From Fundamentals to Spectroscopic Applications of Density Functional Theory

by

Keenan Lyon

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisors:

Zoran Mišković
Professor, Department of Applied Mathematics
University of Waterloo

Duncan John Mowbray
Professor, School of Physical Sciences and Nanotechnology
Yachay Tech Unversity

Internal Member:

Lilia Krivodonova
Associate Professor, Department of Applied Mathematics
University of Waterloo

Mohammad Kohandel
Associate Professor, Department of Applied Mathematics
University of Waterloo

Internal-External Member:

Roger Melko
Professor, Department of Physics and Astronomy
University of Waterloo

External Examiner:

Mahi R. Singh
Professor, Department of Physics
Western University

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Density functional theory (DFT) and its time-dependent counterpart (TDDFT) are crucial tools in material discovery, drug design, biochemistry, catalysis, and nanoscience. However, despite its exact theoretical basis, approximations are necessary throughout, from the description of electron exchange and correlation (xc) interactions to the representation of wavefunctions for ever larger systems and the use of calculated quantities to explain and predict real-world phenomena. To address long-standing problems related to the speed and accuracy of approximations to the xc functional, we develop neural networks to emulate two such approximations, the local density (LDA) and generalized gradient (PBE) approximations, within the DFT code GPAW. We present a strategy for retraining the network and assess which training data is necessary to optimize performance for total energies over a wide class of molecules and crystals. While certain classes of materials proved difficult to describe, neural network implementations were able to reproduce the LDA and PBE xc functionals with high accuracy and a reasonable computation time. In an effort to develop a more efficient, robust, and accurate method for predicting the optical properties of low-dimensional systems, we introduce the LCAO-TDDFT-k-ω code within GPAW, where a linear combination of atomic orbitals (LCAO) representation of the Kohn-Sham wavefunctions and TDDFT implementation in wavenumber k and frequency ω space provides substantial memory and time savings, and a first order derivative discontinuity correction to the electronic gap brings the optical spectra in line with experimental measurements. Convergence of the basis set, the use of low-dimensional response functions, and different ways to incorporate the energy correction are explored for a series of materials across all dimensions: 0D fullerene and chlorophyll monomers, 1D single-walled carbon nanotubes, 2D graphene and phosphorene monolayers, and 3D anatase and rutile titanium dioxide. We develop a set of visualization tools for resolving the energetic, spatial, and reciprocal space distributions of excitations, and find LCAO-TDDFT-k-ω yields qualitative and semi-quantitative agreement with other TDDFT methods and implementations at a fraction of the time and memory cost. Finally, we introduce a phenomenological hydrodynamic model for the optical conductivity of graphene, with contributions due to universal conductivity, Pauli blocking, and intraband transitions included in a systematic way, is fit empirically with results from TDDFT, and manages to reproduce experimental spectra across a wide range of energies within energy loss equations derived for 2D materials. We find experimental parameters such as the amount of doping in graphene, the size of the collection aperture, and the energy of incoming electrons influence the shape of the spectra in important ways, especially in the energy region accessible to higher resolution probing techniques.

Acknowledgements

Thank you to my supervisors Zoran Mišković and Duncan Mowbray. You have lent your expertise and guidance to all my projects, provided opportunities and support that have helped me grow as a researcher, and ensured I made the most of my degree. I am grateful that our collaboration proved to be so fruitful these past three years, both here and in Ecuador. Thank you to my committee members Mohammad Kohandel, Lilia Krivodonova, Roger Melko, and Mahi Singh for agreeing to serve on my PhD committee and taking the time to read and review my thesis.

I have been lucky to collaborate with exceptional and talented researchers during my time here. Special thanks to Kamran Akbari, Aaron Baier-Reinio, Lindsey Daniels, Vito Despoja, Ask Larsen, and Maria Rosa Preciado-Rivas for your insights, your good company, and your help in getting me across the finish line. Thank you to my friends in Waterloo and abroad who have made this program such a joyous time.

Thank you to my family on both sides of the Pacific who have been endlessly supportive throughout my PhD. The love and support they have shown me has made this journey possible. Lastly, a big heartfelt thank you to my wife Laura for cheering me on every day.

Table of Contents

List of Figures

List of Tables

List of Acronyms

BSE	Bethe-Salpeter Equation
BZ	Brillouin Zone
C60	Carbon-60 Fullerene
Chl	Chlorophyll
CNN	Convolutional Neural Network
DFT	Density Functional Theory
DZ(P)	Double Zeta (Polarized)
EELS	Electron Energy Loss Spectroscopy
eHD	Extended Hydrodynamic
GGA	Generalized Gradient Approximation
GLLB-SC	Gritsenko - van Leeuwen - van Lenthe - Baerends, Solid Corrected
FD	Finite Difference
FFT	Fast Fourier Transform
G_0W_0-BSE	G_0W_0-Bethe-Salpeter Equation
HD	Hydrodynamic
HEG	Homogeneous Electron Gas
HOMO	Highest Occupied Molecular Orbital
IBZ	Irreducible Brillouin Zone
KS	Kohn-Sham
LCAO	Linear Combination of Atomic Orbitals
LDA	Local Density Approximation
LHC	Light Harvesting Complex
LUMO	Lowest Unoccupied Molecular Orbital
NF	Noise Factor
NN	Neural Networks
NNLDA	Neural Network Local Density Approximation

NNPBE	Neural Network Perdew, Burke, and Ernzerhof
PBC	Periodic Boundary Conditions
PBE	Perdew, Burke, and Ernzerhof
PBEsol	PBE, Solid
PW	Plane Wave
QP	Quasiparticle
QZ(P)	Quadruple Zeta (Polarized)
RBZ	Reducible Brillouin Zone
ReLU	Rectified Linear Unit
RPA	Random Phase Approximation
RS	Real Space
SC	Self-Consistent
SP	Single-Point
STEM	Scanning Transmission Electron Microscope
SWCNT	Single-Walled Carbon Nanotubes
SZ(P)	Single Zeta (Polarized)
TEM	Transmission Electron Microscope
TDDFT	Time-Dependent Density Functional Theory
TZ(P)	Triple Zeta (Polarized)
xc	Exchange-Correlation
ZLP	Zero-Loss Peak

List of Symbols

α	Polarizability
χ	Density response function
Δ_x	Derivative discontinuity correction
e	Elementary charge
ε_F	Fermi energy
ε_{xc}	Exchange correlation energy density
$\varepsilon_{n\mathbf{k}}$	Kohn-Sham eigenvalue
ε	Dielectric function
$\mathbf{E}$	Electric field
E	Energy
η	*ab initio* broadening
F	Energy loss probability density
f	Fermi occupation
$f_{nn'\mathbf{k}}^{\hat{\mathbf{q}}}$	Oscillator strength
$\mathbf{G}$	Reciprocal lattice vector
γ	Phenomenological broadening
$\hat{H}$	Hamiltonian
$\hbar$	Reduced Planck's constant
$\mathbf{J}$	Current density
k_B	Boltzmann constant
k_F	Fermi wavenumber
m_e	Electron mass
N_e	Number of valence electrons
n	Electron density
ω	Frequency
P	Energy loss density

ϕ	Electric potential
$\phi(\mathbf{r})$	Atomic orbitals
$\psi(\mathbf{r})$	KS Wavefunction
q_c	Maximum in-plane momentum transfer
$\rho_{e/h}$	Electron/hole spectral density
ρ	Charge density
σ	Conductivity
ϱ	Dimensionless electron density gradient
T	Temperature
θ_c	Collection angle
$\hat{V}$	Potential operator
v_F	Fermi velocity
$w_{\mathbf{k}}$	IBZ k-point weight

Chapter 1

Introduction

When it comes to modeling materials, quantum phenomena, and the properties of many-body systems, density functional theory (DFT) is in many ways the theoretical and computational workhorse for understanding what goes on at the nanoscale. DFT provides a framework where atomic structure optimization, band structure calculations, and intermolecular interactions can be computed for systems as diverse as small molecules and huge crystalline systems. Although the theory behind DFT dates back to the 1960s [1], when it was firmly established that the electronic density of a system is all that is required to exactly predict any ground state physical observable, it was only with an increase in computational power, milestones like the development of pseudopotential methods [2, 3], and the introduction of exchange-correlation (xc) functionals beyond the local-density approximation (LDA) [4, 5] that DFT could be relied upon to make predictions in line with experimental observation. Even for systems that are currently unsuitable for prediction within the DFT framework, such as excited-state systems or structures with highly-localized electrons [6, 7], DFT often acts as the starting point that other methods build on. Time-dependent density functional theory (TDDFT) builds on the same fundamentals as DFT and allows for an analysis of time-dependent properties including optical absorption [8, 9, 10], excitation energies of optical modes [11, 12], and other frequency-dependent response properties.

The Venn diagram shown in Fig. 1.1 serves to illustrate the many facets of DFT, and consequently many of the areas that can be improved upon to increase the speed, expand the scope, and enhance the accuracy of its output. Implementing DFT is not as simple as throwing a crystal system into a black box and instantly obtaining physical results, and approximations occur at every stage, either out of computational necessity or due to unknown quantities like the exact exchange and correlation (xc) functional. The words in bold in Fig. 1.1 highlight elements of the DFT framework tackled in this thesis. Firstly, improved xc functionals are fundamental towards better describing systems that depend strongly on the energy spectrum of the material [7, 13, 14]. Secondly, the choice of basis set for the Kohn-Sham wavefunctions has important consequences for the convergence [15] and accuracy [16, 17] of the iterative solution. Thirdly, the optical response functions link results from TDDFT to fundamental quantities in electromagnetic theory [18] but must be tailored to the dimension of the crystal [19, 20]. Finally, measurable quantities like the

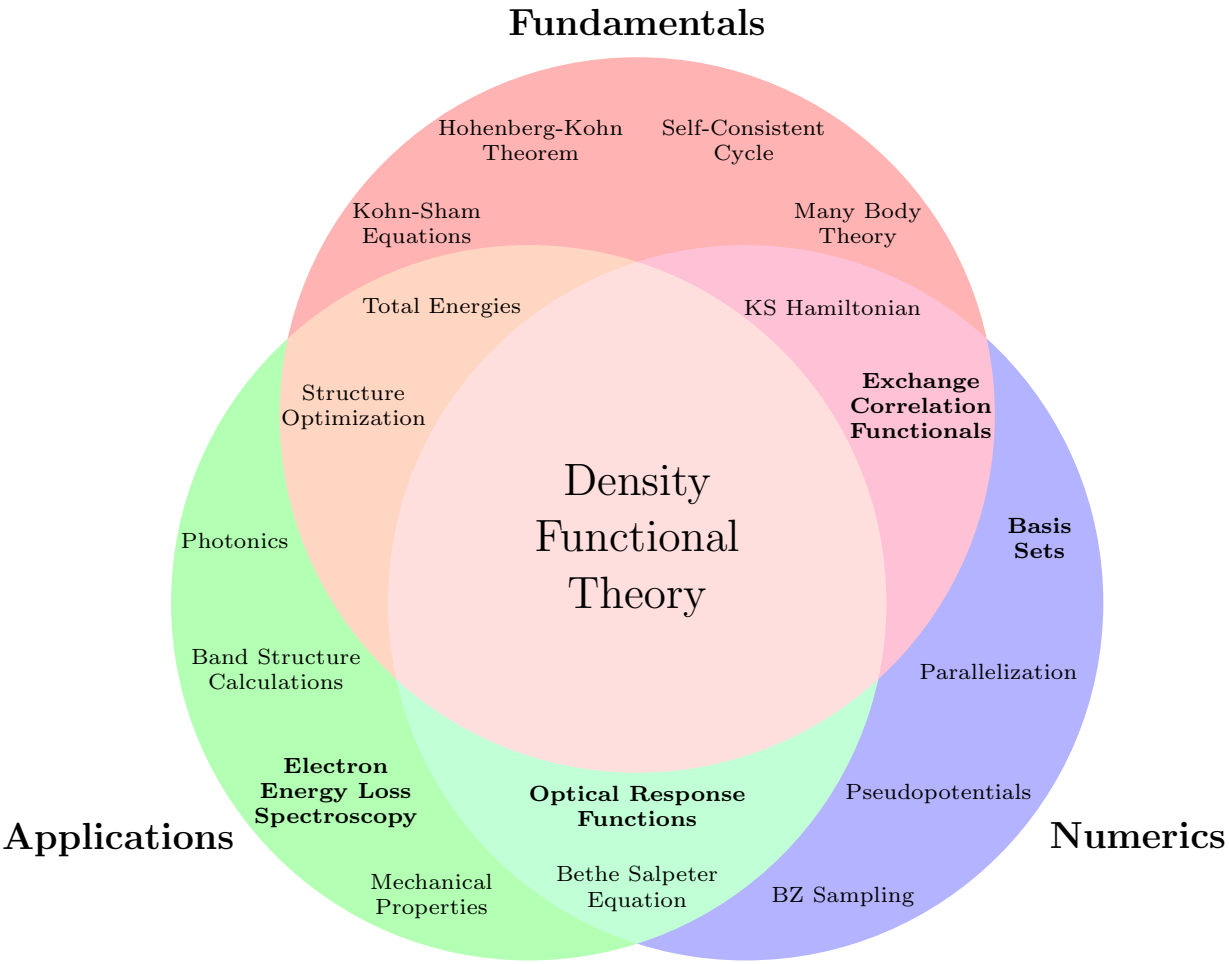

Figure 1.1: A Venn diagram showing the main research directions contained in this work relating to Density Functional Theory. Each circle shows topics that fall within the scope of either fundamentals, numerics, or applications of the DFT framework. Topics that are in bold constitute the focus of this work.

electron energy loss spectra must utilize these response functions while properly accounting for parameters related to the experimental setup [21, 22, 23]. A brief introduction to advances made in each of these areas follows below.

1.1 Density Functional Theory: A Machine Learning Perspective

From a computational standpoint, two major roadblocks in the application of DFT are apparent. The first is that, for atomistic systems with upwards of thousands of atoms, the time and memory

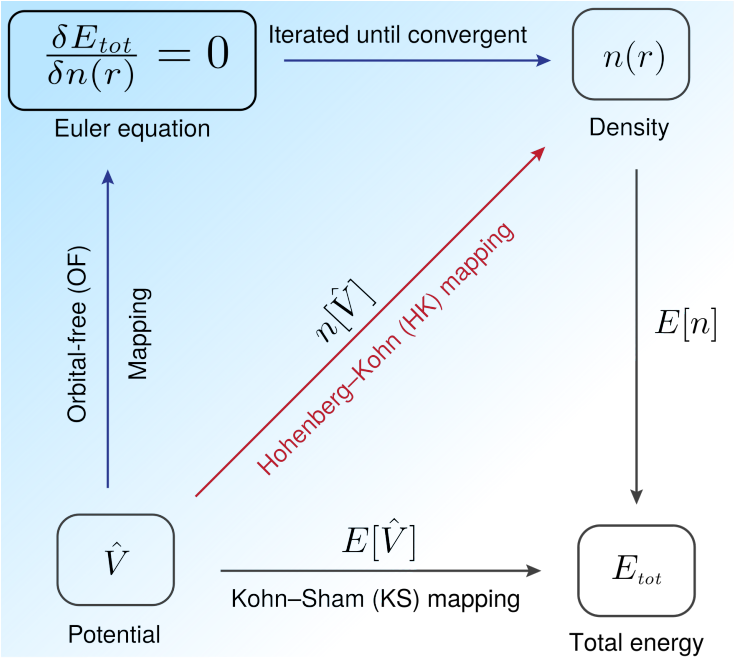

Figure 1.2: Mappings within the DFT framework between the energy minimization (Euler) equation, the potential, the electron density, and the total energy. Most DFT implementations obtain the ground state total energy by solving the Kohn-Sham equations starting from the potential (bottom arrow), although alternative solution methods that rely on the other mappings may be better suited for machine learning applications. Adapted from [24].

costs become too prohibitive for most researchers. Second, higher-tier xc functionals like hybrid functionals [25, 26], that are better suited to describing strongly-correlated electron systems often require on the order of 10 to 100 times the computational resources of conventional DFT xc functionals [27]. It has been suggested that machine-learning techniques could be used to replace aspects of conventional DFT calculations to overcome these barriers, as machine-learned models are known to match the accuracy of the method they are trained on while being less demanding to evaluate.

In deep learning, weights and biases of neurons within a network are trained with large amounts of data to learn hierarchical features without any prior knowledge of the form the mapping is supposed to take. The quality of such networks depends on the quality of the training data available along with a suitable tuning of various hyperparameters and proper pre-processing of the input. In effect, while creating the training data and updating the network are computationally expensive tasks, evaluating the mapping trained by the network is expected to yield accurate results at a fraction of the computational cost. Artificial neural networks (ANNs) that focus on feature engineering for the potential-energy surfaces of molecular systems [28, 29] have gained traction in the chemical physics community, while ANNs that directly represent energy functionals [30, 31, 32], circumvent the usual Kohn-Sham self-consistent implementation [24, 33, 34] (as

shown in Fig. 1.2), or make predictions from densities converged with lower-tier xc functionals about energy gaps calculated with higher-tier xc functionals [35] have all been put forward as novel approaches for determining the electronic and atomic properties of materials more accurately and in a shorter time frame.

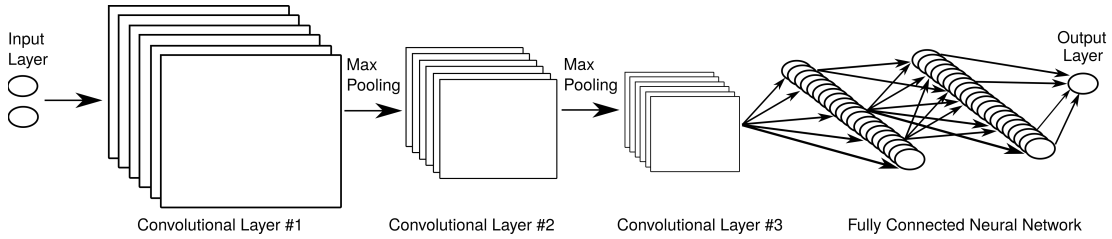

Figure 1.3: Example of a convolutional feed-forward neural network, where tools like convolution and max pooling, often used in series, are used to extract high-level features from the data and pre-process it before it is sent through a fully connected network with weights and biases that are updated with training data.

As many of the approaches set forth remain quite novel, when it comes to eventual integration of these neural network based methods into DFT-based software packages, such as GPAW [36, 37], considerations regarding scalability, universality, and practicality come into play. For example, Brockherde et al. [24] focused on developing an approach which provides a more uniform sampling of the parameter space used for training a neural network model with relatively few parameters, and was shown to be successful in predicting charge densities and total energies for small molecules in 3D and for 1D random Gaussian potentials. However, their use of kernel ridge regression prevents a larger and more diverse set of training points from being employed efficiently. Central to their approach is the training of the Hohenburg-Kohn mapping in Fig. 1.2 to avoid issues related to the reduced parameter space when taking the functional derivative within neural networks [31], but it also prevents an easy and direct integration of the developed network into commonly used DFT software packages. Snyder et al. [30] working in the framework of orbital-free DFT used fully-connected neural networks to predict the interacting kinetic energy of diatomic molecular systems starting from the electron density while Yao et al. [38] accomplished the same for hydrocarbons using a deep convolutional neural network, similar to that displayed in Fig. 1.3, although both methods in their current implementation lack the capacity to predict the full range of properties of most systems. Work by Kolb et al. [32] shows the development of the promising PROPhet software package, which has shown that ANNs can make predictions about energies and band gaps calculated using hybrid functionals such as HSE [35] and quasiparticle methods such as G_0W_0 [7] starting from electron densities calculated with simpler xc functionals, although its applicability in a self-consistent scheme still requires proof-of-concept. One of the more interesting approaches in this new field is given by Ryczko et al. [33, 39], where extensive deep neural networks are applied to images of electron densities and external potentials, replacing all the calculations in the conventional Kohn-Sham DFT scheme via techniques popularized by the image pattern recognition community.

Our work on this topic is motivated by these same issues of scalability, speed and practicality. The Hohenburg-Kohn theorem [1] as stated earlier proves that every physical quantity can be expressed as a functional of the electron density alone. The possibility must then exist for functionals which rely on quantities like the Kohn-Sham orbitals or the kinetic energy, defined implicitly via the electron density, to be approximated to a high degree of accuracy by a neural network which only takes the electron density as input. In contrast to approaches that have only been tested on small molecular systems [24], artificial potentials [34], or two dimensional (2D) systems [39], we are motivated to verify our approach with a class of systems used to benchmark DFT implementations against each other [16, 40, 41]. Analysis of the kinds of networks needed to eventually realize the aforementioned goal provides motivation to "start small" with networks that are capable of first predicting relatively simple but commonly used xc functionals such as the local density approximation (LDA) and the generalized gradient approximation (GGA) as implemented by Perdew, Burke, and Ernzerhof (PBE) [5].

1.2 LCAO-TDDFT-k-ω: Spectroscopy in the Optical Limit

The scalability problem of DFT brought up in the context of machine learning remains ever-present for its time-dependent counterpart TDDFT. Large systems with thousands of atoms and no symmetries that speed up calculations, a common situation when studying biomolecules for example [42, 43], may only reach a convergent state after using approximations unsuitable for studying the optical properties of these systems. For example, optical absorption processes, electron-hole separation, and exciton gemination are of great significance for obtaining accurate optical spectra of materials, but the xc functionals required to describe these are prohibitively expensive for medium sized systems, let alone large macromolecules. In addition, depending on the choice of basis set for the Kohn-Sham wavefunctions, which directly affects the size of the Hamiltonian in the Kohn-Sham eigenvalue problem [44], these large systems may have computational memory requirements beyond the capacity of most researchers. The motivation here is clear: there is a need for a robust and efficient method for determining the optical properties of systems such as biomolecules while retaining the accuracy of computationally unfeasible methods. The code we have developed within the DFT-based software package GPAW [36], LCAO-TDDFT-k-ω [45], has been shown to accomplish this for systems such as chlorophyll monomers [46] and single-walled carbon nanotubes (SWCNTs) [9], with the greatest cost savings of our implementation coming for low-dimensional materials.

When it comes to choosing a TDDFT approach that is most appropriate for large macromolecular systems, the pros and cons of each must be weighed carefully. For example, TDDFT implemented in real space (RS) and frequency domains, TDDFT-r-ω [47], scales as $O(N^5)$ where N is the number of Kohn-Sham wavefunctions, which itself scales with the number of atoms in the system. Likewise, RS-TDDFT-k-ω [48] has high memory costs associated with the size of the grid needed for convergent calculations. RS-TDDFT-r-t [49] alternatively works in the time domain, but stability of this method requires time steps much smaller than needed from alter-

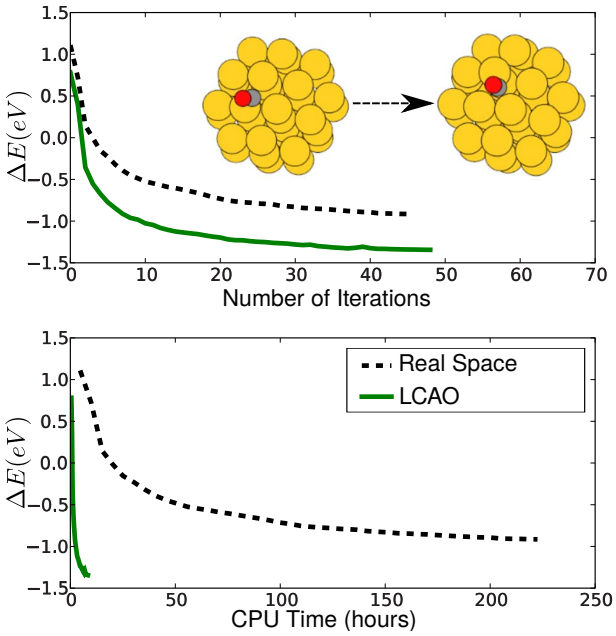

Figure 1.4: The energy of a 38-atom truncated octahedral gold cluster with CO adsorbed, showing the energy as a function of both the number of iterations (top) and CPU time (bottom) for structure optimizations using real space grids (black dashed) and LCAO (green solid) in GPAW. Adapted from [15].

nate approaches [50]. A real space basis set representation is oftentimes inappropriate for the systems of interest. Although plane wave (PW) allows for improved stability without the large grid sizes, high memory costs again become a problem when very large unit cells are employed, the exact type needed to model 0D materials such as macromolecules. A linear combination of atomic orbitals (LCAO) representation [15, 51] has the lowest memory costs of the three representations. The compactness of the basis helps with convergence, while working in reciprocal and frequency steps avoids the pitfalls related to time steps and scaling present in methods like TDDFT-r-ω. Fig. 1.4 indicates just how big a speedup the use of the LCAO basis set has compared to grid-based methods for structural optimization. The very compactness of the LCAO basis set [17, 52] may, however, cause a poor description of certain physical properties. Working with LCAO instead of PW also means that the fast Fourier transform (FFT) that generally makes the PW representation quite efficient is unavailable, so calculations require that optical response functions are calculated purely in the optical limit (i.e. zero wavenumber).

While LCAO-TDDFT-k-ω seems to help resolve the memory and time bottleneck for low-dimensional systems, our motivation also stems in having our robust and efficient method be accurate. Empirical methods that essentially splice the difference between experimental or G_0W_0-BSE band gaps onto the unoccupied Kohn-Sham eigenenergies calculated within TDDFT may prove useful as a quick fix, but lack the predictive quality that DFT is built for. The derivative discontinuity correction to the Gritsenko, van Leeuwen, van Lenthe, Baerends solid corrected

(GLLB-SC) exchange functional [53, 54, 55] provides a first-principles method for calculating this difference, avoiding *ad hoc* changes to the band gap and hence the profile of the optical spectra. Trying to understand how exactly this correction should be introduced into the spectra provides motivation for testing this approach over a wide class of materials. This testing is further motivated by our need to benchmark our LCAO-TDDFT-k-ω code to ensure that the issues related to convergence of the LCAO basis set can be mitigated and that its application to materials of any dimension is justified. Doing so should demonstrate the reliability and applicability of the LCAO-TDDFT-k-ω approach integrated with the GLLB-SC correction for determining the optical properties of the kinds of systems poorly or inefficiently described by other TDDFT methods.

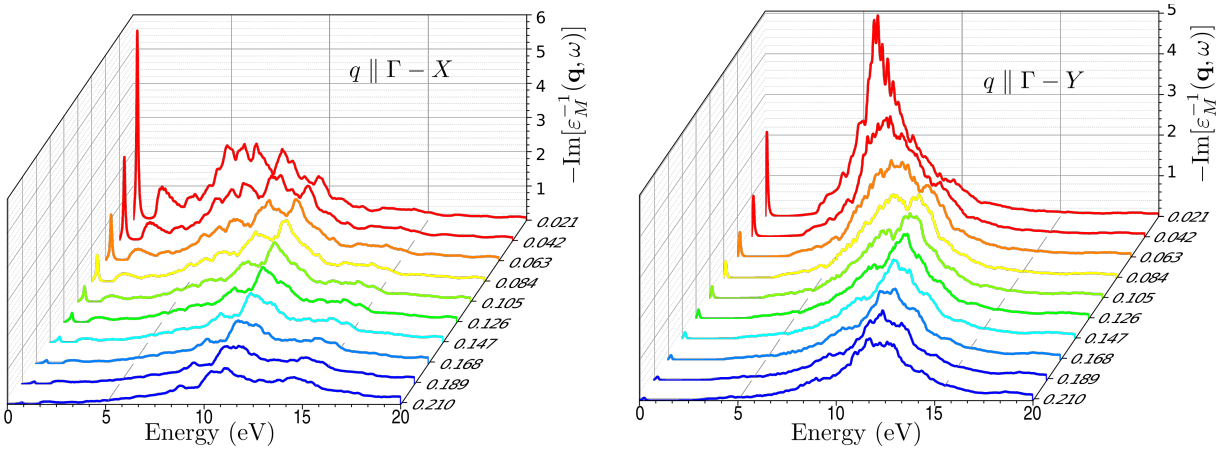

Figure 1.5: The energy loss function $-\operatorname{Im}\varepsilon_M^{-1}(\mathbf{q},\omega)$ of doped monolayer phosphorene in directions parallel to the x- and y-axes from PW-TDDFT-k-ω, showing the difficulty in calculating the optical limit $\lim_{\mathbf{q}\to 0}\varepsilon(\mathbf{q},\omega)$. For atomically thin materials, the energy loss of a fast-moving electron in an EELS setup should take into account the dimensionality of the system, the aperture size, and the importance of relativistic effects. Adapted from [56].

1.3 Conductivity Models and the Electron Energy Loss of Graphene Monolayers

For many experimentalists, DFT is only as useful as its ability to make reliable predictions and reproduce what is measured in the lab. While the standard output of most TDDFT codes is the 3D dielectric function in terms of wavenumber $\mathbf{q}$ and frequency ω, and methodologies exist for plugging in this *ab initio* output to obtain estimates for the observed energy loss in experimental setups, refinements are required to properly account for the different electromagnetic properties of 2D materials, with graphene and phosphorene (see Fig. 1.5) being two well-known examples. The dimensionality of the monolayer can be accounted for directly in the TDDFT calculation

stage [18, 20] using a mean-field approach [19], although in the case of graphene, analytic models for the conductivity have been derived from either a tight-binding approach [57] or empirically [22, 58] after being fit to *ab initio* data.

From both a theoretical and computational standpoint, interactions of monolayer materials with externally moving charged particles have been explored thoroughly in the literature. For example, Electron Energy Loss Spectroscopy (EELS) in a Transmission Electron Microscope (TEM) setup has been used to study plasmon excitations in small-scale structures over a wide frequency range [59] and to probe energy losses in the $\hbar\omega \gtrsim 1$ eV range for monolayer graphene [60, 61, 62], while the addition of aberration correctors [63] to the setup has provided access to the region of mid-infrared energy excitations [64, 65]. Graphene in particular has received a good share of attention, with ample literature regarding the measurements of terahertz (THz) radiation [66, 67, 68], transition radiation (TR) [69], and the plasmonic inverse Doppler effect [70] resulting from interactions of a graphene sheet with fast moving electrons. Given the breadth of literature on the subject, we are motivated to consider how differences in experimental setups, such as the size of the collection aperture, the velocity of incoming electrons, and the range of energies considered, influence the spectra.

Results for the fully relativistic energy loss have been explored for fast electrons incident at an oblique angle [71] and at a normal angle [21, 22] for a single graphene monolayer as well as for multilayer graphene under normal incidence [72, 73, 74]. All these approaches treat the graphene monolayer as a boundary layer in the context of the macroscopic Maxwell's equations described by a conductivity which is purely frequency dependent, either from TDDFT or from empirical models as mentioned above. Such an approach allows for the derivation of analytic forms for the electron energy loss taking only the two-dimensional conductivity and certain experimental parameters as input [21, 22, 58]. These simplifications all allow for a thorough analysis of the different energy loss mechanisms available to incident electrons across a wide range of frequencies. We are in this way motivated to explore how well different models for both the energy loss and in-plane conductivities fare against experimental results from the literature [12, 60, 62, 75].

1.4 Outline of Thesis

This thesis is divided as follows: Chapter 2 provides necessary theoretical background, including discussion and equations related to density functional theory (DFT) and its time-dependent implementation (TDDFT), electron energy loss spectroscopy (EELS), and artificial neural networks (ANNs). Chapter 3 provides details related to the low-dimensional materials under consideration, with special emphasis on the optical properties of the graphene monolayer. Chapter 4 presents the results of our efforts to use a feed-forward ANN to learn the commonly used LDA and PBE xc functionals, showing their performance in modeling physical systems and outlining major challenges any ANN approximation to these quantities is expected to face. Chapter 5 demonstrates the performance of our LCAO-TDDFT-k-ω code, a TDDFT implementation with considerable computational efficiency optimized for modeling the optical absorption of low-dimensional and

large macromolecular systems, providing benchmarking results for materials of every dimension and presenting visualization tools we developed unique to our implementation. Chapter 6 compares our phenomenological model for the conductivity of graphene to that computed using TDDFT methods, and contrasts EELS derived from both experimental spectra before analyzing how the experimental setup and theoretical assumptions affect the profile of the energy loss. Chapter 7 presents a summary of results and future directions for expanding the scope of the research outlined in this thesis.

Chapter 2

Theoretical Background

2.1 Density Functional Theory

In the context of condensed matter physics, all materials can be characterized as systems of interacting electrons and ions. The full many-body Hamiltonian, excluding purely quantum mechanical terms like electron spin and angular momentum coupling, can be written as

$$\mathcal{H} = \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_I \frac{\mathbf{P}_I^2}{2M_I} - \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R_J}|}, \tag{2.1}$$

where ions with charge $Z_I e$ and mass M_I are located at positions $\mathbf{R}_I$ carrying momenta $\mathbf{P}_I$ and electrons with charge $-e$ and mass m are located at positions $\mathbf{r}_i$ while carrying momenta $\mathbf{p}_i$. We note that atomic units are used in all the following equations.

The first two terms in the Hamiltonian correspond to the kinetic energy of the electrons and ions, respectively. The third, fourth and fifth terms in the Hamiltonian denote the electron-ion, electron-electron and ion-ion interactions, respectively. The ground state of the system, corresponding to a solution of the time-independent Schrödinger's equation $\mathcal{H}\Psi(\mathbf{R}_I, \mathbf{P}_I, \mathbf{r}_i, \mathbf{p}_i) = E\Psi(\mathbf{R}_I, \mathbf{P}_I, \mathbf{r}_i, \mathbf{p}_i)$, can theoretically be obtained by minimizing the total energy (the expectation value of the Hamiltonian) with respect to all possible electronic and ionic degrees of freedom. However, when considering systems with even a few ions and electrons, solving this equation analytically becomes intractable. This chapter presents density functional theory (DFT) as one of the most successful methods used to determine the quantum mechanical properties of systems that can be well described by the Hamiltonian above.

The first candidate for simplification of the problem comes about from the large difference between electron masses m and ion masses M_I of about three orders of magnitude. It is therefore often assumed that the electrons are moving much faster than the ions do and thereby respond to their movement close to instantaneously. Known as the Born-Oppenheimer approximation

[76], this effectively decouples the ionic and electronic degrees of freedom, so the ionic positions R_I can be considered classical and therefore fixed. Both the kinetic energy of the ions and the last term in Eq. 2.1 can be safely neglected when determining the electronic portion of the wavefunctions, being constant for a set configuration of ions. It is, however, needed to determine the correct total energy of moving electrons and fixed ions. The many-body Hamiltonian of N interacting electrons in a solid can finally be written in the form

$$\mathcal{H} = \sum_i \frac{\mathbf{p}_i^2}{2m} - \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.2}$$

The solution of this eigenvalue problem is still a daunting task - the next approach examined towards solving the eigenvalue problem for the Hamiltonian in Eq. 2.2 for the ground state of the electrons is known as Hartree-Fock (HF) theory [77], where the N-electron ground state wavefunction of a non-relativistic system is represented by the Slater determinant of N one-electron spin-orbitals,

$$\Psi_0(\mathbf{r}_1 \cdots \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_n} \left(sgn(\sigma) \prod_{i=1}^{N} \psi_i(\mathbf{r}_{\sigma_i}) \right), \tag{2.3}$$

where the sum is over all permutations σ in the symmetric group of n elements S_n in order for the resulting wavefunction product to properly take into account the correct anti-symmetry character in the spatial coordinates (also in spin although the indices are repressed), noting that this approach does neglect correlation of the electrons by effectively decoupling all electrons in the system.

The variational principle states that the N-electron wavefunction describing our quantum mechanical system is obtained by minimizing the total energy $E_0 = \langle \Psi_0 | \mathcal{H} | \Psi_0 \rangle$ with respect to the constituent single-electron wavefunctions ψ_i under the constraint that the ψ_i form an orthonormal set (i.e. $\langle \psi_i | \psi_j \rangle = \delta_{ij}$). The resulting system of non-linear integro-differential equations, known as the canonical HF equations [77], are written as

$$\left[\frac{\mathbf{p}^2}{2m} + \hat{V}_{\text{nuclear}} + \hat{V}_{\text{coulomb}} + \hat{V}_{\text{exchange}} \right] \psi_i = \varepsilon_i \psi_i \tag{2.4}$$

$$\hat{V}_{\text{coulomb}} \psi_i(\mathbf{r}) = e^2 \sum_j^{\text{occ}} \psi_i(\mathbf{r}) \int \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}') d\mathbf{r}' \tag{2.5}$$

$$\hat{V}_{\text{exchange}} \psi_i(\mathbf{r}) = -e^2 \sum_j^{\text{occ}} \psi_j(\mathbf{r}) \int \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}') d\mathbf{r}' \tag{2.6}$$

where the Lagrange multipliers ε_i result from the orthonormality constraint, the nuclear potential in Eq. 2.2 operates on the one-electron wavefunction in a trivial way, and the sums run over

all states occupied by electrons. The so-called exchange potential shown in these equations is a non-local integral operator, implying that the solutions ψ_i and the corresponding eigenvalues ε_i need to be determined self-consistently. All in all, the many-body ground state problem has been simplified down to a set of single-particle equations reminiscent of the time-independent Schrödinger equation.

Understanding that the auxillary wavefunctions and HF eigenenergies were obtained through a variational method, it is clear that the HF ground state energy E_0 will always be higher than the true ground state energy of the system, which can be attributed to the neglect of correlation effects in the above equations, both between electrons and their spins, the Slater determinant proving insufficient to model the complex many-body wavefunction. While methods exist for the exact calculation of the correlation energy through configuration interaction methods utilizing a weighted sum of Slater determinants [78], the computational effort required makes this intractable for most systems. The next section shows that density functional theory can take correlation effects of the quantum mechanical system under study into consideration in a computationally feasible way.

2.1.1 Hohenberg-Kohn Theorem and the Kohn Sham Ansatz

The many body Hamiltonian of interacting electrons written in Eq. 2.2 can be written as

$$\hat{H} = \hat{T} + \sum_i \hat{v}_{\text{ext}}(\mathbf{r}_i) + \hat{V}_{\text{coulomb}} \tag{2.7}$$

composed of the kinetic potential, the external potential, and the electron-electron interaction, or Coulomb potential. As mentioned before, although the electronic degrees of freedom being decoupled makes the problem more tractable, this operator will still be acting on a wavefunction $\Psi(\mathbf{r}_i, \mathbf{p}_i)$ of the positions and momenta of all electrons in the system.

The Hohenberg-Kohn Theorem [1] states that knowledge of the external potential $\hat{V}_{\text{ext}}(\mathbf{r}) = \sum_i \hat{v}_{\text{ext}}(\mathbf{r}_i)$, taking into account the atomic structure of our system, is sufficient to determine exactly the ground-state electronic density $n(r)$ of the system. More precisely, the theorem states that there exists a functional that links the external potential to the ground state electronic density of the system. Likewise, it can be shown that the ground state density determines the external potential $\hat{V}_{\text{ext}}$ uniquely up to a constant, and consequently the entire full Hamiltonian in Eq. 2.1, from which the eigenenergies and eigenfunctions of the many-electron system are derived [77]. While this electronic density $n(r)$ is not known *a priori*, a variational method can be applied allowing for its self-consistent calculation. Ultimately, the Hohenberg-Kohn Theorem shows that the total energy of the system in its ground state will be determined by the external potential $\hat{V}_{\text{ext}}$ defined as a functional of the electronic density. In functional form,

$$E_{GS}\left[n(\mathbf{r}), \hat{V}_{\text{ext}}(\mathbf{r})\right] = T[n(\mathbf{r})] + E_{\text{coulomb}}[n(\mathbf{r})] + \int \hat{V}_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \tag{2.8}$$

The first two terms depend explicitly on the ground state density of the interacting system. It can be shown that the kinetic energy of a non-interacting electron gas in many cases is already a good approximation for that of the fully interacting system, while the electron-electron interaction is also well represented by the classical Coulomb interaction [79]. Eq. 2.8 can therefore be reformatted as

$$E_{GS}\left[n(\mathbf{r}), \hat{V}_{\text{ext}}(\mathbf{r})\right] = T_0[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] + \int \hat{V}_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \tag{2.9}$$

where

$$T_0[n(\mathbf{r})] = -\frac{1}{2}\sum_{i=1}^{N_e/2} \langle \psi_i^*(\mathbf{r}) | \nabla^2 | \psi_i(\mathbf{r}) \rangle \tag{2.10}$$

and the classical long-range Hartree energy is defined as

$$E_H[n(\mathbf{r})] = \frac{1}{2}\int\int n(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \tag{2.11}$$

The last remaining term in Eq. 2.9, called the exchange-correlation (xc) functional [80], includes all interaction terms between the electrons of the fully interacting system and is defined as

$$E_{xc}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_0[n(\mathbf{r})] + E_{\text{coulomb}}[n(\mathbf{r})] - E_H[n(\mathbf{r})]. \tag{2.12}$$

The second key component of the theoretical underpinning of DFT lies in the Kohn-Sham (KS) ansatz [79], which posits that the electron density for the original fully interacting system can be represented by an artificial collection of non-interacting particles, allowing for independent equations for these artificial particles where all many-body effects are collected into the xc functional of the density in Eq. 2.12. This representation can be understood as a mapping from the interacting electron problem to one where a collection of non-interacting electrons act according to some effective potential

$$\hat{V}_{\text{eff}}(\mathbf{r}) = \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}_H(\mathbf{r}) + \hat{V}_{\text{xc}}(\mathbf{r}), \tag{2.13}$$

where

$$\hat{V}_{\text{xc}}(\mathbf{r}) = \frac{\partial E_{xc}[n(\mathbf{r})]}{\partial n(\mathbf{r})} \tag{2.14}$$

and the electron density is decomposed into a weighted sum of orthogonal one-electron wavefunctions

$$n(\mathbf{r}) = \sum_{i=1}^{N} \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r}), \tag{2.15}$$

where N represents the number of occupied states. We can apply a variational procedure on the energy in Eq. 2.9 with respect to the wavefunctions of our non-interacting artificial system ψ_i,

known as the KS wavefunctions. This results in the KS equations

$$\left(-\frac{\nabla^2}{2} + \hat{V}_{\text{eff}}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \tag{2.16}$$

using that $\mathbf{p}_i = i\vec{\nabla}$. Altogether, the original many-electron problem has now been replaced by an auxiliary non-interacting particle problem, and the input for the operator on these KS wavefunctions now depend only on the electron density instead of the collective spatial and momentum profiles of every electron in the system. All issues related to the complex interacting many-electron problem, including the indistinguishability of electrons and other correlation effects, are left to the determination of a reasonable xc potential.

2.1.2 Exchange-Correlation Functionals

We start with the simplest and most well-known of the xc functionals, known as the local density approximation (LDA) [4]. We first decompose the xc total energy via

$$E_{xc}^{\text{LDA}}[n] = E_x^{\text{LDA}}[n] + E_c^{\text{LDA}}[n] = \int \varepsilon_{xc}^{\text{LDA}}[n]n(\mathbf{r}')d\mathbf{r}', \tag{2.17}$$

where $\varepsilon_{xc}^{\text{LDA}}[n]$ represents the xc energy per particle, known as the xc energy density functional. For the LDA, there is an explicit formula for the xc potential shown in Eq. 2.14

$$\hat{V}_{xc}^{\text{LDA}}(\mathbf{r}) = \frac{\delta E_{xc}^{\text{LDA}}[n]}{\delta n(\mathbf{r})} = \varepsilon_{xc}^{\text{LDA}}(n(\mathbf{r})) + n(\mathbf{r})\frac{\partial\varepsilon_{xc}^{\text{LDA}}(n(\mathbf{r}))}{\partial n(\mathbf{r})}. \tag{2.18}$$

It is this function that contributes to the effective potential $\hat{V}_{\text{eff}}(\mathbf{r})$ defined in Eq. 2.13 for the KS Hamiltonian, solved to give the KS wavefunctions. This explicit formula becomes increasingly complicated as the range of functional parameters increases, to the point of being undefined for xc functionals with explicit dependence on quantities that depend implicitly on the electron density $n(\mathbf{r})$.

While LDA formally refers to any xc functional that only depends on the density, its most natural implementation is motivated by physical considerations. Specifically, since the LDA approximation assumes dependence only on the local electron density, $\varepsilon_x^{\text{LDA}}$ and $\varepsilon_c^{\text{LDA}}$ should converge to analytic values for the xc energy density of the homogeneous electron gas (HEG) [1], an artificial material where positive nuclei (and therefore the electron density) are assumed to be evenly distributed in space. An analytic formula for $\varepsilon_x^{\text{LDA}}$ can then be derived and goes as

$$\varepsilon_x^{\text{LDA}}[n(\mathbf{r})] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\mathbf{r})^{\frac{1}{3}}, \tag{2.19}$$

while $\varepsilon_c^{\text{LDA}}$ can be approximated by treating the HEG model perturbatively. Thanks to this bench-

mark and its success in describing many different systems, LDA serves as a starting point for more complicated functionals, many of which are defined as correction terms to this well-tested xc functional. LDA does however have many shortcomings as an xc functional. For example, as a function of space the LDA potential decays asymptotically as an exponential, whereas the exact xc potential should theoretically decay in a slower Coulomb-like manner, given that it describes the correlation effects between electrons interacting via the Coulomb potential [81]. This can cause molecular systems with physically stable solutions to dissociate in a DFT calculation implementing LDA, as the number of bound KS orbitals is underestimated and Rydberg states and ionization potentials are ill-described [4]. Much of the success of the LDA xc functional lies in fortuitous error cancellation [81], as the correlation energy E_c is underestimated and the exchange energy E_x is overestimated. Ultimately, LDA proves to work best when the materials described most resemble the HEG, like bulk metals, but for a larger class of systems it is worthwhile to investigate more complicated xc functionals.

Perdew, Burke, and Ernzerhof (PBE) [5] developed one of the most well-known of the generalized gradient approximation (GGA) functionals, a family of functionals defined by

$$E_{xc}^{\mathrm{GGA}}[n,\nabla n] = E_{xc}^{\mathrm{LDA}}[n] + \left(E_{xc}^{\mathrm{GGA}}[n,\nabla n] - E_{xc}^{\mathrm{LDA}}[n]\right) = \int \varepsilon_{xc}^{\mathrm{GGA}}[n,\nabla n]n(\mathbf{r}')d\mathbf{r}', \tag{2.20}$$

now including the gradient of the electron density as an additional parameter in the energy density functional. The formula for the xc potential changes accordingly to

$$\begin{aligned}\hat{V}_{xc}^{\mathrm{GGA}}(\mathbf{r}) &= \frac{\delta E_{xc}^{\mathrm{GGA}}[n,\nabla n]}{\delta n(\mathbf{r})} \\ &= \varepsilon_{xc}^{\mathrm{GGA}}(n(\mathbf{r}),\nabla n(\mathbf{r})) + n(\mathbf{r})\frac{\partial \varepsilon_{xc}^{\mathrm{GGA}}(n(\mathbf{r}),\nabla n(\mathbf{r}))}{\partial n(\mathbf{r})} - \nabla\cdot n(\mathbf{r})\frac{\partial \varepsilon_{xc}^{\mathrm{GGA}}(n(\mathbf{r},\nabla n(\mathbf{r})))}{\partial \nabla n(\mathbf{r})}.\end{aligned} \tag{2.21}$$

As with LDA, the form of the GGA xc functional is rooted in certain natural constraints that a dependence on the gradient of the electron density introduces. For example, the functional can be chosen such that the correction term $\left(E_{xc}^{\mathrm{GGA}}[n,\nabla n] - E_{xc}^{\mathrm{LDA}}[n]\right)$ goes to zero in the limit of $\nabla n \to 0$, or that it is exact for a reference system like a helium atom. Since atoms and small molecular systems are farthest from the HEG and have strong variation in the term ∇n, good GGA implementations will fare better in describing the properties of these systems [16, 79]. Further constraints and limits can be imposed depending on the type of system we are modeling.

Next in the hierarchy of xc functionals comes the meta-GGA methods [82], so called because they depend on the occupied KS orbitals implicitly through a kinetic energy density labelled $\tau(\mathbf{r})$

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_i [\nabla\psi_i(\mathbf{r})]\cdot[\nabla\psi_i(\mathbf{r})]. \tag{2.22}$$

These meta-GGAs can be formulated to follow theoretical constraints as in the typical LDA implementation, or by following a more empirical approach that fits parameters in the model based on large molecular databases. To determine the xc potential for a meta-GGA theoreti-

cally involves a computation of $\delta\tau(\mathbf{r})/\delta n(\mathbf{r}')$ and likewise $\delta\psi_i(\mathbf{r})/\delta n(\mathbf{r}')$. These quantities are theoretically well-defined, as the Hohenburg-Kohn theorem guarantees that knowledge of $n(\mathbf{r})$ is sufficient to reconstruct the KS potential, which is consequently used to find the KS orbitals $\psi_i[n](\mathbf{r})$ and their energies through a constrained search. In practice, an iterative procedure using the inverse KS equations (for example, the orbital-averaged one shown in Eq. 2.37) is run until a convergent electron density and set of orbitals is generated.

The next and final approximate xc functional under consideration in the hierarchy are hyper-GGAs. Such xc functionals, on top of taking the electron density and its gradient as input, use the KS orbitals as input directly, in contrast to their indirect use in the kinetic energy part of the meta-GGAs [82]. These two approaches together fall under the umbrella term of hybrid functionals. One family of commonly used hyper-GGAs [83] include HF exchange terms, entering the xc functional as

$$E_x^{\text{exact}}[n,\phi] = -\frac{1}{2}\sum_{jk}\iint d\mathbf{r}d\mathbf{r}'\frac{\psi_j^*(\mathbf{r})\psi_k^*(\mathbf{r}')\psi_k(\mathbf{r})\psi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.23}$$

As with the meta-GGA, functional derivatives in closed form are no longer possible. Approximations exist for determining the potentials $\hat{V}_{xc}$ for these hyper-GGA methods [79, 83], including a variation on the Slater potential method [84] for the exact xc functional, but the computational power required to implement these hybrid functionals in addition to the regular self-consistent DFT cycle limits the use of these costly functionals.

For functionals that can be expressed purely in terms of the electron density, we can utilize the framework developed by Levy and Zahariev [85], where

$$\hat{V}_{xc}^*([n],\mathbf{r}) = \hat{V}_{xc}([n],\mathbf{r}) + c[n] = \hat{V}_{xc}([n],\mathbf{r}) + \frac{E_{xc}[n] - \int \hat{V}_{xc}([n],\mathbf{r})n(\mathbf{r})d\mathbf{r}}{\int n(\mathbf{r})d\mathbf{r}} \tag{2.24}$$

where $c[n]$ is an additive constant that depends only on the electron density. Multiplying this equation by $n(\mathbf{r})$ and integrating over space yields

$$E_{xc}[n] = \int \hat{V}_{xc}^*([n],\mathbf{r})n(\mathbf{r})d\mathbf{r}. \tag{2.25}$$

This form can also be used to reformulate the KS equations in Eq. 2.16 as

$$\left(-\frac{\nabla^2}{2} + \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}_{\text{Coulomb}}(\mathbf{r}) + \hat{V}_{xc}^*(\mathbf{r})\right)\psi_i(\mathbf{r}) = (\varepsilon_i + c[n])\psi_i(\mathbf{r}) = \varepsilon_i^*\psi_i(\mathbf{r}), \tag{2.26}$$

where multiplying on the left by $\psi_i^*(\mathbf{r})$, summing over i, and integrating over space gives an expression equivalent to Eq. 2.9, yielding

$$E_{GS} = \sum_i \varepsilon_i^*. \tag{2.27}$$

Ultimately, this provides an alternate method to relate the xc energy and a potential used in a variant of the KS equation, one where knowledge of the potential $\hat{V}^{*}_{xc}([n], \mathbf{r})$ can be used to construct the xc energy functional rather than the original method where the functional derivative of the xc energy function is taken to determine $\hat{V}_{xc}([n], \mathbf{r})$. This method also has the advantage that the ground state energy can be directly computed from the sum of these adjusted KS eigenenergies.

From a computational standpoint, it is often convenient to express the first and second gradients of the electron density as dimensionless parameters

$$\varrho = n^{-4/3}|\nabla n| \qquad \vartheta = n^{5/3}\left(\frac{|\nabla n|^2}{n} - \nabla^2 n\right). \tag{2.28}$$

For many implementations of GGA, including the PBE exchange potential [5], the Becke asymptotic exchange potential [25], and the Lee-Yang-Parr correlation potential [83], the regions where these dimensionless parameters are near unity provide the greatest variation in the output values for the total xc energy, so efforts to replicate these kinds of functionals will also need to capture the extent to which changes in ϱ and/or ϑ affect the xc energy.

2.1.3 The GLLB-SC xc functional

The xc functional hierarchy presented above involves successive corrections and additional parameters included in the computation of E_{xc} and $\hat{V}_{xc}$ to try to obtain a converged electron density as close as possible to the exact one. A good approximation means that the resulting eigenenergies from the self-consistent solution of the KS equations will result in valence and conduction bands resembling experimentally measured ones, but it is the case in certain schemes that the band gap calculated with these methods is not in agreement with experiment. One reason for this is attributed to the derivative discontinuity of the xc potential [53], where the xc potential jumps for integer occupations of electrons. More precisely, for an N-electron system the xc potential is described by

$$\hat{V}_{xc}(\mathbf{r}, N) = \left.\frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}\right|_{N}, \tag{2.29}$$

which will be continuous for fractional numbers of electrons, but at integer number results in a discontinuity in the energy levels, meaning for this kind of system the true KS-DFT quasiparticle band gap is given by

$$E_g = E_g^{KS} + \Delta_x = \varepsilon_{N+1} - \varepsilon_N + \left\langle \lim_{\delta\to 0}\left[\hat{V}_{xc}(\mathbf{r}, N+\delta) - \hat{V}_{xc}(\mathbf{r}, N-\delta)\right]\right\rangle \tag{2.30}$$

where ε_N represents the N^{th} KS eigenenergy and the chevrons emphasize the fact that the discontinuity is a constant function of $\mathbf{r}$. While several computational tools, including the optimized effective potential method, can be used to compute this discontinuity [55], many turn out to be complicated and computationally demanding and therefore unsuitable for larger systems which require more robust methods [54].

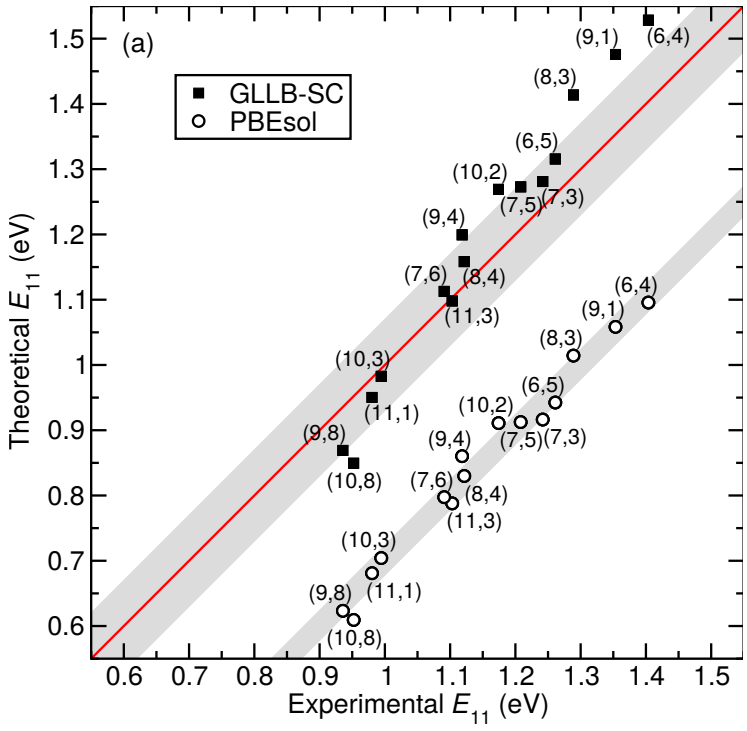

Figure 2.1: Theoretical versus experimental E_{11} transition energies in eV for 15 different chiral semiconducting single-walled carbon nanotubes, E_{11} being the energy of the first bright excitonic peak. Optical absorbance is calculated using the GPAW LCAO-TDDFT-k-ω software package. Filled squares are for eigenenergies calculated with the GLLB-SC derivative discontinuity correction Δ_x and open circles for energies calculated using only PBEsol. Experimental data is from the optical absorbance and electron energy loss measurements of Refs. 23 and 24, respectively. The average errors for GLLB-SC ($\epsilon \approx 0 \pm 70$ meV) and for PBEsol ($\epsilon \approx -300 \pm 3$ meV) transitions are shown as grey regions. The red line, provided to guide the eye, shows the extent to which the correction Δ_x improves the description of the band gap. Adapted from [9]

The GLLB-SC method [55] is one such robust method, requiring only a single calculation with an already-minimized electron density (known as a single-point or Harris calculation) to yield the derivative discontinuity Δ_x. This is a refinement on the original GLLB method, which would run the calculation on a system converged with a GGA xc functional best suited for smaller atoms, by instead running with a system converged with PBEsol, a GGA xc functional suited for solids and larger systems. The GLLB-SC works by adding to the exchange energy functional for the system a simple orbital-weighted approximation that yields the desired discontinuity of the potential at integer particle numbers

$$\Delta_x = \langle \Psi_{N+1} | \Delta_{x,resp}(\mathbf{r}) | \Psi_{N+1} \rangle. \tag{2.31}$$

$$\Delta_x(\mathbf{r}) = \frac{8\sqrt{2}}{3\pi^2}\sum_i^N \left(\sqrt{\varepsilon_{N+1}-\varepsilon_i} - \sqrt{\varepsilon_N-\varepsilon_i}\right)\frac{|\psi_i(\mathbf{r})|^2}{n(\mathbf{r})}, \tag{2.32}$$

where the reference energy ε_i takes on the value of the highest occupied molecular orbital (HOMO) for a given integer occupation. This corrects in an approximate way for the derivative discontinuity problem present in the band gap difference for the material in question by allowing for a rigid upshift in the KS eigenenergies for unoccupied states. As an example, Fig. 2.1 shows just how well the GLLB-SC correction works in matching theoretical and experimental predictions for the first bright excitonic peak for a collection of 15 semiconducting nanotubes.

Some of the success of the derivative discontinuity correction towards describing the quasiparticle band gap can be attributed to its linear correlation with the inverse dielectric constant ε_∞^{-1}, the static limit of the macroscopic dielectric function shown in Eq. A.5, as presented in Fig. 2.2 [9]. Previous studies have shown that the quasiparticle (QP) correction to the electronic band gap $E_g = E_g^{KS} + \Delta_{QP}$, calculated through methods like G_0W_0-BSE [14, 86], is also directly correlated with the inverse dielectric constant of the material. This relates to the static dielectric constant describing the effective screening of electron levels which then renormalizes the band gap correction according to the adjusted eigenenergies [86], as explained in Section 2.2.7. It stands to reason that a linear relationship may exist between the corrections calculated through the two methods, and that they both provide reasonable estimations for the true quasiparticle band gap.

2.1.4 Inverse Kohn-Sham Equations

While the types of xc functionals that can be chosen are abundant, it is important to understand exactly how knowledge of a set of canonical KS orbitals, orbital energies, and an external potential for the many-electron system can actually be used to determine this potential. However, care must be taken depending on whether $\hat{V}_{\text{xc}}(\mathbf{r})$ depends solely on the electron density or, as is done for more complex schemes that are used to better describe electron correlation effects, on the KS orbitals themselves.

One option for obtaining an expression for $\hat{V}_{\text{eff}}(\mathbf{r})$ is to simply rearrange the terms in Eq. 2.16

$$\hat{V}_{\text{eff}}(\mathbf{r}) = \frac{1}{2}\frac{\nabla^2\psi_i(\mathbf{r})}{\psi_i(\mathbf{r})} + \varepsilon_i. \tag{2.33}$$

While this is formally valid for each orbital $\psi_i(\mathbf{r})$, in practice it could used only for a nodeless orbital, which would occur in the lowest eigenvalue state, and could be successfully employed for studying the exact xc potential for one- or two-electron systems, but in calculations for many-electron systems using finite basis set representations for the orbitals it leads to severe numerical difficulties [87]. Instead, starting with Eq. 2.16, we multiply both sides on the left by $\psi_i^*(\mathbf{r})$, sum

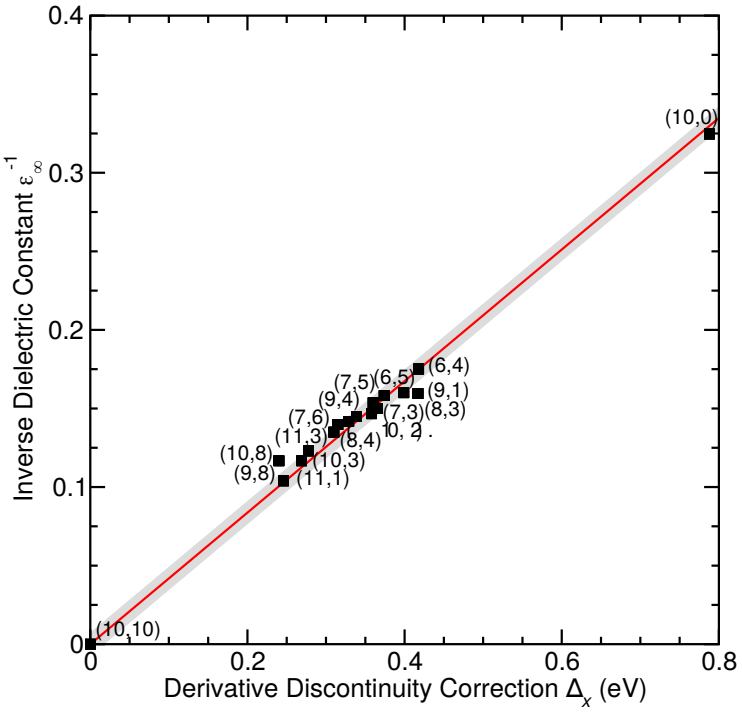

Figure 2.2: Inverse dielectric constant ε_∞^{-1} calculated within LCAO-TDDFT-k-ω versus the GLLB-SC derivative discontinuity correction Δ_x for 15 achiral and 2 chiral SWCNTs. A linear fit $\varepsilon_\infty^{-1} \approx 0.418\Delta_x$ ($r \approx 0.995$) is shown as a red line with standard deviation $\sigma \approx 0.007$ depicted in the grey region. Adapted from [9]

over states i from 1 to N, and divide by the density $n(\mathbf{r})$, resulting in an expression for $\hat{V}_{\text{eff}}(\mathbf{r})$

$$\hat{V}_{\text{eff}}(\mathbf{r}) = \frac{1}{n(\mathbf{r})} \sum_{i=1}^{N} \left[\frac{1}{2} \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) + \varepsilon_i \left| \psi_i(\mathbf{r}) \right|^2 \right] \tag{2.34}$$

from which an expression for $\hat{V}_{xc}(\mathbf{r})$ can be obtained by subtracting off the calculated external potential and Hartree portion.

When the xc functional has explicit dependence on the orbitals rather than just the electron density, we follow the approach of Kananenka et al. [87] by employing the orbital-dependent KS formalism, replacing Eq. 2.16 with

$$\left[-\frac{1}{2}\nabla^2 + \hat{V}_{\text{ext}}(\mathbf{r}) + \hat{V}_{\text{H}}(\mathbf{r}) + \frac{\partial E_{xc}[\{\psi_i^*(\mathbf{r})\}]}{\partial \psi_i^*(\mathbf{r})} \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \tag{2.35}$$

giving a new result for the xc potential

$$\hat{V}_{\mathrm{xc}}(\mathbf{r}) = \frac{1}{n(\mathbf{r})} \sum_{i=1}^{N} \left[\frac{1}{2} \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) + \varepsilon_i \left| \psi_i(\mathbf{r}) \right|^2 \right] - \hat{V}_{\mathrm{ext}}(\mathbf{r}) - \hat{V}_{\mathrm{H}}(\mathbf{r}), \tag{2.36}$$

effectively redefining it to an orbital-averaged form given by

$$\hat{V}_{\mathrm{xc}}(\mathbf{r}) = \frac{1}{n(\mathbf{r})} \sum_{i=1}^{N} \psi_i^*(\mathbf{r}) \frac{\partial E_{xc}[\{\psi_i^*(\mathbf{r})\}]}{\partial \psi_i^*(\mathbf{r})} \psi_i(\mathbf{r}). \tag{2.37}$$

For density functionals with explicit dependence on the electron density this expression reduces to the functional derivative shown in Eq. 2.14 but results in an orbital-averaged potential when using orbital-dependent functionals, thereby providing an efficient way to construct orbital-averaged xc potentials for hybrid functionals.

2.1.5 Wavefunction Representations in DFT

Fig. 2.3 presented the self-consistent cycle that allows the solution of an electron density which minimizes our total energy in Eq. 2.9 while satisfying the KS equations presented in Eq. 2.16. Computationally, the third step in the cycle, solving the single-electron Schrödinger equation, requires a choice of basis for the KS wavefunction $\psi_i(\mathbf{r})$. Three common representations, each with their own pros and cons, are presented below.

The first and most obvious case is to represent the wavefunctions in real space, with all calculations done on a finite real space grid through a finite difference method (FD). Solving the KS equations this way has the advantage of being easily parallelizable and efficient for larger systems. The accuracy and convergence of the method will largely be controlled by the size of grid spacing used. However, for systems with substantial amounts of empty space in their unit cells, as in the case of macromolecules or lower-dimensional systems, a real space representation requires substantial memory to store the wavefunctions. For these smaller systems, a plane wave (PW) implementation tends to converge faster, resulting in a significant time speedup [44]. In PW, all quantities are represented by their Fourier transforms on the periodic super cell so that Eq. 2.16 can be solved in Fourier space. For example, the periodic potential is represented by

$$v(\mathbf{r}) = \sum_{\mathbf{G}} \hat{v}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{2.38}$$

where $\mathbf{G}$ represents a reciprocal lattice vector associated with the Brillouin Zone (BZ) of the crystal. Higher values of $\mathbf{G}$ correspond to solutions with higher kinetic energies. The convergence parameter E_{cut} determines the maximum number of reciprocal lattice vectors included in the plane wave expansion of the KS wavefunctions.

In LCAO, the KS wavefunctions are expanded as a set of atomic-like orbitals $\phi_{nlm}(\mathbf{r}) = \phi_\mu(\mathbf{r})$,

which are constructed as products of radial functions and spherical harmonics which are classical solutions to the exactly solvable hydrogen atom [15]. The expansion goes as

$$\psi_n(\mathbf{r}) = \sum_{\mu} c_{\mu n}\phi_{\mu}(\mathbf{r}), \tag{2.39}$$

where n represents the band index. When Eq. 2.16 is represented as an eigenvalue problem with the total energy Hamiltonian diagonalized in the basis of these atomic orbitals for each atom in the unit cell, the coefficients $c_{\mu n}$ are the variational parameters.

A minimal basis set consists of one atomic orbital-like function for each valence state per atom in the unit cell of the crystal, although expanding the span beyond this basis is necessary to obtain well-converged and physical results. There are two common ways to accomplish this: the first is to include extra radial functions (represented by ζ), corresponding to higher-order orbital solutions of the hydrogen atom. The second way is by including spherical harmonics corresponding to unoccupied angular momentum quantum numbers, referred to as polarization, effectively accounting for possible hybridization of orbitals in the basis set representation [15]. For example, a double-ζ-polarized basis set for a carbon atom, the kind utilized in Chapter 5 to represent the fullerene molecule, will have 2s and 2p valence states with two radial functions each. Including an additional polarization function with d-orbital character yields a total of 5 distinct radial functions. The consequences of a proper choice of LCAO basis set are brought up throughout Chapter 5.

For larger systems and for lower-dimensional materials that represent some of the most promising systems for nanoscience applications, an LCAO basis set representation will have significant speedup over both real space and plane wave schemes [15], as the degrees of freedom in the diagonalization scheme will be limited to the number of distinct radial functions per atom. This does however come at the cost of precision, as the accuracy of real space and plane wave representations can be systematically optimized by either increasing the grid spacing or the plane wave cutoff E_{cut}, respectively [88]. Increasing the number of basis functions after a certain point does not necessarily result in more accurate results, and including too many will result in diminishing returns due to the significant spatial overlap of the higher-order orbital solutions to the hydrogen atom.

2.1.6 The Self-Consistent Cycle

A computational implementation of the KS system, consisting of a variational procedure to minimize the energy in Eq. 2.9 while satisfying Eqs. 2.13, 2.15 and 2.16, requires a self-consistent iterative procedure.

Fig. 2.3 outlines one way of implementing this minimization, starting with an ionic structure which has already converged to a stable equilibrium. After (1) making an initial guess for the electron density, the (2) effective potential is determined by taking the functional derivative of Eq. 2.9 with respect to the electron density, as shown in Eq. 2.13. This effective potential can

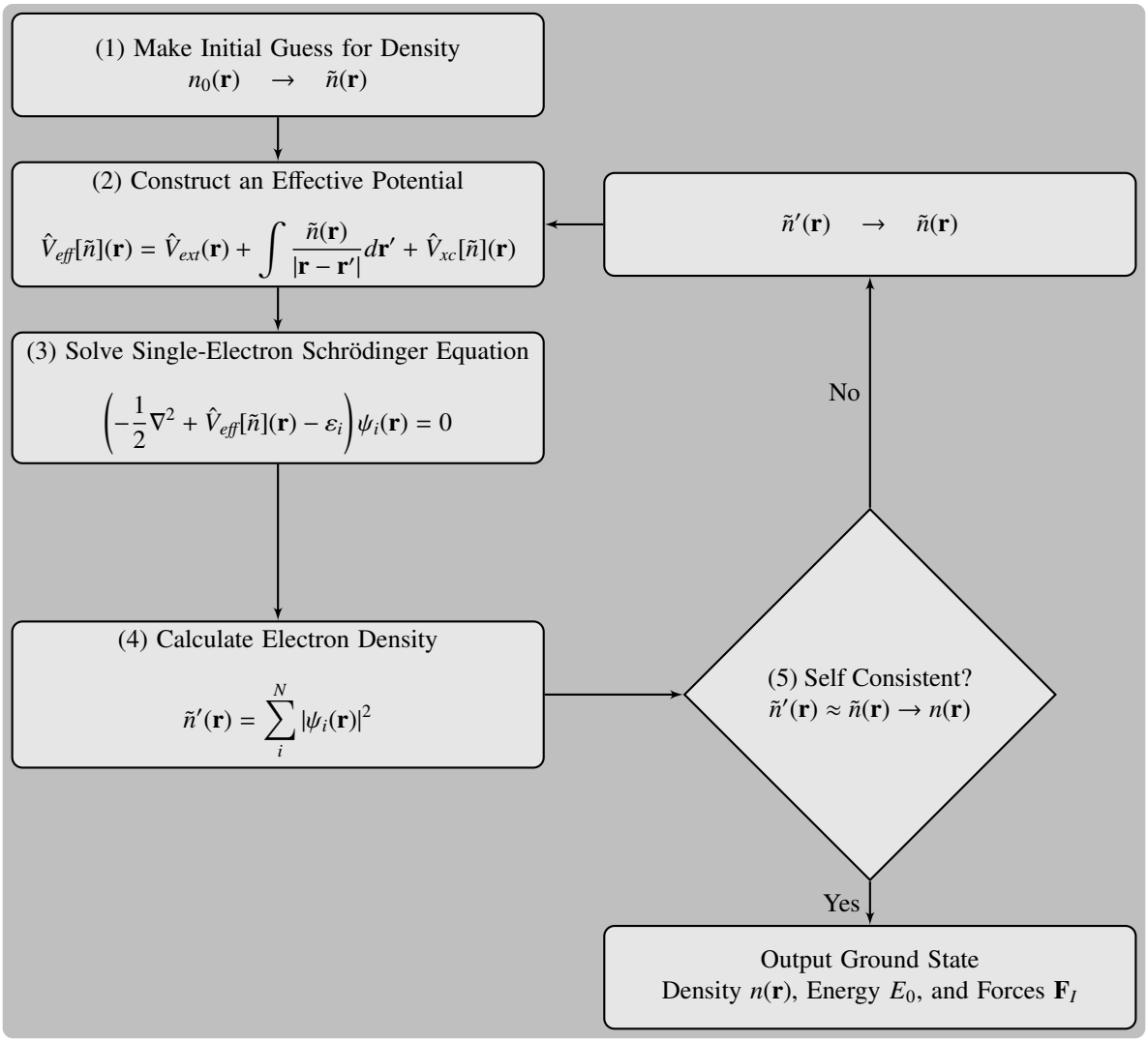

Figure 2.3: Outline of the self-consistent cycle commonly used in density functional theory codes. Adapted from [89]

.

be used to (3) solve the KS eigenvalue problem shown in Eq. 2.16 , after which the (4) electron density can be computed through Eq. 2.15. This cycle is (5) repeated until the total energy reaches a self-consistent local minimum, requiring the electron density and the square of the residuals in the KS equations to satisfy certain convergence cycle criteria.

2.2 Time-Dependent DFT (TDDFT)

While DFT is applicable and useful as an efficient method for determining the ground-state properties of many-electron systems such as solids and molecules, excited state properties like optical response and excitation spectra, coming about from time-dependent potentials, require more than a naïve application of the KS method presented above.

The time-dependent analogue, known as time-dependent density functional theory (TDDFT), extends the time-independent DFT formalism by first invoking the Runge-Gross theorem [89], showing that there is a one-to-one correspondence between the time-dependent density $n(\mathbf{r}, t)$ and time-dependent potentials $\hat{V}_{\text{ext}}(\mathbf{r}, t)$ given a proper initial condition in time. As with time-independent DFT, the TDDFT framework can be understood as an extension of the time-dependent formalism for HF theory where the one-electron orbitals follow the solution of the time dependent Schrödinger equation. Formally, what the Runge-Gross theorem states is that if two external potentials $\hat{V}(\mathbf{r}, t)$ and $\hat{V}'(\mathbf{r}, t)$ differ by more than a purely time-dependent function, they cannot produce the same time-dependent density $n(\mathbf{r}, t)$. The uniqueness and one-to-one correspondence of the external potential and density means that we can take the KS non-interacting electron density to be a valid approximation to the interacting density of the real system, yielding time-dependent KS wavefunctions that satisfy the time-dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\psi_i(\mathbf{r}, t) = \left[-\frac{1}{2}\nabla^2 + \hat{V}_{\text{ext}}(\mathbf{r}, t) + \int \frac{n(\mathbf{r}, t)}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \hat{V}_{xc}(\mathbf{r}, t)\right]\psi_i(\mathbf{r}, t) \tag{2.40}$$

where as before

$$n(\mathbf{r}, t) = \sum_{i=1}^{N} \psi_i^*(\mathbf{r}, t)\psi_i(\mathbf{r}, t) \tag{2.41}$$

For the purposes of this thesis, the use of TDDFT is of particular interest regarding its application towards the calculation of optical response functions, yielding information about the optical and dielectric properties of the materials under question. The density-density response function, the dielectric function, and the conductivity are all connected through constitutive relations and are all examples of linear response functions that can be calculated within the TDDFT formalism, as shown in the following section.

2.2.1 Optical Response Functions

The calculation of the non-interacting density response function is the first step towards a derivation of other more commonly used optical response functions. In real space it is written as [90, 91]

$$\chi^0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{\mathbf{k},\mathbf{q}}^{\text{BZ}} \sum_{n,n'} \frac{f(\varepsilon_{n\mathbf{k}}) - f(\varepsilon_{n'\mathbf{k}+\mathbf{q}})}{\omega + \varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}+\mathbf{q}} + i\eta}\psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{n'\mathbf{k}+\mathbf{q}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}')\psi^*_{n'\mathbf{k}+\mathbf{q}}(\mathbf{r}'), \tag{2.42}$$

where $\varepsilon_{n\mathbf{k}}$ are the eigenvalues associated with energy level n and wavevector $\mathbf{k}$, $\psi_{n\mathbf{k}}(\mathbf{r})$ are the eigenfunctions normalized to the crystal volume Ω, η is the electronic broadening, f is the Fermi-Dirac occupation, and the summation is over wavevectors $\mathbf{k}, \mathbf{q}$ in the BZ. The sum of occupations $f_{n\mathbf{k}} = f(\varepsilon_{n\mathbf{k}})$ should be the total number of electrons in the crystal. For periodic systems, χ^0 can be expanded in plane-wave basis as

$$\chi^0(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{\Omega} \sum_{\mathbf{q}}^{\mathrm{BZ}} \sum_{\mathbf{G}\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \tag{2.43}$$

where $\mathbf{q}$ is the Bloch vector. This can be understood as a Fourier transform in space, yielding a function χ^0 that depends on the reciprocal lattice vectors $\mathbf{G}$, $\mathbf{G}'$ with coefficients given by

$$\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{n,n'} \frac{f(\varepsilon_{n\mathbf{k}}) - f(\varepsilon_{n'\mathbf{k}+\mathbf{q}})}{\omega + \varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}+\mathbf{q}} + i\eta} \times \langle \psi_{n\mathbf{k}} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \psi_{n'\mathbf{k}+\mathbf{q}} \rangle \langle \psi_{n'\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} | \psi_{n\mathbf{k}} \rangle, \tag{2.44}$$

noting that the dipole transition elements $\langle \psi_{n\mathbf{k}} | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} | \psi_{n'\mathbf{k}+\mathbf{q}} \rangle$ are evaluated over the volume of the unit cell.

While the formula above yields the non-interacting density response function, the full interacting density response function, required to properly describe the effects of local fields and inhomogeneities in the system, is obtained by solving Dyson's equation in reciprocal space [91],

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) + \sum_{\mathbf{G}_1\mathbf{G}_2} \chi^0_{\mathbf{G}\mathbf{G}_1}(\mathbf{q}, \omega) K_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q}) \chi_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}, \omega). \tag{2.45}$$

where the kernel is composed of the Fourier-transformed sum of the xc and Coulomb potentials

$$K_{\mathbf{G}_1\mathbf{G}_2}(\mathbf{q}) = \frac{4\pi}{|\mathbf{q}+\mathbf{G}_1|^2} \delta_{\mathbf{G}_1\mathbf{G}_2} + \frac{1}{\Omega} \int d\mathbf{r} \left. \frac{\partial^2 E_{xc}[n]}{\partial n^2} \right|_{n_0(\mathbf{r})} e^{-i(\mathbf{G}_1-\mathbf{G}_2)\cdot\mathbf{r}}. \tag{2.46}$$

The dielectric matrix is related to the interacting density response function by

$$\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \delta_{\mathbf{G}\mathbf{G}'} + \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) \tag{2.47}$$

In the random phase approximation (RPA) [79], where electrons are assumed to respond only to the total electric potential and the external perturbing potential is assumed to oscillate at a single frequency ω, a dielectric function denoted by $\varepsilon_{\mathrm{RPA}}(\mathbf{q}, \omega)$ can be expressed in terms of the non-interacting density response function

$$\varepsilon^{\mathrm{RPA}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \delta_{\mathbf{G}\mathbf{G}'} - \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} \chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega). \tag{2.48}$$

The dependence on $\mathbf{q}$ within RPA stems from the assumption that the effect of the total electric potential on the dielectric function is expected to average out over space. The macroscopic dielectric function is defined by

$$\varepsilon_M(\mathbf{q},\omega) = \frac{1}{\varepsilon_{00}^{-1}(\mathbf{q},\omega)} \tag{2.49}$$

and unless stated otherwise, references to $\epsilon(\mathbf{q},\omega)$ in the proceeding text refer to the macroscopic dielectric function shown here. Considering different directions of $\hat{\mathbf{q}}$ will yield a dielectric matrix.

2.2.2 Optical Limit

In the above section the optical response function $\varepsilon(\mathbf{q},\omega)$, also known as the Lindhard longitudinal dielectric function, was derived explicitly. This longitudinal response is most relevant when determining a material's response to light absorption or electron scattering. Of particular note is the calculation of absorption, which is defined by the formula

$$\mathcal{A} = \frac{4\pi}{c}\,\mathrm{Im}[\lim_{q\to 0}\varepsilon_M(\mathbf{q},\omega)] = \frac{4\pi}{c}\,\mathrm{Im}\left[\lim_{q\to 0}\frac{1}{\varepsilon_{00}^{-1}(\mathbf{q},\omega)}\right]. \tag{2.50}$$

Trying to use the approximation $\mathbf{q}=0$ in Eq. 2.48 fails because the Coulomb kernel $4\pi/|\mathbf{q}+\mathbf{G}|^2$ diverges at $\mathbf{q}=\mathbf{G}=0$. However, it can be shown that in the limit of $\mathbf{q}\to 0$ and $\mathbf{G}=0$ the dielectric matrix can be evaluated. The dipole transition element $\langle\psi_{n\mathbf{k}}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n'\mathbf{k}+\mathbf{q}}\rangle$ with reciprocal lattice vector $\mathbf{G}=0$ becomes [92]

$$\langle\psi_{n\mathbf{k}}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n'\mathbf{k}+\mathbf{q}}\rangle = \langle u_{n\mathbf{k}}|u_{n'\mathbf{k}+\mathbf{q}}\rangle. \tag{2.51}$$

where $u_{n\mathbf{k}}$ is the periodic part of the Bloch wave in the expression $\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$. First order perturbation theory allows us to write $u_{n'\mathbf{k}+\mathbf{q}}$ in terms of other orbitals defined at wavevector $\mathbf{k}$ as

$$|u_{n'\mathbf{k}+\mathbf{q}}\rangle = |u_{n'\mathbf{k}}\rangle + \sum_{m\neq n'}\frac{\langle u_{m\mathbf{k}}|\tilde{V}|u_{n'\mathbf{k}}\rangle}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{m\mathbf{k}}}|u_{m\mathbf{k}}\rangle, \tag{2.52}$$

where the perturbation $\tilde{V}$, reflecting the change in kinetic energy of the two different periodic Bloch waves, is obtained via $\mathbf{k}\cdot\mathbf{p}$ perturbation theory as

$$\tilde{V} = H(\mathbf{k}+\mathbf{q}) - H(\mathbf{k}) = -i\mathbf{q}\cdot(\nabla + i\mathbf{k}) \tag{2.53}$$

Substituting $\tilde{V}$ into Eq. 2.52, multiplying by $\langle u_{n\mathbf{k}}|$ on the left, and applying the orthonormality requirement on the Bloch states $\langle u_{n\mathbf{k}}|u_{m\mathbf{k}}\rangle = \delta_{nm}$ gives

$$\langle\psi_{n\mathbf{k}}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\psi_{n'\mathbf{k}+\mathbf{q}}\rangle_{\mathbf{q}\to 0,\mathbf{G}=0} = -i\mathbf{q}\cdot\frac{\langle u_{n\mathbf{k}}|\nabla + i\mathbf{k}|u_{n'\mathbf{k}}\rangle}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}} = -i\mathbf{q}\cdot\frac{\langle\psi_{n\mathbf{k}}|\nabla|\psi_{n'\mathbf{k}}\rangle}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}}. \tag{2.54}$$

Using the optical form for the dipole transition element, the dielectric function in Eq. 2.48 can be written as

$$\varepsilon(\hat{\mathbf{q}},\omega) = 1 - \frac{4\pi}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n,n'}\frac{w_{\mathbf{k}}\left[f(\varepsilon_{n\mathbf{k}}) - f(\varepsilon_{n'\mathbf{k}})\right]|f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2}{\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}}) + i\eta}. \tag{2.55}$$

where $w_{\mathbf{k}}$ is the weight of each wavevector and the oscillator strengths $f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}$ of the $n \rightarrow n'$ transition at k-point $\mathbf{k}$ in the direction of the Bloch vector $\mathbf{q}$ are defined as

$$f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}} = \lim_{\mathbf{q}\to 0}\frac{\left\langle\psi_{n'\mathbf{k}+\mathbf{q}}\left|e^{i\mathbf{q}\cdot\mathbf{r}}\right|\psi_{n\mathbf{k}}\right\rangle}{|\mathbf{q}|} \approx i\hat{\mathbf{q}}\cdot\frac{\langle\psi_{n'\mathbf{k}}|\nabla|\psi_{n\mathbf{k}}\rangle}{\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}}}. \tag{2.56}$$

It is worthwhile noting that the matrix elements $\langle\psi_{n'\mathbf{k}}|\nabla|\psi_{n\mathbf{k}}\rangle$ are already calculated in DFT when obtaining the forces during structural relaxation, so calculation of the dielectric function in Eq. 2.55 can be accomplished using previously calculated values. This allows for calculations of the optical response functions with this approach to scale as $O(NM^2)$ or better [93], where N is the number of KS wavefunctions and $M \geq N$ is the number of basis functions used in the LCAO calculation over all atoms [9, 46]. Within DFT software packages that employ projector-augmented waves (PAW), this matrix element is given by [10]

$$\langle\psi_n|\nabla|\psi_{n'}\rangle = \sum_{\mu\nu} c^*_{\nu n}c_{\mu n'}\langle\tilde{\phi}_\nu|\mathcal{T}^\dagger\nabla\mathcal{T}|\tilde{\phi}_\mu\rangle, \tag{2.57}$$

where $\tilde{\phi}_\mu$ are the localized basis functions for the n^{th} KS wavefunction $|\tilde{\psi}_n\rangle = \sum_\mu c_{\mu n}|\tilde{\phi}_\mu\rangle$, with coefficients $c_{\mu n}$ as shown in Eq. 2.39, and $\mathcal{T}$ is the PAW transformation operator [2, 3, 83]

$$\mathcal{T} = 1 + \sum_{ai}\left(|\varphi^a_i\rangle - |\tilde{\varphi}^a_i\rangle\right)\langle\tilde{p}^a_i|, \tag{2.58}$$

where $|\tilde{p}^a_i\rangle$ are the smooth PAW projector functions, $\tilde{\varphi}^a_i$ and φ^a_i are the pseudo and all-electron partial waves, and the sum runs over all states i and atoms a as defined within the PAW formalism.

2.2.3 Applying the Derivative Discontinuity Correction

The application of a shift to the eigenenergies of unoccupied orbitals, like that shown in Eq. 2.31, needs to be handled with care when considering its application to the calculation of response functions, like the optical dielectric function shown in Eq. 2.55. Letting $\psi^{\Delta_x}_{n\mathbf{k}}$ refer to conduction band orbitals whose energies have been shifted to $\varepsilon_{n\mathbf{k}} + \Delta_x$, Eq. 2.55 can be rewritten as

$$\varepsilon(\hat{\mathbf{q}},\omega) = 1 - \frac{4\pi}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n,n'}\frac{w_{\mathbf{k}}\left[f(\varepsilon_{n\mathbf{k}}) - f(\varepsilon_{n'\mathbf{k}})\right]|f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2}{\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x) + i\eta}, \tag{2.59}$$

where the oscillator strengths in Eq. 2.56 may be adjusted according to when the discontinuity correction is included. In the *a priori* scissors approach,

$$f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}} = \lim_{\mathbf{q}\to 0} \frac{\left\langle \psi^{\Delta_x}_{n'\mathbf{k}+\mathbf{q}} \left| e^{i\mathbf{q}\cdot\mathbf{r}} \right| \psi_{n\mathbf{k}} \right\rangle}{|\mathbf{q}|} \approx i\hat{\mathbf{q}} \cdot \frac{\langle \psi_{n'\mathbf{k}} | \nabla | \psi_{n\mathbf{k}} \rangle}{\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x}, \tag{2.60}$$

where it is assumed that the spatial distribution of the wavefunction $\psi^{\Delta_x}_{n'\mathbf{k}}$ is unchanged by Δ_x. The transition energies are increased by Δ_x in Eq. 2.59, altering both the positions and intensities of the peaks. It is important to note that the application of $\mathbf{k}\cdot\mathbf{p}$ perturbation theory in Eq. 2.53 to obtain the optical limit dielectric function assumes the electron's velocity is the momentum divided by the mass, $\mathbf{v} = \mathbf{p}/m_e$, which is only a valid approximation for local potentials [94]. This is the case for the GLLB-SC correction presented in Eq. 2.31, validating this *a priori* scissors approach.

However, the application of a scissors approach for non-local potentials will break gauge invariance [94], for example in the Bethe-Salpeter equation where $\mathbf{v} = \mathbf{p}/m_e + \partial\Sigma(\omega)/\partial\mathbf{p}$ where $\Sigma(\omega) = iGW$ is the quasiparticle self-energy. This gauge invariance may be addressed by first approximating the wavefunctions corresponding to the non-local potential by the KS wavefunctions generated by a local potential, and subsequently performing a perturbative expansion in $\mathbf{q}$. In this *a posteriori* scissors approach, the oscillator strengths $f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}$ will be the ones in Eq. 2.56, independent of Δ_x, meaning that only the peak positions in the resulting expression for the dielectric function will be shifted by Δ_x. Both the *a priori* and *a posteriori* scissors approaches can be construed as valid when describing the corrected dielectric function for either semiconducting or insulating systems.

2.2.4 The Two-Point Excitonic Spectral Density

By working in the optical limit and neglecting LFEs, we can define the two-point excitonic spectral function as [10]

$$\rho_{ex}(\mathbf{r}_e, \mathbf{r}_h, \hat{\mathbf{q}}, \omega) = \frac{4\pi}{\Omega} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{nn'} \frac{\eta w_{\mathbf{k}} |f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2 |\psi_{n\mathbf{k}}(\mathbf{r}_h)|^2 |\psi_{n'\mathbf{k}}(\mathbf{r}_e)|^2}{(\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x))^2 + \eta^2}, \tag{2.61}$$

where $\mathbf{r}_e$ and $\mathbf{r}_h$ represent the real space locations of the electron and hole, respectively, and $f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}$ is the oscillator strength of the $n \to n'$ transition from either the *a priori* (Eq. 2.60) or *a posteriori* (Eq. 2.56) scissors approaches. Analogous to the way eigenstates are plotted for the two-particle Hamiltonian in BSE calculations (see Section 2.2.7), Eq. 2.61 shows the density for a non-interacting KS system via the diagonal RPA polarizability. Since electrons and holes are treated as non-interacting in standard TDDFT, it is logical to define expressions for hole and

electron spectral functions that average over the electron and hole coordinates, respectively, as

$$\begin{aligned}\rho_h(\mathbf{r}_h, \hat{\mathbf{q}}, \omega) &= \int \rho_{ex}(\mathbf{r}_e, \mathbf{r}_h, \hat{\mathbf{q}}, \omega) d\mathbf{r}_e \\ &= \frac{4\pi}{\Omega} \sum_{\mathbf{k}} \sum_{nn'} \frac{\eta w_{\mathbf{k}} |f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2 \, |\psi_{n\mathbf{k}}(\mathbf{r}_h)|^2}{(\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x))^2 + \eta^2}, \end{aligned} \tag{2.62}$$

$$\begin{aligned}\rho_e(\mathbf{r}_e, \hat{\mathbf{q}}, \omega) &= -\int \rho_{ex}(\mathbf{r}_e, \mathbf{r}_h, \hat{\mathbf{q}}, \omega) d\mathbf{r}_h \\ &= -\frac{4\pi}{\Omega} \sum_{\mathbf{k}} \sum_{nn'} \frac{\eta w_{\mathbf{k}} |f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2 \, |\psi_{n'\mathbf{k}}(\mathbf{r}_e)|^2}{(\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x))^2 + \eta^2}, \end{aligned} \tag{2.63}$$

the integration presuming the completeness of the KS wavefunction set. These definitions for the excitonic spectral function and the electron and hole spectral densities all satisfy [10]

$$\iint \rho_{ex}(\mathbf{r}_e, \mathbf{r}_h, \hat{\mathbf{q}}, \omega) d\mathbf{r}_h d\mathbf{r}_e = \mathrm{Im}[\varepsilon(\hat{\mathbf{q}}, \omega)], \tag{2.64}$$

thereby allowing for the exciton, hole, and electron spectral functions to be spatially and energetically resolved. Subtracting the electron and hole spectral densities yields [9]

$$\Delta\rho(\mathbf{r}, \hat{\mathbf{q}}, \omega) = \rho_h(\mathbf{r}, \hat{\mathbf{q}}, \omega) + \rho_e(\mathbf{r}, \hat{\mathbf{q}}, \omega) \tag{2.65}$$

$$= \frac{4\pi}{\Omega} \sum_{\mathbf{k}} \sum_{nn'} \frac{\eta w_{\mathbf{k}} |f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}|^2 \left(|\psi_{n\mathbf{k}}(\mathbf{r}_h)|^2 - |\psi_{n'\mathbf{k}}(\mathbf{r}_e)|^2 \right)}{(\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}} + \Delta_x))^2 + \eta^2}. \tag{2.66}$$

This spectral function in particular gives important information about the real-space distribution of both electron and hole states for the system under investigation. Analyzing this function at prominent peaks in conductivity, the dielectric function, or photoabsorption data can yield useful insight into the nature of orbital overlap and directional dependence at energies of interest.

2.2.5 Mean-Field Response Functions in Lower-Dimensional Systems

While the expression in Eq. 2.55 shows the dielectric function in the context of how it is calculated within DFT, it is prudent to note that the other optical response functions can be calculated through their relation to quantities like the dielectric function and the density response function. However, these relations can become muddled when the optical response functions are calculated for lower-dimensional materials when also invoking the optical limit. Appendix A deals with this in the context of two-dimensional materials where the unscreened Coulomb potential changes depending on the dimension dealt with, but these calculations are only valid for finite wavenumber q. Coming from a computational standpoint, we opt instead to follow approaches based on mean-field theory laid out in Ref. [95] for non-interacting 1D SWCNTs and

in Refs. [18, 19] for non-interacting 2D sheets to compute the polarizability response function $\alpha(\hat{\mathbf{q}}, \omega)$ as a function of the 3D dielectric function presented in Eq. 2.55.

For a given dimension d, the polarizability for light polarized in directions "parallel" to the material, $\hat{\mathbf{q}}_{\parallel}$, and in non-periodic directions "perpendicular" to the material, $\hat{\mathbf{q}}_{\perp}$, are expressed according to

$$\alpha_d(\hat{\mathbf{q}}_{\parallel}, \omega) = \frac{\Omega_d}{4\pi} \left(\varepsilon(\hat{\mathbf{q}}_{\parallel}, \omega) - 1\right), \tag{2.67}$$

$$\alpha_d(\hat{\mathbf{q}}_{\perp}, \omega) = \frac{\Omega_d}{4\pi} \left(1 - \frac{1}{\varepsilon(\hat{\mathbf{q}}_{\perp}, \omega)}\right), \tag{2.68}$$

where Ω_d represents the "cross-section" of the unit cell for a given dimension d, and $\varepsilon(\hat{\mathbf{q}}, \omega)$ is the 3D macroscopic dielectric function obtained from Eq. 2.55, so that the polarization is defined per molecule or per layer rather than per unit volume. Ω_{0D} is the volume of the unit cell, Ω_{1D} is the area of the plane in the unit cell perpendicular to the 1D material, Ω_{2D} is the length of the unit cell perpendicular to the plane of the 2D material, and $\Omega_{3D} = 1$. The conductivity σ for any dimension can then be expressed via

$$\sigma(\hat{\mathbf{q}}, \omega) = -i\omega\alpha(\hat{\mathbf{q}}, \omega), \tag{2.69}$$

a generalization of the definition of the optical limit conductivity $\sigma(\hat{\mathbf{q}}, \omega)$ to lower-dimensional systems.

2.2.6 Interband and Intraband Conductivity

Utilizing the expression for the dipole transition element derived in Eq. 2.54, the definition of conductivity in Eq. 2.69, and the non-interacting density response function in Eq. 2.42, we show equations for the interband conductivity [96]

$$\sigma_{inter}(\hat{\mathbf{q}}, \omega) = -i\omega\frac{1}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n\neq n'} \frac{f(\varepsilon_{n\mathbf{k}}) - f(\varepsilon_{n'\mathbf{k}})}{\hbar\omega - (\varepsilon_{n'\mathbf{k}} - \varepsilon_{n\mathbf{k}}) + i\eta_{inter}} \left|\frac{\hat{\mathbf{q}} \cdot \langle\psi_n|\nabla|\psi_{n'}\rangle}{\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}}}\right|^2. \tag{2.70}$$

and the intraband conductivity [97]

$$\sigma_{intra}(\hat{\mathbf{q}}, \omega) = -\frac{i}{\omega}\frac{1}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n} f'(\varepsilon_{n\mathbf{k}}) \left|\hat{\mathbf{q}} \cdot \langle\psi_n|\nabla|\psi_n\rangle\right|^2, \tag{2.71}$$

the distinction coming about depending on whether the summation is between bands or within the same band. Equivalent expressions for the intraband conductivity are derived in Appendix C. We can rewrite this expression to showcase its tensorial nature as

$$\sigma(\hat{\mathbf{q}}, \omega) = \hat{\mathbf{q}} \cdot (\overleftrightarrow{\sigma}_{inter} + \overleftrightarrow{\sigma}_{intra}) \cdot \hat{\mathbf{q}} \tag{2.72}$$

where

$$\overleftrightarrow{\sigma}_{inter}(\hat{\mathbf{q}},\omega) = -i\omega\frac{1}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n\neq n'}\frac{f(\varepsilon_{n\mathbf{k}})-f(\varepsilon_{n'\mathbf{k}})}{\hbar\omega-(\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}})+i\eta_{inter}}\frac{\langle\psi_n|\nabla|\psi_{n'}\rangle\langle\psi_{n'}|\nabla|\psi_n\rangle}{(\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}})^2} \tag{2.73}$$

$$\overleftrightarrow{\sigma}_{intra}(\hat{\mathbf{q}},\omega) = -\frac{i}{\omega+i\eta_{intra}}\frac{1}{\Omega}\sum_{\mathbf{k}}^{\mathrm{BZ}}\sum_{n}f'(\varepsilon_{n\mathbf{k}})\langle\psi_n|\nabla|\psi_n\rangle\langle\psi_n|\nabla|\psi_n\rangle. \tag{2.74}$$

As conductivity is a linear response function satisfying causality, it will satisfy the Kramers-Kronig relations [18, 98]

$$\mathrm{Re}[\sigma(\omega)] = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{\mathrm{Im}[\sigma(\omega')]}{\omega'-\omega}d\omega' \tag{2.75}$$

$$\mathrm{Im}[\sigma(\omega)] = -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{\mathrm{Re}[\sigma(\omega')]}{\omega'-\omega}d\omega' \tag{2.76}$$

with $\mathcal{P}$ as the Cauchy principal value. Often times it is the imaginary portion of the conductivity which is desired, so treating this numerically is easier when the integral is treated as having a removable singularity

$$\mathrm{Im}[\sigma(\omega)] = -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty}\frac{\mathrm{Re}[\sigma(\omega')]-\mathrm{Re}[\sigma(\omega)]}{\omega'-\omega}d\omega', \tag{2.77}$$

which is possible since the integral over the principal value of $1/(\omega'-\omega)$ vanishes. Using these Kramers-Kronig relations with physical arguments about the behavior of the real and imaginary parts of the response function can be used to establish a set of so-called sum rules for various optical parameters [99]. Since conductivity of a material defines a relationship between an applied electric field and the electrons it causes to move around, and the total number of electrons in the system must remain constant under that type of excitation, the f sum rule for conductivity,

$$\int_0^{\infty}\mathrm{Re}[\sigma(\omega)]d\omega = \frac{\pi}{2}\sum_j\frac{e^2}{m_e}, \tag{2.78}$$

comes as a consequence of this conserved quantity, where the sum is over all particles excited by the response included in the conductivity function. Similar sum rules exist for the other optical response functions shown above [99].

2.2.7 G_0W_0-BSE

The GLLB-SC method shown in Section 2.1.3 presents a robust approach for adjusting the electronic band gap by upshifting the energies of unoccupied orbitals. However, it is imperative in many applications to obtain accurate results for the optical band gap, the localization of energy

states, and level alignment for interfaced materials, all properties reliant on a proper description of the addition and removal of electrons into or out of the system [100]. The GW approximation, by including the screened Coulomb interaction into the eigenenergies, is one tactic for adjusting the electronic band structure to include this screening effect [101]. G refers to the full one-particle Green's function, and W to the dynamic screened Coulomb potential. G_0W_0 is the systematic perturbation of this GW, where G_0 is the independent-particle Green's function calculated from the KS wavefunctions and W_0 is the first-order dynamic screened Coulomb interaction, given by

$$W_0(\mathbf{r}_1, \mathbf{r}_2, \omega) = \int \frac{\varepsilon^{-1}(\mathbf{r}_3, \mathbf{r}_2, \omega)}{|\mathbf{r}_1 - \mathbf{r}_3|} d\mathbf{r}_3, \tag{2.79}$$

where ε is the dielectric function and ω is the energy of the quasiparticle being screened [86]. The HF equations in Eq. 2.4 can be understood as a GV approximation, the V being the bare interaction potential. G_0W_0, like the GLLB-SC method [55], is applied after a converged DFT calculation to adjust the KS eigenvalues, so these energies are approximately given by $\varepsilon^{\mathrm{KS}}_{n\mathbf{k}}$ at a given $\mathbf{k}$-point and energy level n. To save calculation time, the dielectric function used in the G_0W_0 method can be calculated within the plasmon pole approximation [102, 103, 104], modelling it as a peak around some main plasmon frequency. On top of accounting for the dynamic screening of electrons, a G_0W_0 adjustment to the KS eigenvalues accounts for the finite lifetime of single-particle excitations in an interacting system, a feature noticeable in experimentally measured spectra peaks close to zero temperature in pure materials.

The G_0W_0 correction to the KS eigenenergies provides a more accurate picture by including the role of charged excitations, those associated with electrons entering and leaving the system. Going one step further, neutral excitations, i.e. electron-hole interactions, are equally important towards describing excitonic and optical properties in a way that better approximates the many-body nature of our systems, especially for semiconducting and insulating materials [105]. These interactions are included through the solution of the four-point Bethe-Salpeter equation (BSE) [14] when calculating response functions, more specifically the macroscopic dielectric function, the electron-hole interaction included by adding so-called "vertex corrections" beyond the random phase approximation. The term "four-point" refers to the two propagating particles (e.g. two electron hole pairs) involved in the density-density response function, BSE defining the kernel of the Dyson equation used to define this response function implicitly, as in Eq. 2.45.

Ultimately, the two methods described above are often used sequentially to give the G_0W_0-BSE method, which proves to be an excellent if computationally expensive way to include screening and electron-hole interactions into response functions, crucial for matching results with data like that acquired from absorption or electron energy loss spectroscopy experiments [20, 100, 105]. In this method, BSE includes the results of G_0W_0 corrected eigenenergies along with ground-state wavefunctions from a converged DFT calculation. Computationally, the G_0W_0 calculation scales cubically with the number of plane waves [7] while constructing the Bethe-Salpeter matrix system scales with the k-point sampling of the BZ, resulting in time and memory costs on the scale of hybrid functionals [13], many implementations of which help to include excitonic effects by including the exact exchange in the xc functional [83]. Often times, fortu-

itous error cancellation within TDDFT comes about from the counterbalancing effect that the *GW* electronic screening and the BSE excitonic binding have on the spectra [9], although this is less the case for materials like small-gap semiconductors where excitonic effects become quite prominent.

2.3 Electron Energy Loss Spectroscopy

2.3.1 Experimental Setup and Terminology

Electron energy-loss spectroscopy (EELS) measures the change in the energy and direction of fast moving electrons after they have interacted with a material. Depending on the particular setup of the experiment and the relevant energy scales of the probing electrons, EELS can be divided into several basic techniques, although the physics behind the interaction processes are fundamentally the same. In high-resolution EELS, the energy of electrons leaving the beam ranges from several electron volts (eV) up to a few hundred. Electrons in this energy range cannot penetrate bulk materials deeply and instead reflect from the surface, making this EELS technique ideal for probing the surfaces of bulk materials at a high energy resolution [106]. At the higher energy scales, EELS performed in a transmission electron microscope (TEM) or in a dedicated spectrometer, also known as transmission EELS, utilizes electron beams with energies typically between 20 keV and 300 keV [107]. For samples with thicknesses in the range of tens of nanometres, these electrons are energetic enough to transmit through the sample, hence the name.

While the previous section introduced a computational mechanism for determining optical response functions like the dielectric function and the conductivity, it remains to be seen how quantities like these can be made sense of in light of experimental data, in this case in the context of electron energy loss measurements, and it is instructive to understand the experimental setup such equations describe.

Shown in Figure 2.4 is an illustration of the inelastic scattering geometry of an EELS experiment. Fast electrons in the form of plane waves interact with the material and lose energy and momentum in the process, both quantities transferred to the electrons or ions in the sample. A detector placed on the other end of the beam (in the case of transmission EELS) will only collect electrons that have lost an amount of momentum q_{tot} for particles incident to the detector at an angle θ. q_{tot} can be split into components parallel ($q_{\perp}$) and perpendicular ($q_{\parallel}$) to the direction of the electron beam.

The magnitude of $\mathbf{q}_{\text{tot}}$ can be connected to the scattering angle θ and the energy loss E by applying conservation of both energy and momentum to the scattering process. As mentioned before, since electrons with energies in the keV range are used, relativistic effects need to be considered. By knowing the energy lost from scattering, a conservation of energy argument

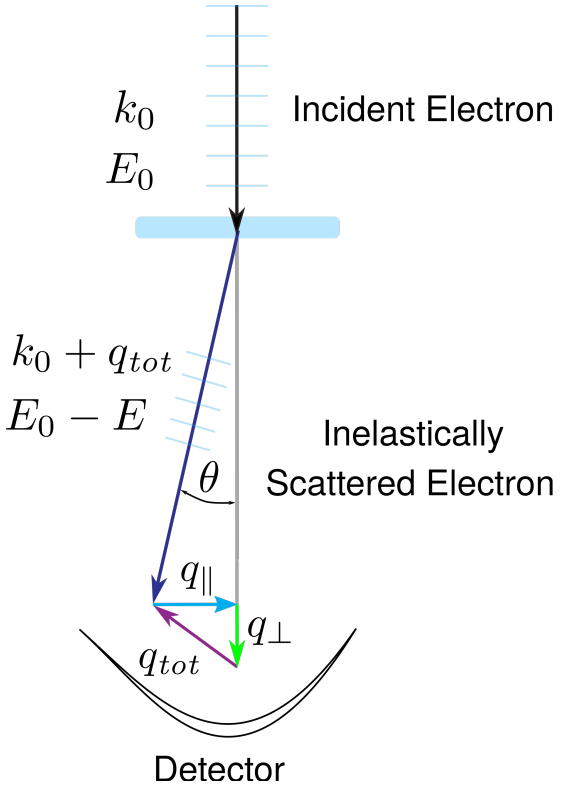

Figure 2.4: Inelastic scattering in a transmission EELS setup. An incident electron wave interacts with a material (light blue), transferring energy E and momentum q_{tot} in the process.

yields a relation between the scattered wavevector $\mathbf{k}_1 = \mathbf{k}_0 + \mathbf{q}_{\text{tot}}$ and the incident wavevector $\mathbf{k}_0$

$$|\mathbf{k}_1|^2 = |\mathbf{k}_0|^2 - \frac{2\gamma m_e E}{\hbar^2} + \frac{E^2}{\hbar^2 c^2} \tag{2.80}$$

while conservation of momentum yields a relationship between $\mathbf{k}_1$ and the incident wavevector $\mathbf{k}_0$ via the scattering angle θ

$$\mathbf{q}_{\text{tot}}^2 = |\mathbf{k}_0|^2 + |\mathbf{k}_1|^2 - 2|\mathbf{k}_0||\mathbf{k}_1|\cos(\theta) \tag{2.81}$$

These two equations in tandem make it possible to calculate values of q_{tot} for any energy loss E and scattering angle θ, and vice versa. For most transmission EELS experiments [62, 12], the energy of the incident beam is much larger than the energy loss E, resulting in small scattering angles, minor differences between wavevectors $\mathbf{k}_0$ and $\mathbf{k}_1$, and having most of the momentum transfer being perpendicular to the incident beam.

Considering the case where the collection aperture is circular and centrally positioned in the STEM device, thereby avoiding the large-angle scattering used in momentum-resolved EELS [12, 75, 107, 108], the range of collection can be classified according to the value q_c, the maxi-

mum in-plane momentum transfer for electrons entering the collection aperture. This is possible with modern TEM configurations that can achieve a sub-nm spatial resolution of the target material by using convergent electron beams. This value is given by

$$q_c = k_0 \theta_c \tag{2.82}$$

where $k_0 = |\mathbf{k}_0|$ is the magnitude of the incident electron momentum and θ_c is the maximum collection angle for electrons scattered by the material. This value can vary significantly depending on the experimental setup, with values much smaller than 1 Å^{-1} for a centrally-placed aperture on top of a narrow slit [12, 75], while measurements with broad aperture angles can yield q_c in the 10 Å^{-1} range [60, 62]. Decreases in the maximum collection angle often accompany a poorer spatial or energy resolution [107].

Since this report deals in part with the experimental spectra generated through EELS, it is important to point out certain features of the spectra and some of the relevant vocabulary useful when trying either to extract the dielectric function of a given material or when developing phenomenological models for the optical response that can be verified against energy loss spectra with the right equations. First it is important to distinguish the low-loss region, covering an energy range from 0 up to 50 eV, while the rest of the spectrum is referred to as the core-loss region. Second, provided the sample is of a certain thickness, the zero loss peak (ZLP) will be the most prominent one in the spectrum [62], generated by electrons that have undergone either no inelastic scattering or have lost insufficient energy for detection. The FWHM of this peak defines the experimental energy resolution, a quantity dependent on the sample thickness, the performance of the microscope, and the speed of electrons exiting the beam [107].

For spectra where losses with relatively high wavenumber are included, the low-loss region will be dominated by plasmon peaks, which at these energy scales constitute the most likely mechanism for inelastic scattering. Spectra associated with vanishing momentum transfers (i.e. $|\mathbf{q}| \to 0$) will correspond to single-electron excitations, for example those seen in both interband and intraband transitions, yielding features in line with optical measurements. The EELS spectrum will be a superposition of all these types of transition across all momenta up to the collection angle of the spectrometer or microscope [106]. In the case of materials that exhibit a bandgap, no intensities other than those grouped under the ZLP would then be measured in the bandgap energy range. Beyond the low-loss region, peaks relating to excitations of inner-shell electrons into unoccupied states will begin to appear in the spectra, but the low-loss region will be the region of primary interest for analyzing the feasibility of materials for optoelectronic or plasmonic applications.

2.3.2 Energy Loss for Bulk Materials

The main mechanism whereby the fast electrons shooting from the beam lose energy to the material is via the Coulomb interaction. The electron acts as an external perturbation and polarizes

the material, thereby inducing electric fields which in turn decelerate the external electrons, leading to energy loss. This energy loss is calculated by treating the incident electron as an external current density $\mathbf{J}_{ext}(z, \mathbf{r}_{\|}, t)$, and the dot product of this quantity with the electric field induced by the perturbation yields the power dissipated per unit volume. Therefore, the total energy loss W resulting from this interaction is [106]

$$W = \int dz \int d\mathbf{r}_{\|} \int dt \frac{\partial P}{\partial V} = \int dz \int d\mathbf{r}_{\|} \int dt \mathbf{J}_{ext}(z, \mathbf{r}_{\|}, t) \cdot \mathbf{E}_{ind}(z, \mathbf{r}_{\|}, t) \tag{2.83}$$

with an in-plane Fourier transform yielding

$$W = \frac{1}{(2\pi)^3} \int dz \int d\mathbf{q}_{\|} \int_{-\infty}^{\infty} d\omega \mathbf{J}_{ext}(z, \mathbf{q}_{\|}, \omega) \cdot \mathbf{E}_{ind}(z, -\mathbf{q}_{\|}, -\omega) \tag{2.84}$$

Since this quantity in real space is a real function, the real part of its Fourier transform is symmetric and the imaginary part is anti-symmetric, giving [72]

$$W = \frac{1}{(2\pi)^3} \int dz \int d\mathbf{q}_{\|} \int_{0}^{\infty} d\omega \left[\mathbf{J}_{ext} \cdot \mathbf{E}^*_{ind} + \mathbf{J}^*_{ext} \cdot \mathbf{E}_{ind}\right] \tag{2.85}$$

Rather than characterize the expression in terms of the loss of a material over a small thickness dz, it is more common to work with a differential scattering probability, the probability of an electron losing energy ω and being scattered according to wavevector $\mathbf{q}_{\|}$. The energy loss W in terms of this differential scattering probability looks like

$$W = \int_0^{\infty} d\omega \omega P(\omega) = \frac{1}{(2\pi)^3} \int d\mathbf{q}_{\|} \int_0^{\infty} d\omega \omega \frac{\partial^2 \sigma}{\partial \mathbf{q}_{\|} \partial \omega} \tag{2.86}$$

where $P(\omega)$ represents the probability of an electron losing energy ω for any wavenumber q. This yields the expression for the differential scattering cross section

$$\frac{\partial^2 \sigma}{\partial \mathbf{q}_{\|} \partial \omega} = \frac{1}{(2\pi)^3 \omega} \left[\mathbf{J}_{ext} \cdot \mathbf{E}^*_{ind} + \mathbf{J}^*_{ext} \cdot \mathbf{E}_{ind}\right]. \tag{2.87}$$

This can be directly related to quantities measurable through the EELS setup, namely the energy loss and the scattering angle, by using

$$d\omega d\mathbf{q}_{\|} = |\mathbf{k}_0|^2 \cos(\theta) d\Omega dE. \tag{2.88}$$

For an incident electron with current density $\mathbf{J}_{ext}(\mathbf{r}, t) = -e\mathbf{v}\delta(\mathbf{r} - \mathbf{v}t)$, Eq. 2.87 can be reformulated in terms of the 3D dielectric function [56]

$$\frac{\partial^2 \sigma}{\partial E \partial \Omega} = \frac{8\pi e^2 |\mathbf{k}_0|^2}{(2\pi)^3 v^2 |\mathbf{q}_{tot}|^2} \mathrm{Im}\left[-\frac{1}{\varepsilon_{3D}(\mathbf{q}_{tot}, \omega)}\right]. \tag{2.89}$$

For this reason, $\mathrm{Im}[-1/\varepsilon_{3D}(\mathbf{q}_{tot},\omega)]$ is called the energy loss function of the material. It can be directly determined via EELS measurements, ultimately yielding the dielectric function with help from the Kramers-Kronig relations shown in Eqs. 2.75, 2.76.

2.3.3 Energy Loss for 2D Materials

The treatment of two-dimensional materials in the context of energy loss, and more specifically with regards to Maxwell's equations, needs to be done with great care, as a theoretically two-dimensional material will enter into these equations as a boundary condition. In addition, as presented in both Appendix A and Section 2.2.5, optical response functions in both an analytical and computational context need to account for the restricted domain electrons can travel in and the methods by which energy can be transmitted and reflected for a monolayer. To start, in the zero-thickness approximation, assuming a monolayer material resting in the xy-plane and simplifying the notation to $\mathbf{q} = \mathbf{q}_{\parallel}$ in the 2D regime, the induced current density is restricted via

$$\mathbf{J}_{ind}(\mathbf{q}, z, \omega) = \delta(z)\mathbf{J}_{2D}(\mathbf{q}, \omega), \tag{2.90}$$

where $\mathbf{J}_{2D}(\mathbf{q},\omega)$ is the Fourier transform of the in-plane induced current, which can be expressed in terms of the 2D Ohm's law and is related to the induced areal charge density via the continuity equation in Eq. A.14. The external current density for an electron travelling at speed $\mathbf{v} = (\mathbf{v}_{\parallel}, v_z)$ in Fourier space goes as

$$\rho_{ext}(\mathbf{q}, z, \omega) = \frac{Ze}{v_z} \exp\left[\frac{i(\omega - \mathbf{q}\cdot\mathbf{v}_{\parallel})z}{v_z}\right]. \tag{2.91}$$

corresponding to a charge density of $\rho_{ext}(\mathbf{r}_{\parallel}, z, t) = Ze\delta(\mathbf{r}_{\parallel} - \mathbf{v}_{\parallel}t)\delta(z - v_z t)$, where $Z = -1$ for incident electrons. The next step is to determine the external potential that comes as a result of this external charge density, which can be determined by solving the inhomogeneous Helmholtz equation for the potential with the source term proportional to the external current density, giving

$$\phi_{ext}(\mathbf{q}, z, \omega) = \frac{4\pi Z e v_z}{(qv_z)^2 + (\omega - \mathbf{q}\cdot\mathbf{v}_{\parallel})^2} \exp\left[\frac{i(\omega - \mathbf{q}\cdot\mathbf{v}_{\parallel})z}{v_z}\right], \tag{2.92}$$

where $q = |\mathbf{q}|$. It is important to note that this equation relies on the quasistatic approximation introduced in Eq. A.10, i.e. that the response of the material to the external charge density is instantaneous.

The induced potential meanwhile can be expressed in terms of the polarizability shown in Eq. A.13, which when transformed to Fourier space yields

$$\phi_{ind}(\mathbf{q}, z, \omega) = -\frac{2\pi e^2}{q}\chi_0(\mathbf{q}, \omega)\left[\phi_{ind}(\mathbf{q}, z, \omega) + \phi_{ext}(\mathbf{q}, z, \omega)\right] e^{-q|z|}, \tag{2.93}$$

noting that this polarizability is associated with the monolayer and therefore has no z-dependence,

and that the factor $2\pi e^2/q$ is the 2D Coulomb potential shown in Eq. A.17. Plugging in Eq. 2.92 and using Eq. A.18 yields the induced potential at $z = 0$

$$\phi_{ind}(\mathbf{q}, 0, \omega) = \left(1 - \frac{1}{\varepsilon(\mathbf{q}, \omega)}\right)\phi_{ext}(\mathbf{q}, 0, \omega), \tag{2.94}$$

exactly as shown in Eq. A.18 for systems of any dimension, where $\varepsilon(\mathbf{q}, \omega)$ is the 2D dielectric function given in Eq. A.21. With expressions for the induced potential and the external charge density, we use the continuity equation to reformulate Eq. 2.83 in terms of these quantities

$$W = -\int dz \int d\mathbf{r}_\parallel \int dt \mathbf{J}_{ext}(z, \mathbf{r}_\parallel, t) \cdot (\nabla\phi_{ind}(z, \mathbf{r}_\parallel, t)) \tag{2.95}$$

$$= \int dz \int d\mathbf{r}_\parallel \int dt (\nabla \cdot \mathbf{J}_{ext}(z, \mathbf{r}_\parallel, t))\phi_{ind}(z, \mathbf{r}_\parallel, t) \tag{2.96}$$

$$= -\int dz \int d\mathbf{r}_\parallel \int dt \left(\frac{\partial}{\partial t}\rho_{ext}(z, \mathbf{r}_\parallel, t)\right)\phi_{ind}(z, \mathbf{r}_\parallel, t). \tag{2.97}$$

Expressing this energy loss in terms of the energy loss probability density in Eq. 2.86 yields [58]

$$P(\omega) = \frac{e^2}{2\pi^2}\int \frac{1}{q}\left[\frac{2qv_z}{(\omega - \mathbf{q}\cdot\mathbf{v}_\parallel) + (qv_z)^2}\right]^2 \mathrm{Im}\left[-\frac{1}{\varepsilon(\mathbf{q}, \omega)}\right] d\mathbf{q}, \tag{2.98}$$

the imaginary component taken in line with Eq. 2.85 and noting that the time derivative of the charge density brings down a factor i. For electrons incident directly perpendicular to the monolayer, $P(\omega)$ can be rewritten as

$$P(\omega) = \frac{4e^2}{\pi v_z^2}\int_0^{q_c} \frac{q^2}{\left[q^2 + \left(\frac{\omega}{v_z}\right)^2\right]^2} \mathrm{Im}\left[-\frac{1}{1 + \frac{2\pi i q}{\omega}\sigma(\mathbf{q}, \omega)}\right] d\mathbf{q} \tag{2.99}$$

where $\sigma(\mathbf{q}, \omega)$ is the 2D conductivity as per Eq. A.21, and q_c refers to the maximum momentum wavenumber transfer collected by the experimental device.

While the conductivity of an isotropic material like graphene generally depends on both $\mathbf{q}$ and ω, for a typical VEELS regime [109] where energy losses fall below $\hbar\omega \lesssim 50$ eV and incident electrons travel normal to the monolayer with speeds comparable to the speed of light, the q-dependent prefactor in Eq. 2.99 suppresses the contributions of wavenumbers exceeding $\omega/c \lesssim 0.015$ Å^{-1} to the integration. For this reason, it is expected that treating the conductivity in the optical limit $\sigma(\mathbf{q}, \omega) \approx \sigma(\omega)$ should yield a good approximate profile for the energy loss. In this limit, the integration over the wavenumber can be done analytically, giving [58]

$$P(\omega) = -\frac{4e^2}{\pi}\mathrm{Im}\left\{\frac{B}{\omega v_z}\left[G\left(\frac{q_c v_z}{\omega}\right) - G(0)\right]\right\} \tag{2.100}$$

where

$$B \equiv -i\frac{v_z}{2\pi\sigma(\omega)} \tag{2.101}$$

and

$$G(x) = \int \frac{x^2}{(x^2+1)^2(x+B)}dx \tag{2.102}$$

$$= -\frac{xB+1}{2(x^2+1)(B^2+1)} + \frac{B(B^2-1)\arctan(x)}{2(B^2+1)^2} + \frac{B^2}{2(B^2+1)^2}\ln\left(\frac{(B+x)^2}{x^2+1}\right). \tag{2.103}$$

If the maximum in-plane momentum q_c is large enough, the limit of integration can be extended to $q_c \to \infty$, since the prefactor to the imaginary part of the inverse dielectric function in Eq. 2.99 peaks strongly at $q = \omega/v_z \ll q_c$, allowing the use of the further approximation

$$G(\infty) = \frac{\pi}{4}\frac{B(B^2-1)}{(B^2+1)^2} \tag{2.104}$$

As mentioned earlier, the above equations are only valid using the quasistatic approximation, in this context referred to as the non-relativistic regime. To incorporate relativistic effects, we follow the formalism outlined in Refs. [21, 22, 71, 72, 73], where the probability density of energy loss $\hbar\omega$ for an incident electron moving at speed $\mathbf{v} = (0, v_z)$ in an isotropic material is expressed via

$$P(\omega) = \frac{1}{\hbar^2}\int_{|\mathbf{q}|<q_c} d\mathbf{q}\, F(\mathbf{q},\omega) = \frac{2\pi}{\hbar^2}\int_0^{q_c} dq\, q\, F(q,\omega). \tag{2.105}$$

$F(q,\omega)$ is called the joint probability density of the energy loss $\hbar\omega$ and the momentum transfer $\hbar q$, analogous to the differential scattering probability for bulk materials introduced in Eq. 2.87. In the relativistic regime, two separate mechanisms of energy transfer to the medium become viable. Ohmic energy loss refers to the electronic excitations in the target that ultimately give rise to Joule heating occurring within the conducting layer, while radiative energy loss refers to the total energy radiated by far-field electromagnetic fields both above and below the conducting layer, and occurs for frequencies $\omega > cq$. This second form of loss only comes to light in the relativistic regime where the speed of light is considered finite.

More formally, in a way analogous to Eq. 2.85, the energy losses associated with Ohmic and radiative loss can be written as [21]

$$W_{ohm} = \int dt \int dz \int d\mathbf{r}_\parallel\ \mathbf{J}_{ind}(z,\mathbf{r}_\parallel,t)\cdot\mathbf{E}(z,\mathbf{r}_\parallel,t) \tag{2.106}$$

$$= \int_0^\infty d\omega\,\omega \int d\mathbf{q}\ F_{ohm}(\mathbf{q},\omega) = \int_0^\infty d\omega\,\omega P_{ohm}(\omega), \tag{2.107}$$

recalling the in-plane $\mathbf{J}_{ind}$ in Fourier space as defined in Eq. 2.90 for a theoretically 2D material. The radiative energy loss, describing the energy radiated away by far field EM fields, is defined via the flux of the Poynting vector due to the induced electric and magnetic fields through a large surface S as

$$W_{rad} = \frac{c}{(2\pi)^2}\int_0^\infty d\omega\ \mathrm{Re}\left\{\oiint_S dS\ \hat{\mathbf{n}}\cdot[\mathbf{E}_{ind}(z,\mathbf{r}_\parallel,\omega)\times\mathbf{H}^*_{ind}(z,\mathbf{r}_\parallel,\omega)]\right\} \tag{2.108}$$

$$= \int_0^\infty d\omega\,\omega\int d\mathbf{q}\ F_{rad}(\mathbf{q},\omega) = \int_0^\infty d\omega\,\omega P_{rad}(\omega). \tag{2.109}$$

As derived in [21] for the case of isotropic materials, the joint probability density can be broken into these Ohmic and radiative contributions , $F(q,\omega) = F_{ohm}(q,\omega) + F_{rad}(q,\omega)$, where the directional wavevector dependence has been dropped. The Ohmic part is given by

$$F_{ohm}(q,\omega) = \frac{4}{\pi}\frac{e^2}{\beta^2\omega}\frac{(cq)^2\ \mathrm{Re}\left[\sigma\right]}{\left[\omega^2\left(\frac{1}{\beta^2}-1\right)+(cq)^2\right]^2}\times\begin{cases}\frac{1}{\left|1+\frac{2\pi}{c}\sigma\sqrt{1-\left(\frac{cq}{\omega}\right)^2}\right|^2}, & \omega > cq\\ \frac{1}{\left|1+i\frac{2\pi}{c}\sigma\sqrt{\left(\frac{cq}{\omega}\right)^2-1}\right|^2}, & \omega < cq\end{cases}, \tag{2.110}$$

while the radiative part given by

$$F_{rad}(q,\omega) = 8\frac{e^2}{\beta^2\omega}\frac{cq^2\sqrt{1-\left(\frac{cq}{\omega}\right)^2}}{\left[\omega^2\left(\frac{1}{\beta^2}-1\right)+(cq)^2\right]^2}\left|\frac{\sigma}{1+\frac{2\pi}{c}\sigma\sqrt{1-\left(\frac{cq}{\omega}\right)^2}}\right|^2, \quad \omega > cq, \tag{2.111}$$

where c is the speed of light, $\beta = v/c$, and $\sigma \equiv \sigma(q,\omega)$ is the in-plane conductivity of the sheet. The total energy loss density of the incident electron can therefore be decomposed into $P(\omega) = P_{\mathrm{ohm}}(\omega) + P_{\mathrm{rad}}(\omega)$. The Ohmic energy loss density has contributions from both below ($\omega < cq$) and above ($\omega > cq$) the light cone, while the radiative contribution consists entirely of contributions above the light cone. The $\omega < cq$ frequency range for a conducting layer corresponds to energy losses from collective modes and interband transitions of valence electrons in the monolayer. In the limit $\beta \to 0$, Eq. 2.105 reduces to Eq. 2.99.

2.4 Artificial Neural Networks

In theory, the human brain performs many fewer computations per second compared to a state of the art computer. Despite this, the brain's ability to take in new information through learning, performing tasks rooted in prior experience and knowledge, gives it an important edge. At its heart, the goal of artificial neural networks is to enable the same benefits that human learning confers into a viable computational framework.

Machine learning in this section will refer to the subcategory of supervised learning, where a set of training data $T = (x_n, t_n) : 1 \leq n \leq N$, along with a target function g with target values $t_n = g(x_n)$, is the input into the problem. The goal is to determine an approximation of this target function for all input values in the domain of g. As we will encounter in Chapter 4, this target function in some cases may even be unknown. Importantly, many classification and regression problems can be formulated as an effort to use an artificial neural network to mimic this mapping starting with a set of training data. These neural networks are made up of a set of connected processing units, each receiving input from either an external source or from other units and computing an output that can then be sent to other processing units, thereby replicating connections in the human brain [110]. Processing units consist of propagation rules which map all incoming input into a single input value, while activation functions are applied onto this input value to determine the output of the unit, the output itself also termed the activation. This is illustrated further in Fig. 2.5, representing a neural network with a single hidden layer, referring to a layer of neurons that only receives and forwards information between other layers of neurons. The output function will match the dimensions determined by the number of input and output units.

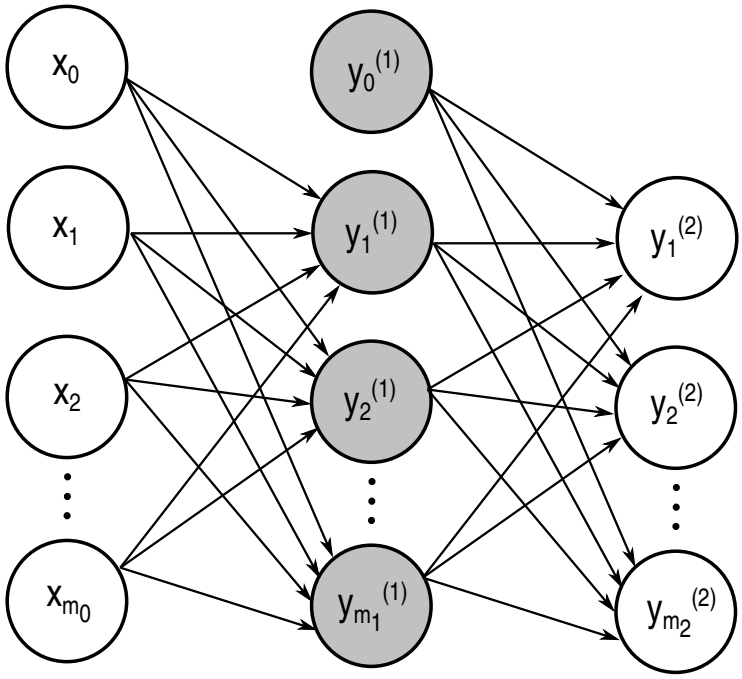

Figure 2.5: A fully connected feed-forward neural network with a single hidden layer. $x_1, x_2, \ldots x_{m_0}$ represent input to the network, with x_0 used as a bias introduced into the activation function. $y_1^{(1)}, y_2^{(1)}, \ldots, y_{m_1}^{(1)}$ are neurons each receiving input from the initial layer with $y_0^{(1)}$ acting as the bias. $y_1^{(2)}, y_2^{(2)}, \ldots, y_{m_2}^{(2)}$ represent the final output of the network.

The neural network shown in Fig. 2.5 is feed-forward, meaning that closed cycles in the network graph are prohibited and each layer only propagates to the one ahead, meaning that the

relationship between input and final output can be described by an explicit function. The basic equation for the input fed to the neuron y_i from m neurons x_k is given by [111]

$$y_i = f(z_i) = f\left(\sum_{k=0}^{m} w_{ik} x_k\right) \tag{2.112}$$

or, extending to a multilayered system

$$y_i^{(l)} = f(z_i^{(l)}) = f\left(\sum_{k=0}^{m_{(l-1)}} w_{ik}^{(l)} y_k^{(l-1)}\right) \tag{2.113}$$

where $m_{(l)}$ represents the number of neurons in layer l, and $w_{ik}^{(l)}$ are the weights for neurons of index k fed into neuron i located in layer l.

The activation function f in Eq. 2.113 determines the output of the unit. As will be shown, these activation functions will play a role in error backpropagation, so being nicely differentiable is important. Monotonic functions are also desirable as they prevent different inputs from mapping to the same output, which may lead to additional extrema in the error surface. The three most common choices for activation functions [110, 112] are

$$\sigma_{sigmoid}(z) = \frac{1}{1+\exp(-z)} \longrightarrow \frac{\partial\sigma(z)}{\partial z} = \sigma(z)(1-\sigma(z)) \tag{2.114}$$

$$\sigma_{softmax}(z,i) = \frac{\exp(z_i)}{\sum_k \exp(z_k)} \longrightarrow \frac{\partial\sigma(z,i)}{\partial z_j} = \sigma(z,i)(\delta_{ij} - \sigma(z,j)) \tag{2.115}$$

$$\sigma_{relu}(z) = \max(0,z) \longrightarrow \frac{\partial\sigma(z)}{\partial z} = H(x). \tag{2.116}$$

The logistic sigmoid and softmax functions are smooth and monotonic, and in addition they allow a probabilistic interpretation with all output restricted between 0 and 1. The derivatives of these two functions also offer a clean implementation without the sharp features of the Heaviside function that the rectified linear unit (ReLU) yields. For a long time, these two were the default activation functions, but more modern implementations of neural networks, especially convolutional neural networks (CNNs), favour the ReLU, thanks to its ability to truly output a zero value and the advantages of a linear activation function for optimization problems [113]. While proper calibration is required, ReLU will be the default choice for the networks explored in Chapter 4. With the activation functions in place, the problem of network training boils down to the determination of the weights w_{ik} in Eq. 2.113 starting with our set of training data. Assuming that, given enough neurons, there is some set of weights which will yield the best approximation to the target function, it is useful to evaluate the progressive performance of the network as more data is fed in and the weights are gradually adjusted. The distance measure between the approximation and the target function will be given by our choice of error term.

There are two primary choices for the evaluation of the error [114]. Sum of squared errors

go as

$$E = \sum_n E_n = \frac{1}{2} \sum_{nk} (y_k(x_n) - t_{nk})^2 \,, \tag{2.117}$$

where t_{nk} is the k^{th} entry of the n^{th} target value, and the cross-entropy error function

$$E = \sum_n E_n = - \sum_{nk} t_{nk} \log(y_k(x_n)). \tag{2.118}$$

For both of these error definitions, using that $y_i(x_n) = f(z_i)$, we have

$$\frac{\partial E_n}{\partial z_i^{L+1}} = \frac{\partial E_n}{\partial y_i^{L+1}} \frac{\partial y_i^{L+1}}{\partial z_i^{L+1}} = (y_i(x_n) - t_{ni}) f'(z_i^{L+1}) \tag{2.119}$$

The variable E in the above two equations is the error resulting from batch training, where a set of n input values $x_1, \ldots x_n$ are run through the network and the total error is used to update the weights in the network. This is in contrast to stochastic training, where the weights are updated based on the individual error E_n. As many training sets tend to have many redundancies, batch training tends to be the preferable approach [115]. Overall, this error function E can be considered as a surface across the space of all weights w_{ik} for which we desire to find a global minimum, the necessary criterion for a local minimum being $\nabla E_n = 0$. The following section outlines one of the most common methods for achieving a weight optimization for any feed-forward neural network.

2.4.1 Gradient Descent and Error Backpropagation

As an analytical solution is not usually possible for complex networks, an iterative approach is often best for minimizing the error. In each iteration step we choose an weight update Δw_t where

$$w_{t+1} = w_t + \Delta w_t \tag{2.120}$$

and t represents the current time step, or epoch, one epoch representing one update of the weights after a batch error has been computed. Gradient descent is a basic first-order optimization algorithm, and presents a nice balance between accuracy and computational cost compared to second-order algorithms like the Hessian [115]. In gradient descent, the gradient ∇E is used in the weight update Δw_t, determined by taking a step into the direction of negative gradient at position Δw_t on the error surface, such that

$$\Delta w_t = -\gamma \frac{\partial E}{\partial w_t} + \lambda \Delta w_{t-1} \tag{2.121}$$

where γ is the learning rate and λ is the momentum [116]. These hyper-parameters determine how quickly the weights update at each iteration and how much of previous iterations are in-

cluded in the weight update, respectively. The choice of γ is especially important, as it needs to be large enough that the network updates in a meaningful way to input data, but not so much as to cause large oscillations which destroy the network performance. The inclusion of momentum is one strategy towards reducing oscillations, although it is also common to start with a relatively fast learning rate and gradually decrease it after a certain error threshold is crossed or a plateau has been reached [117]. In the case of stochastic training the weights are updated according to the error E_n instead [115].

Error backpropagation proves to be a useful way to compute ∇E_n for the weight updates. For a feed-forward network with L hidden layers, $L+1$ being the output layer, the derivative of an error function E_n at input value x_n with respect to the weight w_{ij} is given by

$$\frac{\partial E_n}{\partial w_{ij}^{L+1}} = \frac{\partial E_n}{\partial z_i^{L+1}} \frac{\partial z_i^{L+1}}{\partial w_{ij}^{L+1}} = \delta_i^{L+1} y_j^L \tag{2.122}$$

where δ_i^{L+1} is the contribution of the i^{th} output unit to the error E_n. Going down to a hidden layer $l \in (1, L)$ we can follow a similar procedure

$$\frac{\partial E_n}{\partial w_{ij}^l} = \frac{\partial E_n}{\partial z_i^l} \frac{\partial z_i^l}{\partial w_{ij}^l} = \delta_i^l y_j^{l-1} = \left(\sum_k \frac{\partial E_n}{\partial z_k^{l+1}} \frac{\partial z_k^{l+1}}{\partial z_i^l} \right) y_j^{l-1} \tag{2.123}$$

where we utilize the definition of the error at a given layer to write out a formula for δ_i^l

$$\delta_i^l = f'(z_i^l) \sum_k w_{ik}^{l+1} \delta_k^{l+1}. \tag{2.124}$$

This final equation presents a recursive algorithm for calculating the gradient ∇E_n at the output layer by propagating the error at layer l back through the network to layer $l-1$ and evaluating the derivatives $\partial E_n / \partial w_{ij}^l$ that allow the weights to be updated via gradient descent. This backpropagation algorithm is presented graphically in Fig. 2.6.

Error backpropagation can be a computationally costly endeavour, as it scales with the total number of weights w_{ij} and the number of neurons in each layer, shallower networks scaling better than deeper ones due to the iterative nature of backpropagation [117]. A sufficiently large multilayer network where every neuron in one layer is connected to every neuron in the next layer means the cost of propagating an input value is dominated by the number of weights used. As a result, its just as important to attempt to reduce the amount of input flowing through the network as it is to keep the number of neurons as low as possible. This input flow can be reduced by methods like convolution [111] and pooling [115], which either compact input data before it comes into the network or group it in such a way that the connections between layers need not be from every neuron to every other neuron in the next and preceding layers.

Theoretically, by using a sufficiently high number of hidden neurons, a neural network can model any continuous target function to a sufficient degree of approximation [24]. After training

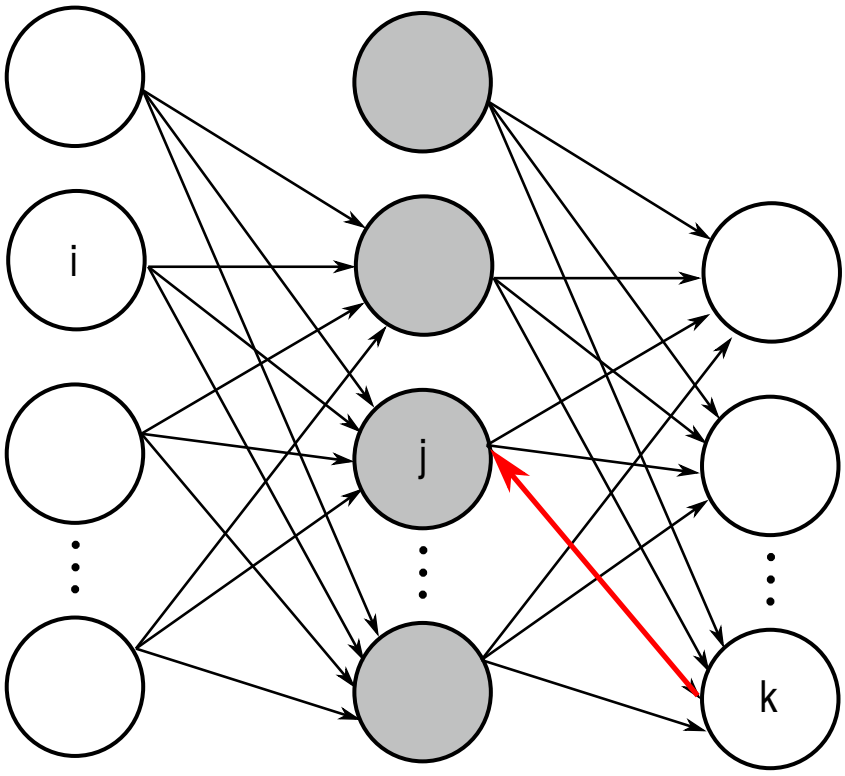

Figure 2.6: Visualization of error propagation through a single hidden layer neural network. Once the error term is computed and used to adjust the weights at the layer of interest, the error terms from the previous layer can be used to determine error terms downstream.

and minimizing the error function with respect to the network weights, it is important that the network is able to generalize beyond the training data set, and the easiest way to determine that is to test its performance on test data not used for training, called the validation set. A network is said to generalize if reduction in the error over the training data can in turn result in an overall improvement in the network's ability to predict the output of validation data. Regularization is one method which tries to avoid over-fitting and provide better agreement with data beyond the training data. Since over-fitting is symptomatic of having more hidden layer neurons than needed for the problem, the addition of a penalty to the error function, as in the case of L_2-regularization [114]

$$E(\omega)^* = E(\omega) + \eta w^T w, \tag{2.125}$$

helps ensure that weights of the network w tend exponentially to zero, effectively deactivating any neurons which do not contribute sufficiently to the overall error. Choosing the number of hidden layers, the number of neurons to use in each layer, the activation and error functions, and hyperparameters like γ, λ, and η are important considerations which often requires a substantial amount of trial and error before a computationally efficient network that well approximates the

desired target function is obtained.

Beyond the classical gradient descent method, one of the most often used weight update methods, known as optimizers, is Adam [118], the name referencing adaptive moment estimation. In Adam, the learning rate γ in Eq. 2.120 takes into account estimations of the first and second moments of the gradient of the error $g_t = \partial E/\partial w[t]$ at each time step t. These first and second moments of the gradient within the Adam optimizer, m_t and v_t respectively, are computed using exponentially moving averages using the gradient evaluated in the current batch,

$$m_{t+1} = \beta_1 m_t + (1-\beta_1) g_{t+1} \qquad v_{t+1} = \beta_2 v_t + (1-\beta_2) g_{t+1}^2, \tag{2.126}$$

where the values for the hyperparameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$ are the default in many software implementations of neural networks [112], and the initial moments are set to zero. Because of this initialization, the estimators are biased towards zero, so it is necessary to implement a bias correction so that the expected value of the moments match with those of powers of the gradient. This is done by

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \qquad \hat{v}_t = \frac{v_t}{1-\beta_2^t}. \tag{2.127}$$

Lastly, the weight update in Eq. 2.121 using the Adam optimizer becomes

$$\Delta w_t = -\gamma \hat{m}_t / \sqrt{\hat{v}_t + \epsilon}, \tag{2.128}$$

where ϵ is a small value introduced to prevent division by zero errors. We note that using the moving average of the gradient rather than the gradient itself usurps the role played by λ in Eq. 2.120. While Adam has some pitfalls related to not generalizing well to all problems that the classical gradient descent method is suited to handle [119], the savings in training time make it an appropriate choice for many neural network implementations, including the ones used in this thesis.

2.4.2 Convolutional Neural Networks and Max Pooling

Convolution [111] and pooling [115] are two ways to pre-process input data before it is fed into a feed-forward neural network. Convolution acts to manipulate the input to highlight important features, while pooling selects the most relevant input or an average from a group. Working in tandem they can help ensure that the input fed into the network is as small as possible while retaining the features of the input. In the terminology of convolutional networks, the input in the form of a 2D matrix I (for example) is manipulated by a smaller 2D matrix known as the kernel K, with the output referred to as a feature map S [111]

$$S(i,j) = (K * I)(i,j) = \sum_{mn} I(i-m, j-n) K(m,n). \tag{2.129}$$

This resulting matrix S will be smaller than the original input image array I, and this type of convolution operation can of course be applied to any dimension of data. Since convolution also captures information about many neighbouring elements in an array, convolution allows for the development of networks where not each element is connected to every other, allowing for a much quicker runtime of the network. The main objective of this operation is to extract higher-level features from the input data, for example edges in the case of pattern recognition, and it is common for convolutional neural networks to employ many layers of convolution with many different kernels, each kernel (often times very sparse matrices) corresponding to different expected features in the data.

If the presence of features is more relevant than their location in the data set, invariance to translation, meaning that translation of the input by a small amount leaves the values of the outputs unchanged, is a desirable property for our network to have. Pooling is one approach that helps to make this possible by grouping together nearby data, either by giving the maximum output in a given neighbourhood or through some sort of averaging and then mapping of output values to a reduced output [115]. On top of ensuring invariance, pooling is essential for handling inputs of varying size, as one might see for differently sized images run through an image classifier. Having a different neural network for each size of image is a much worse approach than a combined convolution/pooling approach that pre-processes the input to a usable input for the feed-forward network. Lastly, pooling serves to reduce the size of the input by feeding the neural network averaged out features, an important tool when training computationally expensive networks [33, 39].

Chapter 3

Graphene and Other Low-Dimensional Materials

3.1 Graphene

Despite having initially been predicted to be experimentally unviable due to quantum fluctuations, two-dimensional materials have increasingly become an active topic of research due to their peculiar properties. Graphene is likely the most famous example, and it came onto the scene in 2004 through the application of scotch tape to graphite, which is composed of many layers of spaced graphene. It follows the discovery of other assemblies of carbon atoms into different configurations, with 0D fullerene, stable molecules of 60 carbon atoms arranged in a geodesic structure, coming in 1985 [120], and carbon nanotubes, discovered sometime between 1952 and 1991, which are cylindrical lattices of rolled-up graphene with small radii and considerable tensile strength. Among these peculiar properties, many of which can be traced to the influence of unbonded sp^2 orbitals above and below each carbon atom, graphene exhibits a linear dispersion relation for frequencies between the visible and THz range, incredibly high room temperature electrical conductivity, and strong electronic tunability as a result of either doping or gating.

Graphene is a honeycomb lattice made up of sp^2-hybridized carbon atoms arranged in a 2-dimensional structure, each spaced approximately 0.142 nm from each other, presented later in Fig. 5.10. At each carbon atom lie three evenly spaced σ-bonds and one delocalized π bond, which points perpendicular to the sheet. Both the unit cell and the reciprocal cell are rhombuses, the direct lattice being composed of two overlapping triangular lattices. Labelled in the reciprocal lattice in Figure 3.1 are four key symmetry points. In the band structure of graphene, these will show up as either saddle points in the energy dispersion or self-crossings, and become relevant in later chapters as we seek to explore peaks in the EELS spectrum of monolayer graphene. At the symmetry point M, transitions occur between π and π^* energy bands, while interband transitions at the Γ symmetry point occur at even higher energies. A plot of the band structure can be found in Fig. 5.8. These transitions can be modeled within different approximations, in particular

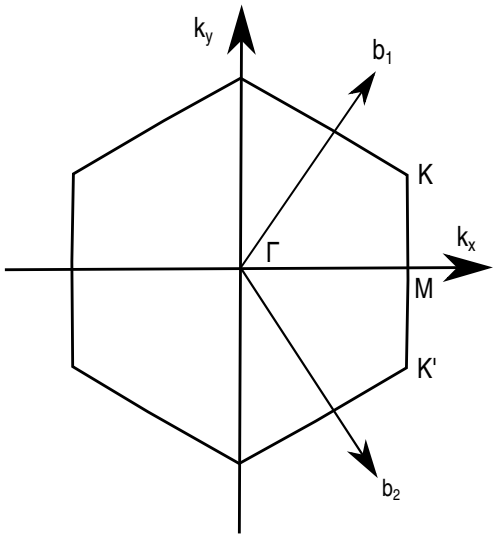

Figure 3.1: The first BZ of graphene, the momentum space dual to the elementary Wigner-Seitz cell. Labelled are the main symmetry points Γ, M, K, and K', the last two being indistinguishable due to the hexagonal symmetries present in the real space graphene lattice.

the *low-energy* spectrum, where electronic interactions are dominated by the unusual energy dispersion near the Dirac point. The relatively high electron density of the material suggests that classical electromagnetic theory will work well here, allowing for an electronic description of the graphene layer in terms of optical response functions.

The Dirac points K, K' are locations that, within a tight-binding model for nearest-neighbour interactions, yield a meeting point between the valence and conduction bands for undoped graphene under ideal conditions. The linear band dispersion near these points follows

$$\varepsilon_{K,K'}(\vec{k}) = \hbar k v_F \tag{3.1}$$

where ε is the energy derived from the tight-binding Hamiltonian, and $v_F \approx 10^6 m/s$ is the Fermi velocity. This kind of linear dispersion mimics the dispersion seen for photons, leading many to research graphene for effects like Klein tunneling and the quantum Hall effect. The density of states for energies near the Dirac point will therefore also follow a linear relationship unique to two-dimensional systems

$$D(E) = \frac{2|E|}{\pi(\hbar v_F)^2}, \tag{3.2}$$

which has important consequences for the absorption of graphene at lower energies.

The difference in scale between the wavelengths of probes in graphene and the characteristic wavenumber scales of the BZ shown in Fig. 3.1 and the mean free electron path suggests a local frequency response is often sufficient towards describing the optical properties of graphene for low energies. An approximation for the optical conductivity (i.e. the conductivity in the optical,

or $q \to 0$, limit) can be derived starting from the tight-binding approximation [57]. This derived conductivity is broken into two parts, the first being the interband contribution $\sigma_I = \sigma_r + i\sigma_i$, with

$$\sigma_r = \sigma_0 \left(1 + \frac{1}{\pi} \tan^{-1} \frac{\hbar\omega - 2E_F}{\hbar\gamma_{\text{inter}}} - \frac{1}{\pi} \tan^{-1} \frac{\hbar\omega + 2E_F}{\hbar\gamma_{\text{inter}}}\right) \tag{3.3}$$

and

$$\sigma_i = -\sigma_0 \frac{1}{2\pi} \ln \frac{(\hbar\omega + 2E_F)^2 + \hbar^2\gamma_{\text{inter}}^2}{(\hbar\omega - 2E_F)^2 + \hbar^2\gamma_{\text{inter}}^2} \tag{3.4}$$

where $\sigma_0 = e^2/4\hbar$ is called the universal conductivity, γ_{inter} is the damping rate for interband transitions, and E_F is the Fermi energy, which will be non-zero for doped graphene. The term σ_0 is known as the universal conductivity as it represents an absorbance of all frequencies below a certain threshold in the low-energy range. The optical conductivity in the case of doped graphene also receives a contribution from the Drude conductivity term

$$\sigma_D = \sigma_0 \frac{4E_F}{\pi} \frac{1}{\hbar\gamma_D - i\hbar\omega} \tag{3.5}$$

derived from the intraband contribution. It is important to note that the damping rate γ_D in the intraband case does not necessarily need to equal γ_{inter} as the mechanisms for dissipation for the two types of excitations may be different. Chapter 6 goes deeper into showing computational and phenomenological models for describing the conductivity in the low-loss region accessible to EELS.

3.1.1 Empirical Models for the Optical Conductivity of Graphene

An empirical model for the optical in-plane conductivity of graphene over a wide energy range needs to account for the many different ways electrons in the monolayer can be excited. For doped graphene, the Drude intraband contribution and Pauli blocking of interband transitions becomes relevant at low frequencies. Three types of interband transitions are relevant in the $0 - 20$ eV range for monolayer graphene: the $\pi \to \pi^*$ transitions near the BZ K points (see Fig. 3.1), which lead to the universal value of absorption for undoped graphene, along with $\pi \to \pi^*$ and $\sigma \to \sigma^*$ high-energy transitions near the M points of the BZ (see Fig. 5.8). These last two, as will be shown in Chapter 6, correspond to peaks in the EELS spectrum of monolayer graphene.

Starting with free-standing, intrinsic, undoped graphene, a model for the conductivity can be constructed by treating the electrons of the π and σ transitions in a two-fluid hydrodynamic

model [11], as shown in Appendix B. Using Eq. A.15, this conductivity is given by

$$\sigma_{\text{hd}}(\omega) = -i\omega\left[\frac{n_\pi^0/m_\pi^*}{\omega_{\pi r}^2 - \omega(\omega + i\gamma_\pi)} + \frac{n_\sigma^0/m_\sigma^*}{\omega_{\sigma r}^2 - \omega(\omega + i\gamma_\sigma)}\right], \tag{3.6}$$

where $\nu = \sigma, \pi$ are the two fluids. The two restoring frequencies $\omega_{\pi r}$ and $\omega_{\sigma r}$ empirically refer to the location of the high-energy interband transitions near the M points, and it has been shown that the in-plane energy loss for 2 eV $\lesssim \hbar\omega \lesssim$ 30 eV [96, 121] can be well described by letting $\hbar\omega_{\pi r} \approx 4$ eV and $\hbar\omega_{\sigma r} \approx 14$ eV, respectively. The widths of the peaks in the EELS spectrum will be given by the broadening constants γ_π and γ_σ, which can be chosen based on either TDDFT calculations of the optical conductivity or on experimental data [22, 58].

As seen in Eq. 3.3, more work is required to have this empirical model describe the conductivity for frequencies ranging from zero up to the visible range, where graphene exhibits its trademark universal absorbance character, resulting from the $\pi \to \pi^*$ transitions near the BZ K points mentioned before and calculated within the Dirac-cone approximation [122]. To ensure this levelling-off to σ_0 in the low frequency region $0 \leq \hbar\omega \lesssim 2$ eV, a correction to Eq. 3.6, called the Dirac correction, is included. However, as seen in Fig. 5.8, electrons from the Dirac transition and the $\pi \to \pi^*$ transition come from the same band, so to preserve the number density of electrons from this band participating in interband transitions, a factor $f \in [0, 1]$ is included multiplicatively in $\sigma_{\text{hd},\pi}(\omega)$, while the Dirac correction of the form $\sigma_0/[1 + (\omega/\omega_c)^4]$ is added to $\text{Re}\,[\sigma_{\text{hd}}(\omega)]$, where ω_c represents a cutoff frequency for the correction.

To codify this preservation of electron number density for all participating electrons, we rewrite the f-sum rule in Eq. 2.78 as

$$\int_0^\infty \text{Re}[\sigma(\omega)]d\omega = \frac{\pi e^2}{2m_e} n_{at} N_e, \tag{3.7}$$

where $N_e = N_\sigma + N_\pi = 4$ is the total number of valence electrons per carbon atom. Choosing m_π^* and m_σ^* to equal the free electron masses will ensure that $N_\sigma = 3$ and $N_\pi = 1$ when $\sigma_{\text{hd}}(\omega)$ is evaluated, as expected based on the occupation of orbitals associated with each band. Given that the other free parameters in the model will be chosen to give best results with experimental or computational data, this f-sum rule ensures a relation between the cutoff frequency ω_c and the reduction factor f, given by $f = 1 - \omega_c m_e \sqrt{2}/(8 n_{at})$. Eq. 3.7 can be rearranged to define the number of valence electrons at energies up to ω' that contribute to the response function

$$N_e(\omega) = \frac{2m_e}{\pi e^2 n_{at}} \int_0^\omega \text{Re}[\sigma(\omega')]d\omega', \tag{3.8}$$

with the expected result that increasing the frequency range allows for the inclusion of more valence electrons to participate in the excitation.

By including the universal conductivity phenomenologically, we have derived an extended hydrodynamic (eHD) model, which factors in all three main interband transitions of intrinsic

graphene. The real part is

$$\mathrm{Re}\left[\sigma_{\mathrm{ehd}}(\omega)\right] = \sigma_0 \frac{\omega_c^4}{\omega_c^4 + \omega^4} + e^2\omega^2 \left[f \frac{\gamma_\pi n_\pi^0 / m_\pi^*}{(\omega_{\pi r}^2 - \omega^2)^2 + \gamma_\pi^2 \omega^2} + \frac{\gamma_\sigma n_\sigma^0 / m_\sigma^*}{(\omega_{\sigma r}^2 - \omega^2)^2 + \gamma_\sigma^2 \omega^2} \right], \tag{3.9}$$

whereas its imaginary part is obtained via the Kramers-Kronig relations shown in Eq. 2.76 [58].

Turning now to doped graphene, the Fermi energy ε_F will be shifted away from the Dirac K point to some finite position inside the conduction or valence π bands, depending on the type of doping. Typical doping in graphene is on the order of $\varepsilon_\mathrm{F} \lesssim 1$ eV, corresponding to an excess of charge carriers on the order of $n_i \lesssim 10^{14}$ cm^{-2} according to the Dirac cone approximation [123]

$$\varepsilon_\mathrm{F} = \hbar v_\mathrm{F} k_\mathrm{F} = \hbar v_\mathrm{F} \sqrt{\pi n_i}. \tag{3.10}$$

At absolute zero temperature, all interband electron transitions with energies $\hbar\omega < 2\varepsilon_\mathrm{F}$ are forbidden according to the Pauli exclusion principle, giving a jump-like real part of the interband conductivity, corresponding to

$$\mathrm{Re}\,\sigma_{\mathrm{inter}}^{\mathrm{step}}(\omega) = \sigma_0\,\Theta(\hbar\omega - 2\varepsilon_\mathrm{F}) \qquad \mathrm{Im}\,\sigma_{\mathrm{inter}}^{\mathrm{step}}(\omega) = \frac{\sigma_0}{2\pi} \ln\left[\frac{(2\varepsilon_\mathrm{F} - \hbar\omega)^2 + \gamma_P^2}{(2\varepsilon_\mathrm{F} + \hbar\omega)^2 + \gamma_P^2} \right], \tag{3.11}$$

where Θ is the Heaviside unit step function [122, 124]. For modeling purposes, this step function can be adjusted via a phenomenological broadening γ_P to describe interband dissipative processes [125] or convergence parameters in TDDFT calculations if they are employed, giving [57, 126]

$$S(\omega) = 1 - \frac{1}{\pi}\left[\arctan\left(\frac{2\varepsilon_\mathrm{F} - \hbar\omega}{\gamma_P} \right) + \arctan\left(\frac{2\varepsilon_\mathrm{F} + \hbar\omega}{\gamma_P} \right) \right]. \tag{3.12}$$

To introduce the aforementioned Pauli blocking into the eHD model, the smoothed step function in Eq. 3.12 is adopted, and we postulate that the real part of the optical conductivity can be faithfully expressed by

$$\mathrm{Re}\left[\sigma_{\mathrm{inter}}(\omega)\right] = S(\omega)\,\mathrm{Re}\left[\sigma_{\mathrm{ehd}}(\omega)\right], \tag{3.13}$$

with $S(\omega)$ given in Eq. 3.12 and $\mathrm{Re}\left[\sigma_{\mathrm{ehd}}(\omega)\right]$ in Eq. 3.9. The imaginary part of the interband contribution can again be evaluated using the KK relations in Eq. 2.78.

The final addition to the model for doped graphene comes from the low-energy intraband electron transitions that occur near the Fermi level, which has been shifted into either the conduction or valence band. Since these transitions are low energy, we again invoke the Dirac cone approximation for the π and π^* bands. This contribution to the optical conductivity takes the form

$$\sigma_{\mathrm{intra}}(\omega) = \frac{i}{\pi} v_B \frac{\varepsilon_\mathrm{F}}{\hbar\omega + i\gamma_D}, \tag{3.14}$$

where the phenomenological damping constant γ_D for intraband transitions may be different from the broadening γ_P used for smoothing or the γ_ν used in the interband transitions. Altogether,

a phenomenological expression for the conductivity of doped graphene over a wide range of frequencies can be written as $\sigma(\omega) = \sigma_{\text{intra}}(\omega) + \sigma_{\text{inter}}(\omega)$, referred to hereafter as the "Drude+ eHD" model, with the caveat that the eHD has been truncated at low frequencies as a result of the smoothing function. A simpler version of this model that is of use for applications in the THz to MIR frequency band would be to only add together the intraband and Drude contributions neglecting the hydrodynamic portion of the conductivity entirely, referred to as the "Drude+ step" model for the optical conductivity.

3.2 Phosphorene

Another 2D material which has attracted significant attention is single layer black phosphorus, known as phosphorene. As with graphene, phosphorene has a tunable direct bandgap with strong layer dependence and a high mobility of charge carriers, reaching $1000 cm^2/Vs$ at room temperature [56], both desirable quantities for integration into optoelectronic devices. The band gap, ranging from 0.3eV in the bulk to 1.7eV for monolayer phosphorene and tunable through either strain, doping, or the application of electric fields, covers a broad frequency range which is slightly higher than graphene. The atomic structure is shown in Fig. 3.2, showing that a single layer of phosphorene is made up of two sublayers of phosphorus atoms each covalently connected to three neighbouring atoms. In the case of multilayered black phosphorus each of these layers would be held together by weak van der Waals interactions. As a result of both armchair and zigzag directions appearing in the atomic structure, as seen in Fig. 3.2, phosphorene reveals a unique in-plane anisotropy which exhibits itself in the electrical, optical, thermal, and mechanical properties of this material [127].

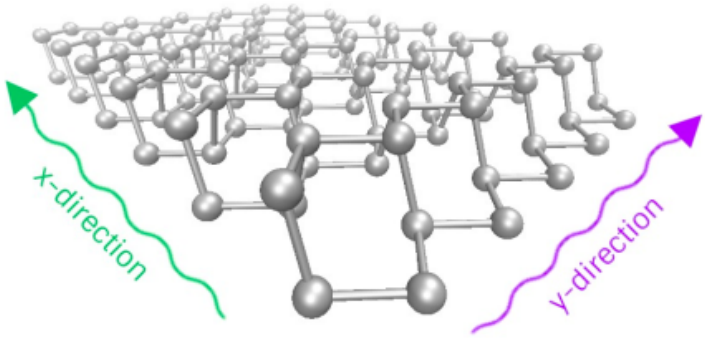

Figure 3.2: 3D visualization of a monolayer phosphorene lattice in real space, with arrows showing the x- and y- directions. The lattice is formed of two sublayers of phosphorus atoms, each atom covalently connected to three neighbouring atoms, giving rise to both armchair and zigzag features.

While most attention is given to the interband transitions of phosphorene, namely the band gap, the plasmonic response from free carriers in a naturally anisotropic material such as phos-

phorene also prove interesting. Theoretical studies of the plasmonic properties of black phosphorus [128] show that the band anisotropy can lead to hyperbolic plasmon polaritons, i.e. the frequency contours of the collective excitation modes in planar wavenumber space has a hyperbolic shape, as a result of a coupling between interband and intraband conductivities. Such a plasmon topology may prove promising for the directional propagation of plasmonic rays and in planar photonics thanks to the resulting large photonic density of states resulting from these hyperbolic modes. Combining phosphorene with heterostructures or influencing the band structure through gating and strain allows further tunability of these hyperbolic modes.

The appearance of hyperbolic modes requires an anisotropy present in the conductivity tensor, unlike the case of graphene where $\sigma_{xx} = \sigma_{yy}$, giving an isotropic dispersion of plasmon modes. In anisotropic 2D materials like phosphorene, the conductivity along the two principle axes will have different values. It has been shown theoretically [128] that for a given frequency ω,

$$\frac{q_x^2}{\mathrm{Im}[\sigma_{yy}(\omega)]} + \frac{q_y^2}{\mathrm{Im}[\sigma_{xx}(\omega)]} = \frac{1}{2\pi}\left|\sqrt{q_x^2+q_y^2}\right|\omega\left(\frac{1}{\mathrm{Im}[\sigma_{xx}(\omega)]\Im[\sigma_{yy}(\omega)]} - \frac{4\pi^2}{c^2}\right) \tag{3.15}$$

describes the plasmon modes for low energies in phosphorene. Hyperbolic modes will appear for $\mathrm{Im}[\sigma_{xx}(\omega)]\,\mathrm{Im}[\sigma_{yy}(\omega)] < 0$, which DFT has verified occurs for frequencies in the infrared region for phosphorene [56, 129].

3.3 Single-Walled Carbon Nanotubes

Single-walled carbon nanotubes (SWCNTs) have drawn attention across many physical disciplines due to their unique physical properties, including ballistic conductance [89], tunable band gaps[130], photoluminescence [10], and high optical absorbance [9], the last three coming about because of the same unbonded sp^2 orbitals that lead to many of graphene's interesting properties. With methods available for separating manufactured SWCNTs based on their chirality, there are many applications, from their use as additives in organic photovoltaic devices due to their intense absorption peak [9]s, or as electrodes due to their ballistic properties [89], or for addition into solar cells due to their thermal stability [10].

With experimentally measured radii ranging from 0.7 to 1.0 nm [95], and a length to diameter ratio on the order of 100, SWCNTs can be classified as 1D materials. Depending on how the carbon nanotube is rolled up, the orientation of the six-atom carbon ring within the lattice relative to the axis of the nanotube will change, leading to different lattice symmetries which change the absorption and conduction properties of the SWCNT [11]. These lattice symmetries can be either zigzag, armchair, or chiral, the former two presented in Fig. 3.3. The chiral vector C_h in terms of the real space unit vectors a_1 and a_2 of the hexagonal graphene lattice follows

$$C_h = n\mathbf{a}_1 + m\mathbf{a}_2 \equiv (n, m) \tag{3.16}$$

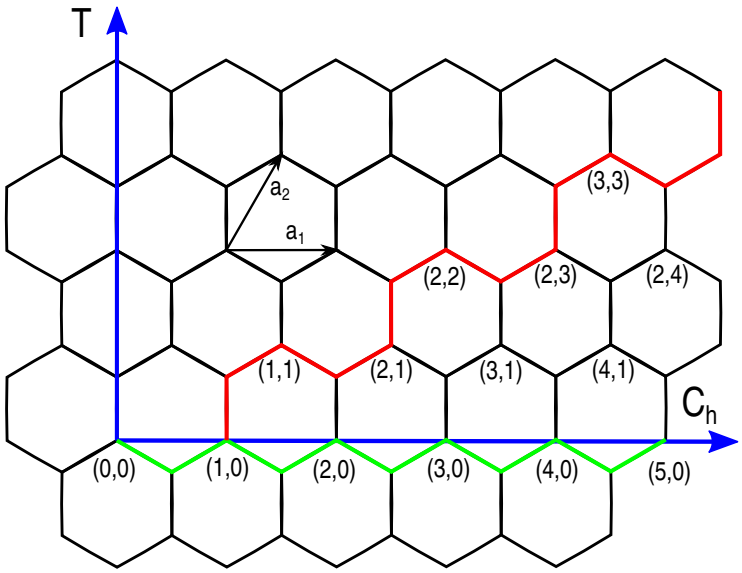

Figure 3.3: Chiral (C_h) and translational (T) vectors projected onto the atomic structure of a graphene monolayer. C_h points along the circumference of the tube while T points along the nanotube axis. Zigzag $(n, 0)$ (green) and armchair (n, n) (red) lattice symmetries occur for angles parallel to the chiral vector and 30 degrees from the chiral vector, respectively.

where $n > m$ are integers. Armchair nanotubes correspond to the case of $n = m$ and zigzag to $m = 0$. The circumference of the nanotube is given by

$$C = a\sqrt{m^2 + n^2 + nm} \tag{3.17}$$

where the lattice constant of graphene $a \approx 2.49$Å. The indices (n, m) uniquely determine the geometry of a given nanotube, including the radius and the number of graphene unit cells used to construct the unit cell of the SWCNT. From a computational standpoint, armchair and zigzag nanotubes admit many more symmetries and require far fewer atoms per unit cell, meaning calculations with nanotubes with chirality (n, n) or $(n, 0)$ will tend to require fewer resources [9, 89]. The effect of chirality is also seen directly in the band gap of the nanotube: $n-m \equiv 1\text{mod}3$ is a sufficient condition for semiconducting nanotubes while $n - m \equiv 0\text{mod}3$ will yield a metallic nanotube, although metallic features can also be seen for nanotubes with small diameters. This modular condition comes about from the number of graphene unit cells used to construct the unit cell, determining how different symmetry points fold into each other.

As with graphene and phosphorene, SWCNTs find many uses in optoelectronic applications, and it is therefore worthwhile to use the DFT techniques shown in Section 2.1 to make predictions about the axial dielectric function and axial conductivity, computational results which can be directly compared and contrasted with photoabsorption data, where light parallel to the principal nanotube axis is used to excite intense transitions between the valence and conduction bands.

3.4 Fullerene

Fullerene (C60), also known as a buckyball, is a zero-dimensional carbon allotrope, made up of sixty sp^2 hybridized carbon atoms, and first discovered accidentally during a study of galaxy clusters using mass spectroscopy. It has 32 faces, including 20 hexagons and 12 pentagons, forming a soccer ball pattern, with double bonds within the hexagonal rings but not in the pentagonal ones. Notable for being the first of the discovered carbon allotropes [131], its high stability means it can withstand high temperature and pressure since it is extremely stable, and has strong solubility and reactivity properties [132]. Additionally, doping fullerene molecules can change the electronic profile enough to make them electrically insulating, conducting, semi-conducting or even superconducting. It can be incorporated into optical devices, photovoltaics, and lubricants, among others [133].

Ultimately, as a stable, spherical, molecular system with prominent excitonic peaks resulting from $\pi \rightarrow \pi^*$ transitions, fullerene provides an excellent test case for the application of DFT to materials of the lowest dimension, especially in composite systems where the activation of modes within the fullerene molecules can hybridize or couple with those of nearby materials [133, 134], enabling the prediction, design, and optimization of relevant technologies. It is important to note that care must be taken in the application of DFT to molecular systems, where the scale of the molecule can be much smaller than the unit cell used to contain it, and relations between the atomic polarizability and the dielectric function may become more complicated than those derived in Eq. 2.67 [95].

3.5 Chlorophyll

All life on Earth depends on photosynthesis, the mechanism through which organisms absorb energy from photons by converting carbon dioxide and water into organic compounds, and is the main natural process for converting solar into chemical energy. Photosynthetic organisms such as algae, plants, and bacteria produce the oxygen we breathe in, the coal we burn, and the food we eat. The process starts with solar energy transfer to a reaction center, often performed by a network of chlorophyll pigments, which then aggregates together with the pigments to form what are known as light-harvesting complexes [135, 137]. Since different environments with a multitude of living organisms span the globe, chlorophyll molecules in complexes such as these can show small variations in their absorption properties to help adapt to the range of available solar spectrum and to facilitate energy transfer in the organism [50]. Acknowledging these variations, the major light-harvesting complex in green plants is known as the light-harvesting complex II (LHC II) [135].

Chlorophylls in photosensitive systems harvest incident photons at distinct wavelengths and transfer their energy as excitons to reaction centers, porphyrin rings consisting of a magnesium atom in the center with four nitrogen atoms surrounding it [46, 138]. Diversity of chlorophylls

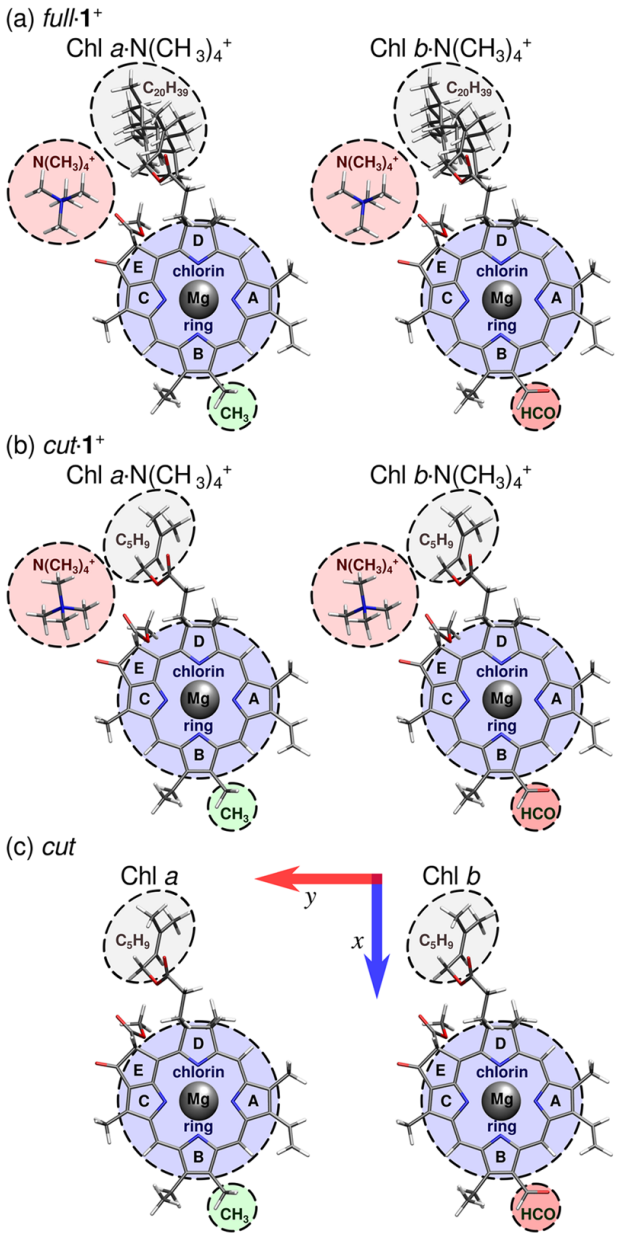

Figure 3.4: Depiction of (a) *full*·$\mathbf{1}^+$, (b) *cut*·$\mathbf{1}^+$, and (c) *cut* chlorophyll *a* (Chl *a*) and chlorophyll *b* (Chl *b*) structures with Mg atom, chlorin ring (blue), methyl group (CH_3, green) or aldehyde (HCO, red) groups, full ($C_{20}H_{39}$, grey) or cut (C_5H_9, grey) hydrocarbon chain, and a tetramethylammonium charge tag ($N(CH_3)_4^+$, red). Mg, O, C, N, and H atoms are depicted in silver, red, grey, blue, and white, respectively. Orientation of the *x* and *y* polarization axes (blue and red arrows) and labelling of the rings (**A**–**E**) follow IUPAC-IUB nomenclature [135, 136]. The structures in (a) *full*·$\mathbf{1}^+$ are based on those provided in Bruce et al. [137].

in nature will come about from the side chains attahced to this porphyrin ring. Chl a and Chl b, with chemical formulas $C_{55}H_{72}MgN_4O_5$[138] and $C_{55}H_{70}MgN_4O_6$[137], consist of 137 and 136 atoms, respectively, the difference being the addition of either a methyl group on the side chains of Chl a or an aldehyde group for Chl b. These two molecules form the fundamental functional units of LHC II, so determining their photoexcitation process has provoked interest in the organic photovoltaic cell [139], optoelectronic device [46], and food production communities [135]. As will be shown in Chapter 5, DFT, with the proper implementation [140], can play a role in determining the absorption characteristics of monomers like Chl a and Chl b and their larger roles within the full light-harvesting complex.

Chapter 4

Training Artificial Neural Networks with xc-Functionals

The Kohn-Sham ansatz, explained in Section 2.1.1, shows that a many body system of interacting electrons can be mapped to a system of non-interacting electrons that interact via an effective potential. This turns what was previously a differential equation that depends on the positions and momenta of all electrons in the system into a set of KS equations whose solution is computationally feasible. This effective potential, defined in Eq. 2.13, contains external, Coulomb, and xc components. The latter of these accounts for all quantum mechanical interactions, as well as any effects left out by the other two terms and the non-interacting kinetic energy operator. We can consider this xc potential in the context of the Hohenberg-Kohn theorem [1], which states that this effective potential uniquely defines the ground-state electron density and vice versa. This means that there must exist an exact form for the xc functional which depends solely on the electron density $n(\mathbf{r})$. It is clear that determining an exact form for this functional is very important, but it also turns out to be a very challenging task, and has been an open problem for over half a century.

Section 2.1.2 showed that many higher levels of xc functionals are needed depending on the type of material under consideration. The LDA xc functional in Eq. 2.17 often works well for predicting the properties of metals with a reasonable accuracy but fails for systems with much sharper variations in the electron density such as atoms. Generalized gradient approximations (GGAs) such as PBE in Eq. 2.20 use the local gradient of the electron density as an additional input to provide a more accurate and complete picture of the properties of molecular systems. However, they suffer from their own problems related to overbinding of atoms and the effects of static correlation in Hartree-Fock theory. At even higher orders, terms which only depend on the electron density in an implicit way such as the KS wavefunctions are mixed in with semi-local methods such as PBE to provide an even better approximation to the material properties of a system. As explored in Section 2.1.2, these methods can be quite computationally demanding, due to both the nested self-consistent cycle required to determine the xc potential and to the prohibitive costs of using the full orbital information, possibly even unoccupied orbitals, in the

functional calculation. A slightly less ambitious task than finding the exact form of the xc functional for all materials may be to determine xc functionals which can consistently take in the full electron density explicitly and return xc energies in agreement with the higher-tier functionals, thereby circumventing the major computational roadblock of these methods.

Section 2.4 introduced the concept of artificial neural networks, which in the context of supervised learning provide a method to learn the mapping between inputs and outputs by providing a set of training points. In fact, according to the universal approximation theorem [141, 142], feed-forward neural networks with sufficient width and depth can approximate any continuous function of finite dimensional input. The quality of this approximation of course depends on there being a sufficient quantity of useful data for the network to train on, a careful determination of how the data is pre-processed, and a network with hyperparameters chosen so as to best benefit from the data available. We outline in this chapter our progress in developing neural network approximations that can emulate two commonly used xc functionals, LDA and PBE, with sufficient accuracy to allow these learned approximations to be used interchangeably with their originals within a fully-functional DFT code. The challenges and obstacles that come from learning these relatively basic xc functionals offer some insight into tackling both of the tasks mentioned above. In what follows, NNLDA and NNPBE are used to refer to neural network approximations to the spin-paired LDA and PBE xc functionals, shortened to LDA and PBE throughout the text. The TensorFlow framework using Keras neural network libraries [112] is utilized throughout, while all DFT calculations in this section are done using GPAW.

4.1 Assessing the Effects of Noise

Before using neural networks to approximate the exact xc functional, it is instructive to first determine the level of accuracy required to achieve the goal of "interchangeability" and whether this is even feasible. For the LDA and PBE approximations, with well-defined mappings between the electron density, the xc energy density defined in Eq. 2.17 and Eq. 2.20, and the xc potential defined in Eq. 2.18 and Eq. 2.21, one clear measure of the accuracy of our prediction is the relative error of our learned xc functionals for every possible density input. For more complicated xc functionals where non-local dependence on the density or the KS orbitals is introduced, the ultimate test will be in how well the learned xc functional is able to predict material properties relative to its benchmark standard method. While certain DFT codes are specifically designed to excel at predicting chemically interesting properties such as atomization or ionization energies, we focus in this work on predicting the minimized ground state energy of a system. Generally, 50 meV is the range of error expected [40] between both implementations of LDA and PBE across different DFT software packages and DFT calculations themselves. For this reason we have chosen it as the golden standard for our NNLDA and NNPBE implementations.

To determine what impact errors in our implementation of PBE might have on the resulting energies, we have applied both unbiased and biased noise to the PBE xc energies and considered the deviations in resulting total energies for a benchmark system. Table 4.1 presents the mean

Error Percentage	Noise Factor Function	
	$\sin(x)$ error (meV)	$\sin^2(x)$ error (meV)
0.5%	131.9 ± 96.1	929.4 ± 644.1
0.1%	25.9 ± 18.6	189.1 ± 130.3
0.05%	13.4 ± 9.2	96.95 ± 65.6

Table 4.1: The mean standard error and standard deviation of the final total energy (meV) using the noisy PBE for 46 molecular compounds in the GPAW G2-1 database [16], where the error percentage defines the amplitude A of either sinusoidal (zero mean) or square sinusoidal (non-zero mean) noise introduced to PBE.

standard error and standard deviation of the ground state energy of 46 molecular compounds in the GPAW G2-1 database for noise added to both ε_{xc} and $\hat{V}_{xc}$ of the sinusoidal form $\mathrm{NF}_{A,f} = Af\sin(|f + 1/f| + \phi)$ or the square sinusoidal form $\mathrm{NF}_{A,f} = Af\sin^2(|f + 1/f| + \phi)$, where A is the amplitude of the noise, f refers to either the xc energy density or the xc potential, and ϕ is a randomly chosen angle. For each molecular compound in the database, an amplitude was chosen and the total energy of the system was computed for three random values of ϕ. The mean standard error was then computed as the average absolute distance of this total energy from the original PBE, $\sum_i |E_{PBE} - E^{noisy}_{PBE}(\phi_i)|/3$. The standard deviation reported is computed over all the materials shown in Appendix D. These noise functions are chosen to have either zero mean in the sinusoidal case or a non-zero mean for the square sinusoidal noise, while the argument of each ensures high frequency oscillations across the whole range of input.

Considering the first column of Table 4.1, we see that as long as our approximation is within 0.1% of the actual PBE functional, and that this approximation is zero on average, most materials should fall under the 50 meV threshold for the error in the total energy. As Appendix D shows, a relative error of 0.05% is required to guarantee that all materials in the database fall under this threshold, thereby providing our first indication of the tolerance of an NNLDA or NNPBE approach. The second column of Table 4.1 shows the effect a biased error, which does not average to zero, has on the total energy. The result is that even a relative error of 0.05% is insufficient to bring the majority of molecular compounds in the database within our stated tolerance of 50 meV.

The reason for this relates to how total energies in DFT programs such as GPAW represent their results, which is always relative to the reference energy of the system. This quantity equals the sum of the total energies of isolated atoms in space. So, for example, the output total energy of an H_2O calculation has implicitly subtracted reference energies associated with two hydrogen atoms and one oxygen atom. For non-empirical codes such as GPAW, the reference energies of all atoms have already been evaluated. However, any modifications to an xc functional require a re-evaluation of these reference energies that Table 4.1 shows are crucial to obtaining the desired tolerance. Computationally, this reference energy calculation can be done by evaluating the component elements in separate very large cells each with a very high grid point density. Since

the output of this result will also depend on the pre-calculated xc functional reference energy, latter calculations of the standard error between two methods will be computed via

$$\left|\left(E^{LDA/PBE} - E_{ref}^{LDA/PBE}\right) - \left(E^{NNLDA/NNPBE} - E_{ref}^{NNLDA/NNPBE}\right)\right|. \tag{4.1}$$

4.2 Training the NNLDA

We begin with an evaluation of our neural network approach in the case of the LDA xc functional, with the xc energy density ε_{xc} defined in Eq. 2.17 and the xc potential $\hat{V}_{xc}$ defined in Eq. 2.17. Twenty thousand uniformly distributed random values for the density n are chosen between 10^{-20} and 10^{10} from which training data using $\varepsilon_{xc}^{LDA}[n]$ and $\hat{V}_{xc}^{LDA}[n]$ are determined. While this range for n is uncharacteristically large and in many ways nonphysical, it allows us to assess from a theoretical standpoint whether the NNLDA is valid in the limit of infinitely-weak and infinitely-strong correlation, for which exact formulae for the correlation part of the xc functional exist [81]. For the purposes of testing the NNLDA on physical systems, we will see that a substantially reduced range could be utilized to achieve similar accuracy.

Feed-forward networks with three hidden layers each consisting of 50 neurons, which we refer to as a 50×3 network, are trained with an input of $\log_{10} n$ and an output of either $\log_{10}(-\varepsilon_{xc} n^{-1/3})$ or $\log_{10}(-\hat{V}_{xc})$. These choices for the output reflect the fact that networks perform best when the relationship between the input and output is quasi-linear [115]. As we see in Eq. 2.19, the exchange part of the LDA energy density is proportional to the cube root of the electron density. The rectified linear units defined in Eq. 2.114 are chosen for the activation functions, the error is mean squared as defined in Eq. 2.117, and weights are updated according to the Adam optimizer in Eq. 2.128, first for 500 epochs at a learning rate of $\gamma = 10^{-4}$ and then for 10000 epochs at a learning rate of $\gamma = 10^{-5}$. Five percent of the data is set aside as a validation set to ensure that overfitting to the training data is not a problem.

Fig. 4.1 shows the relative error between ε_{xc}^{LDA} and ε_{xc}^{NNLDA}, $\hat{V}_{xc}^{LDA}$ and $\hat{V}_{xc}^{NNLDA}$, and $\partial\varepsilon_{xc}^{LDA}/\partial n$ and $\partial\varepsilon_{xc}^{NNLDA}/\partial n$ for three separate implementations of NNLDA all trained according to the aforementioned procedure. Fig. 4.1(a,b) both reveal certain features that we expect all neural network implementations of xc functionals to follow. First, downward spikes mark densities for which NNLDA predicts the energy density or potential to a very high degree of accuracy, which contrast with the more parabolic shapes that edge along the 10^{-3} threshold. This is purely a property of plotting on a logarithmic scale, with the deepest peaks also being the least broad. The oscillations in the relative error also become more dense in the $10^{-2} \leq n \leq 10^{2}$ range, which is exactly where the LDA xc functional fluctuates the most, although it is more prominent in the energy density than in the potential. Fig. 4.1(a,b) are plotted with three separate implementations but all exhibit these same features in addition to consistently having a relative error below the 10^{-3} or 0.1% threshold for all input densities n.

Fig. 4.1(c) shows how the partial derivative of the energy density with respect to the electron density compares using LDA and NNLDA for each of the networks trained above. Taking partial

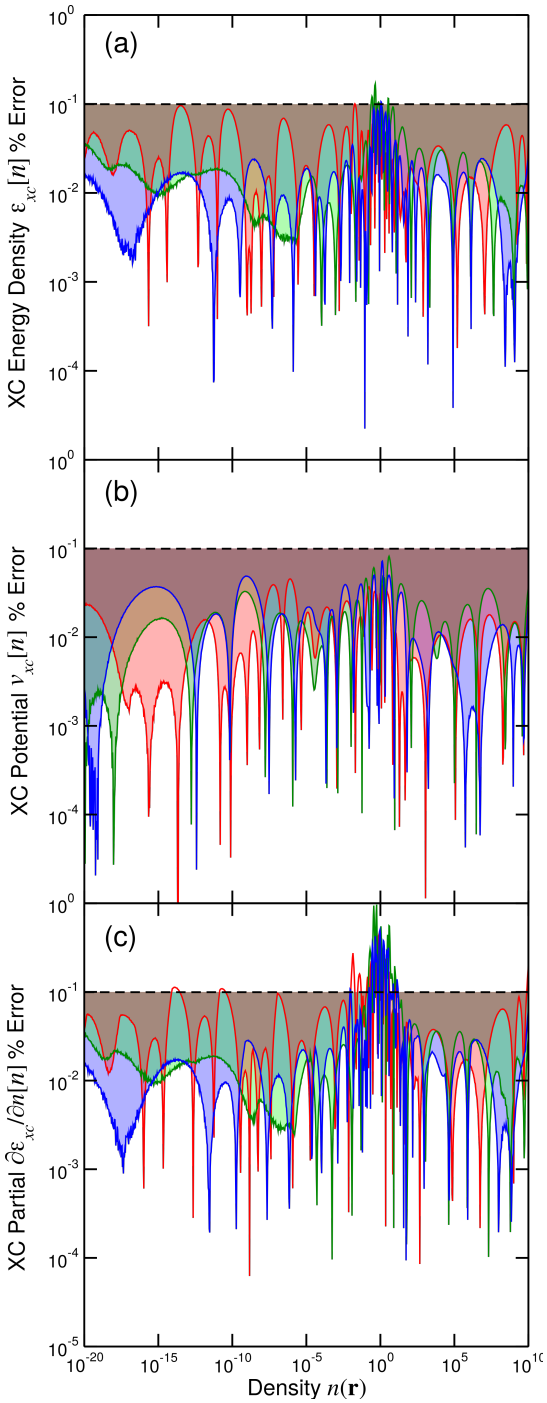

Figure 4.1: Relative error between (a) ε_{xc}^{LDA} and ε_{xc}^{NNLDA}, (b) $\hat{V}_{xc}^{LDA}$ and $\hat{V}_{xc}^{NNLDA}$, and (c) $\partial\varepsilon_{xc}^{LDA}/\partial n$ and $\partial\varepsilon_{xc}^{NNLDA}/\partial n$ for three separate implementations of NNLDA each using 50×3 neural networks trained on the same range of data.

derivatives within a feed-forward network that does not use input pre-processing techniques such as convolution or max pooling (see Section 2.4.2) is a relatively straightforward task since such networks can be expressed as continuous analytic functions. Following Eq. 2.18, being able to predict this quantity accurately means that a network would only be required to train the xc energy density. It would also mean that the learned xc functional is properly self-consistent in the KS framework, which is not the case when the xc energy density and xc potential are trained separately. Interestingly, the peaks and troughs in Fig. 4.1(c) follow an almost identical pattern to those of Fig. 4.1(a), suggesting that the gradient evaluated within the network is well-defined as we predicted. In the $10^{-2} \leq n \leq 10^{2}$ range, where the xc energy density has the highest density of oscillations, the partial derivative crosses over the 10^{-3} or 0.1% threshold, suggesting that the network is either less or more responsive to changes in the electron density in this range. If the xc potential is determined in this fashion and we consider a relative error of 0.1% is necessary, as in Table 4.1, for the xc potential, Fig. 4.1(c) suggests that the network would need to be trained to an even higher accuracy. Although not shown in Fig. 4.1, no noticeable difference was seen when the activation function was changed to either softmax or sigmoid as defined in Eq. 2.114.

The xc potential generated from the functional derivative in the $10^{-2} \leq n \leq 10^{2}$ is both far more oscillatory and performs worse than directly using a neural network to train the xc potential. This suggests that an alternate approach that starts from an altered xc potential, as in the reformulated KS framework outlined in Eq. 2.26, and from which the total xc energy is derived, may be a more appropriate route towards obtaining a self-consistent xc functional [34, 85].

4.3 Training the NNPBE

Moving on to the evaluation of our neural network approach in the case of the PBE xc functional, the norm of the gradient of the electron density $|\nabla n|$ will now be an additional factor in the input to our neural network trained xc functional, the NNPBE. Fifty thousand uniformly distributed random pairs of the density and the gradient norm $[n, |\nabla n|]$ are chosen such that each is allowed to range freely between 10^{-20} and 10^{10}, following the range previously used for the NNLDA. Training data is then determined by evaluating $\varepsilon_{xc}^{PBE}[n, |\nabla n|]$ and $\hat{V}_{xc}^{PBE}[n, |\nabla n|]$. A 50×3 network is used with input of $\log_{10} n$ and $\log_{10} |\nabla n|$ and an output of either $\log_{10}(-\varepsilon_{xc} n^{-1/3})$ or $\log_{10}(-\hat{V}_{xc})$, the renormalization of the output being done in the same manner as with LDA. Rectified linear units are chosen for the activation functions, the error is mean squared, and the Adam optimizer is used for 30000 epochs with a learning rate that decreases exponentially from 5×10^{-4} to 5×10^{-7}. As before, 5% of the data is set aside as a validation set.

Fig. 4.2 shows contour maps of the relative error between ε_{xc}^{PBE} and ε_{xc}^{NNPBE}, $\hat{V}_{xc}^{PBE}$ and $\hat{V}_{xc}^{NNPBE}$, and $\partial\varepsilon_{xc}^{PBE}/\partial n$ and $\partial\varepsilon_{xc}^{NNPBE}/\partial n$. The sharp downward spikes seen in Fig. 4.1 have their counterpart in the dark-coloured rings appearing in the contour maps. Fig. 4.2(a,b) show a more rippled relative error in the $10^{-2} \leq n \leq 10^{2}$ range as with the NNLDA case but for all values of the gradient, and there is an additional region with slope $\log_{10} |\nabla n| = 4/3 \log_{10} n$ where the oscillations are more pronounced. As shown in Eq. 2.28, this is because the dimensionless quantity

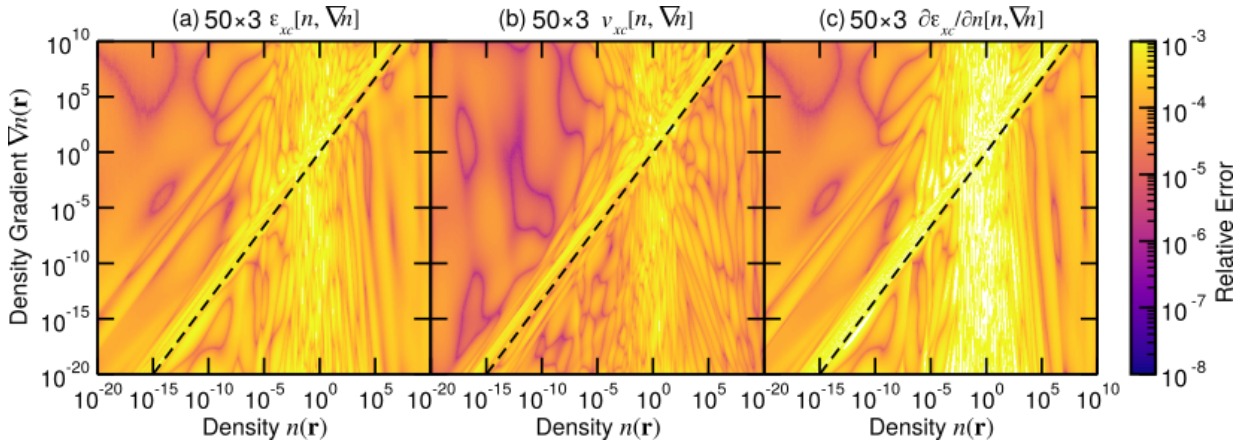

Figure 4.2: Contour maps of the relative error between (a) ε_{xc}^{PBE} and ε_{xc}^{NNPBE}, (b) $\hat{V}_{xc}^{PBE}$ and $\hat{V}_{xc}^{NNPBE}$, and (c) $\partial\varepsilon_{xc}^{PBE}/\partial n$ and $\partial\varepsilon_{xc}^{NNPBE}/\partial n$. A dashed line showing where $\varrho = n^{-4/3}|\nabla n| = 1$ is also plotted on all three contour maps.

$\varrho = n^{-4/3}|\nabla n|$ is a major component of many GGA xc functionals, and there is a range for ϱ that leads to the highest fluctuations with respect to the input parameters. We plot $\varrho = n^{-4/3}|\nabla n| = 1$ for all three contour maps of Fig. 4.2 so as to highlight this region. The higher relative error in the $10^{-2} \leq n \leq 10^{2}$ range can be understood in light of PBE being an improvement to the LDA functional as seen in Eq. 2.20, so that areas of difficulty for the lower-tier functional will manifest in the neural network approximation to its higher-tier counterpart.

Fig. 4.2(c) reveals similar problems to those seen in Fig. 4.1. Specifically, errors in the partial derivative of the xc energy density match closely with those of the xc energy density from which it was generated, except that the error again goes beyond the 10^{-3} or 0.1% threshold in all regions where the xc energy density showed the most fluctuation in the relative error. These fluctuations are not present in the xc potential, which is trained directly. This again suggests that a self-consistent formalism that uses neural network-based xc functionals may be best served by training the xc potential and then extracting the xc energy density from it. It is worth noting that in the formula for $\hat{V}_{xc}^{PBE}$ from Eq. 2.21 there is also a dependence on $\partial\varepsilon_{xc}^{PBE}/\partial|\nabla n|$. However, due to convergence issues for this quantity in GPAW over the large range of densities and density gradients chosen, it is not shown.

4.3.1 Rough Model and Difficulty Zones

One of the reasons that the LDA and PBE xc functionals fluctuate considerably for densities in a certain range or along a line given by the dimensionless parameter ϱ is because these are the regions where for physical systems most of the $(n, |\nabla n|)$ pairs are going to lie. This can be understood qualitatively for the case of ϱ. If we consider that molecular systems in a large unit cell can be roughly described by regions of uniform empty space punctuated by sharp increases

in the electron density in and around the KS orbitals, it is logical that regions of low electron density but a high density gradient or vice versa will not commonly occur in most systems. While for the LDA or PBE cases, it is simple enough visually to ascertain these regions of high fluctuation and limit the training range accordingly, we want to implement a method where the most relevant region of training data is assessed automatically by the neural network if this formalism is extended to networks that take in a larger range of input. Towards that goal we introduce our rough model and difficulty zones.

The goal of the rough model approach is to first train a small network that provides a rough approximation to the xc energy density and xc potential, which can then be divided out from the exact functional to yield a target for a correction model. The idea is that the small network will be able to extract any major features of the functional while the correction model will be better suited to extracting the higher-level features of the functional. In formulaic form this goes as

$$e_{rough} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow \log_{10}\left(-\varepsilon_{xc} n^{-1/3}\right) \tag{4.2}$$

$$e_{correction} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow -\varepsilon_{xc} n^{-1/3} 10^{-e_{rough}} - 1 \tag{4.3}$$

$$\varepsilon_{xc}^{NNPBE} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow -n^{1/3}(1 + e_{correction}) 10^{e_{rough}}. \tag{4.4}$$

and

$$v_{rough} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow \log_{10}\left(-\hat{V}_{xc}\right) \tag{4.5}$$

$$v_{correction} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow -\hat{V}_{xc} 10^{-v_{rough}} - 1 \tag{4.6}$$

$$\hat{V}_{xc}^{NNPBE} : [\log_{10} n, \log_{10} |\nabla n|] \rightarrow -(1 + v_{correction}) 10^{v_{rough}}. \tag{4.7}$$

Our second strategy for getting better performance in the regions that are currently hardest for the network to predict is to first run the rough and correction models, randomly generate new test data, and choose from a random set of test points those that are within a certain radius of the test data that lays above the threshold value in the original model, these points defining a difficulty zone for the network. These test points can then be mixed in with the uniformly distributed random data at different ratios to have the network focus more or less on the areas with the highest relative error, which we have posited are the regions that are most physically relevant. Finally, the correction neural network can be rerun with the new set of training data.

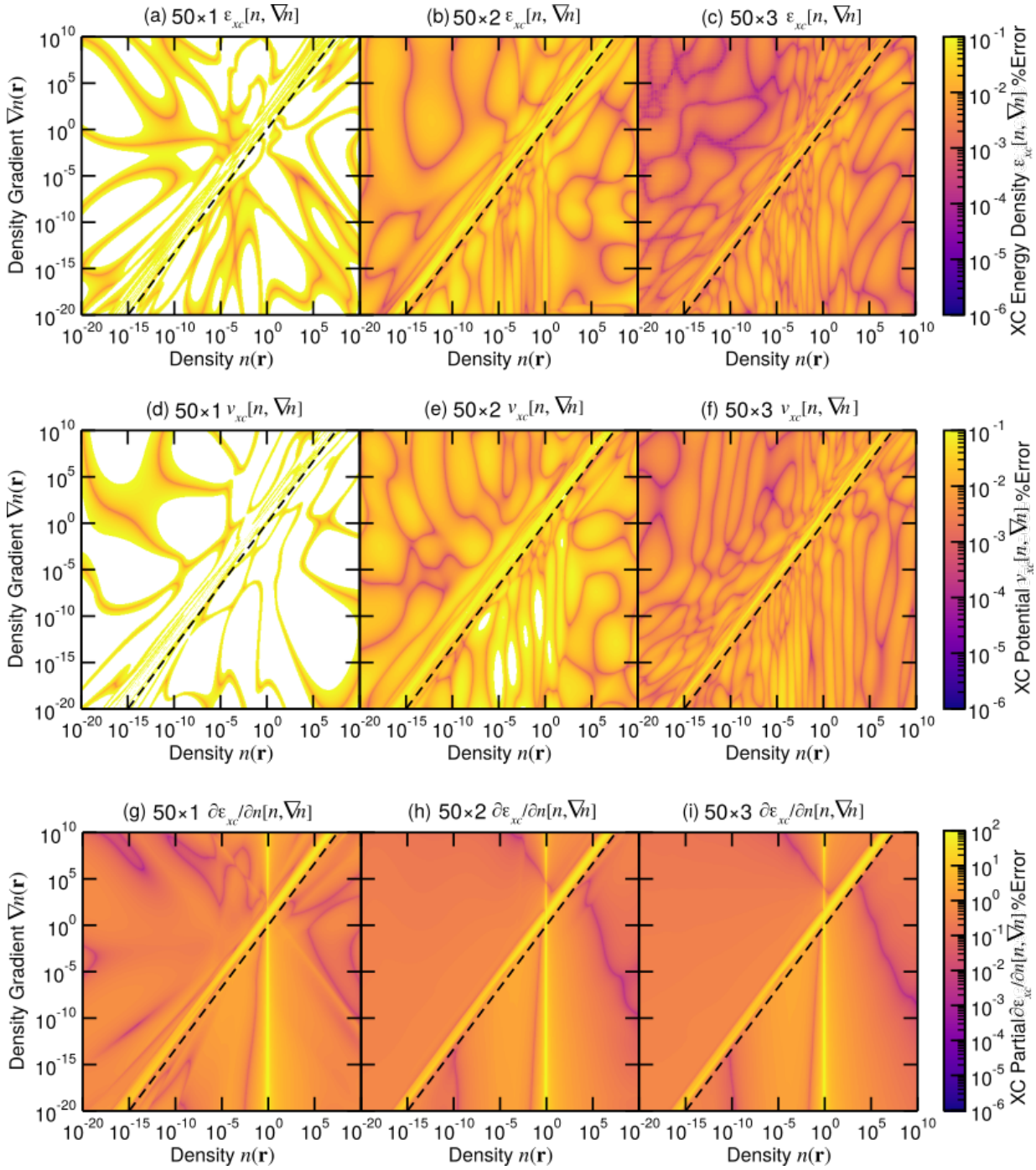

Figure 4.3: Contour maps of the relative error between (a,b,c) ε_{xc}^{PBE} and ε_{xc}^{NNPBE}, (d,e,f) $\hat{V}_{xc}^{PBE}$ and $\hat{V}_{xc}^{NNPBE}$, and (g,h,i) $\partial\varepsilon_{xc}^{PBE}/\partial n$ and $\partial\varepsilon_{xc}^{NNPBE}/\partial n$ for neural networks with (a,d,g) one, (b,e,h) two, and (c,f,i) three hidden layers of 50 neurons each, using the rough model and difficulty zone approach. The colour scale for contour maps (g,h,i) has been adjusted for better visualization. A dashed line showing where $\varrho = n^{-4/3}|\nabla n| = 1$ is also plotted on all nine contour maps.

For the networks in Fig. 4.3 and Fig. 4.4, the rough model is trained with 1000 points in $(n, |\nabla n|)$ space in a 10×1 network for 100 epochs and a 0.01 learning rate. The first correction model is trained with 50000 uniformly distributed random points and 30000 epochs as in the approach for Fig. 4.2. Thirty-five thousand random points are then generated under the condition that they are within a radius of 1 in $(n, |\nabla n|)$ space from test points that yielded a errors much higher than the average relative error, and this data set is mixed with 15000 uniformly distributed random points. The second trained correction model follows the choice of other hyperparameters given in Fig. 4.2 and uses the mixed data to train a network to predict the functional ε_{xc}^{NNPBE}.

Fig. 4.3 and Fig. 4.4 both show contour maps of the relative error between ε_{xc}^{PBE} and ε_{xc}^{NNPBE}, $\hat{V}_{xc}^{PBE}$ and $\hat{V}_{xc}^{NNPBE}$, and $\partial\varepsilon_{xc}^{PBE}/\partial n$ and $\partial\varepsilon_{xc}^{NNPBE}/\partial n$, as in Fig. 4.2. Fig. 4.3 shows the effect that increasing the depth of the network from one hidden layer to three with each layer having 50 neurons each has on this relative error, meaning these networks have 50, 100, and 150 hidden neurons, respectively. It is clear that increasing the network's depth yields a commensurate decrease in the relative error for all points in $(n, |\nabla n|)$ space, while the distribution of areas of high and low relative error also becomes more scattered with more dark-coloured rings corresponding to contours with extremely low relative error. The overall decrease in the error is equivalent for the xc energy density and the xc potential. Although not shown, a 50×4 network that also uses the rough model and difficulty zone performs equally well as the 50×3 network shown in Fig. 4.3(c,f,i). While having overly deep networks can cause issues when it comes to updating weights nearest the input layer, using the Adam optimizer over a sufficiently high number of epochs with the right training rate mitigates this issue [143, 144]. This explains why increasing the network's depth beyond three layers, even with adjusted hyperparameters, does not yield a further decrease in the relative error.

With the implementation of the rough model and the difficulty zone, any vestige of the fluctuating behaviour around the $10^{-2} \leq n \leq 10^{2}$ range has disappeared while there remains a thinner but still prominent increased relative error along the line where our dimensionless electron density gradient ϱ is constant. The increased depth appears to lead the relative error along this slope to become more localized, with the 50×3 network in Fig. 4.3(c,f) using the rough model and difficulty zone to also give more localization compared to the implementation shown in Fig. 4.2(a,b).

Fig. 4.3(g,h,i) shows the derivative of the xc energy density with respect to the electron density has a very different profile from that seen in Fig. 4.2(c). This is an unfortunate side effect of the rough model implementation shown in Eq. 4.4, where the partial derivative of ε_{xc}^{NNPBE} computed within this approach will most strongly depend on the partial derivative of e_{rough} in Eq. 4.2. The contour maps in Fig. 4.3(g,h,i) indirectly reveal that this rough model has captured the main features of the PBE xc functional, that is, the areas that our original network showed the most fluctuation. The change in this partial derivative contour map with increasing depth, namely, a decrease in the number of dark-coloured rings, is attributed to a coincidental error cancellation for the rough model and the low-depth neural network which becomes less prominent as more hidden layers are added. Dark features that remain at higher depths are regions where the rough model managed to predict the PBE xc energy density to a high degree of accuracy.

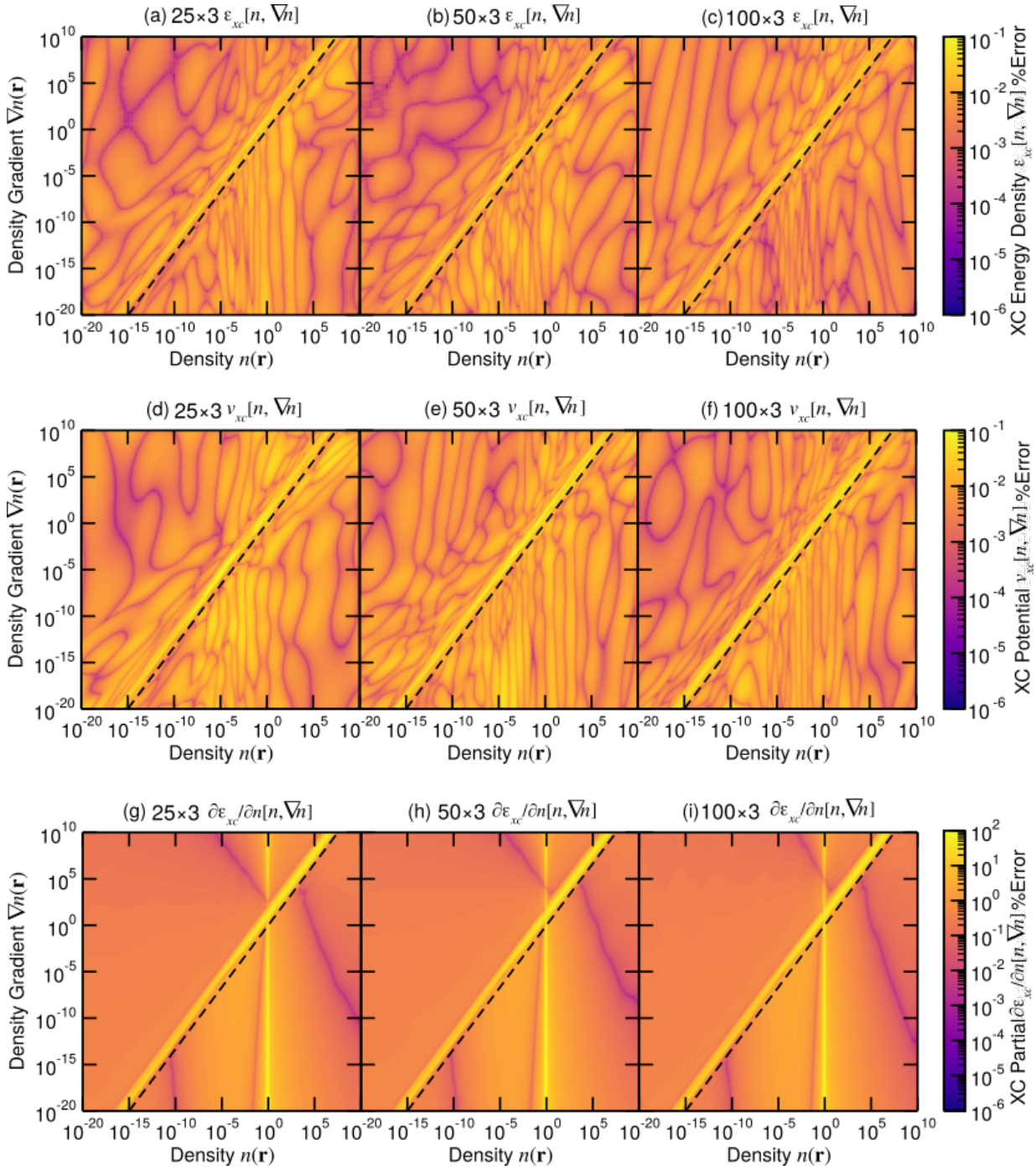

Figure 4.4: Contour maps of the relative error between (a,b,c) ε_{xc}^{PBE} and ε_{xc}^{NNPBE}, (d,e,f) $\hat{V}_{xc}^{PBE}$ and $\hat{V}_{xc}^{NNPBE}$, and (g,h,i) $\partial\varepsilon_{xc}^{PBE}/\partial n$ and $\partial\varepsilon_{xc}^{NNPBE}/\partial n$ for neural networks with three hidden layers of (a,d,g) 25, (b,e,h) 50, and (c,f,i) 100 neurons each, using the rough model and difficulty zone approach. The colour scale for contour maps (g,h,i) has been adjusted for better visualization. A dashed line showing where $\varrho = n^{-4/3}|\nabla n| = 1$ is also plotted on all nine contour maps.

Fig. 4.4 shows the effect that increasing the number of neurons in each hidden layer of the network, from 25 neurons to 100 neurons in a three hidden layer network, corresponding to networks with 75, 150, and 300 neurons each, has on the relative error of the xc energy density, the xc potential, and the partial derivative of the xc energy density with respect to the electron density. It is clear that other than an increased scattering between regions of low and relatively high relative errors that results from having more neurons and hence more weights and biases shaping the profile of the network, there is no noticeable increase in the accuracy with increasing numbers of neurons. Although not shown, a test of a 15×3 network begins to reveal the necessity of having some width to our networks with a colour profile similar to the 50×2 network shown in Fig. 4.3(b,e). It is worth noting the increased performance that the 25×3 network with 75 neurons total has over the 50×2 network with 100 neurons total, revealing the ability of deep-layered networks with the right choice of activation function to perform significantly better than their shallow counterparts, as mentioned in Section 2.4. Having low-width networks is an important consideration for the integration of neural network implementations of the xc functional into DFT software packages such as GPAW, where fewer connections between neurons means faster computations.

4.4 Testing the NNLDA and NNPBE

As mentioned in Section 4.1, only testing the relative error of our NNLDA and NNPBE approximations with respect to the analytic forms of these xc functionals is insufficient to determine how well these approximations fare when modelling molecular and crystal systems. While the G2-1 molecular database [16] was used to estimate the effects that noise added to the PBE functional would have on the final ground state energy in Section 4.1, we opt for the DeltaCodesDFT database [40, 41] of material systems here for two reasons. First, the types of materials under consideration in this database span both the molecular and the crystalline, ideal for testing standard xc functionals like LDA and PBE. Second, the systems all consist solely of one element, allowing us to see the effect that changes in the element properties may have on the accuracy of the neural network approximation. In this way, Eq. 4.1 may be used to yield the desired error in the ground state energies resulting from our neural network approach, accounting for the reference energies.

Before delving into the results calculated within GPAW, we mentioned in Section 4.2 that the training range used for the density and later the norm of the gradient of the density stretched into a range that is unreachable for any physical system, but may prove useful for ensuring that the approximation converges to certain well-defined theoretical values. The discussion around Eqs. 4.2–4.4 also emphasized that the region where our xc functionals fluctuate the most is also where most physical systems will reside.

Fig. 4.5(a) shows a contour map of the relative error between ε_{xc}^{PBE} and ε_{xc}^{NNPBE} if the correction model for the second iteration of the network is trained entirely with points within the difficulty zone. Somewhat contrary to expectations, this trained network reflects a difficulty zone located

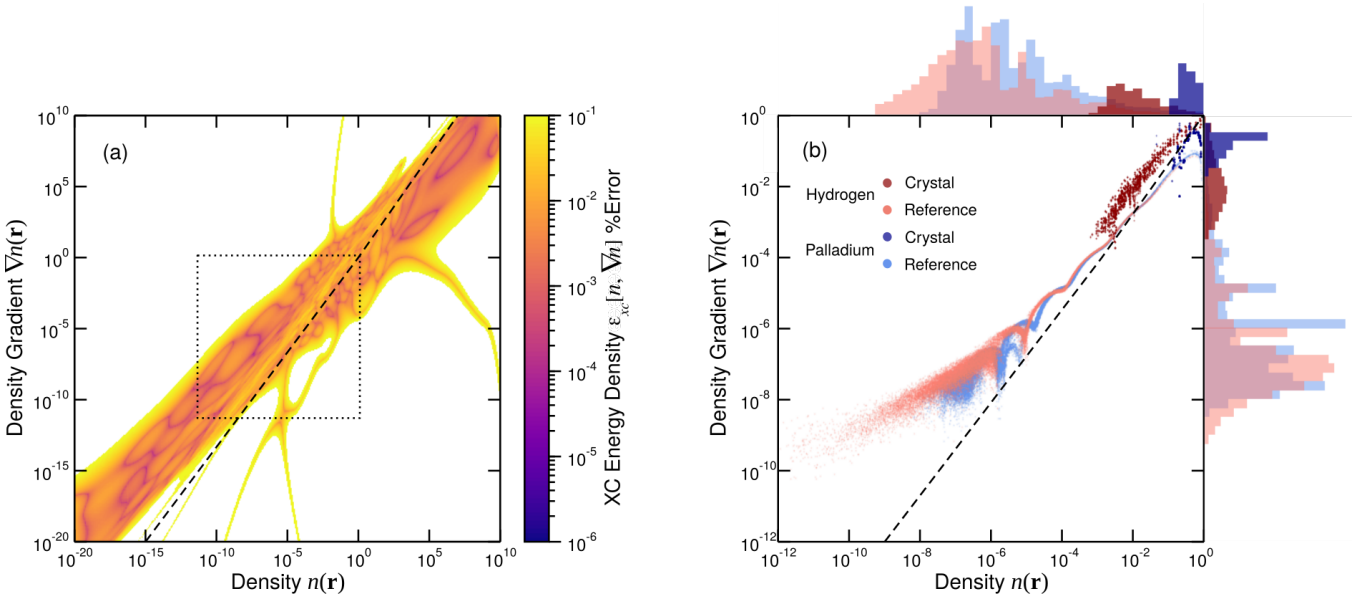

Figure 4.5: (a) Contour map of the relative error between ε_{xc}^{PBE} and ε_{xc}^{NNPBE} for a 50×3 network using the rough model and difficulty zone approach, with the correction model trained entirely on points in the difficulty zone with a dashed line showing $\varrho = n^{-4/3}|\nabla n| = 1$ (dashed line) and the region $(n, |\nabla n|) \in [10^{-12} \dots 10^{0}, 10^{-12} \dots 10^{0}]$ (dotted lines) and (b) a scatter plot of points in $(n, |\nabla n|) \in [10^{-12} \dots 10^{0}, 10^{-12} \dots 10^{0}]$ from converged calculations of the reference system of a hydrogen atom (light red), the reference system of a palladium atom (light blue), and the Delta Codes database crystals for hydrogen (dark red) and palladium (dark blue), with histograms of bin width 0.25 on the log-log scale showing the distribution of electron densities n and the norm of the gradient of electron densities $|\nabla n|$.

entirely on a strip around the diagonal of the contour map, although close inspection shows that the constant ϱ region has not fully disappeared. The white region surrounding this strip is a direct consequence of overfitting the data to give a reduced relative error in the region of interest. While having the network perform so poorly in this region may seem concerning, Fig. 4.5(b) shows for two systems, hydrogen and palladium structures from the Delta Codes database along with their reference systems, all pairs of points in $(n, |\nabla n|)$ space given by these converged calculations fall into the trained region. These two elements were chosen as the overlap between their sets of points was among the lowest, allowing for a better visual representation, although no element in the system diverged considerably from the pattern presently shown for these test cases. The reason for the reference systems having values of $(n, |\nabla n|)$ much lower than for the database structures is because of the amount of empty space that needs to be evaluated, whereas the higher density of these points reflects the fine grid used in the calculations. The range of densities and gradients in each run of the self-consistent cycle (see Section 2.1.6) is not expected to vary considerably, as methods such as density mixing [44] help ensure steady convergence in the iterative scheme.

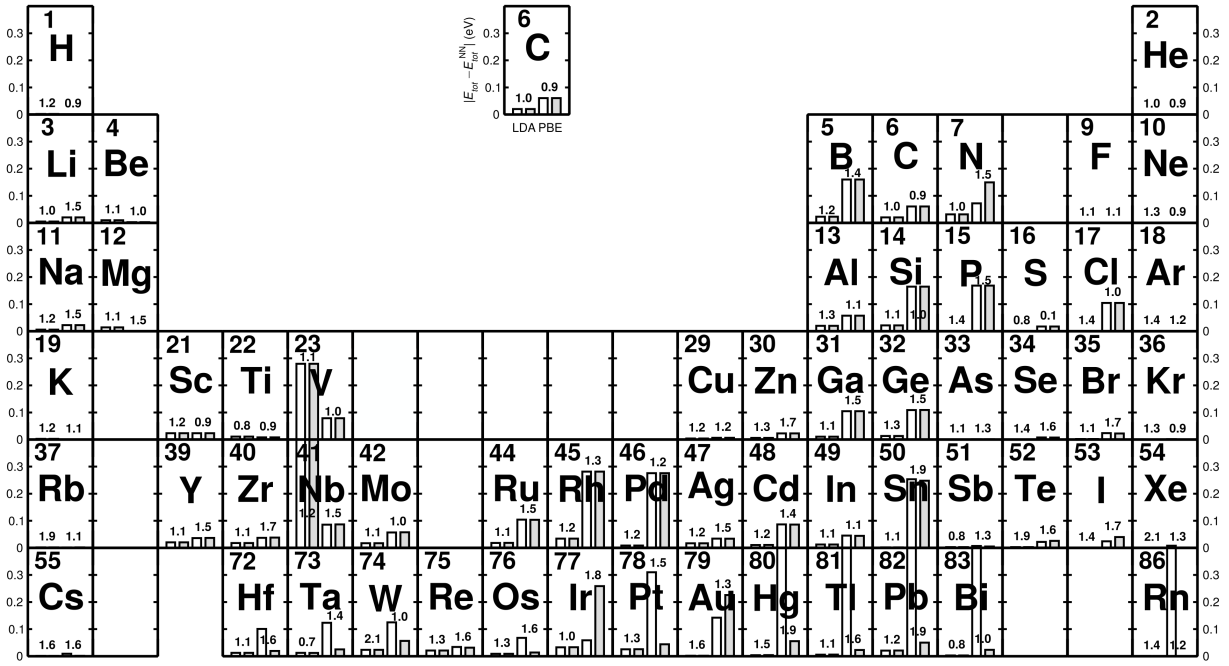

Figure 4.6: Periodic table display of total energy errors calculated using Eq. 4.1 for 59 structures from the Delta Codes DFT database, where each box contains the element symbol and atomic number while the columns in order from left to right show the Harris value for NNLDA, the self-consistent calculation for NNLDA, the Harris value for NNPBE, and the self-consistent calculation for NNPBE, all in units of eV. The numbers above the NNLDA and NNPBE columns are the calculation time ratios between the neural network implementations and the default LDA and PBE methods, respectively.

For each system in the Delta Codes DFT database, which includes the crystal structure and unit cell parameters customized for each system, a plane wave basis set with $E_{cut} = 200$ eV was used with an autogenerated k-point density of 1.0, electronic temperature of $k_B T \approx 1$ meV with all energies extrapolated to $T \to 0$, and a grid spacing of $h \approx 0.2$ Å. For all reference systems, a plane wave basis set with $E_{cut} = 200$ eV was used for a single atom placed centrally in a $10 \times 10 \times 10$ Å^3 unit cell, with a Γ-point sampling of the Brillouin zone, a grid spacing of $h \approx 0.05$ Å, and the same electronic temperature as in the database calculations. Calculations were done using the LDA and PBE xc functionals, as well as the NNLDA presented in Fig. 4.1 and the NNPBE presented in Fig. 4.5, using xc energy densities and xc potentials both trained by neural networks separately. Both self-consistent calculations, where the full self-consistent cycle is executed starting from the initial setup, and Harris (single-point) calculations, where the neural network calculations are executed with a density fixed by the converged result of the LDA or PBE calculations, are performed. Complete details of the results are provided in Appendix D.

Fig. 4.6 displays the result of the above mentioned calculations for 59 structures from the

Delta Codes DFT database, showing the total energy error given in Eq. 4.1 for self-consistent and Harris calculations of the NNLDA and NNPBE. Some elements are absent from the table because the system had non-zero magnetic moments (e.g. oxygen and iron), because the convergence could not be achieved using either LDA or PBE (e.g. calcium) in a reasonable amount of time, or because GPAW does not have the setups for that particular element (e.g. technetium). Employing the periodic table highlights how both the filling of orbitals and the atomic number play a role in the convergence of our neural network implementations. We see that for all systems with the exception of vanadium and niobium, which both have three d electrons, the NNLDA performs exceptionally well, with many systems giving results for the Harris and self-consistent energies with an accuracy within 10^{-5} eV. This also suggests that subtraction of the reference energy is important to the noisy LDA shown in Table 4.1 even when the noise averages to zero. Although the reason behind these high errors for V and Nb is unclear, both the Harris and neural network approaches give similar errors, suggesting the issue may lie with a poorly described reference energy for this particular electron filling.

For the NNPBE implementation, elements with fully or partially empty d-orbitals, i.e., the first four columns of the table, or nearly full p-orbitals, i.e., the last three, we see that both the self-consistent and Harris calculations give total energies below the 50 meV threshold for the majority of systems with few exceptions. This suggests that elements without partially filled p or d-orbitals are described consistently well by our neural network xc functionals, reflecting the difficulty of describing these orbitals. NNPBE encounters considerable difficulty for elements with few p electrons, i.e., in columns 5, 6, or 7, while its performance for transition metals beyond the aforementioned columns consistently yields energy errors in the 100 to 300 meV range, far beyond the threshold NNLDA has almost no difficulty meeting. Results within these transition metals are generally better for the $5d$ elements of row 6 than the $4d$ elements of row 5. This suggests elements with lower gradients of the electron density, that is, valence electrons spread over the more delocalized orbitals of heavier atoms, are easier to describe.

With the exception of nitrogen, the Harris and self-consistent energies of the NNPBE give nearly identical results for all elements without f electrons, i.e., in the first five rows of the table. The difference between the two becomes gradually larger as we fill the $5d$ orbitals, with the single-point calculations for all five elements with filled d bands ($Z \geq 80$) having total energy errors over 0.4 eV, even when the self-consistent calculation gives energy errors near the threshold. One explanation for this may lie in the use of the pseudo-density in the PAW formalism, computed using Eq. 2.15 in tandem with the PAW transformation operator in Eq. 2.58, rather than the all-electron density. However, this does not explain the observed agreement with the NNLDA for the very same elements.

Lastly, we note that the time ratios in the periodic table figure reveal that across all elements, our neural network implementation ranges anywhere from being ten times faster in the case of NNPBE applied to sulfur to twice as slow in the case of NNLDA applied to tungsten, with the networks generally taking slightly longer than their analytic counterparts. The time ratios that fall below one are serendipitous in that the neural-network trained xc-functional managed to converge in far fewer iterations, even if each individual iteration took slightly longer. Despite

this, the similar time scales for each implementation suggests that more complex functionals can be made wider or deeper without incurring substantial computational costs. The highest time cost is in fact incurred by the reference energy calculations, where the high grid density and large unit cell made the NNLDA and NNPBE calculations on average 5 times slower to converge than LDA and PBE, respectively. Luckily, for a given network, these reference energy calculations only need to be performed once for each element. This means the amount of time this convergence takes is immaterial so long as when the network is applied to most systems it is reasonably fast and accurate.

Chapter 5

LCAO-TDDFT-k-ω : Spectroscopy in the Optical Limit

Ab initio time dependent density functional theory (TDDFT) methods constitute the standard computational approach for modelling interactions between light and matter, including optical absorption processes, conduction, and exciton generation and recombination. Starting from a converged ground-state calculation, many different methods are available to compute the quantity of interest, from the quantitatively accurate but computationally inefficient G_0W_0-BSE approach of Section 2.2.7 to less computationally demanding methods starting from the non-interacting density response function of Eq. 2.42. By employing a highly efficient linear combination of atomic orbitals (LCAO) representation of the Kohn-Sham (KS) orbitals (Section 2.1.5) within TDDFT in the reciprocal space (k) and frequency (ω) domains, as implemented within our LCAO-TDDFT-k-ω code, and applying the derivative discontinuity correction of the exchange functional Δ_x to the KS eigenenergies (Section 2.1.3), we are able to provide a semi-quantitative description of the optical absorption, conductivity, and polarizability spectra for 0D, 1D, 2D, and 3D systems within the optical limit ($|\mathbf{q}| \to 0^+$). As we will show, this approach is best suited for modelling the optical absorbance of low dimensional macro-molecular systems, for which robust and efficient methods capable of properly describing all the relevant effects have remained elusive.

Regarding the choice of real or reciprocal space and time, different choices provide advantages and disadvantages depending on the system under consideration. For example, TDDFT in real space (r) and frequency (ω) domains, known as the Casida method or TDDFT-r-ω [47], provides a balance between accuracy and efficiency. However, its $O(N^5)$ scaling, where N is the number of KS wavefunctions, makes it either unsuitable or unusable for calculating the properties of macromolecules and other low-dimensional systems, no matter the wavefunction basis choice. Given a real-space representation, TDDFT in real space (r) and time (t) domains (RS-TDDFT-r-t) [49] requires extremely short time steps to give wavefunction stability, much shorter than the period required for resolved frequency spectra [50], whereas TDDFT in reciprocal space k and frequency ω (RS-TDDFT-k-ω) [48] often has quite high memory costs associated with the grid

spacing required for convergence. PW is an alternative to the RS representation, and allows for a systematic reduction in the computational cost that plagues these other RS methods while improving stability. However, for systems with very large unit cells, the high memory cost can once again make this an unfeasible choice. Using non-periodic or mixed boundary conditions, often seen in low-dimensional systems, also means that a good deal of memory is allocated to describing empty space. As a result, we employ the LCAO representation for the KS wavefunctions, performing TDDFT calculations in reciprocal space (k) and frequency (ω) (LCAO-TDDFT-k-ω), offering an efficient and stable representation with similar accuracy to PW-TDDFT-k-ω.

One caveat of this representation, as previously mentioned, is that LCAO does not allow for systematic convergence the way decreasing the grid spacing h in RS or increasing the plane-wave cutoff energy E_{cut} in PW does. The choice of LCAO basis set must also be sufficient to describe occupied and unoccupied states equally well [52], and beyond including more radial functions or polarization, p-valence and completeness-optimized basis sets have been suggested to overcome both this and the absolute convergence issue [17, 52]. The other caveat, since an efficient Fast Fourier Transform (FFT) is not available in LCAO to efficiently compute the transition dipole elements in the dielectric function in Eq. 2.44, is that the LCAO-TDDFT-k-ω calculations are restricted to the head of the dielectric function ($\mathbf{G} = \mathbf{G}' = 0$) and to the optical limit $|\mathbf{q}| \to 0^+$.

In fact, for applications relating to light absorption or for systems where spatial dispersion may not be a well-defined quantity, the dielectric function in the optical limit is the desired quantity. A TDDFT implementation which evaluates in this limit exactly avoids problems related to the use of dense k-point meshes to approach the $|\mathbf{q}| \to 0^+$ limit. In situations where models require response functions at low-wavenumbers, as in low-energy plasmon-phonon systems [145, 146], it may also be preferable to include spatial dispersion *post hoc* rather than to use these aforementioned k-point meshes with spacing on the order of the wavenumber crossing. As mentioned in Section 2.2.4, working directly in the optical limit also allows for the calculation of the electron and hole spectral densities at a given excitation, providing important information about the real-space distribution of these states for the material in question.

Lastly, in order to obtain a better description of the band gap in semiconducting or insulating materials, which is crucial for obtaining quantitatively accurate absorption and energy loss spectra, our LCAO-TDDFT-k-ω code employs the GLLB-SC derivative discontinuity correction to the eigenenergies of conduction band KS states. This approach, as described in Section 2.2.3, is applied in either an *a priori* or *a posteriori* fashion.

As GPAW includes plane wave and real space implementations of TDDFT-k-ω, we can directly compare the performance of the LCAO-TDDFT-k-ω code with well-established TDDFT implementations. For the remainder of this chapter, we compare the photoabsorption cross section, conductivity, and dielectric function obtained from LCAO-TDDFT-k-ω (using either the *a priori* or *a posteriori* scissors correction), PW-TDDFT-k-ω, G_0W_0-BSE calculations, and experimental measurements where applicable, thereby providing a benchmark for the energies and intensities of the resulting spectra.

5.1 Computational Details

5.1.1 Calculation Parameters

All DFT calculations presented herein employ the PAW method code GPAW [36, 37] within the atomic simulation environment ASE [147, 148]. The PBEsol functional, a type of GGA xc functional tailored for solids and surfaces, was used unless stated otherwise, as it allows a self-consistent calculation of the derivative discontinuity correction Δ_x of Eq. 2.31. Each calculation used a grid spacing of $h \approx 0.2$ Å and an electronic temperature of $k_B T \approx 1$ meV with all energies extrapolated to $T \to 0$. All LCAO-TDDFT-k-ω computations employed a double-ζ-polarized (DZP) basis set [149], with basis sets going up to quadruple-ζ-polarized (QZP) employed to test convergence of the optical spectra, DZP being an standard default for its balance of computational cost and accuracy. All PW-TDDFT-k-ω computations are converged with an energy cutoff of $E_{cut} \approx 340$ eV.

For the C_{60} fullerene system and chlorophyll monomers, the atomic structures were relaxed until the maximum force was below 0.03 eV/Å. Gas phase structures for the *full*·$\mathbf{1}^+$ Chl *a* and Chl *b* monomers, shown in Fig. 3.4, follow models developed in the literature derived from mass spectrometry in tandem with molecular dynamics simulations [137, 138]. To see the influence of the hydrocarbon chain and the tetramethylammonium charge tag ($N(CH_3)_4^+$, or $\mathbf{1}^+$), structures which cut the hydrocarbon chain from $C_{20}H_{39}$ to C_5H_9, both with (*cut*·$\mathbf{1}^+$) and without (*cut*) the charge tag, are also considered. The electron density $n(\mathbf{r})$ and KS wavefunctions ψ_n were set to zero at the cell boundaries, which prevents long-range interactions between repeated images, crucial for modeling charged 0D materials. C_{60} used a $20 \times 20 \times 20$ Å^3 unit cell, while all chlorophyll superstructures were computed in supercells ensuring at least 6 Å of vacuum in every direction. For 0D materials, only Γ-point calculations (i.e. the BZ sampling only includes the Γ-point) were performed.

The (10,0) zigzag and (10,10) armchair SWCNTs' atomic structures were relaxed until the maximum force was below 0.05 eV/Å. Since carbon nanotubes extend along the z-axis, the unit cell was relaxed parallel to this axis until unit cell parameters of $L_z \approx 4.30$ and 2.46 Å, respectively, were obtained, including 10 Å of vacuum perpendicular to the nanotube axis. BZ samplings of $1 \times 1 \times 281$ for the (10,0) SWCNT and $1 \times 1 \times 489$ for the (10,10) SWCNT were employed, and periodic boundary conditions (PBCs) were used only along the z-axis. As with the macromolecules, the electron density and KS wavefunctions were set to zero at the cell boundaries perpendicular to the z-axis.

All Gr calculations use a unit-cell constant of $a = 4.651\ a_0 \approx 2.46$ Å, with Gr layers stacked in a periodic super-lattice along the z-axis perpendicular to each monolayer, separated by a distance of $L = 5a \approx 23.255\ a_0 \approx 12.3$ Å. Unless stated otherwise, the first BZ is sampled with a dense Monkhorst-Pack $301 \times 301 \times 1$ k-point mesh [150]. As this work presents a comparison between monolayer graphene systems computed for basis sets ranging from single zeta-polarized (SZP) to QZP, a total of 18 bands were employed for all LCAO calculations to accommodate the

number of basis functions that the SZP basis set can represent. This number of bands, corresponding to seven unoccupied bands per atom, was previously found [121] to yield converged optical spectra within RS-TDDFT-RPA.

All PN calculations use the already relaxed crystal structure found in the GPAW Computational 2D Materials Database [88]. As with GR, PN layers are stacked along the z-axis with supercells separated by a distance of $L \approx 17.1$ Å. Unless stated otherwise, the first BZ is sampled with a dense Monkhorst-Pack $603 \times 603 \times 1$ k-point mesh [150] with 32 bands, sufficient to converge the main peaks up to 20 eV.

For A-TiO_2, the first BZ is sampled with an $11 \times 11 \times 5$ k-point mesh and approximately 9 unoccupied bands per atom [100], whereas for R-TiO_2 the first BZ is sampled with a $7 \times 7 \times 11$ k-point mesh and about 9 unoccupied bands per atom[105]. It is important to note that adding more bands than is needed to converge the calculation is unnecessary unless one is seeking to study the very high energy part of the optical spectra.

As mentioned before, when electron–hole interactions, such as excitonic binding, strongly effect the shape and intensity of the optical spectra, G_0W_0-BSE is the gold standard for obtaining both a qualitative and quantitative picture the role these effects play. Systems where excitonic effects are most relevant include ones where electron-hole screening is reduced, such as in lower-dimensional systems [130]. As a result, G_0W_0-BSE spectra are shown for C_{60} [134], (10,0) SWCNT [9], PN, A-TiO_2 [100], and R-TiO_2 [105]. G_0W_0-BSE calculations for the (10,0) SWCNT and PN have been performed as part of this work, using the PW implementation within GPAW [101]. Reduced $1 \times 1 \times 64$ and $33 \times 47 \times 1$ k-point samplings of the first BZ were used respectively. We used the Godby-Needs plasmon-pole approximation [102, 103, 104] to describe the screening W, while either 1D or 2D truncation schemes for the Coulomb kernel [151] were employed to remove supercell interactions orthogonal to the (10,0) SWCNT or perpendicular to monolayer phosphorene.

5.1.2 Derivative Discontinuity Correction

The GLLB-SC derivative discontinuity correction, as described in Section 2.1.3, provides an *ab initio* first-order correction to the conduction band KS eigenenergies equal to Δ_x in Eq. 2.31. This correction can be applied as either an *a priori* or *a posteriori* scissors correction as shown in Section 2.2.3, which determines whether this correction influences the relative intensities of the peaks in addition to the blueshift of the spectra by Δ_x. Table 5.1 shows that for the materials considered herein, the correction lies between 0.5 and 0.9 eV, yielding a substantial qualitative change in the position of the optical spectra. This correction is included for all semiconducting materials studied in this chapter, with Δ_x calculated within LCAO-TDDFT-k-ω or PW-TDDFT-k-ω adjusting the spectra calculated from the respective TDDFT implementations.

Table 5.1: Derivative discontinuity correction Δ_x in eV obtained from LCAO or PW representations of the KS wave functions for fullerene (C_{60}), chlorophyll a and b (Chl *a,b*) monomers, (10,0) SWCNT, phosphorene (PN), anatase (A-TiO_2), and rutile (R-TiO_2).

Material	Δ_x (eV)	
	LCAO	PW
C_{60}	0.77	—
Chl *a,b*	0.70	—
(10,0) SWCNT	0.79	—
PN	0.58	0.55
A-TiO_2	0.85	0.89
R-TiO_2	0.67	0.72

5.1.3 Irreducible Brillouin Zone (IBZ)

In crystal systems, any symmetries in the lattice of the crystal manifest as symmetric relations between wavefunctions at different k-points. The irreducible BZ (IBZ) is defined as the most compact splice of the BZ such that every k-point in the full reducible BZ (RBZ) can be mapped through these symmetry relations to points in the IBZ. For high-symmetry crystals such as graphene, this results in substantial time and memory savings, allowing for a k-mesh spread only over points in the IBZ. Quantities calculated that depend only on the magnitude of the KS state wavenumber can simply be multiplied by the weight $w_{\mathbf{k}}$ of each k-point in the IBZ. However, as can be seen in Eq. 2.44 and Eq. 2.55, the sum in the optical response function $\varepsilon(\hat{\mathbf{q}}, \omega)$ is explicitly dependent on the vector components of the wavenumber $\mathbf{k}$ within the BZ. The result is that naïvely applying the LCAO-TDDFT-k-ω method using the IBZ for a crystal system that uses non-trivial symmetries may yield incorrect response functions. Note that 1D and 0D systems utilize trivial k-meshes and are therefore exempt from this consideration.

Table 5.2 shows one simple measure for the in-plane anisotropy of the dielectric function, integrating between 0 eV and 20 eV the real and imaginary parts of the absolute dielectric function difference between x- and y- components for graphene (GR), phosphorene (PN), anatase (A-TiO_2) and rutile (R-TiO_2), for computations done over both the IBZ and RBZ. With the exception of PN, all these materials should yield a symmetric dielectric tensor in the xy-plane, corresponding to planar isotropy. Floating point errors and the incompleteness of the LCAO basis set may partially explain the observed non-zero differences in RBZ calculations for what should be planar isotropic materials. The main import of Table 5.2 is that caution must be exercised when calculating optical quantities with explicit vectorial matrix element dependence. It can be proven geometrically that for rectangular and hexagonal unit cells for isotropic systems, the average of the two directional components, $\varepsilon_{\parallel} = \frac{1}{2}\left(\varepsilon_{xx} + \varepsilon_{yy}\right)$, will yield the appropriate xy-response func-

Table 5.2: In-plane directional dependence of the dielectric function $\varepsilon(\omega)$ from 0 eV to 20 eV over irreducible (IBZ) and reducible (RBZ) Brillouin zones for graphene (Gr), phosphorene (Pn), anatase (A-TiO_2) and rutile (R-TiO_2).

	$\mathrm{Re}\left[\int \lvert\varepsilon_{xx} - \varepsilon_{yy}\rvert\, d\omega\right]$		$\mathrm{Im}\left[\int \lvert\varepsilon_{xx} - \varepsilon_{yy}\rvert\, d\omega\right]$	
Material	IBZ	RBZ	IBZ	RBZ
Gr	18.505	3.037	13.817	1.824
Pn	10.698	11.481	9.250	9.763
A-TiO_2	13.643	0.766	12.126	0.695
R-TiO_2	11.889	0.494	10.776	0.367

tion.

5.2 0D Fullerene and Chlorophyll

The first material we will consider is the isolated molecule of the 0D carbon allotrope fullerene (C_{60}). Here, in addition to calculations done using both LCAO-TDDFT-k-ω and G_0W_0-BSE, we have data for the experimental photoabsorption cross section for fullerene molecules in both gas phase and in a hexane solution. To compare with these measured optical absorptions both in gas phase and in solution, defined as the logarithm of the ratio of the input and output intensities $\log_{10}(I_0/I)$, we use the Chako–Linder relation [152, 131]

$$\frac{[n^2(\omega)+9]^2}{9n(\omega)}\frac{\ln(I_0/I)}{CN_A d} \approx \frac{4\pi}{c}\,\mathrm{Re}[\sigma_{0D}(\hat{\mathbf{q}},\omega)], \tag{5.1}$$

directly relating the intensity ratio to the photoabsorption cross section, as defined in Eq. 2.50, with σ_{0D} the 0D conductivity, as defined through the polarizability via Eq. 2.67. Here $n(\omega)$ is the refractive index of the solution, N_A is Avogadro's number, C is the molar concentration in solution, d is the optical path length, and c is the speed of light, since the refractive index for gases is approximately one. Using this equation we can directly compare both peak positions and intensities of the photoabsorption cross sections calculated within LCAO-TDDFT-k-ω and G_0W_0-BSE with those extracted from experiments in gas phase and solution.

Table 5.3 shows measured and calculated energies for C_{60}'s third bright $\pi - \pi^*$ exciton, $\hbar\omega_\pi$, corresponding to the third peak in the photoabsorption spectra, from experiments in gas phase and hexane solution [131], G_0W_0-BSE [134, 133] calculations, and LCAO-TDDFT-k-ω calculations neglecting and applying the GLLB-SC correction Δ_x from Table 5.1. Neglecting the derivative discontinuity correction leads to an underestimation of the peak position by about 1 eV. This

Table 5.3: Measured and calculated energies of the third bright $\pi - \pi^*$ exciton of fullerene (C_{60}) $\hbar\omega_\pi$ in eV.

Method	$\hbar\omega_\pi$ (eV)
Measurement in Gas Phase	5.97[a]
Measurement in Hexane	5.86[b]
G_0W_0-BSE (VASP)	6.04[c]
G_0W_0-BSE (QE)	6.50[d]
LCAO-TDDFT-k-ω ($\varepsilon_{n'} - \varepsilon_n$)	4.79[e]
LCAO-TDDFT-k-ω ($\varepsilon_{n'} - \varepsilon_n + \Delta_x$)	5.56[e]

[a] Ref. [153] [b] Ref. [132] [c] Ref. [134] [d] Ref. [133]
[e] This work.

shows how essential this correction is for reproducing the measured spectra of C_{60}.

In Fig. 5.1(a) and (b) we show the electron and hole spectral densities in real space at the energy associated with the third bright $\pi - \pi^*$ exciton, $\hbar\omega_\pi \approx 5.6$ eV for LCAO-TDDFT-k-ω and $\hbar\omega_\pi \approx 6.0$ eV with G_0W_0-BSE [134]. For LCAO-TDDFT-k-ω, these spectral densities are calculated using Eqs. 2.62, 2.63, while a similar formalism yields the spatial distribution for G_0W_0-BSE. In both figures, the electron spectral density $\rho_e(\mathbf{r}, \omega_\pi)$ rests largely on the π anti-bonding orbitals on the exterior of the fullerene molecule, while the hole spectral density $\rho_h(\mathbf{r}, \omega_\pi)$ is concentrated on the interior π bonding orbitals, yielding the desired semi-quantitative agreement between the two methods.

Fig. 5.1(c) compares photoabsorption cross sections obtained from experimental photoexcitation data [131], G_0W_0-BSE calculations that explicitly include excitonic effects present in fullerene, and LCAO-TDDFT-k-ω calculations using the *a priori* and *a posteriori* scissors corrections. Both G_0W_0-BSE and the LCAO-TDDFT-k-ω implementations yield single-transition peaks that largely agree with the well-separated ones observed in both gas phase and hexane solution experiments, noting that the measured photoabsorption cross sections in these experiments have uncertain relative intensities as seen in Ref. [153]. While G_0W_0-BSE gives peaks that increase in intensity gradually, consistent with experimental spectra, LCAO-TDDFT-k-ω presents the opposite behaviour, the two methods agreeing in intensity only for the first bright exciton. This is perhaps a direct effect of the neglect of the excitonic effects that are described within a G_0W_0-BSE framework. Fig. 5.1(c) also shows firsthand the significant reduction in the intensities of the lower energy peaks when the scissors correction is applied *a priori*, giving somewhat better agreement with experiment and G_0W_0-BSE calculations for the first bright exciton than the *a posteriori* scissors approach. This is a direct result of the factor Δ_x included in the denominator of the dipole transition elements $f^{\hat{\mathbf{q}}}_{nn'\mathbf{k}}$ defined in Eq. 2.60. Overall we see that LCAO-TDDFT-k-ω and G_0W_0-BSE both capture the peak locations and spatial distribution qualitatively and in line with experiment, with the intensities of the peaks in disagreement at higher energies.

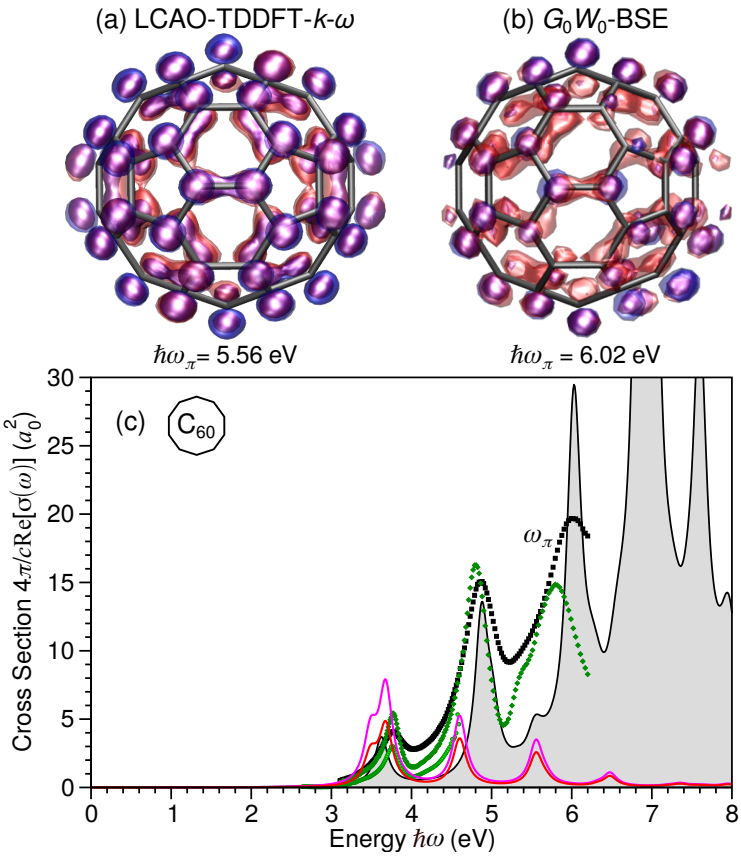

Figure 5.1: Fullerene's (C_{60}) electron (blue) and hole (red) spectral densities for the third bright $\pi - \pi^*$ exciton ω_π from (a) LCAO-TDDFT-k-ω and (b) G_0W_0-BSE [134] and (c) photoabsorption cross section $\frac{4\pi}{c}\mathrm{Re}[\sigma(\hat{\mathbf{q}}_\perp, \omega)]$ in units of a_0^2 versus energy $\hbar\omega$ in eV from G_0W_0-BSE (black thin solid line) [134], LCAO-TDDFT-k-ω *a priori* (red) and *a posteriori* (magenta) scissors calculations, and measurements in gas phase (black squares) [131] and hexane from Ref. [131] (dark green diamonds) and this work (light green circles), shifted by the Chako-Linder factor in Eq. 5.1.

The second set of materials we will consider that fall into the 0D category are the chlorophyll *a* and chlorophyll *b* (Chl *a* and Chl *b*) monomers. These molecules serve as the basic building blocks of the light harvesting complex (LHC) in green plants (LHCII), and are a prime example of the need for computationally efficient codes such as LCAO-TDDFT-k-ω to model the absorption properties of larger macromolecular systems. Although much progress has been made in the experimental measurement of the optical spectra of individual Chl *a* and Chl *b* monomers [137, 154, 138] and in their description at the TDDFT level [155, 156, 140, 157], the extension of these theoretical methods to the description of large chlorophyll-containing biomolecules and the role of surrounding proteins in photosynthesis requires computational resources beyond the capacity of most researchers [139, 158]. Since many of the previously mentioned TDDFT methods become too computationally expensive for such large biomolecules, we are unable to

provide G_0W_0-BSE as a benchmark. In addition, while we consider both LCAO-TDDFT-k-ω and PW-TDDFT-k-ω for the description of the relatively small *cut* monomers, *cut*·$\mathbf{1}^+$ and *full*·$\mathbf{1}^+$ are only considered in the LCAO-TDDFT-k-ω approach, as these larger systems become intractable without the memory savings provided by LCAO's reduced degrees of freedom.

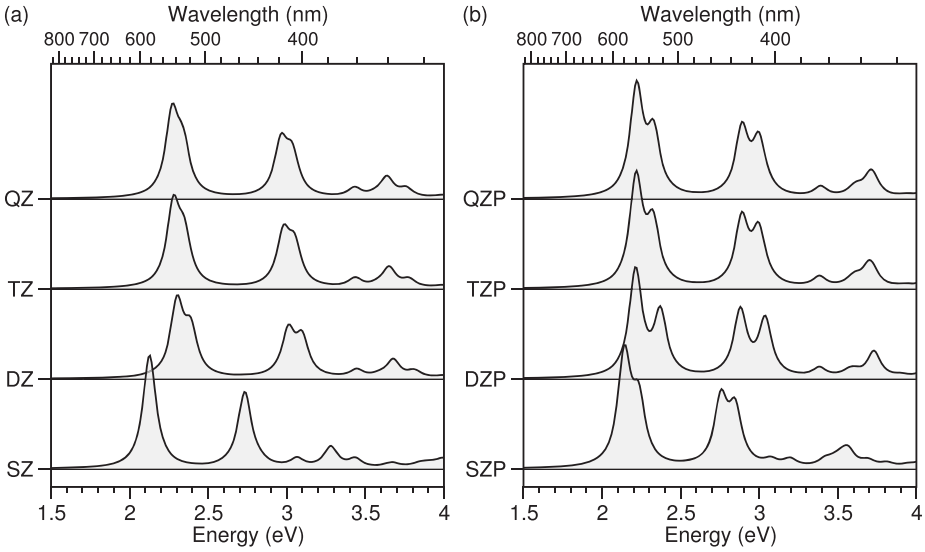

Figure 5.2: Dependence of the LCAO-TDDFT-k-ω optical absorption spectra of Chl *a* on the single-ζ (SZ), double-ζ (DZ), triple-ζ (TZ), and quadruple-ζ (QZ) LCAO basis sets (a) without and (b) with polarization (P) functions.

As stated in Section 2.1.5, a disadvantage of the LCAO method is that the addition of more radial functions or polarization does not necessarily yield a converged spectra. This is because increasing the number of radial functions or adding polarization does not provide a complete basis, e.g., for describing Rydberg or vacuum states. For this reason, it may be necessary to test smaller sample systems for convergence using a wide set of LCAO basis functions. In addition, comparing the result of this basis set convergence test with well-converged PW results can provide an additional check on the LCAO-TDDFT-k-ω method for smaller systems, *cut* Chl *a* and Chl *b* monomers in this case. In Fig. 5.2 we present the optical absorption spectrum of the *cut* Chl *a* monomer using basis sets with ζ ranging from single to quadruple, and with polarization turned off (SZ, DZ, TZ, and QZ) and on (SZP, DZP, TZP, and QZP). SZ is the smallest basis and therefore takes the least memory of the eight basis sets, while QZP represents the case where each atom's valence states are represented by four radial functions each along with a polarization function with higher-order character.

Fig. 5.2 (a) shows that this most reduced SZ basis gives a red-shifted spectrum compared to DZ, TZ, and QZ, the latter two yielding nearly identical spectra. DZ gives the same peak intensities as these latter two methods but with the higher-ζ counterparts giving more separation between the two peaks, excitations corresponding to the Q and Soret bands, respectively. Fig. 5.2 (b) shows that including polarization tends to increase the overall intensity of the spectra, and

that the effect of increasing the number of radial functions matches the unpolarized basis set choice, with red-shifted spectra for the SZP and a slight splitting of the two peaks relative to TZP and QZP. While Fig. 5.2 suggests that convergence in the radial functions is achieved with a DZP basis set, it is unclear whether the inclusion of other basis functions, e.g., Gaussians, would provide spectra in better agreement with experiment. For this reason a comparison with a fully-converged PW-TDDFT-k-ω calculation is essential to determine whether a DZP basis set is in fact sufficient within LCAO-TDDFT-k-ω.

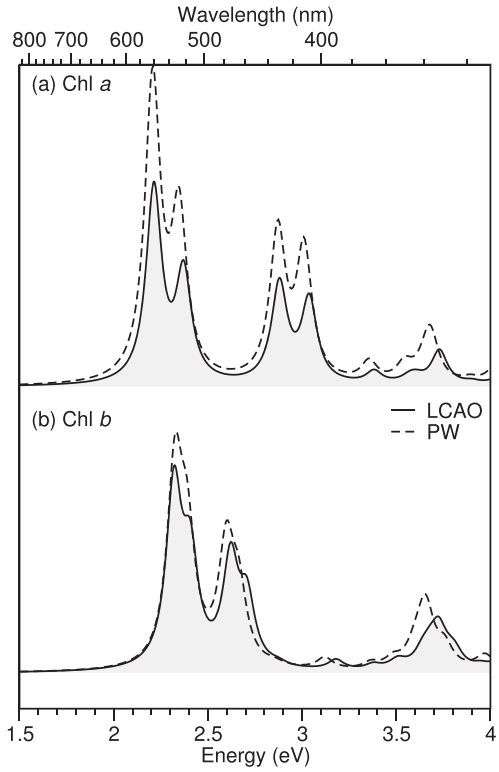

Figure 5.3: Optical absorption of *cut* Chl *a* and *b* monomers from LCAO-TDDFT-k-ω (solid lines) and PW-TDDFT-k-ω (dashed lines).

Fig. 5.3 shows the optical absorption spectra for *cut* Chl *a* and *b* monomers using DZP for LCAO-TDDFT-k-ω and a sufficient energy cutoff for convergence of the PW-TDDFT-k-ω calculations. The locations of the peaks obtained within the two methods agree to within 0.1 eV, with the higher energy part of the LCAO spectra blue-shifted by this amount. Additionally, the intensity of the spectra is significantly lower with LCAO-TDDFT-k-ω for *cut* Chl *a* whereas it is in semi-quantitative agreement for *cut* Chl *B*. As Fig. 5.2 showed that polarization for this system increases the intensity, DZP is the logical choice to obtain better agreement with the converged PW calculation. However, as we will show, the choice of LCAO or PW basis set does not always yield higher or lower intensity spectra in general.

Fig. 5.4 shows spectra obtained from LCAO-TDDFT-k-ω including the *a posteriori* scissors correction Δ_x applied to Chl *a* and Chl *b* monomers shown in Fig. 3.4, along with experimental

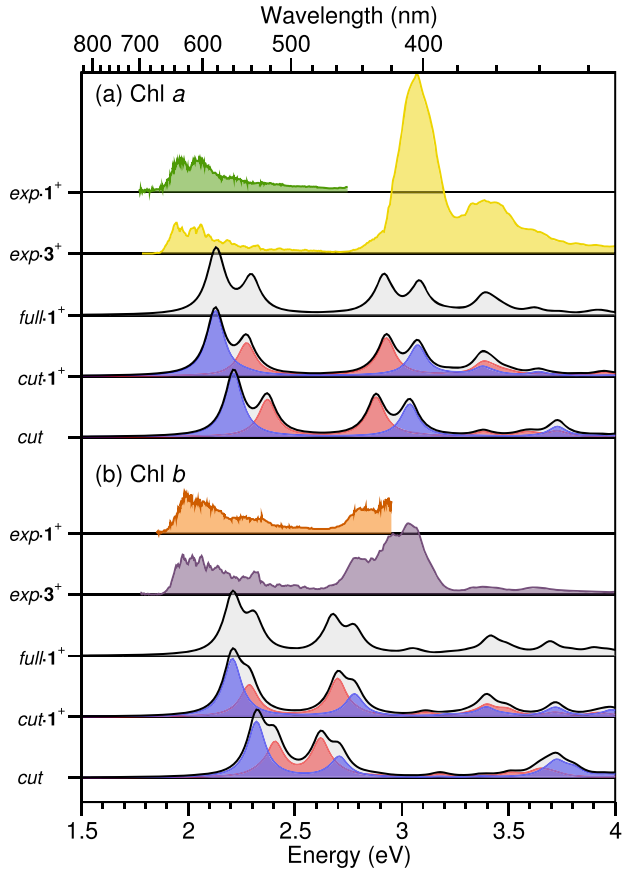

Figure 5.4: Optical absorbance spectra Im $\varepsilon(\omega)$ calculated in LCAO-TDDFT-k-ω for (a) Chl a and (b) Chl b monomers for *cut*, *cut*·$\mathbf{1}^+$, and *full*·$\mathbf{1}^+$ structures, with spectra broken along the x (blue) and y (red) polarization directions shown in Fig. 3.4. These spectra are compared with experimentally measured spectra for monomers with monocationic tetramethylammonium $\mathbf{1}^+$ [137] and acetylcholine $\mathbf{3}^+$ [154] charge tags, written as *exp*·$\mathbf{1}^+$ and *exp*·$\mathbf{3}^+$ respectively.

spectra for these monomers with either monocationic tetramethylammonium $\mathbf{1}^+$ [137] or acetylcholine $\mathbf{3}^+$ [154] charge tags. For both monomers the shape of the spectra shows the same qualitative behaviour between the *full*·$\mathbf{1}^+$ and *cut*·$\mathbf{1}^+$ structures, suggesting that the influence of the carbon chain on the optical absorption spectra is minimal. The effect of the charge tag is more apparent, as it causes a widening of the gap between the Q and Soret band peaks, blue-shifting the former and red-shifting the latter.

For Chl a, the locations of the maxima of the Q band peaks in the experimental spectra and the first excitation energy from LCAO-TDDFT-k-ω calculations for the *full*·$\mathbf{1}^+$ structure differ by less than 0.2 eV, while the locations of the maxima of the Soret band peaks and the fourth excitation energy from LCAO-TDDFT-k-ω differ by less than 0.1 eV. Chl b *full*·$\mathbf{1}^+$ exhibits similar but poorer agreement, with the Q band blue-shifted by 0.23 eV and the Soret band red-shifted by 0.26 eV relative to experiment.

The experimental data also suggests that the relative intensity of the Q band should be less than that of the Soret band, whereas the opposite behaviour is seen with LCAO-TDDFT-k-ω. This disparity in the computed and measured intensities may lie in the neglect of charge transfer excitations within the LCAO-TDDFT-k-ω methodology, an effect also seen in Fig. 5.1 for the higher-energy excitations of C_{60}. Since these types of excitonic effects have been put forward in the literature [137] as a possible absorption mode, the inability of LCAO-TDDFT-k-ω to capture the intensity near the Soret band peak lends further credence to this possibility. In addition, this underestimation of the higher energy intensities further justifies the use of the *a priori* scissors correction over its *a posteriori* counterpart, the former having a diminishing effect on spectral intensity, bringing the Q band to Soret band intensity ratio into better agreement with the experimentally observed one.

5.3 1D Single-Walled Carbon Nanotubes

We continue our evaluation of the LCAO-TDDFT-k-ω code by considering single-walled carbon nanotubes (SWCNTs), which are periodic along a single axis. For a set of fifteen semiconducting and four metallic SWCNT chiralities, the LCAO-TDDFT-k-ω implementation, including the derivative discontinuity correction Δ_x, has previously been shown [9] to semi-quantitatively reproduce experimental spectra over a wide range of energies. We will consider herein the optical conductivity $\sigma_{1D}(\hat{\mathbf{q}}_{\parallel}, \omega)$ for $\hat{\mathbf{q}}_{\parallel}$ along the SWCNT's axis, as shown in Eq. 2.69, for two prototypical SWCNTs, namely, the metallic armchair (10,10) SWCNT and the semiconducting zigzag (10,0) SWCNT. As discussed in Section 3.3, nanotubes can be either armchair, zigzag, or chiral, with metallic or semiconducting properties, depending on the choice of chiral and translational vectors.

Fig. 5.5(a,b) shows the real and imaginary parts of the axial conductivity of the semiconducting (10,0) armchair SWCNT using PW-TDDFT-k-ω, G_0W_0-BSE, LCAO-TDDFT-k-ω with the *a priori* scissors correction for basis sets ranging from SZP to QZP , and LCAO-TDDFT-k-ω with the *a posteriori* scissors correction for the DZP basis set. There is semi-quantitative agreement between the *a posteriori* LCAO-TDDFT-k-ω and PW-TDDFT-k-ω for both intensities and peak positions, with LCAO slightly underestimating the intensities relative to PW. As with chlorophyll, the DZP basis set choice is shown to yield converged spectra. The G_0W_0-BSE calculations show how excitonic effects influence the axial conductivity, expected to be stronger for 1D systems, such as SWCNTs, due to the reduced screening between charge carriers. We see that both PW-TDDFT-k-ω and LCAO-TDDFT-k-ω overestimate the energy for the first transition E_{11} by approximately 0.2 eV relative to G_0W_0-BSE. The explanation for this close agreement has been previously explored [9] in the discussion for Fig. 2.2. Regarding the intensity of the spectra, G_0W_0-BSE and LCAO-TDDFT-k-ω *a posteriori* scissors are in semi-qualitative agreement, with G_0W_0-BSE giving a relatively lower value for the E_{11} transition and a relatively higher one for the E_{22} transition.

The difference between calculating the spectra with LCAO-TDDFT-k-ω using *a priori* or

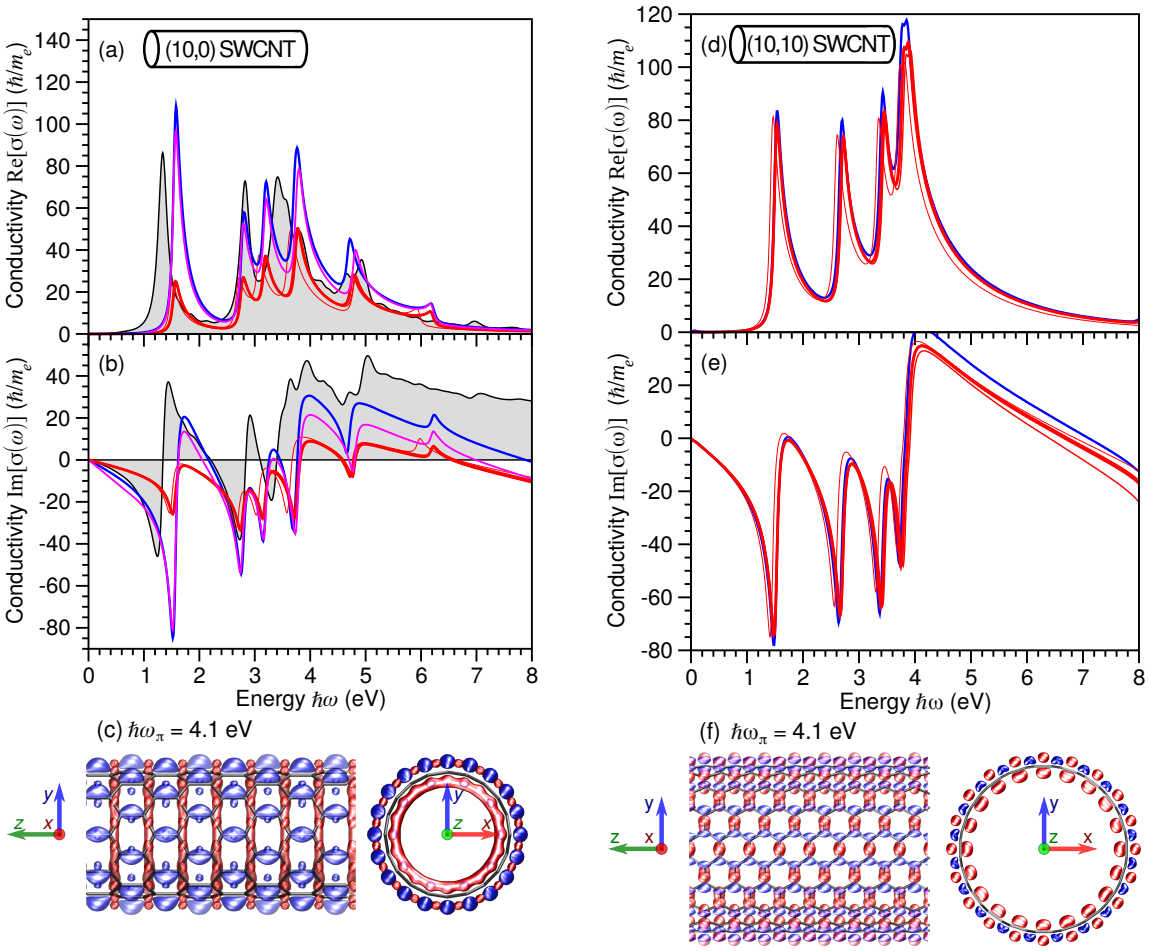

Figure 5.5: Semiconducting (10,0) SWCNT (a) real and (b) imaginary parts of the axial optical conductivity $\sigma(\hat{\mathbf{q}}_{\|},\omega)$ in units of $\hbar/m_e$ versus energy $\hbar\omega$ in eV for $\hat{\mathbf{q}}$ along the nanotube axis, with G_0W_0-BSE (black), PW-TDDFT-k-ω (blue), and LCAO-TDDFT-k-ω *a priori* (red) with single (SZP), double (DZP), triple (TZP), and quadruple (QZP) ζ-polarized basis sets (in order of increasing thickness) and *a posteriori* (magenta) calculations with DZP and (c) positive (red) and negative (blue) isosurfaces for the electron-hole spectral density difference at the $\hbar\omega_\pi \approx 4.1$ eV π plasmon, with axes shown as insets. These same curves are shown for metallic (10,10) SWCNT (d) real and (e) imaginary parts of the axial optical conductivity along with (f) positive (red) and negative (blue) isosurfaces for the density difference, neglecting G_0W_0-BSE and scissors-corrected calculations for this zero-gap material.

a posteriori scissors is seen in the substantial change in both absolute and relative intensities. The derivative discontinuity correction, Δ_x, for the zigzag SWCNT is approximately 0.788 eV, half the energy of the E_{11} transition at $\hbar\omega \approx 1.55$ eV. This means that for this first transition the *a priori* intensity is lower by a factor of $\frac{\omega+\Delta_x}{\omega} \approx 2$. In contrast to fullerene, *a posteriori* scissors LCAO-TDDFT-k-ω provides a better description of the optical spectra for 1D SWCNTs,

reproducing the G_0W_0-BSE spectra, relative to *a priori* scissors.

Fig. 5.5(c) shows the electron-hole spectral density difference for the $\pi - \pi^*$ transition located at $\hbar\omega_\pi \approx 4.1$ eV, calculated using Eq. 2.66. We see that the hole spectral density is more prominent on π orbitals of the carbon-carbon bonds within the nanotube while the spectral density of the excited electron lies predominantly outside the (10,0) SWCNT above the carbon-carbon bonds. The hole spectral density also reveals itself in rings around the circumference of the nanotube in contrast to the more localized electron spectral density.

Fig. 5.5(d,e) shows the real and imaginary parts of the axial conductivity of the metallic (10,10) zigzag SWCNT using PW-TDDFT-k-ω and LCAO-TDDFT-k-ω for basis sets ranging from the SZP to the QZP representation, noting that G_0W_0-BSE is not required for metallic systems and that the *a priori* and *a posteriori* scissors approaches are equivalent when $\Delta_x = 0$. As with the (10,0), there is semi-quantitative agreement between the LCAO-TDDFT-k-ω and PW-TDDFT-k-ω methods for both intensities and peak positions, with LCAO slightly underestimating the intensities relative to PW across all choices of basis set. The real part of the (10,10) SWCNT's conductivity $\mathrm{Re}[\sigma(\hat{\mathbf{q}}_\parallel, \omega)]$ is made up of intense peaks located at 1.5, 2.7, 3.4, and 3.9 eV, with intensities on the order of $\sim 100\ \hbar/m_e$ matching with those of the (10,0) SWCNT and with the intensity of the ballistic conductance expected for a metallic armchair SWCNT.

Figure 5.5(f) shows the electron-hole spectral density difference for the $\pi - \pi^*$ transition located at $\hbar\omega_\pi \approx 4.1$ eV, with the distribution of the electron and hole spectral densities matching with what is seen in the (10,0) SWCNT. The main difference is in the arrangement of the electron spectral density in stripes running alongside the exterior of the nanotube parallel to the axis, with a relatively more localized hole spectral density. The locations of the electron and hole densities predominantly outside and within the nanotube matches with what is seen in Fig. 5.1(c) for fullerene, suggesting this may be a more general property of carbon allotropes.

5.4 2D Graphene and Phosphorene Monolayers

We move on to the evaluation of our LCAO-TDDFT-k-ω method for isolated surface layers that are only non-periodic in a single direction, starting with graphene, which in the case of zero doping shows no band gap and therefore no Δ_x scissors correction.

5.4.1 Optical Conductivity

Fig. 5.6 shows Gr's in-plane 2D conductivity from PW-TDDFT-k-ω using two different software packages, Quantum Espresso (QE) and GPAW, along with LCAO-TDDFT-k-ω calculations in GPAW using the SZP, DZP, TZP, and QZP basis sets. Results between PW-TDDFT-k-ω and LCAO-TDDFT-k-ω show excellent agreement for the $\pi \rightarrow \pi^*$ transition near 4.2 eV for all basis sets, although the LCAO mode shows a longer-tailed peak leading away from this transition. The

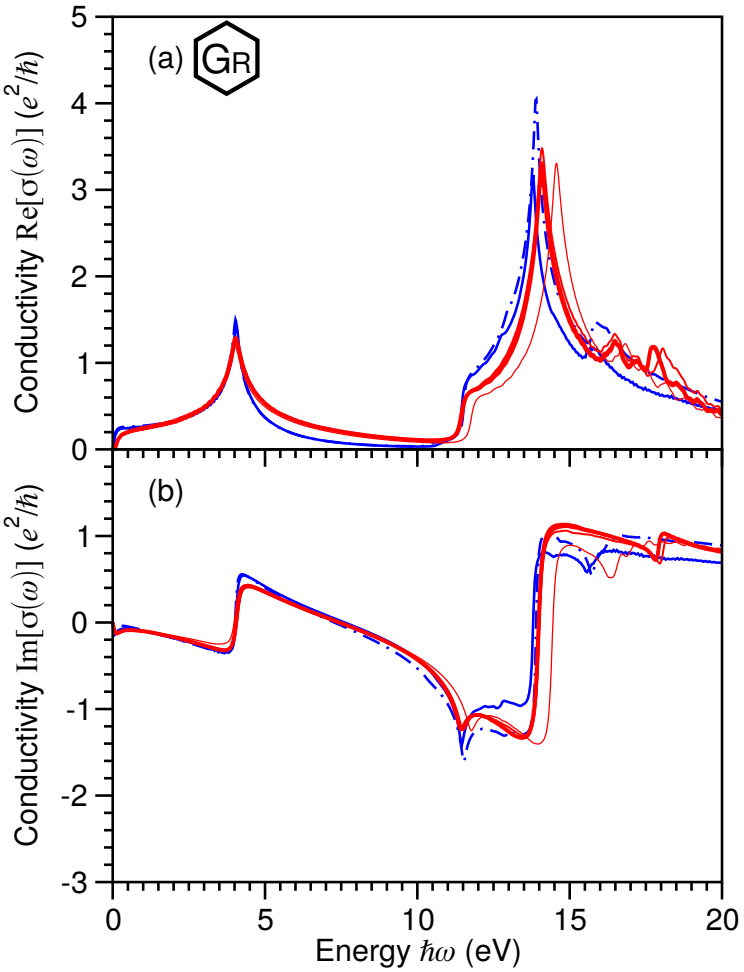

Figure 5.6: Graphene (GR) in-plane (a) real and (b) imaginary parts of the conductivity $\sigma(\hat{\mathbf{q}}_{\parallel}, \omega)$ in units of $e^2/\hbar$ versus energy $\hbar\omega$ in units of eV evaluated with PW-TDDFT-k-ω (blue lines) in both GPAW (solid) and QE (dash-dotted) and LCAO-TDDFT-k-ω (red solid lines) for SZP, DZP, TZP, and QZP basis sets shown in increasing order of thickness.

$\sigma \rightarrow \sigma^*$ transition near 13.5 eV for LCAO-TDDFT-k-ω using the DZP, TZP, or QZP basis sets agree closely with PW-TDDFT-k-ω, noting that the peak location error between the LCAO and PW implementations is approximately equal to the error between two separate PW-TDDFT-k-ω calculations. GR's conductivity shows additional peaks beyond this transition, a small one appearing near 16 eV in all approaches and a secondary peak near 18 eV for calculations done with LCAO-TDDFT-k-ω for DZP, TZP, or QZP basis sets. Whether this latter peak has a physical origin or is an artifact or the basis set choice will be explored when we analyze the spatial distribution of the associated exciton through the spectral density. Once again, DZP is found to be sufficient to describe the location and intensities of the peaks to sufficient accuracy.

Fig. 5.7 presents the in-plane conductivity for PN using PW-TDDFT-k-ω, G_0W_0-BSE, and

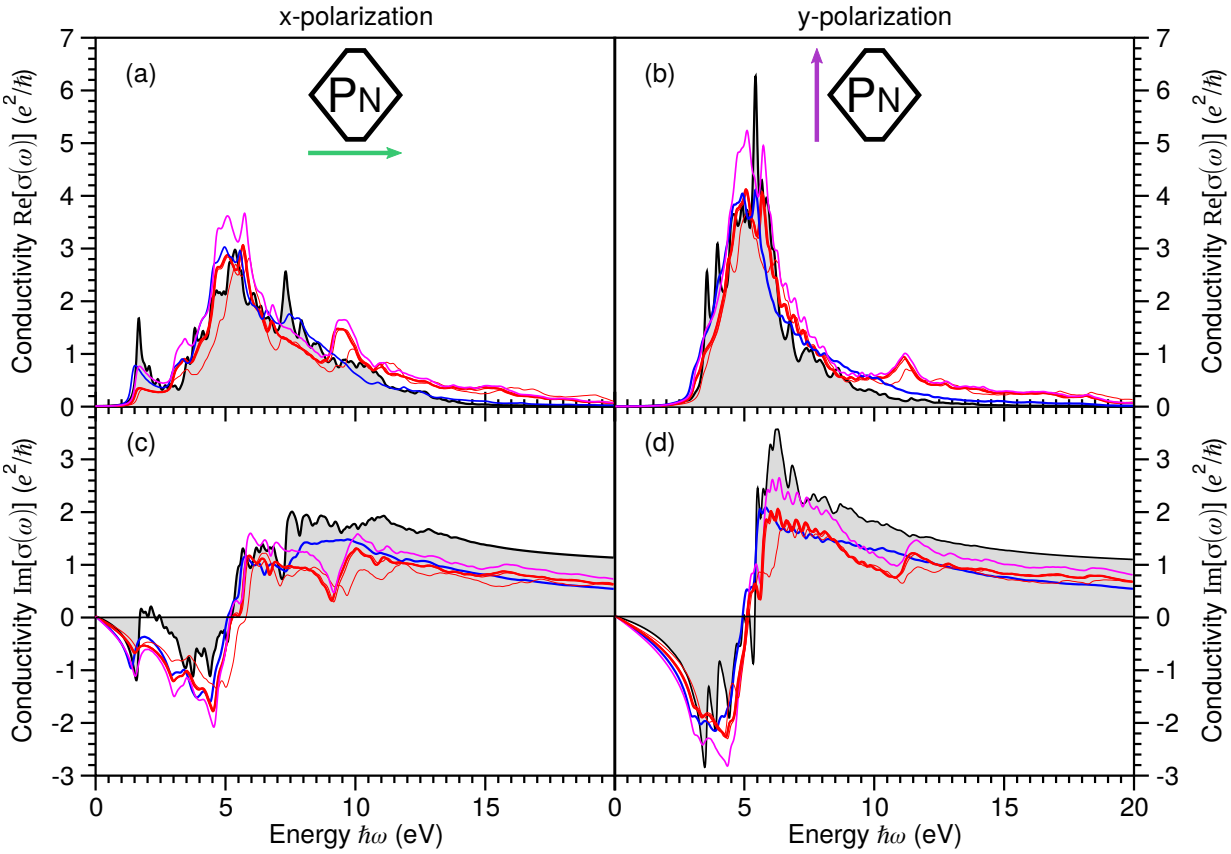

Figure 5.7: Phosphorene (PN) in-plane (a,b) real and (c,d) imaginary parts of the conductivity $\sigma(\hat{\mathbf{q}}_{\|}, \omega)$ in units of $e^2/\hbar$ versus energy $\hbar\omega$ in units of eV from PW-TDDFT-k-ω (blue), G_0W_0-BSE (black), and LCAO-TDDFT-k-ω *a priori* (red) calculations for SZP, DZP, TZP, and QZP basis sets shown in increasing order of thickness and LCAO-TDDFT-k-ω *a posteriori* (magenta) calculations for light polarized along (a,c) $\hat{\mathbf{q}}_x$ and (b,d) $\hat{\mathbf{q}}_y$.

LCAO-TDDFT-k-ω, showing the *a priori* calculation for SZP, DZP, TZP, and QZP basis sets and the *a posteriori* calculation for DZP, along the two orthogonal in-plane polarization directions to verify PN's well-known planar anisotropy. All these approaches capture the anisotropic peak near 1.8 eV in the x-direction, G_0W_0-BSE yielding the greatest conductivity with PW-TDDFT-k-ω and LCAO-TDDFT-k-ω *a posteriori* scissors calculations in semi-quantitative agreement, and the LCAO-TDDFT-k-ω *a priori* scissors, as for the (10,0) SWCNT in Fig. 5.5, yielding a much lower intensity than the G_0W_0-BSE conductivity. In addition, all methods capture the higher intensity peak situated near 5 eV in the y-direction. LCAO-TDDFT-k-ω with both scissors approaches yields a longer-tailed spectra along with a peak near 9 eV in the x-direction and 11 eV in the y-direction, similar to the peak uniquely seen at 18 eV in Fig. 5.6. The difference in intensities between PW-TDDFT-k-ω and LCAO-TDDFT-k-ω using the two scissors approaches again suggests that there is no way ahead of time to predict which approach will yield intensities most consistent with more accurate TDDFT methods or experiment.

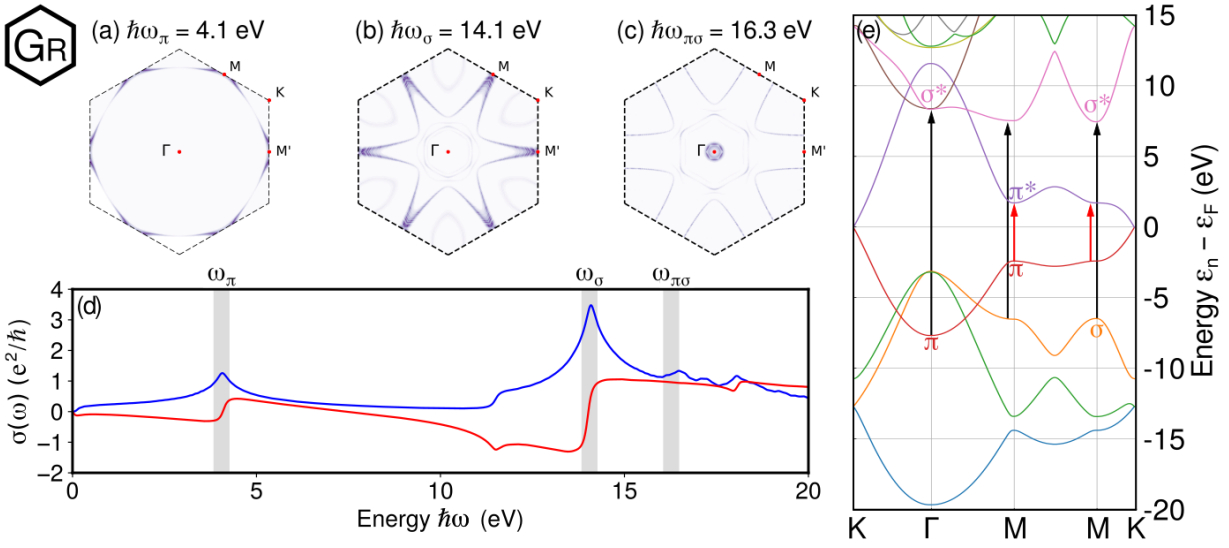

Figure 5.8: BZ profiles of the GR (a) $\pi - \pi^*$ transition $\hbar\omega_\pi \approx 4.1$ eV, (b) $\sigma - \sigma^*$ transition $\hbar\omega_\sigma \approx 14.1$ eV, and (c) $\pi - \sigma^*$ transition $\hbar\omega_{\pi\sigma*} \approx 16.3$ eV, (d) real (blue) and imaginary (red) parts of the conductivity $\sigma(\hat{\mathbf{q}}_\parallel, \omega)$ in units of $e^2/\hbar$ versus energy $\hbar\omega$ in units of eV from LCAO-TDDFT-k-ω, and (e) band structure with arrows showing the three transitions under consideration.

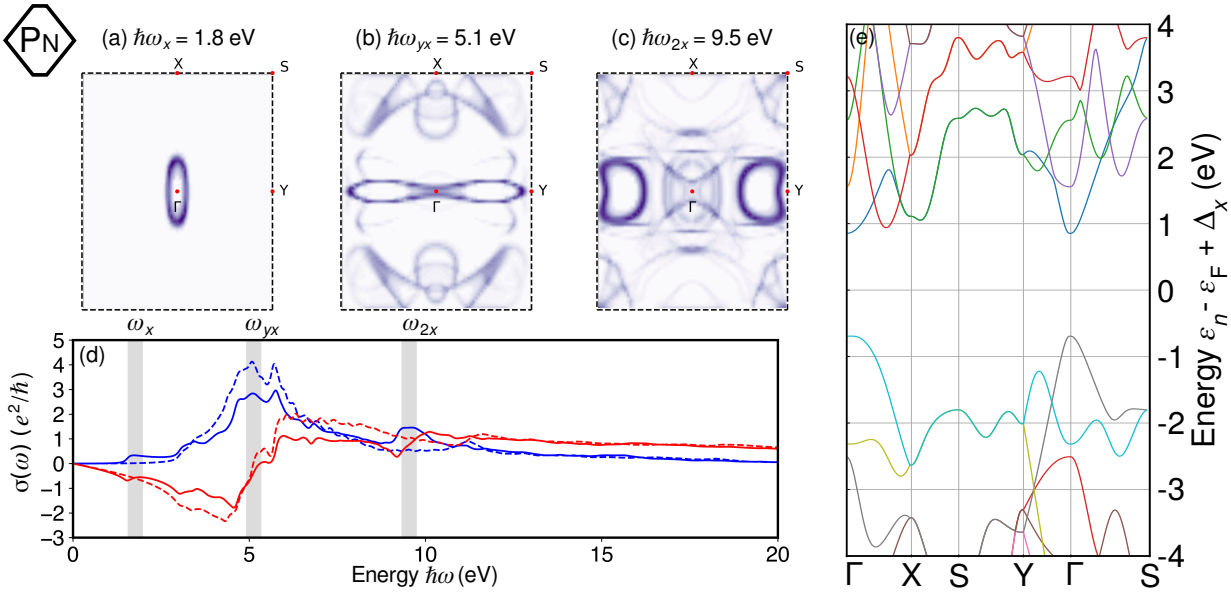

Figure 5.9: BZ profiles of the PN (a) x-polarized transition $\hbar\omega_x \approx 1.8$ eV, (b) y and x-polarized transition $\hbar\omega_{yx} \approx 5.1$ eV, and (c) second x-polarized transition $\omega_{2x} \approx 9.5$ eV, (d) real (blue) and imaginary (red) parts of the conductivity $\sigma(\hat{\mathbf{q}}_\parallel, \omega)$ in units of $e^2/\hbar$ versus energy $\hbar\omega$ in units of eV from LCAO-TDDFT-k-ω for $\hat{\mathbf{q}}_x$ (solid lines) and $\hat{\mathbf{q}}_y$ (dashed lines) polarized light, and (e) band structure including the Δ_x scissors correction.

5.4.2 Brillouin Zone Transition Profiles

As discussed above, calculations using the LCAO-TDDFT-k-ω method require that all calculations of optical response functions are done using only the head, i.e. $\mathbf{G} = \mathbf{G}' = 0$. One effect of this is that every transition computed within the LCAO-TDDFT-k-ω framework is a direct transition between states of the same wavenumber, so at each wavenumber in the BZ of a material the LCAO-TDDFT-k-ω code yields the bands involved in a transition at a given excitation energy, allowing a visualization of the exciton's distribution in reciprocal k-space by creating a contour map over the first BZ at a given energy. In Figs. 5.8(a,b,c) and 5.9(a,b,c) we show for a given k-point and energy $\hbar\omega$ the sum of oscillator strengths $f^{\hat{\mathbf{q}}_x}_{nn'\mathbf{k}}$ defined in Eq. 2.56 for transitions $n \rightarrow n'$ within a $\hbar\Delta\omega = \pm 0.2$ eV range of $\hbar\omega$ weighted by a Lorentzian broadening of 0.1 eV in order to smooth out the reciprocal space spectra. Choosing oscillator strengths for excitations polarized in-plane allows for a visualization of Gr's and Pn's most relevant spectral contributions, although such diagrams can be tailored for any polarization direction, or to include only certain bands. These contour maps give more information at a given energy than the band structure alone, as intense peaks that go beyond the linear cross-sections lying between high-symmetry points are included, and the connections between said points are also highlighted.

Fig. 5.8(a), (b), and (c) show Gr's transitions in the hexagonal BZ for the $\pi - \pi^*$, $\sigma - \sigma^*$, and $\pi - \sigma^*$ excitations seen previously as the three main peaks in the in-plane conductivity in Fig. 5.6. The transitions shown via the sum of the oscillator strengths, which follow the symmetry of the reciprocal lattice, must reflect the same symmetric properties, which in the case of Gr yield a six-fold symmetry. The 4.1 eV and 14.1 eV reciprocal space spectra are largely made up of transitions centred around the high symmetry M-point, with the $\pi - \pi^*$ transition directly connecting along the M→M' direction and the $\sigma - \sigma^*$ transition following a more star-shaped pattern. The peak at 16.3 eV shown in Fig. 5.8(c) is seen to occur largely at the Γ point, which given its energy and symmetry point, we attribute to a $\pi - \sigma^*$ as seen in the band structure in Fig. 5.8(e). A small annulus surrounds the Γ point at this peak, possibly a result of the choice of energy range surrounding the peak. These contour maps are expected to become more noisy and spread out as the number of bands available for transitions at each k-point increase in number.

Fig. 5.9(a), (b), and (c) shows three prominent transitions in Pn's orthorhombic BZ. Symmetry over this BZ means that transitions should exhibit a four-fold symmetry, as we clearly see in Fig. 5.9(a–c). Fig. 5.9(a) shows that the oscillator strengths associated with the anisotropic excitation near 1.9 eV also reveal anisotropic behaviour in the form of an elliptical annulus in the k_x and k_y wavenumbers, with the energy range in Fig. 5.9(d) chosen slightly off the peak center to better show this effect. This transition being centred around the Γ point matches with what is expected from the band structure in Fig. 5.9(e) where the conduction and valence bands are separated by approximately 1.8 eV after accounting for the derivative discontinuity correction.

Fig. 5.9(b) and (c) show BZ profiles for higher energy transitions: the $\omega_{yx} \approx 5.1$ eV prominent for y-polarized light and the $\omega_{2x} \approx 9.5$ eV transition uniquely found in our LCAO-TDDFT-k-ω code. Ellipse-like shapes appear as before, indicating that there is some structure to these excitations, but the scattered nature of the k-points relevant for these transitions suggests no particular

high-symmetry point is associated with these peaks in the in-plane conductivity spectra.

5.4.3 Exciton Density Projection Contours

The oscillator strengths presented in Eq. 2.56, used above in showing the reciprocal space profile of the excitation energies for Gr and Pn, are used again in Eq. 2.61 to define a two-point excitonic spectral function, from which electron and hole spectral densities are defined in Eqs. 2.63, 2.62. This allows us to visualize these spectral densities in real space for a given energy. Since these densities are defined as intensities in three spatial dimensions and at a given energy, we have previously presented the position space data as isosurfaces at a given energy in Fig. 5.1 and Fig. 5.5. However, to infer how the spatial profile of our excitations changes with energy requires projections in real space. In Fig. 5.10 and Fig. 5.11 we present one possible projection method, namely, summing the electron or hole spectral densities over planes normal to an axis of choice in the material. In this way the resulting contour map shows the change in this quantity along the axis over a range of energies. An alternate method of presenting this data is shown in Appendix E, where summation is only done along one direction and snapshots at different voxels give a more comprehensive view of the 2D spatial profile of excitations.

Fig. 5.10 shows four contour maps in real space for both the electron and hole density, projected onto either the z-axis, i.e., summing the spectral density over the xy-plane, or projected onto the C–C bond, i.e., summing over points in the plane normal to this bond. Fig. 5.10(a) and (b) shows the $\hbar\omega_\pi \approx 4.1$ eV transition in real space, with nodal planes visible between areas of brighter intensity, consistent with π^* and π orbitals. $\hbar\omega_\sigma \approx 14.1$ eV instead shows that the excitations are much more localized on the monolayer, consistent with σ^* and σ orbitals. The peak at 16.3 eV, which we inferred in Fig. 5.8 to be of $\pi - \sigma^*$ character, is not as clear as the other two transitions, but with no nodal plane these real space contours suggest the associated excitation will involve σ orbitals. Lastly, the peak near 18 eV seen in Fig. 5.6 for the LCAO-TDDFT-k-ω spectra extends significantly into the vacuum beyond the Gr monolayer. While this may suggest a Rydberg-like state [159], the mechanism by which these states come about is not visible in the spectra from standard TDDFT implementations. Given that Fig. 5.6 showed that this peak is absent in all other TDDFT methods suggests strongly that the 18 eV peak is an artifact of an incomplete basis set.

Fig. 5.10(c) and (d) shows contour maps integrated along planes normal to the carbon bonds in graphene. π^* and π orbital features are again present close to the peak at 4.1 eV, with the electron spectral density from (c) shared between the orbitals and the hole spectral density in (d) avoiding the space between them. The opposite is seen for the σ orbital peak with a reversal in the areas where the electron and hole spectral densities are most intense. This peak is also seen to be more diffuse in-plane, consistent with the fact that other carbon atoms in neighbouring cells also share bonds with the carbon atoms shown in the contour plot.

The upper inset of Fig. 5.10 uses isosurfaces to show the spectral density difference for the $\pi - \pi^*$ exciton labelled along the contour maps. It is important to note that the directional depen-

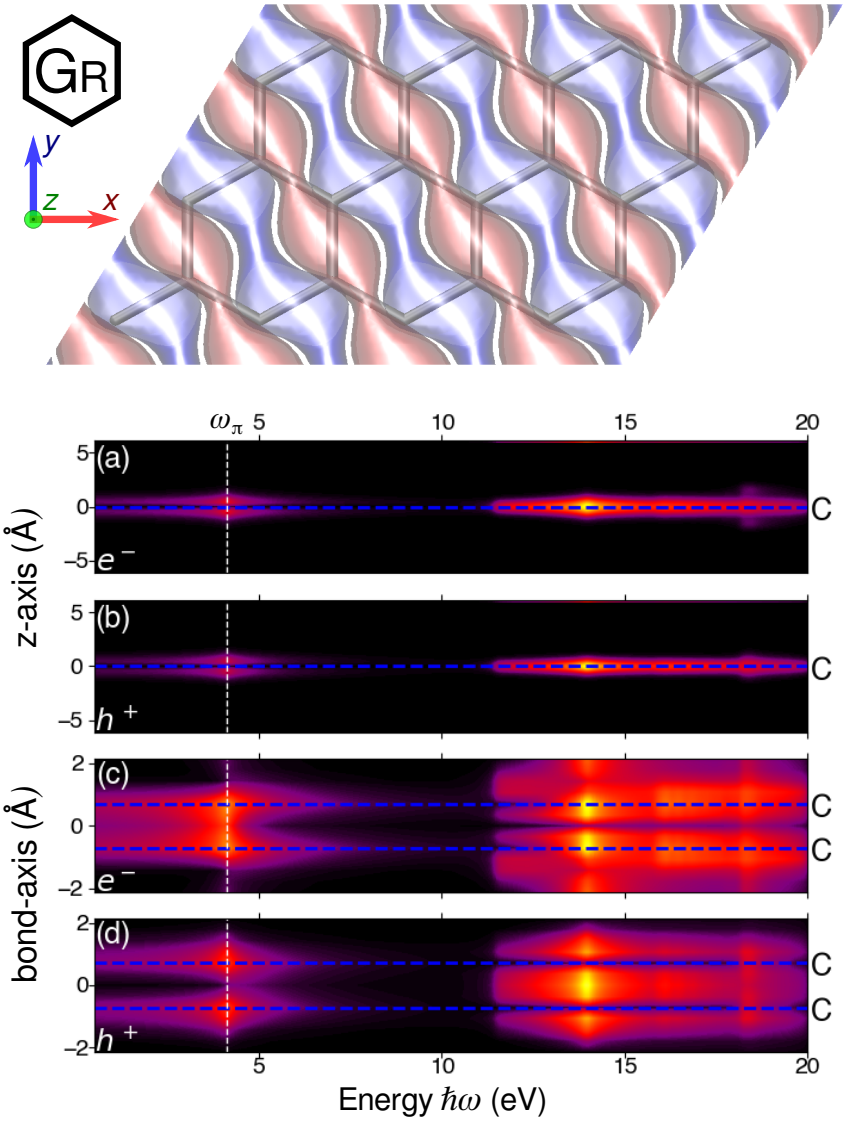

Figure 5.10: (a,c) electron and (b,d) hole spectral densities ρ_e and ρ_h for Gr calculated with LCAO-TDDFT-k-ω and C atoms (blue dashed lines) projected onto the (a,b) z or (c,d) bond axis versus energy $\hbar\omega$ in units of eV for light polarized in-plane, and positive (red) and negative (blue) isosurfaces $\pi - \pi^*$ exciton spectral density difference at $\hbar\omega_\pi = 4.1$ eV (white dashed line) with axes shown as an inset.

dence visible in the isosurface is an expected effect of the choice of polarization direction $\hat{\mathbf{q}}$, and calculating the same isosurface for a difference choice of $\hat{\mathbf{q}}$ must necessarily lead to an isomorphic density profile. This is unlike the contour maps which preserve the in-plane and out-of-plane symmetries of the Gr crystal structure. This inset reveals isosurfaces with striped electron spec-

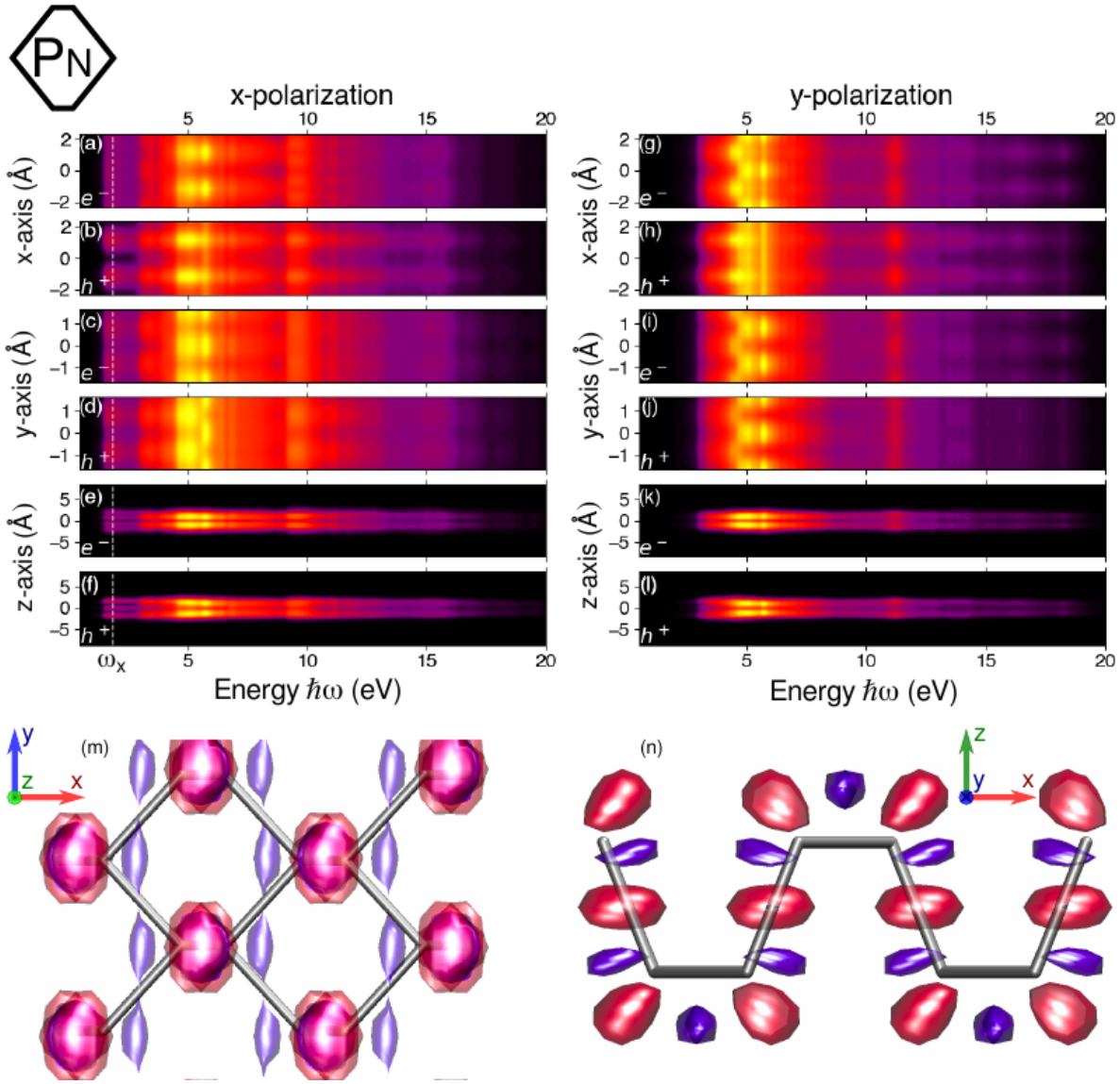

Figure 5.11: (a,c,e,g,i,k) electron (e^-) and (b,d,f,h,j,l) hole (h^+) spectral densities ρ_e and ρ_h for PN calculated with LCAO-TDDFT-k-ω *a priori* scissors correction, projected onto the (a,b,g,h) x, (c,d,i,j) y, or (e,f,k,l) z axis versus energy $\hbar\omega$ in units of eV for (a–f) x and (g–l) y polarized light, the directions depicted in the above inset, and (m,n) positive (red) and negative (blue) isosurfaces of the spectral density difference for the x-polarized exciton located at $\hbar\omega_x = 1.8$ eV (white dashed line) with axes shown as separate insets.

tral density similar to those seen in Fig. 5.5 for the (10,0) semiconducting SWCNT, more so than the localized pockets of electron spectral density seen for the (10,10) metallic SWCNT.

Due to the anisotropy of the phosphorene crystal structure, Fig. 5.11 shows projections onto

the x, y, and z axes for light polarizaed along both the x and y directions, as shown in the lower insets. Both the electron and hole spectral densities spread out over the entire unit cell, a result of the higher atomic number of phosphorene atoms. The x-axis projection for x-polarized light shows nodes in the electron spectral density and anti-nodes for the hole, reminiscent of the $\pi - \pi^*$ transition in Gr. This suggests these may also be the orbitals participating in the anisotropic peak. Beyond this, the two polarization directions match closely, as expected from the spectra shown in Fig. 5.7. The projections along the z-axis show that the electron and hole spectral densities do not extend far beyond the Pn crystal structure, while the nodal behaviour suggests a low carrier density between the two layers of phosphorene atoms.

Fig. 5.11(m) and (n) shows top and side vies of isosurfaces of the electron hole spectral density difference at $\hbar\omega_x \approx 1.8$ eV for x-polarized light. We neglect y-polarized light due to the low density of excitations at that energy. The isosurface reveals that the electron spectral density has stripes along the y-direction while the hole is localized along out-of-plane bonds. Dipoles, which point between areas of opposite charge densities, would therefore point in the x-direction. This gives us a real-space explanation for the anisotropic conductivity peak at this point.

5.4.4 Spectral Convergence with k-Point Spacing

While convergence with respect to the basis set choice has been emphasized throughout our evaluation of the LCAO-TDDFT-k-ω code as one of the most important factors to consider when using this method to determine the optical properties of materials, it is pertinent to also show that LCAO-TDDFT-k-ω converges with respect to another important parameter, the k-point grid spacing. Gr and Pn have the highest band dispersion of all the materials under consideration, so they provide the best test cases.

Fig. 5.12 provides contour plots starting at $\Delta k_x = \Delta k_y \approx 0.05$ nm^{-1} spacing, which for Gr corresponds to a Monkhorst-Pack $541 \times 541 \times 1$ k-point mesh. For Gr, the two main peaks, along with the likely artificial 18 eV peak, are all already converged for relatively coarse grid spacings on the order of $\Delta k \lesssim 0.5$ nm^{-1}, whereas the 16.3 eV peak, which corresponds to transitions at the Γ point, is seen to require a very fine grid spacing on the order of $\Delta k \lesssim 0.1$ nm^{-1}. This can be attributed to the high symmetry nature of this Γ point, which makes it more sensitive to tiny changes in the spacing of the k-point mesh. For even coarse grid spacings, the spectra for Pn is seen to converge with no substantial changes for denser k-point meshes for any region of the energy spectrum.

5.4.5 Implementing Low-Dimensional Response Functions

Beyond the computational point of view, another factor to consider when applying LCAO-TDDFT-k-ω to lower dimensional materials is the accuracy of the mean field response functions as outlined in Section 2.2.5 for describing experimental spectra, where materials do not precisely fall into the theoretical categories of 2D or 1D materials. Such issues can already be seen in the

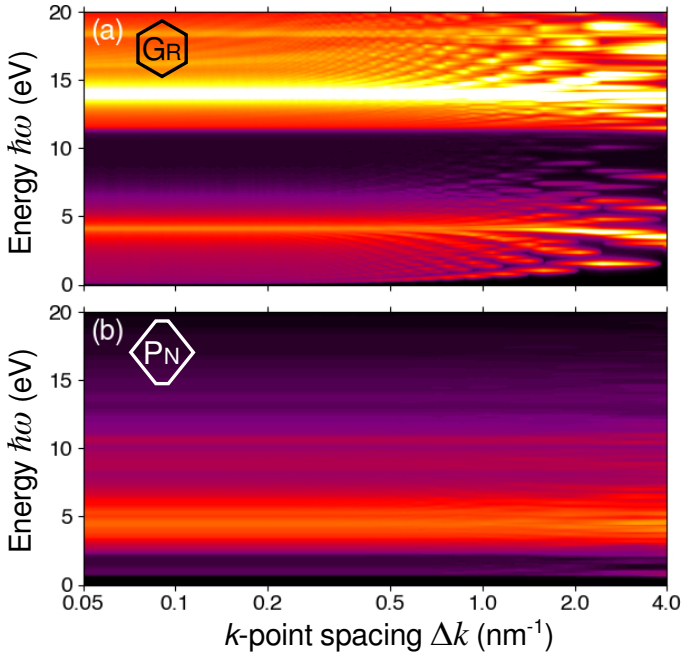

Figure 5.12: Convergence of the spectra of the real in-plane conductivity $\mathrm{Re}[\sigma(\hat{\mathbf{q}}_{\parallel}, \omega)]$ in units of eV with k-point spacing Δk in units of nm^{-1} for both (a) graphene (GR) and (b) phosphorene (PN)

difference in approaches for energy loss spectra for 3D (Section 2.3.2) and 2D (Section 2.3.3) materials. Essentially, this amounts to whether a monolayer in a phenomenological model should be treated as a thin 3D layer or as a 2D boundary condition that utilizes the 2D in-plane conductivity.

Fig. 5.13 shows experimental reflection spectra for x and y polarized light $\Delta R/R$ for PN laid on a sapphire (Al_2O_3) substrate with insulating hexagonal boron nitride (hBN) flakes of width 15 nm laid on top [127]. These are compared to two separate phenomenological models both using the dielectric function as in Eq. 2.59. The first treats monolayer phosphorene as part of a 3D thin layer air-hBN-PN-Al_2O_3 system with $\varepsilon_{3D}(\omega)$ included as input. The second model uses the 2D in-plane conductivity calculated using Eq. 2.69 in 2D as input into a system of Fresnel's equations that use the transfer matrix formalism [160] to account for the multilayer that light needs to traverse and reflect off.

The two-dimensional approach gives a more "peak-like" profile compared to the 3D model for x-polarized light, which better resembles the experimental spectra. However, the 3D model better predicts the reflection ratio for energies beyond the 1.8 eV peak. The overall trend for the reflection ratio is correct for y-polarized light in both models but highly underestimated relative to experiment. While it is clear that neither model provides an accurate enough description of the reflection of monolayer phosphorene, the 3D model requires as input the thickness of the phosphorene layer, to which the reflection ratio is directly proportional. We use 0.5 nm for

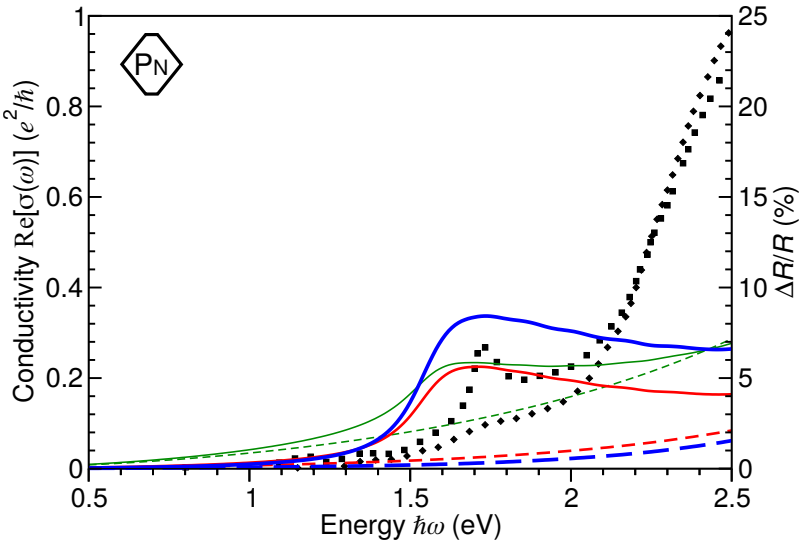

Figure 5.13: Real part of the conductivity $\sigma(\omega)$ calculated in LCAO-TDDFT-k-ω in units of $e^2/\hbar$ (blue thick lines) for 2D (red lines) and 3D (green thin lines) models alongside measurements (symbols) [127] for light polarized in the x (solid, squares) and y (dashed, diamonds) directions.

the thickness based on the spacing between layers in multilayer phosphorene [127]. This is in contrast to the 2D approach, which makes no empirical assumptions about the thickness of the monolayer. This means the phenomenological model manages to describe the experimental spectra with fewer free parameters. This suggests that the use of mean-field optical response functions for lower-dimensional materials may be more suitable, at least for phenomenological purposes.

5.5 3D Anatase and Rutile TiO_2

Although we have mentioned that many of the memory and time savings involved in using the LCAO-TDDFT-k-ω method make it most suitable for lower-dimensional materials, especially large macromolecular systems, it remains important to ensure that the code manages to describe systems which are periodic in every direction. As test cases we consider two photocatalytic materials: anatase (A-TiO_2) and rutile (R-TiO_2) titania. These materials are ideal candidates considering their technological relevance in photovoltaic applications [161] and the challenges both computationally and theoretically in modeling their optical spectra. This includes the role that excitonic effects are expected to play and the effect of unoccupied d levels on the band gap and absorbance [105, 100].

Figs. 5.14 and 5.15 show the dielectric function components parallel ($\varepsilon_{\parallel}(\omega)$) and perpendicular ($\varepsilon_{\perp}(\omega)$) to the tetragonal c-axis for A-TiO_2 and R-TiO_2, respectively. Both these figures compare experimental reflectometry [162, 165] or ellipsometry [166] measurements with G_0W_0-

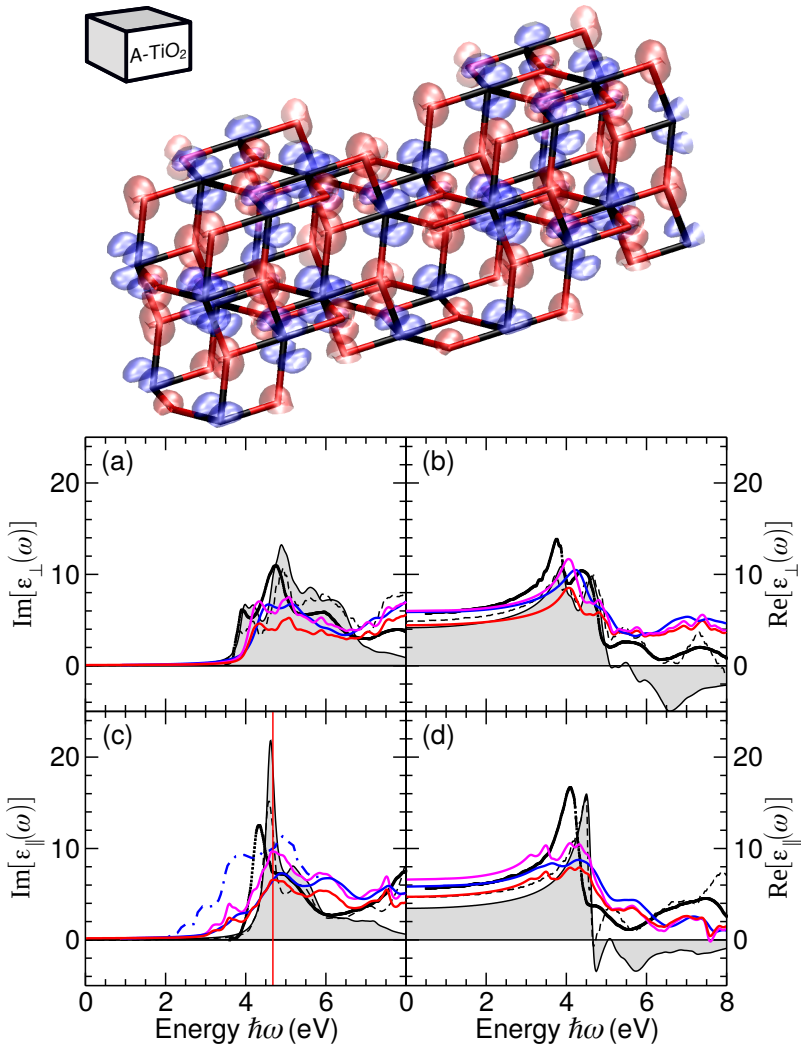

Figure 5.14: Anatase (A-TiO_2) (a,c) imaginary and (b,d) real parts of the dielectric function (a,b) perpendicular ($\perp$) and (c,d) parallel ($\parallel$) to the tetragonal c-axis versus energy $\hbar\omega$ in units of eV derived from reflectometry (black square) [162], G_0W_0-BSE (thin solid filled [100] and dashed [163] black lines), PW-TDDFT-k-ω (dash-dotted [164] and solid blue lines), LCAO-TDDFT-k-ω *a priori* (red solid lines), and *a posteriori* (magenta solid lines) calculations and positive (red) and negative (blue) isosurfaces of the electron-hole spectral density difference for the bright exciton at $\hbar\omega_{ex} \approx 4.8$ eV (red vertical line in (c)) with O and Ti atoms colored red and grey, respectively.

BSE, PW-TDDFT-k-ω, and LCAO-TDDFT-k-ω with both *a priori* and *a posteriori* scissors calculations for the dielectric function. The former two TDDFT implementations are included as excitonic and crystal local field effects are expected to be important for determining an accurate picture for A-TiO_2 and R-TiO_2's optoelectronic response functions, noting that LCAO-TDDFT-k-ω does not include either of these effects in its implementation. However, we expect [164] that DFT should be sufficient to describe the ground state properties of the system, including TiO_2's

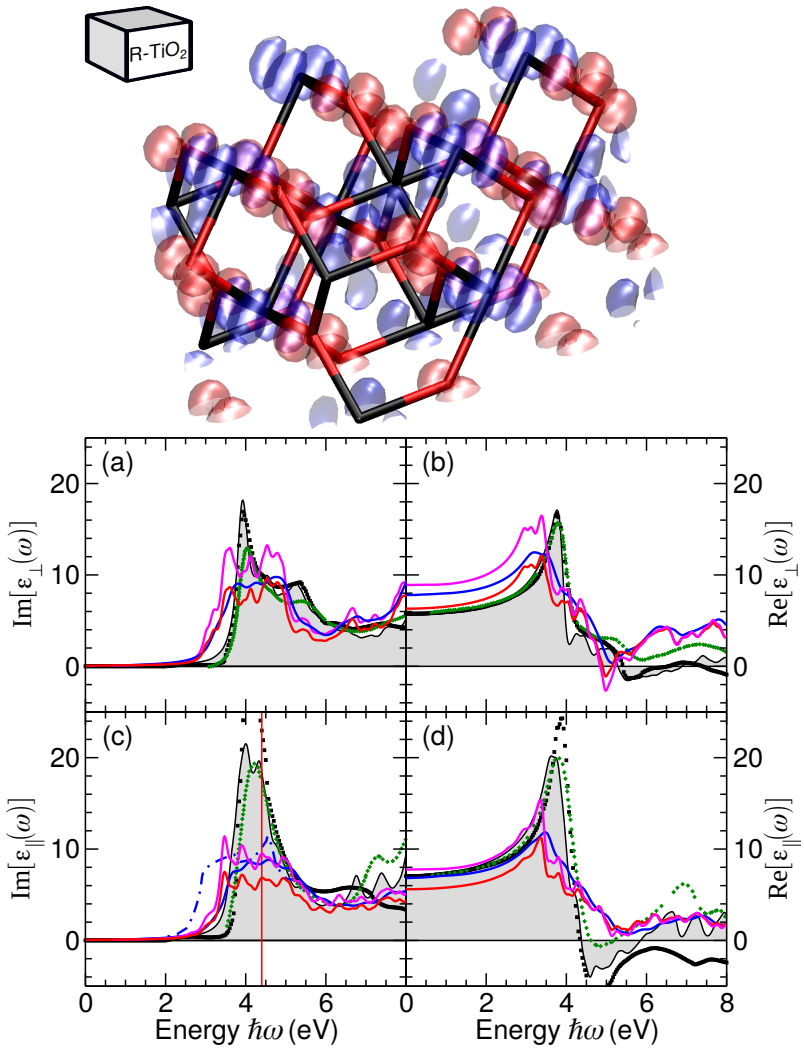

Figure 5.15: Rutile (R-TiO_2) (a,c) imaginary and (b,d) real parts of the dielectric function $\varepsilon(\omega)$ (a,b) perpendicular ($\perp$) and (c,d) parallel ($\parallel$) to the tetragonal c-axis versus energy $\hbar\omega$ in units of eV from reflectometry (black squares) [165], ellipsometry (green diamonds) [166], G_0W_0-BSE (thin solid filled lines) [105], PW-TDDFT-k-ω (dash-dotted [164] and solid blue lines), LCAO-TDDFT-k-ω *a priori* (red solid lines), and *a posteriori* (magenta solid lines) and positive (red) and negative (blue) isosurfaces of the electron-hole spectral density difference for $\hbar\omega_{ex} \approx 4.4$ eV (red vertical line in (c)) with O and Ti atoms colored red and grey, respectively.

atomic structure.

The PW-TDDFT-k-ω and LCAO-TDDFT-k-ω curves in both Fig. 5.14 and Fig. 5.15 show a strong step-like behaviour at the band gap resulting from the strong overlap between valence and conduction band wavefunctions for these titania crystals. This effect is visible as well at the G_0W_0-BSE level, where excitonic binding is relatively weak ($\sim$ 0.1 eV) in both A-TiO_2 [100]

Table 5.4: Energy of the first bright exciton $\hbar\omega_{ex}$ in eV for anatase (A-TiO_2) and rutile (R-TiO_2).

Method	$\hbar\omega_{ex}$ (eV)	
	A-TiO_2	R-TiO_2
Reflectometry	4.31[a]	4.09[b]
Ellipsometry	—	4.22[c]
G_0W_0-BSE	4.63[d]	4.01[e]
G_0W_0-BSE	4.61[f]	4.12[f]
PW-TDDFT-k-ω	4.93[g]	4.55[g]
PW-TDDFT-k-ω	4.90[h]	4.56[h]
LCAO-TDDFT-k-ω	4.81[h]	3.48[h]

[a] Ref. [162] [b] Ref. [165] [c] Ref. [166]
[d] Ref. [100] [e] Ref. [105] [f] Ref. [163]
[g] Ref. [164] [h] This work.

and R-TiO_2 [105], yielding relatively consistent peak energies with the less computationally intensive LCAO-TDDFT-k-ω and PW-TDDFT-k-ω approaches, as shown in Table 5.4. These two methods qualitatively match the peak and trough behaviour of the experimental data from both ellipsometry and reflectometry, but ultimately calculations at the G_0W_0-BSE level are needed to capture the more pronounced peaks visible in experimental measurements [162, 165, 166]. For both A-TiO_2 and R-TiO_2 the dielectric function calculated in the LCAO-TDDFT-k-ω *a posteriori* scissors approach is generally more intense than both PW-TDDFT-k-ω and LCAO-TDDFT-k-ω *a priori* scissors methods, the former two methods yielding the best results for both A-TiO_2 and R-TiO_2 relative to experiment and G_0W_0-BSE.

The electron hole spectral density difference is shown in the upper inset for both Figs. 5.14 and 5.15 at the bright excitonic peak marked by red dashed lines in the plots of $\mathrm{Im}[\varepsilon(\hat{\mathbf{q}}_{\parallel}, \omega)]$. As with graphene, the symmetries of the crystal manifest as an isomorphism between $\Delta\rho(\mathbf{r}, \hat{\mathbf{q}}_y, \omega_{ex})$ and $\Delta\rho(\mathbf{r}, \hat{\mathbf{q}}_x, \omega_{ex})$, the spectral density difference in two separate polarization directions. Both figures show that the electron spectral density is localized to the d-orbitals of the titanium atoms, all pointing in the same direction in the case of A-TiO_2 but pointing in separate directions depending on the atom site for R-TiO_2. The hole spectral density is centred on the oxygen atoms, with a mixture of s and p-orbital character for A-TiO_2 and $2p$-character for R-TiO_2. These isosurface plots are consistent with G_0W_0 results [100] regarding the makeup of orbitals involved in the bright exciton transition.

Chapter 6

Understanding the Energy Loss Profile of Monolayer Graphene

Getting a better understanding of the electrical and optical properties of graphene has been the focus of many theoreticians and experimentalists over the last decade, and many strategies have been put forward towards gaining further insight into this material. While most of the focus has been on the THz regime, EELS (Section 2.3) has proven to be one effective method for targeting samples of single- and multi-layer graphene [167, 168] beyond that energy range. The tuning of the collection aperture for scattered electrons allows for a resolution in the features of the optical spectra, with the $\pi - \pi^*$ and $\sigma - \sigma^*$ interband transitions seen in Fig. 5.8 appearing as massive peaks in the spectra at any resolution, while features like the universal conductivity or Drude contributions will only be visible with finely tuned STEM devices that utilize aberration correctors and monochromators to probe graphene [169, 170, 171, 65]. Given the interest in the material, the ways it can be tuned, the different energy regimes under consideration, and effects of experimental setup, it is important to consider from a phenomenological point of view the interplay between these parameters and their effect on the experimental EELS spectra.

Section 3.1.1 in conjunction with Appendix B has presented phenomenological models at different levels of approximation for both neutral and doped graphene, while *ab initio* TDDFT packages like LCAO-TDDFT-k-ω and PW-TDDFT-k-ω are able to give computationally accurate models for our 3D dielectric function from which we can extract the in-plane conductivity via Eq. 2.69. Section 2.3.3 shows that the conductivity of the monolayer can be used directly as input to determine the EELS spectra, with both non-relativistic and relativistic models available. Changes in the doping of graphene, the size of the collection aperture, the incoming velocity of incident electrons, and broadening effects are all controllable parameters within our model, and we present below a thorough analysis of the influence these have on the EELS spectra and as a result what the spectra can tell us about systems that incorporate graphene monolayers.

To this effect, it helps to first investigate neutral graphene over the broad energy range accessible with EELS, checking the performance of the phenomenological model against *ab initio*

calculations and comparing the theoretical energy loss output to experimental spectra subject to the above mentioned parameter changes. Afterwards, the inclusion of doping into the graphene monolayer is explored; as the features of a doped monolayer like Pauli blocking and intraband transitions will only be apparent in the THz and infrared energy ranges, these are the regions where the changes in experimental setup parameters are explored most thoroughly.

6.1 Neutral Single Layer Graphene

Section 3.1.1 explored phenomenological conductivity models for both neutral and doped graphene, the former in Eq. 3.9 accounting for the universal conductivity of graphene from the Dirac cone approximation and two spectral peaks with hydrodynamic features. The figures and discussion that follow explore how this model compares against results computed within TDDFT, the features in the output energy loss spectra, and how changing values like the collection angle of the EELS setup yield significant changes in the location and intensity of energy loss spectral peaks.

Fig. 6.1(a,b) show the real and imaginary parts of the optical conductivity versus the energy loss $\hbar\omega$, obtained via PW-TDDFT-k-ω from the QUANTUM ESPRESSO simulation package with the PBE xc functional, a 600 eV energy cutoff, and a $601 \times 601 \times 1$ Monkhorst-Pack k-point mesh, and using the eHD model presented in Eq. 3.9, which accounts for the two interband transition peaks as well as the universal conductivity of graphene at energy values below a certain cutoff. The phenomenological parameters in this equation are chosen to yield the best overall fit with the *ab initio* spectra, giving $n_\sigma^0 \approx 115\,\mathrm{nm}^{-2}$ and $n_\pi^0 \approx 38\,\mathrm{nm}^{-2}$, $\omega_{\sigma r} = 14.15$ eV, $\omega_{\pi r} = 4.19$ eV, $\gamma_\sigma = 2.18$ eV, $\gamma_\pi = 2.04$ eV, and $\omega_c = 3.54$ eV (corresponding to $f \approx 0.785$). With these parameters chosen, we see very good agreement between the computational and theoretical approaches for the conductivity, with both giving $\mathrm{Re}\,\sigma(\omega \to 0) = e^2/4\hbar$, negligible conductivity in the region above $\hbar\omega \sim 30$ eV, and quantitatively similar peak heights and widths for the real conductivity spectra at two major interband transitions. The imaginary conductivity shows substantial jumps from negative to positive values at both peaks before decreasing towards zero beyond $\hbar\omega \sim 20$ eV.

In addition, Fig. 6.1(c) plots $N_e(\omega)$, derived in Eq. 3.8, versus the energy loss, showing in both approaches a linear increase in the number of electrons participating in the conductivity, as expected from the Dirac cone approximation, followed by substantial jumps up to $N_e = 1$ beyond the $\pi - \pi^*$ interband transition and a jump up to approximately $N_e = 4$ beyond the peak associated with the $\sigma - \sigma^*$ transition, as expected from the number of valence electrons in each of these types of orbitals. Working without any broadening parameters in the model would yield step-like curves in Fig. 6.1(c).

Fig. 6.2 presents energy loss spectra from three different experimental setups, with incident electron energies normal to the free-standing graphene monolayer of 40, 60, and 100 keV and aperture sizes of $q_c = 0.1\text{Å}^{-1}, 4.3\text{Å}^{-1}$, and 3.2Å^{-1}, respectively. These experimental spectra can be predicted through the use of the non-relativistic probability density from Eq. 2.99 using

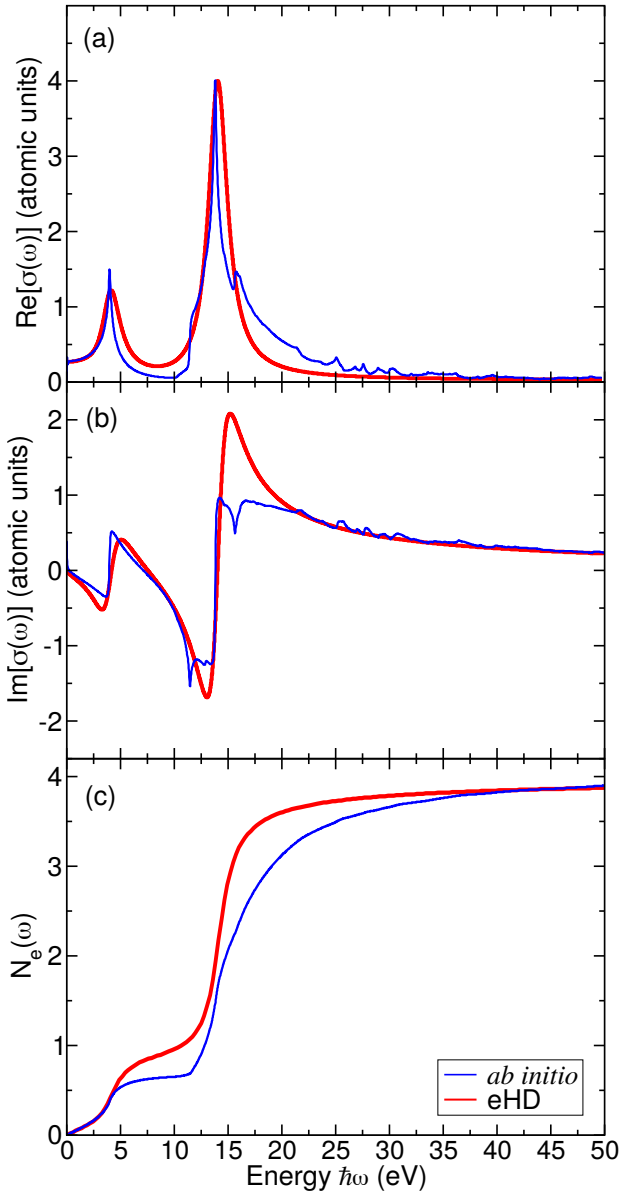

Figure 6.1: (a) Real and (b) imaginary parts of the optical conductivity in atomic units and (c) the number of valence electrons $N_e(\omega)$ participating in excitations up to ω versus the energy loss $\hbar\omega$ in units of eV, calculated within PW-TDDFT-k-ω (blue) and using the eHD model (red).

$\sigma_{2D}(\omega)$ calculated using either PW-TDDFT-k-ω or with the eHD model as shown in Fig. 6.1, noting that the output of these models change based on the parameters $v_\perp$ and q_c, which are different in each experimental setup. The use of a purely optical conductivity has previously been justified in Section 2.3.3.

Fig. 6.2(a) shows that the $\pi - \pi^*$ peak is shifted to the left by 1 eV in the experimental spectra relative to the eHD and *ab initio* models while the location for the $\sigma - \sigma^*$ peak matches

closely. The ratio between the two peaks using PW-TDDFT-k-ω is much larger than the ratio found from the eHD model and in the experimental spectra, the latter two in close agreement. We note that experimental data is not shown in Fig. 6.2(a) for energies below 1.5 eV due to

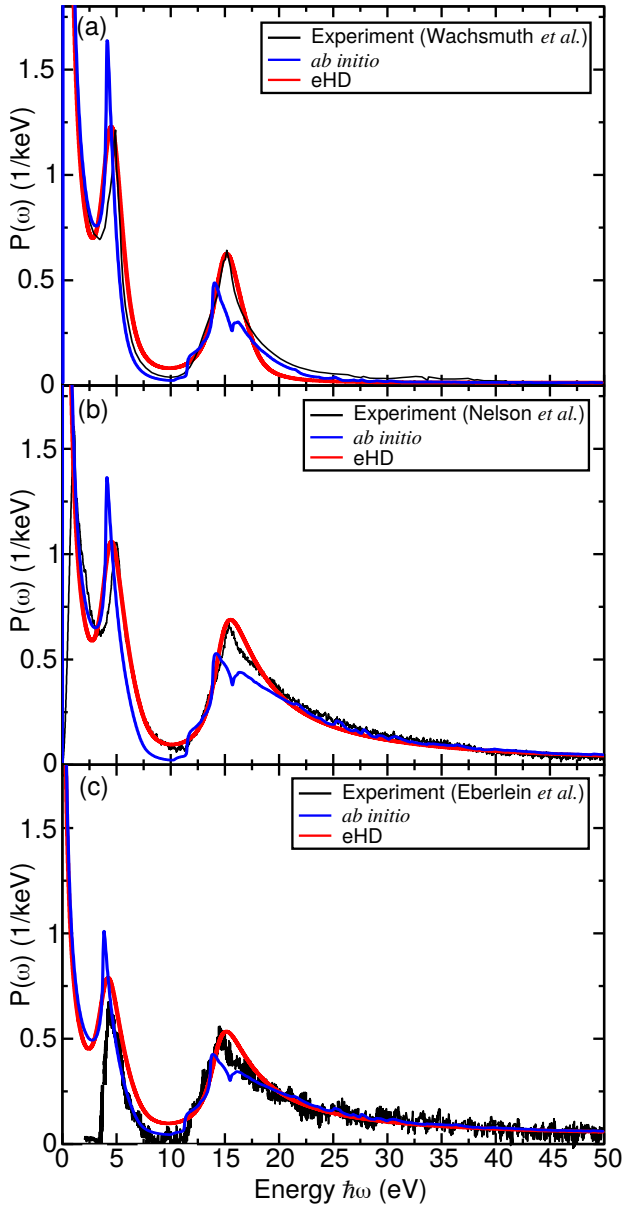

Figure 6.2: Experimental EEL spectra in the $q \to 0$ limit for (a) a 40 keV incident electron with $q_c = 0.1\text{Å}^{-1}$ [12], (b) a 60 keV incident electron with $q_c = 4.3\text{Å}^{-1}$ [62], and (c) a 100 keV incident electron with $q_c = 3.2\text{Å}^{-1}$ [60] alongside probability densities $P(\omega)$ in units of 1/keV versus the energy loss $\hbar\omega$ using the planar conductivity $\sigma_{2D}(\omega)$ calculated using PW-TDDFT-k-ω (blue) and with the eHD model (red) as input.

the removal of the large zero-loss peak (see Section 2.3) for low energies and low momentum transfers. The subtraction of the ZLP may justify some of the disagreement between the eHD model and experimental spectra for energies in the $5 \leq \hbar\omega \leq 12$ eV range, although this region is also better predicted by the probability density using the LCAO-TDDFT-k-ω approach. The quick drop-off in the theoretical spectra versus the longer tail seen in experiment may be related to the small collection aperture or to the excitation of out-of-plane modes not considered in the planar 2D conductivity model.

Fig. 6.2(b) shows an experimental setup with much larger collection aperture and similar electron velocity, with the $\pi-\pi^*$ and $\sigma-\sigma^*$ peaks now shifted to the right by 1 eV in the experimental spectra relative to the eHD and *ab initio* models. These small differences can be attributed to the exclusion of excitonic effects and screening in our application of the PW-TDDFT-k-ω method, although using higher-order xc functionals (see Section 2.1.2) provides one way to include these effects. The behaviour of the peak ratio follows Fig. 6.2(a), although the agreement between the eHD model and experimental spectra for energies beyond $\hbar\omega \sim 20$ eV is far better. The sharpness of the peaks using the *ab initio* conductivity may be the result of performing calculations at very low temperature or not including sufficient broadening in the density-density response function elements, as seen in Eq. 2.42. Notable in the experimental spectra in this figure is the cubic interpolation of the subtracted ZLP tail, ensuring that the energy loss goes to zero at zero frequency, although this subtraction as previously mentioned has important consequences for the intensities and therefore the ratio between the two interband transition peaks.

The patterns seen in Fig. 6.2(c) closely follow those of Fig. 6.2(b), although the ZLP subtraction is extended to all energies below 3.5 eV, substantially diminishing the size of the $\pi - \pi^*$ peak, emphasizing further the importance of properly removing from the spectra the effect of these non-interacting electrons. As in Fig. 6.2(a) the experimental spectra for energies in the $5 \leq \hbar\omega \leq 12$ eV range match with *ab initio* results better than the eHD model but the ratio from PW-TDDFT-k-ω is again too high relative to experiment. Once ZLP subtraction is accounted for, it is clear that the eHD model for neutral graphene provides an excellent description of the intensities and peaks of EELS spectra over a wide range of energy loss values.

Fig. 6.3 shows both how the change in the aperture collection and the use of either the non-relativistic energy loss in Eq. 2.99 or the relativistic energy loss in Eq. 2.105 affect the shape and intensity of the probability density of energy loss $P(\omega)$ using the eHD model $\sigma(\omega)$ as input. No matter the value of the collection wavenumber, the relativistic energy loss consistently exceeds that given by the non-relativistic model by around 10% for all energies, consistent with previous evaluations of the relativistic model [21], along with a slight red-shift in the peak positions. A gradual rightward shift in the peak energies accompanies an increase in q_c, with a nearly 1 eV shift between spectra collected for q_c between 0.006 Å^{-1} and 0.096 Å^{-1}. Increased q_c also tends to gives broader peaks, resulting in relatively longer tails. Most notably, the ratio of the heights of the $\pi - \pi^*$ and $\sigma - \sigma^*$ peaks shrinks with increasing aperture size as transfers resulting from higher wavenumbers begin to have an influence on the intensity distribution, an effect which is also seen experimentally [172, 75]. The spectra shown in Fig. 6.2 all pertain to experiments with very large collection apertures, the lowest on the order of the largest wavenumber cutoff used

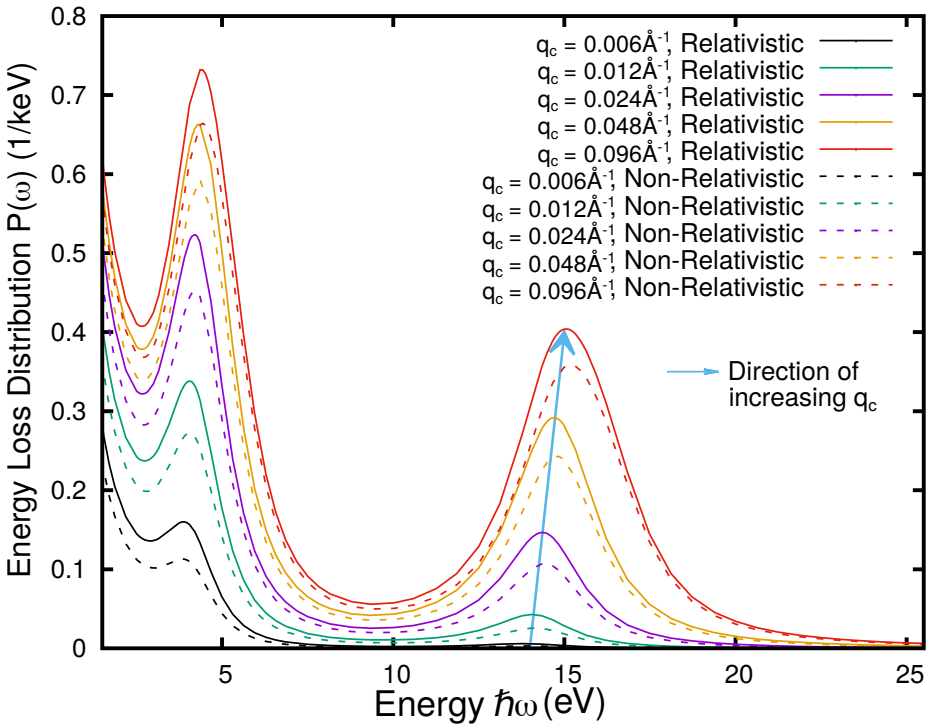

Figure 6.3: Probability density of energy loss $P(\omega)$ from the eHD model using relativistic (solid lines) and non-relativistic energy loss (dashed lines)expressions for q_c ranging from 0.006 Å^{-1} to 0.096 Å^{-1}, where q_c is the maximum in-plane scattering momentum.

in Fig. 6.3, so the three effects mentioned here are not expected to have a significant influence on the probability densities plotted in that figure and would instead be most relevant for higher resolution EELS experimental setups, as shown in Fig. 6.4.

Fig. 6.4(a,b) shows experimental EELS data [75] for monolayer graphene with an incident electron energy of 120 keV while using an EEL setup with a k-resolution of $\sim$ 0.001 Å^{-1}. As the spectra are claimed to have been integrated over the 0 – 0.012 Å^{-1} energy range, we can contrast the spectra with the probability energy loss density as in Fig. 6.2, this time considering both eHD and *ab initio* models using both the non-relativistic and relativistic energy loss implementations. This data also has the advantage of not using the ZLP subtraction method, so our phenomenological models can directly compare against raw data, especially important since the ZLP shape changes substantially with the relatively low energy resolution afforded by the experiment, a direct consequence of the high wavenumber resolution [107]. As the experimental data is given in arbitrary units, it is resized along the vertical direction to provide the best fit with our models. Plots are given for aperture sizes of both $q_c = 0.012\,\text{Å}^{-1}$ and $q_c = 0.024\text{Å}^{-1}$ as the aperture collection angle for this particular EELS experimental setup may be larger than quoted, depending on the positioning of the probe and the alignment of the graphene BZ relative to the aperture. Given the energy resolution, Figure 6.4(c,d) shows the same curves as in Figure 6.4(a,b) but with an applied Gaussian smoothing with a standard deviation of $\hbar\Delta\omega = 0.6$ eV

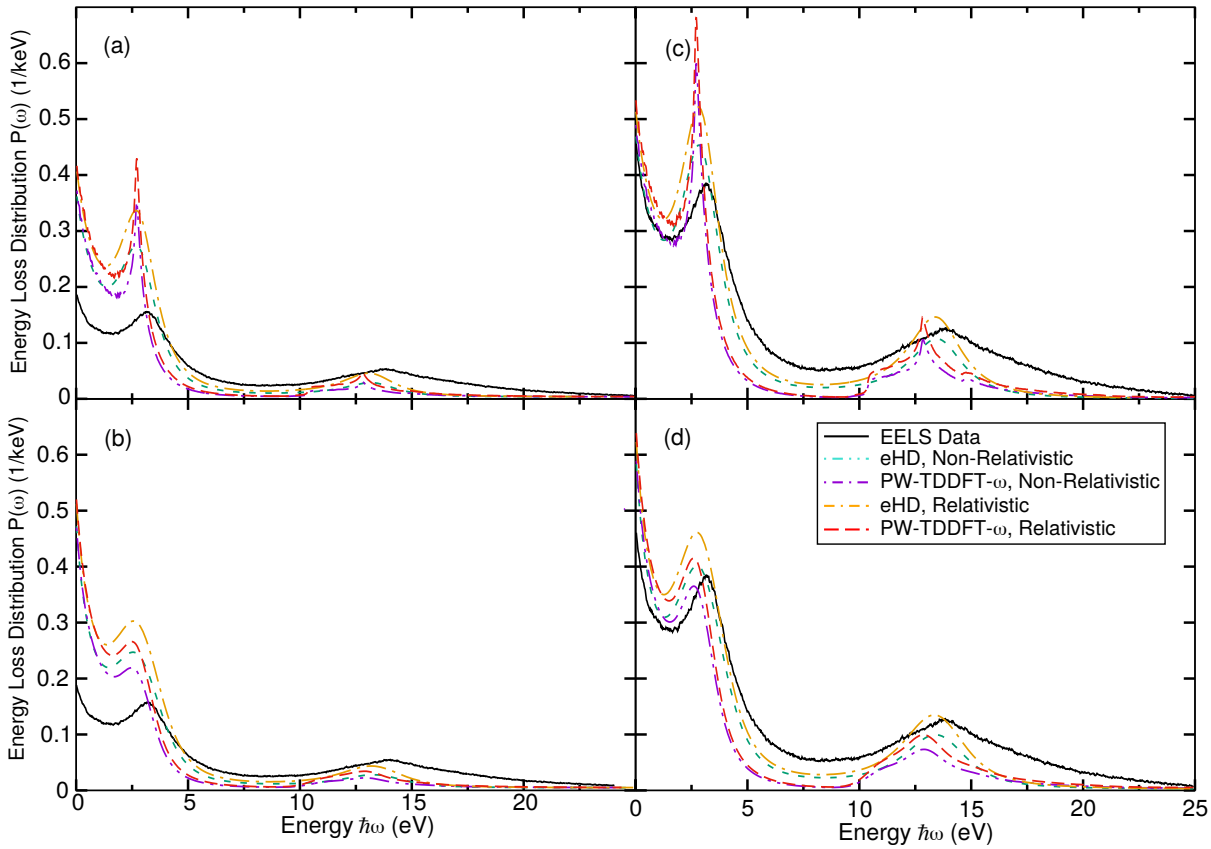

Figure 6.4: Experimental EEL spectra [75] plotted alongside the probability density $P(\omega)$ calculated using the non-relativistic energy loss model with eHD (cyan, dot-dashed) and PW-TDDFT-k-ω (purple, dot-dashed) in-plane conductivities as well as the relativistic energy loss model with eHD (orange, dot-dashed) and PW-TDDFT-k-ω (red, dashed) in-plane conductivities for collection apertures of either (a,c) $q_c = 0.012\,\text{Å}^{-1}$ or (b,d) $q_c = 0.024\text{Å}^{-1}$, plotted with either (a,b) no Gaussian smoothing or (c,d) a Gaussian smoothing with standard deviation equal to the EELS energy resolution [75] of $\hbar\Delta\omega = 0.6$ eV.

equal to the experimental energy broadening.

While the overall shape of the theoretical and experimental curves follow the same general pattern, with peaks near 4 and 15 eV and tail-like behaviour for $\hbar\omega \geq 20$, the experimental data differs in three important ways. The $\sigma - \sigma^*$ peak has an experimentally much longer tail that would correspond to interband transitions unaccounted for by PW-TDDFT-k-ω. Peak values are also shifted rightward by approximately 1 eV relative to theory, and the ratio between the $\pi - \pi^*$ and $\sigma - \sigma^*$ peaks is significantly lower than what is expected theoretically from such a small aperture, no matter the type of energy loss equation used or the input for the in-plane conductivity into those models. The experimental data fits best with the model curves shown in Fig. 6.4(d), where the maximum wavenumber collected by the aperture is double what is quoted and the Gaussian smoothing is applied to imitate the effects of the low energy resolution of the

experimental setup, again seeing that the four models provide roughly equivalent approximations to the spectra. This matches with expectations based on Fig. 6.3, where higher values of q_c directly effect the long-tail behaviour of the observed EEL spectra.

6.2 Doped Single Layer Graphene

While in Section 6.1 we discussed modeling of four independent sets of experimental data for EELS at high energies, in this section we focus on making theoretical predictions for the low-energy region of those spectra for doped graphene, which should be accessible with tools like the modern monochromated Transmission Electron Microscope.

6.2.1 Modeling the Optical Conductivity

Calculations for doped graphene at $\varepsilon_F = 0.1$ eV and $\varepsilon_F = 0.5$ eV are done using both LCAO-TDDFT-k-ω with a converged DZP basis set and PW-TDDFT-k-ω with a converged energy cutoff of $E_{cut} \approx 340$ eV, the doping accomplished by adding excess charges to the unit cells of monolayer graphene. A dense Monkhorst-Pack $301 \times 301 \times 1$ k-point mesh is used, and the electronic temperature is set to $k_B T \approx 1$ meV to ensure integer occupations for all electronic levels.

As we are considering a doped system, the interband conductivity from Eq. 2.70 is calculated for both the LCAO and PW TDDFT implementations, while the intraband conductivity in Eq. 2.71 is calculated solely using LCAO-TDDFT-k-ω, noting that for graphene the intraband conductivity for relatively low doping is expected to follow the equation

$$\sigma^{\parallel}_{\text{2D, intra}}(\omega) = \frac{iD}{\pi(\omega + i\gamma_D)} \tag{6.1}$$

where D is defined as the Drude weight [173], which should theoretically tend to $e^2\varepsilon_F/\hbar^2$ within the RPA approximation. As the matrix elements in Eq. 2.71 are already calculated within the LCAO-TDDFT-k-ω, what remains is the computation of $\frac{\partial f_{n\mathbf{k}}}{\partial \varepsilon_{n\mathbf{k}}}$. If the code is run in the zero temperature limit, this quantity approaches $-\delta(\varepsilon_F - \varepsilon_{n\mathbf{k}})$, which is far too sensitive to the size and resolution of the employed k-grid. Instead, a fictitious electronic temperature of 1 meV $\approx$ 11.6 K is used to smear out the Fermi surface to increase the number of k-points that contribute to the summation in Eq. 2.71. After this temperature is introduced, approximately 10000 k-points are chosen randomly with the condition that $|\varepsilon_{n\mathbf{k}} - \varepsilon_F| < 0.1$ eV, which for graphene at relatively low doping corresponds to an annulus around the K high-symmetry point, and a fixed density calculation is rerun in the LCAO-TDDFT-k-ω regime to give the dipole elements for the randomized k-mesh. A similar technique has previously been shown to give intraband conductivities for metals in agreement with theoretical expectations [174, 175]. Ultimately the value computed acts as a verification that the Drude weight D in Eq. 6.1 gives an intraband conductivity in accordance with *ab initio* results. Calculations of the intraband conductivity

yield Drude weights D in accordance with values of $\varepsilon_F = 0.51$ eV and $\varepsilon_F = 0.104$ eV for electron concentrations of 1.84×10^{13} cm^{-2} and 7.34×10^{11} cm^{-2} in each graphene unit cell, electron densities theoretically corresponding to $\varepsilon_F = 0.5$ eV and $\varepsilon_F = 0.1$ eV, respectively.

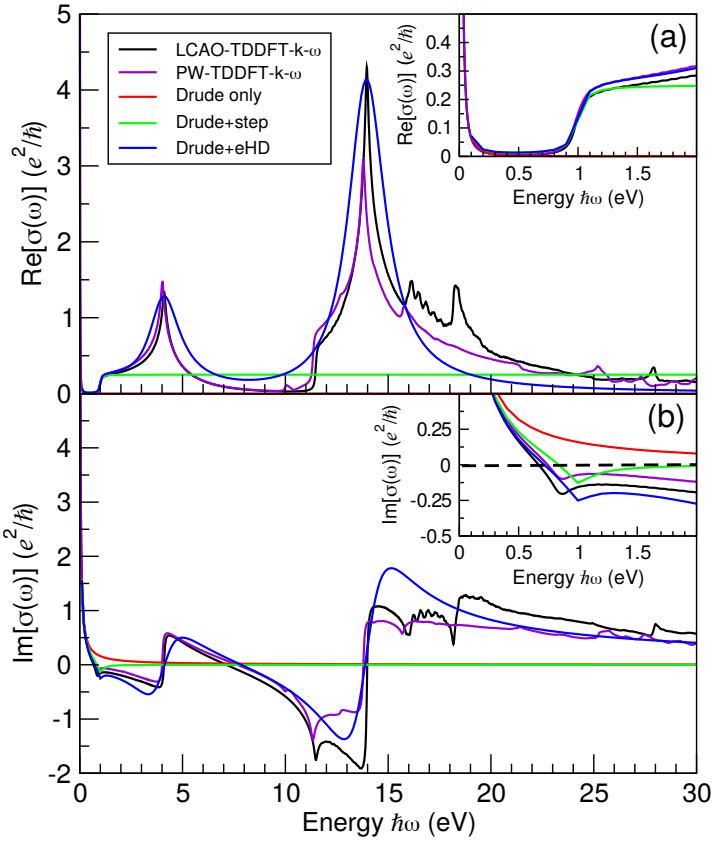

Figure 6.5: (a) Real and (b) imaginary parts of the optical in-plane conductivity $\sigma(\omega)$ in units of $e^2/\hbar$ versus energy $\hbar\omega$ in units of eV for monolayer doped graphene with a Fermi energy of $\varepsilon_F = 0.5$ eV calculated using LCAO-TDDFT-k-ω (black) and PW-TDDFT-k-ω (purple) *ab initio* methods and the Drude only (red), Drude+step (green), and Drude+eHD (blue) conductivity models, with the components of the planar conductivity in the $\hbar\omega \leq 2$ eV energy range shown as insets.

Fig. 6.5 shows the real and imaginary parts of the conductivities obtained from LCAO-TDDFT-k-ω and PW-TDDFT-k-ω calculations and for three different empirical models of the optical conductivity, "Drude only" defined in Eq. 3.5 and the "Drude+step" and "Drude+eHD" models defined in Section 3.1.1, each valid approximations to the optical conductivity at different levels of application. Both the LCAO-TDDFT-k-ω and PW-TDDFT-k-ω results for the conductivity, previously shown in Fig. 5.6 and calculated within GPAW, yield spectra in quantitative agreement with those from QUANTUM ESPRESSO used in Fig. 6.1 [58], although in order to give a Drude+eHD model which matches with the LCAO-TDDFT-k-ω method we use to calculate the intraband conductivity, the parameters in the model in Eq. 3.9 are adjusted to $n_\sigma^0 \approx 118\,\mathrm{nm}^{-2}$,

$n_\pi^0 \approx 35\,\mathrm{nm}^{-2}$, $\omega_{\sigma r} = 13.95$ eV, $\omega_{\pi r} = 4.12$ eV, $\gamma_\sigma = 2.18$ eV, $\gamma_\pi = 1.80$ eV. $\omega_c = 3.54$ eV, corresponding to $f \approx 0.785$, is again chosen to be the cut-off frequency since neither the total number of participating valence electrons nor the relative weights of interband transitions should change substantially with doping according to Eq. 3.8. A value of $\gamma_D = 0.004$ eV in Eq. 2.71 is chosen as the broadening value for this intraband contribution, which is on the order of the Fermi surface temperature smearing used in the LCAO-TDDFT-k-ω intraband conductivity calculation. Lastly, the smoothing parameter for the Pauli blocking in Eq. 3.12 is chosen to be $\gamma_P = \eta = 0.05$ eV, equivalent to the value η_{inter} used in Eq. 2.70.

Overall, the Drude+eHD model shows good agreement between both TDDFT calculations for the optical conductivity. Unique to doped graphene, the slightly broadened step located at $2\varepsilon_\mathrm{F} = 1$ eV is properly captured by the Drude+eHD and Drude+step and TDDFT models. These four models also find close agreement in the IR region of frequencies for the real part of the conductivity, accessible with more modern STEM devices, an area hollowed out by the Pauli blocking step at $2\varepsilon_\mathrm{F}$ with the conductivity only spiking at very low energies thanks to the intraband conductivity given in Eq. 6.1. The location of the step and the intensity of the intraband conductivity are both controlled by the level of doping in the graphene monolayer. The difference between the two TDDFT and three empirical approaches is more apparent in the inset to the imaginary part of the conductivity, where the Drude+step provides worse agreement with TDDFT results than the Drude+eHD. This can be explained through the KK relations in Eq. 2.76 which evaluate the real conductivity over the full energy range to determine the imaginary conductivity component, even for energies within the Pauli blocking gap.

6.2.2 Modeling Energy Loss Densities

Fig. 6.6 shows how the three empirical models along with the LCAO-TDDFT-k-ω method from Fig. 6.5 yield different energy loss densities for the Ohmic part, given by Eq. 2.106, and the radiative part, given by Eq. 2.108. As mentioned earlier, the three empirical models are each valid approximations to the optical graphene conductivity in different energy regimes, which can be directly visualized in Fig. 6.6. The Drude only model is applicable for $\hbar\omega \lesssim \varepsilon_\mathrm{F}/2$, the Drude+step model providing good agreement for energy losses below 2 eV where the π–π^* transition starts to appear, and finally the Drude+eHD model giving a phenomenological treatment for all relevant excitations in the $\hbar\omega \lesssim 20$ eV range. The choice to use the relativistic energy loss equations for doped graphene is motivated by the expected role that the radiative losses will play for very low energies when intraband transitions are active, more so than the simple intensity shift seen for neutral graphene in Fig. 6.3. The energy loss densities are also plotted on the log-log scale to better visualize the region of energy losses below ~2 eV where the comparison between Ohmic and radiative losses are most prominent, and to visualize the relative magnitude of the main spectral features. All energy loss densities are evaluated assuming a maximal collection aperture $q_c \to \infty$ with an incident electron energy of $E_0 = 100$ keV for either $\varepsilon_\mathrm{F} = 0.5$ eV or 0.1 eV.

Fig. 6.6 shows that for energy losses beyond 2 eV there is no perceptible change between doping levels of 0.1 and 0.5 eV. Since this is the range traditionally blotted out by traditional

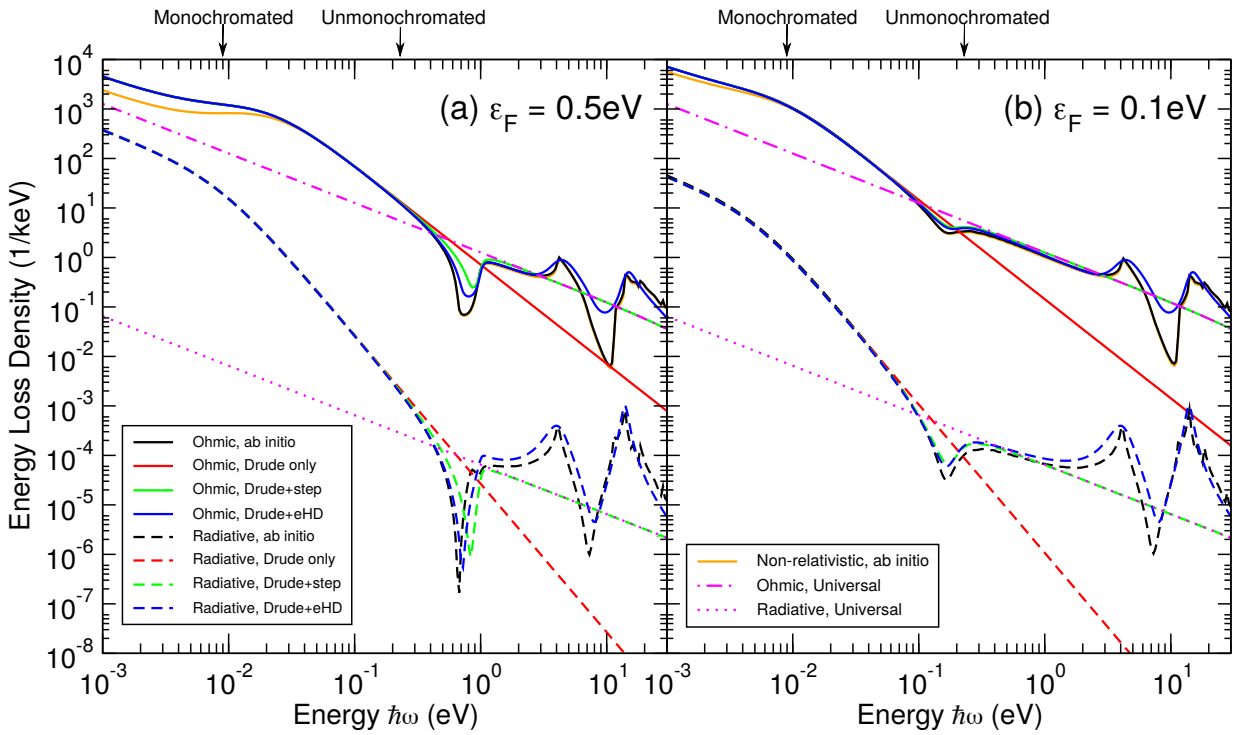

Figure 6.6: Ohmic (solid lines) and radiative (dashed lines) probability energy loss densities in units of 1/keV versus energy $\hbar\omega$ in units of eV using conductivity inputs from LCAO-TDDFT-k-ω (black), Drude only (red), Drude+step (green), Drude+eHD (blue), and the universal conductivity (magenta) for an incident electron energy of $E_0 = 100$ keV and ε_F equal to (a) 0.5 eV and (b) 0.1 eV. The non-relativistic Ohmic energy loss density using the LCAO-TDDFT-k-ω conductivity as input (orange) is also shown.

valence-EELS with large ZLP, as seen in Fig. 6.2, any changes due to doping are not expected to be visible to methods that cannot probe below this range. For both doping scenarios a dip appears between ε_F and $2\varepsilon_F$, matching with the optical gap due to Pauli blocking seen in Fig. 6.5. Below $\varepsilon_F/2$ the Ohmic and radiative energy loss densities broadly match the dominant intraband transitions, a feature expected based on the Dirac plasmon polariton being located in this range [21]. In addition to the three empirical models, Fig. 6.6 also shows the energy loss spectra assuming that only the universal conductivity $\sigma(\omega) = e^2/4\hbar$ leads to electron energy losses at all frequencies, extending the Dirac cone approximation to the full energy range under consideration. These lines reveal the shape of the energy loss spectra for energies below 2 eV for undoped graphene, confirming the $P(\omega) \propto 1/(\hbar\omega)$ behaviour of the spectra observed experimentally for low energy losses [12, 62]. In addition, the location of the crossing between this universal conductivity energy loss density and the one from the Drude only model is seen to directly indicate the location of the Pauli blocking dip for both doping cases, the intersection occurring when the energy loss is close to ε_F. The vertical separation of the lines in this universal conductivity case is about four orders of magnitude, providing an estimate of the degree to which radiative losses

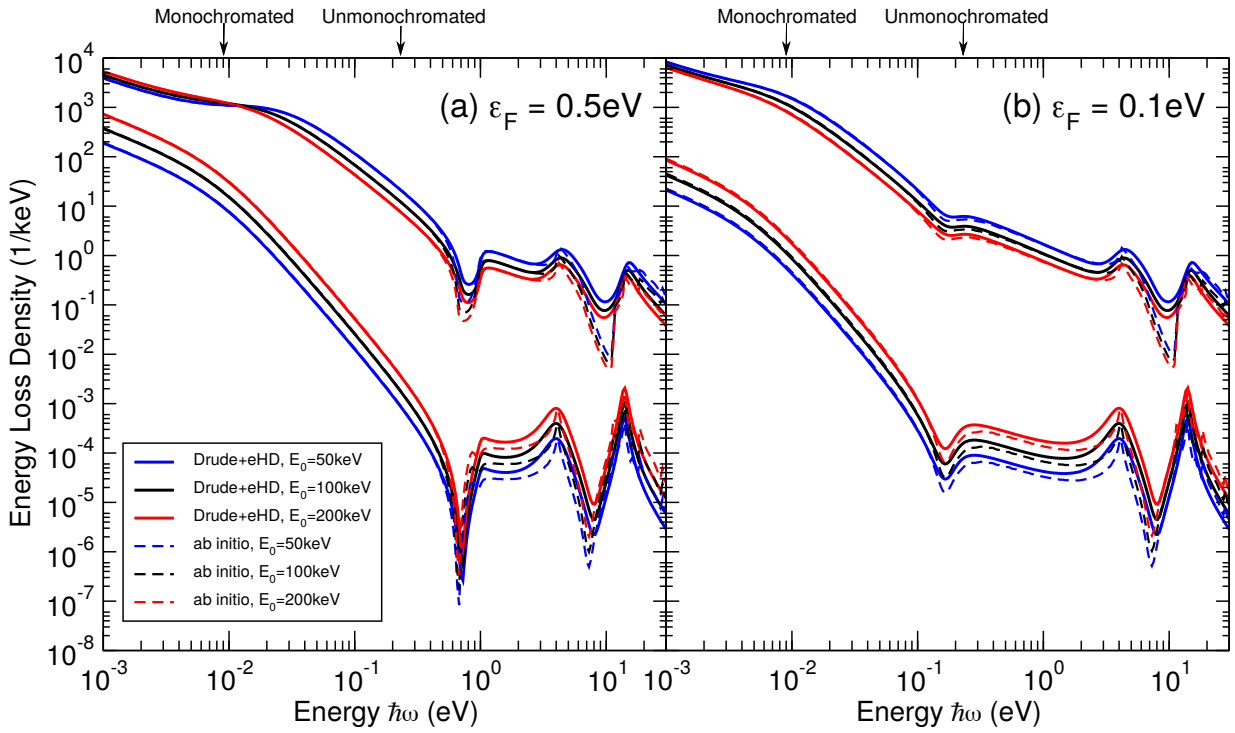

Figure 6.7: Ohmic (top curves) and radiative (bottom curves) energy loss densities in units of 1/keV versus energy $\hbar\omega$ in units of eV with the *ab initio* LCAO-TDDFT-k-ω (dashed lines) and the Drude+eHD model (solid lines) conductivities used as input for incident electron energies of E_0 = 50 (blue), 100 (black), and 200 (red) keV at Fermi energies of (a) 0.5 eV and (b) 0.1 eV, assuming $q_c \to \infty$.

are suppressed relative to the Ohmic loss for undoped graphene [21, 71, 22]. Lastly, the prominence of the radiative loss is especially visible for the higher doping case and for extremely low frequencies, with the radiative energy loss within one order of magnitude of the Ohmic.

Fig. 6.6 also shows the Ohmic energy loss obtained from Eq. 2.99 in the non-relativistic regime for both doping cases with the LCAO-TDDFT-k-ω in-plane conductivity, noting that the only differences come about for energy losses $\lesssim \varepsilon_F/20$, which is expected given the importance of relativistic effects in the THz to MIR frequency range [21]. Two arrows in Fig. 6.6 are located at energy losses of 9 meV and 250 meV, corresponding to FWHM values of the zero-loss peak [169] for monochromated and non-monochromated electron beams, respectively, as described in Section 2.3. These arrows highlight the edges of the intervals accessible to each of these two types of EELS setup, showing the energy range that may be accessible to more advanced STEM devices. The location of the monochromated arrow for ε_F = 0.5 eV covers an area where the relativistic and non-relativistic regimes diverge, showing that high resolution EELS setups should in theory be able to probe the region where relativistic formulations become important.

To see the effects that changes in the incident electron energy E_0 have on the Ohmic and

radiative energy losses, Fig. 6.7 shows that for both doping cases the effect of increasing electron velocity is to increase the intensity of the radiative spectra while decreasing that of the Ohmic spectra, with the exception being for $\varepsilon_F = 0.5$ eV below the monochromated arrow, although even for $\varepsilon_F = 0.1$ eV the intensity increase is far less prominent. This effect relates to the role that the damping parameter γ_D in Eq. 6.1 plays in describing the relativistic effects of intraband electron transitions [21]. As this last effect is slightly beyond the range of a monochromated EELS beam there is no certainty that it can be observed experimentally.

To explore in further detail how the choice of γ_D in Eq. 6.1 influences the relative ratios of the Ohmic and radiative energy loss contributions and the difference between non-relativistic and relativistic models for the energy loss, Fig. 6.8(a,b) shows the total, Ohmic, radiative, and non-relativistic energy loss densities normalized by $P_c = 4/(\pi\varepsilon_F)$, against the normalized energy loss $\overline{\omega} = \omega/\omega_c$ with $\hbar\omega_c = \alpha\varepsilon_\mathrm{F}$, where α is the fine structure constant. This normalization factor is chosen to be on the order of 1 THz for typical graphene doping densities. The Drude only model is used for the input conductivity, as we showed in Fig. 6.6 that for frequencies below $\varepsilon_\mathrm{F}/2$, corresponding to $\overline{\omega} \approx 60$, the intraband transitions described by this model were by far the dominant effect. Fig. 6.8(a,b) shows that the Ohmic and radiative energy loss densities actually cross at low enough frequencies when using such low values of γ_D, the radiative component becoming the dominant factor in describing the relativistic energy loss, an effect not seen in Fig. 6.6 and Fig. 6.7 where $\gamma_D = 0.004$ eV for both $\varepsilon_F = 0.5$ eV and $\varepsilon_F = 0.1$ eV scenarios. These figures also show how at these very low energy losses the non-relativistic model fails to properly match with the energy loss density given in the relativistic model, an effect seen to a lesser extent in Fig. 6.6.

The energy loss where the Ohmic and radiative energy loss densities cross, which we refer to as $\overline{\omega}^*$, represents an important transition point that shows when the non-relativistic model is no longer applicable. Fig. 6.8(c) shows how this crossing point changes relative to the normalized incoming velocity of the electron $\beta = v/c$ for different normalized intraband dampings $\overline{\gamma} = \gamma_D/\omega_c$ equal to 0, 0.01, 0.1, and 0.5. For high damping, the crossing only occurs for very fast electrons and at very low frequencies, but as the damping parameter falls the velocity threshold to induce the crossing drops as well. The highest possible energy loss where the crossing can occur for low energy incident electrons happens near $\overline{\omega}^* = 0.3$ for the case of zero damping, where the strong dependence of the crossing on the value β is no longer seen.

6.2.3 Effects of Small Collection Angle

As a parallel to Fig. 6.3, it is worthwhile to explore the effect that smaller collection apertures, described according to the maximum collection momentum q_c through Eq. 2.82, have on the peak intensities and locations of the monolayer energy loss spectra. Reducing the size of the aperture has been explored previously [176] in the context of eliminating the effects that hybridization of plasmon and optical modes along with relativistic effects have on the EEL spectra of monochromated electron beams. For graphene a similar argument can be made that a smaller

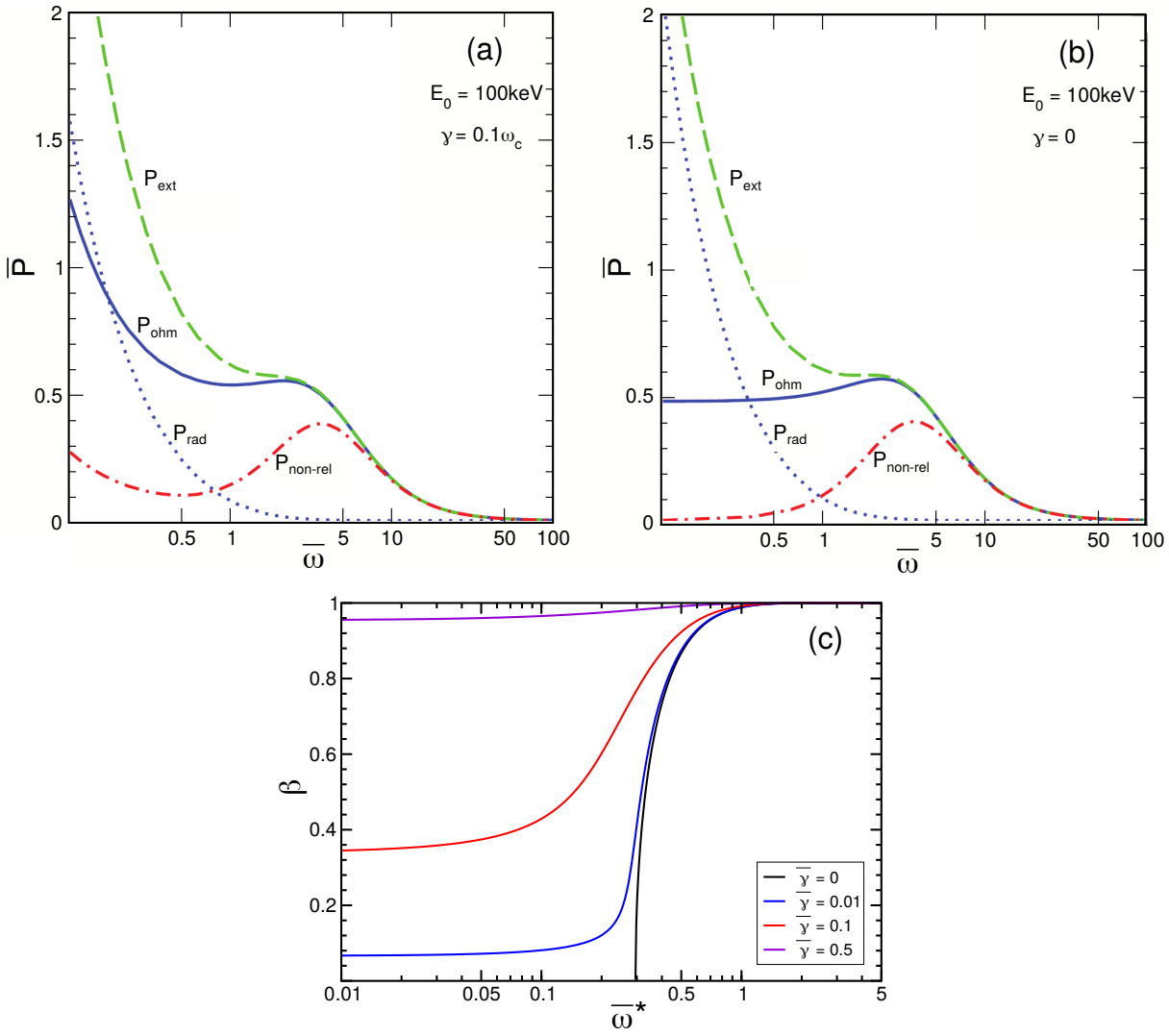

Figure 6.8: Normalized energy loss densities $\overline{P} = P/P_c$ with $P_c = 4/(\pi\varepsilon_F)$ versus normalized energy loss $\overline{\omega} = \omega/\omega_c$ with $\hbar\omega_c = \alpha\varepsilon_F$ for normal incident electron energy $E_0 = 100$ keV showing the external (total) (green dashed), Ohmic (blue solid), radiative (blue dashed), and non-relativistic (red dash-dotted) energy loss densities for damping constants γ_D equal to (a) $0.1\,\omega_c$ and (b) 0 using the Drude only model for the input conductivity, and (c) the normalized electron speed $\beta = v/c$ versus the frequency $\overline{\omega}^*$ where $P_{ohm}(\overline{\omega}^*) = P_{rad}(\overline{\omega}^*)$ for $\overline{\gamma} = \gamma_D/\omega_c$ equal to 0 (black), 0.01 (blue), 0.1 (red) and 0.5 (purple).

collection aperture will reduce the overlap that the EEL spectra may have with the zero-loss peak, while hybridization with the Dirac plasmon [177, 178] with its $\propto \sqrt{q}$ dispersion may also become important for relatively large momentum transfers in the material. Previous investigations [98, 176, 63] of the role that the band gap plays in the profile of the EEL spectra for semiconductors have parallels with the optical gap resulting from Pauli blocking in graphene,

further motivating the exploration of the effects that different aperture sizes have.

Fig.6.9 shows the Ohmic energy loss density using both the Drude+eHD and the LCAO-TDDFT-k-ω conductivities as input, for Fermi levels of $\varepsilon_F = 0.5$ and 0.1 eV and an incident electron energy of $E_0 = 100$ keV with q_c values ranging across a broad spectrum, noting as in Fig. 6.3 the necessity of using the proper limits of integration in Eq. 2.105 rather than the large collection angle limit. As in neutral graphene, the effects of the reduced collection angle appear most strongly as a reduction in the intensity of the interband peaks and the surrounding spectra, although the shape remains intact. Interestingly at smaller q_c values the onset of the effect of the maximum collection momentum reduction occurs at an energy loss value $\hbar\omega_*$ located in the optical gap caused by Pauli blocking for doped graphene, gradually moving into the region of intraband electron transitions for decreasing maximal momentum transfer. The occurrence of this sudden drop is well within the range that a monochromated electron beam can access, suggesting that high energy resolution techniques can probe this region to assess the accuracy of this model towards describing graphene's optical response.

Given the wide difference between the spectra collected with different q_c, there exists the

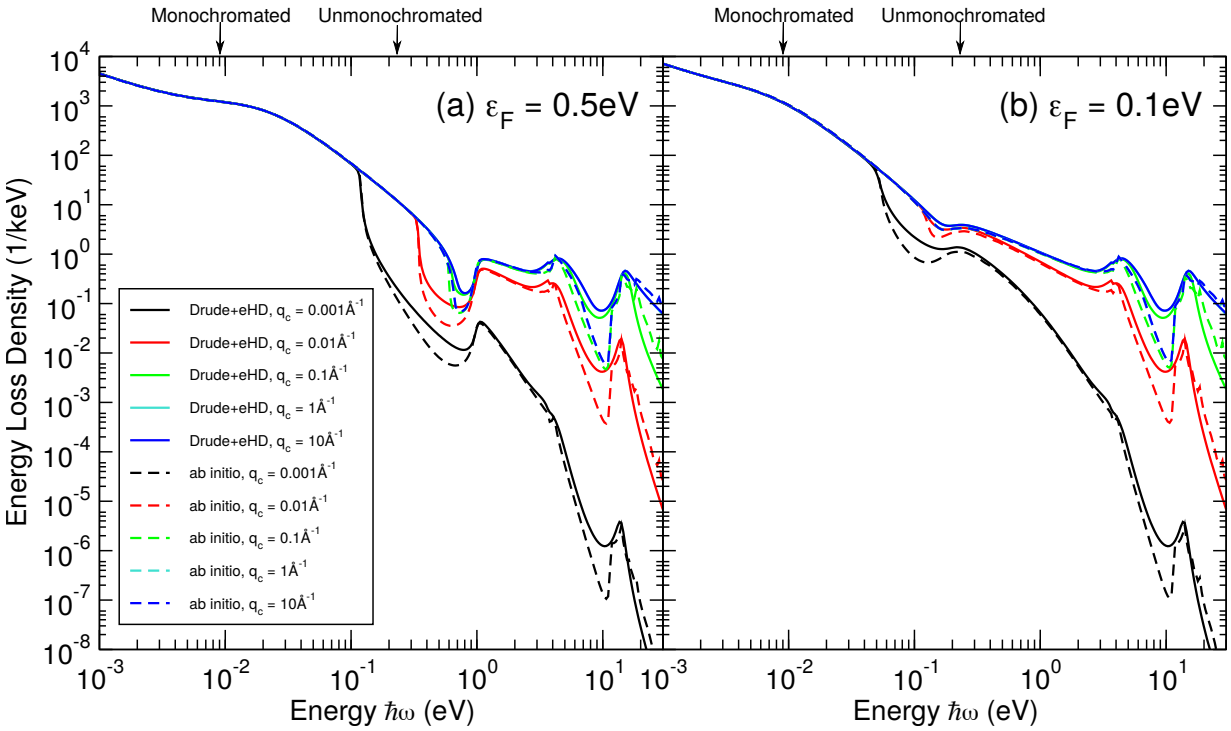

Figure 6.9: Ohmic energy loss density in units of 1/keV using the Drude+eHD (solid lines) or the LCAO-TDDFT-k-ω (dashed lines) conductivities as input versus energy $\hbar\omega$ in units of eV with an incident electron energy of $E_0 = 100$ keV for maximum in-plane collected momenta q_c equal to 0.001 (black), 0.01 (red), 0.1 (green), 1 (cyan), and 10 (blue) Å^{-1} at Fermi levels ε_F equal to (a) 0.5 and (b) 0.1 eV.

possibility of recovering the optical conductivity of graphene given enough experimental EELS data collected at different collection angles [98, 176, 177], noting that the formulae given in Eq. 2.100 for the non-relativistic energy loss are analytic and only require the incident electron energy, the maximum collection angle, the energy loss frequency, and the real and imaginary components of the optical conductivity at a given energy as input, noting that in our discussion for Fig. 6.6 and Fig. 6.8 we have emphasized that the non-relativistic limit is sufficient to describe energy losses for frequencies above the intraband dominated regime. It is therefore theoretically possible to numerically determine the full complex optical conductivity using optical spectra collected using two sufficiently different aperture sizes, noting in Fig.6.9 that spectra collected for $q_c \leq 0.1\text{Å}^{-1}$ begin to show the most substantial intensity change.

One caveat of this approach is that below a certain frequency $\hbar\omega_*(q_{c1})$, where q_{c1} is the smaller of the two maximal collection momenta, the two energy loss spectra converge to the same value, limiting the ability of our extraction procedure to deduce the complex conductivity below this value. However, if it is assumed that the Drude only model describes this region sufficiently well, the only parameters needed to calculate the conductivity in this region are those in Eq. 3.5, but since ε_F is determined from the location of the dip due to Pauli blocking, this means that γ_D, even if it contains some energy dependence, can be directly inferred from any spectral calculation done with high enough energy resolution. Once σ_{intra} is determined this way, the extraction procedure can be done as before to determine σ_{inter}, including all relevant peaks and Pauli blocking. We can directly estimate the q_c necessary to ensure that frequencies below the value $\hbar\omega_*(q_{c1})$ are well described by the intraband model by inputting the Drude+step conductivity into the second line of Eq. 2.110 and assuming that the real component of the conductivity is vanishingly small. This factor will cause a Delta-function spike when

$$1 - \frac{2\pi}{c}\,\mathrm{Im}\,\sigma(\omega)\sqrt{\left(\frac{cq}{\omega}\right)^2 - 1} = 0. \tag{6.2}$$

Given that $\hbar\omega_*$ represents where the intraband transition becomes dominant, using $\mathrm{Im}\,\sigma(\omega)$ from the Drude only model in Eq. 3.5 in this dispersive relation yields a Dirac plasmon-like dispersion $\hbar\omega_* = \sqrt{2e^2 q_c \varepsilon_F}$ [21]. This kind of square root dependence can be seen directly in Fig. 6.9(a), where reducing q_c from 0.01Å^{-1} to 0.001Å^{-1}, a factor of 10, moves the location of $\hbar\omega_*$ from 0.33 eV to 0.11 eV, a factor of approximately $\sqrt{10}$. The applicability of this dispersion relation no longer works for higher wavenumber as the conductivity is dominated by the Pauli blocking optical gap, to the point where $\hbar\omega_*$ saturates near $\hbar\omega_* \approx 5\varepsilon_F/3$ in the limit of a broad collection angle $q_c \to \infty$ [71].

Chapter 7

Conclusions and Future Work

7.1 Summary

This research focused on density functional theory and its applications from three separate but connected perspectives. Recognizing the opportunity that novel deep learning tools have towards gaining a better understanding of the exchange-correlation energy in the Kohn-Sham equations, we explored the application of feed-forward networks to predicting certain universal functionals and their accuracy for real-world material structures, finally positing how extending these methods can provide either more accurate or less computationally expensive DFT calculations. We next explored a formulation of TDDFT which leverages an LCAO basis set representation of the Kohn-Sham equations in wavenumber and frequency space, providing accurate results for lower-dimensional materials and high computational savings for large macromolecular systems, thereby expanding the range of materials that can be assessed using TDDFT. We finally built up a phenomenological model for the energy loss of graphene, analyzing the effect that theoretical and experimental setup parameters had on the expected energy loss spectra, thereby linking quantities that can be computed using TDDFT to actual experimentally measureable values.

In Chapter 4, noise of different amplitudes is added to the PBE functional to assess the required accuracy of a neural network which trains the LDA and PBE xc functionals. The performance of the NNLDA where the xc energy density and xc potential are trained separately is contrasted with the case of an xc potential derived from a functional derivative passed through a trained network. The NNPBE is trained likewise and runs into similar problems as the NNLDA, but across a separate range of input parameters. The so-called "rough model" and "difficulty zone" are factored into a retrained NNPBE approach with training guided towards areas of poor performance, and changes in the width and depth on the network performance are evaluated. The range of training data for real-world systems along with reference energies for calculations done in GPAW is determined, and the network performance is evaluated over a wide range of molecular and crystal systems. Overall the NNLDA is shown to be able to reach a high degree of accuracy for nearly all systems while certain facets of the DFT code need to be explored before NNPBE

can be deemed successful.

A series of benchmarking calculations and visualization tools are presented in Chapter 5 for our LCAO-TDDFT-k-ω code. The pros and cons of using the LCAO-TDDFT-k-ω methodology are explored in depth and the kinds of materials used to benchmark the code are chosen accordingly. The effect of the derivative discontinuity correction, the convergence with respect to the choice of LCAO basis set, and the use of low-dimensional response functions are all explored in depth. Fullerene and chlorophyll a and b are shown to give optical spectra in line with approaches like G_0W_0-BSE and experimental spectra, while studying SWCNTs reveal that LCAO-TDDFT-k-ω gives a similar performance to PW-TDDFT-k-ω and G_0W_0-BSE at a fraction of the computational cost. Graphene and phosphorene monolayers are analyzed the same way along with additional visualization tools for the distribution of points in reciprocal space and the spatial distribution of electrons and hole excitations, yielding insight into the spatial nature of graphene's conductivity peaks and the driving mechanism behind phosphorene's anisotropic response functions. Analysis of graphene's optical spectra is left to the next chapter but the phosphorene spectra are tested against experimental reflection data to verify the use of mean-field approximations to quantities like the polarizability. Anatase and rutile titanium dioxide, two systems well-known to be tricky from a DFT perspective, are analyzed to show the applicability of the LCAO-TDDFT-k-ω code to systems of any dimension.

To relate quantities like the dielectric function computed via TDDFT to experimental measureables like electron energy loss spectra, Chapter 6 highlights the applicability of a phenomenological hydrodynamic model for the optical conductivity of monolayer graphene, factoring in the universal conductivity, Pauli blocking, and intraband transitions in a systematic way. This model is fit empirically to results calculated using TDDFT and is shown to reproduce experimental optical spectra over a wide energy range. Experimental parameters like the size of the aperture, the energy of incoming electrons, and the tuning of the monolayer are shown to all play a role influencing the shape of the spectra, especially in a region of energies only recently available for probing with refined spectroscopy techniques. Theoretical assumptions about the importance of relativistic effects, empirically determined damping rates, and the validity of simpler models for lower energy regimes are tested against results from conductivities computed empirically within our model and using the LCAO-TDDFT-k-ω approach from Chapter 5. Lastly, a procedure based on high-resolution EELS measurements taken at different collection angles is developed that allows for a reconstruction of the optical conductivity of graphene over a broad range of frequencies.

7.2 Future Work

The results of Chapter 4 reveal that our neural network trained PBE xc functional requires certain refinements before it can be considered a valid approximation to the original PBE. Future work in this respect includes limiting the training range to only electron densities and gradients relevant to the materials under consideration, exploring the importance of better converged reference

energies on the perceived quality of the network approximation, and analyzing factors that lead single-point and self-consistent calculations to diverge from each other for materials containing elements of high atomic number. The possibility of shifting the framework so that the xc potential is the only quantity trained may also provide better converged results [85, 34]. Afterwards, future work would move on towards predicting the total energies of systems utilizing meta-GGA or hybrid functionals with networks that take in the electron density as input non-locally, thereby leveraging the full predictive capabilities of deep learning methods.

While many of the results presented in Chapter 5 constitute benchmarking of the LCAO-TDDFT-k-ω implementation against other well-established TDDFT approaches using materials already thoroughly researched in the literature, determining the optical properties of chlorophyll monomers both alone and within the larger protein matrix remains an important problem, and future work can utilize the computational efficiency of the LCAO-TDDFT-k-ω to model the optical spectra of the full light harvesting complex to learn the full mechanism whereby light is absorbed and the energy transported to living organisms. Regarding improving the code itself, future work may explore how refinements on the LCAO basis sets, using either diffuse Gaussians[52] or adjusting the role of polarization in describing unoccupied orbitals[17], affects the convergence of our implementation. Future work may also explore how the spatially resolved electron and hole densities derived from the LCAO-TDDFT-k-ω code can be used to describe results from atomically resolved EELS for low-dimensional nanostructures[179].

Much of the formalism developed and used in Chapter 6 can be readily applied to other atomically thin materials. Future work on an empirical model for doped phosphorene for the $0 \leq \hbar\omega \leq 3$ eV energy range that can then by plugged into our 2D energy loss formalism and analyzed much in the same way as the graphene monolayer. Recent experimental EELS results [129] suggest that the out-of-plane response function, which is almost entirely quashed in graphene due to local field effects[8], may prove relevant for monolayer phosphorene, and future work can formulate the way in which the response normal to the surface affects the energy loss equations [71]. An investigation on whether the model developed for the energy loss works for alternative EELS setups, low-energy electron microscopy for example, also warrants further investigation.

Bibliography

[1] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964. 1, 5, 12, 14, 59

[2] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, Dec 1994. 1, 27

[3] Peter E. Blöchl, Clemens J. Först, and Johannes Schimpl. Projector augmented wave method: *Ab Initio* molecular dynamics with full wave functions. *Bull. Mater. Sci.*, 26(1):33–41, Jan 2003. 1, 27

[4] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, May 1981. 1, 14, 15

[5] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996. 1, 5, 15, 17

[6] A. G. Petukhov, I. I. Mazin, L. Chioncel, and A. I. Lichtenstein. Correlated metals and the lda+u method. *Phys. Rev. B*, 67:153106, Apr 2003. 1

[7] M. Shishkin and G. Kresse. Implementation and performance of the frequency-dependent *gw* method within the paw framework. *Phys. Rev. B*, 74:035101, Jul 2006. 1, 4, 32

[8] Dino Novko, Marijan Šunjić, and Vito Despoja. Optical absorption and conductivity in quasi-two-dimensional crystals from first principles: Application to graphene. *Phys. Rev. B*, 93:125413, Mar 2016. 1, 120

[9] María Rosa Preciado-Rivas, Victor Alexander Torres-Sánchez, and Duncan J. Mowbray. Optical absorption and energy loss spectroscopy of single-walled carbon nanotubes. *Phys. Rev. B*, 100:235429, Dec 2019. 1, 5, 18, 19, 20, 27, 29, 33, 54, 55, 78, 86

[10] Livia Noëmi Glanzmann, Duncan John Mowbray, Diana Gisell Figueroa del Valle, Francesco Scotognella, Guglielmo Lanzani, and Angel Rubio. Photoinduced absorption within single-walled carbon nanotube systems. *J. Phys. Chem. C*, 120(3):1926–1935, 2015. 1, 27, 28, 29, 54

[11] D.J. Mowbray, S. Segui, J. Gervasoni, Z.L. Miskovic, and N.R Arista. Plasmon excitations on a single-wall carbon nanotube by external charges: Two-dimensional two-fluid hydrodynamic model. *Phys. Rev. B*, 82:035405, 2010. 1, 51, 54

[12] P. Wachsmuth, R. Hambach, M. K. Kinyanjui, M. Guzzo, G. Benner, and U. Kaiser. High-energy collective electronic excitations in free-standing single-layer graphene. *Phys. Rev. B*, 88:075433, 2013. 1, 8, 34, 35, 105, 112

[13] F. Henneke, L. Lin, C. Vorwerk, C. Draxl, R. Klein, and C. Yang. Fast optical absorption spectra calculations for periodic solid state systems, 2019. 1, 32

[14] E. E. Salpeter and H. A. Bethe. A relativistic equation for bound-state problems. *Phys. Rev.*, 84:1232–1242, Dec 1951. 1, 19, 32

[15] 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. *Phys. Rev. B*, 80:195112, Nov 2009. 1, 6, 22

[16] Joachim Paier, Robin Hirschl, Martijn Marsman, and Georg Kresse. The perdewburkeernzerhof exchange-correlation functional applied to the g2-1 test set using a plane-wave basis set. *The Journal of Chemical Physics*, 122(23):234102, 2005. 1, 5, 15, 61, 70, 147, 148

[17] Tuomas P. Rossi, Susi Lehtola, Arto Sakko, Martti J. Puska, and Risto M. Nieminen. Nanoplasmonics simulations at the basis set limit through completeness-optimized, local numerical basis sets. *J. Chem. Phys.*, 142:094114, 2015. 1, 6, 76, 120

[18] L. Matthes, O. Pulci, and F. Bechstedt. Influence of out-of-plane response on optical properties of two-dimensional materials: First principles approach. *Phys. Rev. B*, 94:205408, 2016. 1, 8, 30, 31

[19] John F. Dobson, Tim Gould, and Sébastien Lebègue. Layer response theory: Energetics of layered materials from semianalytic high-level theory. *Phys. Rev. B*, 93:165436, Apr 2016. 1, 8, 30

[20] Keenan Lyon, Mara Rosa Preciado-Rivas, Duncan John Mowbray, and Vito Despoja. Lcao-tddft-*k*-: Spectroscopy in the optical limit, 2019. 1, 8, 32

[21] Z. L. Miskovic, S. Segui, J. Gervasoni, and N. R. Arista. Energy losses and transition radiation produced by the interaction of charged particles with a graphene sheet. *Phys. Rev. B*, 94:125414, 2016. 2, 8, 39, 40, 106, 112, 113, 114, 117

[22] K. Lyon, D. J. Mowbray, and Z. L. Miskovic. Modelling relativistic effects in momentum-resolved electron energy loss spectroscopy of graphene. *Radiation Effects and Defects in Solids*, 173(1-2):8–21, 2018. 2, 8, 39, 51, 113

[23] Keenan Lyon, Kamran Akbari, and Zoran L. Miskovic. Relativistic effects in the interaction of fast charged particles with graphene. *Radiation Effects and Defects in Solids*, 175(1-2):84–93, 2020. 2

[24] Felix Brockherde, Leslie Vogt, Li Li, Mark E. Tuckerman, Kieron Burke, and Klaus-Robert Müller. Bypassing the kohn-sham equations with machine learning. *Nature Communications*, 8(1):872, 2017. 3, 4, 5, 44

[25] Axel D. Becke. Densityfunctional thermochemistry. iii. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993. 3, 17

[26] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, 1999. 3

[27] Franois Gygi and Ivan Duchemin. Efficient computation of hartreefock exchange using recursive subspace bisection. *Journal of Chemical Theory and Computation*, 9(1):582–587, 2013. PMID: 26589056. 3

[28] Kristof T. Schütt, Farhad Arbabzadah, Stefan Chmiela, Klaus R. Müller, and Alexandre Tkatchenko. Quantum-chemical insights from deep tensor neural networks. *Nature Communications*, 8(1):13890, 2017. 3

[29] Changjian Xie, Xiaolei Zhu, David R. Yarkony, and Hua Guo. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. iv. coupled diabatic potential energy matrices. *The Journal of Chemical Physics*, 149(14):144107, 2018. 3

[30] John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, and Kieron Burke. Finding density functionals with machine learning. *Phys. Rev. Lett.*, 108:253002, Jun 2012. 3, 4

[31] Li Li, Thomas E. Baker, Steven R. White, and Kieron Burke. Pure density functional for strong correlation and the thermodynamic limit from machine learning. *Phys. Rev. B*, 94:245129, Dec 2016. 3, 4

[32] Brian Kolb, Levi C. Lentz, and Alexie M. Kolpak. Discovering charge density functionals and structure-property relationships with prophet: A general framework for coupling machine learning and first-principles methods. *Scientific Reports*, 7(1):1192, 2017. 3, 4

[33] Kevin Ryczko, Kyle Mills, Iryna Luchak, Christa Homenick, and Isaac Tamblyn. Convolutional neural networks for atomistic systems. *Computational Materials Science*, 149:134 – 142, 2018. 3, 4, 47

[34] Ryo Nagai, Ryosuke Akashi, Shu Sasaki, and Shinji Tsuneyuki. Neural-network kohn-sham exchange-correlation potential and its out-of-training transferability. *The Journal of Chemical Physics*, 148(24):241737, 2018. 3, 5, 64, 120

[35] Levi C Lentz and Alexie M Kolpak. Predicting HSE band gaps from PBE charge densities via neural network functionals. *Journal of Physics: Condensed Matter*, 32(15):155901, jan 2020. 4

[36] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B*, 71:035109, 2005. 4, 5, 77

[37] J Enkovaara, C Rostgaard, J J Mortensen, J Chen, M Dułak, L Ferrighi, J Gavnholt, C Glinsvad, V Haikola, H A 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 K Nørskov, M Puska, T T Rantala, J Schiøtz, K S Thygesen, and K W Jacobsen. Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method. *J. Phys.: Condens. Matter*, 22:253202, 2010. 4, 77

[38] Kun Yao and John Parkhill. Kinetic energy of hydrocarbons as a function of electron density and convolutional neural networks. *Journal of Chemical Theory and Computation*, 12(3):1139–1147, 2016. PMID: 26812530. 4

[39] Kevin Ryczko, David A. Strubbe, and Isaac Tamblyn. Deep learning and density-functional theory. *Phys. Rev. A*, 100:022512, Aug 2019. 4, 5, 47

[40] K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier. Error estimates for solid-state density-functional theory predictions: An overview by means of the ground-state elemental crystals. *Critical Reviews in Solid State and Materials Sciences*, 39(1):1–24, 2014. 5, 60, 70, 149, 150

[41] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E. Castelli, Stewart J. Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, Marcin Dułak, Olle Eriksson, José A. Flores-Livas, Kevin F. Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, Oscar Grånäs, E. K. U. Gross, Andris Gulans, François Gygi, D. R. Hamann, Phil J. Hasnip, N. A. W. Holzwarth, Diana Iuşan, Dominik B. Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, Emine Küçükbenli, Yaroslav O. Kvashnin, Inka L. M. Locht, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, Lars Nordström, Taisuke Ozaki, Lorenzo Paulatto, Chris J. Pickard, Ward Poelmans, Matt I. J. Probert, Keith Refson, Manuel Richter, Gian-Marco Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, Patrik Thunström, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J. van Setten, Veronique Van Speybroeck, John M. Wills, Jonathan R. Yates, Guo-Xu Zhang, and Stefaan Cottenier. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280), 2016. 5, 70, 149, 150

[42] Fuqiang Ban, Kathryn N. Rankin, James W. Gauld, and Russell J. Boyd. Recent applications of density functional theory calculations to biomolecules. *Theoretical Chemistry Accounts*, 108(1):1–11, 2002. 5

[43] Wu Qin, Xin Li, Wen-Wen Bian, Xiu-Juan Fan, and Jing-Yao Qi. Density functional theory calculations and molecular dynamics simulations of the adsorption of biomolecules on graphene surfaces. *Biomaterials*, 31(5):1007 – 1016, 2010. 5

[44] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996. 5, 21, 71

[45] See gitlab.com/lcao-tddft-k-omega/lcao-tddft-k-omega where the LCAO-TDDFT-k-ω code is available free of charge. 5, 151

[46] María Rosa Preciado-Rivas, Duncan John Mowbray, Keenan Lyon, Ask Hjorth Larsen, and Bruce Forbes Milne. Optical excitations of chlorophyll *a* and *b* monomers and dimers. *J. Chem. Phys.*, 151:174102, 2019. 5, 27, 56, 58

[47] Mark E. Casida. *Time-Dependent Density Functional Response Theory for Molecules*, pages 155–192. World Scientific, Singapore, 1995. 5, 75

[48] Jun Yan, Kristian S. Thygesen, and Karsten Wh. Jacobsen. Nonlocal screening of plasmons in graphene by semiconducting and metallic substrates: First-principles calculations. *Phys. Rev. Lett.*, 106:146803, 2011. 5, 75

[49] K Yabana and G F. Bertsch. Time-dependent local-density approximation in real time. *Phys. Rev. B*, 54:4484–4487, 09 1996. 5, 75

[50] Joaquim Jornet-Somoza, Joseba Alberdi-Rodriguez, Bruce F. Milne, Xavier Andrade, Miguel A. L. Marques, Fernando Nogueira, Micael J. T. Oliveira, James J. P. Stewart, and Angel Rubio. Insights into colour-tuning of chlorophyll optical response in green plants. *Phys. Chem. Chem. Phys.*, 17(40):26599–26606, 2015. 6, 56, 75

[51] José M Soler, Emilio Artacho, Julian D Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. The SIESTA method for ab initio order-n materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745–2779, mar 2002. 6

[52] M. Kuisma, A. Sakko, T. P. Rossi, A. H. Larsen, J. Enkovaara, L. Lehtovaara, and T. T. Rantala. Localized surface plasmon resonance in silver nanoparticles: Atomistic first-principles time-dependent density-functional theory calculations. *Phys. Rev. B*, 91:115431, Mar 2015. 6, 76, 120

[53] Oleg Gritsenko, Robert van Leeuwen, Erik van Lenthe, and Evert Jan Baerends. Self-consistent approximation to the Kohn-Sham exchange potential. *Phys. Rev. A*, 51(3):1944–1954, March 1995. 7, 17

[54] Oleg V. Gritsenko, Robert van Leeuwen, and Evert Jan Baerends. Direct approximation of the long- and short-range components of the exchange-correlation Kohn-Sham potential. *Int. J. Quantum Chem.*, 61(2):231–243, 1997. 7, 17

[55] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala. Kohn-Sham potential with discontinuity for band gap materials. *Phys. Rev. B*, 82:115106, Sep 2010. 7, 17, 18, 32

[56] Barun Ghosh, Piyush Kumar, Anmol Thakur, Yogesh Singh Chauhan, Somnath Bhowmick, and Amit Agarwal. Anisotropic plasmons, excitons, and electron energy loss spectroscopy of phosphorene. *Phys. Rev. B*, 96:035422, Jul 2017. 7, 36, 53, 54

[57] Yu. V. Bludov, Aires Ferreira, N. M. R. Peres, and M. I. Vasilevskiy. A primer on surface plasmon-polaritons in graphene. *International Journal of Modern Physics B*, 27(10):1341001, 2013. 8, 50, 52

[58] T. Djordjevic, I. Radovic, V. Despoja, K. Lyon, D. Borka, and Z.L. Miskovic. Analytical modeling of electron energy loss spectroscopy of graphene: Ab initio study versus extended hydrodynamic model. *Ultramicroscopy*, 184:134–142, 2018. 8, 38, 51, 52, 110

[59] Albert Polman, Mathieu Kociak, and F. Javier García de Abajo. Electron-beam spectroscopy for nanophotonics. *Nature Materials*, 18(11):1158–1171, 2019. 8

[60] T. Eberlein, U. Bangert, R. R. Nair, R. Jones, M. Gass, A.L. Bleloch, K.S. Novoselov, A. Geim, and P.R. Briddon. Plasmon spectroscopy of free-standing graphene films. *Phys. Rev. B*, 77:233406, 2008. 8, 35, 105

[61] Antonio Politano and Gennaro Chiarello. Plasmon modes in graphene: status and prospect. *Nanoscale*, 6:10927–10940, 2014. 8

[62] F.J. Nelson, J.-C. Idrobo, J.D. Fite, Z.L. Miskovic, S.J. Pennycook, S.T. Pantelides, J.U. Lee, and A.C. Diebold. Electronic excitations in graphene in the 1.50 ev range: The π and $\sigma + \pi$ peaks are not plasmons. *Nano Lett.*, 14:3827–3831, 2014. 8, 34, 35, 105, 112

[63] M. Stöger-Pollach. Optical properties and bandgaps from low loss eels: Pitfalls and solutions. *Micron*, 39(8):1092 – 1110, 2008. 8, 115

[64] Peter A. Crozier. Vibrational and valence aloof beam eels: A potential tool for non-destructive characterization of nanoparticle surfaces. *Ultramicroscopy*, 180:104 – 114, 2017. Ondrej Krivanek: A research life in EELS and aberration corrected STEM. 8

[65] O.L. Krivanek, N. Dellby, J.A. Hachtel, J.-C. Idrobo, M.T. Hotz, B. Plotkin-Swing, N.J. Bacon, A.L. Bleloch, G.J. Corbin, M.V. Hoffman, C.E. Meyer, and T.C. Lovejoy. Progress in ultrahigh energy resolution eels. *Ultramicroscopy*, 203:60 – 67, 2019. 8, 102

[66] Tianrong Zhan, Dezhuan Han, Xinhua Hu, Xiaohan Liu, Siu-Tat Chui, and Jian Zi. Tunable terahertz radiation from graphene induced by moving electrons. *Phys. Rev. B*, 89:245434, Jun 2014. 8

[67] Shenggang Liu, Chao Zhang, Min Hu, Xiaoxing Chen, Ping Zhang, Sen Gong, Tao Zhao, and Renbin Zhong. Coherent and tunable terahertz radiation from graphene surface plasmon polaritons excited by an electron beam. *Applied Physics Letters*, 104(20):201104, 2014. 8

[68] Kai-Chun Zhang, Xiao-Xing Chen, Chang-Jian Sheng, Kelvin J. A. Ooi, Lay Kee Ang, and Xue-Song Yuan. Transition radiation from graphene plasmons by a bunch beam in the terahertz regime. *Opt. Express*, 25(17):20477–20485, Aug 2017. 8

[69] V M Sukharev, M N Strikhanov, and A A Tishchenko. Transition radiation from graphene. *Journal of Physics: Conference Series*, 357:012015, may 2012. 8

[70] Xihang Shi, Xiao Lin, Ido Kaminer, Fei Gao, Zhaoju Yang, John D. Joannopoulos, Marin Soljacic, and Baile Zhang. Superlight inverse doppler effect. *Nature Physics*, 14(10):1001–1005, 2018. 8

[71] Kamran Akbari, Silvina Segui, Zoran L. Mišković, Juana L. Gervasoni, and Néstor R. Arista. Energy losses and transition radiation in graphene traversed by a fast charged particle under oblique incidence. *Phys. Rev. B*, 98:195410, Nov 2018. 8, 39, 113, 117, 120

[72] Kamran Akbari, Zoran L. Miskovic, Silvina Segui, Juana L. Gervasoni, and Nestor R. Arista. Energy losses and transition radiation in multilayer graphene traversed by a fast charged particle. *ACS Photonics*, 4(8):1980–1992, 2017. 8, 36, 39

[73] Kamran Akbari, Zoran L Miskovic, Silvina Segui, Juana L Gervasoni, and Nestor R Arista. Excitation of hybridized dirac plasmon polaritons and transition radiation in multilayer graphene traversed by a fast charged particle. *Nanotechnology*, 29(22):225201, 2018. 8, 39

[74] Kamran Akbari, Silvina Segui, Juana L. Gervasoni, Zoran L. MiÅkovi, and Nstor R. Arista. Production of plasmons in two layers of graphene with different doping densities traversed by swift electrons. *Applied Surface Science*, 446:191 – 195, 2018. 12th International Conference on Surfaces, Coatings and Nanostructured Materials - NANOSMAT 2017. 8

[75] S.C. Liou, C.-S. Shie, C.H. Chen, R. Breitwieser, W.W. Pai, G.Y. Guo, and M.-W. Chu. π-plasmon dispersion in free-standing graphene by momentum-resolved electron energy-loss spectroscopy. *Phys. Rev. B*, 91:045418, 2015. 8, 34, 35, 106, 107, 108

[76] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, January 1927. 11

[77] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., 1996. 11, 12

[78] C. David Sherrill and III Schaefer, Henry F. The Configuration Interaction Method. *Advances in Quantum Chemistry*, 34:143–269, January 1999. 12

[79] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules (International Series of Monographs on Chemistry)*. Oxford University Press, USA, 1994. 13, 15, 16, 25

[80] B. Y. Tong and L. J. Sham. Application of a self-consistent scheme including exchange and correlation effects to atoms. *Phys. Rev.*, 144:1–4, Apr 1966. 13

[81] Randolph Q. Hood, M. Y. Chou, A. J. Williamson, G. Rajagopal, R. J. Needs, and W. M. C. Foulkes. Quantum monte carlo investigation of exchange and correlation in silicon. *Phys. Rev. Lett.*, 78:3350–3353, Apr 1997. 15, 62

[82] F. Zahariev, S. S. Leang, and Mark S. Gordon. Functional derivatives of meta-generalized gradient approximation (meta-gga) type exchange-correlation density functionals. *The Journal of Chemical Physics*, 138(24):244108, 2013. 15, 16

[83] Carsten Rostgaard. Exact exchange in density functional calculations: An implementation in the projector augmented wave method. Master's thesis, Technical University of Denmark, Kongens Lyngby, Denmark, 2006. 16, 17, 27, 32

[84] Alex P. Gaiduk, Ilya G. Ryabinkin, and Viktor N. Staroverov. Modified slater exchange potential with correct uniform electron gas limit. *Canadian Journal of Chemistry*, 93(1):91–97, 2015. 16

[85] Mel Levy and Federico Zahariev. Ground-state energy as a simple sum of orbital energies in kohn-sham theory: A shift in perspective through a shift in potential. *Physical Review Letters*, 113(11), Sep 2014. 16, 64, 120

[86] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. *Phys. Rev.*, 139:A796–A823, Aug 1965. 19, 32

[87] Alexei A. Kananenka, Sviataslau V. Kohut, Alex P. Gaiduk, Ilya G. Ryabinkin, and Viktor N. Staroverov. Efficient construction of exchange and correlation potentials by inverting the kohnsham equations. *The Journal of Chemical Physics*, 139(7):074112, 2013. 19, 20

[88] Sten Haastrup, Mikkel Strange, Mohnish Pandey, Thorsten Deilmann, Per S Schmidt, Nicki F Hinsche, Morten N Gjerding, Daniele Torelli, Peter M Larsen, Anders C Riis-Jensen, Jakob Gath, Karsten W Jacobsen, Jens Jørgen Mortensen, Thomas Olsen, and Kristian S Thygesen. The computational 2d materials database: high-throughput modeling and discovery of atomically thin crystals. *2D Materials*, 5(4):042002, sep 2018. 22, 78

[89] Livia Noemi Glanzmann. *Modelling of polymer-carbon nanotube heterojunctions for photovoltaic applications*. PhD thesis, Universidad del País Vasco (UPV/EHU), January 2017. 23, 24, 54, 55

[90] Stephen L. Adler. Quantum theory of the dielectric constant in real solids. *Phys. Rev.*, 126:413–420, Apr 1962. 24

[91] Nathan Wiser. Dielectric constant with local field effects included. *Phys. Rev.*, 129:62–69, Jan 1963. 24, 25

[92] Mark S. Hybertsen and Steven G. Louie. Ab initio static dielectric matrices from the density-functional approach. i. formulation and application to semiconductors and insulators. *Phys. Rev. B*, 35:5585–5601, Apr 1987. 26

[93] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley. *ScaLAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997. 27

[94] R. Del Sole and Raffaello Girlanda. Optical properties of semiconductors within the independent-quasiparticle approximation. *Phys. Rev. B*, 48:11789–11795, Oct 1993. 28

[95] Lorin X. Benedict, Steven G. Louie, and Marvin L. Cohen. Static polarizabilities of single-wall carbon nanotubes. *Phys. Rev. B*, 52:8541–8549, Sep 1995. 29, 54, 56

[96] V. Despoja, D. Novko, K. Dekanic, M. Sunjic., and L. Marusic. Two-dimensional and π plasmon spectra in pristine and doped graphene. *Phys. Rev. B*, 87:075447, 2013. 30, 51

[97] René Petersen, Thomas Garm Pedersen, and F. Javier García de Abajo. Nonlocal plasmonic response of doped and optically pumped graphene, mos_2, and black phosphorus. *Phys. Rev. B*, 96:205430, Nov 2017. 30, 143

[98] M. Stöger-Pollach, H. Franco, P. Schattschneider, S. Lazar, B. Schaffer, W. Grogger, and H.W. Zandbergen. Cerenkov losses: A limit for bandgap determination and kramerskronig analysis. *Micron*, 37(5):396 – 402, 2006. 31, 115, 117

[99] Martin Dressel. *Electrodynamics of solids: optical properties of electrons in matter*. Cambridge University Press, (2002). 31, 139

[100] Huijuan Sun, Duncan J. Mowbray, Annapaola Migani, Jin Zhao, Hrvoje Petek, and Angel Rubio. Comparing quasiparticle H_2O level alignment on anatase and rutile TiO_2. *ACS Catal.*, 5(7):4242–4254, 2015. 32, 78, 98, 99, 100, 101

[101] Falco Hüser, Thomas Olsen, and Kristian S. Thygesen. Quasiparticle *GW* calculations for solids, molecules, and two-dimensional materials. *Phys. Rev. B*, 87:235132, Jun 2013. 32, 78

[102] R. W. Godby and R. J. Needs. Metal-insulator transition in kohn-sham theory and quasiparticle theory. *Phys. Rev. Lett.*, 62:1169–1172, Mar 1989. 32, 78

[103] A. Oschlies, R. W. Godby, and R. J. Needs. Gw self-energy calculations of carrier-induced band-gap narrowing in n-type silicon. *Phys. Rev. B*, 51:1527–1535, Jan 1995. 32, 78

[104] Paul Larson, Marc Dvorak, and Zhigang Wu. Role of the plasmon-pole model in the *gw* approximation. *Phys. Rev. B*, 88:125205, Sep 2013. 32, 78

[105] A. Migani, D. J. Mowbray, J. Zhao, H. Petek, and A. Rubio. Quasiparticle level alignment for photocatalytic interfaces. *J. Chem. Theory Comput.*, 10:2103–2114, 2014. 32, 78, 98, 100, 101

[106] Philipp Wachsmuth. *Momentum-resolved electron energy-loss spectroscopy of graphene*. PhD thesis, Ulm University), Mar 2014. 33, 35, 36

[107] P.A. Midgley. A simple new method to obtain high angular resolution $\omega - q$ patterns. *Ultramicroscopy*, 76:91–96, 1999. 33, 34, 35, 107

[108] Ryosuke Senga, Kazu Suenaga, Paolo Barone, Shigeyuki Morishita, Francesco Mauri, and Thomas Pichler. Position and momentum mapping of vibrations in graphene nanostructures. *Nature*, 573(7773):247–250, 2019. 34

[109] R F Egerton. Electron energy-loss spectroscopy in the tem. *Reports on Progress in Physics*, 72(1):016502, 2009. 38

[110] Simon Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, USA, 1st edition, 1994. 41, 42

[111] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 42, 44, 46

[112] Martn Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016. 42, 46, 60

[113] Vinod Nair and Geoffrey Hinton. Rectified linear units improve restricted boltzmann machines vinod nair. volume 27, pages 807–814, 06 2010. 42

[114] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., USA, 1995. 42, 45

[115] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. http://www.deeplearningbook.org. Accessed: 2019-04-19. 43, 44, 46, 47, 62

[116] Ning Qian. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 12(1):145 – 151, 1999. 43

[117] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 edition, 2007. 44

[118] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014. 46

[119] Ashia C. Wilson, Rebecca Roelofs, Mitchell Stern, Nathan Srebro, and Benjamin Recht. The marginal value of adaptive gradient methods in machine learning, 2017. 46

[120] H. Kroto. The birth of c60: Buckminsterfullerene. In Hans Kuzmany, Jörg Fink, Michael Mehring, and Siegmar Roth, editors, *Electronic Properties of Fullerenes*, pages 1–7, Berlin, Heidelberg, 1993. Springer Berlin Heidelberg. 48

[121] D. J. Mowbray. Theoretical electron energy loss spectroscopy of isolated graphene. *Phys. Status Solidi B*, 251:2509–2514, 2014. 51, 78

[122] L. A. Falkovsky and S. S. Pershoguba. Optical far-infrared properties of a graphene monolayer and multilayer. *Phys. Rev. B*, 76:153410, Oct 2007. 51, 52

[123] E.G. Mishchenko. Dynamic conductivity in graphene beyond linear response. *Phys. Rev. Lett.*, 103:246802, 2009. 52

[124] Frank H. L. Koppens, Darrick E. Chang, and F. Javier García de Abajo. Graphene plasmonics: A platform for strong light-matter interactions. *Nano Letters*, 11(8):3370–3377, 2011. 52

[125] S. Das Sarma, Shaffique Adam, E. H. Hwang, and Enrico Rossi. Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.*, 83:407–470, May 2011. 52

[126] Qiaoliang Bao and Kian Ping Loh. Graphene photonics, plasmonics, and broadband optoelectronic devices. *ACS Nano*, 6(5):3677–3694, 2012. 52

[127] Likai Li, Jonghwan Kim, Chenhao Jin, Guo Jun Ye, Diana Y. Qiu, Felipe H. da Jornada, Zhiwen Shi, Long Chen, Zuocheng Zhang, Fangyuan Yang, Kenji Watanabe, Takashi Taniguchi, Wencai Ren, Steven G. Louie, Xian Hui Chen, Yuanbo Zhang, and Feng Wang. Direct observation of the layer-dependent electronic structure in phosphorene. *Nat. Nanotechnol.*, 12:21, Sep 2016. 53, 97, 98

[128] Andrei Nemilentsau, Tony Low, and George Hanson. Anisotropic 2d materials for tunable hyperbolic plasmonics. *Phys. Rev. Lett.*, 116:066804, Feb 2016. 54

[129] Etienne Gaufrès, Frédéric Fossard, Vincent Gosselin, Lorenzo Sponza, François Ducastelle, Zhenglu Li, Steven G. Louie, Richard Martel, Michel Côté, and Annick Loiseau. Momentum-resolved dielectric response of free-standing mono-, bi-, and trilayer black phosphorus. *Nano Letters*, 19(11):8303–8310, Nov 2019. 54, 120

[130] Catalin D. Spataru, Sohrab Ismail-Beigi, Lorin X. Benedict, and Steven G. Louie. Excitonic effects and optical spectra of single-walled carbon nanotubes. *Phys. Rev. Lett.*, 92:077402, Feb 2004. 54, 78

[131] Allan L Smith. Comparison of the ultraviolet absorption cross section of buckminsterfullerene in the gas phase and in hexane solution. *J. Phys. B*, 29(21):4975–4980, nov 1996. 56, 80, 81, 82

[132] J. P. Hare, H. W. Kroto, and R. Taylor. Reprint of: Preparation and uv/visible spectra of fullerenes C_{60} and C_{70}. *Chem. Phys. Lett.*, 589:57 – 60, 2013. 56, 81

[133] Vito Despoja and Duncan John Mowbray. Using surface plasmonics to turn on fullerene's dark excitons. *Phys. Rev. B*, 89:195433, 2014. 56, 80, 81

[134] Duncan John Mowbray and Vito Despoja. Tailoring a molecule's optical absorbance using surface plasmonics. *The Journal of Physical Chemistry C*, 123(43):26498–26508, Oct 2019. 56, 78, 80, 81, 82

[135] Hugo Scheer. *Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications*, volume 25 of *Advances in Photosynthesis and Respiration*, chapter 1, pages 1–26. Springer, Dordrecht, The Netherlands, 2006. 56, 57, 58

[136] Gerard P. Moss. Nomenclature of tetrapyrroles. *Euro. J. Biochem.*, 178(2):277–328, 1988. 57

[137] Bruce Forbes Milne, Yoni Toker, Angel Rubio, and Steen Brøndsted Nielsen. Unraveling the intrinsic color of chlorophyll. *Angew. Chem., Int. Ed. Engl.*, 54(7):2170–2173, February 2015. 56, 57, 58, 77, 82, 85, 86

[138] Bruce F. Milne, Christina Kjær, Jørgen Houmøller, Mark H. Stockett, Yoni Toker, Angel Rubio, and Steen Brøndsted Nielsen. On the exciton coupling between two chlorophyll pigments in the absence of a protein environment: Intrinsic effects revealed by theory and experiment. *Angew. Chem., Int. Ed. Engl.*, 55:6248–6251, 2016. 56, 58, 77, 82

[139] Joaquim Jornet-Somoza, Joseba Alberdi-Rodriguez, Bruce Forbes Milne, Xavier Andrade, Miguel A. L. Marques, Fernando Nogueira, Micael J. T. Oliveira, James J. P. Stewart, and Angel Rubio. Insights into colour-tuning of chlorophyll optical response in green plants. *Phys. Chem. Chem. Phys.*, 17:26599–26606, 2015. 58, 82

[140] Andreas Dreuw and Martin Head-Gordon. Failure of Time-Dependent density functional theory for Long-Range Charge-Transfer excited states: The Zincbacteriochlorin–Bacteriochlorin and Bacteriochlorophyll–Spheroidene complexes. *J. Am. Chem. Soc.*, 126(12):4007–4016, March 2004. 58, 82

[141] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251 – 257, 1991. 60

[142] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989. 60

[143] Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of adam and beyond. In *International Conference on Learning Representations*, 2018. 68

[144] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization, 2017. 68

[145] Vito Despoja, Tijana Djordjević, Lazar Karbunar, Ivan Radović, and Zoran L. Mišković. Ab initio study of the electron energy loss function in a graphene-sapphire-graphene composite system. *Phys. Rev. B*, 96:075433, Aug 2017. 76, 143

[146] Chen Pai-Yen, Argyropoulos Christos, Farhat Mohamed, and Gomez-Diaz J. Sebastian. Flatland plasmonics and nanophotonics based on graphene and beyond. *Nanophotonics*, 6:1239, 2017. 6. 76

[147] S. R. Bahn and K. W. Jacobsen. An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.*, 4:56, 2002. 77

[148] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jn Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S. Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W. Jacobsen. The atomic simulation environment—a python library for working with atoms. *J. Phys.: Condens. Matter*, 29(27):273002, 2017. 77

[149] Mikkel Strange, Iben S. Kristensen, Kristian S. Thygesen, and Karsten W. Jacobsen. Benchmark density functional theory calculations for nanoscale conductance. *J. Chem. Phys.*, 128:114714, 2008. 77

[150] Hendrik J Monkhorst and James D Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188, 1976. 77, 78

[151] Carlo A. Rozzi, Daniele Varsano, Andrea Marini, Eberhard K. U. Gross, and Angel Rubio. Exact coulomb cutoff technique for supercell calculations. *Phys. Rev. B*, 73:205119, May 2006. 78

[152] Bruno Linder and Suheil Abdulnur. Solvent effects on electronic spectral intensities. *J. Chem. Phys.*, 54(4):1807–1814, 1971. 80

[153] J. Berkowitz. Sum rules and the photoabsorption cross sections of C_60. *J. Chem. Phys.*, 111(4):1446–1453, 1999. 81

[154] Mark H. Stockett, Lihi Musbat, Christina Kjær, Jørgen Houmøller, Yoni Toker, Angel Rubio, Bruce Forbes Milne, and Steen Brønsted Nielsen. The soret absorption band of isolated chlorophyll *a* and *b* tagged with quaternary ammonium ions. *Phys. Chem. Chem. Phys.*, 17:25793–25798, 2015. 82, 85

[155] Leila Hedayatifar, Elnaz Irani, Mahmood Mazarei, Soroush Rasti, Yavar T. Azar, Ali T. Rezakhani, Alireza Mashaghi, Farzaneh Shayeganfar, Mehrnaz Anvari, Tiam Heydari, Ali Rahimi Tabar, Nasser Nafari, Mohammad Ali Vesaghi, Reza Asgari, and Mohammad Reza Rahimi Tabar. Optical absorption and electronic spectra of chlorophylls a and b. *RSC Adv.*, 6:109778–109785, 2016. 82

[156] Juha Linnanto and Jouko Korppi-Tommola. Quantum chemical simulation of excited states of chlorophylls, bacteriochlorophylls and their complexes. *Phys. Chem. Chem. Phys.*, 8:663–687, 2006. 82

[157] Andreas B. J. Parusel and Stefan Grimme. A theoretical study of the excited states of chlorophyll a and pheophytin a. *J. Phys. Chem. B*, 104(22):5395–5398, 2000. 82

[158] Carolin Konig and Johannes Neugebauer. First-principles calculation of electronic spectra of light-harvesting complex ii. *Phys. Chem. Chem. Phys.*, 13:10475–10490, 2011. 82

[159] P M Echenique and J B Pendry. The existence and detection of rydberg states at surfaces. *J. Phys. C*, 11(10):2065–2075, may 1978. 93

[160] Keenan Lyon. Analysis of plasmons sustained on the surface of graphene. Master's thesis, University of Waterloo, Waterloo, Canada, 2014. 97

[161] Jenny Schneider, Masaya Matsuoka, Masato Takeuchi, Jinlong Zhang, Yu Horiuchi, Masakazu Anpo, and Detlef W. Bahnemann. Understanding TiO_2 photocatalysis: Mechanisms and materials. *Chem. Rev.*, 114:9919–9986, 2014. 98

[162] Noriko Hosaka, Takao Sekiya, Chikatoshi Satoko, and Susumu Kurita. Optical properties of single-crystal anatase TiO_2. *J. Phys. Soc. Jpn.*, 66(3):877–880, 1997. 98, 99, 101

[163] M Landmann, E Rauls, and W G Schmidt. The electronic structure and optical response of rutile, anatase and brookite TiO_2. *J. Phys.: Condens. Matter*, 24(19):195503, apr 2012. 99, 101

[164] Letizia Chiodo, Juan Maria García-Lastra, Amilcare Iacomino, Stefano Ossicini, Jin Zhao, Hrvoje Petek, and Angel Rubio. Self-energy and excitonic effects in the electronic and optical properties of TiO_2 crystalline phases. *Phys. Rev. B*, 82:045207, Jul 2010. 99, 100, 101

[165] Manuel Cardona and Gunther Harbeke. Optical properties and band structure of wurtzite-type crystals and rutile. *Phys. Rev.*, 137:A1467–A1476, Mar 1965. 98, 100, 101

[166] Mathias Schubert Thomas E. Tiwald. Measurement of rutile TiO_2 dielectric tensor from 0.148 to 33 micrometres using generalized ellipsometry. *Proc. SPIE*, 4103, 2000. 98, 100, 101

[167] A. Politano, G. Chiarello, and C. Spinella. Plasmon spectroscopy of graphene and other two-dimensional materials with transmission electronmicroscopy. *Materials Science in Semiconductor Processing*, 65:88–99, 2017. 102

[168] J.-C. Idrobo and W. Zhou. A short story of imaging and spectroscopy of two- dimensional materials by scanning transmission electron microscopy. *Ultramicroscopy*, 180:156–162, 2017. 102

[169] Ondrej L. Krivanek, Tracy C. Lovejoy, Niklas Dellby, Toshihiro Aoki, R. W. Carpenter, Peter Rez, Emmanuel Soignard, Jiangtao Zhu, Philip E. Batson, Maureen J. Lagos, Ray F. Egerton, and Peter A. Crozier. Vibrational spectroscopy in the electron microscope. *Nature*, 514(7521):209–212, 2014. 102, 113

[170] Fredrik S. Hage, Rebecca J. Nicholls, Jonathan R. Yates, Dougal G. McCulloch, Tracy C. Lovejoy, Niklas Dellby, Ondrej L. Krivanek, Keith Refson, and Quentin M. Ramasse. Nanoscale momentum-resolved vibrational spectroscopy. *Science Advances*, 4(6), 2018. 102

[171] Jordan A. Hachtel, Andrew R. Lupini, and Juan Carlos Idrobo. Exploring the capabilities of monochromated electron energy loss spectroscopy in the infrared regime. *Scientific Reports*, 8(1):5637, 2018. 102

[172] M.K. Kinyanjui, C. Kramberger, and T. Pichler. Direct probe of linearly dispersing 2d interband plasmons in a free-standing graphene monolayer. *Europhys. Lett.*, 97:57005, 2012. 106

[173] Saeed H. Abedinpour, G. Vignale, A. Principi, Marco Polini, Wang-Kong Tse, and A. H. MacDonald. Drude weight, plasmon dispersion, and ac conductivity in doped graphene sheets. *Phys. Rev. B*, 84:045429, Jul 2011. 109

[174] Marco Cazzaniga, Lucia Caramella, Nicola Manini, and Giovanni Onida. Ab initio intra-band contributions to the optical properties of metals. *Phys. Rev. B*, 82:035104, Jul 2010. 109

[175] Andrea Marini, Giovanni Onida, and Rodolfo Del Sole. Plane-wave dft-lda calculation of the electronic structure and absorption spectrum of copper. *Phys. Rev. B*, 64:195125, Oct 2001. 109

[176] M. Stöger-Pollach and P. Schattschneider. The influence of relativistic energy losses on bandgap determination using valence eels. *Ultramicroscopy*, 107(12):1178 – 1185, 2007. 114, 115, 117

[177] Rolf Erni and Nigel D. Browning. The impact of surface and retardation losses on valence electron energy-loss spectroscopy. *Ultramicroscopy*, 108(2):84 – 99, 2008. 115, 117

[178] I. Villo-Perez, Z. L. Miskovic, and N. R. Arista. *Plasmon Spectra of Nano-Structures: A Hydrodynamic Model, in Trends in Nanophysics, V. Barsan and A. Aldea (eds)*, chapter 4, pages 217–254. Springer, Berlin, Heidelberg, 2010. 115

[179] Myron D. Kapetanakis, Wu Zhou, Mark P. Oxley, Jaekwang Lee, Micah P. Prange, Stephen J. Pennycook, Juan Carlos Idrobo, and Sokrates T. Pantelides. Low-loss electron energy loss spectroscopy: An atomic-resolution complement to optical spectroscopies and application to graphene. *Phys. Rev. B*, 92:125147, Sep 2015. 120

[180] John David Jackson and John D Jackson. *Classical electrodynamics*, volume 3. Wiley New York etc., 1962. 138

APPENDICES

Appendix A

Electromagnetic Theory and Material Response Functions

When dealing with materials with a high density of free charge carriers, resulting in an almost continuous spectrum of electron energy levels [180], we are justified in using the classical electromagnetic framework. Maxwell's equations in Gaussian units are below, with boldface representing vector quantities in 3D space and time-dependence.

$$\nabla \cdot \mathbf{D} = 4\pi\rho_{ext} \tag{A.1}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \tag{A.2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{A.3}$$

$$\nabla \times \mathbf{H} = \frac{1}{c}\left(4\pi\mathbf{J}_{ext} + \frac{\partial \mathbf{D}}{\partial t}\right) \tag{A.4}$$

where, for isotropic materials, $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$, ε and μ are the relative permittivities that describe the response of a given material to electric and magnetic fields respectively, ρ_{ext} is the external charge density, $\mathbf{J}_{ext}$ is the external current density, $\mathbf{D}$ is the electric displacement vector, $\mathbf{H}$ is the magnetizing field, $\mathbf{B}$ is the magnetic field, and $\mathbf{E}$ is the electric field. Throughout this work we assume non-magnetic ($\mu = 1$) materials. External contributions (e.g. from incident light) refer to excitations that cause a change in the system, while anything labelled as induced respond to these stimuli.

In addition to Maxwell's equations, these quantities are related by constitutive relations, relating the different functions that describe the movement of free and bound charge carriers and

external excitations

$$\mathbf{J}_{ind}(\mathbf{r},t) = \int d\mathbf{r}'dt'\sigma(\mathbf{r}-\mathbf{r}',t-t')\mathbf{E}(\mathbf{r}',t') \tag{A.5}$$

$$\mathbf{D}(\mathbf{r},t) = \frac{1}{4\pi}\int d\mathbf{r}'dt'\varepsilon(\mathbf{r}-\mathbf{r}',t-t')\mathbf{E}(\mathbf{r}',t'), \tag{A.6}$$

where the dielectric function $\varepsilon(\mathbf{r},t)$ relates the electric field induced by free charges to the total electric field, and $\sigma(\mathbf{r},t)$ describes the linear conductivity, relating the current density to the electric field. These response functions are presumed valid for linear media with little temporal or spatial dispersion [99]. Applying the Fourier transform $\int dtd\mathbf{r}e^{-i(\mathbf{q}\cdot\mathbf{r}-\omega t)}(\cdots)$ gives the fields in terms of their plane-wave components

$$\mathbf{J}_{ind}(\mathbf{q},\omega) = \sigma(\mathbf{q},\omega)\mathbf{E}(\mathbf{q},\omega) \tag{A.7}$$

$$\mathbf{D}(\mathbf{q},\omega) = \frac{1}{4\pi}\varepsilon(\mathbf{q},\omega)\mathbf{E}(\mathbf{q},\omega). \tag{A.8}$$

These electric and magnetic fields can be reformulated in terms of a vector potential $\mathbf{A}$ and scalar potential ϕ

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla\times\mathbf{A}, \tag{A.9}$$

where

$$\phi(\mathbf{r},t) = \int\frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' \tag{A.10}$$

$$\mathbf{A}(\mathbf{r},t) = \frac{1}{c}\int\frac{\mathbf{J}(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'. \tag{A.11}$$

The electric field can be reformulated in Fourier space as

$$\mathbf{E} = -i\mathbf{q}\phi + i\omega\mathbf{A}. \tag{A.12}$$

The vector potential $\mathbf{A}$ is gauge invariant, so we are able to choose $\nabla\cdot\mathbf{A} = 0$ within what is called the Coulomb gauge, and in this gauge the vector potential disappears from the divergence of the electric field. It is important to note that any equations derived this way will only be valid within this gauge, or if we invoke an electrostatic approximation where $\partial\mathbf{A}/\partial t = 0$ always.

A third response function, the polarizability χ, relates the internal charge density to the total electric potential, and in wavenumber-frequency space goes as

$$\rho_{ind} = \chi(\mathbf{q},\omega)\phi \tag{A.13}$$

with $(\mathbf{q},\omega)$ included in all other variables implicitly. This is also often called the density-density response function.

The response functions $\sigma, \chi, \varepsilon$ are all intertwined, but as the dimensionality of the material in question changes, so does the dimensionality of σ and χ, while ε relates different electric fields and therefore remains dimensionless. To derive a general relation between $\sigma(\mathbf{q}, \omega)$ and $\chi(\mathbf{q}, \omega)$, we start with the continuity equation in Fourier space

$$\omega \rho_{ind} = \mathbf{q} \cdot \mathbf{J}_{ind} = \sigma \mathbf{q} \cdot \mathbf{E}, \tag{A.14}$$

which combined with Eqs. A.12,A.13 yields

$$\sigma(\mathbf{q}, \omega) = \frac{i\omega}{q^2} \chi(\mathbf{q}, \omega), \tag{A.15}$$

where $q = |\mathbf{q}|$, noting that these wavenumbers have units in accordance with the dimensionality of the problem. The above relation between the conductivity σ and the polarizability χ is valid in both 2D and 3D. Another relation between the polarizability χ and the dielectric function ε valid in any dimension goes as

$$\varepsilon(\mathbf{q}, \omega) = 1 - v(q)\chi(\mathbf{q}, \omega) \tag{A.16}$$

where $v(q)$ is the unscreened Coulomb potential, which changes to reflect the restricted motion of electrons in lower-dimensional systems

$$v^{3D}(q) = \frac{4\pi}{q^2} \qquad v^{2D}(q) = \frac{2\pi}{q}. \tag{A.17}$$

This relation can be determined from separating out the potential due to external and internal contributions

$$\phi = \phi_{ext} + \phi_{ind} = \phi_{ext} + v(q)\chi(\mathbf{q}, \omega)\phi \rightarrow \phi = \frac{\phi_{ext}}{1 - v(q)\chi(\mathbf{q}, \omega)} = \frac{\phi_{ext}}{\epsilon(\mathbf{q}, \omega)}. \tag{A.18}$$

Rearranging the above equation yields the useful relation

$$\phi_{ind} = (1 - \varepsilon(\mathbf{q}, \omega))\phi = v(q)\chi(\mathbf{q}, \omega)\phi = v(q)\rho_{ind}. \tag{A.19}$$

The relation in the Coulomb gauge between ε and σ in three dimensions therefore goes as

$$\varepsilon_{3D}(\mathbf{q}, \omega) = 1 - \frac{4\pi}{q^2}\chi_{3D}(\mathbf{q}, \omega) = 1 - \frac{4\pi}{q^2}\left(\frac{q^2}{i\omega}\right)\sigma_{3D}(\mathbf{q}, \omega) = 1 + \frac{4\pi i \sigma_{3D}(\mathbf{q}, \omega)}{\omega} \tag{A.20}$$

while the two dimensional Coulomb potential goes as

$$\varepsilon_{2D}(\mathbf{q}, \omega) = 1 + \frac{2\pi i q}{\omega}\sigma_{2D}(\mathbf{q}, \omega) \tag{A.21}$$

Appendix B

Hydrodynamic Approach for the Conductivity of Graphene

To derive an empirical model for the response of the charge carriers in graphene requires an understanding of how small perturbations, $\delta n(\mathbf{r}, t)$ for the electron density $n(\mathbf{r}, t)$ and $\delta \mathbf{J}(\mathbf{r}, t)$ for the current density $\mathbf{J}(\mathbf{r}, t)$, respond to external sources. We first assume $q \ll k_F = \sqrt{\pi n_0}$, the characteristic Fermi wavenumber of graphene, where n_0 corresponds to the average areal electron density of graphene, which increases with doping. This limit ensures that the response of electrons is macroscopic in the model.

In the elementary Euler-flow hydrodynamic equation for ordinary neutral liquids, oscillations in the particle density are caused by a pressure gradient, leading to a restoring force

$$m\frac{\partial \mathbf{J}(\mathbf{r}, t)}{\partial t} = -\nabla P(\mathbf{r}, t), \tag{B.1}$$

where $P(\mathbf{r}, t)$ is the pressure and m is the mass of a particle in the fluid. In addition to the macroscopic response condition, this equation will work best with slow oscillations so that the system is always in a state of near equilibrium. If this pressure is induced by the density, as is the case for ordinary gases and fluids, a linear approximation for the pressure gradient goes as

$$\nabla P(\mathbf{r}, t) \approx \frac{\partial P}{\partial n}\nabla \delta n(\mathbf{r}, t). \tag{B.2}$$

Taking the divergence of the equation and invoking the continuity equation $\nabla \cdot \mathbf{J} = -\partial n/\partial t$ gives

$$\frac{\partial^2 \delta n(\mathbf{r}, t)}{\partial t^2} - \frac{1}{m}\frac{\partial P}{\partial n}\nabla^2 \delta n(\mathbf{r}, t) = 0, \tag{B.3}$$

which a transformation to Fourier space shows that $\omega = \sqrt{\frac{1}{m}\frac{\partial P}{\partial n}}q$ is a resonance condition, revealing firsthand that similar models based on the hydrodynamic approach will yield similar

resonances, even in the optical $q \to 0$ limit.

To utilize the hydrodynamic framework for charged liquids, the main driving force will be the long-wave electromagnetic repulsion rather than a weakly varying pressure gradient. Assuming in equilibrium all non-perturbed charges have equal and opposing forces, this repulsion will come in the form of an integration over all the charges under the Coulomb potential influence, giving a combined electric force

$$m_\nu \frac{\partial \delta \mathbf{J}_\nu(\mathbf{r}, t)}{\partial t} = -n_\nu^0 \nabla_\mathbf{r} \int d\mathbf{r}' \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \delta n_\nu(\mathbf{r}', t) \tag{B.4}$$

where m_ν is the effective mass and n_ν^0 is the equilibrium charge density of charge carrier type ν. The influence of screening can be included in this equation by adjusting to the screened Coulomb potential. To include an electric external force into this equation, a term that depends on $\nabla^2 \phi_{ext}$ according to Eq. A.12 can be included in the combined electric force. Switching to Fourier space and again using the continuity equation shown in Eq. A.14 results in

$$\left(\omega^2 - \frac{n_\nu^0 q^2}{m_\nu} v_{2\mathrm{D}}(q)\right) \delta n_\nu(\mathbf{q}, \omega) = \left(\omega^2 - \frac{n_\nu^0 q^2}{m_\nu} \frac{2\pi e^2}{q}\right) \delta n_\nu(\mathbf{q}, \omega) = -\frac{n_\nu^0 q^2}{m_\nu} \phi_{ext}, \tag{B.5}$$

where $v_{2\mathrm{D}}(\mathbf{q})$ is the 2D Coulomb potential. As with the original Euler flow equation, a resonant frequency of $\omega = \sqrt{2\pi n_\nu^0 e^2 q/(m_\nu)}$ appears in the case of no external electric field, corresponding to square root dependence on wavenumber in the plasmon frequency, as expected for charged fluids in 2D. For low-energy excitations in graphene, the effective electron/hole mass is given by the so-called plasmon mass, $m_\nu = \hbar k_F / v_F$, where $v_F \approx c/300$ is the Fermi speed from the Dirac cone approximation of the π electron bands near the K points of the Brillouin zone.

To compute the high-energy conductivity response for graphene, Eq. B.5 can be adjusted by treating the electrons within a harmonic approximation for fluids, introducing a restoring frequency $\omega_{\nu r}$ that can be included as a force acting on the charge density perturbation. Eq. B.5 then becomes

$$\left(\omega^2 - \frac{n_\nu^0 q^2}{m_\nu} v_{2\mathrm{D}}(q)\right) \delta n_\nu(\mathbf{q}, \omega) = \omega_{\nu r}^2 \delta n_\nu(\mathbf{q}, \omega) - \frac{n_\nu^0 q^2}{m_\nu} \phi_{ext}, \tag{B.6}$$

which using Eq. A.19 can be rearranged to yield

$$\delta n_\nu(\mathbf{q}, \omega) = \frac{n_\nu^0 q^2 / m_\nu}{\omega_{\nu r}^2 - \omega(\omega + i\gamma_\nu)} \phi = \chi_\nu(q, \omega) \phi, \tag{B.7}$$

where the dampening factor γ_ν is introduced to account for dissipative effects for fluid ν. In total, in constructing an empirical model for graphene, the factors n_ν^0, γ_ν, m_ν, and $\omega_{\nu r}$ are free parameters, along with the freedom to choose the number of different charge carriers, although all these choices should ideally be physically motivated. Considering that $\chi_\nu(q, \omega)$ is a linear response function, multiple types of charge carriers can be included by simply adding the individual polarizabilities of each charge carrier.

Appendix C

Derivation of Intraband Conductivity for *ab initio* Calculations

While alternative expressions exist for the optical intraband conductivity, it is not immediately clear that they are equivalent to the expression derived in Eq. 2.71. Petersen [97] states that σ_{intra} can be written as

$$\sigma_{intra}(\omega) = \frac{1}{\omega + i\gamma}\frac{-2ie^2}{4\pi^2\hbar^2}\sum_n \int_{BZ} f'(E_n(\mathbf{k}))(\nabla_{\mathbf{k}} E_n \cdot \hat{q})^2 d\mathbf{k}. \tag{C.1}$$

Despoja [145] also derives an expression for the intraband conductivity,

$$\sigma_{intra}(\omega) = \frac{1}{\omega + i\gamma}\frac{-2i}{\Omega}\sum_{\mathbf{k},n} f'(E_n(\mathbf{k}))\left|j^x_{n\mathbf{k},n\mathbf{k}}(\mathbf{G}=0)\right|^2, \tag{C.2}$$

where the current vertices are given by

$$j^{\mu}_{n\mathbf{k},m\mathbf{k}+\mathbf{q}}(\mathbf{G}) = \frac{\hbar e}{2im}\int_{\Omega} d\mathbf{r} e^{-i(\mathbf{q}+\mathbf{G})\cdot r}\left[\psi^*_{n\mathbf{k}}(\mathbf{r})\left(\partial_\mu \psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})\right) - \left(\partial_\mu \psi^*_{n\mathbf{k}}(\mathbf{r})\right)\psi_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})\right]$$

and the $\hat{q} = \hat{x}$ direction has already been chosen. Simplifying by letting $m = n$, $\mathbf{G} = 0$, and $\mathbf{q} = 0$, and generalizing $\partial_\mu \longrightarrow \nabla$ gives

$$\mathbf{j}_{n\mathbf{k},n\mathbf{k}}(\mathbf{G}=0) = \frac{\hbar e}{2im}\int_{\Omega} d\mathbf{r}\left[\psi^*_{n\mathbf{k}}(\mathbf{r})\nabla\psi_{n\mathbf{k}}(\mathbf{r}) - \nabla\psi^*_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r})\right]. \tag{C.3}$$

Integration by parts (since the wavefunctions disappear in the $r \to \infty$ limit) leaves us with

$$\mathbf{j}_{n\mathbf{k},n\mathbf{k}}(\mathbf{G}=0) = \frac{\hbar e}{im}\int_{\Omega} d\mathbf{r}\psi^*_{n\mathbf{k}}(\mathbf{r})\nabla\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{\hbar e}{im}\langle\psi_n|\nabla|\psi_n\rangle. \tag{C.4}$$

The relation between this expression and $\nabla_{\mathbf{k}} E_n(\mathbf{k})$ can be found starting with Schrodinger's equation,

$$-\frac{\hbar^2}{2m}\nabla_r^2\psi(k,r) = (E(k) - V)\,\psi(k,r), \tag{C.5}$$

working in one dimension for simplicity. Using the fact that the solution is given by a Bloch wave, $\psi = e^{ikr}u(k,r)$ and $[\nabla_r, \nabla_k] = 0$, we apply ∇_k to the left side of the above equation,

$$\begin{aligned}-\frac{\hbar^2}{2m}\nabla_r^2(ir\psi(k,r) + e^{ik\cdot r}\nabla_k u(k,r)) = \\ (E(k) - V)\,(ir\psi(k,r) + e^{ik\cdot r}\nabla_k u(k,r)) + (\nabla_k E(k))\,\psi(k,r).\end{aligned} \tag{C.6}$$

Since $\nabla_r^2(r\psi) = r\nabla_r^2\psi + 2\nabla_r\psi$, we can collect terms to obtain

$$\begin{aligned}-\frac{\hbar^2}{2m}(ir\nabla_r^2\psi(k,r)) - \frac{i\hbar^2}{m}\nabla_r\psi = (E(k) - V)\,(ir\psi(k,r))+ \\ \left(E(k) - (-\frac{\hbar^2}{2m}\nabla_r^2 + V)\right)e^{ik\cdot r}\nabla_k u(k,r) + (\nabla_k E(k))\,\psi.\end{aligned} \tag{C.7}$$

The terms containing ir disappear due to being equivalent up to a factor to Eq. C.5. Multiplying by ψ^* on the left and integrating over all space will make the second term on the right disappear (since $\psi(k,r)$ is an eigenstate), so we are left with

$$-\frac{i\hbar^2}{m}\int dr\,(\psi^*\nabla_r\psi) = \nabla_k E(k)\int dr\psi^*(k,r)\psi(k,r) = \nabla_k E(k), \tag{C.8}$$

equivalently written as

$$\nabla_k E(k) = -\frac{i\hbar^2}{m}\langle\psi|\nabla|\psi\rangle. \tag{C.9}$$

Rewriting Eq.C.2 using Eq.C.4 and re-introducing the q-vector direction dependence gives

$$\sigma_{intra}(\omega) = \frac{1}{\omega + i\gamma}\frac{-2i}{\Omega}\sum_{\mathbf{k},n} f'(E_n(\mathbf{k}))\left|\frac{\hbar e}{im}\hat{\mathbf{q}}\cdot\langle\psi_n|\nabla|\psi_n\rangle\right|^2 \tag{C.10}$$

$$= \frac{1}{\omega + i\gamma}\frac{-2i\hbar^2e^2}{m^2\Omega}\sum_{\mathbf{k},n} f'(E_n(\mathbf{k}))\,|\hat{\mathbf{q}}\cdot\langle\psi_n|\nabla|\psi_n\rangle|^2\,. \tag{C.11}$$

Rewriting Eq. C.1 using Eq. C.9 and recalling that $\int_{BZ} \sim \frac{4\pi^2}{\Omega} \sum_{\mathbf{k}}$ in 2D yields

$$\sigma_{intra}(\omega) = \frac{1}{\omega + i\gamma} \frac{-2ie^2}{4\pi^2\hbar^2} \frac{4\pi^2}{\Omega} \left(\frac{-i\hbar^2}{m}\right)^2 \sum_{\mathbf{k},n} f'(E_n(\mathbf{k})) \left|\hat{q} \cdot \langle \psi_n | \nabla | \psi_n \rangle\right|^2 \tag{C.12}$$

$$= \frac{1}{\omega + i\gamma} \frac{-2i\hbar^2 e^2}{m^2 \Omega} \sum_{\mathbf{k},n} f'(E_n(\mathbf{k})) \left|\hat{q} \cdot \langle \psi_n | \nabla | \psi_n \rangle\right|^2 , \tag{C.13}$$

showing that both expressions for the intraband conductivity in the optical limit are equivalent to the expression shown in Eq. 2.71.

Appendix D

Performance of Neural Networks on Material Datasets

Name	NF 0.005	NF 0.001	NF 0.0005	Name	NF 0.005	NF 0.001	NF 0.0005
LiH	41.2	7.9	3.9	CO	153.2	29.4	14.6
BeH	32.7	7.1	5.6	HCO	155.3	30.1	15.7
CN	89.4	14.4	6.2	H_2CO	149.0	29.3	15.0
Be	32.7	6.8	5.0	CH_3OH	145.3	29.1	15.1
CH_3	25.3	7.3	5.7	N_2	120.1	19.1	9.0
CH_4	27.2	6.8	4.5	N_2H_4	77.2	14.3	9.1
NH_2	44.6	7.1	4.7	H_2O_2	235.5	46.1	23.1
NH_3	35.9	6.7	4.8	F_2	313.6	61.5	30.3
H_2O	119.2	23.6	12.2	CO_2	274.9	53.2	26.8
HF	159.2	31.3	15.7	Na_2	415.9	82.5	41.0
HCN	84.0	14.2	6.5	P_2	143.5	27.2	12.9
Na	210.5	41.6	20.6	Cl_2	200.3	34.5	14.3
SiH_3	184.2	36.8	18.2	NaCl	323.8	63.1	30.9
SiH_4	122.5	24.0	11.7	SiO	285.6	55.9	27.3
PH_2	82.6	15.8	7.7	CS	64.5	10.3	4.0
PH_3	73.1	14.4	7.2	ClF	223.1	41.9	20.0
SH_2	58.0	11.3	5.5	Si_2H_6	263.8	52	25.6
HCl	118.6	20.6	8.9	CH_3Cl	134.8	24.1	10.9
Li_2	95.3	36.3	31	CH_3SH	66.4	11.5	5.1
LiF	200.5	39.4	19.5	HOCl	199.1	35.7	16.8
C_2H_2	42.0	8.7	5.5	SO_2	307.6	59.7	29.0
C_2H_4	44.9	10.9	7.4	H	0.5	0.1	0.1
C_2H_6	51.8	11.3	6.6	Li	41.2	9.2	5.3

Table D.1: The standard error of the final total energy (shown in meV) using the noisy PBE approach shown for 46 chemical compounds in the GPAW G2-1 molecular database[16], assuming a noise factor (NF) of the form $\mathrm{NF}_{A,f} = Af\sin(|f + 1/f| + \phi)$ with magnitude A equal to either 0.005, 0.001, or 0.0005, and f corresponding to either $E_{xc}[n, \nabla n]$ or $v_{xc}[n, \nabla n]$, is added to both functions f. ϕ is an angle randomly chosen at each run of the GPAW script, and the standard error is computed over three runs with different ϕ. Numbers in dark blue surpass the 50meV total energy tolerance, while numbers in white succeed in being within 10meV of the original PBE result.

Name	NF 0.005	NF 0.001	NF 0.0005	Name	NF 0.005	NF 0.001	NF 0.0005
LiH	106.9	22.7	12.2	CO	668.7	139.3	73.2
BeH	155.2	35.9	21.0	HCO	688.7	145.0	77.1
CN	587.1	123.0	64.9	H_2CO	707.7	147.0	76.9
Be	136.1	31.5	18.4	CH_3OH	745.1	153.7	79.8
CH_3	312.4	66.3	35.5	N_2	663.8	141.8	76.5
CH_4	332.0	68.7	35.8	N_2H_4	740.3	155.9	82.9
NH_2	366.5	78.3	42.3	H_2O_2	859.3	178.6	93.6
NH_3	388.3	82.2	43.9	F_2	997.1	204.9	105.9
H_2O	450	94.2	49.8	CO_2	1087.8	227.5	120
HF	521.3	108.1	56.4	Na_2	1379.6	276.7	138.9
HCN	606.1	126.1	66.1	P_2	2227.9	448.5	226
Na	689.5	138.3	69.5	Cl_2	2715.1	549.4	278.6
SiH_3	1052.6	212.2	107.2	NaCl	2057.7	419	214.2
SiH_4	1072.0	215.6	108.5	SiO	1413.1	286.6	145.8
PH_2	1148.1	231.7	117.2	CS	1487.4	300.8	152.5
PH_3	1167.6	235.3	118.7	ClF	1858.9	378.5	193.5
SH_2	1265.8	253.5	127.0	Si_2H_6	2109.7	423.6	212.8
HCl	1377.1	280	142.9	CH_3Cl	1672.4	339.4	172.8
Li_2	147.4	7.7	9.8	CH_3SH	1564.3	315.7	159.7
LiF	597.2	123.2	64.0	HOCl	1790.2	366.9	189.0
C_2H_2	552.4	114.3	59.6	SO_2	2054.8	415.2	210.3
C_2H_4	591.6	122.6	64.0	H	12.1	2.4	1.2
C_2H_6	627.9	128.4	66.0	Li	83.6	14.8	6.2

Table D.2: The standard error of the final total energy (shown in meV) using the noisy PBE approach shown for 46 chemical compounds in the GPAW G2-1 molecular database[16], assuming a noise factor (NF) of the form $\mathrm{NF}_{A,f} = Af\sin^2(|f + 1/f| + \phi)$ with magnitude A equal to either 0.005, 0.001, or 0.0005, and f corresponding to either $E_{xc}[n, \nabla n]$ or $v_{xc}[n, \nabla n]$, is added to both functions f. ϕ is an angle randomly chosen at each run of the GPAW script, and the standard error is computed over three runs with different ϕ. Numbers in dark blue surpass the 50meV total energy tolerance, while numbers in white succeed in being within 10meV of the original PBE result.

Name	LDA − SP	LDA − SC	LDA − t	PBE − SP	PBE − SC	PBE − t
H4	2.0	2.0	1.2	−0.7	−0.7	0.9
He2	−0.1	−0.1	1.0	0.2	0.2	0.9
Li3	−4.9	−4.9	1.0	20.2	20.2	1.5
Be2	−9.7	−9.7	1.1	2.7	2.7	1.0
B12	22.8	22.8	1.2	160	160	1.4
C4	−20.3	−20.3	1.0	60.6	60.6	0.9
N8	31.3	31.3	1.0	72	149.6	1.5
F8	0.8	0.8	1.1	−0.7	−0.7	1.1
Ne4	0.0	−0.1	1.3	−0.4	−0.4	0.9
Na3	−5.9	−5.9	1.2	22.5	22.5	1.5
Mg2	14.2	14.2	1.1	0.0	0.0	1.5
Al4	19.7	19.8	1.3	57.2	57.2	1.1
Si8	−21.3	−21.3	1.1	164.4	164.4	1.0
P8	0.9	0.9	1.4	168.2	168.2	1.5
S1	0.0	0.0	0.8	17.3	17.3	0.1
Cl8	0.2	0.2	1.4	104.2	104.2	1.0
Ar4	−0.7	−0.6	1.4	−0.9	−1.0	1.2
K2	−2.6	−2.6	1.2	−0.4	−0.4	1.1
Sc2	23.3	23.3	1.2	23.4	23.4	0.9
Ti2	10.8	10.8	0.8	−8.0	−8.0	0.9
V2	−279.6	−279.6	1.1	−78.8	−78.8	1.0
Cu4	3.9	3.8	1.2	−6.4	−6.4	1.2
Zn2	6.0	6	1.3	22.8	22.8	1.7
Ga8	10.7	10.7	1.1	104.7	104.8	1.5
Ge8	−13	−13	1.3	108.9	109.2	1.5
As2	0.7	0.7	1.1	−1.1	−1.0	1.3
Se3	−0.4	−0.4	1.4	−7.7	−7.3	1.6
Br8	2.0	2.0	1.1	−23.8	−22.4	1.7
Kr4	0.0	−0.2	1.3	−1.4	−0.4	0.9
Rb2	−1.9	−1.8	1.9	1.6	2.0	1.1

Table D.3: The total energy error calculated using the signed version of Eq. 4.1 for 30 of the 59 systems in the Delta Codes DFT database[40, 41], showing the element name and number of atoms of that element in the unit cell (Name), the error for the NNLDA single-point (LDA-SP) and self-consistent (LDA-SC) approaches along with the computation time ratio (LDA-t) between the NNLDA and LDA methods, along with the error for the NNPBE single-point (PBE-SP) and self-consistent (PBE-SC) approaches along with the computation time ratio (PBE-t) between the NNPBE and PBE methods. The other 29 systems are shown in Table D.4.

Name	LDA − SP	LDA − SC	LDA − t	PBE − SP	PBE − SC	PBE − t
Y2	20.6	20.6	1.1	36.1	36.8	1.5
Zr2	18.0	18.0	1.1	37.2	38.0	1.7
Nb2	−462.1	−462	1.2	85.9	86.7	1.5
Mo2	−17.2	−17.2	1.1	56.9	57.7	1.0
Ru2	−18.4	−18.5	1.1	−104.3	−103.7	1.5
Rh4	34.1	34.1	1.2	−281.7	−281.6	1.3
Pd4	−8.5	−8.5	1.2	−275.8	−275.8	1.2
Ag4	16.9	16.9	1.2	−34.1	−34.0	1.5
Cd2	10.7	10.8	1.2	86.6	86.2	1.4
In2	12.5	12.4	1.1	45.3	44.6	1.1
Sn8	2.0	1.9	1.1	253.5	248.1	1.9
Sb2	3.0	3.0	0.8	6.9	5.0	1.3
Te3	2.9	2.9	1.9	−22.2	−26.0	1.6
I8	0.2	0.5	1.4	−24.2	−39.9	1.7
Xe4	0.2	−0.2	2.1	8.1	−1.0	1.3
Cs2	−1.2	−1.0	1.6	9.3	1.5	1.6
Hf2	12.4	12.4	1.1	100.8	−19.6	1.6
Ta2	−12.5	−12.3	0.7	123.6	−25.4	1.4
W2	−23.6	−23.7	2.1	125.2	−56.5	1
Re2	−21.1	−21.1	1.3	34.2	−31.8	1.6
Os2	8.8	8.8	1.3	67.9	−13.8	1.6
Ir4	−33.0	−33.2	1.0	−59.0	−258.8	1.8
Pt4	−25.7	−25.9	1.3	310.2	44.3	1.5
Au4	1.8	1.8	1.6	142.4	−225.9	1.3
Hg2	3.8	4.1	1.5	470.3	−55.7	1.9
Tl2	5.9	6.0	1.1	537.4	−22.9	1.6
Pb4	20.7	21	1.2	1501.7	−50.4	1.9
Bi2	3.0	3.0	0.8	978.9	−23.8	1.0
Rn4	0.6	0.4	1.4	4015.7	0.0	1.2

Table D.4: The total energy error calculated using the signed version of Eq. 4.1 for 29 of the 59 systems in the Delta Codes DFT database[40, 41], showing the element name and number of atoms of that element in the unit cell (Name), the error for the NNLDA single-point (LDA-SP) and self-consistent (LDA-SC) approaches along with the computation time ratio (LDA-t) between the NNLDA and LDA methods, along with the error for the NNPBE single-point (PBE-SP) and self-consistent (PBE-SC) approaches along with the computation time ratio (PBE-t) between the NNPBE and PBE methods. The other 30 systems are shown in Table D.3.

Appendix E

Slices of Phosphorene Exciton Density

As mentioned in Section 5.4.3, there are a variety of ways to present the electron and hole spectral densities as defined in Eqs. 2.62, 2.63, as these functions yield the oscillator strength intensity for all points in 3D space and at a given excitation energy. We have already explored two ways of presenting this data, both using isosurfaces that give a 3D picture of the electron and hole densities at a single energy (see Fig. 5.1) and by summing over planes normal to an axis of choice in the material, giving contour maps along one spatial direction and for a range of excitation energies. We present below for the case of phosphorene an alternate method of presenting this data, where summation is only done along one direction and snapshots at different voxels are used to give a more comprehensive view of the 2D spatial profile of excitations. This plotting method is not limited to two-dimensional materials, and can in fact be used for any system that utilizes our LCAO-TDDFT-k-ω code [45].

Shown in the figures below are the electron densities projected onto either the x-axis (Figs. E.1, E.2) or the y-axis (Figs. E.3, E.4) of the phosphorene unit cell for light polarized along either the x-axis (Figs. E.1, E.3) or the y-axis (Figs. E.2, E.4), expecting based on the results of Fig. 5.11 that the hole densities will look relatively similar. The voxels mentioned in these plots run from one edge of an axis to another, so these plots can be imagined to be lined up panel to panel to give a quasi-2D contour map along the energy axis where only one axis is projected, in this case the z-axis. More voxels are shown for projections along the y-axis due to the asymmetry of the phosphorene crystal. While the projection contour shown in Fig. 5.11 in Section 5.4.3 shows comprehensively that the electron density is concentrated at certain locations in the xz-plane or the yz-plane, the quasi-2D contours shown in Figs. E.1–E.4 show exactly where in the xy-plane these localized pockets of carrier density occur. These figures also reveal the symmetries present in the unit cell, where drawing these contours along the x or $-x$ directions or along the y or $-y$ directions yield mirror-reversed contour maps. Ultimately this provides a third way to represent the spatial profile of excitations within LCAO-TDDFT-k-ω.

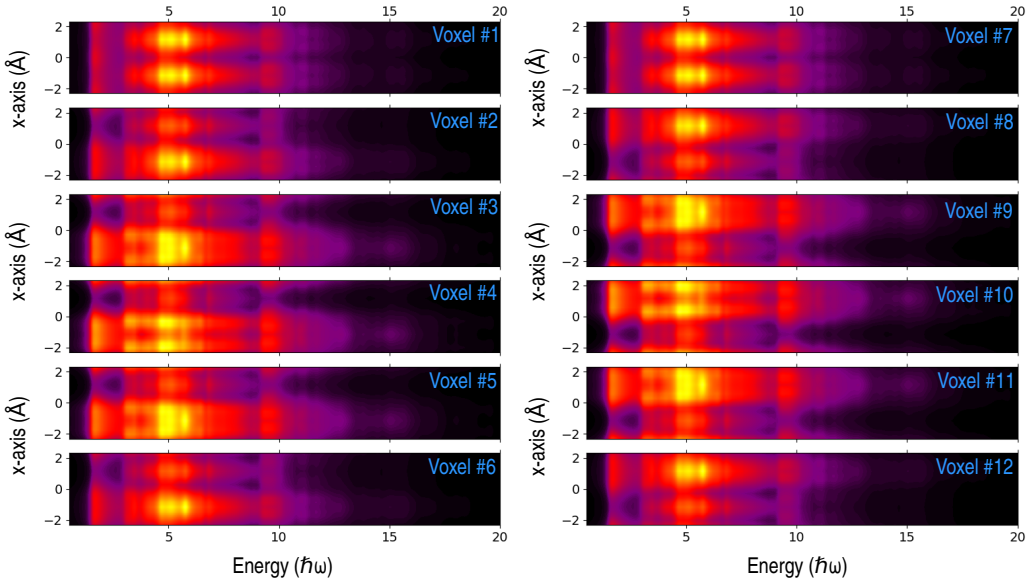

Figure E.1: Phosphorene (PN) LCAO-TDDFT-k-ω calculated electron (e^-) densities ρ_e projected onto the x axis for 12 voxels chosen along the y-axis versus energy $\hbar\omega$ in eV for light polarized along the x-axis.

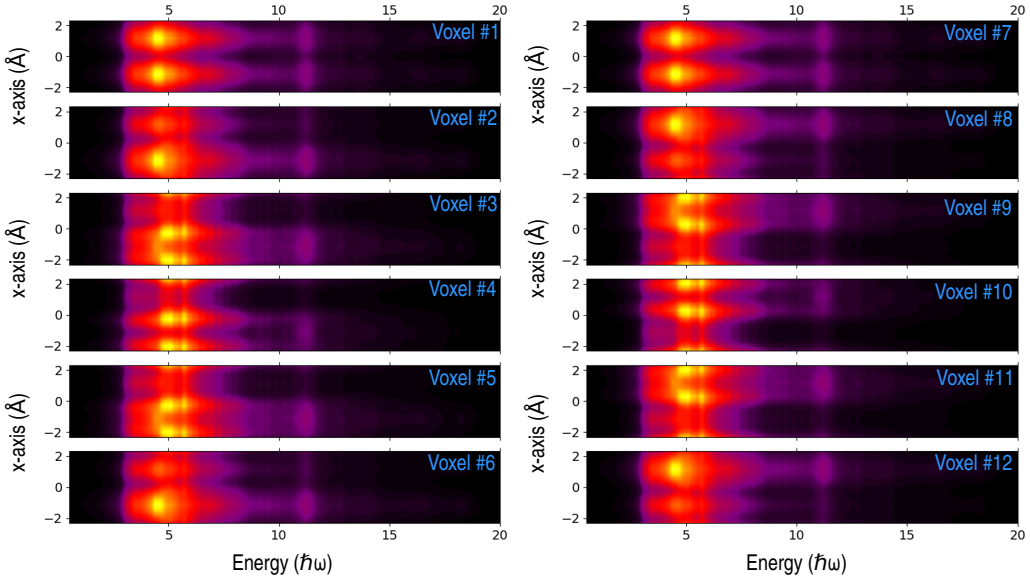

Figure E.2: Phosphorene (PN) LCAO-TDDFT-k-ω calculated electron (e^-) densities ρ_e projected onto the x axis for 12 voxels chosen along the y-axis versus energy $\hbar\omega$ in eV for light polarized along the y-axis.

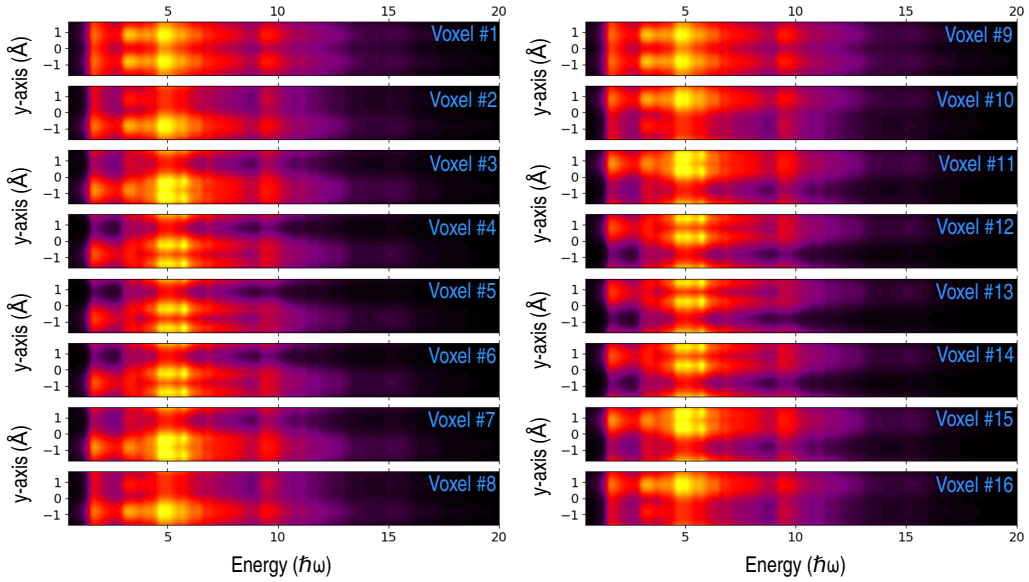

Figure E.3: Phosphorene (PN) LCAO-TDDFT-k-ω calculated electron (e^-) densities ρ_e projected onto the y axis for 16 voxels chosen along the x-axis versus energy $\hbar\omega$ in eV for light polarized along the x-axis.

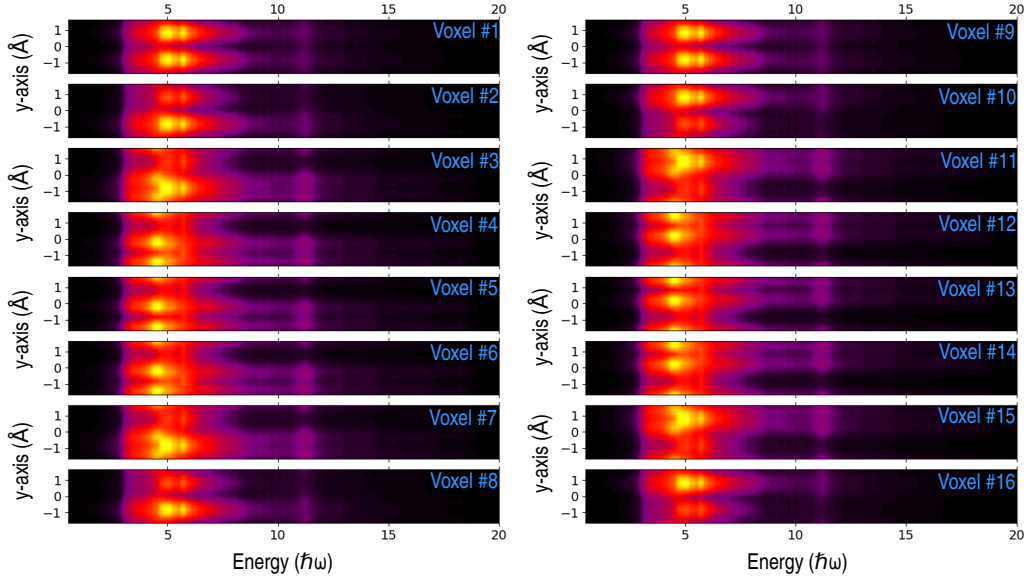

Figure E.4: Phosphorene (PN) LCAO-TDDFT-k-ω calculated electron (e^-) densities ρ_e projected onto the y axis for 16 voxels chosen along the x-axis versus energy $\hbar\omega$ in eV for light polarized along the y-axis.