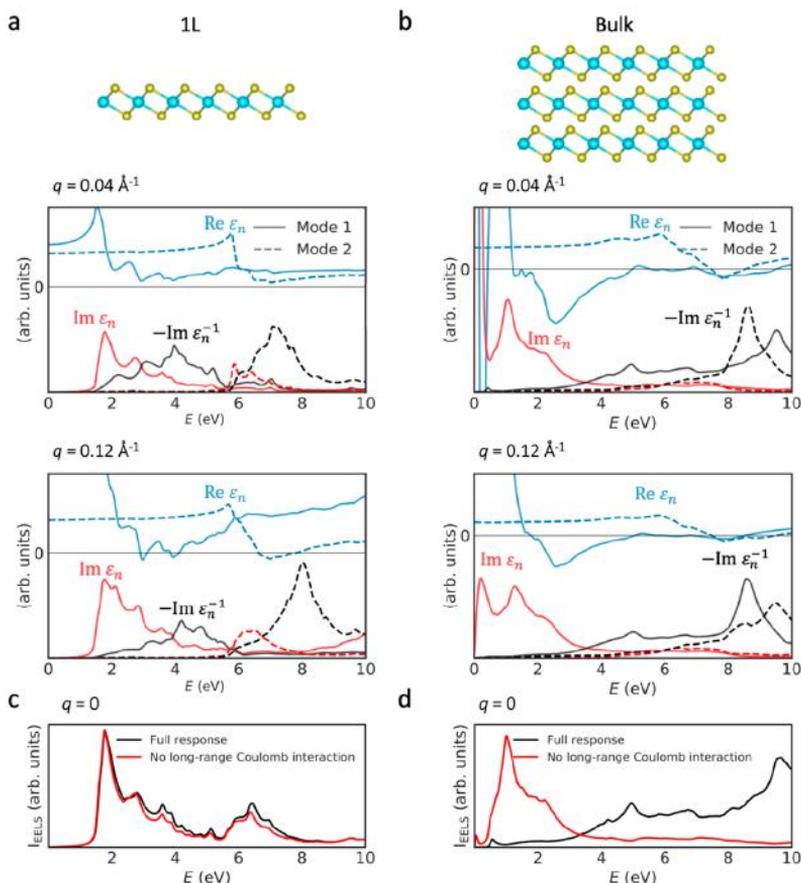


Figure 4. Plasmon formation in 2D vs 3D. (a–b) Evolution of the eigenvalues obtained from the spectral decomposition of the RPA dielectric function for 1L and bulk PtSe₂, as a function of q . The loss spectra of the 1L system are more sensitive to q than those of the bulk because of the gradual turning on of the long-range screening in 2D with increasing q . (c–d) The macroscopic RPA loss function in the $q = 0$ limit for 1L and bulk. Inclusion of the long-range component of the Coulomb interaction (the Fourier component corresponding reciprocal lattice vector $G = 0$) has essentially no effect on the 1L spectrum while the bulk spectrum is affected dramatically due to the formation of plasmons.

However, as q increases, the importance of the exchange interaction grows to dominate the direct interaction, leading to an excitation spectrum dominated primarily by plasmons. As demonstrated in Figure 2, the excitonic peak A dominates at small q while the B and C plasmonic peaks grow dominant as q increases. Although a similar q -dependent exciton–plasmon oscillator strength crossover was shown in theory for MoS₂,³⁹ the momentum integrated EELS employed in that case was insufficient as explicit and direct experimental proof of the effect. In this work, we provide experimental identification of the excitonic or plasmonic nature of excitations in the EEL spectrum using the intricate momentum-dependent characters of collective excitations in 2D materials.

By tracking the peak energies as a function of momentum q (Figure S7), we obtain the dispersion of the plasmon peaks B and C, as shown in Figure 3a. Both the B and C peaks exhibit positive and nonlinear dispersion for all sample thicknesses (Figure 3a–3b). For relatively large q , the dispersion of both peaks is \sqrt{q} -like for all material thicknesses. This behavior is

reminiscent of the acoustic plasmon dispersion in a classical free-electron model in 2D dimensions which is expressed as $\omega = \frac{v_p}{\sqrt{2}} \sqrt{1 - e^{-qd}} \propto \sqrt{q}$ in the finite- q approximation.^{40,41}

At the large- q end, the dispersion curves tend to gradually reach a limit with decreasing slopes as q further increases, in good agreement with the trend in the experimental results. For the monolayer this is well reproduced by our RPA calculations (Figure 3b), while the agreement with theory is less impressive for the 4L and bulk systems (Figure S8) where RPA does predict the plasmon energies correctly to within 0.5 eV, but yields a too flat dispersion. In the small- q range, the experimental peak energies flatten out and deviate from the \sqrt{q} -dispersion. We note that, due to the limited momentum resolution in the diffraction mode, the spectra at $q < 0.04 \text{ \AA}^{-1}$ are influenced by the excitation at the optical limit $q \rightarrow 0$, which may affect the dispersion curve in this q range. We further observe that the B and C peaks blue shift with sample

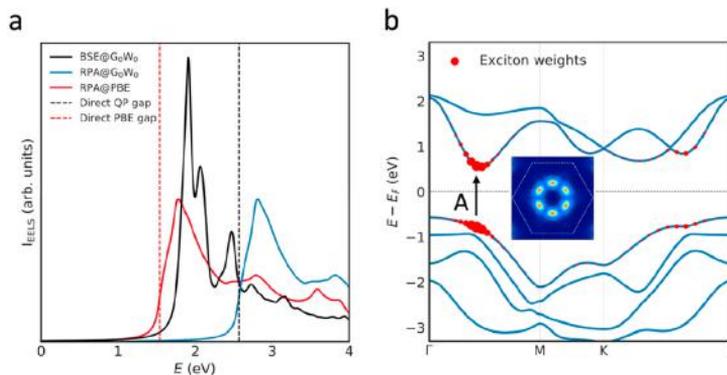


Figure 5. Excitonic origin of the A-peak in monolayers. (a) Comparison of BSE and RPA calculated loss functions of monolayer PtSe₂ in the $q \rightarrow 0$ limit. The BSE spectrum is calculated on top of the G_0W_0 band structure. For comparison, the RPA spectrum is calculated on top of the PBE band structure and the PBE band structure corrected to match the G_0W_0 band gap, respectively. (b) The BSE calculated excitonic weights of the A peak. The colored inset shows the 2D projected exciton wave function distribution in k space. Both the excitonic weights and k -space wave function show that the A exciton originates from direct transitions at the midpoint of ΓM (highlighted by the black arrow).

thickness, which is also reproduced by our RPA calculations (Figure S8).

Eigenmode Analysis of the Plasmon Excitations. In general, the q -EEL spectrum is directly related to the inverse of the macroscopic dielectric function, $\epsilon_M(q, \omega)$. While ϵ_M is thus the relevant physical observable, it provides little insight into the nature of the excitations underlying the different spectral features, e.g. whether they are of single-particle, plasmonic, or excitonic character, whether one or several different excitations contribute to a specific spectral feature, or what the spatial form and symmetry of such excitations are. As already discussed in relation to eq 3, the first question is largely a matter of which level of theory is required to reproduce a given spectral feature. At this point we have established that the B and C peaks are well described at the RPA level and thus have a single-particle and/or plasmonic nature. To investigate the plasmonic nature of the B and C peaks in more detail and to clarify whether they represent one, two, or several different plasmon excitations, we turn to an eigenmode decomposition of the dielectric function. Following Andersen et al.,⁴² the microscopic dielectric function can be decomposed into a set of eigenmodes

$$\epsilon(\mathbf{r}, \mathbf{r}', \omega) = \sum_n \epsilon_n(\omega) \phi_n(\mathbf{r}, \omega) \rho_n(\mathbf{r}', \omega) \quad (5)$$

where $\epsilon_n(\omega)$ is the eigenvalue of the n^{th} mode of the dielectric function and $\phi_n(\mathbf{r}, \omega)$ and $\rho_n(\mathbf{r}', \omega)$ are the left and right eigenvectors which satisfy the Poisson equation $\nabla^2 \phi_n = -4\pi \rho_n$. Here, ϕ_n is the potential and ρ_n is the induced density of the n^{th} eigenmode. The modes satisfying $\text{Re}\epsilon_n(\omega) = 0$ for some frequency $\omega = \omega_p$ are plasmonic as they represent self-sustained charge density oscillations. We note that only modes with a spatially symmetric profile (described by $\rho_n(\mathbf{r}, \omega)$) perpendicular to the PtSe₂ film will contribute to the macroscopic in-plane q -EELS.

Our spectral analysis shows that the full RPA EEL spectra for all material thicknesses and in the relevant (q, ω)-range can be almost entirely accounted for by two dielectric eigenmodes. Figure 4 shows the individual contribution of the two eigenvalues to $\text{Im}\epsilon(\omega)$, $\text{Re}\epsilon(\omega)$, and $\text{Im}\epsilon^{-1}(\omega)$ for the monolayer (left) and bulk (right) for momentum transfers q

$= 0.04 \text{ \AA}^{-1}$ and $q = 0.12 \text{ \AA}^{-1}$, respectively. For the monolayer, we see that the two modes are separately responsible for the B and C peaks, i.e. each peak can be associated with a distinct excitation. As q increases both peaks evolve from single-particle interband transition peaks into true plasmons, as defined by a zero-crossing of $\text{Re}\epsilon_n(\omega)$. For the bulk system, we see that one mode is entirely responsible for the B peak. This mode also provides the dominant contribution to the C peak but there is a contribution from one additional mode. The dielectric modes of the bulk system vary much less with q and are plasmonic for all values of q . We ascribe the stronger plasmonic character of the excitations in bulk versus 1L, to the stronger dielectric screening response in 3D compared to 2D. In particular, as previously discussed, for small values of q , the e-h exchange term, i.e. the V in eq 3 that is responsible for screening and drives plasmon formation, is directly proportional to q . This point becomes particularly clear in the optical limit ($q = 0$) where the long-range dielectric screening vanishes identically in 2D. Figure 4c-d shows a comparison of the $q = 0$ EEL spectrum calculated with and without the long-range component of the Coulomb interaction for the 1L and bulk systems, respectively. Technically, the removal of the long-range screening is achieved by setting the divergent $\mathbf{G} = 0$ component of the Fourier transformed Coulomb potential, $V_G(\mathbf{q}) = 1/|\mathbf{G} + \mathbf{q}|^2$, to zero (here \mathbf{G} denotes a reciprocal lattice vector). From Figure 4c-d, we conclude that, in the optical limit, the monolayer response is not influenced by the long-range interactions while there is a very significant effect in the 3D bulk system. A direct consequence of this result is that for 2D materials the $q = 0$ EEL spectrum coincides with the optical spectrum. It is well-known that excitonic effects strongly influence the optical spectrum of a 2D semiconductor and thus the same must be expected for the $q = 0$ EEL spectrum, and indeed this is the case (see discussion about the A peak below). However, since the RPA does not account for excitonic effects, the RPA spectrum of a 2D material in the $q = 0$ limit simply reduces to the single-particle spectrum renormalized by the local field effects created by screening from the $\mathbf{G} \neq 0$ components of the Coulomb potential.

As previously noted, the B and C plasmons are of interband nature. We now address the question of which bands/orbitals

are involved in those plasmons. For simplicity we focus on the monolayer in the small q limit where long-range Coulomb interactions are negligible making the connection between the single-particle transitions and the spectral peaks straightforward. To make this identification, we note from Figure 4c that the B and C plasmons originate from the two single-particle interband peaks (van Hove singularities) at about 4 and 8 eV, respectively. At finite q the Coulomb interaction blue shifts these peaks creating the B and C plasmons. By considering the projected density of states (PDOS) we infer that the transitions involved in the B peak are from a lower valence band of mainly Pt- d character situated approximately 2 eV below the Fermi level to the lowest conduction band with mainly Se- p orbitals. Similarly, the C peak originates from transitions from Pt- p to Pt- s orbitals (see Figure S9).

Excitonic Nature of the A Peak. We now turn to the origin of the A peak that we have provisionally referred to as an excitonic peak. To support this interpretation, we perform a BSE calculation for the monolayer in the optical limit. Due to the well-known tendency of PBE to underestimate band gaps, we use the G_0W_0 approximation to correct the band gap of the PBE band structure.^{43–46} A comparison of the RPA and BSE results can be seen in Figure 5a. For completeness we show the RPA result based on the PBE energies and the G_0W_0 -band gap corrected PBE energies, respectively. The large difference between the BSE and RPA spectra highlights the importance of the direct e–h interaction, which drives the formation of excitons and leads to a dramatic rearrangement of spectral weights. We obtain a binding energy of the lowest A exciton of 0.51 eV, which compares well to previous calculations.^{22–24} Because the G_0W_0 -correction of the PBE band gap is comparable in size to the exciton binding energy, the peak in the RPA-PBE spectrum around 2 eV (which is essentially a peak in the joint density of states of the PBE single-particle spectrum) coincides almost with the lowest excitonic peak in the BSE- G_0W_0 spectrum. This explains the relatively good agreement between the RPA-PBE spectra and the experimental spectra in Figure 2a, d.

Next, we analyze the wave function of the A exciton as obtained from the BSE calculation of the PtSe₂ monolayer. Figure 5b shows the distribution of expansion coefficients for the e–h pairs forming the lowest exciton, on the band structure and the Brillouin zone (BZ) (inset), respectively. It appears that the exciton is made up of transitions from the highest valence band to the lowest conduction band located at the midpoint of Γ and M in the BZ.

Due to the large number of atoms in the 4L structure, it is difficult to perform accurate BSE calculations for this system. The fact that the A peak is only visible in the experimental EEL spectrum for small q could indicate, according to the analysis presented earlier, that the A peak is also of excitonic origin in the 4L structure. On the other hand, the existence of excitonic states in the 4L structure seems unlikely due to its metallic nature, which intuitively should screen the direct e–h interaction and prevent exciton formation. However, recalling that screening is strongly suppressed in 2D materials in the small- q limit, it cannot be ruled out that the e–h interaction is sufficiently attractive that excitons may form in the optical limit despite the metallic band structure, and that the A peak in the 4L structure could be of excitonic origin. To confirm this, resonant photoluminescence or differential reflectance spectroscopy could be used in future work.

We performed a BSE calculation for bulk PtSe₂ to investigate how the inclusion of the direct e–h interaction affects the loss spectrum. As shown in Figure S10, we find no significant excitonic effects in the bulk system that can account for the presence of the A-peak. This is as one would expect for a 3D metallic system where excitonic effects should indeed be screened out. We speculate that the A-peak observed in the experimental spectrum of a bulk system for small q could represent an intraband plasmon. We obtain an in-plane (intraband) plasma frequency of 1.6 eV, which agrees well with the experimental A-peak position. However, in the full RPA spectrum the intraband plasmon peak is renormalized to 0.56 eV due to the screening by interband transitions. It should, however, be noted that the PBE functional is known to underestimate interband transition energies, which generally leads to an overestimation of interband screening and thus a red-shifting of plasmon energies in general.⁴⁷ However, we have not been able to confirm this effect from our calculations, and consequently, the origin of the A-peak in the bulk system remains an open question.

CONCLUSIONS

We have unraveled the elementary electronic excitations in layered PtSe₂ using a combination of q -EELS measurements and theoretical calculations. Our analysis shows that the excitation spectrum, from monolayer to bulk, is governed by three distinct collective excitations, namely a low energy peak around 2 eV and two plasmons at 4–5 eV and 7–10 eV, respectively. The qualitatively different wave vector dependence of the dielectric function in 2D and 3D is manifested directly in the excitation spectra. In the monolayer and 4-layer samples, the low energy peak represents an exciton and completely dominates the excitation spectrum in the small q limit, and only at larger q do the plasmons develop and eventually dwarf the excitonic feature. In the bulk, the low energy peak and the plasmons exist side by side even in the small q limit. For all material thicknesses, the EEL spectrum becomes dominated by the plasmons as q is increased beyond $\sim 0.05 \text{ \AA}^{-1}$. Based on the distinct momentum dependence characters, we identify the optical fingerprints of the unknown absorption peaks as excitons or plasmons in the PtSe₂ system, inspiring applications in phonon-assisted photoluminescence.⁴⁸ Our work advances the understanding of the connection between dielectric screening and the formation of collective excitations in solids and establishes the fundamental basis for photonics and optoelectronics of low-dimensional PtSe₂.

METHODS

STEM Imaging and q -EELS. An atomically resolved ADF-STEM image was obtained on an aberration corrected scanning TEM (Triple C1, JEOL JEM-2100CF) operated at 60 kV. The convergence angle of the STEM beam was set as 35 mrad and acceptance angle 70–200 mrad, and the beam current was set to 15 pA. The q -EELS data were collected in diffraction mode using parallel beam illumination on a double Wiener Filtered monochromated TEM (Triple C2, JEOL ARM-60M Δ^2) operated at 60 kV, with an energy resolution of 40–50 meV and momentum resolution of 0.03 \AA^{-1} . The illumination area of the samples for a selected area electron diffraction is approximately 300 nm in diameter. The dual EELS mode of the Gatan Quantum 965 spectrometer was used to obtain the low loss spectra and zero loss peak (ZLP) simultaneously, to eliminate the effect of the ZLP drift during the spectrum acquisition. The dispersion of the spectrum was set to 0.005 eV/channel. In the dual-EELS, we set the dwell time for the ZLP to $5 \times 10^{-5} \text{ s}$ –0.005 s and the dwell time for low loss to 0.05

s–5 s as the q selected moved away from the Γ point. To get a good signal-to-noise ratio, 100 spectra were accumulated for each q , which took 1 min–18 min depending on the q values. For each high- q spectrum which took longer, we moved to fresh sample positions to minimize the effects of irradiation damage on the collective excitations.

DFT Calculation. The calculations were performed using the ASR framework⁴⁹ in combination with the GPAW electronic structure code.⁵⁰ To obtain the multilayer structures, we started from the relaxed T-phase PtSe₂ monolayer (space group $P\bar{3}m1(164)$) structure obtained from C2DB,^{22,51} obtaining the N-layer structures by stacking N monolayers in the most favorable AA' stacking and relaxed it using the PBE-D3 scheme⁵² to account for the vdW-interaction between the layers. We performed the ground state calculations for the monolayer, multilayer, and bulk systems using the asr.gs recipe. The calculations were performed in plane wave mode using a double- ζ polarized basis set, an energy cutoff of 800 eV, and a gamma-point centered k-point grid with a density of 12 k-points/Å² for the multilayer systems and a k-point density of 12 k-points/Å³ for the bulk system. From the ground state calculations, we can then get the band structure using the ASR band structure recipe. The spin–orbit coupling is added nonself-consistently.

RPA and BSE Calculations. To perform the EELS calculations within the RPA, we extended the ASR framework interface with the GPAW code linear response module to enable EELS calculations. Moving forward, the EELS module will be available in ASR. To ensure converged results, we interpolate the ground state calculation onto a k-point grid with an in-plane density of 50 k-points/Å² and include conduction bands equal to five times the number of valence bands. We further employ a broadening of 50 meV and a local field cutoff of 50 eV. We employ a nonlinear frequency grid with $\omega_0 = 0.005$ eV using the tetrahedron integration method for the mono- and multilayer systems, while employing the normal integration method for the bulk system.

The BSE calculations were performed on a k-point grid with an in-plane density of 6 k-points/Å². We included the four top valence bands and the lowest four conduction bands in the calculation of the BSE Hamiltonian. The calculation of the electronic screening was performed using 84 bands, all of the conduction bands equal to five times the number of valence bands, with a local field cutoff of 50 eV. The EELS spectrum was then calculated on a 10001-point frequency grid spanning 0–8 eV. To get the G_0W_0 correction to the quasiparticle energies, we used the asr.gw recipe. To generate the excitonic weights plot for the lowest lying exciton, we extracted the excitonic weights from the BSE calculation and then matched the k-points of the band structure calculation to the k-point grid of the BSE calculation to get the projection.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsnano.2c03322>.

Figures showing the low-magnification STEM images, simulated STEM images, momentum selection in q -EELS, thickness determination by EELS, temporal stability of the EEL spectrum, thickness dependent orbital projected band structures and density of states, ZLP removal and peak fitting in the experiments, experimental and calculated peak dispersions in 4L and bulk PtSe₂, interband transitions in 1L and bulk, influence of excitonic effects and intraband screening in the bulk (Figures S1–S10). Supplementary note elaborating on the momentum transfer dependence of the different Coulomb interaction matrix elements. (PDF)

AUTHOR INFORMATION

Corresponding Authors

Kazu Suenaga – Nanomaterials Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8565, Japan; The Institute of Scientific and Industrial Research (ISIR-SANKEN), Osaka University, Ibaraki 567-0047, Japan; Email: suenaga-kazu@sanken.osaka-u.ac.jp

Mark Kamper Svendsen – Computational Atomic-scale Materials Design (CAMD), Department of Physics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark; orcid.org/0000-0001-9718-849X; Email: markas@dtu.dk

Authors

Jinhua Hong – Nanomaterials Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8565, Japan; orcid.org/0000-0002-6406-1780

Masanori Koshino – Nanomaterials Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8565, Japan

Thomas Pichler – Faculty of Physics, University of Vienna, A-1090 Vienna, Austria; orcid.org/0000-0001-5377-9896

Hua Xu – Key Laboratory of Applied Surface and Colloid Chemistry, School of Materials Science and Engineering, Shaanxi Normal University, Xi'an 710119, P. R. China; orcid.org/0000-0001-5277-7901

Kristian S. Thygesen – Computational Atomic-scale Materials Design (CAMD), Department of Physics and Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark; orcid.org/0000-0001-5197-214X

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acsnano.2c03322>

Author Contributions

[†]J.H. and M.S. contributed equally to this work. J.H. and K.S. conceived this project. J.H. performed the q -EELS measurement; M.S. and K.T. did the theoretical calculation; H.X. contributed the samples; J.H., M.S., M.K., K.T., K.S., and T.P. commented and analyzed the data and theory; all authors cowrote the paper with full discussion.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by JST-CREST (JPMJCR20B1, JPMJCR20B5, JPMJCR1993) and JSPS-KAKENHI (JP16H06333, JP19K04434 and JP17H04797) and A3 Foresight program. H. Xu acknowledges the support from the National Natural Science Foundation of China (51972204). The Center for Nanostructured Graphene (CNG) is sponsored by the Danish National Research Foundation, Project DNRF103. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant agreements No 773122 (LIMA) and No 951215 (MORE-TEM). K.S.T. is a Villum Investigator supported by VILLUM FONDEN (Grant No. 37789).

REFERENCES

- (1) Chhowalla, M.; Shin, H. S.; Eda, G.; Li, L. J.; Loh, K. P.; Zhang, H. The Chemistry of Two-Dimensional Layered Transition Metal Dichalcogenide Nanosheets. *Nat. Chem.* **2013**, *5* (4), 263–275.
- (2) Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. Electronics and Optoelectronics of Two-Dimensional Transition Metal Dichalcogenides. *Nat. Nanotechnol.* **2012**, *7* (11), 699–712.
- (3) Wang, Y.; Li, L.; Yao, W.; Song, S.; Sun, J. T.; Pan, J.; Ren, X.; Li, C.; Okunishi, E.; Wang, Y.-Q.; et al. Monolayer PtSe₂, a New Semiconducting Transition-Metal-Dichalcogenide, Epitaxially Grown by Direct Selenization of Pt. *Nano Lett.* **2015**, *15* (6), 4013–4018.
- (4) Zhao, Y.; Qiao, J.; Yu, Z.; Yu, P.; Xu, K.; Lau, S. P.; Zhou, W.; Liu, Z.; Wang, X.; Ji, W.; et al. High-Electron-Mobility and Air-Stable 2D Layered PtSe₂ FETs. *Adv. Mater.* **2017**, *29* (5), 1604230.
- (5) Zhao, Y.; Qiao, J.; Yu, P.; Hu, Z.; Lin, Z.; Lau, S. P.; Liu, Z.; Ji, W.; Chai, Y. Extraordinarily Strong Interlayer Interaction in 2D Layered PtSe₂. *Adv. Mater.* **2016**, *28* (12), 2399–2407.
- (6) Ciarrocchi, A.; Avsar, A.; Ovchinnikov, D.; Kis, A. Thickness-Modulated Metal-to-Semiconductor Transformation in a Transition Metal Dichalcogenide. *Nat. Commun.* **2018**, *9* (1), 919.
- (7) Chia, X.; Adriano, A.; Lazar, P.; Sofer, Z.; Luxa, J.; Pummer, M. Layered Platinum Dichalcogenides (PtS₂, PtSe₂, and PtTe₂) Electrocatalysis: Monotonic Dependence on the Chalcogen Size. *Adv. Funct. Mater.* **2016**, *26* (24), 4306–4318.
- (8) Lin, S.; Liu, Y.; Hu, Z.; Lu, W.; Mak, C. H.; Zeng, L.; Zhao, J.; Li, Y.; Yan, F.; Tsang, Y. H.; et al. Tunable Active Edge Sites in PtSe₂ Films Towards Hydrogen Evolution Reaction. *Nano Energy* **2017**, *42*, 26–33.
- (9) Avsar, A.; Ciarrocchi, A.; Pizzochero, M.; Unuchek, D.; Yazyev, O. V.; Kis, A. Defect Induced, Layer-Modulated Magnetism in Ultrathin Metallic PtSe₂. *Nat. Nanotechnol.* **2019**, *14* (7), 674–678.
- (10) Avsar, A.; Cheon, C.-Y.; Pizzochero, M.; Tripathi, M.; Ciarrocchi, A.; Yazyev, O. V.; Kis, A. Probing Magnetism in Atomically Thin Semiconducting PtSe₂. *Nat. Commun.* **2020**, *11* (1), 4806.
- (11) Xu, H.; Zhang, H.; Liu, Y.; Zhang, S.; Sun, Y.; Guo, Z.; Sheng, Y.; Wang, X.; Luo, C.; Wu, X.; et al. Controlled Doping of Wafer-Scale PtSe₂ Films for Device Application. *Adv. Funct. Mater.* **2019**, *29* (4), 1805614.
- (12) Yu, X.; Yu, P.; Wu, D.; Singh, B.; Zeng, Q.; Lin, H.; Zhou, W.; Lin, J.; Suenaga, K.; Liu, Z.; Wang, Q. J.; et al. Atomically Thin Noble Metal Dichalcogenide: a Broadband Mid-Infrared Semiconductor. *Nat. Commun.* **2018**, *9* (1), 1545.
- (13) García de Abajo, F. J. Optical Excitations in Electron Microscopy. *Rev. Mod. Phys.* **2010**, *82* (1), 209–275.
- (14) Koppens, F. H. L.; Chang, D. E.; García de Abajo, F. J. Graphene Plasmonics: A Platform for Strong Light–Matter Interactions. *Nano Lett.* **2011**, *11* (8), 3370–3377.
- (15) Lourenço-Martins, H.; Gérard, D.; Kociak, M. Optical Polarization Analogue in Free Electron Beams. *Nat. Phys.* **2021**, *17* (5), 598–603.
- (16) Wang, G.; Wang, Z.; McEvoy, N.; Fan, P.; Blau, W. J. Layered PtSe₂ for Sensing, Photonic, and (Opto-)Electronic Applications. *Adv. Mater.* **2021**, *33* (1), 2004070.
- (17) Zeng, L.; Lin, S.; Lou, Z.; Yuan, H.; Long, H.; Li, Y.; Lu, W.; Lau, S. P.; Wu, D.; Tsang, Y. H. Ultrafast and Sensitive Photodetector Based on a PtSe₂/Silicon Nanowire Array Heterojunction with a Multiband Spectral Response from 200 to 1550 nm. *NPG Asia Mater.* **2018**, *10* (4), 352–362.
- (18) Wang, L.; Zhang, S.; McEvoy, N.; Sun, Y.-y.; Huang, J.; Xie, Y.; Dong, N.; Zhang, X.; Kislyakov, I. M.; Nunzi, J.-M.; et al. Nonlinear Optical Signatures of the Transition from Semiconductor to Semimetal in PtSe₂. *Laser Photonics Rev.* **2019**, *13* (8), 1900052.
- (19) Jakhar, A.; Kumar, P.; Moudgil, A.; Dhyani, V.; Das, S. Optically Pumped Broadband Terahertz Modulator Based on Nanostructured PtSe₂ Thin Films. *Adv. Opt. Mater.* **2020**, *8* (7), 1901714.
- (20) Chernikov, A.; Berkelbach, T. C.; Hill, H. M.; Rigosi, A.; Li, Y.; Aslan, O. B.; Reichman, D. R.; Hybertsen, M. S.; Heinz, T. F. Exciton Binding Energy and Nonhydrogenic Rydberg Series in Monolayer WS₂. *Phys. Rev. Lett.* **2014**, *113* (7), 076802.
- (21) Ugeda, M. M.; Bradley, A. J.; Shi, S.-F.; da Jornada, F. H.; Zhang, Y.; Qiu, D. Y.; Ruan, W.; Mo, S.-K.; Hussain, Z.; Shen, Z.-X.; et al. Giant Bandgap Renormalization and Excitonic Effects in a Monolayer Transition Metal Dichalcogenide Semiconductor. *Nat. Mater.* **2014**, *13* (12), 1091–1095.
- (22) Haastrup, S.; Strange, M.; Pandey, M.; Deilmann, T.; Schmidt, P. S.; Hinsche, N. F.; Gjerding, M. N.; Torelli, D.; Larsen, P. M.; Riis-Jensen, A. C.; et al. The Computational 2D Materials Database: High-Throughput Modeling and Discovery of Atomically Thin Crystals. *2D Mater.* **2018**, *5* (4), 042002.
- (23) Thygesen, K. S. Calculating Excitons, Plasmons, and Quasiparticles in 2D Materials and van der Waals Heterostructures. *2D Mater.* **2017**, *4* (2), 022004.
- (24) Sajjad, M.; Singh, N.; Schwingenschlögl, U. Strongly Bound Excitons in Monolayer PtS₂ and PtSe₂. *Appl. Phys. Lett.* **2018**, *112* (4), 043101.
- (25) Bae, S.; Nah, S.; Lee, D.; Sajjad, M.; Singh, N.; Kang, K. M.; Kim, S.; Kim, G.-J.; Kim, J.; Baik, H.; et al. Exciton-Dominated Ultrafast Optical Response in Atomically Thin PtSe₂. *Small* **2021**, *17* (45), 2103400.
- (26) Ghosh, B.; Alessandro, F.; Zappia, M.; Brescia, R.; Kuo, C.-N.; Lue, C. S.; Chiarello, G.; Politano, A.; Caputi, L. S.; Agarwal, A.; et al. Broadband Excitation Spectrum of Bulk Crystals and Thin Layers of PtTe₂. *Phys. Rev. B* **2019**, *99* (4), 045414.
- (27) Habenicht, C.; Knupfer, M.; Büchner, B. Investigation of the Dispersion and the Effective Masses of Excitons in Bulk 2H–MoS₂ Using Transition Electron Energy-Loss Spectroscopy. *Phys. Rev. B* **2015**, *91* (24), 245203.
- (28) Kramberger, C.; Hambach, R.; Giorgetti, C.; Rümeli, M. H.; Knupfer, M.; Fink, J.; Büchner, B.; Reining, L.; Einarsson, E.; Maruyama, S.; et al. Linear Plasmon Dispersion in Single-Wall Carbon Nanotubes and the Collective Excitation Spectrum of Graphene. *Phys. Rev. Lett.* **2008**, *100* (19), 196803.
- (29) Wachsmuth, P.; Hambach, R.; Kinyanjui, M. K.; Guzzo, M.; Benner, G.; Kaiser, U. High-Energy Collective Electronic Excitations in Free-Standing Single-Layer Graphene. *Phys. Rev. B* **2013**, *88* (7), 075433.
- (30) Roth, F.; König, A.; Fink, J.; Büchner, B.; Knupfer, M. Electron Energy-Loss Spectroscopy: a Versatile Tool for the Investigations of Plasmonic Excitations. *J. Electron Spectrosc. Relat. Phenom.* **2014**, *195*, 85–95.
- (31) Qi, R.; Li, N.; Du, J.; Shi, R.; Huang, Y.; Yang, X.; Liu, L.; Xu, Z.; Dai, Q.; Yu, D.; et al. Four-Dimensional Vibrational Spectroscopy for Nanoscale Mapping of Phonon Dispersion in BN Nanotubes. *Nat. Commun.* **2021**, *12* (1), 1179.
- (32) Qi, R.; Shi, R.; Li, Y.; Sun, Y.; Wu, M.; Li, N.; Du, J.; Liu, K.; Chen, C.; Chen, J.; et al. Measuring Phonon Dispersion at an Interface. *Nature* **2021**, *599* (7885), 399–403.
- (33) Li, J.; Kolekar, S.; Ghorbani-Asl, M.; Lehnert, T.; Biskupek, J.; Kaiser, U.; Krashennnikov, A. V.; Batzill, M. Layer-Dependent Band Gaps of Platinum Dichalcogenides. *ACS Nano* **2021**, *15* (8), 13249–13259.
- (34) Wachsmuth, P.; Hambach, R.; Benner, G.; Kaiser, U. Plasmon Bands in Multilayer Graphene. *Phys. Rev. B* **2014**, *90* (23), 235434.
- (35) Cudazzo, P.; Sponza, L.; Giorgetti, C.; Reining, L.; Sottile, F.; Gatti, M. Exciton Band Structure in Two-Dimensional Materials. *Phys. Rev. Lett.* **2016**, *116* (6), 066803.
- (36) Onida, G.; Reining, L.; Rubio, A. Electronic Excitations: Density-Functional versus Many-Body Green's-Function Approaches. *Rev. Mod. Phys.* **2002**, *74* (2), 601–659.
- (37) Onida, G.; Reining, L.; Godby, R. W.; Del Sole, R.; Andreoni, W. Ab Initio Calculations of the Quasiparticle and Absorption Spectra of Clusters: The Sodium Tetramer. *Phys. Rev. Lett.* **1995**, *75* (5), 818–821.

- (38) Rohlfling, M.; Louie, S. G. Electron-Hole Excitations in Semiconductors and Insulators. *Phys. Rev. Lett.* **1998**, *81* (11), 2312–2315.
- (39) Nerl, H. C.; Winther, K. T.; Hage, F. S.; Thygesen, K. S.; Houben, L.; Backes, C.; Coleman, J. N.; Ramasse, Q. M.; Nicolosi, V. Probing the Local Nature of Excitons and Plasmons in Few-Layer MoS₂. *npj 2D Mater. Appl.* **2017**, *1* (1), 2.
- (40) Ferrell, R. A. Predicted Radiation of Plasma Oscillations in Metal Films. *Phys. Rev.* **1958**, *111* (5), 1214–1222.
- (41) Andersen, K.; Thygesen, K. S. Plasmons in Metallic Monolayer and Bilayer Transition Metal Dichalcogenides. *Phys. Rev. B* **2013**, *88* (15), 155128.
- (42) Andersen, K.; Jacobsen, K. W.; Thygesen, K. S. Spatially Resolved Quantum Plasmon Modes in Metallic Nano-Films from First-Principles. *Phys. Rev. B* **2012**, *86* (24), 245129.
- (43) Hedin, L. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.* **1965**, *139* (3A), A796–A823.
- (44) Hybertsen, M. S.; Louie, S. G. Electron Correlation in Semiconductors and Insulators: Band Gaps and Quasiparticle Energies. *Phys. Rev. B* **1986**, *34* (8), 5390–5413.
- (45) Aulbur, W. G.; Jönsson, L.; Wilkins, J. W. Quasiparticle Calculations in Solids. In *Solid State Physics*; Ehrenreich, H., Spaepen, F., Eds.; Elsevier B.V. & Academic Press: San Diego, 2000; Vol. 54, pp 1–218.
- (46) Golze, D.; Dvorak, M.; Rinke, P. The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy. *Front. Chem.* **2019**, *7*, 377.
- (47) Yan, J.; Jacobsen, K. W.; Thygesen, K. S. Conventional and Acoustic Surface Plasmons on Noble Metal Surfaces: A Time-Dependent Density Functional Theory Study. *Phys. Rev. B* **2012**, *86* (24), 241404.
- (48) Lo, P.-Y.; Peng, G.-H.; Li, W.-H.; Yang, Y.; Cheng, S.-J. Full-Zone Valley Polarization Landscape of Finite-Momentum Exciton in Transition Metal Dichalcogenide Monolayers. *Phys. Rev. Res.* **2021**, *3* (4), 043198.
- (49) Gjerding, M.; Skovhus, T.; Rasmussen, A.; Bertoldo, F.; Larsen, A. H.; Mortensen, J. J.; Thygesen, K. S. Atomic Simulation Recipes: A Python Framework and Library for Automated Workflows. *Comput. Mater. Sci.* **2021**, *199*, 110731.
- (50) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dulak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; et al. Electronic Structure Calculations with GPAW: a Real-Space Implementation of the Projector Augmented-Wave Method. *J. Phys. Condens. Mater.* **2010**, *22* (25), 253202.
- (51) Gjerding, M. N.; et al. Recent Progress of the Computational 2D Materials Database (C2DB). *2D Mater.* **2021**, *8*, 044002.
- (52) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.

Recommended by ACS

Sub-10 fs Intervalley Exciton Coupling in Monolayer MoS₂ Revealed by Helicity-Resolved Two-Dimensional Electronic Spectroscopy

Lawson T. Lloyd, Gregory S. Engel, et al.

JUNE 07, 2021

ACS NANO

READ 

Up- and Down-Conversion between Intra- and Intervalley Excitons in Waveguide Coupled Monolayer WSe₂

Yueh-Chun Wu, Jun Yan, et al.

JULY 20, 2020

ACS NANO

READ 

Momentum-Dark Intervalley Exciton in Monolayer Tungsten Diselenide Brightened via Chiral Phonon

Zhipeng Li, Su-Fei Shi, et al.

NOVEMBER 25, 2019

ACS NANO

READ 

Dark-Exciton Driven Energy Funneling into Dielectric Inhomogeneities in Two-Dimensional Semiconductors

Haowen Su, Milan Delor, et al.

MARCH 16, 2022

NANO LETTERS

READ 

Get More Suggestions >

11.4 Paper [IV]: Predicting and machine learning structural instabilities in 2D materials,

Simone Manti, Mark Kamper Svendsen, Nikolaj R. Knøsgaard, Peder M. Lyngby, Kristian S. Thygesen,

“Predicting and machine learning structural instabilities in 2D materials,”

Submitted to npj Computational Materials: [arXiv:2201.08091](https://arxiv.org/abs/2201.08091).

Predicting and machine learning structural instabilities in 2D materials

Simone Manti^{§,*}, Mark Kamper Svendsen[§], Nikolaj R. Knøsgaard[§], Peder M. Lyngby[§], and Kristian S. Thygesen^{§,†}

[§]CAMD, Computational Atomic-Scale Materials Design, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby Denmark

[†]Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, DK - 2800 Kongens Lyngby, Denmark

*Corresponding author: smanti@fysik.dtu.dk

Abstract. We address the problem of predicting the zero temperature dynamical stability (DS) of a periodic crystal without computing its full phonon band structure. Using data for two-dimensional (2D) crystals, we first present statistical evidence that DS can be inferred with good reliability from the phonon frequencies at the center and boundary of the Brillouin zone (BZ). This analysis represents a validation of the DS test employed by the Computational 2D Materials Database (C2DB). For 137 dynamically unstable 2D crystals, we displace the atoms along an unstable mode and relax the structure. This procedure yields a dynamically stable crystal in 49 cases. The elementary properties of these new structures are characterised using the C2DB workflow, and it is found that their properties can differ significantly from those of the original unstable crystals, e.g. band gaps are opened by 0.3 eV on average. All the crystal structures and properties are available in the C2DB. Finally, we train a classification model on the DS data for 3295 2D materials in the C2DB using a representation encoding the electronic structure of the crystal. We obtain an excellent receiver operating characteristic (ROC) curve with an area under the curve (AUC) of 0.89, showing that the classification model can drastically reduce computational efforts in high-throughput studies.

1. Introduction

Computational materials discovery aims at identifying novel materials for specific applications, often employing first principles methods such as density functional theory (DFT) [1]. The potential of a given material for the targeted application is usually evaluated based on elementary properties of the crystal, such as the electronic band gap, the optical absorption spectrum, or the magnetic order. Such properties can be highly sensitive to even small distortions of the lattice that reduce the symmetry of the crystal, and it is therefore important to develop efficient methods for identifying and accounting for such distortions.

Lattice distortions can be classified according to their periodicity relative to the primitive cell of the crystal. Local instabilities conserve the periodicity of the crystal, i.e. they do not enlarge the number of atoms in the primitive cell. Other distortions, accompanied by a modulation of the electronic density known as charge density wave (CDW) [2], lead to an enlargement of the period of the crystal, which can be either commensurate or incommensurate with the high-symmetry phase. A classical example is the Peierls instability in one-dimensional [3] and two-dimensional (2D) [4] systems, where a gap is opened in the CDW state. A universal microscopic theory of the CDW phase is still missing due to the many possible and intertwined driving mechanisms, e.g. electron-phonon interaction [5], Fermi surface nesting, or phonon-phonon interactions [6], which makes a precise clear-cut definition of the CDW phase difficult. In addition, the CDW state is sensitive to external effects such as temperature and doping [7]. As a testimony to the complexity of the problem, different models and concepts are used to describe the CDW phase depending on the dimensionality of the material [8, 9, 10, 11, 12].

The last few years have witnessed an increased interest in CDW states of 2D materials. For example, CDW physics is believed to govern the transition from the trigonal prismatic T-phase to the lower symmetry T'-phase in monolayer MoS₂ [13] as well as the plethora of temperature dependent phases in monolayers of NbSe₂ [14, 15], TaS₂ [16, 17], TaSe₂ [18, 19], and TiSe₂ [20, 21]. In addition, a number of recent studies have investigated the possibility to control CDW phase transitions. For instance, the T-phase of monolayer MoS₂ can be stabilized by argon bombardment [22], exposure to electron beams [13], or Li-ion intercalation [23]. Similar results have been reported for MoTe₂ [24].

Regardless of the fundamental origin of possible lattice distortions, it remains of great practical importance to devise efficient schemes that makes it possible to verify whether or not a given structure is dynamically stable (DS), i.e. whether it represents

a local minimum of the potential energy surface. Structures that are not DS are frequently generated in computational studies, e.g. when a structure is relaxed under symmetry constraints or the chosen unit cell is too small to accommodate the stable phase. Tests for DS are rarely performed in large-scale discovery studies, because there is no established way of doing it apart from calculating the full phonon band structure [25], which is a time-consuming task. At the same time, the importance of incorporating such tests is in fact unclear; that is, it is not known how much symmetry-breaking distortions generally influence the properties of a materials.

A straightforward strategy to generate potentially stable structures from dynamically unstable ones, is to displace the atoms along an unstable phonon mode using a supercell that can accommodate the distortion. This approach has previously been adopted to explore structural distortions in metallic system [26] in bulk perovskites [27, 28] and one-dimensional organometallic chains [29]. However, systematic studies of structural instabilities in 2D materials, have so far been lacking.

In this work, we perform a systematic study of structural distortions across a broad class of 2D crystals, and explore a machine learning-based approach to DS classification. Throughout, we focus on the most common case of small-period, commensurate distortions that can be accommodated in a 2×2 repetition of the primitive cell of the high-symmetry phase. We shall refer to the test for the occurrence of such distortions as the Center and Boundary Phonon (CBP) protocol. The motivation behind the present work is fourfold: (i) To assess the reliability of the CBP protocol (which is currently used for DS classification in the Computational 2D Materials Database (C2DB)[30, 31]). (ii) To elucidate the effect of symmetry-breaking distortions on the basic electronic properties of crystals. (iii) To obtain the DS phases of a set of dynamically unstable 2D materials that were originally generated by combinatorial lattice decoration, and make them available to the community via the C2DB. (iv) To explore the viability of a machine learning based classification scheme for predicting DS using input from a DFT calculation of the undistorted high-symmetry phase.

The paper is structured as follows. In Section 2 we describe the CBP protocol. We first benchmark the CBP protocol against full phonon band structure calculations and evaluate its statistical success rate. For 137 dynamically unstable 2D materials, we further analyse how the small-period distortions that stabilise the materials influence their electronic properties. Section 3 concludes the paper.

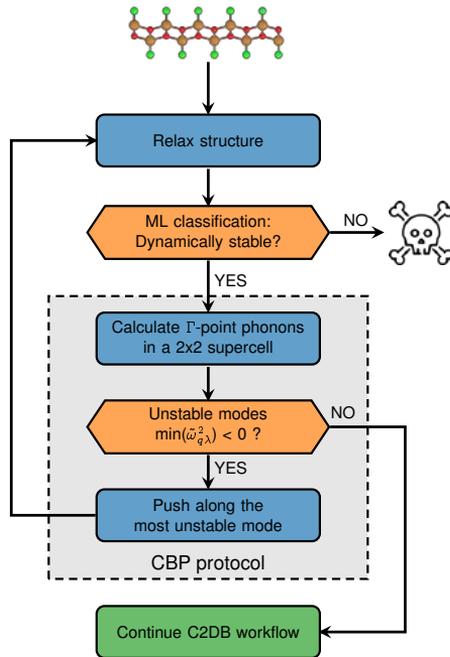


Figure 1. Workflow to test for dynamical stability of (two-dimensional) crystals in high-throughput computational studies. In a first step, a machine learning classification algorithm is used to filter out unstable crystals at minimal computational cost. The stability of the remaining crystals are assessed using the CBP protocol. Only materials predicted to be stable by the CBP protocol continue to the characterisation workflow.

2. Results and discussion

This section presents and discusses the results of the CBP protocol and the connection with the machine learning predictions. Together they will increase the success rate of finding dynamically stable material within the C2DB workflow (see Figure (1)).

2.1. The CBP protocol: Stability test

Given a material that has been relaxed in some unit cell (from hereon referred to as the primitive unit cell), the CBP protocol proceeds by evaluating the stiffness tensor of the material and the Hessian matrix of a supercell obtained by repeating the primitive cell 2×2 times. In the current work, the stiffness tensor is calculated as a finite difference of the stress under an applied strain, while the Hessian matrix is calculated as a finite difference of the forces on all the atoms of the 2×2 supercell under displacement of the atoms in one primitive unit cell (this is equivalent to calculating

the phonons at the center and specific high symmetry points at boundary of the BZ of the primitive cell, see Fig. (4)). Next, the stiffness tensor and the Hessian matrix are diagonalised, and the eigenvalues are used to infer a structural stability. A negative eigenvalue of the stiffness tensor indicates an instability of the lattice (the shape of the unit cell) while a negative eigenvalue of the 2×2 Hessian signals an instability of the atomic structure. The obvious question here, is whether it suffices to consider the Hessian of the 2×2 supercell, or equivalently consider the phonons at the BZ center and boundaries.

We can distinguish three possible outcomes when comparing the CBP protocol against full phonon calculations (see Figure (2)), namely a true positive result where the material is predicted to be stable by both the CBP protocol and a full phonon calculation, a true negative result where the material is predicted to be unstable by both the CBP protocol and a full phonon calculation, and a false positive result where the material is predicted to be stable by the CBP protocol, but a full phonon calculation would find the material to be unstable. The false positive case occurs when a material is stable in a 2×2 supercell, but unstable if allowed to distort in a larger cell. Our results show that such large-period distortions which do not show as distortions in a 2×2 cell, are relatively rare (see Section 2.3). We note that the case of a false negative is not possible, because a material that is unstable in a 2×2 cell is de facto unstable.

2.2. The CBP protocol: Structure generation

Here we outline a simple procedure to generate distorted and potentially stable structures from an initial dynamically unstable structure. The basic idea is to displace the atoms along an unstable phonon mode followed by a relaxation. In practice, the unstable mode is obtained as the eigen function corresponding to a negative eigenvalue of the Hessian matrix of the 2×2 supercell. The procedure is illustrated in Figure (3) for the well known T-T' phase transition of MoS₂ [13]. The left panel shows the atomic structure and phonon band structure of monolayer MoS₂ in the T-phase. Both the primitive unit cell (black) and the 2×2 supercell (orange) are indicated. The CBP method identifies an unstable mode at the BZ boundary (M point). After displacing the atoms along the unstable mode, a distorted structure is obtained, which after relaxation leads to the dynamically stable T'-phase of MoS₂ shown in the right panel.

In this work, we have applied the CBP protocol systematically to 137 dynamically unstable 2D materials. The 137 monolayers were selected from the C2DB according to the following two criteria: First, to ensure that all materials are chemically "reasonable", only

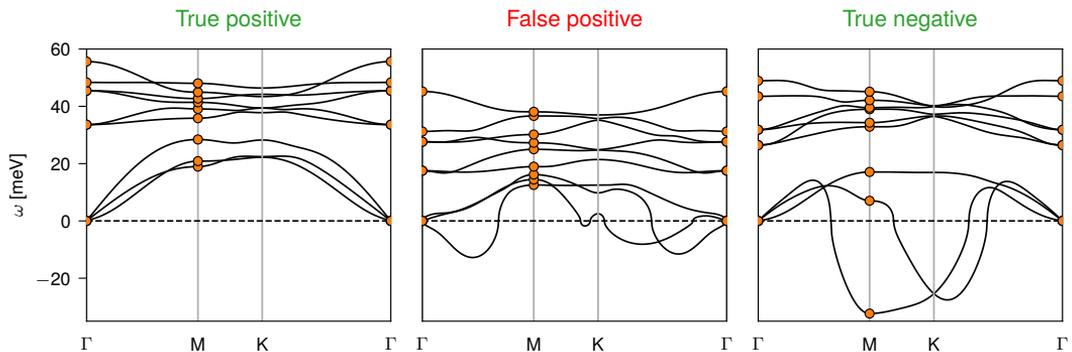


Figure 2. Phonon band structure for monolayer MoS₂ in the H-phase (left), NbSSe (middle), and MoS₂ in the T-phase (right). Note that imaginary phonon frequencies are represented by negative values. The CBP protocol (orange dots) is sufficient to conclude that a material is dynamically stable (unstable) in the situations depicted in the left (right) panels. In contrast, when the relevant distortion requires a supercell larger than a 2×2 , and the phonon frequencies are real at the center and boundary of the BZ, the CBP protocol will result in a false positive result.

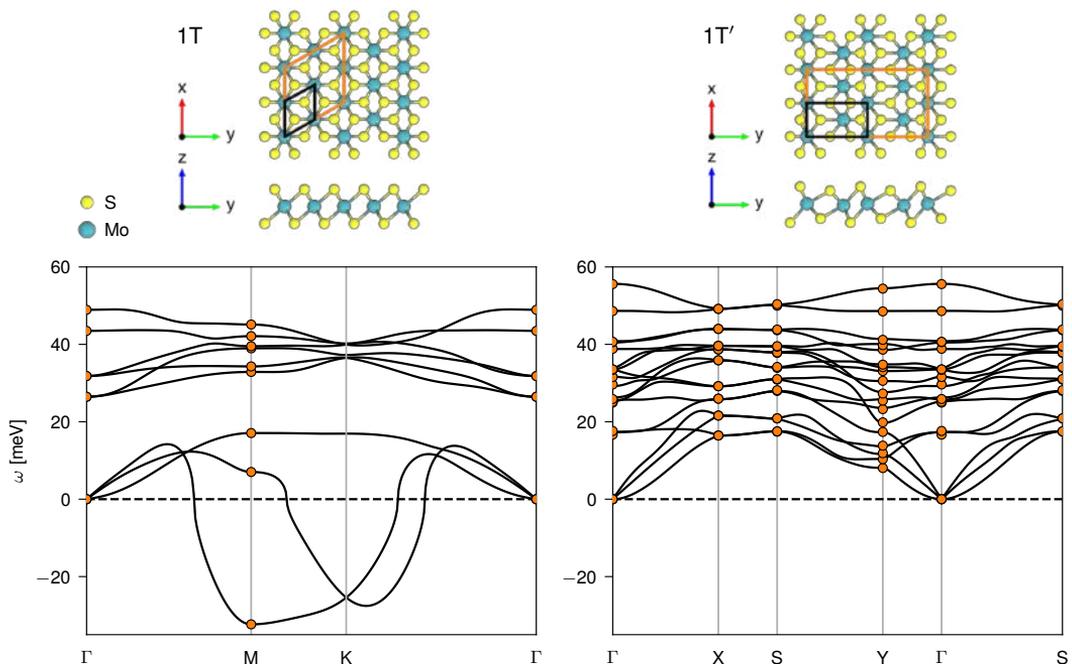


Figure 3. The CBP protocol captures the instability of the T-phase (left) of MoS₂. Both the primitive unit cell (black) and the 2×2 supercell (orange) are shown. Displacing the atoms along the unstable TA mode at the M-point ($\mathbf{q} = (\frac{1}{2}, 0)$), which can be accommodated in the 2×1 supercell, and subsequently relaxing the structure results in the dynamically stable T'-phase (right).

materials with a low formation energy were selected. Specifically, we require that $\Delta H_{\text{hull}} < 0.2$ eV/atom, where ΔH_{hull} is the energy above the convex hull defined by the most stable (possibly mixed) bulk phases

of the relevant composition [32, 31]. Secondly, we consider only materials with exactly one unstable mode, i.e. one negative eigenvalue of the Hessian matrix at a given q -point. We stress that the latter condition is

not strictly necessary but was adopted here to limit the number of materials. When two or more unstable modes exist there is not a unique way to distort the structure. One possibility is to push along the linear combination of modes yielding the distorted structure with the highest symmetry [26]. However, it is not clear that imposing high symmetry is the best strategy for finding a DS structure. For the few cases with multiple unstable modes that we have analysed, we have found that pushing along the most unstable eigenmode (the one with the most negative eigenvalue) often yields a DS structure, like in the case of T-MoS₂.

For the 137 dynamically unstable materials we displaced the atoms along the (unique) unstable mode. The size of the displacement was chosen such that the maximum atomic displacement was exactly 0.1 Å. This displacement size was chosen based on the MoS₂ example discussed above, where it results in a minimal number of subsequent relaxation steps. A smaller value does not guarantee that the system leaves the saddle point, while a larger value creates a too large distortion resulting in additional relaxation steps. During relaxation the unit cell was allowed to change with no symmetry constraints and the relaxation was stopped when the forces on all atoms were below 0.01 eV / Å.

2.3. Assessment of the CBP protocol

To test the validity the CBP protocol, we have performed full phonon calculations for a set of 20 monolayers predicted as dynamically stable by the CBP protocol. The 20 materials were randomly selected from the C2DB and cover 7 different crystal structures. Out of the 20 materials 10 are metals and 10 are insulators/semiconductors. The calculated phonon band structures are reported in the supporting information (SI). For all materials, the phonon frequencies obtained with the CBP protocol equal the frequencies of the full phonon band structure at the q -points $\mathbf{q} \in \{(0,0), (\frac{1}{2},0), (0,\frac{1}{2}), (\frac{1}{2},\frac{1}{2})\}$. This is expected as the phonons at these q -points can be accommodated by the 2×2 supercell.

Within the set of 20 materials, we find three False-positive cases, namely CoTe₂, NbSSe, and TaTe₂. These materials exhibit unstable modes (imaginary frequencies or equivalently negative force constant eigenvalues) in the interior of the BZ (NbSSe and TaTe₂) or at the K-point (CoTe₂), while all phonon frequencies at the q -points covered by the CBP protocol, are real. A simple interpolation of the frequencies is not enough to catch an instability at internal points of BZ (see Figure S1) because a supercell larger than a 2x2 is needed to catch these distortions, like in the case of 2H-CoTe₂ (see Figure S2). This relatively low percentage of False-positives

in our representative samples is consistent with the work by Mounet et al. [25] who computed the full phonon band structure of 258 monolayers predicted to be (easily) exfoliable from known bulk compounds. Applying the CBP protocol to their data yields 14 False-positive cases; half of these are transition metal dichalcogenides (TMDs) with Co, Nb or Ta.

We note that the small imaginary frequencies in the out of plane modes around the Γ -point seen in some of the phonon band structures are not distortions, but are rather due to the interpolation of the dynamical matrix. In particular, these artifacts occur because of the broken crystal point-group symmetry in the force constant matrix and they will vanish if a larger supercell is used or the rotational sum rule is imposed [33, 34] or higher-order multipolar interactions are included [35].

2.4. Stable distorted monolayers

The 137 dynamically unstable materials, which were selected from the C2DB according to the criteria described in Section 2.2, can be divided into two groups depending on whether the eigenvalues of the Hessian at the wave vectors $q_x = (\frac{1}{2}, 0)$, $q_y = (0, \frac{1}{2})$ and $q_{xy} = (\frac{1}{2}, \frac{1}{2})$, are equal or not. Equality of the eigenvalues implies an isotropic Hessian. For such materials, we generate distorted structures by displacing the atoms along the unstable mode at $q_x = (\frac{1}{2}, 0)$, followed by relaxation in a 2×1 supercell. In the case of an anisotropic Hessian, in general, it may exist a particular combination of q -vectors that stabilizes the system. Here, we were interested in finding a general method to generate DS structures in a high-throughput way and therefore, we decided to displace only at $q_{xy} = (\frac{1}{2}, \frac{1}{2})$ in a 2×2 supercell.

After atomic displacement and subsequent relaxation, the CBP protocol was applied again to test for dynamical stability of the distorted structures. Histograms of the minimum eigenvalue, $\min(\tilde{\omega}_{\mathbf{q}\lambda}^2)$, of the Hessian matrix at \mathbf{q} for the unstable mode λ , are shown in Figure (4) with the materials before and after atomic displacement shown in the upper and lower panels, respectively. Negative eigenvalues, corresponding to unstable materials, are shown in red while positive eigenvalues are shown in green. We removed the three translational modes with eigenvalues close to zero before extracting $\min(\tilde{\omega}_{\mathbf{q}\lambda}^2)$. Out of the 137 unstable materials, 49 become dynamically stable (according to the CBP protocol). By far the highest success rate for generating stable crystals was found for the isotropic materials (left panel), where 43 out of 91 materials became stable while only 6 out of the 43 anisotropic materials became stable. In principle, the procedure can be applied many times to increase the number of the DS structures, while here, we applied the protocol only

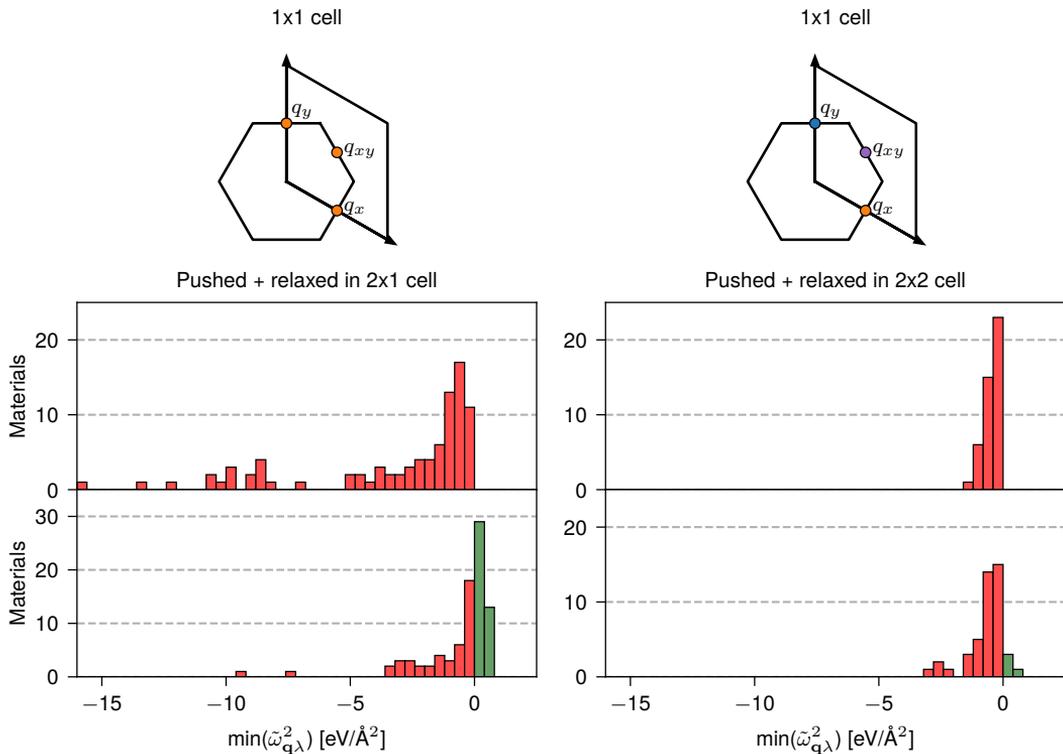


Figure 4. The 137 dynamically unstable 2D materials studied in this work can be divided into two groups depending on whether the minimum eigenvalues of the Hessian matrix $\tilde{\omega}_{\mathbf{q}\lambda}$ at $\mathbf{q} = \{(\frac{1}{2}, 0), (0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\}$ are equal (left panel orange) or different (right panel different colors). For the first group of materials, the atoms are displaced along the mode at $\mathbf{q} = (\frac{1}{2}, 0)$ and relaxed in a 2×1 supercell. Starting from 91 unstable materials (red) in the upper panel, 43 (green) become stable (lower panel). For the second group, displacing along $\mathbf{q} = (\frac{1}{2}, \frac{1}{2})$ and relaxing in a 2×2 supercell, yields a dynamically stable structure in 6/46 cases.

once for computational reasons. An example where applying the protocol two times is the case of the 1T-TiSe₂ monolayer in the SI. In that case, the material is first displaced at $q_{xy} = (\frac{1}{2}, \frac{1}{2})$ in a 2×2 supercell. Then the unstable material obtained is again displaced along one of the two degenerate modes at Γ and the final structure with the experimentally known CDW state [20] is obtained.

A wide range of elementary properties of the 49 distorted, dynamically stable materials were computed using the C2DB workflow (see Table 1 in [31] for a complete list of the properties). The atomic structures together with the calculated properties are available in the C2DB. Table 1 provides an overview of the symmetries, minimal Hessian eigenvalues, total energies, and electronic band gap of the 49 materials before and after the distortion.

Apart from the reduction in symmetry, the distortion also lowers the total energy of the materials.

An important descriptor for the thermodynamic stability of a material is the energy above the convex hull, ΔH_{hull} . Figure 5 shows a plot of ΔH_{hull} before and after the distortion of the 49 materials. The reduction in energy upon distortion ranges from 0 to 0.2 eV/atom. In fact, several of the materials come very close to the convex hull and some even fall onto the hull, indicating their global thermodynamic stability (at $T = 0$ K) with respect to the reference bulk phases. We note that all DFT energies, including the reference bulk phases, were calculated using the PBE xc-functional, which does not account for van der Waals interactions. Accounting for the vdW interactions will downshift the energies of layered bulk phases and thus increase ΔH_{hull} for the monolayers slightly. This effect will, however, not influence the relative stability of the pristine and distorted monolayers, which is the main focus of the current work. Another characteristic trend observed is the opening/increase of the electronic

Material	Space group - Wyckoff		ΔH_{hull} [eV/atom]		$\min(\tilde{\omega}_{\mathbf{q}\lambda}^2)$ [eV/Å ²]		$e_{\text{gap}}^{\text{PBE}}$ [eV]	
	before	after	before	after	before	after	before	after
AgBr ₂	164-bd	14-ae	0.05	0.00	-1.01	0.00	0.00	0.00
AgCl ₂	164-bd	14-ae	0.05	0.00	-2.60	0.06	0.00	0.00
AsClTe	156-ac	156-abc	0.19	0.02	-0.51	0.25	1.29	1.48
CdBr	164-ad	13-d	0.18	0.07	-1.00	0.03	0.00	1.28
CdCl	164-d	164-cd	0.17	0.04	-0.74	0.17	0.00	1.67
CoSe	164-bd	25-acgh	0.05	0.03	-4.82	0.09	0.00	0.00
CrBrCl	156-abc	7-a	0.11	0.06	-0.98	0.15	0.00	0.64
CrBr ₂	164-bd	14-de	0.10	0.06	-0.59	0.07	0.00	0.49
CrCl ₂	164-bd	14-be	0.11	0.05	-1.87	0.15	0.00	0.76
CrSSe	156-abc	6-ab	0.15	0.09	-9.75	0.71	0.00	0.00
CrS ₂	164-bd	11-e	0.18	0.05	-15.62	0.74	0.00	0.00
CrSe ₂	164-bd	11-e	0.14	0.05	-12.17	0.77	0.00	0.00
CrTe ₂	164-bd	11-e	0.02	0.01	-1.71	0.08	0.00	0.00
CrPS ₃	162-dek	11-ef	0.09	0.03	-3.13	0.00	0.00	0.34
FePSe ₃	162-dek	11-fe	0.13	0.12	-0.42	0.04	0.13	0.13
FeSe ₂	187-bi	187-eg	0.15	0.00	-2.04	0.00	0.00	0.00
HfBrCl	156-abc	6-ab	0.14	0.03	-9.61	0.40	0.00	0.82
HfBrI	156-abc	6-ab	0.22	0.05	-10.41	0.39	0.00	0.73
HfBr ₂	164-bd	11-e	0.14	0.04	-10.21	0.41	0.00	0.8
HfCl ₂	164-bd	11-e	0.14	0.04	-8.51	0.38	0.00	0.85
HgSe	164-d	10-mn	0.11	0.04	-0.83	0.04	0.08	0.37
HgTe	164-d	10-mn	0.11	0.04	-0.61	0.04	0.08	0.37
InTe	164-d	129-c	0.18	0.10	-0.41	0.23	0.00	0.00
InBrSe	59-ab	59-ab	0.03	0.02	-0.50	0.04	1.23	1.23
InSe	187-hi	187-hi	0.00	0.00	-0.28	0.01	1.39	1.39
MoSeTe	156-abc	6-ab	0.20	0.08	-10.67	0.69	0.00	0.00
MoTe ₂	164-bd	11-e	0.17	0.01	-13.23	0.67	0.00	0.00
NbS ₂	187-bi	187-eg	0.00	0.00	-1.08	0.06	0.00	0.00
NbTe ₂	187-bi	187-fg	0.00	0.00	-0.37	0.51	0.00	0.00
PdI ₂	164-bd	14-ae	0.17	0.03	-0.56	0.12	0.00	0.59
RhI ₂	164-bd	164-ad	0.17	0.17	-0.64	0.04	0.00	0.00
RhO ₂	164-bd	11-e	0.16	0.15	-2.19	0.65	0.00	0.00
RhTe ₂	164-bd	11-e	0.11	0.07	-1.67	0.36	0.00	0.13
Scl ₃	162-dk	14-e	0.00	0.00	-0.25	0.02	1.85	1.85
TiBrCl	156-ac	6-ab	0.05	0.00	-9.15	0.52	0.00	0.29
TiBr ₂	164-bd	11-e	0.06	0.04	-0.54	0.30	0.00	0.12
TiCl ₂	164-bd	11-e	0.11	0.00	-9.99	0.53	0.00	0.32
TiO ₂	164-bd	11-e	0.14	0.12	-1.96	0.47	2.70	2.85
TiS ₂	187-bi	31-a	0.14	0.14	-0.49	0.00	0.73	0.79
TiPSe ₃	162-dek	2-i	0.16	0.00	-1.24	0.00	0.00	0.00
VTe ₂	164-bd	11-e	0.02	0.00	-1.05	0.38	0.00	0.00
ZrBrCl	156-abc	6-ab	0.12	0.02	-8.05	0.40	0.00	0.59
ZrBrI	156-abc	6-ab	0.15	0.02	-8.89	0.38	0.00	0.48
ZrBr ₂	164-bd	11-e	0.11	0.00	-8.65	0.47	0.00	0.59
ZrClI	156-abc	6-ab	0.19	0.07	-8.69	0.26	0.00	0.48
ZrCl ₂	164-bd	11-e	0.11	0.03	-7.13	0.46	0.00	0.60
ZrI ₂	164-bd	11-e	0.14	0.00	-8.59	0.48	0.00	0.43
ZrS ₂	187-bi	31-a	0.19	0.18	-0.80	0.15	0.96	1.13

Table 1. Some of the calculated properties of the subset of the 137 materials that became dynamically stable after displacing the atoms along an unstable phonon mode. The properties are shown before and after the distortion, i.e. for the original dynamically unstable structures and the final dynamically stable structures, respectively.

band gap. The increase of the single-particle band gap is expected to be related to the total energy gained by making the distortion. Figure 6 shows the relation between the two quantities. Simplified models, for low dimensional systems and weak electron-phonon coupling, predict a proportionality between these two quantities [36]. From our results it is clear that there is no universal relationship between the change in band gap and total energy. In particular, several of the metals show large gain in total energy while the gap

remains zero.

It is interesting that 21 of the distorted and dynamically stable materials exhibit direct band gaps when a tolerance of 0.1 eV is employed for the difference between the direct and indirect gap. Atomically thin direct band gap semiconductors are highly relevant as building blocks for opto-electronic devices, but only a hand full of such materials are known to date e.g. monolayers of the Mo- and W-based transition metal dichalcogenides [37, 38] and

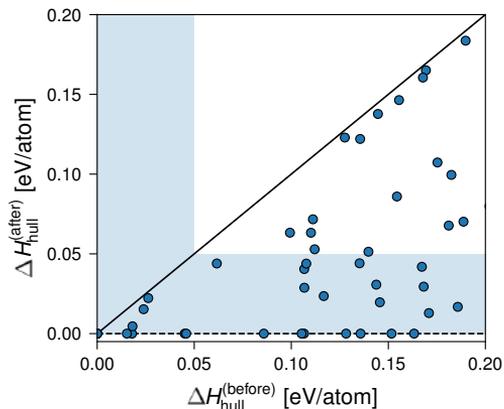


Figure 5. The energy above the convex hull for the 49 monolayers before and after distortion. Materials with a $\Delta H_{\text{hull}}^{\text{(after)}}$ close to zero are expected to be thermodynamically stable. The range up to 0.05 eV/atom above the convex hull has been indicated by a shaded blue region to visualise the importance of structural distortions for assessing the thermodynamic stability.

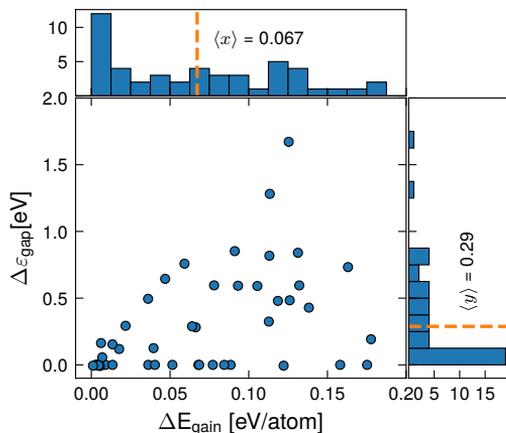


Figure 6. The average energy gain for the new stable materials is 0.067 eV / atom and the average gap opening is 0.29 eV.

black phosphorous [39]. As an example of a monolayer material that drastically changes from a metal to a direct band gap semiconductor upon distortion, we show the band structure of CdBr in Figure 7. The initial unstable metallic phase of the material becomes dynamically stable upon distortion and opens a direct band gap of 1.28 eV at the C point.

2.5. Machine learning dynamical stability

We next attempt to accelerate the dynamic stability prediction using the machine learning model outlined in Section 3.2.

As an introductory exercise we consider the correlation between dynamical stability and five elementary materials properties, namely the energy above the convex hull, the PBE band gap, the DOS at the Fermi level, the total energy per atom, and heat of formation. Figure 8 shows the distribution of these properties over the 3295 2D materials. The materials have been split into dynamically stable (blue) and dynamically unstable materials (orange), respectively. There is a clear correlation between dynamical stability and the first three materials properties shown in panels a-c. In particular, dynamically stable materials are closer to the convex hull, have larger band gap, and lower DOS at E_F as compared to dynamically unstable materials. The observed correlation with ΔH_{hull} is consistent with previous findings based on phonon calculations [31]. In contrast, no or only weak correlation is found for the last two quantities in panels d-e. These five properties were used as a low-dimensional feature vector for training an XGBoost machine learning model that will serve as a baseline for a model trained on the higher dimensional RAD-PDOS representation described in Sec. 3.2.

To evaluate the performance of our model we employ the receiver operating characteristic (ROC) curve. The ROC curve maps out the number of materials correctly predicted as unstable as a function of the number of materials incorrectly labelled as unstable, and it is calculated by varying the classification tolerance of the model. The area under the curve (AUC) is a measure of the performance of the classifier. Random guessing would amount to a linear ROC curve with unit slope, shown in Fig. 9 by the dashed grey line, and correspond to an AUC of 0.5 whereas a perfect classification model would have an AUC of 1. When calculating the ROC curve of our dynamic stability classifier we employ ten fold cross-validation (CV). This allows us to obtain a mean ROC curve and its standard deviation, which we then use to evaluate the performance of our model.

The results from the machine learning model is shown in Fig. 9. The mean ROC curve is shown in blue in Fig.9a; it achieves an excellent 10-fold CV AUC of 0.90 ± 0.01 . This suggests that the XGBoost model is able to efficiently detect the dynamically unstable materials in the C2DB. We quantify the effect of the RAD-PDOS fingerprints by comparing the performance of the full model with a model trained on the low-dimensional fingerprint. We observe that the effect on including the RAD-PDOS in the fingerprint is statistically significant raising the AUC from 0.82 ± 0.01

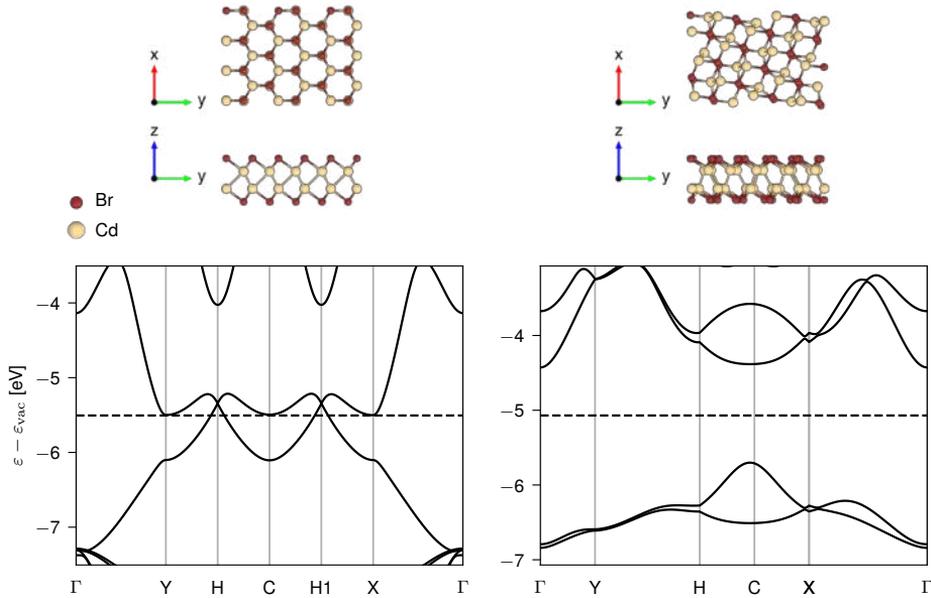


Figure 7. Extreme case of gap opening for the new stable material (CdBr) where the difference in the gap between the initial unstable metallic phase and the final structure is 1.28 eV.

to 0.90 ± 0.01 . The relative impact of the RAD-PDOS fingerprints is shown in the feature importance evaluation in Fig. 9c. Here feature importance refers to how many times a feature is used to perform a split in the decision trees, and the feature importance have been summed for the six different components of the RAD-PDOS fingerprints, i.e. summing the radial distance and energy axes of the fingerprint. The vertical dashed black line shows the feature importance of random noise for reference. We observe that especially the RAD-PDOS ss fingerprint leads to many splits in the gradient boosted trees. Part of the importance is explained by the number of non-zero features in the fingerprints. For the materials without d -orbital electrons, the sd , pd and dd fingerprints will be all zero, while the ss , sp and pp will have fewer zero-valued features and thus more features to use for splits in the trees.

Because of the strong performance of the model, we envision that it can be deployed directly after the initial relaxation step of a high-throughput workflow to reduce the number of phonon calculations needed to remove the dynamically unstable materials as depicted in Figure 1. The ML model does not replace the phonon calculation but is merely used to avoid

performing expensive phonon calculations for materials that can be labelled unstable by the ML model. Depending on the number of stable candidates that one is willing to falsely label as unstable, it is possible to save a significant amount of phonon-calculations by pre-screening with the ML model. The willingness to sacrifice materials is controlled by the classification tolerance. The trade-off between the number of unstable materials removed and the number of stable materials lost is directly mapped out by the ROC curve. In Fig. 9b we have indicated the classification thresholds where we discard 5%, 10% and 20 % of the stable materials, and we observe that we can save $56 \pm 9\%$, $70 \pm 6\%$ and $85 \pm 3\%$ of the computations for the three thresholds, respectively.

As an additional test of the machine learning model we apply it to the set of the 137 dynamically unstable materials that were investigated using the CBP protocol in the first part of the paper. The dynamical stability of the materials is evaluated by the ML model both before and after being pushed along an unstable mode (recall that before the push all the 137 materials are unstable; after the push the subset of 49 materials listed in Table 1 become stable

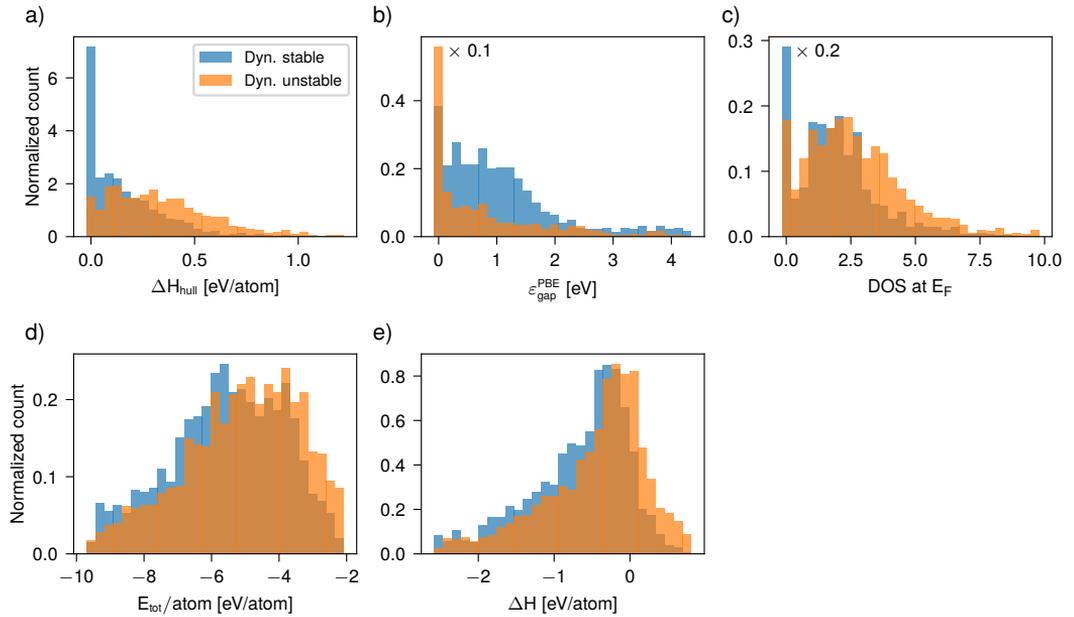


Figure 8. Histograms of electronic features for stable and unstable materials. a) shows the distribution of the energy above the convex hull for high and low stability predicted materials. The stable materials tend to be closer to the convex hull. b) shows that materials predicted to be stable more frequently have a PBE band gap larger than zero. c) shows the DOS at the Fermi level distributions, which supports that stable materials have lower DOS at E_F . d) is for total energy per atom, which shows a slightly better separation between stable and unstable materials as stable materials tend to have a slightly lower total energy per atom. In e) the unstable materials tend to have a slightly larger heat of formation. Note that for b) and c) the peaks at $x = 0$ for both stable and unstable materials have been scaled down by a factor of 10 and 5, respectively.

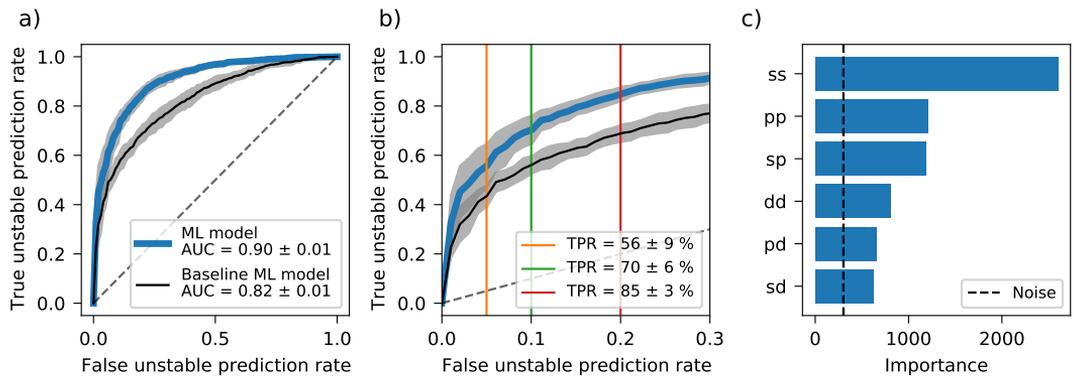


Figure 9. Machine learning results. a) shows the ROC curves for a machine learning model trained on RAD-PDOS fingerprints and a baseline model trained on electronic features from the C2DB. The AUC scores are 0.91 and 0.82, respectively. b) shows the ROC curves zoomed on low false prediction rates with different classification thresholds highlighted with vertical lines. By accepting 5% of the stable materials being falsely characterised as unstable, we can correctly label $56 \pm 9\%$ of the unstable materials. For 10% the true prediction rate is $70 \pm 6\%$ and for 20% it is $85 \pm 3\%$. c) shows feature importances of the different RAD-PDOS fingerprints measured as how many times the fingerprints are used to perform a split in the ML model. The RAD-PDOS ss fingerprint is found to be the most important fingerprint for the ML model.

while the other materials remain unstable). It is found that before the push 56% of the unstable materials are labeled correctly. After the push, only 29% of the unstable materials are labelled as unstable while the accuracy of the stable materials are 72%. Overall, the ML model performs worse on this test set than on a randomly selected test set from the original data set. An obvious explanation is that the 137 materials were selected according to (i) low energy above the convex hull ($\Delta H_{\text{hull}} < 0.2 \text{ eV/atom}$) and (ii) dynamically unstable. As seen from Fig. 8a such materials are highly unusual and not well represented by the set of materials used to train the model.

In conclusion, we have performed a systematic study of structural instabilities in 2D materials. We have validated a simple protocol (here referred to as the CBP protocol) for identifying dynamical instabilities based on the frequency of phonons at the center and boundary of the BZ. The CBP protocol correctly classifies 2D materials as dynamically stable/unstable in 261 out of 278 cases [25] and is ideally suited for high-throughput studies where the computational cost of evaluating the full phonon band structure becomes prohibitive.

For 137 dynamically unstable monolayers with low formation energies, we displaced the atoms along an unstable phonon mode and relaxed the structure in a 2×1 or 2×2 supercell. This resulted in 49 distorted, dynamically stable monolayers. The success rate of obtaining a dynamically stable structure from this protocol was found to be significantly higher for materials with only one unstable phonon mode as compared to cases with several modes. In the latter case, the displacement vector is not unique and different choices generally lead to different, (dynamically unstable) structures. The 49 stable structures were fully characterised by an extensive computational property workflow and the results are available via the C2DB database. The properties of the distorted structures can deviate significantly from the original high symmetry structures, and we found only a weak, qualitative relation between the gain in total energy and band gap opening upon distortion. Finally, we trained a machine learning classification model to predict the dynamical stability using a radially decomposed projected density of states (RAD-PDOS) representation as input and a gradient boosting decision tree ensemble method (XGBoost) as learning algorithm. The model achieves an excellent ROC-AUC score of 0.90 and lends itself to high-throughput assessment of dynamical stability.

3. Methods

3.1. Density functional theory calculations

All phonon calculations were performed using the `asr.phonopy` recipe of the Atomic Simulation Recipes (ASR) [40], which makes use of the Atomic Simulation Environment (ASE) [41] and PHONOPY [42]. The DFT calculations were performed with the GPAW [43] code and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [44]. The BZ was sampled on a uniform k -point mesh of density of 6.0 \AA^2 and the plane wave cutoff was set to 800 eV. To evaluate the Hessian matrix, the small displacement method was used with a displacement size of 0.01 \AA and forces were converged up to 10^{-4} eV/\AA . The non-analytical force constants were not included because we saw that they do not have any effect on the minimum eigenvalues of the hessian. To benchmark the CBP protocol, we compare to full phonon band structures. In these calculations, the size of the supercell is chosen such that the Hessian matrix includes interactions between pairs of atoms within a radius of at least 12 \AA . (This implies that the supercell must contain a sphere of radius 12 \AA). The spacegroup of the materials before and after applying the protocol, were calculated with `spglib` [45], with a symmetry threshold of 0.1 \AA , a value also used in a similar study [26].

3.2. Machine learning method

Below we describe the machine learning algorithm that we have developed and assessed in an attempt to accelerate the prediction of dynamic instabilities. Our choice of machine learning algorithm is the library, XGBoost [46], due to its robustness and flexibility, while being a simpler model compared to neural network methods. XGBoost is a regularized high-performance implementation of gradient tree boosting, which makes predictions based on an ensemble of gradient boosted decision trees. The decision trees of the ensemble are grown sequentially while learning from the mistakes of the previous trees by minimizing the loss function through gradient descent. This loss function is regularized to reduce the complexity of the individual decision trees which reduces the risk of overfitting. In contrast, in the widely used decision tree ensemble model Random Forest, the decision trees are grown independently and without any regularization. The dataset used is a subset of C2DB and consists of 3295 materials, which does not include the 137 distorted materials identified in the first part of the paper (as these will be used as a particularly challenging test case for the model performance). As input for the model we use the radially projected density of states (RAD-PDOS) material fingerprints

[47]. The RAD-PDOS starts from the wave functions projected onto the atomic orbitals (ν) of all the atoms (a) of the crystal, $\rho_{nk}^{\nu\nu'} = |\langle \psi_{nk} | a\nu \rangle|^2$. For each state, these projections are then combined into a radially distributed orbital pair correlation function,

$$\rho_{nk}^{\nu\nu'}(R) = \sum_{aa'} \rho_{nk}^{\nu\nu'} \rho_{nk}^{a'a'} G(R - |R_a - R_{a'}|; \delta_R) \times \exp(-\alpha_R R) \quad (1)$$

Finally, the radial functions are distributed on an energy grid,

$$\rho^{\nu\nu'}(R, E) = \sum_{nk} \rho_{nk}^{\nu\nu'}(R) G(E - (\varepsilon_{nk} - E_F); \delta_E) \times \exp(-\alpha_E R), \quad (2)$$

where $G(x; \delta)$ is a Gaussian of width δ centered at $x = 0$. For the materials in the dataset, the s , p , and d orbitals lead to six unique components of the RAD-PDOS fingerprint ($\nu\nu' = \{ss, sp, sd, pp, pd, dd\}$). The fingerprint involves some hyperparameters for which we use the values $E_{\min} = -10\text{eV}$, $E_{\max} = 10\text{eV}$, $N_E = 25$, $\delta_E = 0.3\text{eV}$, $\alpha_E = 0.2\text{eV}^{-1}$, $R_{\min} = 0$, $R_{\max} = 5\text{\AA}$, $N_R = 20$, $\delta_R = 0.25\text{\AA}$, $\alpha_R = 0.33\text{\AA}^{-1}$.

In addition to the RAD-PDOS fingerprint, we consider a low-dimensional fingerprint consisting of five features, namely the PBE electronic band gap ($\varepsilon_{\text{gap}}^{\text{PBE}}$), crystal formation energy (ΔH), density of states at the Fermi level ($\text{DOS}@E_F$), energy above the convex hull (ΔH_{hull}) and the total energy per atom in the unit cell. The low-dimensional fingerprint is used to train a "baseline" ML model that we use to benchmark the performance of the ML model based on the more involved RAD-PDOS fingerprint. Common to all the features considered is that they are obtained from a single DFT calculation and thus are much faster to compute than the phonon frequencies. The gradient boosting model introduces several hyperparameters such as depth of the trees, learning rate, minimum loss gain to perform a split and minimum weights in tree leaves. These parameters are optimized using Bayesian optimization where a Gaussian process is fitted to the mean test ROC-AUC of a 10-fold cross-validation. The XGBoost classification model is in fact a logistic regression model, i.e. the output of the model is a number between 0 and 1 which is interpreted as a probability. In our case, 0 (1) refers to a dynamically stable (unstable) material.

4. Data availability

All the crystal structures and properties will be available in the C2DB (<https://doi.org/10.11583/DTU.14616660.v1>).

5. Acknowledgments

The Center for Nanostructured Graphene (CNG) is sponsored by the Danish National Research Foundation, Project DNRF103. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant agreement No 773122 (LIMA). K. S. T. is a Villum Investigator supported by VILLUM FONDEN (grant no. 37789).

6. Author contributions

S.M. and K.S.T. developed the initial concept. S.M. performed the benchmark of the CBP protocol and developed the workflow and the analysis for the pushed materials. M.K.S., N.R.K and P.M.L. conducted the machine learning analysis. K.S.T. supervised the work and helped in interpretation of the results. All authors modified and discussed the paper together.

7. Competing interests

The authors declare no competing interests.

References

- [1] Kohn W and Sham L J 1965 *Phys. Rev.* **140**(4A) A1133–A1138
- [2] Grüner G 1988 *Rev. Mod. Phys.* **60**(4) 1129–1181
- [3] Peierls R 1996 *Quantum Theory of Solids* International Series of Monographs on Physics (Clarendon Press) ISBN 9780192670175
- [4] Ono Y and Hamano T 2000 *Journal of the Physical Society of Japan* **69** 1769–1776 (Preprint <https://doi.org/10.1143/JPSJ.69.1769>) URL <https://doi.org/10.1143/JPSJ.69.1769>
- [5] Weber F, Rosenkranz S, Castellan J P, Osborn R, Karapetrov G, Hott R, Heid R, Bohnen K P and Alatas A 2011 *Phys. Rev. Lett.* **107**(26) 266401
- [6] Bianco R, Errea I, Monacelli L, Calandra M and Mauri F 2019 *Nano Letters* **19** 3098–3103 (Preprint <https://doi.org/10.1021/acs.nanolett.9b00504>)
- [7] Zhou J S, Monacelli L, Bianco R, Errea I, Mauri F and Calandra M 2020 *Nano Letters* **20** 4809–4815 (Preprint <https://doi.org/10.1021/acs.nanolett.0c00597>)
- [8] Leroux M, Le Tacon M, Calandra M, Cario L, Méasson M A, Diener P, Borrisenko E, Bosak A and Rodière P 2012 *Phys. Rev. B* **86**(15) 155125
- [9] Bianco R, Errea I, Monacelli L, Calandra M and Mauri F 2019 *Nano Letters* **19** 3098–3103
- [10] Calandra M, Mazin I I and Mauri F 2009 *Phys. Rev. B* **80**(24) 241108
- [11] Zhu X, Cao Y, Zhang J, Plummer E W and Guo J 2015 *Proceedings of the National Academy of Sciences* **112** 2367–2371
- [12] Johannes M and Mazin I 2008 *Physical Review B* **77** 165135
- [13] Lin Y C, Dumcenco D O, Huang Y S and Suenaga K 2014 *Nature Nanotechnology* **9** 391–396 ISSN 1748-3395
- [14] Xi X, Zhao L, Wang Z, Berger H, Forró L, Shan J and Mak K F 2015 *Nature Nanotechnology* **10** 765–769
- [15] Ugeda M, Bradley A, Zhang Y, Onishi S, Chen Y, Ruan W, Ojeda-Aristizabal C, Ryu H, Edmonds M, Tsai H Z,

- Riss A, Mo S, Lee D, Zettl A, Hussain Z, Shen Z X and Crommie M 2016 *Nature Physics* **12** 92–97 ISSN 1745-2473
- [16] Yang Y, Fang S, Fatemi V, Ruhman J, Navarro-Moratalla E, Watanabe K, Taniguchi T, Kaxiras E and Jarillo-Herrero P 2018 *Phys. Rev. B* **98**(3) 035203
- [17] Yu Y, Yang F, Lu X F, Yan Y J, Cho Y H, Ma L, Niu X, Kim S, Son Y W, Feng D, Li S, Cheong S W, Chen X H and Zhang Y 2015 *Nature Nanotechnology* **10** 270–276 ISSN 1748-3395
- [18] Ryu H, Chen Y, Kim H, Tsai H Z, Tang S, Jiang J, Liou F, Kahn S, Jia C, Omrani A A, Shim J H, Hussain Z, Shen Z X, Kim K, Min B I, Hwang C, Crommie M F and Mo S K 2018 *Nano Letters* **18** 689–694
- [19] Ge Y and Liu A Y 2012 *Phys. Rev. B* **86**(10) 104101
- [20] Sugawara K, Nakata Y, Shimizu R, Han P, Hitosugi T, Sato T and Takahashi T 2016 *ACS Nano* **10** 1341–1345
- [21] Wang H, Chen Y, Duchamp M, Zeng Q, Wang X, Tsang S H, Li H, Jing L, Yu T, Teo E H T and Liu Z 2018 *Advanced Materials* **30** 1704382
- [22] Zhu J, Wang Z, Yu H, Li N, Zhang J, Meng J, Liao M, Zhao J, Lu X, Du L, Yang R, Shi D, Jiang Y and Zhang G 2017 *Journal of the American Chemical Society* **139** 10216–10219
- [23] Wang L, Xu Z, Wang W and Bai X 2014 *Journal of the American Chemical Society* **136** 6693–6697
- [24] Krishnamoorthy A, Bassman L, Kalia R K, Nakano A, Shimofuji F and Vashishta P 2018 *Nanoscale* **10**(6) 2742–2747
- [25] Mounet N, Gibertini M, Schwaller P, Campi D, Merkys A, Marrazzo A, Sohier T, Castelli I E, Cepellotti A, Pizzi G and Marzari N 2018 *Nature Nanotechnology* **13** 246–252 ISSN 1748-3395
- [26] Togo A and Tanaka I 2013 *Phys. Rev. B* **87**(18) 184104 URL <https://link.aps.org/doi/10.1103/PhysRevB.87.184104>
- [27] Patrick C E, Jacobsen K W and Thygesen K S 2015 *Phys. Rev. B* **92**(20) 201205
- [28] Yang R X, Skelton J M, da Silva E L, Frost J M and Walsh A 2020 *The Journal of Chemical Physics* **152** 024703
- [29] Kayastha P and Ramakrishnan R 2021 *The Journal of Chemical Physics* **154** 061102
- [30] Hastrup S, Strange M, Pandey M, Deilmann T, Schmidt P S, Hinsche N F, Gjerding M N, Torelli D, Larsen P M, Riis-Jensen A C, Gath J, Jacobsen K W, Mortensen J J, Olsen T and Thygesen K S 2018 *2D Materials* **5** 042002
- [31] Gjerding M, Taghizadeh A, Rasmussen A, Ali S, Bertoldo F, Deilmann T, Holguin U, Knøsgaard N, Kruse M, Manti S *et al.* 2021 *2D Materials* **8** 044002
- [32] <https://cmr.fysik.dtu.dk/c2db/c2db.html>
- [33] Eriksson F, Fransson E and Erhart P *Advanced Theory and Simulations* **2** 1800184
- [34] Carrete J, Li W, Lindsay L, Broido D A, Gallego L J and Mingo N 2016 *Materials Research Letters* **4** 204–211 (Preprint <https://doi.org/10.1080/21663831.2016.1174163>) URL <https://doi.org/10.1080/21663831.2016.1174163>
- [35] Royo M, Hahn K R and Stengel M 2020 *Phys. Rev. Lett.* **125**(21) 217602 URL <https://link.aps.org/doi/10.1103/PhysRevLett.125.217602>
- [36] Rossnagel K 2011 *Journal of Physics: Condensed Matter* **23** 213001 URL <https://doi.org/10.1088/0953-8984/23/21/213001>
- [37] Mak K F, Lee C, Hone J, Shan J and Heinz T F 2010 *Physical Review Letters* **105** 136805
- [38] Manzeli S, Ovchinnikov D, Pasquier D, Yazyev O V and Kis A 2017 *Nature Reviews Materials* **2** 1–15
- [39] Liu H, Neal A T, Zhu Z, Luo Z, Xu X, Tománek D and Ye P D 2014 *ACS Nano* **8** 4033–4041
- [40] Gjerding M, Skovhus T, Rasmussen A, Bertoldo F, Larsen A H, Mortensen J J and Thygesen K S 2021 *Computational Materials Science* **199** 110731
- [41] Larsen A H, Mortensen J J, Blomqvist J, Castelli I E, Christensen R, Dulak M, Friis J, Groves M N, Hammer B, Hargus C *et al.* 2017 *Journal of Physics: Condensed Matter* **29** 273002
- [42] Togo A and Tanaka I 2015 *Scripta Materialia* **108** 1 – 5 ISSN 1359-6462
- [43] Enkovaara J, Rostgaard C, Mortensen J J, Chen J, Dulak M, Ferrighi L, Gavnholt J, Glinsvad C, Haikola V, Hansen H A, Kristoffersen H H, Kuisma M, Larsen A H, Lehtovaara L, Ljungberg M, Lopez-Acevedo O, Moses P G, Ojanen J, Olsen T, Petzold V, Romero N A, Stausholm-Møller J, Strange M, Tritsarlis G A, Vanin M, Walter M, Hammer B, Häkkinen H, Madsen G K H, Nieminen R M, Nørskov J K, Puska M, Rantala T T, Schiøtz J, Thygesen K S and Jacobsen K W 2010 *Journal of Physics: Condensed Matter* **22** 253202 URL
- [44] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77**(18) 3865–3868
- [45] A T 2009 Sgplib URL <https://atztogo.github.io/spglib/>
- [46] Chen T and Guestrin C 2016 Xgboost: A scalable tree boosting system *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '16* (New York, NY, USA: Association for Computing Machinery) p 785–794 ISBN 9781450342322 URL <https://doi.org/10.1145/2939672.2939785>
- [47] Knøsgaard N R and Thygesen K S 2022 *Nature Communications* **13** 468 ISSN 2041-1723 URL <https://doi.org/10.1038/s41467-022-28122-0>

11.5 Paper [V]: Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption,

Jiban Kangsabanik, Mark Kamper Svendsen, Alireza Taghizadeh, Andrea Crovetto, Kristian S. Thygesen,
“Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption,”
Submitted to Journal of the American Chemical Society: arXiv:2207.09569.

Indirect band gap semiconductors for thin-film photovoltaics: High-throughput calculation of phonon-assisted absorption

Jiban Kangsabanik,* Mark Kamper Svendsen, Alireza Taghizadeh, and Kristian S. Thygesen†
*CAMD, Computational Atomic-Scale Materials Design, Department of Physics,
Technical University of Denmark, 2800 Kgs. Lyngby Denmark*

Andrea Crovetto
*National Centre for Nano Fabrication and Characterization (DTU Nanolab),
Technical University of Denmark, 2800 Kgs. Lyngby Denmark*

(Dated: September 28, 2022)

Discovery of high-performance materials remains one of the most active areas in photovoltaics (PV) research. Indirect band gap materials form the largest part of the semiconductor chemical space, but predicting their suitability for PV applications from first principles calculations remains challenging. Here we propose a computationally efficient method to account for phonon assisted absorption across the indirect band gap and use it to screen 127 experimentally known binary semiconductors for their potential as thin film PV absorbers. Using screening descriptors for absorption, carrier transport, and nonradiative recombination, we identify 28 potential candidate materials. The list, which contains 20 indirect band gap semiconductors, comprises both well established (3), emerging (16), and previously unexplored (9) absorber materials. Most of the new compounds are anion rich chalcogenides (TiS_3 , Ga_2Te_5) and phosphides (PdP_2 , CdP_4 , MgP_4 , BaP_3) containing homoelemental bonds, and represent a new frontier in PV materials research. Our work highlights the previously underexplored potential of indirect band gap materials for optoelectronic thin-film technologies.

I. INTRODUCTION

Renewable electricity produced by photovoltaics (PV) is likely to become a cornerstone of a future sustainable and climate-neutral energy system. Although the current PV technology is mainly based on silicon (Si) wafer solar cells (26.7% efficiency)[1, 2], thin film technologies have distinct advantages such as mechanical flexibility, reduced material usage, potential for low-cost, low-temperature processing, and integration with existing Si technology in tandem solar cells. After decades of research and device optimization various thin film absorber materials, like GaAs (29.1%), CdTe (22.1%), InP (24.2%), CIGS (23.4%), etc.[1, 3–5] have reached efficiencies comparable to the Si-based technology. Most recently, hybrid lead-halide perovskite (MAPbI_3) has emerged as an additional PV candidate material with a record efficiency of 25.8% [6].

Although promising, the thin film PV materials still suffer from major drawbacks associated with low material abundance (In, Ga), toxicity (Cd, As, Pb), and long-term device stability under ambient conditions (hybrid perovskites), which hinder their large-scale deployment. As such, the discovery of new (thin-film) PV materials remains a relevant and timely challenge. Within the PV literature, the most widely used selection-metric for absorber materials is the electronic band gap. Indeed, Shockley and Queisser demonstrated that thermodynamic detailed balance sets an upper limit to the power

conversion efficiency (PCE) of a semiconductor given solely by its band gap.[7] More realistic PV descriptors have been developed by invoking further physical properties like the optical absorption coefficient, the carrier mobility, and internal quantum efficiency, along with device related aspects such as film thickness, light-trapping schemes, etc.[8–11]

Recently, the substantial increase in computer power along with improvements in theoretical methods have ushered in an era where it is possible to screen for new materials faster, cheaper, and more rationally than the traditional approach based on experimental trial and error. In particular, high-throughput density functional theory (DFT) calculations have been employed to identify suitable PV candidate materials.[12–18] Indirect band gap semiconductors, which occupy a major portion of the semiconductor space, have mostly been ignored in these studies. This is because indirect gap materials will absorb photons less efficiently around the band gap, which is likely to lead to low PCE in thin film devices. On the other hand, it has been argued that a small indirect gap might be beneficial for PV applications due to reduced radiative recombination.[19, 20] This situation clearly calls for a clarification.

The major bottleneck in evaluating the PV performance of an indirect band gap semiconductor from first-principles, is modeling indirect absorption. This is a higher order process involving one or several phonons (or defects or free carriers) in addition to the photon. Such calculations are very demanding. As a consequence, ab-initio calculations of momentum-indirect absorption has so far been limited to a few materials[21–29].

In this work, we propose an approximate method to

* jibka@dtu.dk

† thygesen@fysik.dtu.dk

evaluate phonon-assisted absorption in indirect band gap semiconductors, which only involves the Γ -point optical phonons and therefore is feasible for high-throughput studies. After establishing the accuracy of the approach for a few prototypical materials, we use it to evaluate the key PV performance parameters for 127 stable, binary semiconductors of which 97 have indirect band gaps. In addition to already known materials, we identify a number of new indirect band gap materials that show promising potential for thin film PV. All the materials along with their calculated properties are available in an open online database. To the best of our knowledge, this is the first high-throughput study evaluating indirect gap materials for PV, which includes calculation of phonon-assisted absorption.

II. METHOD

Phonon assisted absorption is a second order process, which involves a photon-phonon pair satisfying (crystal) momentum conservation (see Fig.1(a)). Because of the involved sum over intermediate states and the need to calculate all the phonons and electron-phonon couplings, its calculation from first-principles is very demanding. In this work we use a simplified approach that involves only the optical Γ -point phonons. Within second order perturbation theory, the phonon-assisted absorption is given by,

$$\alpha(\omega) = \frac{4\pi e^2}{m^2 \epsilon_0 c} \frac{1}{n(\omega)\omega} \frac{1}{N_k^2} \sum_{\mathbf{k}_1 \mathbf{k}_2, c, v, l} |t_{c\mathbf{k}_2 \leftarrow v\mathbf{k}_1}^{l, \pm}|^2 \delta(E_{c\mathbf{k}_2} - E_{v\mathbf{k}_1} - \hbar\omega \pm \hbar\omega_{\mathbf{q}}^l) \quad (1)$$

where,

$$t_{c\mathbf{k}_2 \leftarrow v\mathbf{k}_1}^{l, \pm} = \sqrt{n_{\mathbf{q}}^l} \left[\sum_{\alpha} \frac{\langle \psi_{c\mathbf{k}_2} | V^l(\mathbf{q}) | \psi_{\alpha\mathbf{k}_1} \rangle \langle \psi_{\alpha\mathbf{k}_1} | \hat{\mathbf{e}} \cdot \mathbf{p} | \psi_{v\mathbf{k}_1} \rangle}{E_{v\mathbf{k}_1} - E_{\alpha\mathbf{k}_1} + \hbar\omega} + \sum_{\beta} \frac{\langle \psi_{c\mathbf{k}_2} | \hat{\mathbf{e}} \cdot \mathbf{p} | \psi_{\beta\mathbf{k}_2} \rangle \langle \psi_{\beta\mathbf{k}_2} | V^l(\mathbf{q}) | \psi_{v\mathbf{k}_1} \rangle}{E_{v\mathbf{k}_1} - E_{\beta\mathbf{k}_2} \pm \hbar\omega_{\mathbf{q}}^l} \right] \quad (2)$$

Here $\hbar\omega$ is the energy of the incident photon, l denotes the phonon mode and $\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1$ its momentum, and $n(\omega)$ is the refractive index of the semiconductor. The two sums in the square bracket of Eq. (2) correspond to the two transition pathways denoted '1' and '2' in Fig.1(a). The matrix elements appearing in Eq. (2) contain the electron-phonon potential, $V^l(\mathbf{q})$, and coupling to a photon field of polarisation $\hat{\mathbf{e}}$, respectively. The phonon occupation factors, $n_{\mathbf{q}}^l$, are given by the Bose-Einstein distribution.

In our approximation, we assume that scattering on a (l, \mathbf{q}) -phonon can be replaced by scattering on a $(l, \mathbf{0})$ -phonon (acoustic phonons are disregarded). Specifically, we keep all transitions in Eq. (2), but make the replacements

$$\omega_{\mathbf{q}}^l \rightarrow \omega_{\mathbf{0}}^l \quad (3)$$

$$\langle \psi_{c\mathbf{k}_2} | V^l(\mathbf{k}_2 - \mathbf{k}_1) | \psi_{v\mathbf{k}_1} \rangle \rightarrow \langle u_{c\mathbf{k}_2} | V^l(\mathbf{0}) | u_{v\mathbf{k}_1} \rangle, \quad (4)$$

where u is the periodic part of the Bloch function. Our approximation avoids the computation of the full \mathbf{q} -dependent phonons and electron-phonon couplings and instead requires these quantities only at the Γ -point. The efficiency gain is largest for materials with small unit cells and it becomes exact in the limit of a large unit cell. In particular, due to band folding, the exact result for any

crystal is obtained by using a supercell containing a sufficiently large number of copies of the primitive cell (although this strategy is inefficient as momentum selection rules are not exploited). Below we illustrate the performance of our approach for the case of bulk silicon, and refer to the supplementary information (SI) for further tests (BP, MoS₂, MoSe₂, etc.), all showing encouraging results. Details of the DFT calculations can be found in section V.

Fig. 2(a) shows the phonon-assisted absorption in silicon obtained within our approximate scheme applied to the Si primitive unit cell (orange - 3 optical phonons included) and a supercell containing $2 \times 2 \times 2$ repetitions of the primitive cell (green - 45 optical phonons). Our results are compared to experimental data ($T = 296\text{K}$) and a previous ab-initio result including all phonons. The theoretical spectra have been shifted horizontally to match the experimental absorption onset. We have checked that the use of the LDA and PBE xc-functionals lead to insignificant changes in the spectrum. While our $2 \times 2 \times 2$ supercell Γ -phonon approach shows excellent agreement with both sets of reference data, the primitive unit cell Γ -phonon approach slightly underestimates the absorption coefficient. Next, we add the first order absorption coefficient, α_{dir} , (calculated in the random phase approximation (RPA)) to the indirect absorption coeffi-

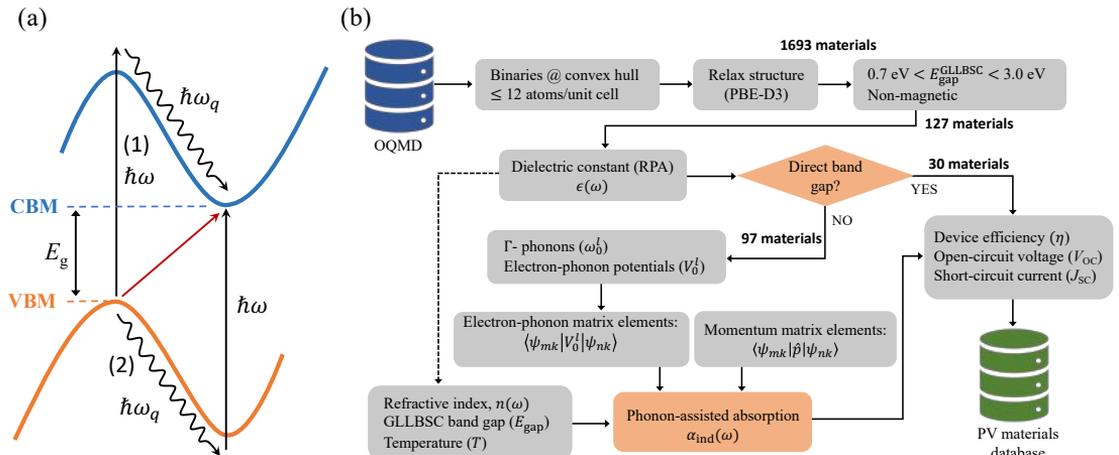


FIG. 1. (a) Schematic diagram of phonon-assisted optical absorption. (b) Computational workflow used to evaluate the PV performance parameters for direct and indirect band gap semiconductors.

cients and evaluate the radiative device PCE, η , as a function of thickness of the Si absorber, see Fig. 2(b). Overall, the different descriptions of the indirect absorption leads to similar PCEs with the most approximate Γ -phonon description underestimating the PCE by maximum 5% relative to experiments. The last two panels of Fig. 2 show the temperature dependence of the indirect absorption as calculated with the Γ -phonons of the $2 \times 2 \times 2$ supercell (c) and primitive cell (d), respectively. The supercell calculations are in excellent agreement with the experimental data (dashed lines), except for small deviations at temperatures below 100K. The primitive cell calculations show larger deviations – also at higher T close to the absorption onset.

The workflow used to calculate the PV performance parameters is illustrated in Fig. 1(b). For 1693 thermodynamically stable elemental and binary compounds from the Open Quantum Materials Database (OQMD) with up to 12 atoms/unit cell, we perform structural relaxations using the PBE[30] xc-functional with D3 corrections[31] to account for van der Waals interactions (more details can be found in computational details section). The data set includes 338 semiconductors[32] for which we evaluate the band gaps using both the PBE and the GLLBSC[33] xc-functional. Keeping only non-magnetic compounds with GLLBSC band gap between 0.7 eV and 3.0 eV, reduces the data set further down to 127 compounds. For these materials, the complex dielectric function is calculated in the RPA (this is a first order response function including only direct momentum-conserving transitions with local field effects accounted for). For the 97 semiconductors with indirect band gap, the phonon assisted absorption coefficient is calculated using the Γ -phonon approximation in the primitive unit cell, c.f. Eqs. (1-2). Finally, the PV parameters are eval-

uated from the total absorption coefficient, including first and second order processes, as described in SI in section S1.

III. RESULTS AND DISCUSSION

A. Descriptors

The main descriptor used to quantify the potential of a PV absorber material, is its power conversion efficiency (PCE), η , which can be expressed as,

$$\eta = \frac{\max[V(J_{\text{sc}} - J_r^{\text{tot}}(e^{\frac{eV}{k_B T}} - 1))]}{\int_0^\infty EI_{\text{sun}}(E)dE} \quad (5)$$

In this expression V is the voltage, J_{sc} is the short-circuit current density, J_r^{tot} is total recombination current density, and T is the temperature. In the denominator, we have the input power density, where $I_{\text{sun}}(E)$ represents the photon flux from the incident solar spectrum (we have used the standard AM1.5G solar spectrum). More details can be found in Methodology section in SI. Here we stress that the only input required to obtain η from Eq. (5) in the radiative limit (i.e. when non-radiative recombination processes are neglected and thus internal quantum efficiency $Q_i = 1$) is the material's absorption spectrum.

In the simplest theory, the absorbance is assumed to be zero below the band gap and one above. In the radiative limit this yields the well known Shockley-Queisser (SQ) limit.[7] Clearly, this description fails to differentiate between different materials with the same band gap. While this is a reasonable approximation for bulk cells, a realistic description of thin film devices must account for the

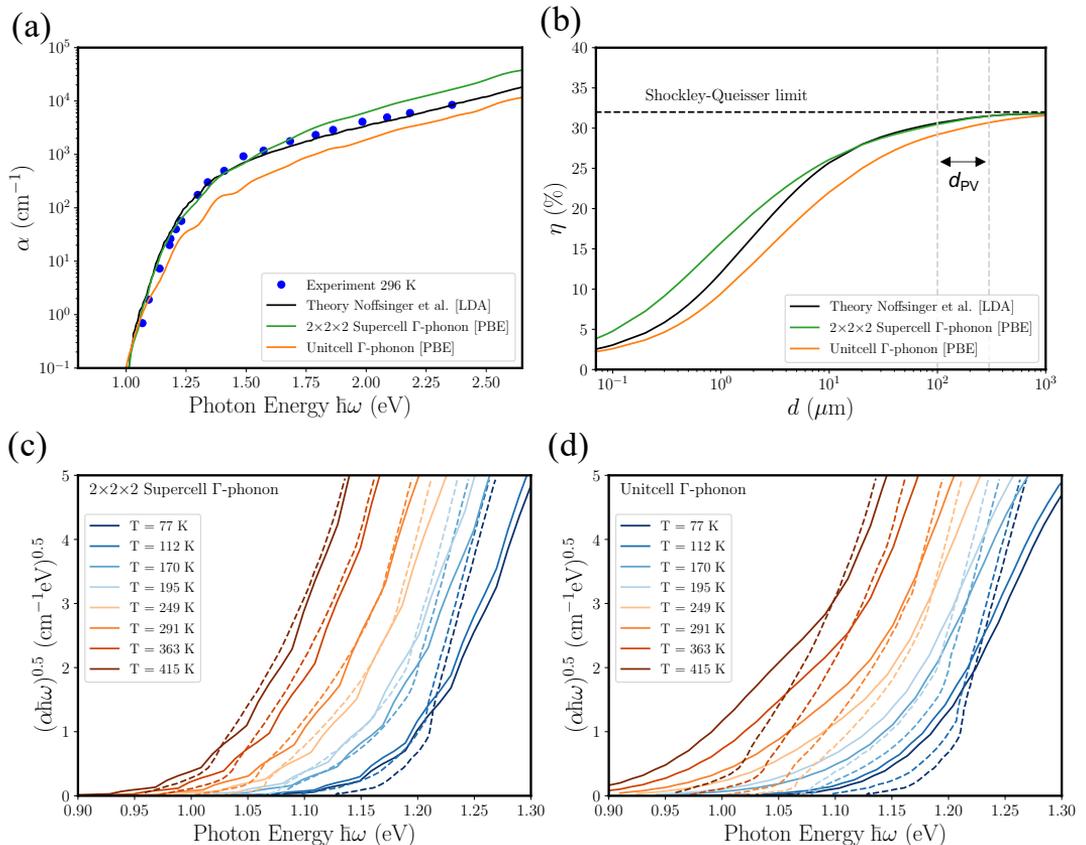


FIG. 2. (a) Comparison of phonon assisted absorption in Si, calculated using the approximation Eqs.(1-2) (green, orange), previous calculations including all phonons (black), and experimental data at $T = 296$ K (blue). The corresponding Shockley-Queisser device efficiencies as a function of film thickness are shown in (b). Here ' d_{PV} ' is the usual thickness of the Si active layer in PV cells. (c-d) Comparison of calculated phonon-assisted absorption onsets (solid lines) with experimental data (dashed lines) at different temperatures. In (c) the Γ -phonons of a $2 \times 2 \times 2$ supercell are used in Eqs.(1-2), while in (d) the Γ -phonons of the primitive Si unit cell are included. The calculated spectra have been shifted horizontally to match the experimental absorption onset.

finite thickness and frequency dependent absorption.[8-10]

A popular PV performance descriptor that invokes the absorption coefficient is the Spectroscopic Limited Maximum Efficiency (SLME).[9] The SLME sets $J_r^{\text{tot}} = J_{\text{rad}} \exp(\Delta E_g^{\text{dir}}/k_B T)$ in Eq. (5) to account for non-radiative recombination. In other words, the effect of an indirect band gap is taken into account via an increased non-radiative loss rate. In fact, this can lead to erroneous conclusions for indirect band gap materials. For example, the calculated SLME for Si at $100 \mu\text{m}$ thickness is only 1%, while the experimental PCE is around 25%. We argue that explicitly accounting for indirect absorption processes is essential for a proper description of

PV performance parameters across all band gap types. In particular, given the exact absorption coefficient, Eq. (5) would yield the exact PCE in the radiative limit. More detailed discussion is provided in the SI section S3. To account for non-radiative recombination we use the formalism proposed by Blank *et al.*[10] as discussed later.

B. Si as a case study

Before discussing our high-throughput results, we illustrate the PV analysis in detail for the well known case of Si. The silicon crystal has an indirect band gap of 1.1 eV, whereas the direct band gap at the Γ -point is 3.0 eV

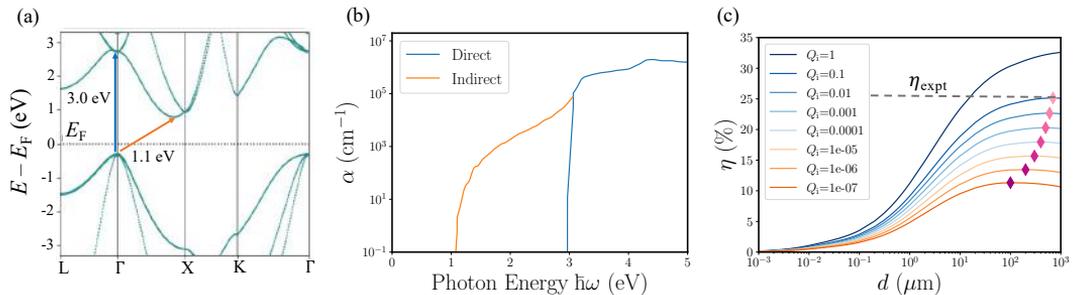


FIG. 3. (a) Electronic band structure of Silicon (Si), (b) Total optical absorption coefficient for Si (including direct and phonon assisted (indirect) absorption). (c) Efficiency (η) vs. thickness (d) plot for various internal (luminescence) quantum efficiencies (Q_i). Maximum η at each Q_i is denoted using \diamond symbol. Here experimental efficiency η_{expt} for a typical Si-solar cell is denoted by black dashed line.

(see Fig. 3(a)). Since the first-order absorption coefficient (e.g. evaluated using the independent particle approximation, the RPA, or the Bethe-Salpeter Equation), only includes vertical transitions, it will fail to account for photon absorption below the direct gap. The effect of including phonon-assisted transitions in the absorption coefficient can be seen in Fig. 3(b) (the orange curve is the same as shown in Fig. 2(a)). While the second-order processes are weak, they are completely essential for a correct description of the opto-electronic properties. In the radiative limit we obtain a PCE of 31.3% (with phonon-assisted processes) and 4.9% (only first-order processes) for a 100 μm thick Si film.

To move beyond the radiative approximation one must account for non-radiative recombination such as Shockley-Read-Hall recombination[34, 35] and Auger processes[36]. These are challenging to evaluate from first principles and to some extent dependent on the sample quality via the concentration and types of defects in the material. In addition, there are additional loss channels that depend on the device architecture, such as surface and interface recombination. Here we will neglect the latter types, and only consider the non-radiative processes taking place within the absorbing layer. To that end, we follow Blank *et al.*[10] who proposed a detailed balance compatible approach where non-radiative processes are included via the internal luminescence quantum efficiency, Q_i (the ratio of the radiative to the total recombination rates inside the material under equilibrium conditions), which becomes a free parameter, see Section S1 in SI for more details.

Fig. 3(c) shows the calculated η vs film thickness for different values of Q_i . As expected, with increasing non-radiative recombination the maximum achievable efficiency decreases. Perhaps less obvious, also the optimal thickness (thickness at which η is maximum) decreases when the non-radiative recombination increases. This effect occurs because the optimal thickness occurs exactly when the increase in absorption due to a further

increase in thickness is balanced by the increase in the non-radiative recombination. The existence of an optimal thickness then follows by noting that the former decays exponentially through the material while the latter is constant. We obtain an efficiency close to the experimental value (η_{expt}) of 25% at optimal film thickness for Q_i around 10^{-1} , see Fig. 3(c). The experimental value of Q_i in the highest-efficiency Si solar cell [2] is 1.6×10^{-2} [4], which implies a maximum efficiency of about 23%. We ascribe this small underestimation of the PCE to the Γ -point phonon approximation and to the fact that the front surface of a real Si solar cell is textured to promote light scattering and increase the average optical path of photons in Si. Both approximations lead to a slight underestimation of the fraction of incident light absorbed in a film of given thickness, see Fig. 2(b).

C. High-throughput screening

Moving on to the main purpose of our work, namely the identification of high-performance indirect band gap PV absorbers, Fig. 4(a) shows the distribution of chemical elements in the 127 semiconductors that were subject to the workflow in Fig. 1(b). There is a relatively homogeneous distribution of cations, with the exception of certain 3d transition metals, which are absent because we have excluded magnetic materials. The anionic elements are dominated by the chalcogens (in particular sulphur) followed by the halogens. We note that we did not filter out materials containing rare or toxic elements at this step. In Fig. 4(b) we show how the 127 compounds are distributed over the dimensionality of their covalently bonded atomic clusters. The dimensionality analysis follows the method of Larsen *et al.*[37]. It can be seen that the materials are almost evenly divided between '3D' and 'non-3D' materials. About 50% of the latter class are van der Waals (vdW) layered 2D materi-

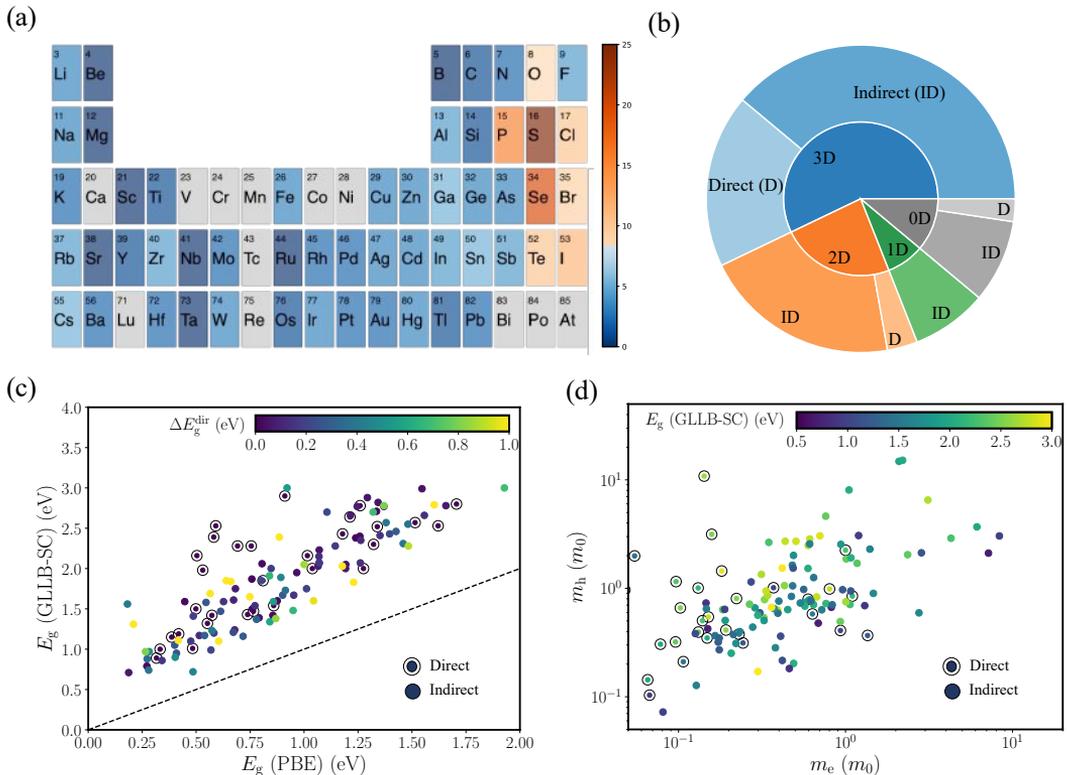


FIG. 4. (a) Elemental distribution of the dataset in the periodic table. Color scale shows number of occurrence of an element in the final dataset. Grey colored boxes represents element which are absent in the dataset. (b) Dimensionality distribution of the dataset as 3D, 2D, 1D, and 0D bulk materials. Associated indirect (ID) and direct (D) gap material distributions are shown in the outer circle. (c) GLLB-SC vs PBE band gap distribution for the selected materials. Color scale represents difference between direct and indirect band gap (ΔE_g^{dir}). (d) Effective mass (in units of electron rest mass m_0) distribution of the compounds. Here the colors scale shows corresponding GLLB-SC bandgaps.

als while the remainder are vdW-bonded 1D or 0D clusters. For each dimensionality class we show the fraction of direct (D) and indirect (ID) band gaps.

In total, more than 75% of the 127 semiconductors possess an indirect band gap. Fig. 4(c) shows the band gap distribution for our dataset. It is well known that the PBE xc-functional systematically underestimates the band gap. The GLLB-SC functional provides more accurate band gaps at little extra computational cost[33, 38]. An important reason behind the good performance of the GLLB-SC is includes a correction of the Kohn-Sham band gap to account for the so-called derivative discontinuity. The GLLB-SC band gaps range from 0.7 eV to 3.0 eV, whereas with PBE band gap lies between 0.2 eV and 2 eV. From the color scale it is clear that most of the indirect gap materials have $\Delta E_g^{\text{dir}} < 0.5$ eV while a only few compounds have $\Delta E_g^{\text{dir}} > 1.0$ eV (all materials with ΔE_g^{dir} above 1 eV have yellow color code) .

High absorption and low non-radiative recombination rates do not guarantee good PV performance. As an additional criterion, the carrier mobility, $\mu = \frac{e\tau}{m^*}$, should be high enough to ensure efficient electron-hole separation and extraction of charges at the contacts. The relaxation time, τ , accounts for scattering on other charge carriers, phonons, and crystal imperfections. While it can be accurately estimated from first principles[39, 40], its calculation remains challenging in high-throughput studies. As a consequence we focus on the carrier effective mass

$$m^* = 3 \left[\frac{1}{m_{xx}^*} + \frac{1}{m_{yy}^*} + \frac{1}{m_{zz}^*} \right]^{-1}, \quad (6)$$

where m_{ii}^* is the effective mass along direction i , which we calculate by fitting a parabola to the band extrema. Fig. 4(d) gives an overview of the hole and electron effective masses in the 127 semiconductors considered. The band gap is indicated by the color code.

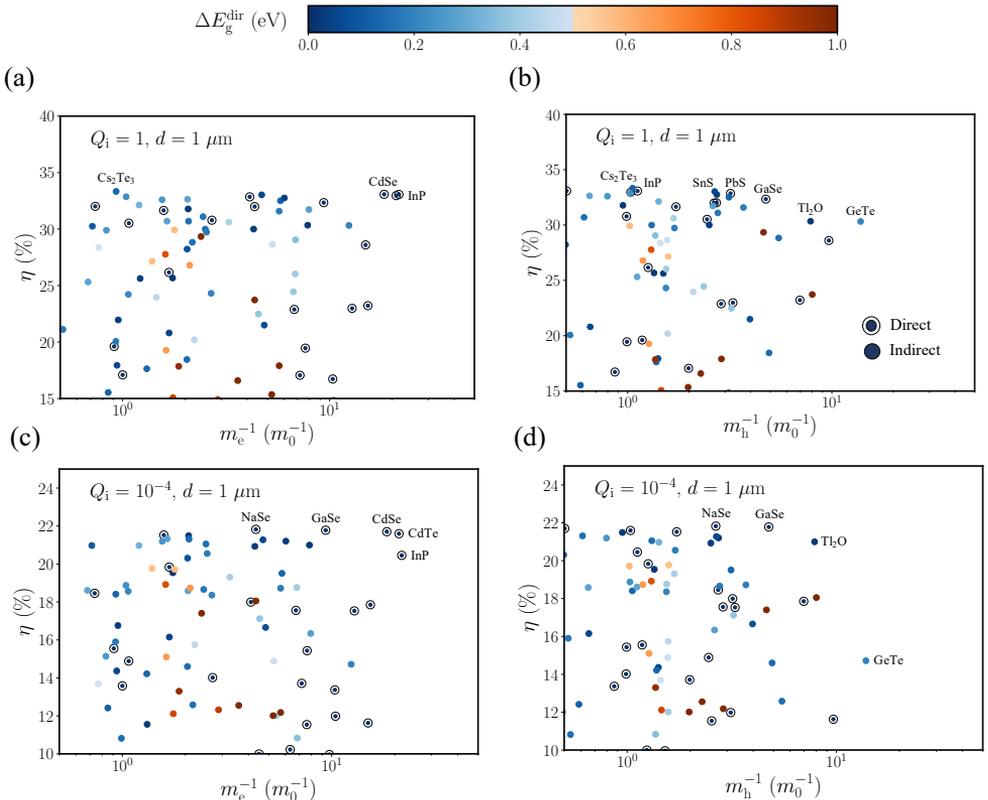


FIG. 5. Pareto distribution for the selected photovoltaic (PV) materials (device efficiency (η) vs inverse effective masses (m_e^{-1} , m_h^{-1})). (a) Radiative ($Q_i=1$) η vs. m_e^{-1} , (b) Radiative ($Q_i=1$) η vs. m_h^{-1} . (c) η (at $Q_i=10^{-4}$) vs. m_e^{-1} , (d) η (at $Q_i=10^{-4}$) vs. m_h^{-1} . Color scale represents difference between direct gap and fundamental electronic band gap (ΔE_g^{dir}). η is calculated at $T=298$ K, and film thickness $d=1$ μm .

To evaluate the potential of a given PV material taking both optical absorption and carrier transport into account, we plot in Fig. 5 the PCE (for film thickness $d=1$ μm) against the inverse carrier effective mass of electrons (left) and holes (right). We label all materials that are "Pareto optimal" [41], i.e. for which no other material has higher values for both η and $1/m^*$. In the upper panels (a-b), η is evaluated in the radiative limit, $Q_i=1$, while we use a more realistic value of $Q_i=10^{-4}$ in the lower panels. The color scale represents the difference in direct and indirect band gap, ΔE_g^{dir} .

Some well known thin film PV materials, like CdTe, PbS and InP, as well as emerging materials, like GaSe, SnS, are found on the Pareto fronts in Fig. 5. Not unexpectedly, most of the promising PV materials have $\Delta E_g^{\text{dir}} < 0.5$ eV (blueish rather than reddish color). In Table I, we list all materials satisfying the criteria: $\eta_{\text{max}} > 20\%$ at $Q_i=1$, $d=1$ μm , $T=298$ K, and $m_e, m_h < m_0$. The materials are ordered according to the size of ΔE_g^{dir} .

While there is a tendency for smaller ΔE_g^{dir} to correlate with higher η , there are also clear deviations from this trend. For example, SnSe₂ with $\Delta E_g^{\text{dir}} = 0.41$ eV has $\eta = 30.61\%$, which is not far from the highest value $\eta = 33.07\%$ obtained for the direct band gap material InP.

In Fig. 6, we show how η varies with the amount of non-radiative recombination for some of the most promising materials. Here we vary Q_i from 1 (radiative limit) to 10^{-7} (dominant non-radiative recombination) and plot the corresponding maximum efficiency, η_{max} , and optimal film thickness, $d@ \eta_{\text{max}}$, indicated by the color code. We also show η_{rad} at 1 μm film thickness. In the SI we provide more detailed plots of η versus film thickness for various values for Q_i (as in Fig. 3(c)) for all the materials in Fig. 6. The order of the materials in Fig. 6 is such that ΔE_g^{dir} increases from left to right, i.e. direct band gap materials, indicated by green shaded region, are found to the left. Considering only η_{rad} when com-

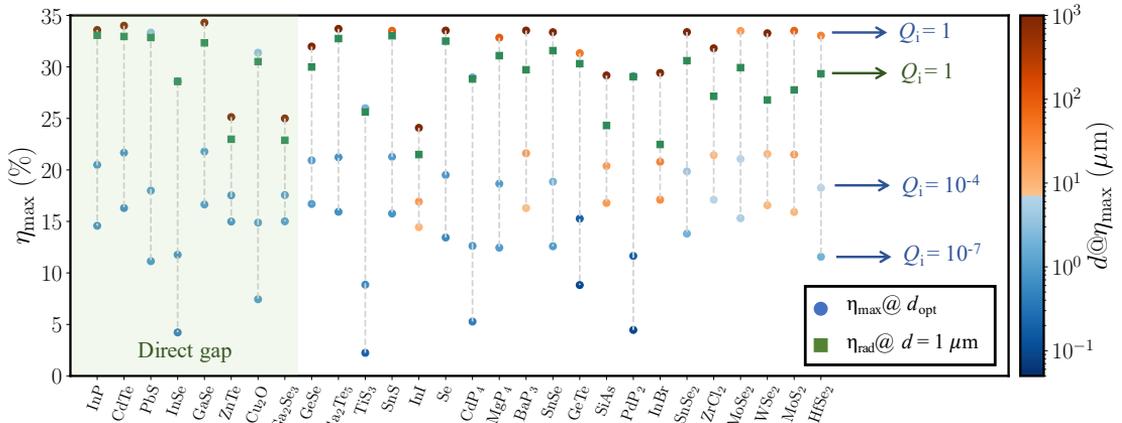


FIG. 6. Maximum radiative photovoltaic device efficiency (η_{\max}) as a function of different internal quantum efficiencies ($\log_{10} Q_i = 0, -4$ and -7) for the selected compounds ($\eta_{\max} > 20\%$ at $Q_i = 1$, $d = 1 \mu\text{m}$, $T = 298 \text{ K}$, $m_e, m_h < m_0$). Color scale represents corresponding film thickness (d) at η_{\max} . Additionally, η_{rad} [η at $Q_i=1$] at $d = 1 \mu\text{m}$ is shown using \square symbol. Light green shaded region represents the direct band gap materials.

paring a direct and indirect gap compound is not sufficient, as Q_i can be different for different materials. For example, CdTe has $\eta_{\text{SQ}} \approx 33\%$, whereas the known highest device efficiency is 22.1%. [1] This efficiency is slightly higher than η_{\max} at $Q_i=10^{-4}$. The measured Q_i on the highest-efficiency CdTe solar cell is 8×10^{-3} [4], indicating that some additional loss mechanisms are present besides non-radiative recombination. At the same time, we can see a number of indirect gap materials which show $\eta_{\max} > 20\%$, for practical Q_i (10^{-2} - 10^{-4}) ranges. The optimal device thickness in case of indirect materials are higher mainly due to low absorption at the indirect gap. We can see a trend that with increasing $\Delta E_{\text{g}}^{\text{dir}}$ the optimal device thickness (at different Q_i values) increases, which means with larger $\Delta E_{\text{g}}^{\text{dir}}$ the suitability for thin film PV technology might reduce. Still, there are a number of indirect gap compounds showing excellent η_{\max} at thin film thickness range. These materials are discussed in the next section.

There has been a lot of discussion whether a small indirect gap can be beneficial for PV application. [19, 20] With this study, we clarify how an indirect and direct gap material compare in the radiative limit for small film thicknesses (see SI section S7 for a detailed discussion). We find that at smaller film thicknesses ($d = 1 \mu\text{m}$), in the radiative limit (high mobility) adding a small indirect gap to a direct gap material will never show higher η_{rad} , as the gain in absorption (higher J_{sc}) will be counteracted by higher radiative loss (lower V_{oc}). However, an indirect gap material with small $\Delta E_{\text{g}}^{\text{dir}}$ can exhibit higher η_{rad} compared to a direct gap material with same fundamental gap, as long as the gap is below the SQ maximum

(1.1 eV). For example, an indirect gap material with $E_{\text{g}} = 0.95 \text{ eV}$ and $\Delta E_{\text{g}}^{\text{dir}} = 0.1 \text{ eV}$ can in principle show higher η_{rad} compared to a material with direct E_{g} of 0.95 eV. This is mainly due to the opposite effect (higher gain in V_{oc} compared to loss in J_{sc}) as discussed for the previous case. With this study, using our formalism we can provide real material examples here. For example, InSe is a direct gap material with $E_{\text{g}} = 0.89 \text{ eV}$ showing η_{rad} of 28.60%. In comparison, CdP₄, which has an indirect gap of 0.86 eV and direct gap of 1.01 eV, shows higher η_{rad} of 28.83% (see Table I). More examples can be found in SI section S7 and related tables.

D. Discussion of final list of materials

Among the materials that pass all screening criteria, we discard the ones that have been shown to be unstable in air (elemental I, Cs₂P₃, Rb₂P₃, RbSb₂, KSb₂, Na₃Sb, NaSe, YN). We also leave out the Tl-containing materials (Tl₂O, Tl₂Te₃) due to serious toxicity concerns. Finally, we discard the Os- and Ru-containing materials (OsAs₂, OsP₂, RuP₂) due to their very low earth abundance.

The final list of screened materials is presented in Table 1, grouped by maturity level. We note that all of the compounds have an ICSD-id, which means that they have been previously synthesized. CdTe, InP, and PbS are established materials which have been particularly successful in single-junction, multijunction, and quantum dot solar cells, respectively.

Among the emerging materials identified by our work,

TABLE I. Various relevant parameters (Compound, space group, Dimensionality (dim), Band gap (E_g) (calculated using GLLB-SC functional), difference between direct and indirect gap (ΔE_g^{dir}), effective masses (m_e and m_h), device efficiencies (η_1 [$d=1 \mu\text{m}$, $Q_i=1$] and η_2 [$d=1 \mu\text{m}$, $Q_i=10^{-4}$]), ICSD-id, etc.) for the selected compounds for thin film PV technology.

Established									
Formula ^[Ref]	space group	dim	E_g	ΔE_g^{dir} (eV)	$m_e(m_0)$	$m_h(m_0)$	η_1 (%)	η_2 (%)	ICSD-id
InP ^[3]	F-43m	3D	1.32	0.00	0.05	0.89	33.07	20.45	41443
CdTe ^[1]	F-43m	3D	1.43	0.00	0.05	0.97	32.96	21.59	31844
PbS ^[42]	Fm-3m	0D	1.15	0.00	0.24	0.31	32.86	17.99	62192
Emerging, PCE > 5%									
Formula ^[Ref]	space group	dim	E_g	ΔE_g^{dir} (eV)	$m_e(m_0)$	$m_h(m_0)$	η_1 (%)	η_2 (%)	ICSD-id
InSe ^[43]	P6 ₃ /mmc	2D	0.89	0.00	0.07	0.10	28.60	11.63	30377
Cu ₂ O ^[44]	Pn-3m	3D	1.01	0.00	0.93	0.41	30.52	14.89	261853
Se ^[45]	P3 ₁ 21	1D	1.24	0.12	0.17	0.32	32.51	19.51	164271
WSe ₂ ^[46]	P6 ₃ /mmc	2D	1.48	0.66	0.47	0.84	26.80	18.74	40752
Emerging, PCE < 5%									
Formula ^[Ref]	space group	dim	E_g	ΔE_g^{dir} (eV)	$m_e(m_0)$	$m_h(m_0)$	η_1 (%)	η_2 (%)	ICSD-id
GaSe ^[47]	P6 ₃ /mmc	2D	1.47	0.00	0.11	0.21	32.34	21.78	41978
ZnTe ^[48]	P3 ₁	3D	2.00	0.00	0.08	0.31	22.98	17.53	80076
Ga ₂ Se ₃ ^[49]	Cc	3D	2.00	0.00	0.15	0.35	22.88	17.56	35028
GeSe ^[50]	Pnma	2D	1.60	0.03	0.23	0.40	30.00	20.93	17006
SnS ^[51]	Pnma	2D	1.39	0.06	0.21	0.37	33.03	21.27	24376
InI ^[52]	Cmcm	0D	2.00	0.07	0.21	0.25	21.49	16.66	38129
SnSe ^[53]	Pnma	2D	1.20	0.18	0.18	0.27	31.58	18.72	16997
GeTe ^[54]	R3m	2D	0.99	0.20	0.08	0.07	30.32	14.72	655497
SnSe ₂ ^[55]	P-3m1	2D	1.23	0.41	0.31	0.60	30.61	19.31	43594
MoSe ₂ ^[56]	P6 ₃ /mmc	2D	1.34	0.58	0.56	0.98	29.93	19.72	49800
MoS ₂ ^[57]	P6 ₃ /mmc	2D	1.38	0.83	0.62	0.77	27.77	18.92	24000
HfSe ₂ ^[58]	P-3m1	2D	1.11	1.02	0.42	0.22	29.34	17.40	638902
Unexplored									
Formula	space group	dim	E_g	ΔE_g^{dir} (eV)	$m_e(m_0)$	$m_h(m_0)$	η_1 (%)	η_2 (%)	ICSD-id
Ga ₂ Te ₅	I4/m	3D	1.42	0.03	0.17	0.37	32.75	21.20	1085
TiS ₃	P2 ₁ /m	2D	0.79	0.03	0.82	0.67	25.62	8.72	42072
CdP ₄	P2 ₁ /c	3D	0.86	0.15	0.46	0.18	28.83	12.58	25605
MgP ₄	P2 ₁ /c	3D	1.10	0.17	0.41	0.36	31.10	18.66	113
BaP ₃	C2/m	3D	1.41	0.18	0.39	0.59	29.73	20.55	23618
SiAs	C2/m	2D	1.75	0.20	0.37	0.65	24.31	18.36	43227
PdP ₂	C2/c	3D	0.88	0.36	0.15	0.73	29.05	10.83	166275
InBr	Cmcm	0D	1.73	0.37	0.22	0.31	22.47	17.12	55188
ZrCl ₂	R3m	2D	1.59	0.57	0.72	0.63	27.16	19.77	20144

there are a few with record PCE above 5%. They are InSe [43], Cu₂O [44], elemental Se[45], and WSe₂ [46]. The other emerging materials have been investigated for PV applications, but their record PCE is below 5% or they have not yet been incorporated in a working solar cell. Except for InI [52], all materials in this

group are chalcogenides, and the large majority of them have 2D (layered) structures. Group IV, Ga-based, and transition-metal chalcogenides are particularly well represented. In many cases, the development of these layered compounds for PV applications has focused on extremely thin films, to take advantage of the enhanced light ab-

sorption that many of these compounds experience in atomically-thin form. ZnTe is a well known hole contact layer in CdTe PV technology, but it has also been tested as a wide-gap photoabsorber for tandem cell applications, although efficiencies are still very low [48].

The "rediscovery" of many already known PV materials lends credibility to our approach. We note that a few binary compounds with notable PV efficiencies, like Sb₂S₃ and Zn₃P₂, are missed by our approach because they have more than 12 atoms in their unit cell. We miss GaAs due to underestimation band gap using GLLB-SC functional (0.68 eV). We also miss CdSe because the calculated hole effective mass is higher than m_0 .

We find nine compounds that have not been previously studied as PV materials. We label them "unexplored" in Table 1. To the best of our knowledge, none of these compounds have been synthesized as thin films, except for InBr [59]. Intriguingly, none of these compounds belongs to traditional semiconductor families. What most of these semiconductors have in common is the existence of homoelement bonds, i.e., bonds between the same elements. This feature is absent in the currently known compound semiconductors for PV (III-V, II-VI, perovskites, and derivatives thereof). Due to this qualitative difference, the properties of these semiconductors with hybrid (homo- and heteroelement) bonding cannot be readily extrapolated from the properties of well-known semiconductors with only heteroelement bonds. Thus, these materials could be a new frontier of PV materials discovery.

Four compounds (CdP₄, MgP₄, BaP₃, and PdP₂) are phosphorus-rich phosphides, containing metal-P as well as P-P bonds. Due to these additional covalent bonds, phosphorus is in a higher oxidation state (ranging from -1 to -0.5) than the classic -3 state found in optoelectronic phosphides, e.g., InP, Zn₃P₂, and the II-IV-P₂ ternaries. In general, phosphorus-rich phosphides have not been studied as potential PV materials before. They had hardly been grown in thin-film form until a very recent report of single-phase CuP₂ films in the P₂₁/c structure by reactive sputtering and soft annealing in an inert atmosphere [60].

In a similar spirit, TiS₃ can be classified as a sulfur-rich sulfide due to the presence of both metal-S and S-S bonds. TiS₃ has only been minimally investigated in photoelectrochemical cells and not in thin-film form [61]. In SiAs, all Si atoms are bonded to 3 As atoms and 1 Si atom, so Si can be considered as in the +3 oxidation state, thus explaining the 1:1 stoichiometry. There is one report of SiAs in a photoelectrochemical cell, but the sample was not in thin-film form [62]. In Ga₂Te₅, 20% of the Te atoms are only bonded to other 4 Te atoms (Te-Te bonds), and the remaining 80% is bonded to one Te atom and 2 Ga atoms.

On the other hand, the two halide compounds InBr and ZrCl₂ do not have homoelement bonds and are quite different from each other. InBr should be qualitatively similar to the already known InI [52], but with a larger

ΔE_g^{dir} and a band gap closer to the SQ optimum. ZrCl₂ is one of the first transition metal halides to be identified for potential PV applications. It has the same structure as 3R-MoS₂ with Zr in the unusual +2 oxidation state. Unlike most of the known PV semiconductors, ZrCl₂ has a lone valence band. It also has a large ΔE_g^{dir} and reasonably low effective masses along its 2D layers.

Considering a combination of crustal abundance and elasticity of production [63], we estimate that only five compounds from the "unexplored" family could potentially be scaled to terawatt levels. These are TiS₃, MgP₄, BaP₃, SiAs, and ZrCl₂. These compounds deserve further investigation.

IV. CONCLUSION

In summary, we propose an efficient methodology to calculate phonon assisted absorption, which enables us to accurately evaluate optical absorption for indirect band gap materials. Using this method, along with various relevant photovoltaic (PV) descriptors (such as device efficiency, effective masses, internal luminescence quantum efficiency, etc.) we evaluate the potential of 127 experimentally known binary semiconductors (of which 97 are indirect band gap) for thin film PV. After screening the materials based on the PV descriptors and considering toxicity, material abundance and air stability, we identify 28 promising materials of which 20 have an indirect gap. This list includes well established PV materials like InP, CdTe, PbS, as well as emerging ones such as, elemental Se, SnS, SnSe, SnSe₂, lending credibility to our approach. Additionally, we discover 9 previously unexplored compounds with potential for thin film PV. Most of these compounds show homoelemental bonds and the constituent elements show deviation from their well known oxidation states. These materials (and their respective stoichiometric classes) have scarcely, if ever, been explored for photovoltaics and could become a new frontier of next generation PV materials research. Further theoretical work on these materials should address their defect tolerance, carrier mobilities and lifetimes. Such studies can benefit from our open database containing the calculated atomic and electronic structures as well as basic PV descriptors for the 127 materials investigated in this work.

V. METHODOLOGY

A. Computational Details

First principles calculations are performed using Density Functional Theory (DFT)[64] within the Projector-Augmented Wave (PAW) formalism[65] as implemented in GPAW code[66, 67], in combination with Atomic Simulation Environment (ASE)[68]. A plane wave cutoff energy of 800 eV and k-mesh density of 8 Å⁻¹ are employed

for geometry optimization. Electronic structure calculations are performed using Perdew-Burke-Ernzerhof (PBE)[30] exchange correlation functional with double zeta polarized (dzp) basis set[69]. The accuracy of dzp basis set was tested in a previous study and reported to produce phonon modes with fair accuracy [70]. The phonon modes are evaluated via calculating the dynamical matrices within the harmonic approximation. Force constant matrices are calculated using small displacement approach[71] as implemented in ASE. Electron-phonon matrix elements are evaluated based on calculation of gradients of effective potential with finite difference technique [39] as implemented in GPAW. As our approximation involves only zone-center (Γ point) phonons, we use primitive unit cells for respective calculations. Next, the momentum matrix elements are obtained using the finite difference technique as implemented in GPAW. A real-space grid spacing of 0.2 Å is chosen for these calculations. We have used $16 \times 16 \times 16$ (8 \AA^{-1}) k-mesh (density) for systems with less (more) than 4 atoms in the unit cell, when evaluating the electron-phonon coupling matrix elements. For phonon calculations, 8 \AA^{-1} (6 \AA^{-1}) k-mesh density is used for unit cells with less (more) than 4 atoms and the forces are converged until 10^{-7} eV/\AA . The width of the Fermi-Dirac smearing is kept at 50 meV. While calculating the probability of transition, the Dirac delta function is replaced with a Gaussian function with a smearing width of 40 meV, which accounts for inhomogeneous spectral broadening.

The choice of PBE exchange-correlation functional which is known to slightly overestimate the lattice parameters and significantly underestimate the band gap values, is expected to have some effect on the calculated absorption coefficient and efficiency. In order to account for the band gap underestimation, we have calculated the band gaps with GLLB-SC functional[33], which is known to produce accurate band gaps[72] and is highly affordable computationally. During the calculation of probability of transition, the conduction bands are scissor shifted to match the GLLB-SC gap. The effect of temperature is incorporated only into the phonon occupation numbers. We have considered $T = 298 \text{ K}$ for all the calculations. Additional finite temperature effects (e.g. lattice expansion and electron-phonon renormalization, etc.) may exhibit effects on the absorption coefficient and as such device efficiency slightly. Optical absorption coefficient beyond the direct gap is obtained via calculating the frequency dependent dielectric function within the Random Phase Approximation (RPA) as implemented in GPAW. The computational workflow was constructed using the Atomic Simulation Recipes (ASR)[73] and executed using the MyQueue[74] task scheduler frontend.

We acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program Grant No. 773122 (LIMA) and Grant agreement No. 951786 (NOMAD CoE). K. S. T. is a Villum Investigator supported by VILLUM FONDEN (grant no. 37789).

-
- [1] M. A. Green, E. D. Dunlop, J. Hohl-Ebinger, M. Yoshita, N. Kopidakis, K. Bothe, D. Hinken, M. Rauer, and X. Hao, Solar cell efficiency tables (version 60), *Progress in Photovoltaics: Research and Applications* **30**, 687 (2022).
- [2] K. Yoshikawa, H. Kawasaki, W. Yoshida, T. Irie, K. Konishi, K. Nakano, T. Uto, D. Adachi, M. Kanematsu, H. Uzu, and K. Yamamoto, Silicon heterojunction solar cell with interdigitated back contacts for a photoconversion efficiency over 26%, *Nature Energy* **2**, 17032 (2017).
- [3] M. Wanlass, Systems and methods for advanced ultra-high-performance in p solar cells, National Renewable Energy Lab.(NREL), Golden, CO (United States) (2017).
- [4] M. A. Green and A. W. Ho-Baillie, Pushing to the limit: radiative efficiencies of recent mainstream and emerging solar cells, *ACS Energy Letters* **4**, 1639 (2019).
- [5] M. Nakamura, K. Yamaguchi, Y. Kimoto, Y. Yasaki, T. Kato, and H. Sugimoto, Cd-free Cu(In,Ga)(Se,S) 2 thin-film solar cell with record efficiency of 23.35%, *IEEE Journal of Photovoltaics* **9**, 1863 (2019).
- [6] H. Min, D. Y. Lee, J. Kim, G. Kim, K. S. Lee, J. Kim, M. J. Paik, Y. K. Kim, K. S. Kim, M. G. Kim, T. J. Shin, and S. Il Seok, Perovskite solar cells with atomically coherent interlayers on SnO₂ electrodes, *Nature* **598**, 444 (2021).
- [7] W. Shockley and H. J. Queisser, Detailed balance limit of efficiency of p-n junction solar cells, *Journal of applied physics* **32**, 510 (1961).
- [8] T. Tiedje, E. Yablonovitch, G. D. Cody, and B. G. Brooks, Limiting efficiency of silicon solar cells, *IEEE Transactions on electron devices* **31**, 711 (1984).
- [9] L. Yu and A. Zunger, Identification of potential photovoltaic absorbers based on first-principles spectroscopic screening of materials, *Physical review letters* **108**, 068701 (2012).
- [10] B. Blank, T. Kirchartz, S. Lany, and U. Rau, Selection metric for photovoltaic materials screening based on detailed-balance analysis, *Physical review applied* **8**, 024032 (2017).
- [11] S. Kim, J. A. Márquez, T. Unold, and A. Walsh, Upper limit to the photovoltaic efficiency of imperfect crystals from first principles, *Energy & Environmental Science* **13**, 1481 (2020).
- [12] X.-G. Zhao, J.-H. Yang, Y. Fu, D. Yang, Q. Xu, L. Yu, S.-H. Wei, and L. Zhang, Design of lead-free inorganic halide perovskites for solar cells via cation-transmutation, *Journal of the American Chemical Society* **139**, 2630 (2017).
- [13] K. Kuhar, M. Pandey, K. S. Thygesen, and K. W. Jacobsen, High-throughput computational assessment of previously synthesized semiconductors for photovoltaic and photoelectrochemical devices, *ACS Energy Letters* **3**, 436 (2018).
- [14] K. Choudhary, M. Bercx, J. Jiang, R. Pachter, D. Lamoen, and F. Tavazza, Accelerated discovery of efficient solar cell materials using quantum and machine-learning methods, *Chemistry of Materials* **31**, 5900 (2019).

- [15] Y. Li and K. Yang, High-throughput computational design of organic–inorganic hybrid halide semiconductors beyond perovskites for optoelectronics, *Energy & Environmental Science* **12**, 2233 (2019).
- [16] S. Luo, T. Li, X. Wang, M. Faizan, and L. Zhang, High-throughput computational materials screening and discovery of optoelectronic semiconductors, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **11**, e1489 (2021).
- [17] D. Dahliah, G. Brunin, J. George, V.-A. Ha, G.-M. Rignanese, and G. Hautier, High-throughput computational search for high carrier lifetime, defect-tolerant solar absorbers, *Energy & Environmental Science* **14**, 5057 (2021).
- [18] Y. Gan, N. Miao, P. Lan, J. Zhou, S. R. Elliott, and Z. Sun, Robust design of high-performance optoelectronic chalcogenide crystals from high-throughput computation, *Journal of the American Chemical Society* **144**, 5878 (2022).
- [19] E. M. Hutter, M. C. Gélvez-Rueda, A. Osharov, V. Bulović, F. C. Grozema, S. D. Stranks, and T. J. Savenije, Direct–indirect character of the bandgap in methylammonium lead iodide perovskite, *Nature materials* **16**, 115 (2017).
- [20] T. Kirchartz and U. Rau, Decreasing radiative recombination coefficients via an indirect band gap in lead halide perovskites, *The Journal of Physical Chemistry Letters* **8**, 1265 (2017).
- [21] J. Noffsinger, E. Kioupakis, C. G. Van de Walle, S. G. Louie, and M. L. Cohen, Phonon-assisted optical absorption in silicon from first principles, *Physical review letters* **108**, 167402 (2012).
- [22] Y. Kang, H. Peelaers, K. Krishnaswamy, and C. G. Van de Walle, First-principles study of direct and indirect optical absorption in basno3, *Applied Physics Letters* **112**, 062106 (2018).
- [23] B. Monserrat, C. E. Dreyer, and K. M. Rabe, Phonon-assisted optical absorption in basno 3 from first principles, *Physical Review B* **97**, 104310 (2018).
- [24] J. Menéndez, D. J. Lockwood, J. C. Zwickels, and M. Noël, Phonon-assisted optical absorption in germanium, *Physical Review B* **98**, 165207 (2018).
- [25] I. Bravić and B. Monserrat, Finite temperature optoelectronic properties of bas from first principles, *Physical Review Materials* **3**, 065402 (2019).
- [26] V.-A. Ha, B. Karasulu, R. Maezono, G. Brunin, J. B. Varley, G.-M. Rignanese, B. Monserrat, and G. Hautier, Boron phosphide as a p-type transparent conductor: Optical absorption and transport through electron-phonon coupling, *Physical Review Materials* **4**, 065401 (2020).
- [27] Y. Ge, W. Wan, X. Guo, and Y. Liu, Direct and indirect optical absorptions of cubic bas and bsb, *Optics Express* **28**, 238 (2020).
- [28] C.-c. Jian, X. Ma, J. Zhang, and X. Yong, Strained mosi2n4 monolayers with excellent solar energy absorption and carrier transport properties, *The Journal of Physical Chemistry C* **125**, 15185 (2021).
- [29] J. Willis, I. Bravić, R. R. Schnepf, K. N. Heinselman, B. Monserrat, T. Unold, A. Zakutayev, D. O. Scanlon, and A. Crovetto, Prediction and realisation of high mobility and degenerate p-type conductivity in cacup thin films, *Chemical Science* **13**, 5872 (2022).
- [30] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Physical review letters* **77**, 3865 (1996).
- [31] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu, *The Journal of chemical physics* **132**, 154104 (2010).
- [32] M. K. Svendsen, H. Sugimoto, A. Assadillayev, D. Shima, M. Fujii, K. S. Thygesen, and S. Raza, Computational discovery and experimental demonstration of boron phosphide ultraviolet nanoresonators, *Advanced Optical Materials* , 2200422 (2022).
- [33] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, Kohn-sham potential with discontinuity for band gap materials, *Physical Review B* **82**, 115106 (2010).
- [34] W. Shockley and W. Read Jr, Statistics of the recombinations of holes and electrons, *Physical review* **87**, 835 (1952).
- [35] R. N. Hall, Electron-hole recombination in germanium, *Physical review* **87**, 387 (1952).
- [36] L. Hultdt, Band-to-band auger recombination in indirect gap semiconductors, *physica status solidi (a)* **8**, 173 (1971).
- [37] P. M. Larsen, M. Pandey, M. Strange, and K. W. Jacobsen, Definition of a scoring parameter to identify low-dimensional materials components, *Physical Review Materials* **3**, 034003 (2019).
- [38] F. A. Rasmussen and K. S. Thygesen, Computational 2d materials database: electronic structure of transition-metal dichalcogenides and oxides, *The Journal of Physical Chemistry C* **119**, 13169 (2015).
- [39] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phonon-limited mobility in n-type single-layer mos 2 from first principles, *Physical Review B* **85**, 115317 (2012).
- [40] K. Kaasbjerg, Atomistic t-matrix theory of disordered two-dimensional materials: Bound states, spectral properties, quasiparticle scattering, and transport, *Physical Review B* **101**, 045433 (2020).
- [41] M. Khoroshiltseva, D. Slanzi, and I. Poli, A pareto-based multi-objective optimization algorithm to design energy-efficient shading devices, *Applied energy* **184**, 1400 (2016).
- [42] M.-J. Choi, F. P. García de Arquer, A. H. Proppe, A. Seifitokaldani, J. Choi, J. Kim, S.-W. Baek, M. Liu, B. Sun, M. Biondi, *et al.*, Cascade surface modification of colloidal quantum dot inks enables efficient bulk homojunction photovoltaics, *Nature communications* **11**, 1 (2020).
- [43] J. Martínez-Pastor, A. Segura, J. L. Valdés, and A. Chevy, Electrical and photovoltaic properties of indium-tin-oxide/ p -InSe/Au solar cells, *Journal of Applied Physics* **62**, 1477 (1987).
- [44] T. Minami, Y. Nishi, and T. Miyata, Efficiency enhancement using a Zn1-xGex-O thin film as an n-type window layer in Cu2O-based heterojunction solar cells, *Applied Physics Express* **9**, 052301 (2016).
- [45] T. K. Todorov, S. Singh, D. M. Bishop, O. Gunawan, Y. S. Lee, T. S. Gershon, K. W. Brew, P. D. Antunez, and R. Haight, Ultrathin high band gap solar cells with improved efficiencies from the world’s oldest photovoltaic material, *Nature communications* **8**, 1 (2017).
- [46] K.-H. Kim, M. Andreev, S. Choi, J. Shim, H. Ahn, J. Lynch, T. Lee, J. Lee, K. N. Nazif, A. Kumar, *et al.*, High-efficiency wse2 photovoltaic devices with electron-

- selective contacts, *ACS nano* **16**, 8827 (2022).
- [47] R. Ishikawa, P. J. Ko, R. Anzo, C. L. Woo, G. Oh, and N. Tsuboi, Photovoltaic characteristics of gase/mose2 heterojunction devices, *Nanoscale Research Letters* **16**, 1 (2021).
- [48] W. Wang, J. D. Phillips, S. J. Kim, and X. Pan, ZnO/ZnSe/ZnTe Heterojunctions for ZnTe-Based Solar Cells, *Journal of Electronic Materials* **40**, 1674 (2011).
- [49] C.-H. Ho, Ga₂Se₃ defect semiconductors: The study of direct band edge and optical properties, *ACS omega* **5**, 18527 (2020).
- [50] S.-C. Liu, Z. Li, J. Wu, X. Zhang, M. Feng, D.-J. Xue, and J.-S. Hu, Boosting the efficiency of gese solar cells by low-temperature treatment of pn junction, *Science China Materials* **64**, 2118 (2021).
- [51] P. Sinsermuksakul, L. Sun, S. W. Lee, H. H. Park, S. B. Kim, C. Yang, and R. G. Gordon, Overcoming efficiency limitations of sns-based solar cells, *Advanced Energy Materials* **4**, 1400496 (2014).
- [52] W. A. Dunlap-Shohl, I. G. Hill, Y. Yan, and D. B. Mitzi, Photovoltaic Effect in Indium(I) Iodide Thin Films, *Chemistry of Materials* **30**, 8226 (2018).
- [53] R. Nandi, P. S. Pawar, K. E. Neerugatti, J. Y. Cho, S. Kim, S. H. Cho, Y. S. Lee, and J. Heo, Vapor-transport-deposited orthorhombic-snse thin films: A potential cost-effective absorber material for solar-cell applications, *Solar RRL* **6**, 2100676 (2022).
- [54] A. M. Jones, R. A. Coutu, and R. A. Lake, Design and analysis of novel ge-gete pn junction for photovoltaics, in *2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS)* (IEEE, 2016) pp. 343–348.
- [55] X. Yu, J. Zhu, Y. Zhang, J. Weng, L. Hu, and S. Dai, Sns₂ quantum dot sensitized solar cells prepared employing molecular metal chalcogenide as precursors, *Chemical communications* **48**, 3324 (2012).
- [56] S. Memaran, N. R. Pradhan, Z. Lu, D. Rhodes, J. Ludwig, Q. Zhou, O. Ogunsolu, P. M. Ajayan, D. Smirnov, A. I. Fernandez-Dominguez, *et al.*, Pronounced photovoltaic response from multilayered transition-metal dichalcogenides pn-junctions, *Nano letters* **15**, 7532 (2015).
- [57] S. Wi, H. Kim, M. Chen, H. Nam, L. J. Guo, E. Meyerhofer, and X. Liang, Enhancement of photovoltaic response in multilayer mos₂ induced by plasma doping, *ACS nano* **8**, 5270 (2014).
- [58] A. M. Afzal, S. Mumtaz, M. Z. Iqbal, M. W. Iqbal, A. Manzoor, G. Dastgeer, M. J. Iqbal, Y. Javed, R. Khan, N. A. Shad, *et al.*, Fast and high photoresponsivity gallium telluride/hafnium selenide van der waals heterostructure photodiode, *Journal of Materials Chemistry C* **9**, 7110 (2021).
- [59] A. D. Brothers, F. J. Feuerborn, and S. E. Bauer, Excitons in Indium Halide Thin Films, *physica status solidi (b)* **158**, 503 (1990).
- [60] A. Crovetto, D. Kojda, F. Yi, K. N. Heinselman, D. A. LaVan, K. Habicht, T. Unold, and A. Zakutayev, Crystallize it before it diffuses: Kinetic stabilization of thin-film phosphorus-rich semiconductor cup₂, *Journal of the American Chemical Society* **0**, null (0), pMID: 35822809, <https://doi.org/10.1021/jacs.2c04868>.
- [61] O. Gorochov, A. Katty, N. Le Nagard, C. Levy-Clement, and D. Schleich, Photoelectrochemical study of TiS₃ in aqueous solution, *Materials Research Bulletin* **18**, 111 (1983).
- [62] H. Lewerenz and H. Wetzell, Photoeffect on sias electrodes, *Journal of the Electrochemical Society* **130**, 1228 (1983).
- [63] P. C. K. Vesborg and T. F. Jaramillo, Addressing the terawatt challenge: scalability in the supply of chemical elements for renewable energy, *RSC Advances* **2**, 7933 (2012).
- [64] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Physical review* **140**, A1133 (1965).
- [65] P. E. Blöchl, Projector augmented-wave method, *Physical review B* **50**, 17953 (1994).
- [66] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Real-space grid implementation of the projector augmented wave method, *Physical Review B* **71**, 035109 (2005).
- [67] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dulak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. Hansen, *et al.*, Electronic structure calculations with gpaw: a real-space implementation of the projector augmented-wave method, *Journal of physics: Condensed matter* **22**, 253202 (2010).
- [68] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, *et al.*, The atomic simulation environment—a python library for working with atoms, *Journal of Physics: Condensed Matter* **29**, 273002 (2017).
- [69] A. H. Larsen, M. Vanin, J. J. Mortensen, K. S. Thygesen, and K. W. Jacobsen, Localized atomic basis set in the projector augmented wave method, *Physical Review B* **80**, 195112 (2009).
- [70] A. Taghizadeh, U. Leffers, T. G. Pedersen, and K. S. Thygesen, A library of ab initio raman spectra for automated identification of 2d materials, *Nature communications* **11**, 1 (2020).
- [71] D. Alfè, Phon: A program to calculate phonons using the small displacement method, *Computer Physics Communications* **180**, 2622 (2009).
- [72] I. E. Castelli, F. Hüser, M. Pandey, H. Li, K. S. Thygesen, B. Seger, A. Jain, K. A. Persson, G. Ceder, and K. W. Jacobsen, New light-harvesting materials using accurate and efficient bandgap calculations, *Advanced Energy Materials* **5**, 1400915 (2015).
- [73] M. Gjerding, T. Skovhus, A. Rasmussen, F. Bertoldo, A. H. Larsen, J. J. Mortensen, and K. S. Thygesen, Atomic simulation recipes: A python framework and library for automated workflows, *Computational Materials Science* **199**, 110731 (2021).
- [74] J. J. Mortensen, M. Gjerding, and K. S. Thygesen, Myqueue: Task and workflow scheduling system, *Journal of Open Source Software* **5**, 1844 (2020).

11.6 Paper [VI]: Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials,

Mark Kamper Svendsen, Sajid Ali, Nicolas Stenger, Kristian Sommer Thygesen, Jake Iles-Smith,
“Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials,”
Submitted to arXiv: arXiv:2207.10630.

Signatures of Non-Markovianity in Cavity-QED with Color Centers in 2D Materials

Mark Kamper Svendsen,¹ Sajid Ali,¹ Nicolas Stenger,^{2,3,4} Kristian Sommer Thygesen,^{1,3} and Jake Iles-Smith^{5,6}

¹*CAMD, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

²*Department of electrical and photonics engineering,*

Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

³*Center for Nanostructured Graphene, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

⁴*NanoPhoton – Center for Nanophotonics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

⁵*Department of Physics and Astronomy, The University of Manchester,*

Oxford Road, Manchester M13 9PL, United Kingdom

⁶*Department of Electrical and Electronic Engineering,*

The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

(Dated: January 2022)

Light-matter interactions of defects in two dimensional materials are expected to be profoundly impacted by strong coupling to phonons. In this work, we combine *ab initio* calculations of a defect in hBN, with a fully quantum mechanical and numerically exact description of a cavity-defect system to elucidate this impact. We show that even at weak light-matter coupling, the dynamical evolution of the cavity-defect system has clear signatures of non-markovian phonon effects, and that linear absorption spectra show the emergence of hybridised light-matter-phonon states in regimes of strong light-matter coupling. We emphasise that our methodology is general, and can be applied to a wide variety of material/defect systems.

Single-photon emission has been observed from a broad range of two dimensional materials (2DM) [1–7]. Of particular interest to quantum technologies is the emission from defect complexes in hexagonal Boron Nitride (hBN) [8–13], owing to the materials wide band-gap [14] and stability of the emitters [10, 15]. Not only does the precise nature of these defect complexes remain under scrutiny [10, 11, 16], but also their quantum optical properties remain largely unexplored, particularly in regimes of strong light-matter coupling [17, 18]. Crucially, the optical properties of condensed matter systems cannot be understood through the standard theory of quantum optics, since strong system-environment interactions lead to a breakdown of the underpinning perturbative methods [19, 20]. Of particular relevance to the optical properties of defect complexes in 2DM are strong electron-phonon interactions. Specifically, defect emission spectra typically show sharp and well resolved phonon sidebands [21–24], a hallmark of vibronic state formation and long lived correlations between electronic and vibrational degrees of freedom. It is therefore crucial that any theory describing the dynamical or optical behaviour of a 2DM defect complex, accurately accounts for electron-phonon interactions.

To this end, we study the cavity quantum electrodynamics (cQED) of defects in 2DM [17, 18], focusing on two colour centres in hBN proposed as single-photon emitters, the $C_B V_N$ [10, 25] and $C_2 C_N$ defect complexes [26, 27]. We accurately account for electron-phonon interactions by combining *ab initio* calculations of the defects, with a fully quantum mechanical and numerically exact description of the dynamics of the defect states interacting with a single mode optical cavity. The former provides an atomistic description of

the phonon modes of the material system. The latter employs the time evolved matrix product operator (TEMPO) algorithm [28], where non-Markovian influences of the phonon environment are encoded within a tensor network. When combined with tensor compression methods, TEMPO provides an efficient approach for calculating the reduced dynamics of the system [29–31], which we extend to extract the linear response spectrum of the defect-cavity system. This hybrid approach is inspired by [32] and allows us to maintain the accuracy and insight of first-principles calculations, while enabling coherent dynamics to be captured in previously inaccessible regimes of light-matter coupling, information that is typically inaccessible to first-principle methods.

A central finding of this work is that electron-phonon interactions in the studied defect complexes lead to highly non-Markovian dynamics, even in regimes of moderate light-matter coupling strengths, which are accessible to state-of-the-art experiments. We attribute this behaviour to coherent coupling between the electronic states of the defect and high-quality (Q) factor phonon modes, leading to long-lived correlations between the system and its phonon environment. This provides a mechanism to optically manipulate mechanical modes of a condensed matter system that would otherwise not directly couple to light. Furthermore, from linear response spectra, we see evidence of hybridisation between cavity-defect polaritons and high- Q phonon modes when entering regimes of strong light-matter coupling. These states bare resemblance to recently predicted exciton-photon-phonon hybridisation in hBN [33], however in this instance they occur in a continuum of modes coupled to a defect state, and therefore will inherit non-linear features from the localised electronic states.

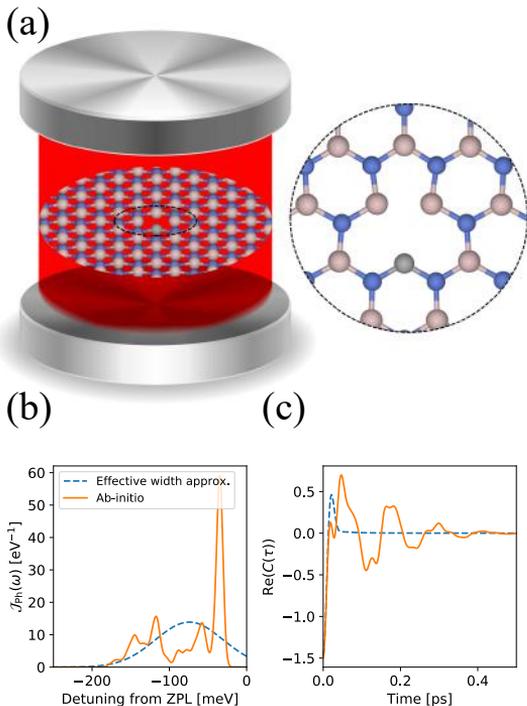


FIG. 1. (a) A schematic figure of the $C_B V_N$ defect in hBN interacting with a single quantised field mode. This mode could be plasmonic or photonic in nature. (b) A plot of the *ab initio* (solid) and effective width (dashed) spectral densities. Using the atomistic simulation, we ascribe the dominant peak at ~ 125 meV to a phonon breathing mode. (c) shows the bath correlation function for the *ab initio* (solid) and effective width (dashed) spectral densities. The sharp peaks of the *ab initio* spectral density leads to long lived oscillations in the correlation function, which are washed out when an effective width approximation is used.

For simplicity we focus on the $C_B V_N$ defect complex in the manuscript, leaving the discussion of $C_2 C_N$ to the supplementary information (SI). The electronic structure of the $C_B V_N$ defect consists of two separate manifolds [34]: the first consists of the $(2)^1 A_1$ and $(1)^1 B_1$ excited states, and a $(1)^1 A_1$ singlet ground state; the second contains the triplet states $(1)^3 B_1$ and $(2)^3 B_1$ [16], separated in energy by $\hbar\omega_e = 2$ eV [34]. By assuming the inter-system crossing due to spin-orbit coupling occurs on a much longer timescale (10 – 100 ns) than cavity-enhanced optical transitions (< 1 ns), we restrict our attention to the triplet state manifold, which reduces to a two level system with ground and excited states, $|g\rangle$ and $|e\rangle$ respectively [35]. The electronic transition interacts with a single mode optical cavity as shown schematically

in Fig. 1a, with a cavity resonance $\hbar\Omega_c$ and described by a Jaynes-Cummings type interaction:

$$H_S = \hbar\omega_e \sigma^\dagger \sigma + \hbar g (\sigma^\dagger a + \sigma a^\dagger) + \hbar\Omega_c a^\dagger a, \quad (1)$$

where $\sigma = |g\rangle\langle e|$ is the transition dipole operator, a (a^\dagger) is the annihilation (creation) operator of the cavity mode, and g is the light-matter coupling strength. Here we limit ourselves to the single excitation subspace, spanned by the basis $\{|g, 0\rangle, |e, 0\rangle, |g, 1\rangle\}$. The composite cavity-emitter system interacts with both vibrational and electromagnetic environments, such that the global Hamiltonian is written as $H = H_S + H_I^{EM} + H_I^{Ph} + H_B^{EM} + H_B^{Ph}$.

The emitter-cavity system couples to an external electromagnetic environment either through direct emission from the defect into free space, or leakage from the cavity mode. In the rotating wave approximation the interaction to the electromagnetic field takes the form $H_I^{EM} = \sum_j \sum_l (\hbar f_{j,1} \hat{A}_j^\dagger c_{j,1} + \text{h.c.})$, where $\hat{A}_1 = \sigma$ and $\hat{A}_2 = a$, and $c_{j,1}$ is the annihilation operator for the l^{th} mode of the j^{th} electromagnetic environment. In the limit of weak coupling to these fields, we can make a flat spectrum approximation, such that $f_{1,1} = \sqrt{2\pi\Gamma}$ and $f_{2,1} = \sqrt{2\pi\kappa}$ [36, 37]. Here Γ is the spontaneous emission rate from the TLE into free space, and κ is the cavity leakage rate. Using these approximations and tracing over the external EM fields, we can describe the optical contribution to the system evolution through the superoperator $\mathcal{V}_t = \exp(t \mathcal{L}_S)$, where $\mathcal{L}_S \rho = -i[H_S, \rho] + \Gamma L_\sigma \rho + \kappa L_a \rho$, and we have introduced the Lindblad dissipators $L_\sigma \rho = \sigma \rho \sigma^\dagger - \{\sigma^\dagger, \rho\}/2$. For a full derivation and discussion of these approximations please refer to the SI.

We take a linear electron-phonon interaction [38] of the form $H_I^{Ph} = \sigma^\dagger \sigma \sum_{\mathbf{k}} \hbar g_{\mathbf{k}} (b_{\mathbf{k}}^\dagger + b_{-\mathbf{k}})$, where $b_{\mathbf{k}}$ is the annihilation operator for a phonon mode with wave vector \mathbf{k} , with the corresponding free field evolution $H_B^{Ph} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} b_{\mathbf{k}}^\dagger b_{\mathbf{k}}$, and $g_{\mathbf{k}}$ is the electron-phonon coupling strength of the \mathbf{k}^{th} phonon mode. The electron-phonon coupling can be fully characterised in terms of the spectral density $\mathcal{J}_{Ph}(\omega) = \sum_{\mathbf{k}} S_{\mathbf{k}} \delta(\omega - \omega_{\mathbf{k}})$, where $S_{\mathbf{k}} = \omega_{\mathbf{k}}^{-2} |g_{\mathbf{k}}|^2$ is the partial Huang-Rhys factor of mode \mathbf{k} . The total Huang-Rhys parameter can then be recovered by integrating the spectral density through $S_{\text{Tot}} = \int_0^\infty d\omega \mathcal{J}_{Ph}(\omega)$.

To determine the partial Huang-Rhys factors, and therefore the spectral density, we treat the electron-phonon coupling from first principles, using density functional theory (DFT) to calculate the normal modes of the ground- and excited states of the hBN lattice with the considered defect complex. Importantly, the normal modes of the ground- and excited states of the system can be very different and can even be of different types. To properly account for this, we employ the method outlined by Duschinsky [39] to calculate the partial Huang-Rhys factors. For completeness we also checked the generating function approach [40] and found almost identical

results, meaning that the modes in the ground and excited states for the system studied in the present work are in fact very similar. The DFT calculations of the phonons were performed for periodically repeated defect complexes in hBN monolayers using the Vienna Ab Initio Simulation Package (VASP) [41]. In order to avoid interactions between the periodic copies of the defects, we used a 9x9x1 supercell for the calculations, which we relax on a 3x3x1 K-point grid, to a force convergence of 10^{-3} eV \AA^{-1} using a plane wave cut-off of 700 eV. All calculations were performed with the HSE06 exchange correlation functional, which is essential for an accurate description of the electron-phonon coupling [42]. The normal modes and the dynamical matrix was calculated at the Γ -point. Note that we relax the defect in the in-plane constrained C_2V configuration.

The resulting spectral density is shown in Fig. 1(b), where we see multiple sharp peaks present. These peaks correspond to high- Q phonon modes present in hBN; notably from the atomistic simulations, we can assign the dominant contributions to the peak ~ 125 meV from delocalized defect breathing modes, in which the atoms surrounding the defect oscillate along the dipole direction of the defect.

To understand the influence of these modes will have on the cavity-defect system, we can calculate the bath correlation function $C(\tau) = \int_0^\infty d\omega \mathcal{J}_{\text{Ph}}(\omega) (\cos(\omega\tau) \coth(\omega/2k_B T) - i \sin(\omega\tau))$, shown in Fig. 1(c), where T is the temperature, k_B is the Boltzmann constant, and we have assumed the phonon bath to be initially in thermal equilibrium. This correlation function encodes the timescales over which memory effects last between the system and environment [43], for a full derivation we refer the reader to the SI. The sharp peaks in the *ab initio* spectral density lead to long lived oscillation in the correlation function, which can be attributed to phonon modes with large Q -factors. To highlight the importance of this structure we can replace the spectral density with a broadened function with the same total Huang-Rhys parameter, as shown by the dashed curve in Fig. 1(b). The corresponding correlation function, shown in Fig. 1(c), has only small oscillations which are rapidly damped.

To study the time evolution of the composite cavity-emitter system, we employ the time-evolving matrix product operator (TEMPO) algorithm initially developed by Strathearn *et al* [28]. TEMPO is a powerful method for the study of open quantum systems in strong coupling regimes [29, 30], and has been applied to study quantum thermodynamics in the strong coupling regime [44], as well as non-additive phenomena in quantum optics [31], and optimal control [45]. The starting point for the TEMPO formalism, alongside other numerically exact path integral methods [46–48], is the Trotter decomposition [49], where for a sufficiently small time increment δt , the propagator for the open sys-

tem can be factorised such that $\mathcal{U}_{\delta t} = e^{(\mathcal{L}_S + \mathcal{L}_B)\delta t} \approx \mathcal{V}_{\delta t}^{1/2} \mathcal{W}_{\delta t} \mathcal{V}_{\delta t}^{1/2} + \mathcal{O}(\delta t^3)$. The superoperator $\mathcal{V}_{\delta t}$ is as defined above, and captures the evolution of the system and dissipation through the external electromagnetic fields. The interaction between the system and phonon environment is captured through $\mathcal{W}_{\delta t} = \exp(\delta t \mathcal{L}_B)$, and is given by $\mathcal{L}_B \rho = -i [H_1^{\text{Ph}} + H_B^{\text{Ph}}, \rho]$.

This partitioning allows one to construct a discrete-time influence functional of Feynman-Vernon type [46, 47], which captures the influence of the environment to all orders in the interaction strength. The reduced state after k -time-steps can be expressed as:

$$\rho^{\alpha_k} = \sum_{\vec{\alpha}, \vec{\beta}} \rho^{\alpha_0} \prod_{j=1}^k \mathcal{V}_{\beta_j}^{\alpha_j} \mathcal{V}_{\alpha_{j-1}}^{\beta_j} \mathcal{F}_{\beta_k \dots \beta_1}, \quad (2)$$

where we have introduced the compound indices $\alpha_k = (s_k, r_k)$ and $\beta_k = (t_k, u_k)$, yielding the density matrix elements $\rho^{\alpha_k} = \langle r_k | \rho | s_k \rangle$, and the system superoperator is given by $\mathcal{V}_{\beta_k}^{\alpha_k} = \langle r_k | \mathcal{V}_{\delta t} [|t_k\rangle \langle u_k|] |s_k\rangle$. The influence of the phonon environment is captured by the influence tensor $\mathcal{F}_{\beta_k \dots \beta_1} = \text{tr}_E(\mathcal{W}^{\beta_k} \dots \mathcal{W}^{\beta_1} [\tau_B])$, where $\mathcal{W}^{\beta_k} = \langle t_k | \mathcal{W}_{\delta t} [|t_k\rangle \langle u_k|] |u_k\rangle$. By taking the initial state of the environment to be in thermal equilibrium $\tau_B = \exp(-\sum_k \nu_k b_k^\dagger b_k / k_B T) / \text{tr}[\exp(-\sum_k \nu_k b_k^\dagger b_k / k_B T)]$, the trace over the environmental degrees of freedom can be done analytically [46, 47].

Crucially, the influence tensor scales exponentially in the number of time-steps taken [46, 47]; while applying a finite time memory approximation can reduce the computational cost of propagating the reduced state of the system out to long times [46, 47, 50], it limits one to scenarios where key dynamics occur on short timescales. A key insight of Strathearn *et al* [28] was that the influence tensor may be represented in matrix product operator (MPO) form. This allows one to encode the exponentially growing tensor as a tensor network, and apply tensor compression [51] to reduce the rank of the elements in the network, thereby circumventing exponential scaling. The reduced state of the electronic system is then calculated by contracting the network down after each time-step. The convergence of the TEMPO algorithm is sensitive to taking a sufficiently small timestep δt , and the degree to which the tensors are compressed. Further details of TEMPO and its convergence properties are discussed in the SI.

We now consider the dynamics of the cavity-defect system initialised in the state $\rho(0) = |e, 0\rangle \langle e, 0|$. Fig. 2 shows the time evolution of the excited state population and the cavity mode occupation for different cavity coupling parameters. We find that across all parameter regimes TEMPO predicts complex oscillations in the cavity occupation. We attribute these oscillations to high- Q phonon modes in the environment, which lead to long-lived oscillations in the bath correlation function shown

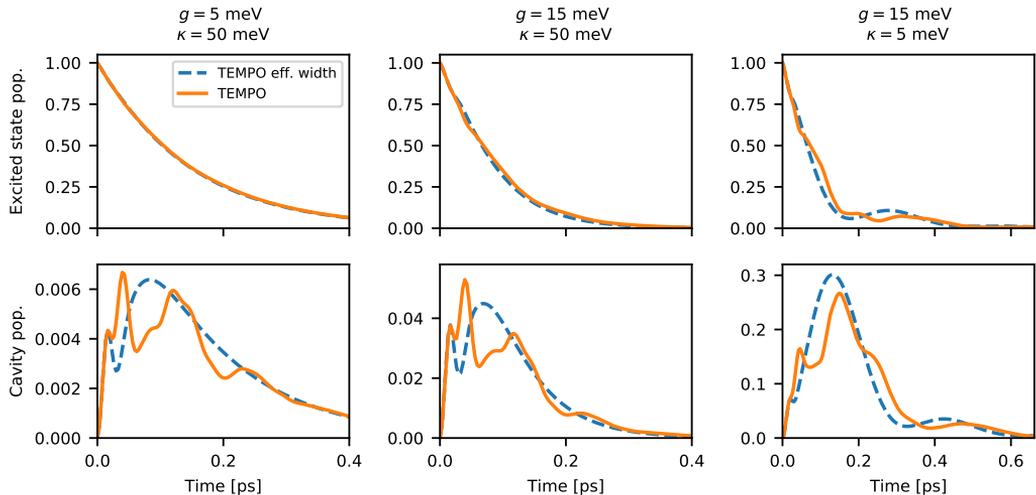


FIG. 2. Time dependent emitter population (top) and cavity occupation (bottom) for different coupling strengths and cavity widths. In contrast to the effective width spectral density (dashed), the *ab initio* spectral density (solid orange) predicts significant oscillations in the cavity occupation over all parameter regimes. Parameters used are $T = 4$ K, $\Gamma = 4$ meV, and $\hbar\Omega_c = \hbar\omega_e - \hbar\lambda$ where $\lambda = \int_0^\infty d\omega\omega^{-2}\mathcal{J}_{\text{Ph}}(\omega)$ is the reorganisation energy. The step size used to obtain convergence was $dt = 3.2$ fs, with SVD cut-off $\epsilon_C = 10^{-6}$.

in Fig. 1c, indeed these oscillations follow closely those observed in the bath correlation function. We confirm this by repeating the TEMPO calculations with the effective width spectral density shown in Fig. 1b. Since this spectral density has the same total Huang-Rhys parameter, naïvely we might expect similar dynamics to emerge. However, the resulting dynamics in Fig. 2 show no oscillations at weak light-matter coupling, suggesting it is indeed the structure in the spectral density that has the dominant contribution to non-Markovian behaviour. Interestingly, phonon induced oscillations are present in the cavity dynamics across all parameter regimes studied, but only emerge in the emitter dynamics in the strong light-matter coupling regime. This is a consequence of phonon processes inducing fluctuations to the defects' excited state energy; the cavity is sensitive to these fluctuations through the dipole coupling with the emitter, leading to the observed oscillations in the occupation. However, these effects do become increasingly visible in the emitter dynamics with increasing coupling strength, with the onset of vacuum Rabi oscillations.

We now turn our attention to the linear spectrum of the cavity-defect system. The linear response of an electronic system coupled to a quantized mode can be formulated in two ways, by an external field coupling to the electronic degrees of freedom, or by an external current pumping the cavity mode [52, 53]. In both cases the driving can be included as a perturbation at the

Hamiltonian level, such that $H'(t) = H + E(t) \cdot \mu$. Here the system transition operator can take on two values $\mu = \{\sigma^\dagger + \sigma, a^\dagger + a\}$. The first is the dipole operator of the defect, describing the scattering of light directly off the two level transition, and forms the basis of standard linear response theory [54]. The second is the quadrature operator of the cavity field, which induces a static polarisation of the cavity mode and excites real photons into the field [53]. $E(t)$ then corresponds to the external field or the external current respectively.

Treating the driving as weak, we can extract the linear response spectrum using density matrix perturbation theory [54]. Taking the semi-impulsive limit such that $E(t) \propto e^{i\omega_D t} \delta(t)$, where ω_D is the driving frequency, we obtain the absorption spectrum:

$$A(\omega) = 2\text{Re} \left[\int_0^\infty dt e^{i\omega t} S^{(1)}(t) \right], \quad (3)$$

where we have introduced the first-order response function $S^{(1)}(t) = \text{tr}(\mu(t)[\mu(0), \chi(-\infty)])$. The global equilibrium density operator is given by $\chi(-\infty) = |g, 0\rangle\langle g, 0| \otimes \tau_B \otimes |\{0}\rangle\langle\{0}\rangle$, where $|\{0}\rangle$ denotes the vacuum state of the electromagnetic field. The above expressions allow us to extend TEMPO to calculate the first-order response function by propagating the initial system state $\rho(0) = \mu(0) |g, 0\rangle\langle g, 0|$.

Figs. 3(a-c) show the absorption spectra of the cavity mode when driven by an external current for various

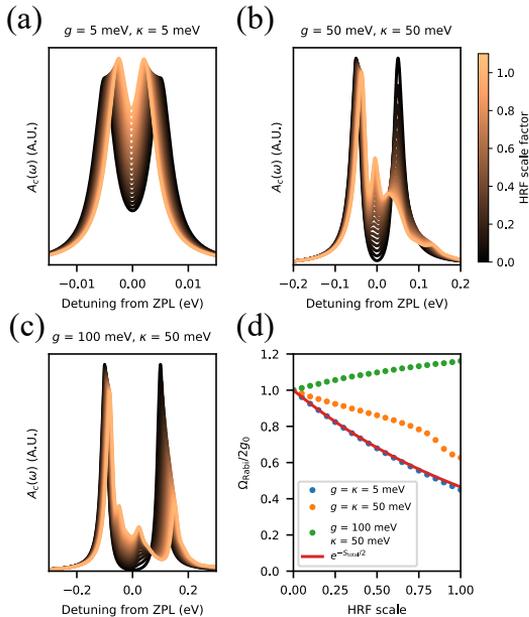


FIG. 3. (a-c) the linear absorption spectra when probing the cavity mode, for various values of the scaling parameter α_{HRF} . Here structure in the absorption spectra becomes increasingly important for large coupling, and broad cavity widths. (d) shows the change in peak position as a function of α_{HRF} (points) for the cavity parameters studied in (a-c). All other parameters are the same as Fig. 2

cavity parameters. To understand the role of phonons in these spectra, we artificially include a scaling parameter to spectral density $\mathcal{J}_{\text{Ph}}(\omega) \rightarrow \alpha_{\text{HRF}}\mathcal{J}_{\text{Ph}}(\omega)$, where $\alpha_{\text{HRF}} \in [0, 1]$, equivalent to scaling the total Huang-Rhys parameter. Fig. 3a shows absorption spectrum at the onset of the strong coupling regime, where $g = \kappa$. In the absence of phonons ($\alpha_{\text{HRF}} = 0$) we see a clear Rabi splitting, with the separation of the peaks determined by the light-matter coupling g . As α_{HRF} increases, the separation of these peaks is reduced. We can understand this reduction by appealing to the polaron formalism commonly used to study the behaviour semiconductor quantum dots [20, 55]: here, the light-matter coupling strength is reduced by the Frank-Condon factor $\mathcal{F} = \exp(-\alpha_{\text{HRF}}S_{\text{Tot}}/2)$, which accounts for the geometric difference of the phonon modes associated to the emitter ground and excited states. Fig. 3d compares the peak separation as a function of the scaling parameter α_{HRF} for the three parameter regimes. For the parameters associated with Fig. 3a, the peak separation follows the Frank-Condon factor \mathcal{F} (solid curve).

In regimes of stronger light-matter coupling show in

Fig. 3b and c, a more complex picture emerges: significant structure becomes apparent in the spectra as α_{HRF} increases. Of particular interest is the departure from the well understood polaronic physics seen in Fig. 3a, which is highlighted in Fig. 3d. Here we see that at stronger coupling regimes ($g = 50$ meV), the renormalisation no longer follows the Frank-Condon factor, and in regime of strong light-matter coupling ($g = 100$ meV), we in-fact see the splitting increase with α_{HRF} . We attribute this behaviour to a hybridisation of the light-matter polariton and high- Q phonon modes. The resultant state is a tri-partite quasi-particle with characteristics of light, matter, and vibrations. This can be seen most clearly in the upper-polariton of Fig. 3c, where an additional splitting emerges when $\alpha_{\text{HRF}} \rightarrow 1$. This interpretation is supported by the C_2C_N calculations shown in the SI; this defect complex has a spectral density with little low energy structure, such that at $g = 50$ meV no hybridisation occurs, and the renormalisation of the Rabi frequency follows closely the Frank-Condon factor \mathcal{F} . It is only at higher light-matter coupling strengths ($g = 100$ meV) for C_2C_N , when the polariton splitting approaches resonance with a high- Q phonon mode that we observe a departure from Frank Condon physics, heralding hybridised polariton-polaron states.

Conclusion— In this letter we have combined atomistic simulations of a defect complex in hBN with TEMPO, a numerically exact and fully quantum mechanical simulation method. Our hybrid approach allows us to study realistic emitters beyond phenomenological and approximate treatments [22], providing a complete description of electron-phonon interactions in condensed matter single photon emitters in optical or plasmonic cavities, with direct relevance to on-going experiments. Furthermore, by considering the cavity quantum electro-dynamics of a defect across the weak and strong light-matter coupling regimes, we have shown that strong coupling to high- Q vibrational modes play a significant role in determining the dynamics of the cavity-defect system, even in the weak light-matter coupling regime. At strong light-matter coupling, the absorption spectrum show clear signatures of hybridisation between the light-matter polaritons and phonon modes inherent to hBN.

The method we present here is general, and not restricted to the material system or specific defect complexes studied here, with potential application to organic polaritons [56]. It is worth noting that we have restricted ourselves to low temperatures in the above discussion; at elevated temperatures phonon processes beyond linear electron-phonon coupling become important through mechanisms such as the Jahn-Teller effect [57].

Acknowledgments— We thank Mortiz Fischer and Emil V. Denning for insightful conversations. The work is supported by the BIOMAG project funded by the Novo Nordisk Foundation (grant no. NNF21OC0066526) We also acknowledge funding from the European Research

Council (ERC) under the European Union's Horizon 2020 research and innovation program Grant No. 773122 (LIMA). K. S. T. is a Villum Investigator supported by Villum Foundation (grant no. 37789). This work was supported by the Danish National Research Foundation through NanoPhoton - Center for Nanophotonics, Grant No. DNRF147 and Center for Nanostructured Graphene, Grant No. DNRF103. NS acknowledges support from the Villum Foundation through grant no. 00028233 and from the Independent Research Fund Denmark - Natural Sciences (Project no. 0135-00403B). J.I.-S. acknowledges support from the Royal Commission for the Exhibition of 1851.

-
- [1] C. Chakraborty, L. Kinnischtzke, K. M. Goodfellow, R. Beams, and A. N. Vamivakas, Voltage-controlled quantum light from an atomically thin semiconductor, *Nature Nanotechnology* **10**, 507 (2015).
- [2] Y.-M. He, G. Clark, J. R. Schaibley, Y. He, M.-C. Chen, Y.-J. Wei, X. Ding, Q. Zhang, W. Yao, X. Xu, C.-Y. Lu, and J.-W. Pan, Single quantum emitters in monolayer semiconductors, *Nature Nanotechnology* **10**, 497 (2015).
- [3] M. Koperski, K. Nogajewski, A. Arora, V. Cherkez, P. Mallet, J.-Y. Vuillen, J. Marcus, P. Kossacki, and M. Potemski, Single photon emitters in exfoliated wse2 structures, *Nature Nanotechnology* **10**, 503 (2015).
- [4] A. Srivastava, M. Sidler, A. V. Allain, D. S. Lembke, A. Kis, and A. Imamoglu, Optically active quantum dots in monolayer WSe₂, *Nature Nanotechnology* **10**, 491 (2015).
- [5] J. Klein, M. Lorke, M. Florian, F. Sigger, L. Sigl, S. Rey, J. Wierzbowski, J. Cerne, K. Müller, E. Mitterreiter, *et al.*, Site-selectively generated photon emitters in monolayer mos2 via local helium ion irradiation, *Nature communications* **10**, 1 (2019).
- [6] S. Michaelis de Vasconcellos, D. Wigger, U. Wurstbauer, A. W. Holleitner, R. Bratschitsch, and T. Kuhn, Single-photon emitters in layered van der waals materials, *physica status solidi (b)*, 2100566 (2022).
- [7] M. Kianinia, Z.-Q. Xu, M. Toth, and I. Aharonovich, Quantum emitters in 2d materials: Emitter engineering, photophysics, and integration in photonic nanostructures, *Applied Physics Reviews* **9**, 011306 (2022).
- [8] T. T. Tran, K. Bray, M. J. Ford, M. Toth, and I. Aharonovich, Quantum emission from hexagonal boron nitride monolayers, *Nature Nanotechnology* **11**, 37 (2016).
- [9] I. Aharonovich and M. Toth, Quantum emitters in two dimensions, *Science* **358**, 170 (2017).
- [10] M. Fischer, J. Caridad, A. Sajid, S. Ghaderzadeh, M. Ghorbani-Asl, L. Gammelgaard, P. Bøggild, K. S. Thygesen, A. Krasheninnikov, S. Xiao, *et al.*, Controlled generation of luminescent centers in hexagonal boron nitride by irradiation engineering, *Science Advances* **7**, eabe7138 (2021).
- [11] N. Mendelson, D. Chugh, J. R. Reimers, T. S. Cheng, A. Gottscholl, H. Long, C. J. Mellor, A. Zettl, V. Dyakonov, P. H. Beton, *et al.*, Identifying carbon as the source of visible single-photon emission from hexagonal boron nitride, *Nature materials* **20**, 321 (2021).
- [12] A. Gottscholl, M. Kianinia, V. Soltamov, S. Orlinskii, G. Mamin, C. Bradac, C. Kasper, K. Krambrock, A. Sperlich, M. Toth, *et al.*, Initialization and read-out of intrinsic spin defects in a van der waals crystal at room temperature, *Nature materials* **19**, 540 (2020).
- [13] M. Hoese, P. Reddy, A. Dietrich, M. K. Koch, K. G. Fehler, M. W. Doherty, and A. Kubanek, Mechanical decoupling of quantum emitters in hexagonal boron nitride from low-energy phonon modes, *Science advances* **6**, eaba6038 (2020).
- [14] G. Cassabois, P. Valvin, and B. Gil, Hexagonal boron nitride is an indirect bandgap semiconductor, *Nature photonics* **10**, 262 (2016).
- [15] T. T. Tran, C. Bradac, A. S. Solntsev, M. Toth, and I. Aharonovich, Suppression of spectral diffusion by anti-stokes excitation of quantum emitters in hexagonal boron nitride, *Applied Physics Letters* **115**, 071102 (2019).
- [16] K. Li, T. Smart, and Y. Ping, C₂C_N as a 2 ev single-photon emitter candidate in hexagonal boron nitride, *arXiv preprint arXiv:2110.01787* (2021).
- [17] T. Vogl, R. Lecamwasam, B. C. Buchler, Y. Lu, and P. K. Lam, Compact cavity-enhanced single-photon generation with hexagonal boron nitride, *ACS Photonics* **6**, 1955 (2019).
- [18] S. Häußler, G. Bayer, R. Waltrich, N. Mendelson, C. Li, D. Hunger, I. Aharonovich, and A. Kubanek, Tunable fiber-cavity enhanced photon emission from defect centers in hbn, *Advanced Optical Materials* **9**, 2002218 (2021).
- [19] E. V. Denning, J. Iles-Smith, N. Gregersen, and J. Mork, Phonon effects in quantum dot single-photon sources, *Opt. Mater. Express* **10**, 222 (2020).
- [20] A. Nazir and D. P. S. McCutcheon, Modelling exciton-phonon interactions in optically driven quantum dots, *Journal of Physics: Condensed Matter* **28**, 103002 (2016).
- [21] T. Vuong, G. Cassabois, P. Valvin, A. Ouerghi, Y. Chasagneux, C. Voisin, and B. Gil, Phonon-photon mapping in a color center in hexagonal boron nitride, *Physical review letters* **117**, 097402 (2016).
- [22] M. A. Feldman, A. Puzos, L. Lindsay, E. Tucker, D. P. Briggs, P. G. Evans, R. F. Haglund, and B. J. Lawrie, Phonon-induced multicolor correlations in hbn single-photon emitters, *Phys. Rev. B* **99**, 020101 (2019).
- [23] P. Khatri, I. J. Luxmoore, and A. J. Ramsay, Phonon sidebands of color centers in hexagonal boron nitride, *Phys. Rev. B* **100**, 125305 (2019).
- [24] G. Grosso, H. Moon, C. J. Ciccarino, J. Flick, N. Mendelson, L. Mennel, M. Toth, I. Aharonovich, P. Narang, and D. R. Englund, Low-temperature electron-phonon interaction of quantum emitters in hexagonal boron nitride, *ACS Photonics* **7**, 1410 (2020).
- [25] A. Sajid and K. S. Thygesen, Vncb defect as source of single photon emission from hexagonal boron nitride, *2D Materials* **7**, 031007 (2020).
- [26] C. Jara, T. Rauch, S. Botti, M. A. Marques, A. Norambuena, R. Coto, J. Castellanos-Águila, J. R. Maze, and F. Muñoz, First-principles identification of single photon emitters based on carbon clusters in hexagonal boron nitride, *The Journal of Physical Chemistry A* **125**, 1325 (2021).
- [27] K. Li, T. J. Smart, and Y. Ping, Carbon trimer as a 2 ev single-photon emitter candidate in hexagonal boron ni-

- tride: A first-principles study, *Physical Review Materials* **6**, L042201 (2022).
- [28] A. Strathearn, P. Kirton, D. Kilda, J. Keeling, and B. W. Lovett, Efficient non-markovian quantum dynamics using time-evolving matrix product operators, *Nature communications* **9**, 1 (2018).
- [29] M. R. Jørgensen and F. A. Pollock, Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-markovian path integrals, *Phys. Rev. Lett.* **123**, 240602 (2019).
- [30] D. Gribben, A. Strathearn, J. Iles-Smith, D. Kilda, A. Nazir, B. W. Lovett, and P. Kirton, Exact quantum dynamics in structured environments, *Phys. Rev. Research* **2**, 013265 (2020).
- [31] D. Gribben, D. M. Rouse, J. Iles-Smith, A. Strathearn, H. Maguire, P. Kirton, A. Nazir, E. M. Gauger, and B. W. Lovett, Exact dynamics of nonadditive environments in non-markovian open quantum systems, *PRX Quantum* **3**, 010321 (2022).
- [32] M. K. Svendsen, Y. Kurman, *et al.*, Combining density functional theory with macroscopic qed for quantum light-matter interactions in 2D materials, *Nature communications* **12**, 1 (2021).
- [33] S. Latini, U. De Giovannini, E. J. Sie, N. Gedik, H. Hübener, and A. Rubio, Phonoritons as hybridized exciton-photon-phonon excitations in a monolayer h – BN optical cavity, *Phys. Rev. Lett.* **126**, 227401 (2021).
- [34] A. Sajid, J. R. Reimers, and M. J. Ford, Defect states in hexagonal boron nitride: Assignments of observed properties and prediction of properties relevant to quantum computation, *Phys. Rev. B* **97**, 064101 (2018).
- [35] This analysis is repeated in the supplementary information for the $C_{2v}N$ defect.
- [36] H. J. Carmichael, *Statistical methods in quantum optics 1: master equations and Fokker-Planck equations*, Vol. 1 (Springer Science & Business Media, 1999).
- [37] H. Maguire, J. Iles-Smith, and A. Nazir, Environmental nonadditivity and franck-condon physics in nonequilibrium quantum systems, *Physical review letters* **123**, 093601 (2019).
- [38] G. D. Mahan, *Many-particle physics* (Springer Science & Business Media, 2013).
- [39] R. Borrelli, A. Capobianco, and A. Peluso, Franck-condon factors—computational approaches and recent developments, *Canadian Journal of Chemistry* **91**, 495 (2013).
- [40] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van de Walle, First-principles theory of the luminescence lineshape for the triplet transition in diamond nv centres, *New Journal of Physics* **16**, 073026 (2014).
- [41] G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, *Physical review B* **47**, 558 (1993).
- [42] J. Heyd and G. E. Scuseria, Efficient hybrid density functional calculations in solids: Assessment of the heyd-scuseria-ernzerhof screened coulomb hybrid functional, *The Journal of chemical physics* **121**, 1187 (2004).
- [43] H.-P. Breuer, F. Petruccione, *et al.*, *The theory of open quantum systems* (Oxford University Press on Demand, 2002).
- [44] M. Popovic, M. T. Mitchison, A. Strathearn, B. W. Lovett, J. Goold, and P. R. Eastham, Quantum heat statistics with time-evolving matrix product operators, *PRX Quantum* **2**, 020338 (2021).
- [45] G. E. Fux, E. P. Butler, P. R. Eastham, B. W. Lovett, and J. Keeling, Efficient exploration of hamiltonian parameter space for optimal control of non-markovian open quantum systems, *Phys. Rev. Lett.* **126**, 200401 (2021).
- [46] N. Makri and D. E. Makarov, Tensor propagator for iterative quantum time evolution of reduced density matrices. i. theory, *The Journal of chemical physics* **102**, 4600 (1995).
- [47] N. Makri and D. E. Makarov, Tensor propagator for iterative quantum time evolution of reduced density matrices. ii. numerical methodology, *The Journal of chemical physics* **102**, 4611 (1995).
- [48] M. Cygorek, M. Cosacchi, A. Vagov, V. Axt, B. Lovett, J. Keeling, and E. Gauger, Simulation of open quantum systems by automated compression of arbitrary environments, *Nature Physics* (2022).
- [49] H. F. Trotter, On the product of semi-groups of operators, *Proceedings of the American Mathematical Society* **10**, 545 (1959).
- [50] A. Strathearn, B. W. Lovett, and P. Kirton, Efficient real-time path integrals for non-markovian spin-boson models, *New Journal of Physics* **19**, 093009 (2017).
- [51] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Annals of Physics* **349**, 117 (2014).
- [52] M. Ruggenthaler, N. Tancogne-Dejean, J. Flick, H. Appel, and A. Rubio, From a quantum-electrodynamical light-matter description to novel spectroscopies, *Nature Reviews Chemistry* **2**, 1 (2018).
- [53] J. Flick, D. M. Welakuh, M. Ruggenthaler, H. Appel, and A. Rubio, Light-matter response in nonrelativistic quantum electrodynamics, *ACS photonics* **6**, 2757 (2019).
- [54] S. Mukamel, *Principles of nonlinear optical spectroscopy*, 6 (Oxford University Press on Demand, 1999).
- [55] J. Iles-Smith, D. P. McCutcheon, A. Nazir, and J. Mørk, Phonon scattering inhibits simultaneous near-unity efficiency and indistinguishability in semiconductor single-photon sources, *Nature Photonics* **11**, 521 (2017).
- [56] J. Del Pino, F. A. Schröder, A. W. Chin, J. Feist, and F. J. Garcia-Vidal, Tensor network simulation of non-markovian dynamics in organic polaritons, *Physical review letters* **121**, 227401 (2018).
- [57] S. White, C. Stewart, A. S. Solntsev, C. Li, M. Toth, M. Kianinia, and I. Aharonovich, Phonon dephasing and spectral diffusion of quantum emitters in hexagonal boron nitride, *Optica* **8**, 1153 (2021).

11.7 Paper [VII]: Recent Progress of the Computational Materials Database (C2DB)

M. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N. R. Knøsgaard, M. Kruse, A. H. Larsen, S. Manti, T. G. Pedersen, U. Petralanda, T. Skovhus, M. K. Svendsen, J. J. Mortensen, T. Olsen and K. S. Thygesen, “Recent Progress of the Computational 2D Materials Database (C2DB),” *2D Mater.* 2021, **8**, 044002.

PAPER • OPEN ACCESS

Recent progress of the Computational 2D Materials Database (C2DB)

To cite this article: Morten Niklas Gjerding *et al* 2021 *2D Mater.* **8** 044002

View the [article online](#) for updates and enhancements.

You may also like

- [High throughput computational screening for 2D ferromagnetic materials: the critical role of anisotropy and local correlations](#)

Daniele Torelli, Kristian S Thygesen and Thomas Olsen

- [Exploring key aspects of art documentation and examination in curation technology-using Cheng-Shiu university conservation center's project as a case study](#)

L C Lin and I C Li

- [Machine learning enabled discovery of application dependent design principles for two-dimensional materials](#)

Victor Venturi, Holden L Parks, Zeeshan Ahmad *et al.*

2D Materials



PAPER

OPEN ACCESS

RECEIVED
18 January 2021

REVISED
31 May 2021

ACCEPTED FOR PUBLICATION
30 June 2021

PUBLISHED
15 July 2021

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Recent progress of the Computational 2D Materials Database (C2DB)

Morten Niklas Gjerding¹, Alireza Taghizadeh^{1,2}, Asbjørn Rasmussen¹, Sajid Ali¹, Fabian Bertoldo¹, Thorsten Deilmann³, Nikolaj Rørbæk Knøsgaard¹, Mads Kruse¹, Ask Hjorth Larsen¹, Simone Manti¹, Thomas Garm Pedersen², Urko Petralanda¹, Thorbjørn Skovhus¹, Mark Kamper Svendsen¹, Jens Jørgen Mortensen¹, Thomas Olsen¹ and Kristian Sommer Thygesen^{1,*}

¹ Computational Atomic-scale Materials Design (CAMD), Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

² Department of Materials and Production, Aalborg University, 9220 Aalborg Øst, Denmark

³ Institut für Festkörperteorie, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany

* Author to whom any correspondence should be addressed.

E-mail: thygesen@fysik.dtu.dk

Keywords: 2D materials, high-throughput, ab-initio, database, density functional theory

Abstract

The Computational 2D Materials Database (C2DB) is a highly curated open database organising a wealth of computed properties for more than 4000 atomically thin two-dimensional (2D) materials. Here we report on new materials and properties that were added to the database since its first release in 2018. The set of new materials comprise several hundred monolayers exfoliated from experimentally known layered bulk materials, (homo)bilayers in various stacking configurations, native point defects in semiconducting monolayers, and chalcogen/halogen Janus monolayers. The new properties include exfoliation energies, Bader charges, spontaneous polarisations, Born charges, infrared polarisabilities, piezoelectric tensors, band topology invariants, exchange couplings, Raman spectra and second harmonic generation spectra. We also describe refinements of the employed material classification schemes, upgrades of the computational methodologies used for property evaluations, as well as significant enhancements of the data documentation and provenance. Finally, we explore the performance of Gaussian process-based regression for efficient prediction of mechanical and electronic materials properties. The combination of open access, detailed documentation, and extremely rich materials property data sets make the C2DB a unique resource that will advance the science of atomically thin materials.

1. Introduction

The discovery of new materials, or new properties of known materials, to meet a specific industrial or scientific requirement, is an exciting intellectual challenge of the utmost importance for our environment and economy. For example, the successful transition to a society based on sustainable energy sources and the realisation of quantum technologies (e.g. quantum computers and quantum communication) depend critically on new materials with novel functionalities. First-principles quantum mechanical calculations, e.g. based on density functional theory (DFT) [1], can predict the properties of materials with high accuracy even before they are made

in the lab. They provide insight into mechanisms at the most fundamental (atomic and electronic) level and can pinpoint and calculate key properties that determine the performance of the material at the macroscopic level. Powered by high-performance computers, atomistic quantum calculations in combination with data science approaches, have the potential to revolutionise the way we discover and develop new materials.

Atomically thin, two-dimensional (2D) crystals represent a fascinating class of materials with exciting perspectives for both fundamental science and technology [2–5]. The family of 2D materials has been growing steadily over the past decade and counts about a hundred materials that have been realised

in single-layer or few-layer form [6–10]. While some of these materials, including graphene, hexagonal boron nitride (hBN), and transition metal dichalcogenides (TMDs), have been extensively studied, the majority have only been scarcely characterised and remain poorly understood. Computational studies indicate that around 1000 already known layered crystals have sufficiently weak interlayer (IL) bonding to allow the individual layers to be mechanically exfoliated [11, 12]. Supposedly, even more 2D materials could be realised beyond this set of already known crystals. Adding to this the possibility of stacking individual 2D layers (of the same or different kinds) into ultrathin van der Waals (vdW) crystals [13], and tuning the properties of such structures by varying the relative twist angle between adjacent layers [14, 15] or intercalating atoms into the vdW gap [16, 17], it is clear that the prospects of tailor made 2D materials are simply immense. To support experimental efforts and navigate the vast 2D materials space, first-principles calculations play a pivotal role. In particular, FAIR⁵ [18] databases populated by high-throughput calculations can provide a convenient overview of known materials and point to new promising materials with desired (predicted) properties. Such databases are also a fundamental requirement for the successful introduction and deployment of artificial intelligence in materials science.

Many of the unique properties exhibited by 2D materials have their origin in quantum confinement and reduced dielectric screening. These effects tend to enhance many-body interactions and lead to profoundly new phenomena such as strongly bound excitons [19–21] with nonhydrogenic Rydberg series [22–24], phonons and plasmons with anomalous dispersion relations [25, 26], large dielectric band structure renormalisations [27, 28], unconventional Mott insulating and superconducting phases [14, 15], and high-temperature exciton condensates [29]. Recently, it has become clear that long range magnetic order can persist [30, 31] and (in-plane) ferroelectricity even be enhanced [32], in the single layer limit. In addition, first-principles studies of 2D crystals have revealed rich and abundant topological phases [33, 34]. The peculiar physics ruling the world of 2D materials entails that many of the conventional theories and concepts developed for bulk crystals break down or require special treatments when applied to 2D materials [26, 35, 36]. This means that computational studies must be performed with extra care, which in turn calls for well-organised and well-documented 2D property data sets that can form the basis for the development, benchmarking, and consolidation of physical theories and numerical implementations.

⁵ FAIR data are data which meet principles of findability, accessibility, interoperability, and reusability.

The Computational 2D Materials Database (C2DB) [6, 37] is a highly curated and fully open database containing elementary physical properties of around 4000 2D monolayer crystals. The data has been generated by automatic high-throughput calculations at the level of DFT and many-body perturbation theory as implemented in the GPAW [38, 39] electronic structure code. The computational workflow is constructed using the atomic simulation recipes (ASR) [40]—a recently developed Python framework for high-throughput materials modelling building on the atomic simulation environment (ASE) [41]—and managed/executed using the MyQueue task scheduler [42].

The C2DB differentiates itself from existing computational databases of bulk [43–45] and low-dimensional [11, 12, 46–50] materials, by the large number of physical properties available, see table 1. The use of beyond-DFT theories for excited state properties (GW band structures and Bethe–Salpeter equation (BSE) absorption for selected materials) and Berry-phase techniques for band topology and polarisation quantities (spontaneous polarisation, Born charges, piezoelectric tensors), are other unique features of the database.

The C2DB can be downloaded in its entirety or browsed and searched online. As a new feature, all data entries presented on the website are accompanied by a clickable help icon that presents a scientific documentation (‘what does this piece of data describe?’) and technical documentation (‘how was this piece of data computed?’). This development enhances the usability of the database and improves the reproducibility and provenance of the data contained in C2DB. As another novelty it is possible to download all property data pertaining to a specific material or a specific type of property, e.g. the band gap, for all materials thus significantly improving data accessibility.

In this paper, we report on the significant C2DB developments that have taken place during the past two years. These developments can be roughly divided into four categories: (1) General updates of the workflow used to select, classify, and stabilise the materials. (2) Computational improvements for properties already described in the 2018 paper. (3) New properties. (4) New materials. The developments, described in four separate sections, cover both original work and review of previously published work. In addition, we have included some outlook discussions of ongoing work. In the last section we illustrate an application of statistical learning to predict properties directly from the atomic structure.

2. Selection, classification, and stability

Figure 1 illustrates the workflow behind the C2DB. In this section we describe the first part of the workflow

Table 1. Properties calculated by the C2DB monolayer workflow. The computational method and the criteria used to decide whether the property should be evaluation for a given material is also shown. A ‘*’ indicates that spin-orbit coupling (SOC) is included. All calculations are performed with the GPAW code using a plane wave basis except for the Raman calculations, which employ a double-zeta polarised basis of numerical atomic orbitals [51].

Property	Method	Criteria	Count
Bader charges	PBE	None	3809
Energy above convex hull	PBE	None	4044
Heat of formation	PBE	None	4044
Orbital projected band structure	PBE	None	2487
Out-of-plane dipole	PBE	None	4044
Phonons (Γ and BZ corners)	PBE	None	3865
Projected density of states	PBE	None	3332
Stiffness tensor	PBE	None	3968
Exchange couplings	PBE	Magnetic	538
Infrared polarisability	PBE	$E_{\text{gap}}^{\text{PBE}} > 0$	784
Second harmonic generation	PBE	$E_{\text{gap}}^{\text{PBE}} > 0$, non-magnetic, non-centrosymmetric	375
Electronic band structure PBE	PBE*	None	3496
Magnetic anisotropies	PBE*	Magnetic	823
Deformation potentials	PBE*	$E_{\text{gap}}^{\text{PBE}} > 0$	830
Effective masses	PBE*	$E_{\text{gap}}^{\text{PBE}} > 0$	1272
Fermi surface	PBE*	$E_{\text{gap}}^{\text{PBE}} = 0$	2505
Plasma frequency	PBE*	$E_{\text{gap}}^{\text{PBE}} = 0$	3144
Work function	PBE*	$E_{\text{gap}}^{\text{PBE}} = 0$	4044
Optical polarisability	RPA@PBE	None	3127
Electronic band structure	HSE06@PBE*	None	3155
Electronic band structure	G ₀ W ₀ @PBE*	$E_{\text{gap}}^{\text{PBE}} > 0, N_{\text{atoms}} < 5$	357
Born charges	PBE, Berry phase	$E_{\text{gap}}^{\text{PBE}} > 0$	639
Raman spectrum	PBE, LCAO basis set	Non-magnetic, dyn. stable	708
Piezoelectric tensor	PBE, Berry phase	$E_{\text{gap}}^{\text{PBE}}$, non-centrosym.	353
Optical absorbance	BSE@G ₀ W ₀ *	$E_{\text{gap}}^{\text{PBE}} > 0, N_{\text{atoms}} < 5$	378
Spontaneous polarisation	PBE, Berry phase	$E_{\text{gap}}^{\text{PBE}} > 0$, nearly centrosym. polar space group	151
Topological invariants	PBE*, Berry phase	$0 < E_{\text{gap}}^{\text{PBE}} < 0.3$ eV	242

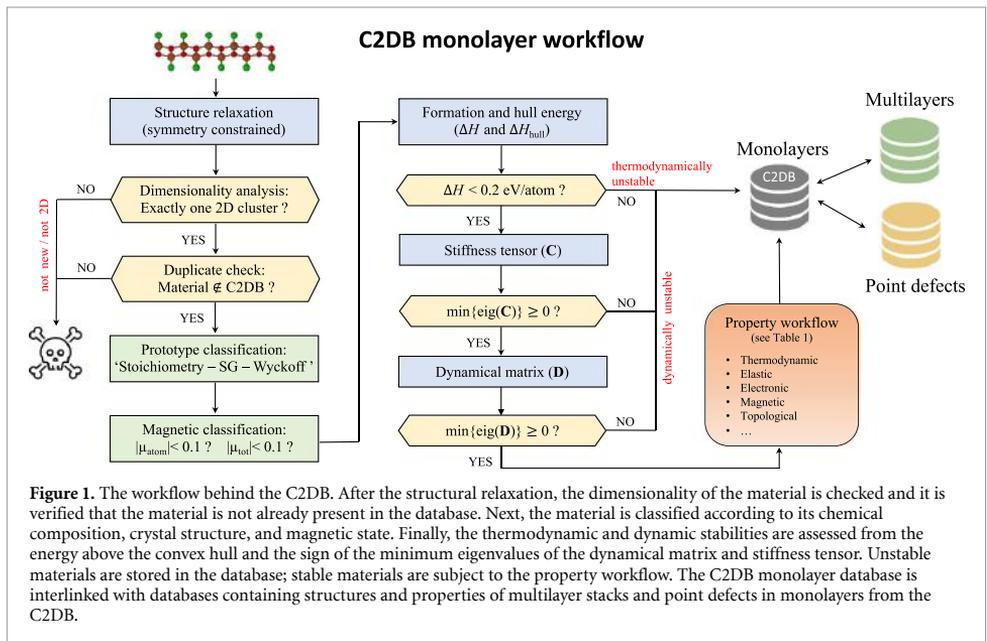


Figure 1. The workflow behind the C2DB. After the structural relaxation, the dimensionality of the material is checked and it is verified that the material is not already present in the database. Next, the material is classified according to its chemical composition, crystal structure, and magnetic state. Finally, the thermodynamic and dynamic stabilities are assessed from the energy above the convex hull and the sign of the minimum eigenvalues of the dynamical matrix and stiffness tensor. Unstable materials are stored in the database; stable materials are subject to the property workflow. The C2DB monolayer database is interlinked with databases containing structures and properties of multilayer stacks and point defects in monolayers from the C2DB.

until the property calculations (red box), focusing on aspects related to selection criteria, classification, and stability assessment, that have been changed or updated since the 2018 paper.

2.1. Structure relaxation

Given a prospective 2D material, the first step is to carry out a structure optimisation. This calculation is performed with spin polarisation and with the symmetries of the original structure enforced. The latter is done to keep the highest level of control over the resulting structure by avoiding ‘uncontrolled’ symmetry breaking distortions. The prize to pay is a higher risk of generating dynamically unstable structures.

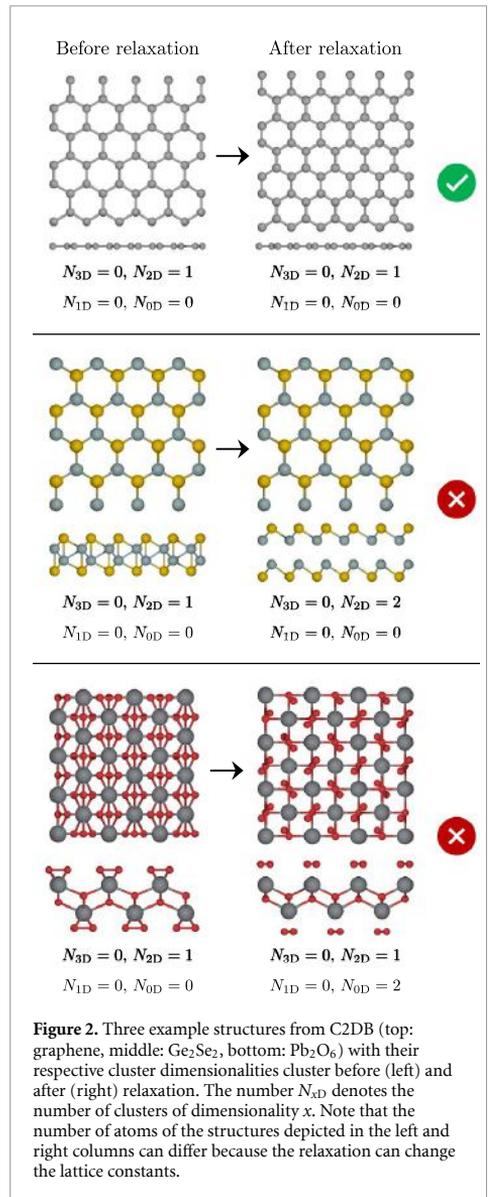
2.2. Selection: dimensionality analysis

A dimensionality analysis [52] is performed to identify and filter out materials that have disintegrated into non-2D structures during relaxation. Covalently bonded clusters are identified through an analysis of the connectivity of the structures where two atoms are considered to belong to the same cluster if their distance is less than some scaling of the sum of their covalent radii, i.e. $d < k(r_i^{\text{cov}} + r_j^{\text{cov}})$, where i and j are atomic indices. A scaling factor of $k = 1.35$ was determined empirically. Only structures that consist of a single 2D cluster after relaxation are further processed. Figure 2 shows three examples (graphene, Ge_2Se_2 , and Pb_2O_6) of structures and their cluster dimensionalities before and after relaxation. All structures initially consist of a single 2D cluster, but upon relaxation Ge_2Se_2 and Pb_2O_6 disintegrate into two 2D clusters as well as one 2D and two 0D clusters, respectively. On the other hand, the relaxation of graphene decreases the in-plane lattice constant but does not affect the dimensionality. According to the criterion defined above only graphene will enter the database.

2.3. Selection: ranking similar structures

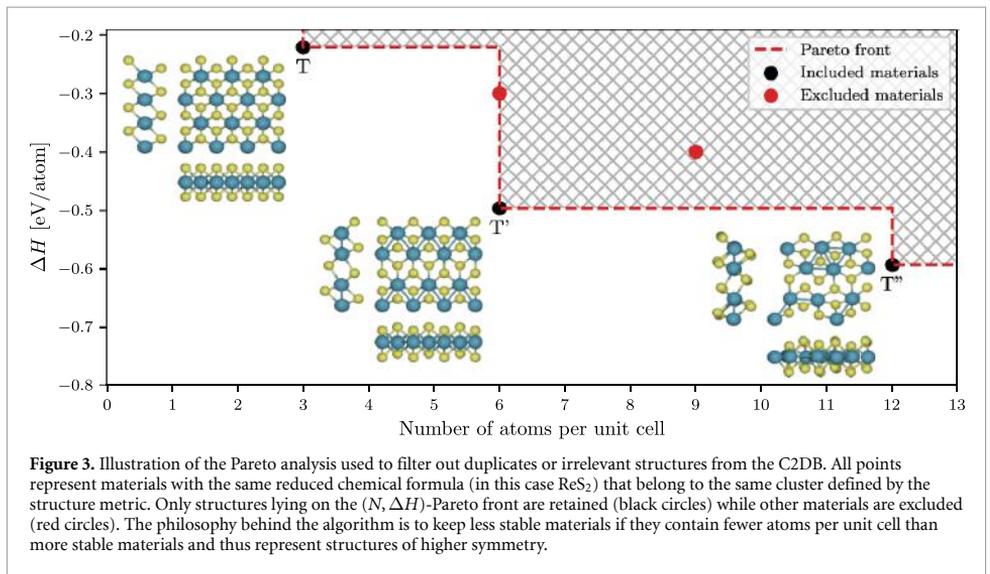
Maintaining a high-throughput database inevitably requires a strategy for comparing similar structures and ranking them according to their relevance. In particular, this is necessary in order to identify different representatives of the same material e.g. resulting from independent relaxations, and thereby avoid duplicate entries and redundant computations. The C2DB strategy to this end involves a combination of structure clustering and Pareto analysis.

First, a single-linkage clustering algorithm is used to group materials with identical reduced chemical formula and ‘similar’ atomic configurations. To quantify configuration similarity a slightly modified version of PyMatGen’s [53] distance metric is employed where the cell volume normalisation is removed to make it applicable to 2D materials surrounded by vacuum. Roughly speaking, the metric measures the maximum distance an atom must be moved (in units of Å) in order to match the two



atomic configurations. Two atomic configurations belong to the same cluster if their distance is below an empirically determined threshold of 0.3 Å.

At this point, the simplest strategy would be to remove all but the most stable compound within a cluster. However, this procedure would remove many high symmetry crystals for which a more stable distorted version exists. For example, the well known T-phase of MoS_2 would be removed in favour of the more stable T'-phase. This is undesired as high-symmetry structures, even if dynamically unstable at $T = 0$, may provide useful information and might in fact become stabilised at higher temperatures [54]. Therefore, the general strategy adopted for the C2DB,



is to keep a material that is less stable than another material of the same cluster if it has fewer atoms in its primitive unit cell (and thus typically higher symmetry). Precisely, materials within a given cluster are kept only if they represent a defining point of the $(N, \Delta H)$ -Pareto front, where N is the number of atoms in the unit cell and ΔH is the heat of formation. A graphical illustration of the Pareto analysis is shown in figure 3 for the case of ReS_2 .

2.4. Classification: crystal structure

The original C2DB employed a *crystal prototype* classification scheme where specific materials were promoted to prototypes and used to label groups of materials with the same or very similar crystal structure. This approach was found to be difficult to maintain (as well as being non-transparent). Instead, materials are now classified according to their *crystal type* defined by the reduced stoichiometry, space group number, and the alphabetically sorted labels of the occupied Wyckoff positions. As an example, MoS_2 in the H-phase has the crystal type: AB2-187-bi.

2.5. Classification: magnetic state

In the new version of the C2DB, materials are classified according to their magnetic state as either *non-magnetic* or *magnetic*. A material is considered magnetic if any atom has a local magnetic moment greater than $0.1 \mu_B$.

In the original C2DB, the *magnetic* category was further subdivided into ferromagnetic (FM) and antiferromagnetic (AFM). But since the simplest antiferromagnetically ordered state typically does not represent the true ground state, all material entries with an AFM state have been removed from the C2DB and replaced by the material in its FM state. Although the latter is less stable, it represents a

more well defined state of the material. Crucially, the nearest neighbour exchange couplings for all magnetic materials have been included in the C2DB (see section 5.8). This enables a more detailed and realistic description of the magnetic order via the Heisenberg model. In particular, the FM state of a material is not expected to represent the true magnetic ground if the exchange coupling $J < 0$.

2.6. Stability: thermodynamic

The heat of formation, ΔH , of a compound is defined as its energy per atom relative to its constituent elements in their standard states [55]. The thermodynamic stability of a compound is evaluated in terms of its energy above the *convex hull*, ΔH_{hull} , which gives the energy of the material relative to other competing phases of the same chemical composition, including mixed phases [6], see figure 4 for an example. Clearly, ΔH_{hull} depends on the pool of reference phases, which in turn defines the convex hull. The original C2DB employed a pool of reference phases comprised by 2807 elemental and binary bulk crystals from the convex hull of the Open Quantum Materials Database (OQMD) [55]. In the new version, this set has been extended by approximately 6783 ternary bulk compounds from the convex hull of OQMD, making a total of 9590 stable bulk reference compounds.

As a simple indicator for the thermodynamic stability of a material, the C2DB employs three labels (low, medium, high) as defined in table 2. These indicators are unchanged from the original version of the C2DB. In particular, the criterion $\Delta H_{\text{hull}} < 0.2 \text{ eV atom}^{-1}$, defining the most stable category, was established based on an extensive analysis of 55 experimentally realised monolayer crystals [6].

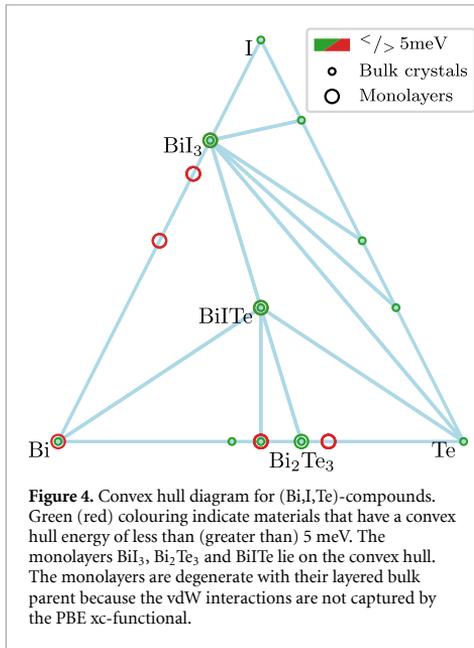


Figure 4. Convex hull diagram for (Bi,I,Te)-compounds. Green (red) colouring indicate materials that have a convex hull energy of less than (greater than) 5 meV. The monolayers BiI_3 , Bi_2Te_3 and BiTe lie on the convex hull. The monolayers are degenerate with their layered bulk parent because the vdW interactions are not captured by the PBE xc-functional.

Table 2. Thermodynamic stability indicator assigned to all materials in the C2DB. ΔH and ΔH_{hull} denote the heat of formation and energy above the convex hull, respectively.

Thermodynamic stability indicator	Criterion (eV atom^{-1})
Low	$\Delta H > 0.2$
Medium	$\Delta H < 0.2$ and $\Delta H_{\text{hull}} > 0.2$
High	$\Delta H < 0.2$ and $\Delta H_{\text{hull}} < 0.2$

It should be emphasised that the energies of both monolayers and bulk reference crystals are calculated with the Perdew-Burke-Ernzerhof (PBE) xc-functional [56]. This implies that some inaccuracies must be expected, in particular for materials with strongly localised *d*-electrons, e.g. certain transition metal oxides, and materials for which dispersive interactions are important, e.g. layered van der Waals crystals. The latter implies that the energy of a monolayer and its layered bulk parent (if such exists in the pool of references) will have the same energy. For further details and discussions see reference [6].

2.7. Stability: dynamical

Dynamically stable materials are situated at a local minimum of the potential energy surface and are thus stable to small structural perturbations. Structures resulting from DFT relaxations can end up in saddle point configurations because of imposed symmetry constraints or an insufficient number of atoms in the unit cell.

In C2DB, the dynamical stability is assessed from the signs of the minimum eigenvalues of (1) the stiffness tensor (see section 3.1) and (2) the Γ -point

Hessian matrix for a supercell containing 2×2 repetitions of the unit cell (the structure is not relaxed in the 2×2 supercell). If one of these minimal eigenvalues is negative the material is classified as dynamically unstable. This indicates that the energy can be reduced by displacing an atom and/or deforming the unit cell, respectively. The use of two categories for dynamical stability, i.e. stable/unstable, differs from the original version of the C2DB where an intermediate category was used for materials with negative but numerically small minimal eigenvalue of either the Hessian or stiffness tensors.

3. Improved property methodology

The new version of the C2DB has been generated using a significantly extended and improved workflow for property evaluations. This section focuses on improvements relating to properties that were already present in the original version of the C2DB while new properties are discussed in the next section.

3.1. Stiffness tensor

The stiffness tensor, C , is a rank-4 tensor that relates the stress of a material to the applied strain. In Mandel notation (a variant of Voigt notation) C is expressed as an $N \times N$ matrix relating the N independent components of the stress and strain tensors. For a 2D material $N = 3$ and the tensor takes the form:

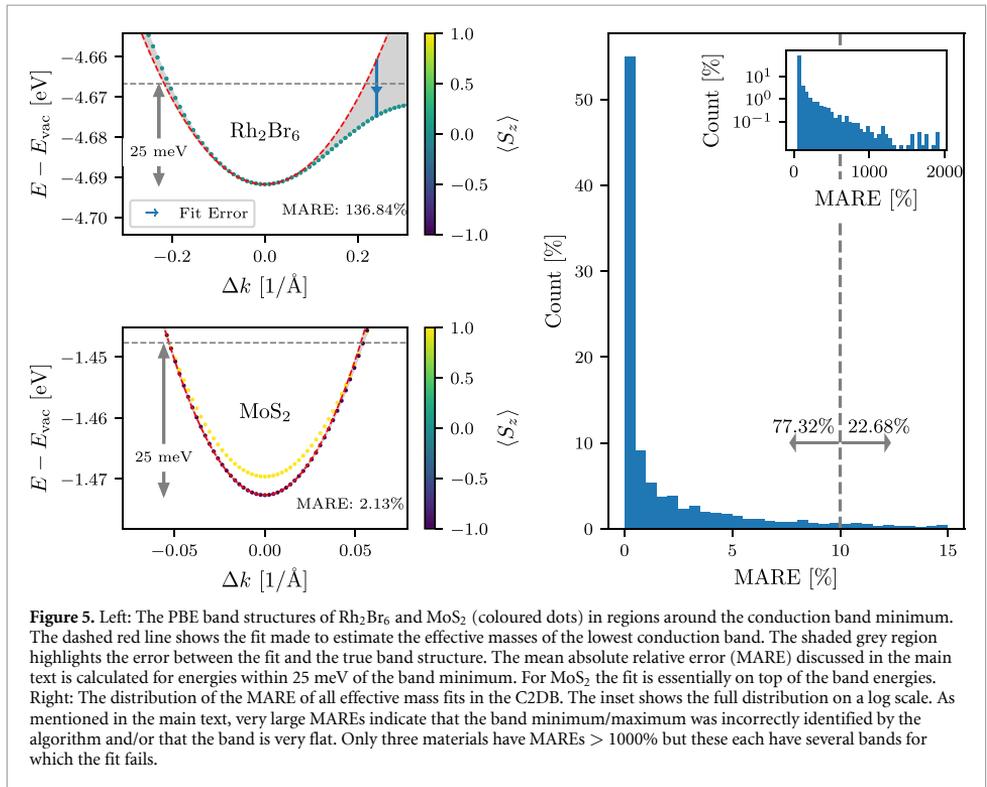
$$C = \begin{bmatrix} C_{xxxx} & C_{xyxy} & \sqrt{2}C_{xxxy} \\ C_{xyxy} & C_{yyyy} & \sqrt{2}C_{yyxy} \\ \sqrt{2}C_{xxxy} & \sqrt{2}C_{yyxy} & 2C_{xyxy} \end{bmatrix}, \quad (1)$$

where the indices on the matrix elements refer to the rank-4 tensor. The factors multiplying the tensor elements account for their multiplicities in the full rank-4 tensor. In the C2DB workflow, C is calculated as a finite difference of the stress under an applied strain with full relaxation of atomic coordinates. A negative eigenvalue of C signals a dynamical instability, see section 2.7.

In the first version of the C2DB only the diagonal elements of the stiffness tensor were calculated. The new version also determines the shear components such that the full 3×3 stiffness tensor is now available. This improvement also leads to a more accurate assessment of dynamical stability [57].

3.2. Effective masses with parabolic estimates

For all materials with a finite band gap the effective masses of electrons and holes are calculated for bands within 100 meV of the conduction band minimum and valence band maximum, respectively. The Hessian matrices at the band extrema (BE) are determined by fitting a second order polynomial to the PBE band structure including SOC, and the effective masses are obtained by subsequent diagonalisation of the Hessian. The main fitting-procedure is unaltered



from the first version of C2DB, but two important improvements have been made.

The first improvement consists in an additional k -mesh refinement step for better localisation of the BE in the Brillouin zone. After the location of the BE has been estimated based on a uniformly sampled band structure with k -point density of 12 \AA , another one-shot calculation is performed with a denser k -mesh around the estimated BE positions. This ensures a more accurate and robust determination of the location of the BE, which can be important in cases with a small but still significant spin-orbit splitting or when the band is very flat or non-quadratic around the BE. The second refinement step is the same as in the first version of C2DB, i.e. the band energies are calculated on a highly dense k -mesh in a small disc around the BE, and the Hessian is obtained by fitting the band energies in the range up to 1 meV from the BE.

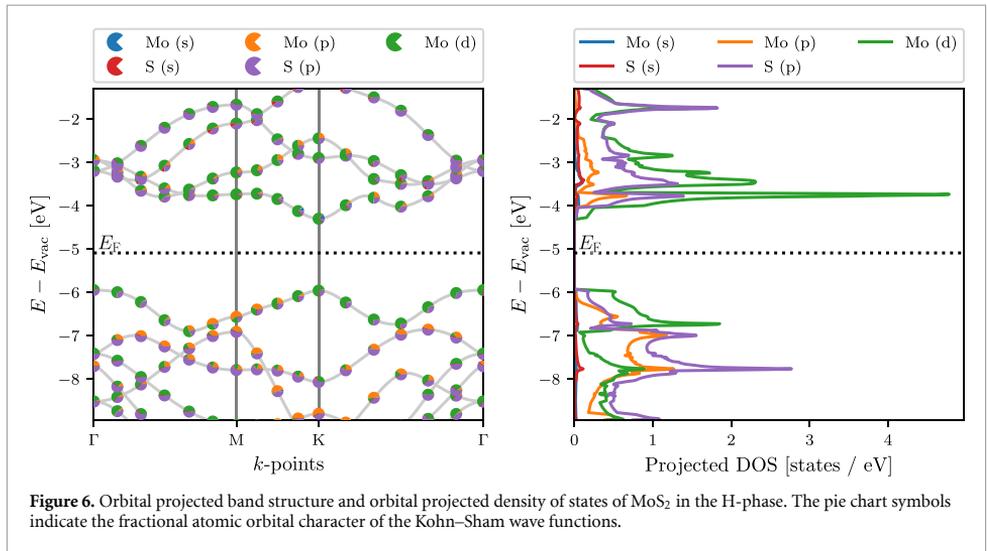
The second improvement is the calculation of the mean absolute relative error (MARE) of the polynomial fit in a 25 meV range from the BE. The value of 25 meV corresponds to the thermal energy at room temperature and is thus the relevant energy scale for many applications. To make the MARE independent of the absolute position of the band we calculate the average energy of the band over the 25 meV and compare the deviation of the fit to this energy scale. The MARE provides a useful measure of the parabolicity

of the energy bands and thus the validity of the effective mass approximation over this energy scale.

Figure 5 shows two examples of band structures with the effective mass fits and corresponding fit errors indicated. Additionally, the distribution of MARE for all the effective mass fits in the C2DB are presented. Most materials have an insignificant MARE, but a few materials have very large errors. Materials with a MARE above a few hundreds of percentages fall into two classes. For some materials the algorithm does not correctly find the position of the BE. An example is Ti_2S_2 in the space group C2/m. For others, the fit and BE location are both correct, but the band flattens away from the BE which leads to a large MARE as is the case for Rh_2Br_6 shown in the figure or Cl_2Ti_2 in the space group P-1. In general a small MARE indicates a parabolic band while materials with large MARE should be handled on a case-by-case basis.

3.3. Orbital projected band structure

To facilitate a state-specific analysis of the PBE Kohn-Sham wave functions, an orbital projected band structure (PBS) is provided to complement the projected density of states (PDOS). In the PAW methodology, the all-electron wave functions are projected onto atomic orbitals inside the augmentation spheres centred at the position of each atom. The PBS resolves these atomic orbital contributions to the



wave functions as a function of band and k -point whereas the PDOS resolves the atomic orbital character of the total density of states as a function of energy. The SOC is not included in the PBS or PDOS, as its effect is separately visualised by the spin-PBS also available in the C2DB.

As an example, figure 6 shows the PBS (left) and PDOS (right) of monolayer MoS₂ calculated with PBE. The relative orbital contribution to a given Bloch state is indicated by a pie chart symbol. In the present example, one can deduce from the PBS that even though Mo- p orbitals and S- p orbitals contribute roughly equally to the DOS in the valence band, the Mo- p orbital contributions are localised to a region in the BZ around the M -point, whereas the S- p orbitals contribute throughout the entire BZ.

3.4. Corrected G_0W_0 band structures

The C2DB contains G_0W_0 quasiparticle (QP) band structures of 370 monolayers covering 14 different crystal structures and 52 chemical elements. The details of these calculations can be found in the original C2DB paper [6]. A recent in-depth analysis of the 61.716 G_0W_0 data points making up the QP band structures led to several important conclusions relevant for high-throughput G_0W_0 calculations. In particular, it identified the linear QP approximation as a significant error source in standard G_0W_0 calculations and proposed an extremely simple correction scheme (the *empirical Z* (empZ) scheme), that reduces this error by a factor of two on average.

The empZ scheme divides the electronic states into two classes according to the size of the QP weight, Z . States with $Z \in [0.5, 1.0]$ are classified as QP consistent (QP-c) while states with $Z \notin [0.5, 1.0]$ are classified as QP inconsistent (QP-ic). With this definition, QP-c states will have at least half of their spectral weight in the QP peak. The distribution of

the 60.000+ Z -values is shown in figure 7. It turns out that the linear approximation to the self-energy, which is the gist of the QP approximation, introduces significantly larger errors for QP-ic states than for QP-c states. Consequently, the empZ method replaces the calculated Z of QP-ic states with the mean of the Z -distribution, $Z_0 \approx 0.75$. This simple replacement reduces the average error of the linear approximation from 0.11 to 0.06 eV.

An illustration of the method applied to MoS₂ is shown in figure 7. The original uncorrected G_0W_0 band structure is shown in blue while the empZ corrected band structure is shown in orange. MoS₂ has only one QP-ic state in the third conduction band at the K -point. Due to a break-down of the QP approximation for this state, the G_0W_0 correction is greatly overestimated leading to a local discontinuity in the band structure. The replacement of Z by Z_0 for this particular state resolves the problem. All G_0W_0 band structures in the C2DB are now empZ corrected.

3.5. Optical absorbance

In the first version of the C2DB, the optical absorbance was obtained from the simple expression [6]

$$A(\omega) \approx \frac{\omega \text{Im}\alpha^{2D}(\omega)}{\epsilon_0 c}, \quad (2)$$

where α^{2D} is the long wavelength limit of the in-plane sheet polarisability density (note that the equation is written here in SI units). The sheet polarisability is related to the sheet conductivity via $\sigma^{2D}(\omega) = -i\omega\alpha^{2D}(\omega)$. The expression (2) assumes that the electric field inside the layer equals the incoming field (i.e. reflection is ignored), and hence, it may overestimate the absorbance.

In the new version, the absorbance is evaluated from $A = 1 - R - T$, where R and T are the reflected and transmitted powers of a plane wave at normal

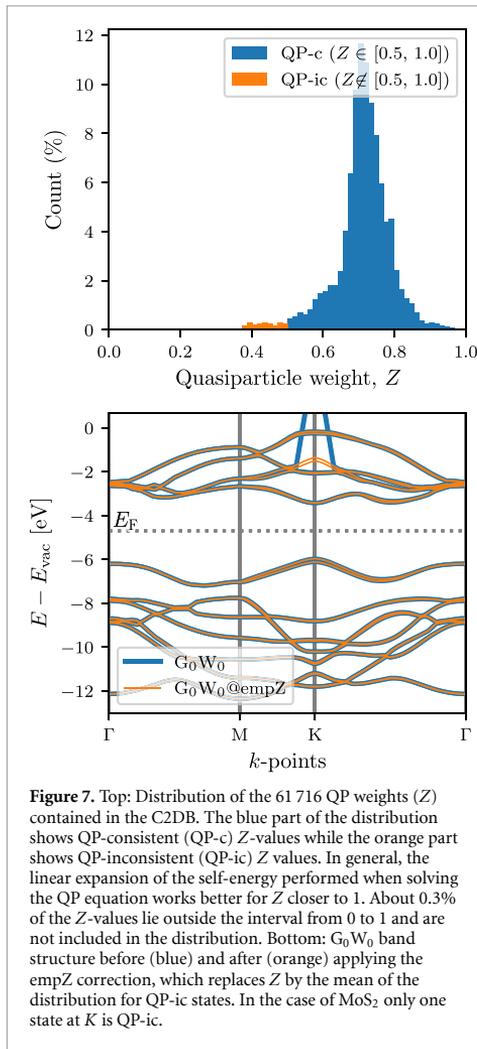


Figure 7. Top: Distribution of the 61 716 QP weights (Z) contained in the C2DB. The blue part of the distribution shows QP-consistent (QP-c) Z -values while the orange part shows QP-inconsistent (QP-ic) Z values. In general, the linear expansion of the self-energy performed when solving the QP equation works better for Z closer to 1. About 0.3% of the Z -values lie outside the interval from 0 to 1 and are not included in the distribution. Bottom: G_0W_0 band structure before (blue) and after (orange) applying the empZ correction, which replaces Z by the mean of the distribution for QP-ic states. In the case of MoS₂ only one state at K is QP-ic.

incidence, respectively. These can be obtained from the conventional transfer matrix method applied to a monolayer suspended in vacuum. The 2D material is here modelled as an infinitely thin layer with a sheet conductivity. Alternatively, it can be modelled as quasi-2D material of thickness d with a ‘bulk’ conductivity of $\sigma = \sigma^{2D}/d$ [58], but the two approaches yield very similar results, since the optical thickness of a 2D material is much smaller than the optical wavelength. Within this model, the expression for the absorbance of a suspended monolayer with the sheet conductivity σ^{2D} reads:

$$A(\omega) = \text{Re} \left\{ \sigma^{2D}(\omega)\eta_0 \right\} \left| \frac{2}{2 + \sigma^{2D}(\omega)\eta_0} \right|^2, \quad (3)$$

where $\eta_0 = 1/(\epsilon_0 c) \approx 377 \Omega$ is the vacuum impedance.

If the light–matter interaction is weak, i.e. $|\sigma^{2D}\eta_0| \ll 1$, equation (3) reduces to equation (2).

Nonetheless, due the strong light–matter interaction in some 2D materials, this approximation is not reliable in general. In fact, it can be shown that the maximum possible absorption from equation (3) is 50%, which is known as the upper limit of light absorption in thin films [59]. This limit is not guaranteed by equation (2), which can even yield an absorbance above 100%.

As an example, figure 8 shows the absorption spectrum of monolayer MoS₂ for in- and out-of-plane polarised light as calculated with the exact equation (3) and the approximate equation (2), respectively. In all cases the sheet polarisability is obtained from the BSE to account for excitonic effects [6]. For weak light–matter interactions, e.g. for the z -polarised light, the two approaches agree quite well, but noticeable differences are observed in regions with stronger light–matter interaction.

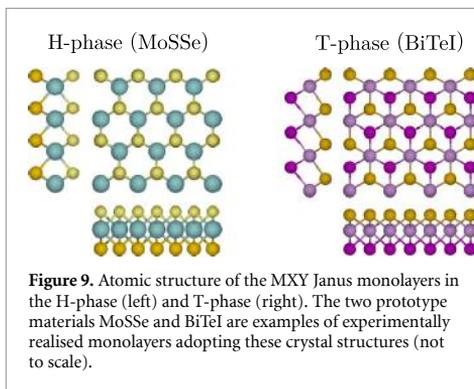
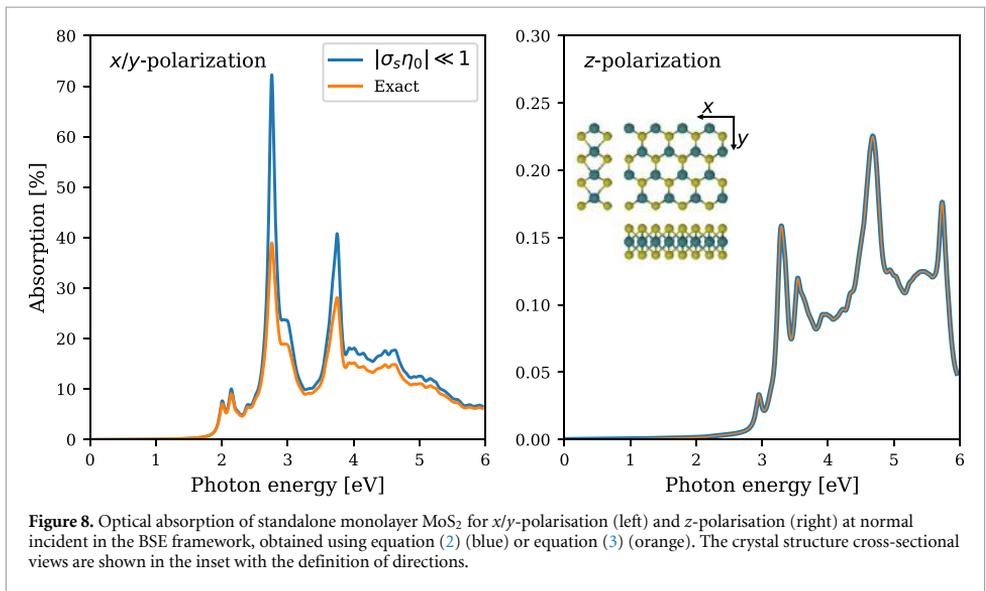
4. New materials in the C2DB

In this section we discuss the most significant extensions of the C2DB in terms of new materials. The set of materials presented here is not complete, but represents the most important and/or well defined classes. The materials discussed in sections 4.1 and 4.2 (MXY Janus monolayers and monolayers extracted from experimental crystal structure databases) are already included in the C2DB. The materials described in sections 4.3 and 4.4 (homo-bilayers and monolayer point defect systems) will soon become available as separate C2DB-interlinked databases.

4.1. MXY Janus monolayers

The class of TMDC monolayers of the type MX₂ (where M is the transition metal and X is a chalcogen) exhibits a large variety of interesting and unique properties and has been widely discussed in the literature [60]. Recent experiments have shown that it is not only possible to synthesise different materials by changing the metal M or the chalcogen X, but also by exchanging the X on one side of the layer by another chalcogen (or halogen) [61–63]. This results in a class of 2D materials known as MXY Janus monolayers with broken mirror symmetry and finite out-of-plane dipole moments. The prototypical MXY crystal structures are shown in figure 9 for the case of MoSSe and BiTeI, which have both been experimentally realised [61–63]. Adopting the nomenclature from the TMDCs, the crystal structures are denoted as H- or T-phase, depending on whether X and Y atoms are vertically aligned or displaced, respectively.

In a recent work [64], the C2DB workflow was employed to scrutinise and classify the basic electronic and optical properties of 224 different MXY Janus monolayers. All data from the study is available in the C2DB. Here we provide a brief discussion of the Rashba physics in these materials and refer the



interested reader to [64] for more details and analysis of other properties.

A key issue when considering hypothetical materials, i.e. materials not previously synthesised, is their stability. The experimentally synthesised MoSSe and BiTeI are both found to be dynamically stable and lie within 10 meV of the convex hull confirming their thermodynamic stability. Out of the 224 initial monolayers 93 are classified as stable according to the C2DB criteria (dynamically stable and $\Delta H_{\text{hull}} < 0.2 \text{ eV atom}^{-1}$). Out of the 93 stable materials, 70 exhibit a finite band gap when computed with the PBE xc-functional.

The Rashba effect is a momentum dependent splitting of the band energies of a 2D semiconductor in the vicinity of a band extremum arising due to the combined effect of spin-orbit interactions and a broken crystal symmetry in the direction perpendicular to the 2D plane. The simplest model used to describe the Rashba effect is a 2D electron gas in a perpendicular electric field (along the *z*-axis). Close to

the band extremum, the energy of the two spin bands is described by the Rashba Hamiltonian [65, 66]:

$$H = \alpha_R (\boldsymbol{\sigma} \times \mathbf{k}) \cdot \hat{\mathbf{e}}_z, \quad (4)$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices, $\mathbf{k} = \mathbf{p}/\hbar$ is the wave number, and the Rashba parameter is proportional to the electric field strength, $\alpha_R \propto E_0$.

Although the Rashba Hamiltonian is only meant as a qualitative model, it is of interest to test its validity on the Janus monolayers. The electric field of the Rashba model is approximately given by $E_0 = \Delta V_{\text{vac}}/d$, where ΔV_{vac} is the shift in vacuum potential on the two sides of the layer (see left inset of figure 10) and d is the layer thickness. Assuming a similar thickness for all monolayers, the electric field is proportional to the potential shift. Not unexpected, the latter is found to correlate strongly with the difference in electronegativity of the X and Y atoms, see left panel of figure 10.

The Rashba energy, E_R , can be found by fitting $E(k) = \hbar^2 k^2 / 2m^* \pm \alpha_R k$ to the band structure (see right inset of figure 10) and should scale with the electric field strength. However, as seen from the right panel of figure 10, there is no correlation between the two quantities. Hence we conclude that the simple Rashba model is completely inadequate and that the strength of the perpendicular electric field cannot be used to quantify the effect of spin-orbit interactions on band energies.

4.2. Monolayers from known layered bulk crystals

The C2DB has been extended with a number of monolayers that are likely exfoliable from experimentally known layered bulk compounds. Specifically, the Inorganic Crystal Structure Database (ICSD) [67] and Crystallography Open Database (COD) [68]

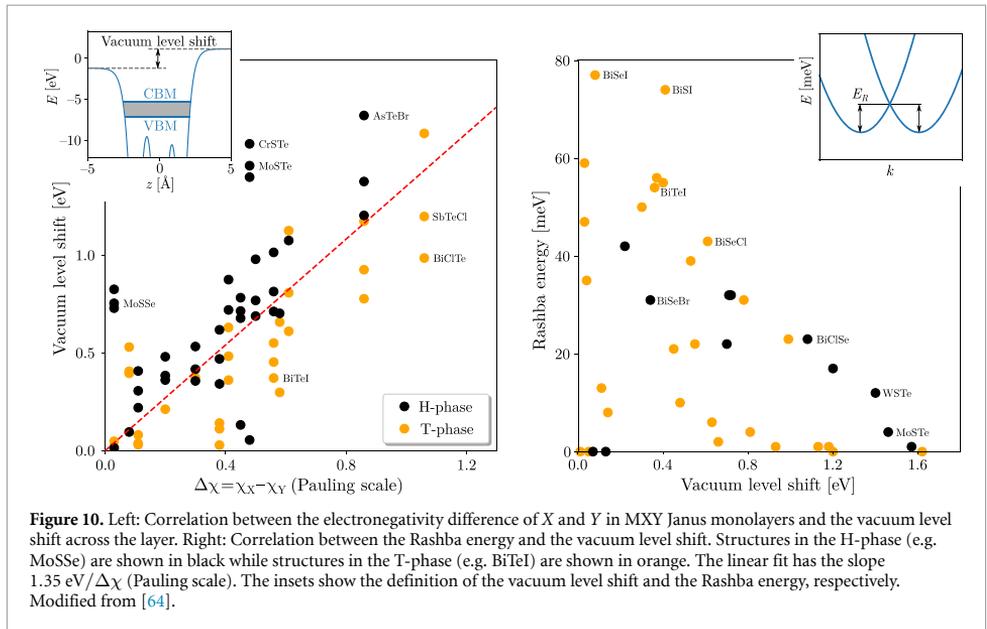


Figure 10. Left: Correlation between the electronegativity difference of X and Y in MX_Y Janus monolayers and the vacuum level shift across the layer. Right: Correlation between the Rashba energy and the vacuum level shift. Structures in the H-phase (e.g. MoSSe) are shown in black while structures in the T-phase (e.g. BiTeI) are shown in orange. The linear fit has the slope 1.35 eV/ $\Delta\chi$ (Pauling scale). The insets show the definition of the vacuum level shift and the Rashba energy, respectively. Modified from [64].

have first been filtered for corrupted, duplicate and theoretical compounds, which reduce the initial set of 585.485 database entries to 167.767 unique materials. All of these have subsequently been assigned a ‘dimensionality score’ based on a purely geometrical descriptor. If the 2D score is larger than the sum of 0D, 1D and 3D scores we regard the material as being exfoliable and we extract the individual 2D components that comprise the material (see also section 2.2). We refer to the original work on the method for details [52] and note that similar approaches were applied in [11, 12] to identify potentially exfoliable monolayers from the ICSD and COD.

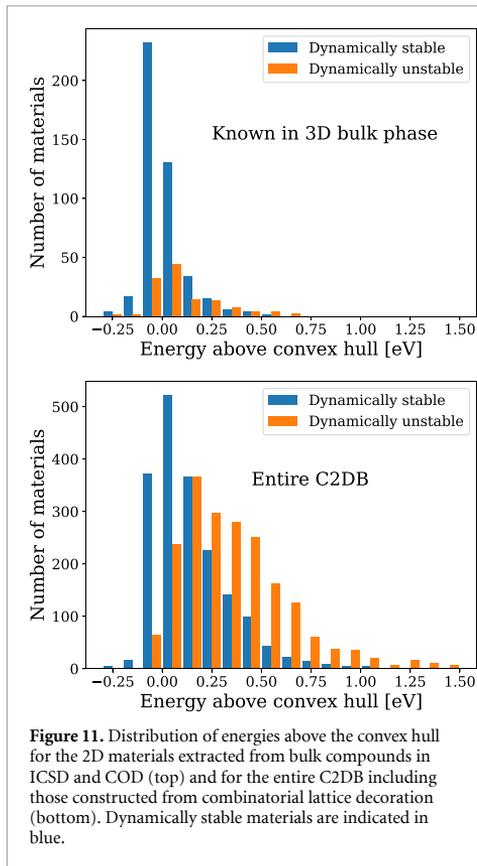
The search has been limited to bulk compounds containing less than six different elements and no rare earth elements. This reduces the set of relevant bulk materials to 2991. For all of these we extracted the 2D components containing less than 21 atoms in the unit cell, which were then relaxed and sorted for duplicates following the general C2DB workflow steps described in sections 2.1–2.3. At this point 781 materials remain. This set includes most known 2D materials and 207 of the 781 were already present in the C2DB prior to this addition. All the materials (including those that were already in C2DB) have been assigned an ICSD/COD identifier that refers to the parent bulk compound from which the 2D material was computationally exfoliated. We emphasise that we have not considered exfoliation energies in the analysis and a subset of these materials may thus be rather strongly bound and challenging to exfoliate even if the geometries indicate van der Waals bonded structures of the parent bulk compounds.

Figure 11 shows the distribution of energies above the convex hull for materials derived from

parent structures in ICSD or COD as well as for the entire C2DB, which includes materials obtained from combinatorial lattice decoration as well. As expected, the materials derived from experimental bulk materials are situated rather close to the convex hull whereas those obtained from lattice decoration extend to energies far above the convex hull. It is also observed that a larger fraction of the experimentally derived materials are dynamically stable. There are, however, well known examples of van der Waals bonded structures where the monolayer undergoes a significant lattice distortion, which will manifest itself as a dynamical instability in the present context. For example, bulk MoS₂ exists in van der Waals bonded structures composed of either 2 H-MoS₂ or 1 T-MoS₂ layers, but a monolayer of the 1 T phase undergoes a structural deformation involving a doubling of the unit cell [69] and is thus categorised as dynamically unstable by the C2DB workflow. The dynamically stable materials derived from parent bulk structures in the ICSD and COD may serve as a useful subset of the C2DB that are likely to be exfoliable from known compounds and thus facilitate experimental verification. As a first application the subset has been used to search for magnetic 2D materials, which resulted in a total of 85 ferromagnets and 61 anti-ferromagnets [70].

4.3. Outlook: multilayers

The C2DB is concerned with the properties of covalently bonded monolayers (see discussion of dimensionality filtering in section 2.2). However, multilayer structures composed of two or more identical monolayers are equally interesting and often have properties that deviate from those of the monolayer. In fact, the synthesis of layered vdW structures with a



controllable number of layers represents an interesting avenue for atomic-scale materials design. Several examples of novel phenomena emerging in layered vdW structures have been demonstrated including direct-indirect band gap transitions in MoS_2 [71, 72], layer-parity selective Berry curvatures in few-layer WTe_2 [73], thickness-dependent magnetic order in CrI_3 [74, 75], and emergent ferroelectricity in bilayer hBN [76].

As a first step towards a systematic exploration of multilayer 2D structures, the C2DB has been used as basis for generating homobilayers in various stacking configurations and subsequently computing their properties following a modified version of the C2DB monolayer workflow. Specifically, the most stable monolayers (around 1000) are combined into bilayers by applying all possible transformations (unit cell preserving point group operations and translations) of one layer while keeping the other fixed. The candidate bilayers generated in this way are subject to a stability analysis, which evaluates the binding energy and optimal IL distance based on PBE-D3 [77] total energy calculations keeping the atoms of the monolayers fixed in their PBE relaxed geometry, see figures 12 and table 3.

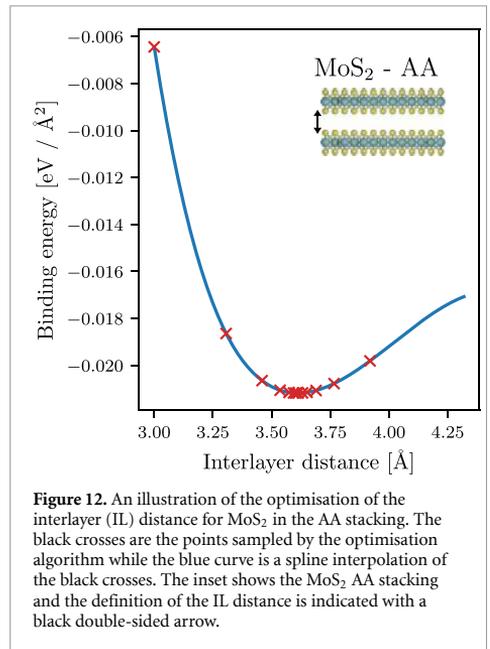


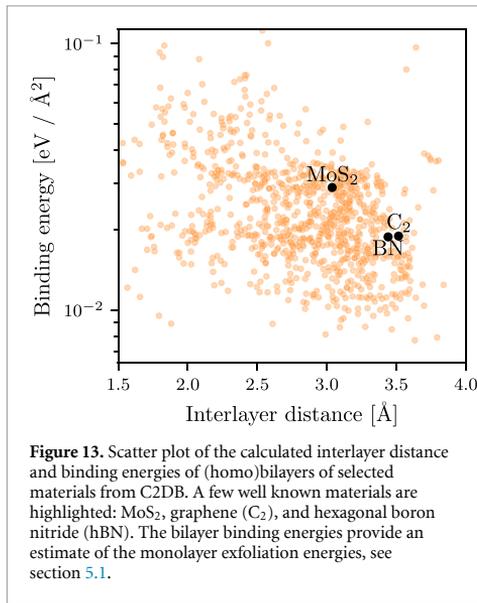
Table 3. Exfoliation energies for selected materials calculated with the PBE+D3 xc-functional as described in section 4.3 and compared with the DF2 and rVV10 results from [11]. The spacegroups are indicated in the column 'SG'. All numbers are in units of $\text{meV } \text{Å}^{-2}$.

Material	SG	PBE + D3	DF2	rVV10
MoS_2	P-6m2	28.9	21.6	28.8
MoTe_2	P-6m2	30.3	25.2	30.4
ZrNBr	Pmmn	18.5	10.5	18.5
C	P6/mmm	18.9	20.3	25.5
P	Pmna	21.9	38.4	30.7
BN	P-6m2	18.9	19.4	24.4
WTe_2	P-6m2	32.0	24.7	30.0
PbTe	P3m1	23.2	27.5	33.0

The calculated IL binding energies are generally in the range from a few to a hundred $\text{meV } \text{Å}^{-2}$ and IL distances range from 1.5 to 3.8 Å . A scatter plot of preliminary binding energies and IL distances is shown in figure 13. The analysis of homobilayers provides an estimate of the energy required to peel a monolayer off a bulk structure. In particular, the binding energy for the most stable bilayer configuration provides a measure of the *exfoliation energy* of the monolayer. This key quantity is now available for all monolayers in the C2DB, see section 5.1.

4.4. Outlook: point defects

The C2DB is concerned with the properties of 2D materials in their pristine crystalline form. However, as is well known the perfect crystal is an idealised model of real materials, which always contain defects in smaller or larger amounts depending on the intrinsic materials properties and growth conditions. Crystal defects often have a negative impact on



physical properties, e.g. they lead to scattering and life time-reduction of charge carriers in semiconductors. However, there are also important situations where defects play a positive enabling role, e.g. in doping of semiconductors, as colour centres for photon emission [78, 79] or as active sites in catalysis.

To reduce the gap between the pristine model material and real experimentally accessible samples, a systematic evaluation of the basic properties of the simplest native point defects in a selected subset of monolayers from the C2DB has been initiated. The monolayers are selected based on the stability of the pristine crystal. Moreover, only non-magnetic semiconductors with a PBE band gap satisfying $E_{\text{gap}} > 1$ eV are currently considered as such materials are candidates for quantum technology applications like single-photon sources and spin qubits. Following these selection criteria around 300 monolayers are identified and their vacancies and intrinsic substitutional defects are considered, yielding a total of about 1500 defect systems.

Each defect system is subject to the same workflow, which is briefly outlined below. To enable point defects to relax into their lowest energy configuration, the symmetry of the pristine host crystal is intentionally broken by the chosen supercell, see figure 14 (a). In order to minimise defect–defect interaction, supercells are furthermore chosen such that the minimum distance between periodic images of defects is larger than 15 Å. Unique point defects are created based on the analysis of equivalent Wyckoff positions for the host material. To illustrate some of the properties that will feature in the upcoming point defect database, we consider the specific example of monolayer CH₂Si.

First, the formation energy [80, 81] of a given defect is calculated from PBE total energies. Next,

Slater–Janak transition state theory is used to obtain the charge transition levels [82, 83]. By combining these results, one obtains the formation energy of the defect in all possible charge states as a function of the Fermi level. An example of such a diagram is shown in figure 14 (b) for the case of the V_C and C_{Si} defects in monolayer CH₂Si. For each defect and each charge state, the PBE single-particle energy level diagram is calculated to provide a qualitative overview of the electronic structure. A symmetry analysis [84] is performed for the defect structure and the individual defect states lying inside the band gap. The energy level diagram of the neutral V_{Si} defect in CH₂Si is shown in figure 14 (c), where the defect states are labelled according to the irreducible representations of the C_s point group.

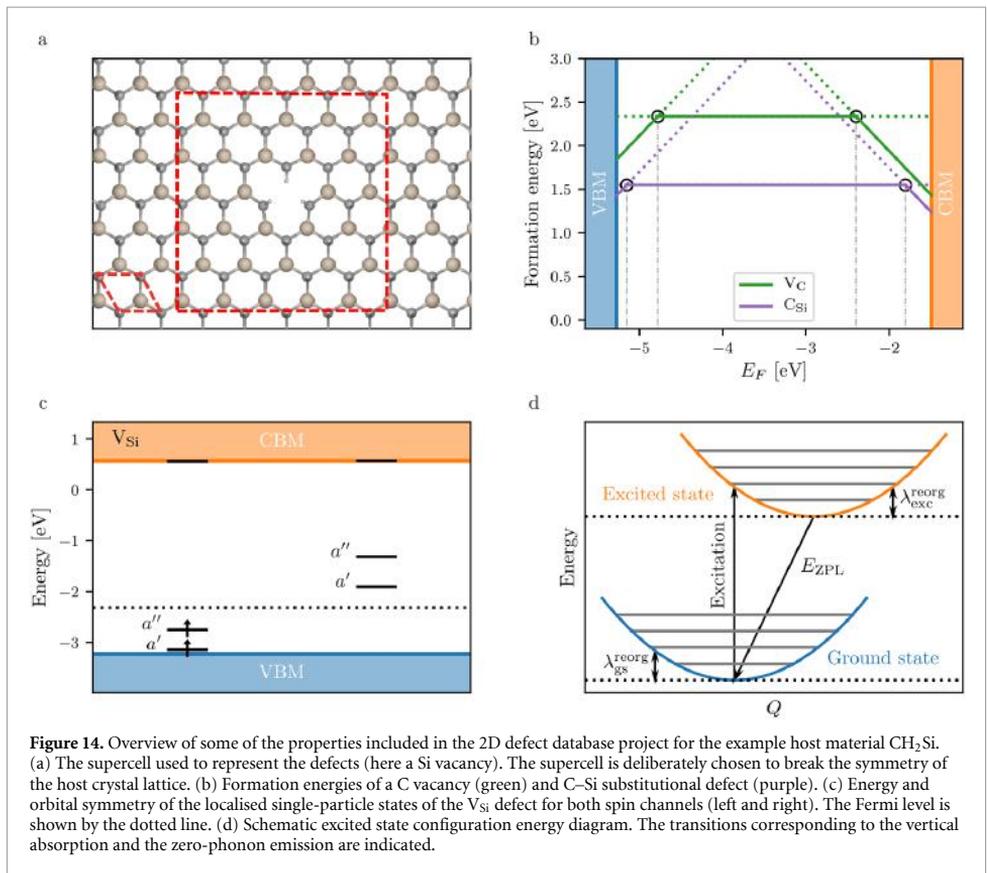
In general, excited electronic states can be modelled by solving the Kohn–Sham equations with non-Aufbau occupations. The excited-state solutions are saddle points of the Kohn–Sham energy functional, but common self-consistent field (SCF) approaches often struggle to find such solutions, especially when nearly degenerate states are involved. The calculation of excited states corresponding to transitions between localised states inside the band gap is therefore performed using an alternative method based on the direct optimisation (DO) of orbital rotations in combination with the maximum overlap method (MOM) [85]. This method ensures fast and robust convergence of the excited states, as compared to SCF. In figure 14 (d), the reorganisation energies for the ground and excited state, as well as the zero-phonon line (ZPL) energy are sketched. For the specific case of the Si vacancy in CH₂Si, the DO-MOM method yields $E_{\text{ZPL}} = 3.84$ eV, $\lambda_{\text{gs}}^{\text{reorg}} = 0.11$ eV and $\lambda_{\text{exc}}^{\text{reorg}} = 0.16$ eV. For systems with large electron-phonon coupling (i.e. Huang–Rhys factor > 1) a one-dimensional approximation for displacements along the main phonon mode is used to produce the configuration coordinate diagram (see figure 14 (d)). In addition to the ZPL energies and reorganisation energies, the Huang–Rhys factors, photoluminescence spectrum from the 1D phonon model, hyperfine coupling and zero field splitting are calculated.

5. New properties in the C2DB

This section reports on new properties that have become available in the C2DB since the first release. The employed computational methodology is described in some detail and results are compared to the literature where relevant. In addition, some interesting property correlations are considered along with general discussions of the general significance and potential application of the available data.

5.1. Exfoliation energy

The exfoliation energy of a monolayer is estimated as the binding energy of its bilayer in the most stable



stacking configuration (see also section 4.3). The binding energy is calculated using the PBE + D3 xc-functional [86] with the atoms of both monolayers fixed in the PBE relaxed geometry. Table 3 compares exfoliation energies obtained in this way to values from Mounet *et al* [11] for a representative set of monolayers.

5.2. Bader charges

For all monolayers we calculate the net charge on the individual atoms using the Bader partitioning scheme [87]. The analysis is based purely on the electron density, which we calculate from the PAW pseudo density plus compensation charges using the PBE xc-functional. Details of the method and its implementation can be found in Tang *et al* [88]. In section 5.4 we compare and discuss the relation between Bader charges and Born charges.

5.3. Spontaneous polarisation

The spontaneous polarisation (\mathbf{P}_s) of a bulk material is defined as the charge displacement with respect to that of a reference centrosymmetric structure [89, 90]. Ferroelectric materials exhibit a finite value

of \mathbf{P}_s that may be switched by an applied external field and have attracted a large interest for a wide range of applications [91–93].

The spontaneous polarisation in bulk materials can be regarded as electric dipole moment per unit volume, but in contrast to the case of finite systems this quantity is ill-defined for periodic crystals [89]. Nevertheless, one can define the formal polarisation density:

$$\mathbf{P} = \frac{1}{2\pi} \frac{e}{V} \sum_l \phi_l \mathbf{a}_l, \quad (5)$$

where \mathbf{a}_l (with $l \in \{1, 2, 3\}$) are the lattice vectors spanning the unit cell, V is the cell volume and e is the elementary charge. ϕ_l is the polarisation phase along the lattice vector defined by:

$$\phi_l = \sum_i Z_i \mathbf{b}_l \cdot \mathbf{u}_i - \phi_l^{\text{elec}}, \quad (6)$$

where \mathbf{b}_l is the reciprocal lattice vector satisfying $\mathbf{b}_l \cdot \mathbf{R}_l = 2\pi$ and \mathbf{u}_i is the position of nucleus i with charge eZ_i . The electronic contribution to the polarisation phase is defined as:

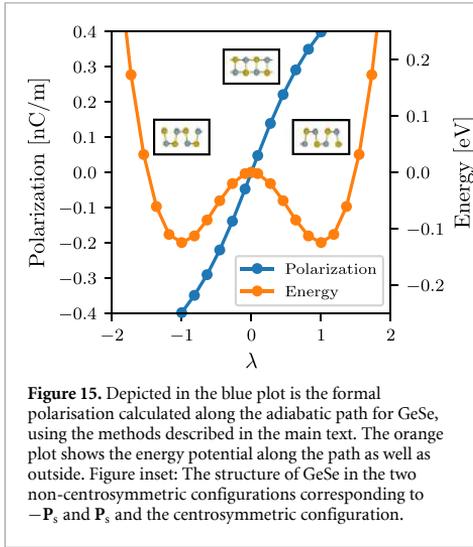


Figure 15. Depicted in the blue plot is the formal polarisation calculated along the adiabatic path for GeSe, using the methods described in the main text. The orange plot shows the energy potential along the path as well as outside. Figure inset: The structure of GeSe in the two non-centrosymmetric configurations corresponding to $-\mathbf{P}_s$ and \mathbf{P}_s and the centrosymmetric configuration.

$$\phi_l^{\text{elec}} = \frac{1}{N_{\mathbf{k} \perp \mathbf{b}_l}} \text{Im} \sum_{\mathbf{k} \in \text{BZ}_{\perp \mathbf{b}_l}} \times \ln \prod_{j=0}^{N_{\mathbf{k} \parallel \mathbf{b}_l} - 1} \det_{\text{occ}} [u_{m\mathbf{k}+j\delta\mathbf{k}} | u_{m\mathbf{k}+(j+1)\delta\mathbf{k}}], \quad (7)$$

where $\text{BZ}_{\perp \mathbf{b}_l} = \{\mathbf{k} | \mathbf{k} \cdot \mathbf{b}_l = 0\}$ is a plane of \mathbf{k} -points orthogonal to \mathbf{b}_l , $\delta\mathbf{k}$ is the distance between neighbouring \mathbf{k} -points in the \mathbf{b}_l direction and $N_{\mathbf{k} \parallel \mathbf{b}_l}$ ($N_{\mathbf{k} \perp \mathbf{b}_l}$) is the number of \mathbf{k} -points along (perpendicular to) the \mathbf{b}_l direction. These expressions generalise straightforwardly to 2D.

The formal polarisation is only well-defined modulo $e\mathbf{R}_n/V$ where \mathbf{R}_n is any lattice vector. However, changes in polarisation are well defined and the spontaneous polarisation may thus be obtained by:

$$\mathbf{P}_s = \int_0^1 \frac{d\mathbf{P}(\lambda)}{d\lambda} d\lambda, \quad (8)$$

where λ is a dimensionless parameter that defines an adiabatic structural path connecting the polar phase ($\lambda = 1$) with a non-polar phase ($\lambda = 0$).

The methodology has been implemented in GPAW and used to calculate the spontaneous polarisation of all stable materials in the C2DB with a PBE band gap above 0.01 eV and a polar space group symmetry. For each material, the centrosymmetric phase with smallest atomic displacement from the polar phase is constructed and relaxed under the constraint of inversion symmetry. The adiabatic path connecting the two phases is then used to calculate the spontaneous polarisation using equations (5)–(8). An example of a calculation for GeSe is shown in figure 15 where the polarisation along the path connecting two equivalent polar phases via the centrosymmetric phase is shown together with the total energy. The

spontaneous polarisation obtained from the path is 39.8 nC m^{-1} in good agreement with previous calculations [94].

5.4. Born charges

The Born charge of an atom a at position \mathbf{u}_a in a solid is defined as:

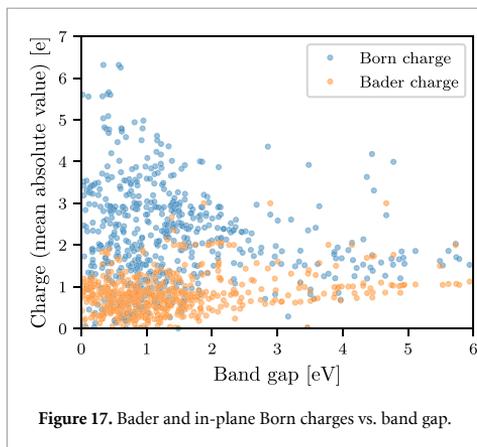
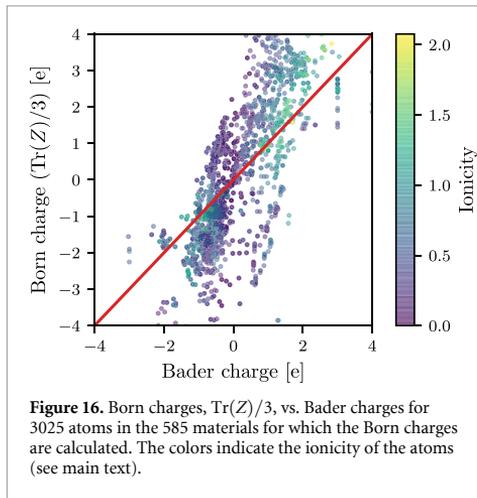
$$Z_{ij}^a = \frac{V}{e} \left. \frac{\partial P_i}{\partial u_{aj}} \right|_{E=0}. \quad (9)$$

It can be understood as an effective charge assigned to the atom to match the change in polarisation in direction i when its position is perturbed in direction j . Since the polarisation density and the atomic position are both vectors, the Born charge of an atom is a rank-2 tensor. The Born charge is calculated as a finite difference and relies on the Modern theory of polarisation [95] for the calculation of polarisation densities, see reference [96] for more details. The Born charge has been calculated for all stable materials in C2DB with a finite PBE band gap.

It is of interest to examine the relation between the Born charge and the Bader charge (see section 5.2). In materials with strong ionic bonds one would expect the charges to follow the atoms. On the other hand, in covalently bonded materials the hybridisation pattern and thus the charge distribution, depends on the atom positions in a complex way, and the idea of charges following the atom is expected to break down. In agreement with this idea, the (in-plane) Born charges in the strongly ionic hexagonal hBN ($\pm 2.71e$ for B and N, respectively) are in good agreement with the calculated Bader charges ($\pm 3.0e$). In contrast, (the in-plane) Born charges in MoS_2 ($-1.08e$ and $0.54e$ for Mo and S, respectively) deviate significantly from the Bader charges ($1.22e$ and $-0.61e$ for Mo and S, respectively). In fact, the values disagree even on the sign of the charges underlining the non-intuitive nature of the Born charges in covalently bonded materials.

Note that the out-of-plane Born charges never match the Bader charges, even for strongly ionic insulators, and are consistently smaller in value than the in-plane components. The smaller out-of-plane values are consistent with the generally smaller out-of-plane polarisability of 2D materials (for both electronic and phonon contributions) and agrees with the intuitive expectation that it is more difficult to polarise a 2D material in the out-of-plane direction as compared to the in-plane direction.

Figure 16 shows the average of the diagonal of the Born charge tensor, $\text{Tr}(Z^a)/3$, plotted against the Bader charges for all 585 materials in the C2DB for which the Born charges have been computed. The data points have been coloured according to the ionicity of the atom a defined as $I(a) = |\chi_a - \langle \chi \rangle|$, where χ_a and $\langle \chi \rangle$ are the Pauling electronegativity of atom a and the average electronegativity of all atoms in the

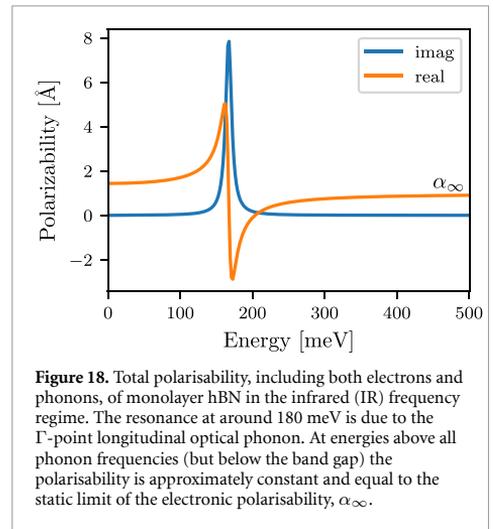


unit cell, respectively. The ionicity is thus a measure of the tendency of an atom to donate/accept charge relative to the average tendency of atoms in the material. It is clear from figure 16 that there is a larger propensity for the Born and Bader charges to match in materials with higher ionicity.

Figure 17 plots the average (in-plane) Born charge and the Bader charge versus the band gap. It is clear that large band gap materials typically exhibit integer Bader charges, whereas there is no clear correlation between the Born charge and the band gap.

5.5. Infrared polarisability

The original C2DB provided the frequency dependent polarisability computed in the random phase approximation (RPA) with inclusion of electronic interband and intraband (for metals) transitions [6]. However, phonons carrying a dipole moment (so-called IR active phonons) also contribute to the polarisability at frequencies comparable to the frequency of optical phonons. This response is described by the IR polarisability:



$$\alpha^{\text{IR}}(\omega) = \frac{e^2}{A} \mathbf{Z}^T \mathbf{M}^{-1/2} \left(\sum_i \frac{\mathbf{d}_i \mathbf{d}_i^T}{\omega_i^2 - \omega^2 - i\gamma\omega} \right) \mathbf{M}^{-1/2} \mathbf{Z}, \quad (10)$$

where \mathbf{Z} and \mathbf{M} are matrix representations of the Born charges and atomic masses, ω_i^2 and d_i are eigenvectors and eigenvalues of the dynamical matrix, A is the in-plane cell area and γ is a broadening parameter representing the phonon lifetime and is set to 10 meV. The total polarisability is then the sum of the electronic polarisability and the IR polarisability.

The new C2DB includes the IR polarisability of all monolayers for which the Born charges have been calculated (stable materials with a finite band gap), see section (5.4). As an example, figure 18 shows the total polarisability of monolayer hexagonal hBN. For details on the calculation of the IR polarisability see reference [96].

5.6. Piezoelectric tensor

The piezoelectric effect is the accumulation of charges, or equivalently the formation of an electric polarisation, in a material in response to an applied mechanical stress or strain. It is an important material characteristic with numerous scientific and technological applications in sonar, microphones, accelerometers, ultrasonic transducers, energy conversion, etc [97, 98]. The change in polarisation originates from the movement of positive and negative charge centres as the material is deformed.

Piezoelectricity can be described by the (proper) piezoelectric tensor c_{ijk} with $i, j, k \in \{x, y, z\}$, given by [99]:

$$c_{ijk} = \frac{e}{2\pi V} \sum_l \frac{\partial \phi_l}{\partial \epsilon_{jk}} a_{li}, \quad (11)$$

which differs from equation (5) only by a derivative of the polarisation phase with respect to the strain tensor

Table 4. Comparison of computed piezoelectric tensor versus experimental values and previous calculations for hexagonal BN and a selected set of TMDCs (space group 187). All numbers are in units of nC/m. Experimental data for MoS₂ is obtained from [102].

Material	Exp.	Theory [101]	C2DB
BN	—	0.14	0.13
MoS ₂	0.3	0.36	0.35
MoSe ₂	—	0.39	0.38
MoTe ₂	—	0.54	0.48
WS ₂	—	0.25	0.24
WSe ₂	—	0.27	0.26
WTe ₂	—	0.34	0.34

ϵ_{jk} . Note that c_{ijk} does not depend on the chosen branch cut.

The piezoelectric tensor is a symmetric tensor with at most 18 independent components. Furthermore, the point group symmetry restricts the number of independent tensor elements and their relationships due to the well-known Neumann's principle [100]. For example, monolayer MoS₂ with point group D_{3h} , has only one non-vanishing independent element of c_{ijk} . Note that c_{ijk} vanishes identically for centrosymmetric materials. Using a finite-difference technique with a finite but small strain (1% in our case), equation (11) has been used to compute the proper piezoelectric tensor for all non-centrosymmetric materials in the C2DB with a finite band gap. Table 4 shows a comparison of the piezoelectric tensors in the C2DB with literature for a selected set of monolayer materials. Good agreement is obtained for all these materials.

5.7. Topological invariants

For all materials in the C2DB exhibiting a direct band gap below 1 eV, the k -space Berry phase spectrum of the occupied bands has been calculated from the PBE wave functions. Specifically, a particular k -point is written as $k_1\mathbf{b}_1 + k_2\mathbf{b}_2$ and the Berry phases $\gamma_n(k_2)$ of the occupied states on the path $k_1 = 0 \rightarrow k_1 = 1$ is calculated for each value of k_2 . The connectivity of the Berry phase spectrum determines the topological properties of the 2D Bloch Hamiltonian [103, 104].

The calculated Berry phase spectra of the relevant materials are available for visual inspection on the C2DB webpage. Three different topological invariants have been extracted from these spectra and are reported in the C2DB: (1) The Chern number, C , takes an integer value and is well defined for any gapped 2D material. It determines the number of chiral edge states on any edge of the material. For any non-magnetic material the Chern number vanishes due to time-reversal symmetry. It is determined from the Berry phase spectrum as the number of crossings at any horizontal line in the spectrum. (2) The mirror Chern number, C_M , defined for gapped materials with a mirror plane in the atomic layer [105]. For such materials, all states may be chosen as mirror

eigenstates with eigenvalues $\pm i$ and the Chern numbers C_{\pm} can be defined for each mirror sector separately. For a material with vanishing Chern number, the mirror Chern number is defined as $C_M = (C_+ - C_-)/2$ and takes an integer value corresponding to the number of edge states on any mirror symmetry preserving edge. It is obtained from the Berry phase spectrum as the number of chiral crossings in each of the mirror sectors. (3) The Z_2 invariant, ν , which can take the values 0 and 1, is defined for materials with time-reversal symmetry. Materials with $\nu = 1$ are referred to as quantum spin Hall insulators and exhibit helical edge states at any time-reversal conserving edge. It is determined from the Berry phase spectrum as the number of crossing points modulus 2 at any horizontal line in the interval $k_2 \in [0, 1/2]$.

Figure 19 shows four representative Berry phase spectra corresponding to the three cases of non-vanishing C , C_M and ν as well as a trivial insulator. The four materials are: OsCl₃ (space group 147)—a Chern insulator with $C = 1$, OsTe₂ (space group 14)—a mirror crystalline insulator with $C_M = 2$, SbI (spacegroup 1)—a quantum spin Hall insulator with $\nu = 1$ and BiTe (spacegroup 156)—a trivial insulator. Note that a gap in the Berry phase spectrum always implies a trivial insulator.

In [106] the C2DB was screened for materials with non-trivial topology. At that point it was found that the database contained 7 Chern insulators, 21 mirror crystalline topological insulators and 48 quantum spin Hall insulators. However, that does not completely exhaust the the topological properties of materials in the C2DB. In particular, there may be materials that can be topologically classified based on crystalline symmetries other than the mirror plane of the layer. In addition, second order topological effects may be present in certain materials, which imply that flakes will exhibit topologically protected corner states. Again, the Berry phase spectra may be used to unravel the second order topology by means of nested Wilson loops [107].

5.8. Exchange coupling constants

The general C2DB workflow described in sections 2.1–2.3 will identify the FM ground state of a material and apply it as starting point for subsequent property calculations, whenever it is more stable than the spin-paired ground state. In reality, however, the FM state is not guaranteed to comprise the magnetic ground state. In fact, AFM states often have lower energy than the FM one, but in general it is non-trivial to obtain the true magnetic ground state. We have chosen to focus on the FM state due to its simplicity and because its atomic structure and stability are often very similar to those of other magnetic states. Whether or not the FM state is the true magnetic ground state is indicated by the nearest neighbour exchange coupling constant as described below.

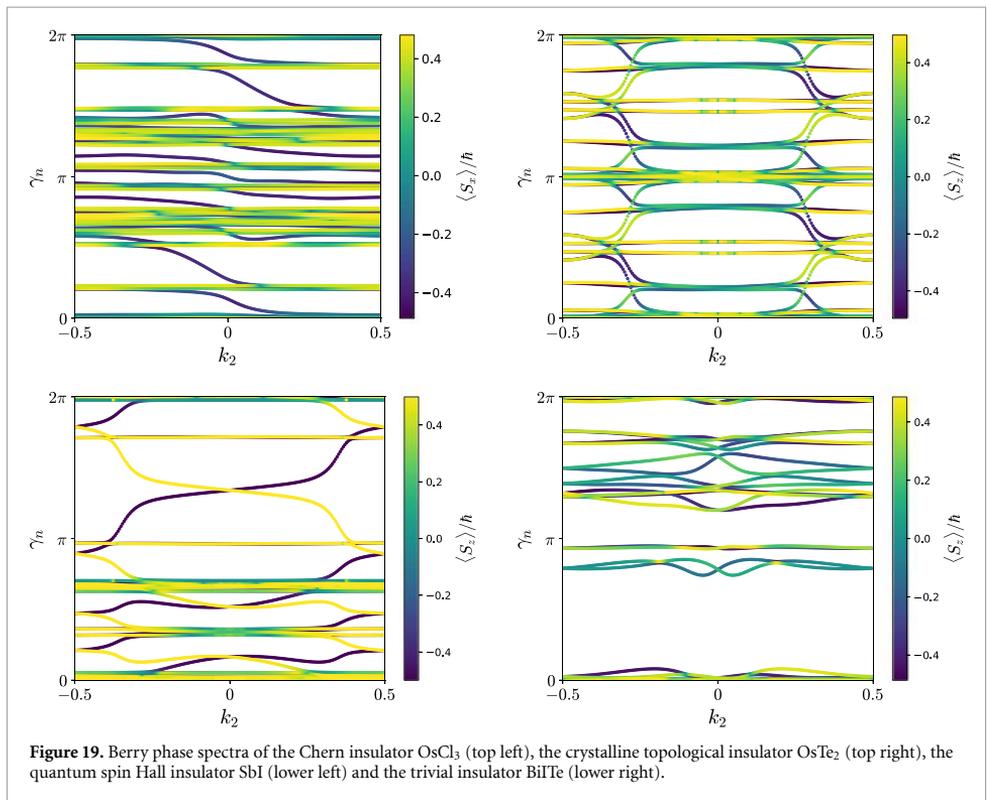


Figure 19. Berry phase spectra of the Chern insulator OsCl₃ (top left), the crystalline topological insulator OsTe₂ (top right), the quantum spin Hall insulator SbI (lower left) and the trivial insulator BiTe (lower right).

When investigating magnetic materials the thermodynamical properties (for example the critical temperatures for ordering) are of crucial interest. In two dimensions the Mermin–Wagner theorem [108] comprises an extreme example of the importance of thermal effects since it implies that magnetic order is only possible at $T = 0$ unless the spin-rotational symmetry is explicitly broken. The thermodynamic properties cannot be accessed directly by DFT. Consequently, magnetic models that capture the crucial features of magnetic interactions must be employed. For insulators, the Heisenberg model has proven highly successful in describing magnetic properties of solids in 3D as well as 2D [109]. It represents the magnetic degrees of freedom as a lattice of localised spins that interact through a set of exchange coupling constants. If the model is restricted to include only nearest neighbour exchange and assume magnetic isotropy in the plane, it reads:

$$H = -\frac{J}{2} \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \frac{\lambda}{2} \sum_{\langle ij \rangle} S_i^x S_j^z - A \sum_i (S_i^z)^2, \quad (12)$$

where J is the nearest neighbour exchange constant, λ is the nearest neighbour anisotropic exchange constant and A measures the strength of single-ion anisotropy. We also neglect off-diagonal exchange coupling constants that give rise to terms proportional to $S_i^x S_j^y$, $S_i^y S_j^z$ and $S_i^z S_j^x$. The out-of-plane direction has

been chosen as z and $\langle ij \rangle$ implies that for each site i we sum over all nearest neighbour sites j . The parameters J , λ and A may be obtained from an energy mapping analysis involving four DFT calculations with different spin configurations [70, 110, 111]. The thermodynamic properties of the resulting ‘first principles Heisenberg model’ may subsequently be analysed with classical Monte Carlo simulations or renormalised spin wave theory [36, 112].

The C2DB provides the values of J , λ , and A as well as the number of nearest neighbours N_{nn} and the maximum eigenvalue of S_z (S), which is obtained from the total magnetic moment per atom in the FM ground state (rounded to nearest half-integer for metals). These key parameters facilitate easy post-processing analysis of thermal effects on the magnetic structure. In [113] such an analysis was applied to estimate the critical temperature of all FM materials in the C2DB based on a model expression for T_C and the parameters from equation (12).

For metals, the Heisenberg parameters available in C2DB should be used with care because the Heisenberg model is not expected to provide an accurate description of magnetic interactions in this case. Nevertheless, even for metals the sign and magnitude of the parameters provide an important qualitative measure of the magnetic interactions that may be used to screen and select materials for more detailed investigations of magnetic properties.

A negative value of J implies the existence of an AFM state with lower energy than the FM state used in C2DB. This parameter is thus crucial to consider when judging the stability and relevance of a material classified as magnetic in C2DB (see section 2.5). Figure 20 shows the distribution of exchange coupling constants (weighted by S^2) of the magnetic materials in the C2DB. The distribution is slightly skewed to the positive side indicating that FM order is more common than AFM order.

The origin of magnetic anisotropy may stem from either single-ion anisotropy or anisotropic exchange and it is in general difficult *a priori* to determine, which mechanism is most important. There is, however, a tendency in the literature to neglect anisotropic exchange terms in a Heisenberg model description of magnetism and focus solely on the single-ion anisotropy. In figure 20 we show a scatter plot of the anisotropy parameters A and λ for the FM materials ($J > 0$). The spread of the parameters indicate that the magnetic anisotropy is in general equally likely to originate from both mechanisms and neglecting anisotropic exchange is not advisable. For ferromagnets, the model (equation (12)) only exhibits magnetic order at finite temperatures if $A(2S - 1) + \lambda N_{mn} > 0$ [113]. Neglecting anisotropic exchange thus excludes materials with $A < 0$ that satisfies $A(2S - 1) + \lambda N_{mn} > 0$. This is in fact the case for 11 FM insulators and 31 FM metals in the C2DB.

5.9. Raman spectrum

Raman spectroscopy is an important technique used to probe the vibrational modes of a solid (or molecule) by means of inelastic scattering of light [114]. In fact, Raman spectroscopy is the dominant method for characterising 2D materials and can yield detailed information about chemical composition, crystal structure and layer thickness. There exist several different types of Raman spectroscopies that differ mainly by the number of photons and phonons involved in the scattering process [114]. The first-order Raman process, in which only a single phonon is involved, is the dominant scattering process in samples with low defect concentrations.

In a recent work, the first-order Raman spectra of 733 monolayer materials from the C2DB were calculated, and used as the basis for an automatic procedure for identifying a 2D material entirely from its experimental Raman spectrum [115]. The Raman spectrum is calculated using third-order perturbation theory to obtain the rate of scattering processes involving creation/annihilation of one phonon and two photons, see reference [115] for details. The light field is written as $\mathcal{F}(t) = \mathcal{F}_{in} \mathbf{u}_{in} \exp(-i\omega_{in}t) + \mathcal{F}_{out} \mathbf{u}_{out} \exp(-i\omega_{out}t) + c.c.$ where $\mathcal{F}_{in/out}$ and $\omega_{in/out}$ denote the amplitudes and frequencies of the input/output electromagnetic fields, respectively. In addition, $\mathbf{u}_{in/out} = \sum_i u_{in/out}^i \mathbf{e}_i$ are the corresponding polarisation vectors, where \mathbf{e}_i denotes the unit

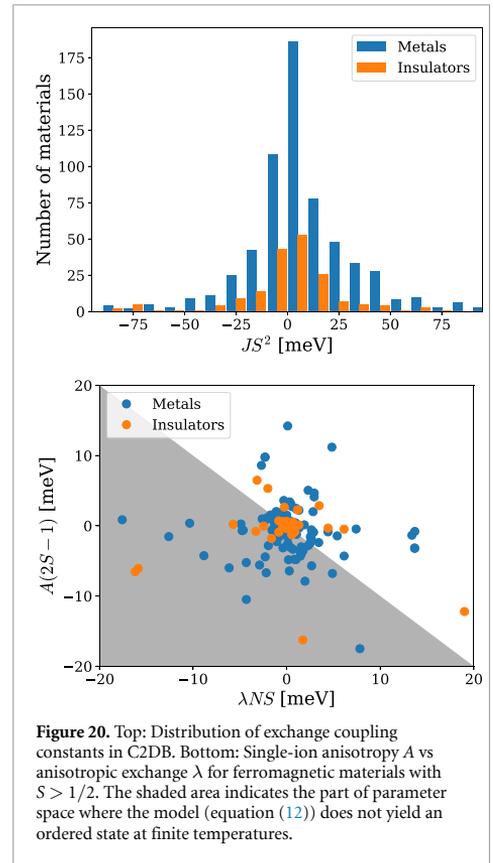
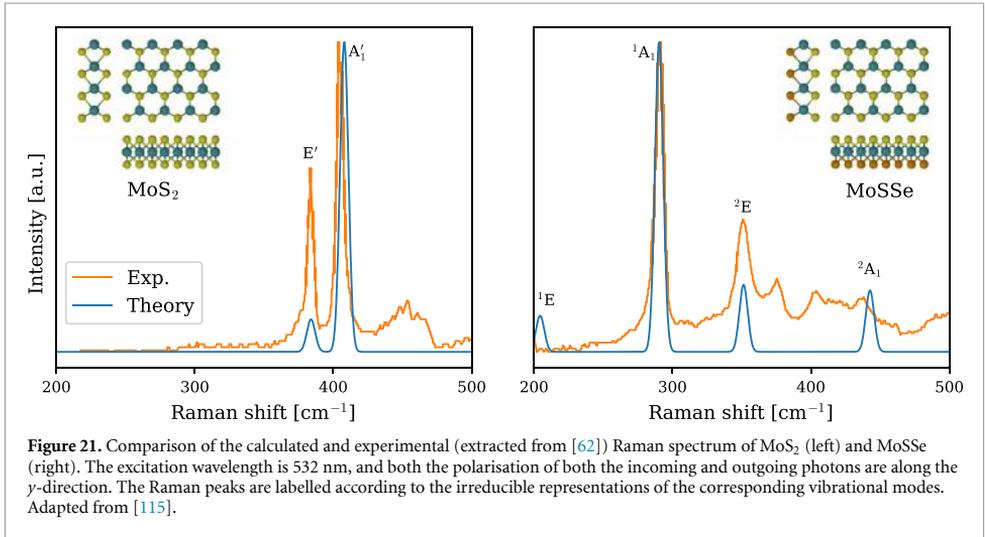


Figure 20. Top: Distribution of exchange coupling constants in C2DB. Bottom: Single-ion anisotropy A vs anisotropic exchange λ for ferromagnetic materials with $S > 1/2$. The shaded area indicates the part of parameter space where the model (equation (12)) does not yield an ordered state at finite temperatures.

vector along the i -direction with $i \in \{x, y, z\}$. Using this light field, the final expression for the Stokes Raman intensity involving scattering events by only one phonon reads [115]:

$$I(\omega) = I_0 \sum_{\nu} \frac{n_{\nu} + 1}{\omega_{\nu}} \left| \sum_{ij} u_{in}^i R_{ij}^{\nu} u_{out}^j \right|^2 \delta(\omega - \omega_{\nu}). \tag{13}$$

Here, I_0 is an unimportant constant (since Raman spectra are always reported normalised), and n_{ν} is obtained from the Bose-Einstein distribution, i.e. $n_{\nu} \equiv (\exp[\hbar\omega_{\nu}/k_B T] - 1)^{-1}$ at temperature T for a Raman mode with energy $\hbar\omega_{\nu}$. Note that only phonons at the Brillouin zone center (with zero momentum) contribute to the one-phonon Raman processes due to momentum conservation. In equation (13), R_{ij}^{ν} is the Raman tensor for phonon mode ν , which involves electron-phonon and dipole matrix elements as well as the electronic transition energies and the incident excitation frequency. Equation (13) has been used to compute the Raman spectra of the 733 most stable, non-magnetic monolayers in C2DB for a range of excitation frequencies and polarisation configurations. Note that the Raman shift $\hbar\omega$ is typically expressed in cm^{-1} with



1 meV equivalent to 8.0655 cm^{-1} . In addition, for generating the Raman spectra, we have used a Gaussian [$G(\omega) = (\sigma\sqrt{2\pi})^{-1} \exp(-\omega^2/2\sigma^2)$] with a variance $\sigma = 3 \text{ cm}^{-1}$ to replace the Dirac delta function, which accounts for the inhomogeneous broadening of phonon modes.

As an example, figure 21 shows the calculated Raman spectrum of monolayer MoS₂ and the Janus monolayer MoSse (see section 4.1). Experimental Raman spectra extracted from reference [62] are shown for comparison. For both materials, good agreement between theory and experiment is observed for the peak positions and relative amplitudes of the main peaks. The small deviations can be presumably be attributed to substrate interactions and defects in the experimental samples as well as the neglect of excitonic effects in the calculations. The qualitative differences between the Raman spectra can be explained by the different point groups of the materials (C_{3v} and D_{3h} , respectively), see reference [115]. In particular, the lower symmetry of MoSse results in a lower degeneracy of its vibrational modes leading to more peaks in the Raman spectrum.

Very recently, the Raman spectra computed from third order perturbation theory as described above, were supplemented by spectra obtained from the more conventional Kramers–Heisenberg–Dirac (KHD) approach. Within the KHD method, the Raman tensor is obtained as the derivative of the static electric polarisability (or equivalently, the susceptibility) along the vibrational normal modes [116, 117]:

$$R_{ij}^{\nu} = \sum_{\alpha l} \frac{\partial \chi_{ij}^{(1)}}{\partial r_{\alpha l}} \frac{v_{\alpha l}^{\nu}}{\sqrt{M_{\alpha}}}. \quad (14)$$

Here, $\chi_{ij}^{(1)}$ is the (first-order) susceptibility tensor, r_{α} and M_{α} are the position and atomic mass of atom

α , respectively, and $v_{\alpha l}^{\nu}$ is the eigenmode of phonon ν . The two approaches, i.e. the KHD and third-order perturbation approach, can be shown to be equivalent [114], at least when local field effects can be ignored as is typically the case for 2D materials [35]. We have also confirmed this equivalence from our calculations. Furthermore, the computational cost of both methods is also similar [115]. However, the KHD approach typically converge faster with respect to both the number of bands and *k*-grid compared to the third-order perturbation method. This stems from the general fact that higher-order perturbation calculations converge slower with respect to *k*-grid and they require additional summations over a complete basis set (virtual states) and hence a larger number of bands [118]. Currently, Raman spectra from both approaches can be found at the C2DB website.

5.10. Second harmonics generation

Nonlinear optical (NLO) phenomena such as harmonic generation, Kerr, and Pockels effects are of great technological importance for lasers, frequency converters, modulators, etc. In addition, NLO spectroscopy has been extensively employed to obtain insight into materials properties [119] that are not accessible by e.g. linear optical spectroscopy. Among numerous nonlinear processes, second-harmonic generation (SHG) has been widely used for generating new frequencies in lasers as well as identifying crystal orientations and symmetries.

Recently, the SHG spectrum was calculated for 375 non-magnetic, non-centrosymmetric semiconducting monolayers of the C2DB, and multiple 2D materials with giant optical nonlinearities were identified [120]. In the SHG process, two incident photons at frequency ω generate an emitted photon at frequency of 2ω . Assume that a mono-harmonic electric

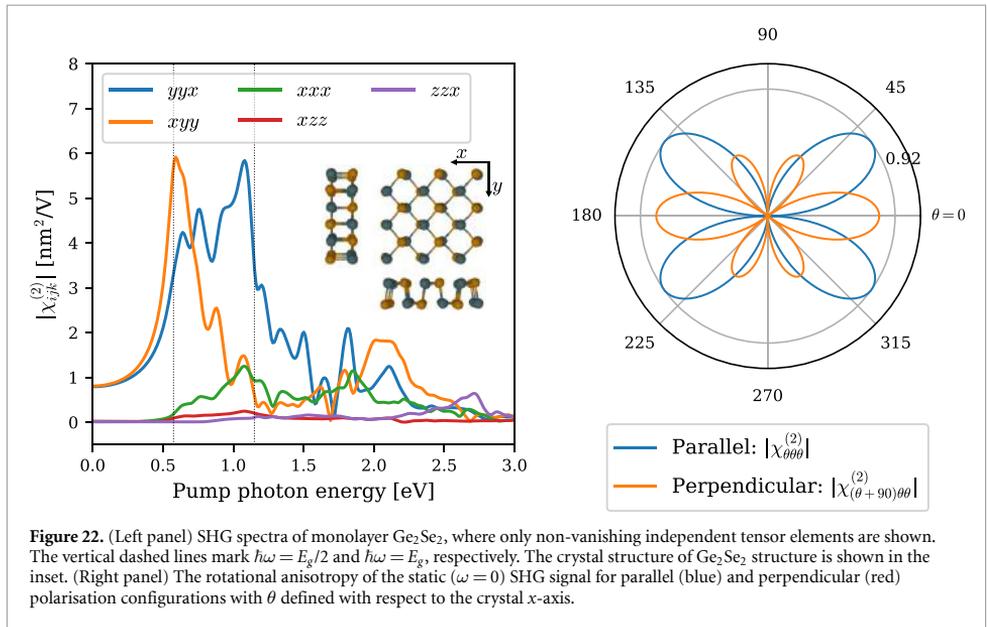


Figure 22. (Left panel) SHG spectra of monolayer Ge_2Se_2 , where only non-vanishing independent tensor elements are shown. The vertical dashed lines mark $\hbar\omega = E_g/2$ and $\hbar\omega = E_g$, respectively. The crystal structure of Ge_2Se_2 structure is shown in the inset. (Right panel) The rotational anisotropy of the static ($\omega = 0$) SHG signal for parallel (blue) and perpendicular (red) polarisation configurations with θ defined with respect to the crystal x -axis.

field written $\mathcal{F}(t) = \sum_i \mathcal{F}_i \mathbf{e}_i e^{-i\omega t} + \text{c.c.}$ is incident on the material, where \mathbf{e}_i denotes the unit vector along direction $i \in \{x, y, z\}$. The electric field induces a SHG polarisation density $\mathbf{P}^{(2)}$, which can be obtained from the quadratic susceptibility tensor $\chi_{ijk}^{(2)}$,

$$P_i^{(2)}(t) = \epsilon_0 \sum_{jk} \chi_{ijk}^{(2)}(\omega, \omega) \mathcal{F}_j \mathcal{F}_k e^{-2i\omega t} + \text{c.c.}, \quad (15)$$

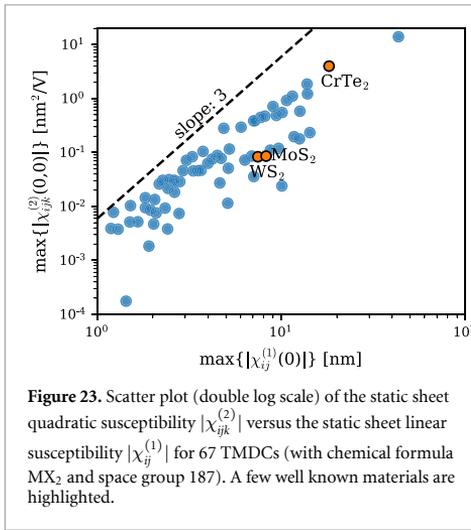
where ϵ_0 denotes the vacuum permittivity. $\chi_{ijk}^{(2)}$ is a symmetric (due to intrinsic permutation symmetry i.e. $\chi_{ijk}^{(2)} = \chi_{jik}^{(2)}$) rank-3 tensor with at most 18 independent elements. Furthermore, similar to the piezoelectric tensor, the point group symmetry reduces the number of independent tensor elements.

In the C2DB, the quadratic susceptibility is calculated using density matrices and perturbation theory [118, 121] with the involved transition dipole matrix elements and band energies obtained from DFT. The use of DFT single-particle orbitals implies that excitonic effects are not accounted for. The number of empty bands included in the sum over bands was set to three times the number of occupied bands. The width of the Fermi–Dirac occupation factor was set to $k_B T = 50$ meV, and a line-shape broadening of $\eta = 50$ meV was used in all spectra. Furthermore, time-reversal symmetry was imposed in order to reduce the \mathbf{k} -integrals to half the BZ. For various 2D crystal classes, it was verified by explicit calculation that the quadratic tensor elements fulfil the expected symmetries, e.g. that they all vanish identically for centrosymmetric crystals.

As an example, the calculated SHG spectra for monolayer Ge_2Se_2 is shown in figure 22 (left panel).

Monolayer Ge_2Se_2 has five independent tensor elements, $\chi_{xxx}^{(2)}$, $\chi_{xyy}^{(2)}$, $\chi_{xzz}^{(2)}$, $\chi_{yyx}^{(2)} = \chi_{yyx}^{(2)}$, and $\chi_{zzx}^{(2)} = \chi_{zzx}^{(2)}$, since it is a group-IV dichalcogenide with an orthorhombic crystal structure (space group 31 and point group C_{2v}). Note that, similar to the linear susceptibility, the bulk quadratic susceptibility (with SI units of mV^{-1}) is ill-defined for 2D materials (since the volume is ambiguous) [120]. Instead, the unambiguous *sheet* quadratic susceptibility (with SI units of $\text{m}^2 \text{V}^{-1}$) is evaluated. In addition to the frequency-dependent SHG spectrum, the angular dependence of the static ($\omega = 0$) SHG intensity at normal incidence for parallel and perpendicular polarisations (relative to the incident electric field) is calculated, see figure 22 (right panel). Such angular resolved SHG spectroscopy has been widely used for determining the crystal orientation of 2D materials. The calculated SHG spectra for all non-vanishing inequivalent polarisation configurations and their angular dependence, are available in the C2DB.

Since C2DB has already gathered various material properties of numerous 2D materials, it provides a unique opportunity to investigate interrelations between different material properties. For example, the strong dependence of the quadratic optical response on the electronic band gap was demonstrated on basis of the C2DB data [120]. As another example of a useful correlation, the static quadratic susceptibility is plotted versus the static linear susceptibility for 67 TMDCs (with formula MX_2 , space group 187) in figure 23. Note that for materials with several independent tensor elements, only the largest is shown. There is a very clear correlation between the two quantities. This is not unexpected as both



the linear and quadratic optical responses are functions of the transition dipole moments and transition energies. More interestingly, the strength of the quadratic response seems to a very good approximation to be given by a universal constant times the linear susceptibility to the power of three (ignoring polarisation indices), i.e.

$$\chi^{(2)}(0,0) \approx A\chi^{(1)}(0)^3, \quad (16)$$

where A is only weakly material dependent. Note that this scaling law is also known in classical optics as semi-empirical Miller's rule for non-resonant quadratic responses [122], which states that the second order electric susceptibility is proportional to the product of the first-order susceptibilities at the three frequencies involved.

6. Machine learning properties

In recent years, material scientists have shown great interest in exploiting the use of machine learning (ML) techniques for predicting materials properties and guiding the search for new materials. ML is the scientific study of algorithms and statistical models that computer systems can use to perform a specific task without using explicit instructions but instead relying on patterns and inference. Within the domain of materials science, one of the most frequent problems is the mapping from atomic configuration to material property, which can be used e.g. to screen large material spaces in search of optimal candidates for specific applications [123, 124].

In the ML literature, the mathematical representation of the input observations is often referred to as a fingerprint. Any fingerprint must satisfy a number of general requirements [125]. In particular, a fingerprint must be:

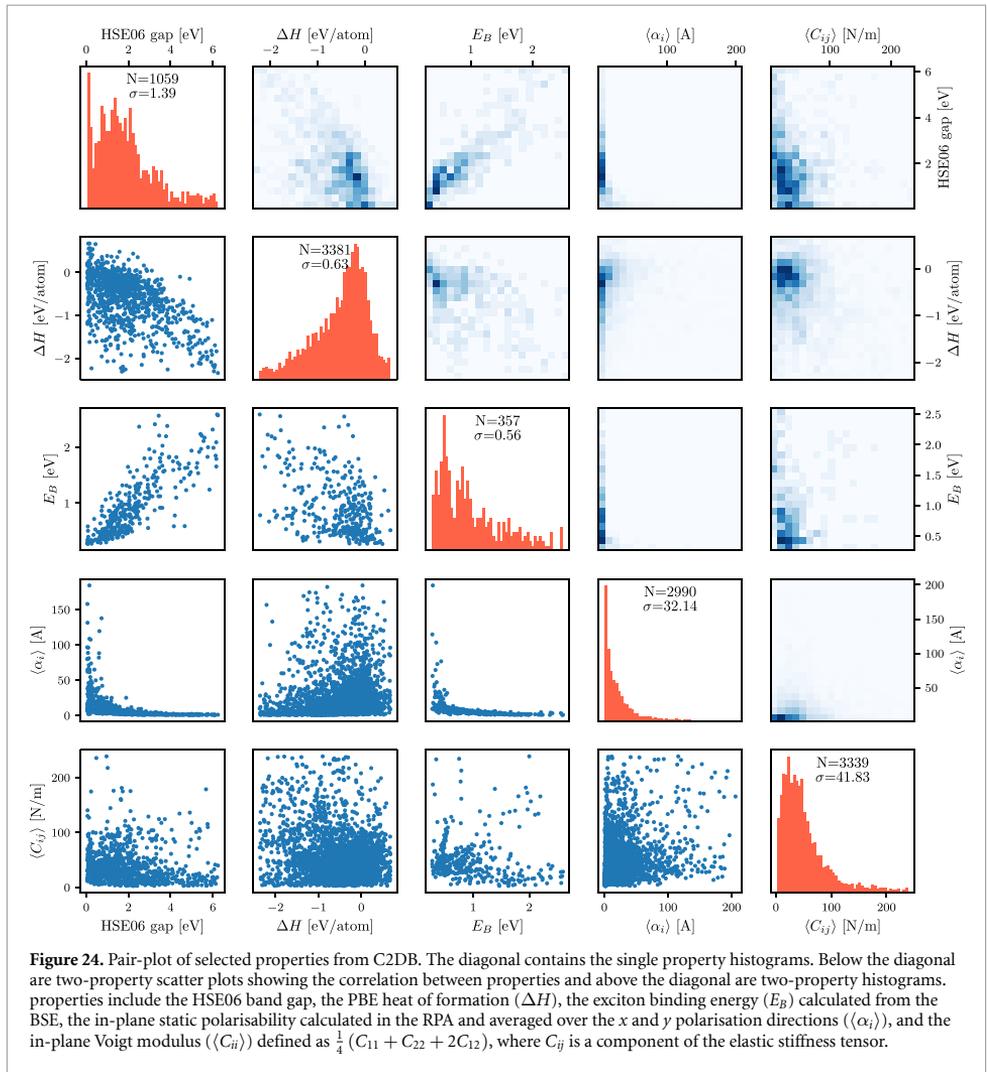
- Complete:* The fingerprint should incorporate all the relevant input for the underlying problem, i.e. materials with different properties should have different fingerprints.
- Compact:* The fingerprint should contain no or a minimal number of features redundant to the underlying problem. This includes being invariant to rotations, translations and other transformations that leave the properties of the system invariant.
- Descriptive:* Materials with similar target values should have similar fingerprints.
- Simple:* The fingerprint should be efficient to evaluate. In the present context, this means that calculating the fingerprint should be significantly faster than calculating the target property.

Several types of atomic-level materials fingerprints have been proposed in the literature, including general purpose fingerprints based on atomistic properties [126, 127] possibly encoding information about the atomic structure, i.e. atomic positions [125, 128, 129], and specialised fingerprints tailored for specific applications (materials/properties) [130, 131].

The aim of this section is to demonstrate how the C2DB may be utilised for ML-based prediction of general materials properties. Moreover, the study serves to illustrate the important role of the fingerprint for such problems. The 2D materials are represented using three different fingerprints: two popular structural fingerprints and a more advanced fingerprint that encodes information about the electronic structure via the PDOS. The target properties include the HSE06 band gap, the PBE heat of formation (ΔH), the exciton binding energy (E_B) obtained from the many-body BSE, the in-plane static polarisability calculated in the RPA averaged over the x and y polarisation directions ($\langle\langle\alpha_i\rangle\rangle$), and the in-plane Voigt modulus ($\langle\langle C_{ii}\rangle\rangle$) defined as $\frac{1}{4}(C_{11} + C_{22} + 2C_{12})$, where C_{ij} is a component of the elastic stiffness tensor in Mandel notation.

To introduce the data, figure 24 shows pair-plots of the dual-property relations of these properties. The plots in the diagonal show the single-property histograms, whereas the off-diagonals show dual-property scatter plots below the diagonal and histograms above the diagonal. Clearly, there are only weak correlations between most of the properties, with the largest degree of correlation observed between the HSE06 band gap and exciton binding energy. The lack of strong correlations motivates the use of ML for predicting the properties.

The prediction models are built using the Ewald sum matrix and many-body tensor representation (MBTR) as structural fingerprints. The Ewald fingerprint is a version of the simple Coulomb matrix fingerprint [128] modified to periodic systems [125]. The MBTR encodes first, second and third order



terms like atomic numbers, distances and angles between atoms in the system [129]. As an alternative to the structural fingerprints, a representation based on the PBE PDOS is also tested. This fingerprint⁶ encodes the coupling between the PDOS at different atomic orbitals in both energy and real space. It is defined as:

$$\rho_{\nu\nu'}(E, R) = \sum_{a \in \text{cell}} \sum_{a'} \rho_{a\nu}(E) \rho_{a'\nu'}(E) G \times (R - |R_a - R_{a'}|), \quad (17)$$

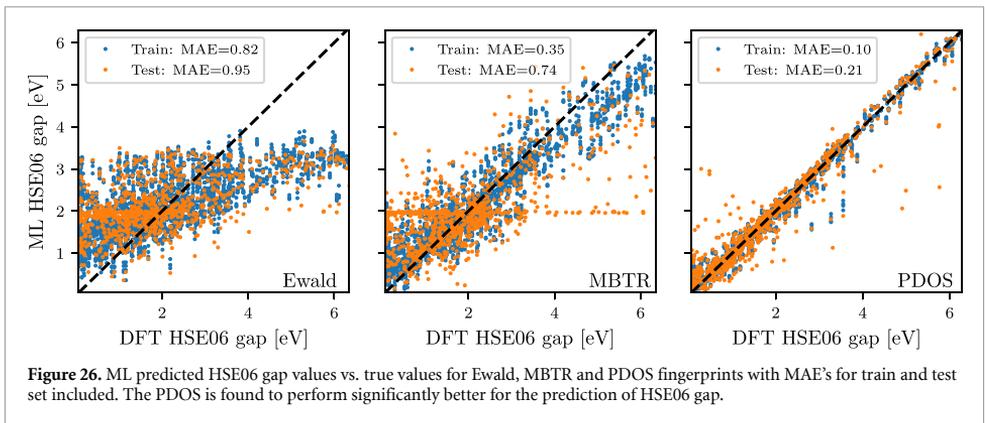
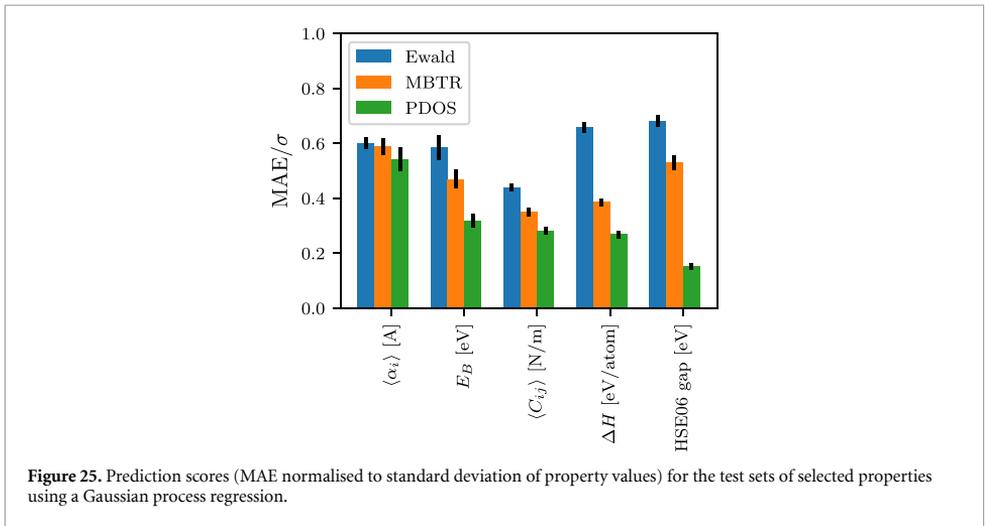
where G is a Gaussian smearing function, a denotes the atoms, ν denotes atomic orbitals, and the PDOS is given by:

$$\rho_{a\nu}(E) = \sum_n |\langle \psi_n | a\nu \rangle|^2 G(E - \epsilon_n), \quad (18)$$

⁶ Details will be published elsewhere.

where n runs over all eigenstates of the system. Since this fingerprint requires a DFT-PBE calculation to be performed, additional features derivable from the DFT calculation can be added to the fingerprint. In this study, the PDOS fingerprint is amended by the PBE band gap. The latter can in principle be extracted from the PDOS, but its explicit inclusion has been found to improve the performance of the model.

A Gaussian process regression using a simple Gaussian kernel with a noise component is used as learning algorithm. The models are trained using 5-fold cross validation on a training set consisting of 80% of the materials with the remaining 20% held aside as test data. Prior to training the model, the input space is reduced to 50 features using principal component analysis (PCA). This step is necessary to reduce the huge number of features in the MBTR fingerprint to a manageable size. Although this is not required for the Ewald and PDOS fingerprints,



we perform the same feature reduction in all cases. The optimal number of features depends on the choice of fingerprint, target property and learning algorithm, but for consistency 50 PCA components are used for all fingerprints and properties in this study.

Figure 25 shows the prediction scores obtained for the five properties using the three different fingerprints. The employed prediction score is the mean absolute error of the test set normalised by the standard deviation of the property values (standard deviations are annotated in the diagonal plots in figure 24). In general, the PDOS fingerprint outperforms the structural fingerprints. The difference between prediction scores is smallest for the static polarisability $\langle\alpha_i\rangle$ and largest for the HSE06 gap. It should be stressed that although the evaluation of the PBE-PDOS fingerprint is significantly more time consuming than the evaluation of the structural fingerprints, it is still much faster than the evaluation of all the target properties. Moreover, structural fingerprints require the atomic structure, which in turns

requires a DFT structure optimisation (unless the structure is available by other means).

The HSE06 band gap shows the largest sensitivity to the employed fingerprint. To elaborate on the HSE06 results, figure 26 shows the band gap predicted using each of the three different fingerprints plotted against the true band gap. The mean absolute errors on the test set is 0.95 and 0.74 eV for Ewald and MBTR fingerprints, respectively, while the PDOS significantly outperforms the other fingerprints with a test MAE of only 0.21 eV. This improvement in prediction accuracy is partly due to the presence of the PBE gap in the PDOS fingerprint. However, our analysis shows that the pure PDOS fingerprint without the PBE gap still outperforms the structural fingerprints. Using only the PBE gap as feature results in a test MAE of 0.28 eV.

The current results show that the precision of ML-based predictions are highly dependent on the type of target property and the chosen material representation. For some properties, the mapping between atomic structure and property is easier to learn while

others might require more/deeper information, e.g. in terms of electronic structure fingerprints. Our results clearly demonstrate the potential of encoding electronic structure information into the material fingerprint, and we anticipate more work on this relevant and exciting topic in the future.

7. Summary and outlook

We have documented a number of extensions and improvements of the C2DB made in the period 2018–2020. The new developments include: (1) A refined and more stringent workflow for filtering prospective 2D materials and classifying them according to their crystal structure, magnetic state and stability. (2) Improvements of the methodology used to compute certain challenging properties such as the full stiffness tensor, effective masses, G_0W_0 band structures, and optical absorption spectra. (3) New materials including 216 MX₂ Janus monolayers and 574 monolayers exfoliated from experimentally known bulk crystals. In addition, ongoing efforts to systematically obtain and characterise bilayers in all possible stacking configurations as well as point defects in the semiconducting monolayers, have been described. (4) New properties including exfoliation energies, spontaneous polarisations, Bader charges, piezoelectric tensors, IR polarisabilities, topological invariants, magnetic exchange couplings, Raman spectra, and SHG spectra. It should be stressed that the C2DB will continue to grow as new structures and properties are being added, and thus the present paper should not be seen as a final report on the C2DB but rather a snapshot of its current state.

In addition to the above mentioned improvements relating to data quantity and quality, the C2DB has been endowed with a comprehensive documentation layer. In particular, all data presented on the C2DB website are now accompanied by an information field that explains the meaning and representation (if applicable) of the data and details how it was calculated thus making the data easier to understand, reproduce, and deploy.

The C2DB has been produced using the ASR in combination with the GPAW electronic structure code and the MyQueue task and workflow scheduling system. The ASR is a newly developed Python-based framework designed for high-throughput materials computations. The highly flexible and modular nature of the ASR and its strong coupling to the well established community-driven ASE project, makes it a versatile framework for both high- and low-throughput materials simulation projects. The ASR and the C2DB-ASR workflow are distributed as open source code. A detailed documentation of the ASR will be published elsewhere.

While the C2DB itself is solely concerned with the properties of perfect monolayer crystals, ongoing efforts focus on the systematic characterisation

of homo-bilayer structures as well as point defects in monolayers. The data resulting from these and other similar projects will be published as separate, independent databases, but will be directly interlinked with the C2DB making it possible to switch between them in a completely seamless fashion. These developments will significantly broaden the scope and usability of the C2DB+ (+ stands for associated databases) that will help theoreticians and experimentalists to navigate one of the most vibrant and rapidly expanding research fields at the crossroads of condensed matter physics, photonics, nanotechnology, and chemistry.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.11583/DTU.14616660>.

Acknowledgments

The Center for Nanostructured Graphene (CNG) is sponsored by the Danish National Research Foundation, Project DNRF103. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program Grant Agreement No. 773122 (LIMA) and Grant Agreement No. 951786 (NOMAD CoE). T D acknowledges financial support from the German Research Foundation (DFG Projects No. DE 2749/2-1).

ORCID iDs

Morten Niklas Gjerding  <https://orcid.org/0000-0002-5256-660X>

Alireza Taghizadeh  <https://orcid.org/0000-0003-0876-9538>

Asbjørn Rasmussen  <https://orcid.org/0000-0001-7110-9255>

Sajid Ali  <https://orcid.org/0000-0001-7865-2664>

Fabian Bertoldo  <https://orcid.org/0000-0002-1219-8689>

Thorsten Deilmann  <https://orcid.org/0000-0003-4165-2446>

Nikolaj Rørbæk Knøsgaard  <https://orcid.org/0000-0003-3709-5464>

Mads Kruse  <https://orcid.org/0000-0002-0599-5110>

Ask Hjorth Larsen  <https://orcid.org/0000-0001-5267-6852>

Simone Manti  <https://orcid.org/0000-0003-3770-0863>

Thomas Garm Pedersen  <https://orcid.org/0000-0002-9466-6190>

Urko Petralanda  <https://orcid.org/0000-0003-0226-0028>

Thorbjørn Skovhus  <https://orcid.org/0000-0001-5215-6419>

Mark Kamper Svendsen  <https://orcid.org/0000-0001-9718-849X>

Jens Jørgen Mortensen  <https://orcid.org/0000-0001-5090-6706>

Thomas Olsen  <https://orcid.org/0000-0001-6256-9284>

Kristian Sommer Thygesen 

<https://orcid.org/0000-0001-5197-214X>

References

- [1] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
- [2] Schwierz F 2010 *Nat. Nanotechnol.* **5** 487
- [3] Novoselov K, Mishchenko A, Carvalho A and Castro Neto A 2016 *Science* **353** 6298
- [4] Ferrari A C et al 2015 *Nanoscale* **7** 4598–810
- [5] Bhimanapati G R et al 2015 *ACS Nano* **9** 11509–39
- [6] Hastrup S et al 2018 *2D Mater.* **5** 042002
- [7] Shivayogimath A et al 2019 *Nat. Commun.* **10** 1–7
- [8] Zhou J et al 2018 *Nature* **556** 355–9
- [9] Anasori B, Lukatskaya M R and Gogotsi Y 2017 *Nat. Rev. Mater.* **2** 1–17
- [10] Dou L et al 2015 *Science* **349** 1518–21
- [11] Mounet N et al 2018 *Nat. Nanotechnol.* **13** 246–52
- [12] Ashton M, Paul J, Sinnott S B and Hennig R G 2017 *Phys. Rev. Lett.* **118** 106101
- [13] Geim A K and Grigorieva I V 2013 *Nature* **499** 419–25
- [14] Cao Y, Fatemi V, Fang S, Watanabe K, Taniguchi T, Kaxiras E and Jarillo-Herrero P 2018 *Nature* **556** 43–50
- [15] Biströtzer R and MacDonald A H 2011 *Proc. Natl Acad. Sci.* **108** 12233–7
- [16] Zhao X et al 2020 *Nature* **581** 171–7
- [17] Wan J, Lacey S D, Dai J, Bao W, Fuhrer M S and Hu L 2016 *Chem. Soc. Rev.* **45** 6742–65
- [18] Wilkinson M D et al 2016 *Sci. Data* **3** 1–9
- [19] Wirtz L, Marini A and Rubio A 2006 *Phys. Rev. Lett.* **96** 126104
- [20] Cudazzo P, Tokatly I V and Rubio A 2011 *Phys. Rev. B* **84** 085406
- [21] Klots A et al 2014 *Sci. Rep.* **4** 6608
- [22] 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
- [23] Olsen T, Latini S, Rasmussen F and Thygesen K S 2016 *Phys. Rev. Lett.* **116** 056401
- [24] Riis-Jensen A C, Gjerding M N, Russo S and Thygesen K S 2020 *Phys. Rev. B* **102** 201402
- [25] Felipe H, Xian L, Rubio A and Louie S G 2020 *Nat. Commun.* **11** 1–10
- [26] Sohler T, Gibertini M, Calandra M, Mauri F and Marzari N 2017 *Nano Lett.* **17** 3758–63
- [27] Ugeda M M et al 2014 *Nat. Mater.* **13** 1091–5
- [28] Winther K T and Thygesen K S 2017 *2D Mater.* **4** 025059
- [29] Wang Z, Rhodes D A, Watanabe K, Taniguchi T, Hone J C, Shan J and Mak K F 2019 *Nature* **574** 76–80
- [30] Gong C et al 2017 *Nature* **546** 265–9
- [31] Huang B et al 2017 *Nature* **546** 270–3
- [32] Chang K et al 2016 *Science* **353** 274–8
- [33] Olsen T, Andersen E, Okugawa T, Torelli D, Deilmann T and Thygesen K S 2019 *Phys. Rev. Mater.* **3** 024005
- [34] Marrazzo A, Gibertini M, Campi D, Mounet N and Marzari N 2019 *Nano Lett.* **19** 8431–40
- [35] Thygesen K S 2017 *2D Mater.* **4** 022004
- [36] Torelli D and Olsen T 2018 *2D Mater.* **6** 015028
- [37] Rasmussen F A and Thygesen K S 2015 *J. Phys. Chem. C* **119** 13169–83
- [38] Enkovaara J et al 2010 *J. Phys.: Condens. Matter.* **22** 253202
- [39] Mortensen J J, Hansen L B and Jacobsen K W 2005 *Phys. Rev. B* **71** 035109
- [40] Gjerding M, Skovhus T, Rasmussen A, Bertoldo F, Larsen A H, Mortensen J J and Thygesen K S 2021 Atomic simulation recipes—a python framework and library for automated workflows (arXiv:2104.13431)
- [41] Larsen A H et al 2017 *J. Phys.: Condens. Matter.* **29** 273002
- [42] Mortensen J J, Gjerding M and Thygesen K S 2020 *J. Open Source Softw.* **5** 1844
- [43] Saal J E, Kirklín S, Aykol M, Meredig B and Wolverton C 2013 *JOM* **65** 1501–9
- [44] Jain A et al 2013 *APL Mater.* **1** 011002
- [45] Curtarolo S et al 2012 *Comput. Mater. Sci.* **58** 218–26
- [46] Ataca C, Sahin H and Ciraci S 2012 *J. Phys. Chem. C* **116** 8983–99
- [47] Lebegue S, Björkman T, Klintonberg M, Nieminen R M and Eriksson O 2013 *Phys. Rev. X* **3** 031002
- [48] Kormányos A, Burkard G, Gmitra M, Fabian J, Zólyomi V, Drummond N D and Fal'ko V 2015 *2D Mater.* **2** 022001
- [49] Zhou J et al 2019 *Sci. Data* **6** 1–10
- [50] Choudhary K, Kalish I, Beams R and Tavazza F 2017 *Sci. Rep.* **7** 1–16
- [51] Larsen A H, Vanin M, Mortensen J J, Thygesen K S and Jacobsen K W 2009 *Phys. Rev. B* **80** 195112
- [52] Larsen P M, Pandey M, Strange M and Jacobsen K W 2019 *Phys. Rev. Mater.* **3** 034003
- [53] Ong S P et al 2013 *Comput. Mater. Sci.* **68** 314–19
- [54] Patrick C E, Jacobsen K W and Thygesen K S 2015 *Phys. Rev. B* **92** 201205
- [55] Kirklín S, Saal J E, Meredig B, Thompson A, Doak J W, Aykol M, Rühl S and Wolverton C 2015 *npj Computat. Mater.* **1** 1–15
- [56] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865–8
- [57] Maździarz M 2019 *2D Mater.* **6** 048001
- [58] Li Y and Heinz T F 2018 *2D Mater.* **5** 025021
- [59] Hadley L N and Dennison D 1947 *J. Opt. Soc. Am.* **37** 451–65
- [60] Wang G, Chernikov A, Glazov M M, Heinz T F, Marie X, Amand T and Urbaszek B 2018 *Rev. Mod. Phys.* **90** 021001
- [61] Lu A Y et al 2017 *Nat. Nanotechnol.* **12** 744–9
- [62] Zhang J et al 2017 *ACS Nano* **11** 8192–8
- [63] Fülöp B et al 2018 *2D Mater.* **5** 031013
- [64] Riis-Jensen A C, Deilmann T, Olsen T and Thygesen K S 2019 *ACS Nano* **13** 13354
- [65] Bychkov Y A and Rashba E I 1984 *J. Phys. C: Solid State Phys.* **17** 6039
- [66] Petersen L and Hedegård P 2000 *Surf. Sci.* **459** 49–56
- [67] Bergerhoff G, Brown I and Allen F et al 1987 *Int. Union Crystallogr., Chester* **360** 77–95
- [68] Gražulis S et al 2012 *Nucleic Acids Res.* **40** D420–7
- [69] Qian X, Liu J, Fu L and Li J 2014 *Science* **346** 1344–7
- [70] Torelli D, Moustafa H, Jacobsen K W and Olsen T 2020 *npj Comput. Mater.* **6** 158
- [71] Mak K F, Lee C, Hone J, Shan J and Heinz T F 2010 *Phys. Rev. Lett.* **105** 136805
- [72] Splendiani A, Sun L, Zhang Y, Li T, Kim J, Chim C Y, Galli G and Wang F 2010 *Nano Lett.* **10** 1271–5
- [73] Xiao J et al 2020 *Nat. Phys.* **16** 1028–34
- [74] Sivasdas N, Okamoto S, Xu X, Fennie C J and Xiao D 2018 *Nano Lett.* **18** 7658–64
- [75] Liu Y, Wu L, Tong X, Li J, Tao J, Zhu Y and Petrovic C 2019 *Sci. Rep.* **9** 1–8
- [76] Yasuda K, Wang X, Watanabe K, Taniguchi T and Jarillo-Herrero P 2020 (arXiv:2010.06600)
- [77] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
- [78] Northrup T and Blatt R 2014 *Nat. Photon.* **8** 356–63
- [79] O'Brien J, Furusawa A and Vuckovic J 2009 *Photonics quantum technologies nat Photonics* **3** 687
- [80] Zhang S and Northrup J E 1991 *Phys. Rev. Lett.* **67** 2339

- [81] Van de Walle C G, Laks D, Neumark G and Pantelides S 1993 *Phys. Rev. B* **47** 9425
- [82] Janak J F 1978 *Phys. Rev. B* **18** 7165
- [83] Pandey M, Rasmussen F A, Kuhar K, Olsen T, Jacobsen K W and Thygesen K S 2016 *Nano Lett.* **16** 2234–9
- [84] Kaappa S, Malola S and Häkkinen H 2018 *J. Phys. Chem. A* **122** 8576–84
- [85] Levi G, Ivanov A V and Jonsson H 2020 *Faraday Discuss.* **224** 448–66
- [86] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
- [87] Bader R F W 1990 *Atoms in Molecules: A Quantum Theory (The Int. Series of Monographs on Chemistry vol 22)* (Oxford: Clarendon)
- [88] Tang W, Sanville E and Henkelman G 2009 *J. Phys.: Condens. Matter.* **21** 084204
- [89] Resta R 1992 *Ferroelectrics* **136** 51–5
- [90] King-Smith R D and Vanderbilt D 1993 *Phys. Rev. B* **47** 3
- [91] Zhang S and Yu F 2011 *J. Am. Ceram. Soc.* **94** 3153–70
- [92] Maeder M D, Damjanovic D and Setter N 2004 *J. Electroceram.* **13** 385–92
- [93] Scott J F 2000 *Ferroelectric Memories* vol 3 (Berlin: Springer)
- [94] Rangel T, Fregoso B M, Mendoza B S, Morimoto T, Moore J E and Neaton J B 2017 *Phys. Rev. Lett.* **119** 067402
- [95] Resta R and Vanderbilt D 2007 *Theory of Polarization: A Modern Approach Phys. Ferroelectr.* vol 105 (Berlin: Springer) pp 31–68
- [96] Gjerding M N, Cavalcante L S R, Chaves A and Thygesen K S 2020 *J. Phys. Chem. C* **124** 11609–16
- [97] Ye Z G 2008 *Handbook of Advanced Dielectric, Piezoelectric and Ferroelectric Materials: Synthesis, Properties and Applications* (Amsterdam: Elsevier)
- [98] Ogawa T 2016 *Piezoelectric Materials* (Croatia: InTech)
- [99] Vanderbilt D 1999 *J. Phys. Chem. Solids* **61** 147–51
- [100] Authier A 2003 *Int. Tables for Crystallography: Volume D: Physical Properties of Crystals* (Dordrecht: Springer)
- [101] Duerloo K A N, Ong M T and Reed E J 2012 *J. Phys. Chem. Lett.* **3** 2871–6
- [102] Zhu H et al 2015 *Nat. Nanotechnol.* **10** 151–5
- [103] Taherinejad M, Garrity K F and Vanderbilt D 2014 *Phys. Rev. B* **89** 115102
- [104] Olsen T 2016 *Phys. Rev. B* **94** 235106
- [105] Fu L 2011 *Phys. Rev. Lett.* **106** 106802
- [106] Olsen T, Andersen E, Okugawa T, Torelli D, Deilmann T and Thygesen K S 2019 *Phys. Rev. Mater.* **3** 024005
- [107] Benalcazar W A, Bernevig B A and Hughes T L 2017 *Phys. Rev. B* **96** 245115
- [108] Mermin N D and Wagner H 1966 *Phys. Rev. Lett.* **17** 1133–6
- [109] Olsen T 2019 *MRS Commun.* **9** 1142–50
- [110] Olsen T 2017 *Phys. Rev. B* **96** 125143
- [111] Torelli D and Olsen T 2020 *J. Phys.: Condens. Matter.* **32** 335802
- [112] Lado J L and Fernández-Rossier J 2017 *2D Mater.* **4** 035002
- [113] Torelli D, Thygesen K S and Olsen T 2019 *2D Mater.* **6** 045018
- [114] Long D A 2002 *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules* (Chichester: Wiley)
- [115] Taghizadeh A, Leffers U, Pedersen T G and Thygesen K S 2020 *Nat. Commun.* **11** 3011
- [116] Lee S Y and Heller E J 1979 *J. Chem. Phys.* **71** 4777
- [117] Umari P and Pasquarello A 2003 *J. Phys. Condens. Matter* **15** S1547–52
- [118] Taghizadeh A, Hipolito F and Pedersen T G 2017 *Phys. Rev. B* **96** 195413
- [119] Prylepa A et al 2018 *J. Phys. D: Appl. Phys* **51** 043001
- [120] Taghizadeh A, Thygesen K S and Pedersen T G 2021 *ACS Nano* **15** 7155
- [121] Aversa C and Sipe J E 1995 *Phys. Rev. B* **52** 14636–45
- [122] Miller R C 1964 *Appl. Phys. Lett.* **5** 17–19
- [123] Schmidt J, Marques M R G, Botti S and Marques M A L 2019 *Computat. Mater.* **5** 83
- [124] Zhuo Y, Mansouri Tehrani A and Brgoch J 2018 *J. Phys. Chem. Lett.* **9** 1668–73
- [125] Faber F, Lindmaa A, von Lilienfeld O A and Armiento R 2015 *Int. J. Quantum Chem.* **115** 1094–101
- [126] Ward L, Agrawal A, Choudhary A and Wolverton C 2016 *Computat. Mater.* **2** 1–7
- [127] Ghiringhelli L M, Vybiral J, Levchenko S V, Draxl C and Scheffler M 2015 *Phys. Rev. Lett.* **114** 105503
- [128] Rupp M, Tkatchenko A, Müller K R and Von Lilienfeld A 2012 *Phys. Rev. Lett.* **108** 058301
- [129] Huo H and Rupp M 2018 Unified representation of molecules and crystals for machine learning (arXiv:1704.06439)
- [130] Jorgensen P B, Mesta M, Shil S, García Lastra J M, Jacobsen K W, Thygesen K S and Schmidt M N 2018 *J. Chem. Phys.* **148** 241735
- [131] Rajan A C, Mishra A, Satsangi S, Vaish R, Mizuseki H, Lee K R and Singh A K 2018 *Chem. Mater.* **30** 4031–8

11.8 Paper [VIII]: Tunable free-electron X-ray radiation from van der Waals materials,

Michael Shentcis, Adam K Budniak, Xihang Shi, Raphael Dahan, Yaniv Kurman, Michael Kalina, Hanan Herzig Sheinfux, Mark Blei, Mark Kamper Svendsen, Yaron Amouyal, Sefaattin Tongay, Kristian Sommer Thygesen, Frank HL Koppens, Efrat Lifshitz, F Javier García de Abajo, Liang Jie Wong, Ido Kaminer, “Tunable free-electron X-ray radiation from van der Waals materials,” *Nat. Photonics* 14, 686–692 (2020).



Tunable free-electron X-ray radiation from van der Waals materials

Michael Shentcic¹✉, Adam K. Budniak², Xihang Shi¹, Raphael Dahan¹, Yaniv Kurman¹, Michael Kalina³, Hanan Herzig Sheinfux⁴, Mark Blei⁵, Mark Kamper Svendsen⁶, Yaron Amouyal³, Sefaattin Tongay⁵, Kristian Sommer Thygesen⁶, Frank H. L. Koppens^{1,4,7}, Efrat Lifshitz², F. Javier García de Abajo^{1,4,7}, Liang Jie Wong^{1,8} and Ido Kaminer¹✉

Tunable sources of X-ray radiation are widely used for imaging and spectroscopy in fundamental science, medicine and industry. The growing demand for highly tunable, high-brightness laboratory-scale X-ray sources motivates research into new fundamental mechanisms of X-ray generation. Here, we demonstrate the ability of van der Waals materials to serve as a platform for tunable X-ray generation when irradiated by moderately relativistic electrons available, for example, from a transmission electron microscope. The radiation spectrum can be precisely controlled by tuning the acceleration voltage of the incident electrons, as well as by our proposed approach: adjusting the lattice structure of the van der Waals material. We present experimental results for both methods, observing the energy tunability of X-ray radiation from the van der Waals materials WSe₂, CrPS₄, MnPS₃, FePS₃, CoPS₃ and NiPS₃. Our findings demonstrate the concept of material design at the atomic level, using van der Waals heterostructures and other atomic superlattices, for exploring novel phenomena of X-ray physics.

The wealth of unique properties of van der Waals (vdW) materials in either bulk or single-atomic-layer form have constituted the basis of many novel physical phenomena and fundamental advances in recent years^{1–3}. For example, graphene, an atom-thick layer of graphite⁴, exhibits ultra-high carrier mobility at room temperature⁵, excellent optical transparency⁶, high Young's modulus⁷, high thermal conductivity⁸ and many other properties of practical utility. Moreover, this material has enabled the observation of novel phenomena such as the room-temperature quantum Hall effect⁹. Transition metal dichalcogenides (TMDs)^{10,11} and transition metal thiophosphates¹² have also emerged as intriguing families of vdW materials. Semiconductors by nature, TMDs can exhibit both an indirect bandgap in the bulk and a direct bandgap as a single atomic layer¹⁰. This unique property, combined with weak dielectric screening in two dimensions¹³, gives rise to strong photoluminescence and large exciton-binding energies, making TMDs attractive materials for light-emitting devices¹⁴. Likewise, transition metal thiophosphates are layered semiconductors additionally characterized by unique magnetic properties that make them potentially useful for application in quantum information devices and in spintronics¹⁵.

Here, we show how vdW materials can be used to explore novel phenomena of X-ray physics, in which free electrons passing through vdW layered structures create tunable X-ray radiation. Our findings thus present a new class of applications for vdW materials, as possible platforms for X-ray sources with a table-top footprint that have the advantages of being tunable and monochromatic relative to all existing compact sources. We demonstrate the unprecedented tunability of X-ray generation using free electrons passing through vdW

materials, and explore the processes by which the X-ray is generated, using both theory and experiments (Fig. 1a–c). The two mechanisms involved are parametric X-ray radiation (PXR) and coherent bremsstrahlung (CBS)^{16–19}. Both mechanisms arise simultaneously as a result of periodic interactions of free electrons propagating through an atomic structure (more detailed descriptions are provided at the beginning of the Results section).

In our experiment, the output photon energy is controlled by changing the incident electron energy, as well as through a method based on adjusting the composition and stacking of the vdW material structure. We present a comparative study demonstrating the highly precise dependence of the X-ray energy spectrum on the choice of the transition metal atom in the vdW material (iron (Fe), cobalt (Co) or nickel (Ni) as the 'X' in XPS₃). Looking at the big picture, the experimental results presented here constitute a proof of principle for our proposed wider design approach: to precisely tailor the radiation energy spectrum and angular distribution of X-ray emission by means of material design at the atomic level. We show that a wide range of superlattice crystalline materials, both natural and artificial, can be used for X-ray generation, and we present a theory for X-ray emission from designed superlattice structures. These findings pave the way towards the use of superlattice atomic structures for realization of tunable and versatile sources of X-ray radiation with a table-top footprint.

The appeal of using vdW materials for X-ray generation is enhanced by the fact that many vdW materials possess high in-plane thermal conductivities²⁰, and some have higher melting temperatures than conventional materials. Moreover, radiation damage can be further reduced by using heterostructures combining different

¹Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. ²Schulich Faculty of Chemistry, Technion—Israel Institute of Technology, Haifa, Israel. ³Department of Materials Science and Engineering, Technion—Israel Institute of Technology, Haifa, Israel. ⁴ICFO—Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Castelldefels, Spain. ⁵School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA. ⁶CAMD and Center for Nanostructuring Graphene (CNG), Department of Physics, Technical University of Denmark, Kongens Lyngby, Denmark. ⁷ICREA—Institut de Recerca i Estudis Avançats, Barcelona, Spain. ⁸School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore. ✉e-mail: michael.shentcic@gmail.com; kaminer@technion.ac.il

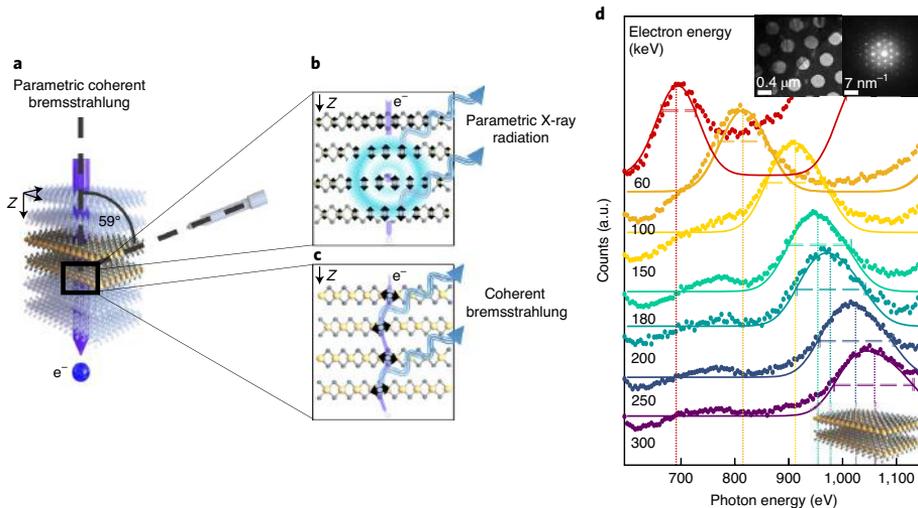


Fig. 1 | Demonstration of free-electron radiation from vdW materials. **a**, The radiation is produced from two combined mechanisms, PXR and CBS, both emitted with the same dispersion relation (equation (1)) from an electron (e^-) propagating through periodic crystal structures (aligned along the Z axis). The angle between the direction of the detector and the negative direction of electron propagation is 59° . **b**, Illustration of the PXR mechanism, produced as the particle's electromagnetic field (light blue halo) is diffracted off the periodic crystal structure. **c**, Illustration of the CBS mechanism, describing the X-ray emission by an undulated electron due to a series of periodic bremsstrahlung interactions with the crystal lattice. **d**, Photon energy spectrum of X-ray radiation created by an electron moving along the [001] zone axis of WSe_2 . The emitted radiation peaks at different photon energies, depending on the electron kinetic energy (60–300 keV). The experimental results (dotted curves) are in good agreement with the theoretical prediction both for the peak energy (equation (1) with $m=2$, vertical dotted lines) and the peak width (equation (2), horizontal dashed lines), using no fitting parameters. We compare the results with simulations of PXR and CBS and present the PXR theory that is found to be stronger in this case (solid curves). The trends match the experimental observation, with the only fitting parameters being the y axis scaling of each curve. We attribute deviations of the experimental results from the simulated curves to emission from core transitions (a broad peak around ~ 780 eV and on the right edge). Insets: sample image and diffraction pattern (top right); 3D model (bottom right). a.u., arbitrary units.

kinds of vdW materials^{21,22}. Other qualities are specifically attractive for PXR and CBS and arise from the weak bonding between the layers in vdW materials, which helps to maintain the crystalline nature of the material that is crucial for both mechanisms. Moreover, the crystal structures of vdW materials are characterized by larger unit cells than those of conventional three-dimensional (3D) bulk crystalline materials. As shown below, the larger unit cell enables X-ray generation with relatively high brightness in spectral ranges such as the water window (useful for biological imaging²³), using electron energies obtainable from laboratory-scale electron sources such as those used in transmission electron microscopes (TEMs) and scanning electron microscopes (SEMs). Furthermore, the wide range of compositions and flexibility in the stacking of vdW materials provide extra tunability to shape the output radiation through controlling the atomic lattice geometry.

Energy tunability of X-ray sources is a key factor in many applications, such as core-level spectroscopy^{24–26} and various X-ray imaging techniques^{27–29}. The energy tunability required for such applications is usually achieved in undulator facilities, such as synchrotrons and free-electron lasers^{30–32}. However, the operation of such facilities necessitates immense resources in terms of space, energy and safety measures, which limit their accessibility. Most laboratory-scale sources take the form of an X-ray tube, in which free electrons are used to induce bremsstrahlung or core transitions that produce X-rays. Such sources do not provide energy tunability or radiation directionality capabilities, which have a lot of potential for a wide variety of uses.

These limitations motivate research into new physical mechanisms for X-ray generation with the potential to create

laboratory-scale X-ray sources that are tunable and directional. In particular, research advances in X-ray-generation mechanisms have led to the development of laser-driven Compton-based X-ray sources³³ that use the micrometre-scale periodicity of light, which is much smaller than the millimetre/centimetre scales of conventional undulator facilities. Such smaller-scale undulators enable the levels of electron acceleration to be reduced, while still resulting in tunable X-ray generation. More recent proposals rely on shrinking the undulating periodicity even further by leveraging the high electromagnetic confinement in graphene surface plasmons³⁴, metasurfaces³⁵, metamaterials³⁶ and nanophotonic vacuum fluctuations³⁷. Our study explores different phenomena that lead to X-ray generation from free electrons, in which the electron undulation is done at the ultimate periodicity: that of the atomic crystal lattice.

Results

Tunable X-ray radiation from vdW materials. The vdW-material-based atomic undulator that we present here relies on two radiation mechanisms, CBS and PXR, using modest electron energies that are available in relatively simple systems such as a TEM. The mechanisms of CBS and PXR have not yet been studied in vdW materials. Both mechanisms take place as a propagating electron undergoes coherent interaction with the intrinsic periodicity of a crystalline material (Fig. 1).

PXR is emitted from the modulation induced by the incident electron on the bound electrons of the material's atoms, creating polarization currents from which directional X-ray radiation is emitted. The PXR mechanism can also be described as diffraction of the incident electron's Coulomb field off the periodical atomic

arrangement of the structure (Fig. 1b). We model PXR by treating the periodic atomic structure as a discrete dipole array (described by the black arrows in Fig. 1b), with bound electrons around each atom being modelled as an effective dipole quantified through an associated X-ray atomic polarizability.

CBS is emitted from the incident electron as a result of a periodic series of bremsstrahlung interactions with the nuclei and bound electrons in the material, where each interaction involves acceleration of the electron, causing it to emit radiation (Fig. 1c). The modulation of the free-electron velocity by the periodic potential results in interference of the emitted radiation and causes it to become directional. We model CBS by considering the far-field radiation from electrons whose trajectories are calculated by using the relativistic Newton–Lorentz equations in the presence of an atomic potential, which is in turn obtained from density functional theory (DFT; Methods). We can equivalently describe the CBS emission by using a propagating radiating dipole excited as a result of the field induced by the electrons of the crystal structure (described by the black arrows in Fig. 1c).

The radiation emitted from each electron through the CBS and PXR mechanisms maintains temporal and spatial coherence (in contrast to coherence from multiple electrons, which requires bunching of the electrons). That is, the radiation emitted from multiple locations along the trajectory of the incident electron interferes constructively to produce angle-dependent output radiation (coherent interference from multiple locations is possible because the final quantum state of the material is not changed). The radiation from both mechanisms follows the same dispersion relation¹⁹

$$E_m = \hbar\omega_m = m \frac{2\pi\hbar c\beta \cos(\theta)}{d[1-\beta\cos(\varphi)]} \quad (1)$$

where $\beta = v/c$ is the normalized speed of the electron (where v is the velocity of the electron and c is the velocity of light); E_m and ω_m are the photon energy and angular frequency, respectively, corresponding to radiation of order m (where $m = 1, 2, 3, \dots$); \hbar is the reduced Planck constant; d is the lattice constant; and θ and φ are the angles defining the electron velocity relative to the reciprocal lattice of the crystal and the photon emission direction, respectively.

Each of the PXR and CBS mechanisms becomes of greater influence at different regimes of emission energies and angles. However, they both share the same dispersion relation (equation (1)), and thus they can be effectively considered as a single combined effect^{38,39}, which we refer to as parametric coherent bremsstrahlung (PCB).

The unique structure and large variety of vdW materials make them promising candidates for PCB radiation for several reasons. One reason arises from the unique layered structure of vdW materials, which are made of strong covalently bonded atomic layers joined together by vdW forces. Consequently, the effect of integrating certain atoms (for example, tungsten) into a vdW structure (for example, WSe₂) can be seen as effectively ‘stretching’ the lattice constant d , thus redshifting the radiation energy peaks E_m . For example, the lattice constant of bulk tungsten, 3.16 Å, is increased to 12.98 Å (unit cell size in an AB stacking, that is, the distance between two alternate layers) when it is embedded into a WSe₂ TMD structure, giving rise to PCB X-ray radiation with a different energy–angle profile for the same incident electron velocities (Fig. 1d). The longer effective period of the vdW crystal allows soft X-ray photons (for example, in the water window) to be created with higher incident electron energy than in regular 3D bulk crystals—seen by the trade-off of d and β in equation (1). This regime of parameters leads to higher output power and brightness when using electrons from TEMs and SEMs.

Figure 2a,b presents PCB radiation from additional vdW structures, namely, MnPS₃ and CrPS₃, oriented along the [103] and [001] zone axes, respectively. As in Fig. 1d, we note a good agreement of the experimental results with the predicted energy values calculated from equation (1).

Using the PCB theory, we also predict the width $\Delta\omega$ of the spectral peaks, which for thin TEM samples and highly directional electron beams yields (Supplementary Section 1)

$$\frac{\Delta\omega}{\omega_m} = \frac{\Delta E}{E_m} \approx \sqrt{0.8 \frac{d^2}{m^2 L^2} + \Delta\varphi_D^2 \frac{\beta^2 \sin^2 \varphi}{[1-\beta\cos\varphi]^2} + \Delta\theta_\epsilon^2 \tan^2 \theta} \quad (2)$$

with ΔE being the energy width, L the electron interaction length in the crystal, $\Delta\varphi_D$ the angular aperture of the detector and $\Delta\theta_\epsilon$ the angular spread of the incident electron beam. In the present study, the width of the spectral lines is primarily determined by the angular aperture of the energy-dispersive X-ray spectrometer, 16°, which collects the emission for a range of angles $\varphi = 113^\circ$ – 129° . In contrast, the left term in the square root of equation (2) is comparatively negligible in our case because the crystal thicknesses of the different materials ($L \approx 100$ nm) do not introduce notable broadening relative to the angular aperture of the detector. Similarly, $\Delta\theta_\epsilon < 0.1$ mrad, estimated to be the electron beam divergence angle in our experiment, is small and does not substantially alter $\Delta\omega$. An additional effect of considerable broadening is the detector energy resolution (~ 80 eV), which has to be combined with the result of equation (2) (Methods).

We can use equation (2) to estimate the monochromaticity of the emitted radiation in a way that is independent of the detector parameters. Neglecting $\Delta\varphi_D$, we find that PCB radiation generated by a collimated electron beam results in $\Delta\omega/\omega_m = 0.9d/mL$, which means $\frac{1.2}{m}$ % for an interaction length of 100 nm in WSe₂. Therefore, when collected over a small angle, the X-rays produced are indeed very monochromatic, possessing a narrow bandwidth below 1% for $m \geq 2$.

The positions of the experimental peak maxima in Figs. 1d and 2a–c show small deviations from the theoretical values (marked by vertical dotted lines), possibly due to small uncertainties in the detector angle or in the reported values of the interlayer distances used here as input parameters. Additional effects that explain deviations between theory and experiment include the background of incoherent bremsstrahlung radiation, different interaction lengths for each incident electron energy, divergence of the incident electron beam and further subsequent diffraction of the radiation by the crystal and the detector (Supplementary Section 7). The last effect should be sensitive to the properties of the electron beam because it essentially consists of Kossel-line-like diffraction from a coherent source comprising the emission from all of the atoms exposed to each incident electron wave. Additionally, our PXR simulation shows that since the detector is located in the backward direction relative to the electron velocity, there is a small Doppler shift that causes the energy peak to be slightly redshifted, in qualitative agreement with our measurements (seen in the theory and experiments in Fig. 1d).

We use the simulation tools here developed for CBS and PXR to quantify both mechanisms and compare them with our experimental results. We find that in our regime of interest (moderately relativistic electrons available from TEMs), the PXR mechanism dominates the radiation emission as it is two orders of magnitude larger than CBS. These models enable the average brightness values of PCB radiation sources from vdW materials to be theoretically predicted. For example, we examine the brightness values obtainable in a TEM set-up such as that used in this work: a relatively low electron current of ~ 1 nA and a ~ 1 nm electron beam diameter (at the plane of the sample) passing through a WSe₂ sample of ~ 100 nm thickness. For the detector orientation of $\varphi = 121^\circ$ with respect to the electron velocity, we estimate a brightness value of $\sim 1 \times 10^9$ photons s^{-1} mrad^{-2} mm^{-2} 0.1BW^{-1} within our range of energy tunability of ~ 700 – $1,100$ eV (where BW is the bandwidth). This level of brightness compares favourably with state-of-the-art X-ray tubes, while the input power is smaller by a factor of 10^{-5} – 10^{-8} (because

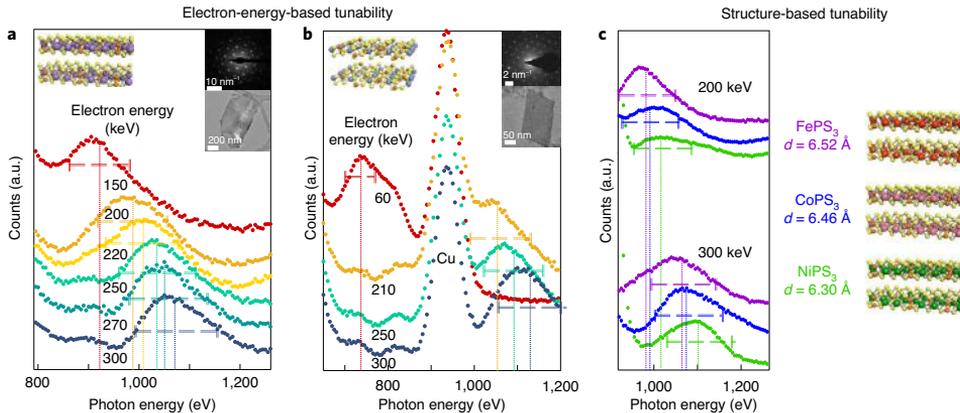


Fig. 2 | Tunability of X-ray radiation from vdW materials. **a, b**, X-ray radiation spectra collected from MnPS₃ (**a**) and CrPS₃ (**b**). The photon energy peak is tuned by varying the incident electron kinetic energy in the range 60–300 keV, as indicated. Theoretically predicted peak energy (equation (1) for $m = 2$, vertical dotted lines) and peak width (equation (2), horizontal dashed lines) both show a good match with the experimental results. The constant energy peak in **b** is the characteristic radiation peak emitted from copper (Cu). Top right insets: sample image and diffraction pattern for crystal orientations of [001] (CrPS₃) (**b**) and [103] (MnPS₃) (**a**). Top left insets: 3D models of the layered vdW structures. **c**, Radiation from three vdW materials with the same crystal structure (FePS₃ (purple), CoPS₃ (blue) and NiPS₃ (green)), but differing by a picometre-scale difference in d (right inset). The resulting radiation energy tuning through structural modification is demonstrated at two incident electron kinetic energies, showing the possibility for combined X-ray energy tunability via variation of the structure and of the electron acceleration voltage. Sensitivity of the photon energy tuning is high, being limited only by the energy resolution and angular aperture of the detector. a.u., arbitrary units.

the electron spot diameter in our experiment is much smaller than that of X-ray tubes). Moreover, our PCB radiation is directional and tunable, in contrast to the radiation from X-ray tubes (which is either characteristic or bremsstrahlung). The estimated numbers of photons collected at the solid angle of the detector are 4.0×10^3 photons $s^{-1} eV^{-1}$, 2.6×10^3 photons $s^{-1} eV^{-1}$, 1.8×10^3 photons $s^{-1} eV^{-1}$, 1.5×10^3 photons $s^{-1} eV^{-1}$, 1.4×10^3 photons $s^{-1} eV^{-1}$, 1.1×10^3 photons $s^{-1} eV^{-1}$ and 1.0×10^3 photons $s^{-1} eV^{-1}$, at each peak from 60 keV to 300 keV (Fig. 1d and Methods). The typical number of photons per second is $\sim 10^5$ (integrating over the widths of the peaks).

The brightness of the PCB source can be further improved by optimizing parameters such as the detector orientation and size, the electron acceleration voltage and the sample thickness (similarly to other undulation-based emission mechanisms, the PCB brightness is quadratic with the interaction length). It can also be increased by using larger currents and reducing the electron spot size, for which an optimum compromise can be found by keeping in mind the trade-off enforced by Coulomb repulsion (space charge), which leads to greater beam divergence (hence smaller interaction length) with larger electron density. When the electron beam divergence exceeds about 0.1° , the heights of the radiation peaks start to broaden and decrease notably ($>5\%$ from their original values). When considering space charge, we show in Supplementary Fig. 5 that the potential increase in brightness of PCB from vdW materials can be as much as 10^7 times the value reported above (for an optimal combination of electron current and beam size).

The X-ray generation from the PCB mechanism can be extended to the hard X-ray regime with higher electron energies, as we show in Supplementary Section 4, where we simulate the output radiation from 1 MeV and 5 MeV electrons. Electron beams generated by relatively compact sources such as photoemission injectors based on radio-frequency guns and d.c. high-voltage guns^{40,41} enable higher electron currents of up to tens of milliamperes in acceleration voltages of a few megaelectronvolts, which can strongly improve the emission brightness from thicker crystal samples.

The performance of PCB radiation from vdW materials can also be viewed from the perspective of energy transfer and efficiency. The total probability of each 60 keV electron going through a 100 nm interaction length in WSe₂ to produce PCB emission is $\sim 1 \times 10^{-4}$, resulting in an average electron energy loss to PCB of 0.25 eV. Out of this, the probability of radiation in the direction of the detector is $\sim 1 \times 10^{-5}$, that is, electron energy loss to ‘useful X-ray photons’ of 0.025 eV. This probability is of the same order of magnitude as is found in related processes such as Smith–Purcell radiation. The efficiency can be further improved because the electron energy can be reused. Such an approach is used in many free-electron-based applications, such as travelling-wave tubes. Recycling the electron energy means that the absolute efficiency depends on competing channels of energy loss in the sample. In the case of PCB radiation, such processes mainly come from Coulomb collisions that result in the ionization of other electrons, excitation of atoms and non-coherent bremsstrahlung radiation. We use a numerical simulator developed by the National Institute of Standards and Technology to estimate the total energy loss of a 60 keV electron in a 100 nm sample from our materials, and find an average energy loss of ~ 300 eV. Therefore, we predicted the efficiency of the mechanism to be $\sim 0.1\%$ (60 keV electrons, 100 nm WSe₂).

At higher energies, the efficiency improves as the radiation is more directional due to relativistic contraction. In addition, higher electron energies have longer penetration depths in thicker samples (that is, lower competing loss channels). Overall, such efficiency can be tolerable when considering that X-ray sources are never generally efficient and bearing in mind that unlike any other compact X-ray source, PCB-based sources are tunable.

Discussion

We now discuss methods to control the emitted spectra from PCB radiation. The X-ray energy is tuned in Figs. 1d and 2a,b by adjusting the incident electron energy. However, such a method necessitates realignment to ensure that different conditions, such as the quality of the electron beam collimation, remain unchanged, thus adding a degree of difficulty to that approach. Another method for

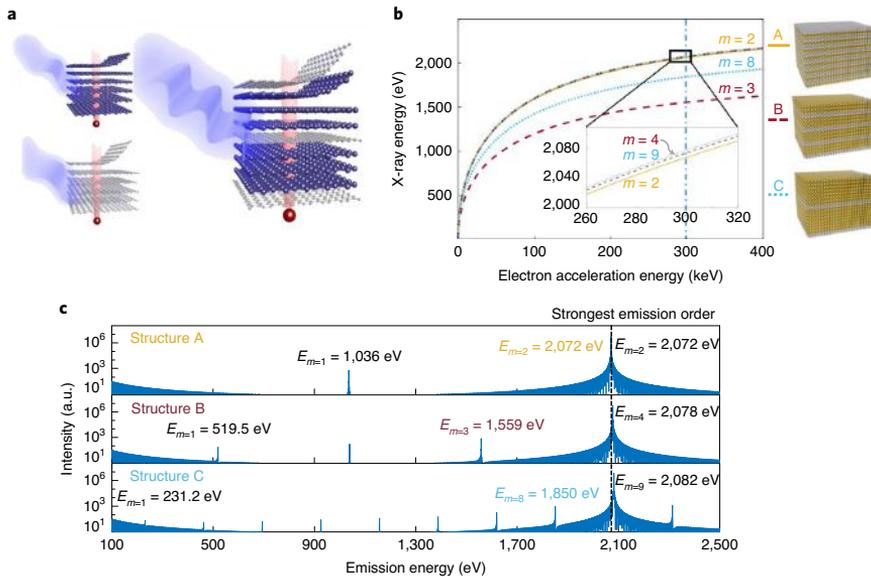


Fig. 3 | Spectral shaping of X-ray radiation via customized superlattices. **a**, Illustration of the proposed design concept for tuning PCB X-ray radiation energy, combining two crystalline materials into a heterostructure whose periodicity is customized by the choice of layers; vdW stacking and molecular beam epitaxy provide possible approaches to these heterostructures. **b**, X-ray emission peak energy of superlattice structures formed by alternating layers of graphite ($d_1=0.3350$ nm) and hexagonal boron nitride ($d_2=0.3308$ nm), used as an illustrative example. Each superlattice contains a different number of layers from each material, as shown in the models on the right: $n_1=1, n_2=1$ (structure A, yellow line); $n_1=1, n_2=3$ (structure B, red dashed line); $n_1=1, n_2=8$ (structure C, blue dotted line). The different choice of superlattice enables direct control over the number and energy of the resulting emission peaks. The choice of different emission orders enables coarse or fine energy tunability, as demonstrated in the plot and inset, respectively. **c**, Full emission spectra simulated under the conditions indicated by the vertical blue dashed-dotted line in **b** (300 keV kinetic energy), showing that the most dominant order of emission for each structure is $m=n_1+n_2$ (see equation (3)). a.u., arbitrary units.

adjusting the output photon profile involves changing the orientation of the crystal with respect to the electron beam, which, unfortunately, compromises the emission efficiency and other properties of the output radiation.

Taking a different approach, we present in Fig. 2c the results of our method for controlling the output radiation spectra by precise manipulation of the lattice structure. Using several vdW materials with similar basic structures in which the main element is altered, we show that the resulting small but precise changes in the lattice constant lead to corresponding precision control over the spectral peaks of the emitted PCB radiation. Figure 2c shows three vdW materials sharing the formula XPS₃, where ‘X’ represents the transition metals Fe/Co/Ni. The lattice constant differs slightly between these vdW compounds (Fig. 2c, right inset), enabling fine energy tuning of the emission through lattice variation. This tuning method provides an extra handle of versatility on top of the conventional PCB tuning mechanisms of electron energy and orientation, further supporting the potential of vdW materials as compact platforms for high-quality X-ray generation.

Tunability by the structural design of vdW materials is the first step towards our proposed broader design concept, in which a specially engineered superlattice structure can be used to spectrally shape the emitted X-ray radiation. Analogous structural customizations for manipulating radiation have been suggested for other applications at optical wavelengths in metasurfaces⁴² and in designer Smith–Purcell gratings^{43,44}. However, this approach has never been considered for X-ray radiation manipulation.

Figure 3a presents our design approach by exploring a general superlattice configuration in which two crystalline materials are

combined into a superlattice. Mathematically, we derive a corrected dispersion relation where the lattice factor d in equation (1) is replaced by a weighted sum of lattice factors $d_1n_1+d_2n_2$, with d_1 and d_2 denoting the lattice constants and n_1 and n_2 the number of layers in each period of the two combined crystals. We thus have (generalizing equation (1))

$$E'_m = \hbar\omega'_m = m \frac{2\pi\hbar\beta\cos\theta}{(d_1n_1+d_2n_2)(1-\beta\cos\theta)} \quad (3)$$

PCB radiation from monocrystalline materials is usually strongest for the first emission order ($m=1$). However, in superlattices made of two materials, the dominant emission would occur at $m=n_1+n_2$ (the total number of layers in a single superlattice period).

Figure 3b,c shows that changing the atomic composition of the superlattice period allows us to control the range of emitted radiation energy for the same incident electron velocities. The curves in Fig. 3b are calculated by combining graphite ($d_1=0.3350$ nm) and hexagonal boron nitride ($d_2=0.3308$ nm) into a superlattice structure and presenting several emission orders of such structures. The theoretical estimate of the intensity relation between emission orders in Fig. 3c is based on a model that uses a one-dimensional array of radiating dipoles. In this model, each dipole represents a lattice plane of the superlattice. As the PCB energy profile is dependent on the constructive interference of emission from consecutive lattice planes, we can calculate the relation between the emission peaks by considering the electromagnetic field emitted from a row of point dipoles with the corresponding relative phases. We note an excellent match between the emitted energies presented in this model and the dispersion relation in equation (3).

Outlook

We envision the use of vdW heterostructures^{45,46} and other artificial superlattice crystalline materials to optimize the emission of PCB X-ray radiation with user-defined spectral and spatial properties. These sources are also attractive due to practical considerations, such as the promise of higher stability and smaller dependence on recalibrations and electron beam alignments, because it is possible to tune the emission by changing between a series of superlattices without modifying the electron beam properties. Fabrication of such superlattice–PCB sources can benefit from a variety of new crystal growth techniques, as well as the large wealth of possible chemical compositions and crystal structures offered by vdW materials. Additionally, molecular beam epitaxy⁴⁷, metal organic chemical vapour deposition⁴⁸ and other established material-growth techniques could be used to enable superlattice design incorporating other material families (for example, III-Vs, II-VIs) for X-ray generation and manipulation.

Another meaningful improvement would be the use of nano-modulated electrons that could make the radiation coherent from multiple electrons (microbunching). These nano-modulated electrons can be generated via emittance exchange techniques^{49,50}, laser–plasma interactions⁵¹ or electromagnetic intensity gratings⁵². Future work could develop structures that achieve resonant PCB conditions for interaction between the propagating electrons and periodic structure to improve the emission efficiency of PCB radiation.

Our theoretical and experimental results show the advantages of vdW materials in engineering the emission of energy-tunable X-ray radiation. VdW materials add versatility to X-ray energy tuning of PCB radiation based on our proposed lattice modification technique, which can be combined with previously explored tuning techniques or used as a separate method. Additionally, this technique demonstrates the advantages of designer superlattices for energy-tunable emission of monochromatic and directional X-ray radiation.

In contrast to state-of-the-art tunable X-ray sources, PCB radiation can be generated from modestly relativistic electrons. This makes the superlattice–PCB X-ray sources relatively compact and more accessible for various applications, such as X-ray spectroscopy^{24–26}. For example, PCB radiation from vdW materials can be used for compact and tunable X-ray photoemission spectroscopy methods. Such an application is already within reach, using the emission flux and brightness of PCB from the low current of a basic TEM set-up, as described above.

Additional research tools such as transmission X-ray microscopy and scanning photoelectron microscopy⁵³ are also potential applications for future PCB sources. X-ray imaging techniques can benefit from the use of energy-tunable sources, as the image contrast is affected by the energy of the incident radiation^{27–29}. This concept may be used, for instance, to improve the yield and accuracy of mammography techniques, in addition to reducing the radiation dose delivered to a patient²⁸. Future imaging techniques that use the tunability of superlattice–PCB X-ray sources can create compact variants of K-edge imaging techniques⁵⁴ with limited brightness and flux, yet at accessible laboratory scales.

In conclusion, we theoretically predict and experimentally demonstrate tunable, high-brightness X-ray emission from vdW materials when irradiated by moderately relativistic electrons available, for example, from a TEM. The radiation spectrum can be precisely controlled by tuning the acceleration voltage of the incident electrons, as well as by our proposed approach—adjusting the lattice structure of the vdW material. We present experimental results for both methods, observing the energy tunability of X-ray radiation from the vdW materials WSe₂, CrPS₃, MnPS₃, FePS₃, CoPS₃ and NiPS₃. Our findings demonstrate the concept of material design at the atomic level, using vdW heterostructures and other atomic superlattices, for exploring novel phenomena of X-ray physics. Looking forward,

the control of atomic layers in both the longitudinal and transverse directions would enable spatial, angular and spectral shaping of output X-ray radiation. We envision the customization of superlattices for user-specific applications via inverse design techniques that optimize the desired output radiation characteristics, given the electron beam conditions and the specific geometrical constraints.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41566-020-0689-7>.

Received: 29 December 2019; Accepted: 7 August 2020;

Published online: 14 September 2020

References

- Novoselov, K. S. et al. Two-dimensional atomic crystals. *Proc. Natl Acad. Sci. USA* **102**, 10451–10453 (2005).
- Basov, D. N., Fogler, M. M. & García de Abajo, F. J. Polaritons in van der Waals materials. *Science* **354**, aag1992 (2016).
- Low, T. et al. Polaritons in layered two-dimensional materials. *Nat. Mater.* **16**, 182–194 (2016).
- Novoselov, K. S. et al. Electric field effect in atomically thin carbon films. *Science* **306**, 666–669 (2004).
- Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–162 (2009).
- Hao, J. et al. High performance optical absorber based on a plasmonic metamaterial. *Appl. Phys. Lett.* **96**, 251104 (2010).
- Lee, C., Wei, X., Kysar, J. W. & Hone, J. Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science* **321**, 385–388 (2008).
- Balandin, A. A. Thermal properties of graphene and nanostructured carbon materials. *Nat. Mater.* **10**, 569–581 (2011).
- Kane, C. L. & Mele, E. J. Quantum spin Hall effect in graphene. *Phys. Rev. Lett.* **95**, 226801 (2005).
- Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nat. Nanotechnol.* **7**, 699–712 (2012).
- Chhowalla, M. et al. The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. *Nat. Chem.* **5**, 263–275 (2013).
- Evain, M., Brec, R. & Wbango, M. H. Structural and electronic properties of transition metal thiophosphates. *J. Solid State Chem.* **71**, 244–262 (1987).
- Latini, S., Olsen, T. & Thygesen, K. S. Excitons in van der Waals heterostructures: the important role of dielectric screening. *Phys. Rev. B* **92**, 245123 (2015).
- Jariwala, D., Sangwan, V. K., Lauhon, L. J., Marks, T. J. & Hersam, M. C. Emerging device applications for semiconducting two-dimensional transition metal dichalcogenides. *ACS Nano* **8**, 1102–1120 (2014).
- Susner, M. A., Chyasnavichyus, M., McGuire, M. A., Ganesh, P. & Maksymovych, P. Metal thio- and selenophosphates as multifunctional van der Waals layered materials. *Adv. Mater.* **29**, 1602852 (2017).
- Überall, H. High-energy interference effect of bremsstrahlung and pair production in crystals. *Phys. Rev.* **103**, 1055–1067 (1956).
- Korobochko, Y. S., Kosmach, V. F. & Mineev, V. I. On coherent electron bremsstrahlung. *Sov. Phys. JETP* **21**, 834–839 (1965).
- Baryshevsky, V. G. & Feranchuk, I. D. Parametric X-rays from ultrarelativistic electrons in a crystal: theory and possibilities of practical utilization. *J. Phys. France* **44**, 913–922 (1983).
- Baryshevsky, V. G., Feranchuk, I. D. & Ulyanenko, A. P. *Parametric X-ray Radiation In Crystals* (Springer, 2005).
- Jiang, P., Qian, X., Gu, X. & Yang, R. Probing anisotropic thermal conductivity of transition metal dichalcogenides MX₂ (M = Mo, W and X = S, Se) using time-domain thermoreflectance. *Adv. Mater.* **29**, 1701068 (2017).
- Zan, R. et al. Control of radiation damage in MoS₂ by graphene encapsulation. *ACS Nano* **7**, 10167–10174 (2013).
- Lehnert, T., Lehtinen, O., Algara-Siller, G. & Kaiser, U. Electron radiation damage mechanisms in 2D MoSe₂. *Appl. Phys. Lett.* **110**, 033106 (2017).
- Kirz, J., Jacobsen, C. & Howells, M. Soft X-ray microscopes and their biological applications. *Q. Rev. Biophys.* **28**, 33–130 (1995).
- de Groot, F. & Kotani, A. *Core Level Spectroscopy of Solids* (CRC Press, 2008).
- Hitchcock, A. P. Soft X-ray spectroscopy and ptychography. *J. Electron Spectrosc. Relat. Phenom.* **200**, 49–63 (2015).

26. Agarwal, B. K. *X-ray Spectroscopy: an Introduction* Vol. 15 (Springer, 2013).
27. Hayakawa, Y. et al. X-ray imaging using a tunable coherent X-ray source based on parametric X-ray radiation. *J. Instrum.* **8**, C08001 (2013).
28. Carroll, F. E. Tunable monochromatic X rays: a new paradigm in medicine. *Am. J. Roentgenol.* **179**, 583–590 (2002).
29. Okada, H. et al. Basic study of parametric X-ray radiation for clinical diagnosis using 125 MeV linear particle accelerator. *J. Hard Tissue Biol.* **24**, 299–302 (2015).
30. Saldin, E. L., Schneidmiller, E. A. & Yurkov, M. *The Physics of Free-Electron Lasers* (Springer, 2000).
31. Ackermann, W. et al. Operation of a free-electron laser from the extreme ultraviolet to the water window. *Nat. Photon.* **1**, 336–342 (2007).
32. Winick, H. & Doniach, S. *Synchrotron Radiation Research* (Springer, 2012).
33. Powers, N. D. et al. Quasi-monoenergetic and tunable X-rays from a laser-driven Compton light source. *Nat. Photon.* **8**, 28–31 (2014).
34. Wong, L. J., Kaminer, I., Ilic, O., Joannopoulos, J. D. & Soljačić, M. Towards graphene plasmon-based free-electron infrared to X-ray sources. *Nat. Photon.* **10**, 46–52 (2016).
35. Rosolen, G. et al. Metasurface-based multi-harmonic free-electron light source. *Light Sci. Appl.* **7**, 64 (2018).
36. Pizzi, A. et al. Graphene metamaterials for intense, tunable, and compact extreme ultraviolet and X-ray sources. *Adv. Sci.* **7**, 1901609 (2019).
37. Rivera, N., Wong, L. J., Joannopoulos, J. D., Soljačić, M. & Kaminer, I. Light emission based on nanophotonic vacuum forces. *Nat. Phys.* **15**, 1284–1289 (2019).
38. Blazhevich, S. V. et al. First observation of interference between parametric X-ray and coherent bremsstrahlung. *Phys. Lett. A* **195**, 210–212 (1994).
39. Feranchuk, I. D., Ulyanenkova, A., Harada, J. & Spence, J. C. H. Parametric X-ray radiation and coherent bremsstrahlung from nonrelativistic electrons in crystals. *Phys. Rev. E* **62**, 4225–4234 (2000).
40. Fraser, J. S., Sheffield, R. L. & Gray, E. R. A new high-brightness electron injector for free electron lasers driven by RF linacs. *Nucl. Instrum. Methods Phys. Res. A* **250**, 71–76 (1986).
41. Dunham, B. et al. Record high-average current from a high-brightness photoinjector. *Appl. Phys. Lett.* **102**, 034105 (2013).
42. Li, X. et al. Dispersion engineering in metamaterials and metasurfaces. *J. Phys. D* **51**, 054002 (2018).
43. Kaminer, I. et al. Spectrally and spatially resolved Smith–Purcell radiation in plasmonic crystals with short-range disorder. *Phys. Rev. X* **7**, 011003 (2017).
44. Remez, R. et al. Spectral and spatial shaping of Smith–Purcell radiation. *Phys. Rev. A* **96**, 061801 (2017).
45. Geim, A. K. & Grigorieva, I. V. Van der Waals heterostructures. *Nature* **499**, 419–425 (2013).
46. Gjerding, M. N., Petersen, R., Pedersen, T. G., Mortensen, N. A. & Thygesen, K. S. Layered van der Waals crystals with hyperbolic light dispersion. *Nat. Commun.* **8**, 320 (2017).
47. Herman, M. A. & Sitter, H. *Molecular Beam Epitaxy: Fundamentals and Current Status* (Springer, 2012).
48. Dapkus, P. D. Metalorganic chemical vapor deposition. *Annu. Rev. Mater. Sci.* **12**, 243–269 (1982).
49. Graves, W., Kärtner, F., Moncton, D. & Piot, P. Intense superradiant X rays from a compact source using a nanocathode array and emittance exchange. *Phys. Rev. Lett.* **108**, 263904 (2012).
50. Nanni, E. A., Graves, W. S. & Moncton, D. E. Nanomodulated electron beams via electron diffraction and emittance exchange for coherent X-ray generation. *Phys. Rev. Accel. Beams* **21**, 014401 (2018).
51. Naumova, N. et al. Attosecond electron bunches. *Phys. Rev. Lett.* **93**, 195003 (2004).
52. Lim, J., Chong, Y. & Wong, L. J. Terahertz-optical intensity grating for creating high-charge, attosecond electron bunches. *New J. Phys.* **21**, 033020 (2019).
53. Attwood, D. T. *Soft X-Rays and Extreme Ultraviolet Radiation* (Cambridge Univ. Press, 2000).
54. Roessler, E. et al. Sensitivity of photon-counting based K-edge imaging in X-ray computed tomography. *IEEE Trans. Med. Imaging* **30**, 1678–1690 (2011).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Experimental. The radiation was generated by highly collimated electron beams in an FEI Titan Themis G2 TEM with acceleration voltage in the range 60–300 kV, used to control the velocity of the incident electrons. The energy spectrum was measured using a Dual Bruker XFlash 6 | 100 energy-dispersive X-ray spectroscopy detector oriented at $\varphi = 121^\circ$ with respect to the electron velocity vector and with $\Delta\varphi_0 = 16^\circ$. The energy resolution of the detector in the energy range of the measurements is ~80 eV full-width at half-maximum. The resolution was determined by fitting a Gaussian distribution to the characteristic radiation energy peaks of W and Se in proximity to the PCB radiation peaks. All spectra were obtained by aligning the collimated electron beams along the crystal zone axis ([103] and [001]), verified by their diffraction patterns. The input power in our experiment was of the order of ~0.1 mW (~100 kV × 1 nA), much lower than that of commercial X-ray tubes, which lies in the range of tens to thousands of watts. The relatively low input power required relatively long accumulation times of 16 min during the experiment. However, no visible damage was detected on the sample after up to ten such measurements.

Theoretical. The theoretical prediction of the peak emission (vertical dotted lines in Figs. 1d and 2) were obtained using equation (1) without any fitting parameters. We developed two simulation frameworks. In the CBS simulator, the trajectories of the electrons travelling in the vdW material were obtained by solving the relativistic Newton–Lorentz equations using a fifth-order Runge–Kutta algorithm^{34,35}. The impinging electrons were assumed to be normally incident on the plane of the vdW layers. The atomic potential of the vdW material was obtained from DFT computations that take into consideration all inner-shell electrons^{36,37}. The DFT calculations were performed with the GPAW (Grid-based Projector-Augmented Wave) code³⁸, which uses a plane-wave basis set to expand the electron wave functions. The radiation from the electrons was calculated from their trajectories via the Liénard–Wiechert formula. Since the electron beam width is far larger than the transverse lattice cell size, the total radiation spectrum was obtained by averaging over the radiation spectra of individual electrons spanning one unit cell of the atomic structure.

In the PXR simulator, the atomic structure was modelled by periodic dipole arrays, and the incident electrons were assumed to travel along straight lines with uniform velocity. The response of the bound electrons around each atom was modelled by effective dipoles, with the atomic polarizability derived from the experimental scattering factor³⁹.

The widths of the peaks in Figs. 1 and 2 were calculated by combining ΔE (calculated using equation (2)) with the energy broadening due to the detector energy resolution ΔE_{res} , for an approximate energy width of $\Delta E_{\text{tot}} = \sqrt{\Delta E^2 + \Delta E_{\text{res}}^2}$. The energy widths were also found in another more precise way, using the PXR simulation, and convolving the spectra with a Gaussian (full-width at half-maximum of ΔE_{res}) that models the response of the detector. The results yield the theoretical spectral line presented in Fig. 1. The energy widths obtained in both methods approximate well the experimental data.

Sample preparation. Chromium thiophosphate (CrPS₄) was synthesized by vapour transport synthesis with a temperature gradient of 750 °C/710 °C and a reaction time of 4 days, as described in detail elsewhere⁶⁰. Transition metal phosphorous trisulfides were synthesized by the same technique, but the reaction time was longer (1 week), and the temperature gradient was different for each of the specific compounds—MnPS₃: 650 °C/600 °C; FePS₃: 650 °C/550 °C; CoPS₃: 625 °C/575 °C; NiPS₃: 650 °C/600 °C. The vdW materials were mechanically exfoliated and directly transferred to a TEM grid using a protocol described in ref. ⁶⁰.

Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding authors upon reasonable request.

Code availability

The codes that support the plots within this paper and other findings of this study are available from the corresponding authors upon reasonable request.

References

- Wong, L. J. et al. Laser-induced linear-field particle acceleration in free space. *Sci. Rep.* **7**, 11159 (2017).

- Wang, W. L. & Kaxiras, E. Efficient calculation of the effective single-particle potential and its application in electron microscopy. *Phys. Rev. B* **87**, 085103 (2013).
- Susi, T. et al. Efficient first principles simulation of electron scattering factors for transmission electron microscopy. *Ultramicroscopy* **197**, 16–22 (2019).
- Enkovaara, J. et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys. Condens. Matter* **22**, 253202 (2010).
- Henke, B. L., Gullikson, E. M. & Davis, J. C. X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E = 50\text{--}30,000$ eV, $Z = 1\text{--}92$. *At. Data Nucl. Data Tables* **54**, 181–342 (1993).
- Budniak, A. K. et al. Exfoliated CrPS₄ with promising photoconductivity. *Small* **16**, 1905924 (2020).

Acknowledgements

We thank Y. Kauffmann for advice and discussions. This work was supported by the ERC (Starter Grant no. 851780), the ISF (Grant no. 830/19) and the European Commission via the Marie Skłodowska-Curie Action Phonsi (H2020-MSCA-ITN-642656). H.H.S. also acknowledges the support of Marie Skłodowska-Curie Actions (H2020-MSCA-IF-2018-843830). K.S.T. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant no. 1773122, LIMA). The Center for Nanostructured Graphene is sponsored by the Danish National Research Foundation, Project DNRF103. F.H.L.K. acknowledges financial support from the Government of Catalonia through the SGR grant, and from the Spanish Ministry of Economy and Competitiveness, through the “Severo Ochoa” Programme for Centres of Excellence in R&D (SEV-2015-0522), and Explora Ciencia FIS2017-91599-EXP. F.H.L.K. also acknowledges support by Fundació Cellex Barcelona, Generalitat de Catalunya through the CERCA program, and the Mineco grants Plan Nacional (FIS2016-81044-P) and the Agency for Management of University and Research Grants (AGAUR) 2017 SGR 1656. Furthermore, the research leading to these results has received funding from the European Union's Horizon 2020 under grant agreement no. 785219 (Core2) and no. 881603 (Core3) Graphene Flagship, and no. 820378 (Quantum Flagship). This work was supported by the ERC TOPONANOP under grant agreement no. 726001. L.J.W. acknowledges the support of the Agency for Science, Technology and Research (A*STAR) Advanced Manufacturing and Engineering Young Individual Research Grant (A1984c0043), and the Nanyang Assistant Professorship Start-up Grant. F.J.G.A. acknowledges support from the Spanish MINECO (Grant nos. MAT2017-88492-R and SEV2015-0522), ERC (Advanced Grant no. 789104-eNANO), the Catalan CERCA Program and Fundació Privada Cellex. I.K. was also supported by an Azrieli Faculty Fellowship.

Author contributions

M.S. spearheaded the project, designed and performed the electron microscopy experiments, prepared the samples, analysed the data and developed the superlattice theory. A.K.B. contributed to the measurements and performed electron microscopy experiments. A.K.B., H.H.S., M.B., Y.A., S.T., F.H.L.K. and E.L. synthesized the vdW materials and prepared the TEM samples. R.D. and M.K. advised on experimental aspects. X.S., Y.K. and F.J.G.A. developed and executed the PXR simulations. M.K.S. and K.S.T. performed the DFT simulations. L.J.W. developed and executed the CBS simulations. F.J.G.A., L.J.W., X.S. and I.K. contributed to the discussion of the experimental results, to the comparative analysis of the different theoretical mechanisms and to the overall conclusions. M.S. and I.K. conceived the idea. I.K. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41566-020-0689-7>.

Correspondence and requests for materials should be addressed to M.S. or I.K.

Reprints and permissions information is available at www.nature.com/reprints.

Our understanding of light-matter interactions enables a wide range of the technologies that define modern life. Examples include lasers, solar cells, and light-emitting diodes, all of which are likely to play a key role in our transition towards a more sustainable society.

Light is an electromagnetic wave. Matter on the other hand is made of atoms that consist of negatively charged electrons and positively charged protons. Because of the charge of these constituents, light and matter can interact. However, the charged particles also interact with each other, and this interaction is what ultimately defines the properties of a material. These properties are called the electronic structure of the material, and they can be calculated using so-called first-principles quantum mechanical calculations. First-principles means that the only information going into the calculation is the atoms making up the material and their positions.

This thesis presents examples of how these types of calculations can be used to both interpret novel experimental results on emerging 2D materials and actively drive materials discovery for different technological applications. Light can also be used to change material properties with potential in for example chemistry. A particularly interesting version of this is how resonant electromagnetic environments can be used to alter the electronic structure of materials, even in the absence of actual light. This happens as a result of the altered quantum fluctuations of the electromagnetic field in such environments, and the effect is directly related to the fundamental quantum nature of both light and matter. It also presents a uniquely challenging computational problem because it becomes necessary to quantum mechanically treat both light and matter at the same high level of accuracy. The development of methods capable of this is another central topic of this thesis.

Technical University of Denmark
DTU Physics
Department of Physics

Fysikvej 311
2800 Kongens Lyngby, Denmark
Phone: +45 4525 3344
info@fysik.dtu.dk

www.fysik.dtu.dk

