

Ab initio theory of polarons



Weng Hong Sio
Jesus College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity Term, 2019

Abstract

Ab initio theory of polarons

Weng Hong Sio, Jesus College

A thesis submitted for the degree of
Doctor of Philosophy, Trinity Term 2019

The polaron is a quasiparticle formed by collectively dressing an electron with a phonon cloud. It emerges in a crystal lattice when an extra electron is trapped at the lattice distortion induced by itself. Due to its unique role and technological implications, this subject has consistently received attention for many years. To date, first-principle studies of polarons in materials are still based on direct calculations with large simulation cells. In this thesis we develop a theoretical and computational framework that allows for the modelling of lattice polarons in different materials of interest, but using unit-cell calculations of electronic, vibrational properties and electron-phonon couplings. We explain the architecture of the implementation in detail, and propose a self-interaction corrected DFT scheme for the validation of our methodology and implementation. We present extensive tests and benchmark for the code. We apply this new approach to investigate polaron properties in LiF and Li₂O₂ as prototypical bulk materials that host large and small polarons. We demonstrate that our method can describe both large and small polarons on the same footing, and capture both Fröhlich-type polar electron-phonon coupling and non-Fröhlich couplings to acoustic and optical phonons. Furthermore, we report calculations of formation energy, excitation energy, and wavefunction of polarons. To establish a quantitative understanding of which KS states and phonon modes contribute to the formation of polarons, we analysis the polaron spectral decomposition in a spirit similar to the Eliashberg spectral function. The last part of the thesis is devoted to explore the dimensionality effect on the polaron properties in h-BN. In this case, we compare the polaron energies for a localised hole between the h-BN in bulk and in monolayer, and show that the two-dimensional hole polaron trapped in the h-BN monolayer exhibits a layer formation energy in comparison to the bulk one. This work will open the door to predictive calculations of polarons in more complicated systems, such as van der Waals heterostructure and interfacial systems, and constitutes as a first step toward complete *ab initio* many-body calculations of polarons in real materials.

Contents

1	Introduction	1
1.1	Background and motivations	1
1.2	Overview of the thesis	5
2	Early studies of polarons	8
2.1	Semi-classical theory of self-trapped polarons	8
2.2	Modern polaron models	15
2.2.1	The Fröhlich model	16
2.2.2	Ground-state solutions of the Fröhlich model	20
2.2.3	The Holstein model	22
2.2.4	Toward model studies into <i>ab initio</i> formalism	26
2.3	Experimental signatures of polarons	27
2.3.1	Scanning Tunneling Microscopy	28
2.3.2	The ARPES technique	30
2.3.3	Spectral feature of polarons in ARPES	32
3	First principle methods for extended systems	39
3.1	Electronic structure in solids	39
3.1.1	The electronic structure problem	39
3.1.2	Density functional theory	43
3.1.3	Periodicities in solids	45
3.2	Vibrational properties of solids	51
3.2.1	Adiabatic lattice potentials	51
3.2.2	Density-functional perturbation theory	53
3.2.3	Longitudinal optical phonons	56
3.3	Electron-phonon interaction	61
3.3.1	<i>Ab initio</i> calculations of electron-phonon interactions	61
3.3.2	Fourier-Wannier interpolation	63
4	<i>Ab initio</i> theory of polarons	74
4.1	Polarons in density-functional theory	74
4.1.1	Derivation of the polaron equations	75
4.1.2	The formation energy of a polaron and the meaning of the polaron eigenvalue	80

4.1.3	Polaron equations in the basis of Kohn-Sham states and phonon modes	82
4.1.4	Lattice distortion in the polaronic ground state	84
4.1.5	Formation energy in the basis of Kohn-Sham states and phonon modes	85
4.1.6	Visualisation of the polaron wavefunction	87
4.1.7	Link with the Landau-Pekar model	88
4.2	Self-interaction corrections to polarons	89
5	Practical implementation	94
5.1	Density-functional theory calculations	94
5.2	Solution of the polaron equations	95
5.3	The implementation	101
5.3.1	Code development	101
5.3.2	Benchmark	108
5.4	Validation of the SIC functional	113
5.4.1	α -Quartz	113
5.4.2	Benchmark	114
6	Application I: bulk materials	116
6.1	Test systems	116
6.1.1	Lithium fluoride	116
6.1.2	Lithium peroxide	118
6.2	Electron polaron in LiF	119
6.2.1	Polaron energy vs. supercell size and Mott transition	119
6.2.2	Polaron wavefunctions	121
6.2.3	Spectral decomposition of the polaron	123
6.3	Hole polaron in LiF	125
6.3.1	Polaron energy vs. supercell size and Mott transition	125
6.3.2	Polaron wavefunction	126
6.3.3	Spectral decomposition of the polaron	128
6.4	Electron polaron in Li ₂ O ₂	131
6.4.1	Polaron energy vs. supercell size and Mott transition	131
6.4.2	Polaron wavefunction	132
6.4.3	Cross validation of the SIC functional	132
6.4.4	Spectral decomposition of the polaron	134

7	Application II: 2D materials	136
7.1	Hexagonal boron nitride	136
7.2	Computational details	139
7.3	Preliminary analyses	140
7.3.1	Convergence of band structures against vacuum size	142
7.3.2	Polaron energies	143
7.3.3	Polaron wavefunctions	143
7.3.4	Spectral decompositions of the polaron	145
7.4	Formation energy of polarons in monolayer h-BN	150
7.4.1	Polaron energies and metal-insulator transitions	151
7.4.2	Convergence of polaron energies with vacuum size	153
7.5	Comparisons between bulk and monolayer	155
7.6	Further work	161
8	Conclusion	163
8.1	Summary	163
8.2	Future work	165
A	Truncation of the long-range Fröhlich matrix element	168
B	Average phonon number	173
	Bibliography	177

Chapter 1

Introduction

The polaron is a quasiparticle that can be found in many crystalline solids, such as semiconductors [1], insulators [2] and molecular crystals [3]. It emerges when a slow moving electron interacts with the surrounding lattice via electron-lattice (phonon) interaction, which may introduce a distortion to the lattice and eventually the electron becomes localised at the distorted region. Owing to the effect of lattice deformation driven by the electron itself, the terms “self-trapped electrons” and “self-trapping” have been frequently used to imply polarons and the related electron-phonon physics in many literatures in the past [4–6]. The composite particle consisting of the self-trapped electron and the lattice distortion is known as an electron polaron, as firstly introduced by Landau in 1933 [7]. Despite such polaron concepts had been proposed in theory for many years, in contrast, the first experimental footprint of self-trapped holes were discovered in alkali halides using spin resonance techniques only starting from 1955 [8]. The richness and diversity of self-trapping phenomena makes polarons one of the most active research themes in condensed matter physics.

1.1 Background and motivations

In a simple idealised picture of crystalline solids in the absence of any defects at zero temperature, the lattice consists of valence electrons and nuclei in the ground state, is perfectly ordered, and exhibits a spatial translational symmetry. For example, this symmetry in the crystal is key to preserve the periodicity for all the electronic states in the crystal as demonstrated in the Bloch’s theorem. However, the symmetry may

be broken if the crystal is doped by introducing additional charge carriers. When a doped charge carrier is coupled to the atomic nuclei in the lattice via electron-phonon interaction (EPI), the nuclei are displaced in response to screen the additional electric field. This, in turn, considerably modifies the electronic and transport properties of the excess charge carriers [9]. This phenomenon is accompanied by several effects, such as renormalisation of electronic bands, broadening of spectral functions, enhancement of effective masses [10], and polaron satellites found in angle-resolved photoemission spectroscopy (ARPES) measurements [11]. These initial steps provide a blueprint for ultimately understanding the role of polaron formation in various classes of materials.

Furthermore, studies of polarons have been attracting significant attentions from different areas of physics, starting from the innovative work by pioneers of condensed matter theory such as Landau [7], Pekar [10], Fröhlich [12], Lee, Low and Pines [9], Feynman [13], Holstein [14], Bogoliubov [15] and others [16–19]. Not only these investigations involved formulating theoretical models for lattice polarons across different scales, such as the classic Landau-Pekar model for continuum polarons, the Fröhlich model for large polarons [12] and the Holstein model [14] for small polarons, but also various advanced techniques were developed to analyse these models and their solutions for different coupling regimes [9, 13, 15]. Beyond early work on the development of analytical techniques, more recent work included numerical investigations of the Fröhlich model using the diagrammatic Monte Carlo method [20], path integral Monte Carlo [21] and the renormalisation group approach [22]. For a more detailed presentation about this vast research area one might consider up-to-date reviews [23–25].

Over the years, polaron concepts have been adapted to several different fields outside the traditional condensed matter and material physics communities. For instance, the Landau-Pekar polaron model found applications in quantum gravity

[26] and quantum state of reduction [27]. As the polaron problem constitutes as an important chapter of in soliton physics and nonlinear physics, there have been studies on the stabilities and possible solutions from the mathematical physics perspective [28–31]. Very recently, research on polarons has moved onto the emergent field of ultracold atoms and Bose-Einstein condensates (BEC) [22, 32, 33]. For example, by combining a BEC with an optical lattice, it is possible to explore a quantum system analogous to electrons in a crystalline solid but with tunable and precise control of the lattice properties [34]. In such systems the concept of lattice polarons can be generalised to the “Bose polaron”, in which an impurity atom is an analogue to the extra electron but instead localised in the optical lattice immersed by an uniform ultracold gas [35], and the interaction between the impurity atom and the bosonic bath consists of ultracold atoms is an analogue to the e-ph interaction in solids. This advancement sheds light on studying emergent states of quantum matter and offers an alternative way to validate polaron theories, using novel experimental techniques in atomic physics [34, 36, 37].

Despite the successes of the polaron models and growing interest in applying these methods to study the quasi-particle properties of the impurity in ultracold atoms [22, 32, 33], there are obstacles toward the application of these techniques to the study of lattice polarons in *real* solids. For example, it is unclear how to select the appropriate form of couplings that should appear in the model for a specific compound. Furthermore, in most complicated materials of technological interest such as perovskites [38, 39] and transition metal oxides [40], interesting couplings emerge that are not well captured by the extreme descriptions of Fröhlich and Holstein polarons [12, 14]. Although polaron models have been advanced and are getting more complex recently [41], the lack of comprehensive insights for specific compounds generally limit the quality and the accuracy of the model. In contrast to the profuse application of models [12, 14, 42] in studying various sizes of polarons across different

scales, *ab initio* studies of lattice polarons are limited to small and intermediate-size systems, which involves few hundreds up to about two thousands atoms [43]. In order to address these drawbacks, considerable effort is being made to expand the model Hamiltonians using advanced *ab initio* formalisms that enable precise EPI mechanisms in *real* solids to be captured [44, 45], where the EPI is computed from first principle without making any *priori* assumptions.

Due to the fast-growing computational power in the recent decades, calculations based on density-functional theory and periodic supercell have become an extremely promising tool for studying polarons in various systems, including alkali halides, metal oxides, superconductors and even molecules [46–49]. However, the DFT supercell approach also encounters some limitations in the study of polarons. Since such calculations consider adding or removing an electron in a supercell, it is impractical to perform the calculations for a system that hosts intermediate-size polarons and beyond (i.e. supercells containing up to a few thousand atoms) [43]. In addition to these computational challenges, standard DFT calculations with local functional also suffer from the well-known self-interaction error [50], which is believed to be the causes of failure on predicting localised states. Early attempts involved modifying the electronic structure using the Hubbard-corrected DFT [51, 52], hybrid functionals [51–53] and specialised self-interaction correction (SIC) schemes [46]. Even if these technical problems will be resolved in the future, DFT calculations based on supercell calculations offer limited physical insight into the EPI mechanisms that drive polaron formation. Therefore, it is important to establish a general formalism starting from practical considerations of the *real* lattice structure and their interactions with electrons, in order to bridge the gap between such calculations and various model approaches.

In this thesis, we formulate an *ab initio* theory of polarons and develop the corresponding computational framework. This work enables for the first time to access

not only the wavefunction for small polarons that has already been probed by the supercell approach [54, 55], but also for large polarons that were predicted in model calculations. Our method also offers an unique opportunity to calculate atomic displacements associated to the polaron ground state, polaron eigenvalues and formation energies, as well as several interesting quantities that cannot be explored directly by using the supercell approach, such as the spectral decompositions of polaron wavefunctions and atomic displacements, polaron effective masses, and the average number of excited phonons. One of the potential applications of our theory is the investigation of polarons in systems with complicated structures, such as various perovskites and low-dimensional systems including two-dimensional (2D) materials [56–60], van der Waals heterostructures [61, 62], surfaces and interfaces [63–65].

In order to demonstrate the capability of our method, we concentrate on two simple bulk materials as a starting point: lithium fluoride LiF, which is a prototypical system for studying Fröhlich type polarons, and the hexagonal lithium peroxide Li₂O₂, which exhibits a Holstein type EPI that leads to small polaron formation. In addition to the investigation of the distinctive EPI mechanisms and the leading polaronic effects in these compounds, the latter system is important for validating the present theory against explicit DFT supercell calculations.

1.2 Overview of the thesis

Despite many models and solutions based on the plane-wave electron and the single phonon coupling were developed to study polaronic effects in the last century, these models lack sufficient atomistic insight. In order to accurately describe polaronic properties in real materials, it is important to consider both electrons and phonons using a proper atomistic treatment. This thesis is devoted to this aim, and is organised as follows. In Chapter 2 we outline several representative works for polarons starting

from the pioneer studies in the last century to the recent developments. In particular, we introduce the classic Landau-Pekar model, followed by the quantum mechanical description of large polarons by Fröhlich and small polarons by Holstein. We emphasize the significance of these models and explain what is missing in order to accurately describe polarons in real materials. Finally, we present some recent experimental discoveries on polarons in transition metal oxides.

In Chapter 3 we review the theoretical basis for enabling *ab initio* calculations of electronic and vibrational properties of solids. Here we begin with density-functional theory (DFT) in the planewave and pseudopotential formalism, and we describe the extension of DFT to density-functional perturbation theory for computing vibrational properties. These methods constitute as a stepping stone for studying electron-phonon interactions in real materials. Finally, we present the Fourier-Wannier interpolation method that provides efficient calculations of electron-phonon matrix elements from first-principles.

In Chapter 4 we develop a theoretical framework for calculating self-trapped polarons in semiconductors and insulators from first-principles, and we propose a DFT functional with self-interaction correction (SIC) for the corresponding validation. In Chapter 5 we show the computational details of all unit cell calculations, and describe how to solve the polaron equations. We also explain the technical aspect and difficulties for the implementation. Finally, we give details of the code validation and benchmark.

Chapters 6 and 7 present applications of the method to explore polarons in bulk crystals and 2D systems respectively. In Chapter 6 we focus on two bulk materials, namely the lithium fluoride (LiF) and lithium peroxide (Li_2O_2). In particular, we show a cross validation by comparing direct calculations using the SIC functional with previous work of Li_2O_2 . Then we discuss the dependence of the polaron energies on supercell size, and show calculations of polaron wavefunctions and lattice distortions.

tions. Lastly, we present the spectral decomposition of the polaron into KS states and phonon modes. In Chapter 7 we proceed to investigate polarons in hexagonal boron nitride (h-BN), we present first-principles calculations of polaron formation energies, eigenvalues, wavefunctions and spectral decompositions for both the hole and electron polarons in h-BN monolayer. We demonstrate the concept of two-dimensional polarons and discuss how to converge polaron energies in monolayer calculations. We compare the polaron energies between the monolayer and the bulk form for the same compound. Finally, Chapter 8 presents the summary, the forecast of future works and a conclusion for this thesis.

Chapter 2

Early studies of polarons

In this introductory chapter, we review the notion of lattice polarons and the historical development of various polaron approaches based on models and analytical techniques. We begin with the Landau-Pekar (LP) polaron model and discuss the underlying assumptions and limitations. Next we derive the quantum Fröhlich model, and explain its significance and the interconnection to the classic LP model. Similarly we introduce the concept of Holstein polarons, and the distinction between Fröhlich and Holstein type EPI is emphasised. Furthermore, we review experimental evidence for the polaron footprints, such as the polaron satellites found by angle-resolved photoemission spectroscopy (APRES) [11] and electron paramagnetic resonance (EPR) measurements [66]. The purpose of this section is to give a preliminary overview on polaron physics so that we can proceed to a more elaborate discussion in later chapters.

2.1 Semi-classical theory of self-trapped polarons

The first polaron concept was proposed by Landau in 1933, who provides a qualitative discussion in his seminar article about a slow moving electron in an ionic crystal. He suggested that the electron may perturb the crystal and eventually become trapped at a distorted region of the host lattice induced by itself to form the so-called polaron.

The phenomenological model for describing such a nonlinear behaviour was missing for a decade, and was only developed by Landau and Pekar in the mid 1940s. When this problem was revisited again, they realised that there was an internal con-

tradition in the well-established Bloch theory, namely: while considering the band state of a free electron, the ions are always fixed at their equilibrium positions. In this case, the electron turns out to be sitting in a periodic potential well and hence appears to be delocalised. However, Pekar considered an alternative scenario where the ions are able to move, and assumed that the electron should follow the ionic motion adiabatically. He worked out the first polaron theory by treating the ionic crystal as a continuum and polarisable dielectric. In this new point of view, the band electron can polarise the ions with its own electric field. The ions are then displaced to screen this additional electric field and form dipoles, which creates a secondary dipole potential acting back on the electron. Through this feedback mechanism, the electron and the ions can find an equilibrium configuration that breaks translational invariance; that is a polaronic ground state. Because of the loss of periodicity, the trapped electron exhibits a discrete energy spectrum, and the ground state energy is lowered in comparison to the perfect lattice hosting a band electron. In the following, we review this fascinating model and explain its significance in the historical development of the polaron theory. This model provides a very useful starting point to understand our novel *ab initio* formalism described in Chapter 4.

The Landau-Pekar (LP) model is a simple framework for studying self-trapped phenomena in solids, which describes how a single electron added to a polar insulator becomes localised. The basic assumption considered in the model is that the electron wavefunction spreads over large spatial dimensions spanning many crystal unit cells. As a consequence, the atomistic information of the crystal are smoothed out; the interaction of the added electron with the valence electrons is described via the effective-mass approximation and thus only enters to the kinetic energy term; the interaction of this electron with the ionic lattice is described by a continuum

electrostatic model. The total energy of the LP model is expressed as:

$$E_{\text{LP}} = \frac{\hbar^2}{2m^*} \int d\mathbf{r} |\nabla\psi|^2 + \frac{1}{2} \int d\mathbf{r} \mathbf{E} \cdot \mathbf{D}, \quad (2.1)$$

where $\psi(\mathbf{r})$ is the wavefunction of the added electron, $\mathbf{E}(\mathbf{r})$ is the self-consistent electric field and $\mathbf{D}(\mathbf{r})$ is the electric displacement field. The first term on the r.h.s. of Eq. (2.1) describes the band energy of the extra electron, which includes electron-electron interactions via the conduction band effective mass m^* . The second term describes the total electrostatic energy of the dielectric [67].

The electric displacement field $\mathbf{D}$ is related to the charge density of free carriers, and therefore to the wavefunction of the excess electron, via the Maxwell's equations $\nabla \cdot \mathbf{D} = -e|\psi(\mathbf{r})|^2$ (e is the elementary charge), or can be equivalently expressed in an integral form:

$$\mathbf{D} = \frac{e}{4\pi} \nabla \int d\mathbf{r}' \frac{|\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.2)$$

The electric displacement field can also be related to the self-consistent electric field via $\mathbf{D} = \epsilon_0 \epsilon^0 \mathbf{E}$, where ϵ_0 is the vacuum permittivity and ϵ^0 is the static dielectric constant. By replacing Eq. (2.2) into (2.1) we obtain the total electrostatic energy as:

$$\frac{1}{2} \int d\mathbf{r} \mathbf{E} \cdot \mathbf{D} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{\epsilon^0} \int d\mathbf{r} d\mathbf{r}' \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.3)$$

In this expression, both the electronic screening and the lattice screening are included in the electric field $\mathbf{E}$. Since the electronic screening energy is already accounted for in the band structure term in Eq. (2.1), we need to subtract this contribution from Eq. (2.3). The electronic-only contribution is determined by evaluating Eq. (2.3) with the ionic screening turned off, i.e. by using the high-frequency (electronic) permittivity ϵ^∞ instead of the static (electronic and ionic) permittivity ϵ^0 . After removing this

contribution the electrostatic energy reads:

$$\frac{1}{2} \int d\mathbf{r} \mathbf{E} \cdot \mathbf{D} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{\epsilon^0} - \frac{1}{\epsilon^\infty} \right) \int d\mathbf{r} d\mathbf{r}' \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.4)$$

By defining $1/\kappa = 1/\epsilon^\infty - 1/\epsilon^0$ [23], we can rewrite Eq. (2.1) as a functional of the polaron wavefunction:

$$E_{\text{LP}}[\psi] = \frac{\hbar^2}{2m^*} \int d\mathbf{r} |\nabla \psi(\mathbf{r})|^2 - \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \int d\mathbf{r} d\mathbf{r}' \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.5)$$

The ground-state energy of the LP polaron is obtained by minimising this functional with respect to ψ , subject to the constraint provided by the normalisation condition $\int d\mathbf{r} |\psi(\mathbf{r})|^2 = 1$. This problem can be solved by transforming it into an unconstrained minimisation with the normalisation incorporated via the Lagrange multiplier ε :

$$E'_{\text{LP}}[\psi, \varepsilon] = \frac{\hbar^2}{2m^*} \int d\mathbf{r} |\nabla \psi(\mathbf{r})|^2 - \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \int d\mathbf{r} d\mathbf{r}' \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} - \varepsilon \left(\int d\mathbf{r} |\psi(\mathbf{r})|^2 - 1 \right). \quad (2.6)$$

By setting to zero the two functional derivatives $\delta E'_{\text{LP}}/\delta \psi^*$ and $\delta E'_{\text{LP}}/\delta \varepsilon$, one obtains a Schrödinger-type eigenvalue problem for ψ :

$$\left\{ -\frac{\hbar^2}{2m^*} \nabla^2 \psi(\mathbf{r}) - \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \int d\mathbf{r}' \frac{|\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}), \right. \quad (2.7)$$

$$\left. \int d\mathbf{r} |\psi(\mathbf{r})|^2 = 1. \right. \quad (2.8)$$

Here the eigenvalue ε has the dimension of energy but it does not represent the total energy of the polaron. This is seen by projecting Eq. (2.7) onto ψ^* and comparing with Eq. (2.5):

$$E_{\text{LP}} = \varepsilon + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \int d\mathbf{r} d\mathbf{r}' \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.9)$$

Equation (2.7) provides an intuitive understanding of the nature of polaron self-

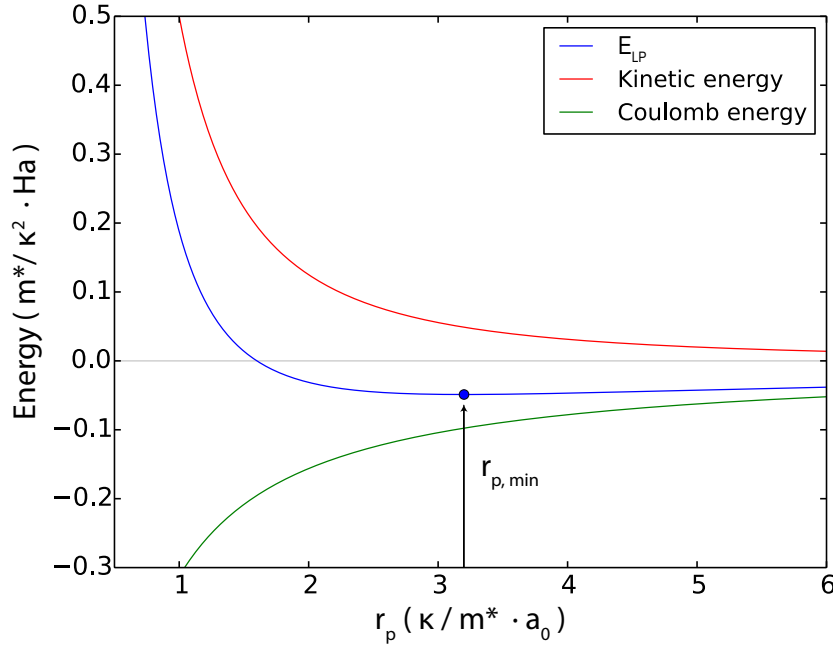


Figure 2.1: Plot of energy vs polaron radius r_p using the trial function in Eq. (2.10)

trapping in the LP model. Let us consider for example a normalised trial wavefunction [23]:

$$\psi(\mathbf{r}) = (\pi r_p^3)^{-1/2} \exp(-|\mathbf{r}|/r_p). \quad (2.10)$$

Using this trial function in the LP functional in Eq. (2.5) as shown in Fig. 2.1, we see that that minimisation of the kinetic energy term in Eq. (2.7) favors delocalisation (larger r_p), while the minimisation of the Coulomb term favors localisation (smaller r_p). The polaron size r_p results from a trade-off between these competing effects. By replacing the above exponential ansatz in Eq. (2.5), one obtains a simple estimate for the energy as a function of the polaron size r_p :

$$E_{\text{LP}}(r_p) = \frac{\hbar^2}{2m^*r_p^2} - \frac{5}{16} \frac{1}{\kappa} \frac{e^2}{4\pi\epsilon_0 r_p}. \quad (2.11)$$

The minimum of this function is given by:

$$\frac{r_{\text{p,min}}}{a_0} = \frac{16}{5} \frac{\kappa}{m^*/m_e}, \quad (2.12)$$

where a_0 denotes the Bohr radius and m_e is the free electron mass. By replacing Eq. (2.12) inside Eq. (2.11) we find the ground-state energy:

$$\frac{E_{\text{LP,min}}}{E_{\text{Ha}}} = -\frac{25}{512} \frac{m^*/m_e}{\kappa^2}, \quad (2.13)$$

with $E_{\text{Ha}} = 27.211$ eV is known as the Hartree energy. Furthermore, by using Eq. (2.9) we obtain the polaron eigenvalue [23]:

$$\frac{\varepsilon}{E_{\text{Ha}}} = -\frac{75}{512} \frac{m^*/m_e}{\kappa^2}. \quad (2.14)$$

The energy given by Eq. (2.13) can also be expressed by using the standard polaron coupling constant α , which is defined as [23]:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar} \sqrt{\frac{m^*}{2\hbar\omega_{\text{LO}}}} \frac{1}{\kappa}, \quad (2.15)$$

where ω_{LO} is the characteristic frequency of longitudinal optical phonons. By combining Eqs. (2.13) and (2.15) one obtains the standard result:

$$E_{\text{LP,min}} = -\frac{50}{512} \alpha^2 \hbar\omega_{\text{LO}}, \quad (2.16)$$

which is very close to the original variational solution by Pekar [10]. Much work has been done to improve on the simple exponential ansatz in Eq. (2.10) employed in this overview of the LP model. However, apart from obtaining a more accurate prefactor in front of the term $\alpha^2 \hbar\omega_{\text{LO}}$ in Eq. (2.16), these improvements do not change the qualitative features of the solution. This is because Eq. (2.5) can be written in a

scale-invariant form by defining:

$$\psi(\mathbf{r}) = a^{-3/2} \phi(\mathbf{r}/a), \quad (2.17)$$

with $a = \kappa a_0 / (m^*/m_e)$ so that:

$$\frac{E_{\text{LP}}}{\alpha^2 \hbar \omega} = \int d\mathbf{r} |\nabla \phi(\mathbf{r})|^2 - \int \frac{|\phi(\mathbf{r})|^2 |\phi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (2.18)$$

subject to the normalisation condition $\int d\mathbf{r} |\phi(\mathbf{r})|^2 = 1$. The direct numerical solution of Eq. (2.18) yields a wavefunction which is very close to the original variational result found by Pekar using a modified exponential [10, 23, 68, 69].

From Eq. (2.13) we see that the formation of a localised polaron is only possible when $\epsilon^0 > \epsilon^\infty$, that is in polar crystals. This leaves out those polarons that can form in non-polar semiconductors. In addition, since the electron-phonon (e-ph) coupling mechanism is related to the ionic dielectric response, i.e. to the long-range Fröhlich potential generated by lattice distortions, the LP model also leaves out acoustic and piezo-acoustic polarons. Further limitations of the model are that it assumes an isotropic dielectric, and does not take into account the atomistic nature of the crystal lattice. In Ref. [69] it was pointed out that the LP model is essentially never valid, because it relies on the assumption of large polarons in order to use continuum electrostatics, but its results tend to be accurate in the regime of strong coupling, that is for small polarons, in contrast with the starting hypothesis. The LP model is said to describe strong-coupling polarons because the energy given by Eq. (2.16) is almost the same as that obtained in the strong-coupling limit of the Feynman theory [13], $-\alpha^2 \hbar \omega / 3\pi$, as shown in Fig. 2.2.

In Chapter 4, we show how the essential physics of the LP model can be retained by moving to an *ab initio* formalism based on density-functional theory, and that most of the intrinsic limitations of the model can be overcome in this new framework.

2.2 Modern polaron models

Short after the classic treatment of polarons by Landau and Pekar in late 1940s, Fröhlich provided a complete microscopic description of polarons in 1950 using a proper quantum mechanical Hamiltonian, namely the Fröhlich model. Although the first polaron theory was founded and developed by Landau and Pekar, the beginning of real interest in this field was marked by the pioneering work of Fröhlich. This is demonstrated by the substantial amount of work on determining the analytical solutions for the Fröhlich model between the year 1950 and 1955 [9, 13, 16, 70]. In the fully quantum mechanical picture of Fröhlich, the polaron was described as a composed (quasi-) particle consisting of an electron and the induced lattice distortion nearby. As opposed to the LP model, which represents the deformed lattice as a rigid and static polarisable dielectric, the lattice distortion was captured precisely by using harmonic oscillators in the Fröhlich model. This refinement was possibly the first time that the quantum field picture of a non-relativistic electron interacting with a scalar bosonic field (phonons) found its way in solid state physics. The Fröhlich model considers the limit of a slow electron moving in a crystalline solid with a broad spatial extent, so that the discrete nature of the atomic structure becomes insignificant, and the continuum approximation could be used to simplify the model as much as possible. The model retains sufficient ingredients of polaron physics, such as the band electron as a simple structureless plane wave, dispersive phonons as a single Einstein mode, and EPIs as a screened dipole interaction. This strategy turns out to be useful when theorists develop simple models in order to understand the underlying mechanism and their interrelations. In addition to its successes in explaining the properties of an electron in polar crystal, the model constitutes as an extraordinary playground for many different fields of physics as mentioned earlier. Therefore, it is useful to discuss this model in detail.

2.2.1 The Fröhlich model

To derive the Fröhlich Hamiltonian for describing a large polaron hosted by polar crystals, we start from the lattice field theory of Fröhlich [70]. Let us firstly define the total polarisation density $\mathbf{P}_{\text{tot}}$ at the coordinate $\mathbf{r}$ in the host lattice as the linear superposition of the ionic polarisations $\mathbf{P}_{\text{ion}}$ and electronic polarisations $\mathbf{P}_{\text{ele}}$:

$$\mathbf{P}_{\text{tot}}(\mathbf{r}, t) = \mathbf{P}_{\text{ion}}(\mathbf{r}, t) + \mathbf{P}_{\text{ele}}(\mathbf{r}, t). \quad (2.19)$$

Here these two different polarisations in nature are responsible for the broken translational symmetry of the perfect lattice due to the ionic motion. The ionic polarisation originates from the displacements of ions which may be induced thermally; the vibrational frequencies lie in the infrared region. The ionic displacements result in the electronic polarisation, which oscillates much faster, with frequencies in the visible and ultraviolet region. When we consider an extra slow electron sitting in the conduction band, the electron interacts with the phonon field associated with the polarisation and this defines the so-called electron-phonon Hamiltonian. The slow electron interacts with such induced ionic dipoles via long-range Coulomb interactions to form a large polaron, its wavefunction will spread over many crystal unit cells. In this case one may ignore effects of the lattice discreteness. This is known as the continuum approximation and all the quantities in the Fröhlich model can be expressed using continuous functions. Using the relations from electrostatics:

$$\begin{aligned} \mathbf{D} &= \mathbf{P} + \epsilon_0 \mathbf{E}, \\ \mathbf{D} &= \epsilon_0 \epsilon(\omega) \mathbf{E}, \end{aligned} \quad (2.20)$$

where $\epsilon(\omega)$ is the frequency-dependent dielectric constant of the medium. Rearranging Eq. (2.20):

$$\mathbf{P} = \left[1 - \frac{1}{\epsilon(\omega)} \right] \mathbf{D}. \quad (2.21)$$

In the low-frequency limit $\epsilon(\omega \rightarrow 0) \rightarrow \epsilon^0$ and the total polarisation is given by:

$$\mathbf{P} = \left(1 - \frac{1}{\epsilon^0} \right) \mathbf{D}. \quad (2.22)$$

Now we consider the high-frequency limit in order to eliminate the electronic polarisation from the total polarisation. The so-called high-frequency limit here means that the frequency is much higher than the infrared range, but much lower than the visible-ultraviolet range, i.e. $\omega_{\text{ion}} \ll \omega_{\infty} \ll \omega_{\text{ele}}$. The electronic polarisation responses with this frequency and we have:

$$\mathbf{P}_{\text{ele}} = \left(1 - \frac{1}{\epsilon^{\infty}} \right) \mathbf{D}. \quad (2.23)$$

We can extract the ionic polarisation by combining Eqs. (2.22) and (2.23):

$$\mathbf{P}_{\text{ion}} = \mathbf{P} - \mathbf{P}_{\text{ele}} = \left(\frac{1}{\epsilon^{\infty}} - \frac{1}{\epsilon^0} \right) \mathbf{D} \equiv \frac{1}{\kappa} \mathbf{D}. \quad (2.24)$$

If we place an electron with coordinate $\mathbf{r}_{\text{ele}}$ into the system, and the electron creates a source field $\mathbf{D}(\mathbf{r}, \mathbf{r}_{\text{ele}})$, the equation of motion for describing the ionic polarisation density is given by:

$$\ddot{\mathbf{P}}_{\text{ion}}(\mathbf{r}) + \omega_{\text{ion}}^2 \mathbf{P}_{\text{ion}}(\mathbf{r}) = \frac{\omega_{\text{ion}}^2}{\kappa} \mathbf{D}(\mathbf{r}, \mathbf{r}_{\text{ele}}). \quad (2.25)$$

Similarly, the equation of motion for the electron is:

$$m^* \ddot{\mathbf{r}}_{\text{ele}} = -\frac{e}{\epsilon_0} \mathbf{P}_{\text{ion}}. \quad (2.26)$$

From these equations of motion, we are able to write down the Hamiltonian $\hat{H}$ which can be partitioned into three terms as:

$$\hat{H} = \hat{H}_{\text{ele}} + \hat{H}_{\text{ph}} + \hat{H}_{\text{int}}, \quad (2.27)$$

namely the non-interacting electron term $\hat{H}_{\text{ele}}$, the non-interacting phonon term $\hat{H}_{\text{ph}}$ and the EPI term $\hat{H}_{\text{int}}$. They are given by:

$$\begin{aligned} \hat{H}_{\text{ele}} &= \frac{\mathbf{p}_{\text{ele}}^2}{2m^*}, \\ \hat{H}_{\text{ph}} &= \frac{\kappa}{2\omega_{\text{ion}}^2} \left(\frac{1}{\epsilon_0} \right) \int d\mathbf{r} \left[\dot{\mathbf{P}}_{\text{ion}}^2(\mathbf{r}) + \omega_{\text{ion}}^2 \mathbf{P}_{\text{ion}}^2(\mathbf{r}) \right], \\ \hat{H}_{\text{int}} &= -\frac{1}{\epsilon_0} \int d\mathbf{r} \mathbf{D}(\mathbf{r}, \mathbf{r}_{\text{ele}}) \cdot \mathbf{P}_{\text{ion}}(\mathbf{r}), \end{aligned} \quad (2.28)$$

where the conjugate momenta for the electron and the ionic polarisation are given by $\mathbf{p}_{\text{ele}} = m^* \dot{\mathbf{r}}_{\text{ele}}$ and $\mathbf{P}_{\text{ion}} = \frac{\kappa}{\omega_{\text{ion}}^2} \dot{\mathbf{P}}_{\text{ion}}$. Note that we have neglected the contribution from the electronic polarisation as $\mathbf{P}_{\text{ele}} \ll \mathbf{P}_{\text{ion}}$ because of $\omega_{\text{ion}}^2 \ll \omega_{\text{ele}}^2$. It is useful to recall several electrostatics properties for the system beforehand. The source field $\mathbf{D}$ generated by the electron satisfies the Maxwell's equations $\nabla \cdot \mathbf{D} = -e\delta(\mathbf{r} - \mathbf{r}_{\text{ele}})$ and $\nabla \times \mathbf{D} = 0$. Thus:

$$\mathbf{D} = -\frac{e}{4\pi} \nabla \left[\frac{1}{|\mathbf{r} - \mathbf{r}_{\text{ele}}|} \right]. \quad (2.29)$$

Similarly, we can also express the ionic polarisation using a scalar potential $\Phi_{\text{pol}}(\mathbf{r})$ because of $\nabla \times \mathbf{P} = 0$:

$$\mathbf{P}_{\text{ion}} = \epsilon_0 \nabla \Phi_{\text{pol}}(\mathbf{r}). \quad (2.30)$$

Using Eqs. (2.29) and (2.30), we can evaluate the EPI term defined in Eq. (2.28) as:

$$\hat{H}_{\text{int}} = \frac{e}{4\pi} \left[\int d\mathbf{r} \nabla \cdot \left(\Phi_{\text{pol}} \nabla \frac{1}{|\mathbf{r} - \mathbf{r}_{\text{ele}}|} \right) - \Phi_{\text{pol}} \nabla^2 \frac{1}{|\mathbf{r} - \mathbf{r}_{\text{ele}}|} \right] = e\Phi_{\text{pol}}(\mathbf{r}_{\text{ele}}). \quad (2.31)$$

Until now the derived Hamiltonian is still a classical object. We proceed by performing a canonical quantisation on the polarisation field. The canonical commutation between the field operator and its conjugate momentum is:

$$[P_\alpha(\mathbf{r}), \frac{\kappa}{\omega_0^2} \dot{P}_\beta(\mathbf{r}')] = i\hbar \delta_{\alpha\beta} \delta^{(3)}(\mathbf{r} - \mathbf{r}'), \quad (2.32)$$

where the indices α, β denote the Cartesian components, $\delta^{(3)}(\mathbf{r})$ is the Dirac delta function in three-dimensional space. This enables us to transform the conjugate field operators into the creation and annihilation operator pair such that $[B_\alpha(\mathbf{r}), B_\beta^\dagger(\mathbf{r}')] = \delta_{\alpha\beta} \delta^{(3)}(\mathbf{r} - \mathbf{r}')$ is satisfied:

$$\begin{aligned} \mathbf{B}(\mathbf{r}) &= \left(\frac{\kappa}{2\hbar\omega_0} \right)^{\frac{1}{2}} \left[\mathbf{P}_{\text{ion}}(\mathbf{r}) + \frac{i}{\omega_0} \dot{\mathbf{P}}_{\text{ion}}(\mathbf{r}) \right], \\ \mathbf{B}^\dagger(\mathbf{r}) &= \left(\frac{\kappa}{2\hbar\omega_0} \right)^{\frac{1}{2}} \left[\mathbf{P}_{\text{ion}}(\mathbf{r}) - \frac{i}{\omega_0} \dot{\mathbf{P}}_{\text{ion}}(\mathbf{r}) \right]. \end{aligned} \quad (2.33)$$

The inverse relations are:

$$\begin{aligned} \mathbf{P}_{\text{ion}}(\mathbf{r}) &= \left(\frac{\hbar\omega_0}{2\kappa} \right)^{\frac{1}{2}} [\mathbf{B}(\mathbf{r}) + \mathbf{B}^\dagger(\mathbf{r})], \\ i\dot{\mathbf{P}}_{\text{ion}}(\mathbf{r}) &= \left(\frac{\hbar\omega_0}{2\kappa} \right)^{\frac{1}{2}} \omega_0 [\mathbf{B}(\mathbf{r}) - \mathbf{B}^\dagger(\mathbf{r})]. \end{aligned} \quad (2.34)$$

It is possible to express the operator $\mathbf{B}(\mathbf{r})$ as a Fourier series by considering $\nabla \times \mathbf{P} = 0$:

$$\mathbf{B}(\mathbf{r}) = \left(\frac{\epsilon_0}{\Omega} \right)^{\frac{1}{2}} \sum_{\mathbf{q}} \frac{\mathbf{q}}{|\mathbf{q}|} b_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}}, \quad (2.35)$$

where Ω is the volume of the primitive unit cell. Combining Eqs. (2.28), (2.34) and (2.35), the non-interacting phonon term is simplified to:

$$\hat{H}_{\text{ph}} = \hbar\omega_0 \sum_{\mathbf{q}} b_{\mathbf{q}}^\dagger b_{\mathbf{q}} + \frac{1}{2} \hbar\omega_0. \quad (2.36)$$

Combining Eqs. (2.30), (2.34) and (2.35), the scalar potential is obtained as:

$$\Phi_{pol}(\mathbf{r}_{ele}) = i\sqrt{\frac{1}{\Omega} \frac{\hbar\omega_0}{2\epsilon_0\kappa}} \sum_{\mathbf{q}} \frac{1}{|\mathbf{q}|} \left[b_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}_{ele}} - b_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}_{ele}} \right]. \quad (2.37)$$

Inserting the scalar potential into Eq. (2.31) we obtain the EPI term as:

$$\hat{H}_{int} = i\sqrt{\frac{1}{\Omega} \frac{e^2}{\epsilon_0} \frac{\hbar\omega_0}{2\kappa}} \sum_{\mathbf{q}} \frac{1}{|\mathbf{q}|} \left[b_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}_{ele}} - b_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}_{ele}} \right]. \quad (2.38)$$

At this point, we arrived at a fully quantum mechanical description of large continuum polarons, for which the Hamiltonian takes the form:

$$\hat{H}_{\text{Fröhlich}}(\mathbf{r}_{ele}) = \frac{\mathbf{p}_{ele}^2}{2m^*} + \hbar\omega_0 \sum_{\mathbf{q}} b_{\mathbf{q}}^\dagger b_{\mathbf{q}} + \sum_{\mathbf{q}} g(\mathbf{q}) b_{\mathbf{q}}^\dagger e^{-i\mathbf{q}\cdot\mathbf{r}_{ele}} + \text{h.c.} \quad (2.39)$$

where

$$g(\mathbf{q}) = \frac{i}{|\mathbf{q}|} \sqrt{\frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \frac{\hbar\omega_0}{2} \frac{1}{\kappa}} \quad (2.40)$$

is the matrix element of the Fröhlich type EPI and exhibits a long-range character.

2.2.2 Ground-state solutions of the Fröhlich model

The Fröhlich model is perhaps one of the most popular problems in theoretical physics, it provides the earliest and simplest framework to study particle-field interactions using a Galilean invariant Hamiltonian. Given that the model is potentially transferable to more complicated phenomena in particle physics, quantum field theorists such as Bogoliubov, Lee and Feynman have extensively worked on solutions for this model during the early 1950s. The problem belongs to a nonlinear system, and there is no exact analytical solution; only perturbative or variational solutions were available in that period of time. The first numerical solution based on diagrammatic Monte Carlo was only achieved in 1998 [20]. In many of these pioneer studies [9, 13, 70], the

ground-state solution of the model was determined by using the variational method. However, it turns out that the choice of the variational ansatz was tricky because the Hamiltonian contains the coupling strength α . While the variational ansatz considered by Lee only works in the weak coupling regime and leads to a ground-state energy $E_0 = -\alpha\hbar\omega_0$ [9], the different ansatz considered by Pekar yields a rather different ground-state energy $E_0 = -\frac{25}{256}\alpha^2\hbar\omega_0$, which is however valid only in the strong coupling regime ($\alpha \gg 1$) [70]. This leads to the hypothesis that the ground-state solution of the model may exhibit a first order phase transition as a function of α , but it was later proven that such transition does not exist at least for the ground-state solution of the Fröhlich model [71]. On top of these “direct variational” approaches, Feynman derived an effective “phonon free” action for the Fröhlich model by using a similar analogy in quantum electrodynamics. Using his path-integral formulation, he worked out an upper bound of the ground-state energy [13]:

$$\frac{E_0}{\hbar\omega_0} \leq \min_{v,w} \frac{3(v-w)^2}{4v} - \frac{\alpha v}{\sqrt{\pi}} \int_0^\infty \left[\frac{v^2 - w^2}{v} (1 - e^{-vx}) \right]^{-\frac{1}{2}} e^{-x} dx, \quad (2.41)$$

where the arguments u, v are variational parameters. Since the Feynman solution Eq. (2.41) carries an implicit integral with two minimisation arguments, one needs to solve a coupled integral equation. Therefore, the ground-state energy by Feynman cannot be expressed analytically in coupling strength α . However, it can be easily solved by performing a numerical minimisation, and we illustrate this result in Fig. 2.2. It is interesting to see that the ground-state solution of the Fröhlich model simply has three different characters with respect to the α . The variational method by Lee *et al.* [9] agrees perfectly with Feynman in small α , whereas the Landau-Pekar agrees very well at large α but with a small offset of $1.9 \hbar\omega_0$. For $\alpha \approx 5 - 8$ only the Feynman approach provides accurate solutions to the model. From this discussion, we learn that the Landau-Pekar model can be regarded as a particular solution to

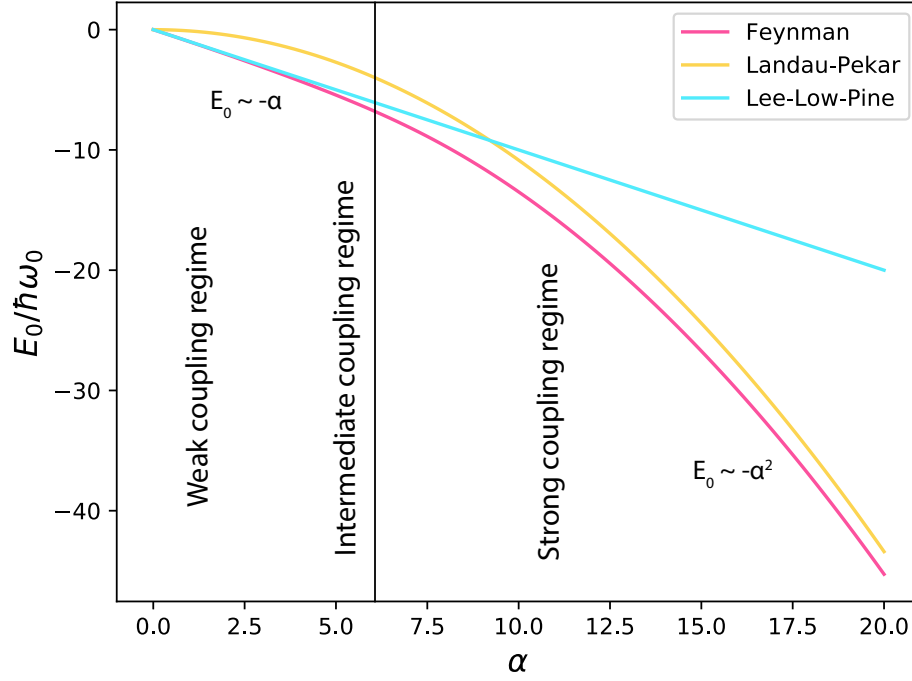


Figure 2.2: Ground-state energy of the Fröhlich model against coupling strength α

the quantum Fröhlich model at the strong coupling limit, i.e. large number of excited phonons coupled to the added electron makes it effectively trapped in the potential [24]. Since the DFT formalism also uses the adiabatic approximation in order to simplify the many-body problem, it is expected that the polaron energies calculated by DFT would also inherit the same character as the LP solution.

2.2.3 The Holstein model

In addition to the primary studies of continuum polaron models [9, 10, 12, 13, 70, 72–74] developed from the 1950s, there continues to be substantial attention to consider the short-range EPs present in molecular crystals [14, 75–77], leading to the formation of a small-radius polaron localised on a few atomic sites. In such systems, the electrons are tightly bound by the atomic potential to form narrow bands. The electronic bandwidth is determined by the hopping between the atomic orbitals in the

nearest neighbour atoms. The necessity to consider the discrete polaron model (e.g. Holstein model) here lies on the fact that the polaron radius becomes comparable to the lattice constant when the coupling is sufficiently large, and thus the effective mass approximation used in the continuum polaron models is no longer valid in this scenario. Therefore, the description of the small polaron problem is better considered by using a tight-binding model, as firstly introduced by Holstein in 1959 [14]. Here we briefly review the original work by Holstein to get an intuition on how the small polaron behaves. The Hamiltonian describing a tightly-bound electron which interacts with a one-dimensional molecular crystal is written as [14]:

$$H(\mathbf{r}, \{R_n\}) = -\frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}}^2 + \sum_{n=1}^N \left[-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial R_n^2} + \frac{1}{2} M \omega_0^2 R_n^2 \right] + V(\mathbf{r}, \{R_n\}), \quad (2.42)$$

where the first term is the kinetic energy of the electron, with $\mathbf{r}$ denoting the electron coordinate, the second term describes the vibrational motion of molecules, with R_n denoting the displacement of the n -th molecular site. Since only short-range interactions are present, the potential experienced by the electron can be decoupled as $V(\mathbf{r}, \{R_n\}) = \sum_n U(\mathbf{r} - n\mathbf{a}, R_n)$, where $\mathbf{a}$ is a lattice vector. Using the tight-binding approximation, the electron wavefunction should obey the stationary equation for the n -th site:

$$\left[-\frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}}^2 + U(\mathbf{r} - n\mathbf{a}, R_n) \right] \phi(\mathbf{r} - n\mathbf{a}, R_n) = E(R_n) \phi(\mathbf{r} - n\mathbf{a}, R_n), \quad (2.43)$$

where $\phi(\mathbf{r} - n\mathbf{a}, R_n)$ is the localised atomic orbital at the n -th molecular site, and $E(R_n)$ is the corresponding electron energy eigenvalue that depends on the displacement R_n . By constructing the polaron wavefunction Ψ for the Hamiltonian (2.42) using the atomic orbitals as the basis set $\Psi(\mathbf{r}, \{R_1, \dots, R_N\}) = \sum_n a_n(\{R_1, \dots, R_N\}) \phi(\mathbf{r} - n\mathbf{a}, R_n)$, following similar steps in the tight-binding model as set out in Ref. [14] one

can obtain a eigenvalue difference equation for the expansion coefficients a_n :

$$\left[\sum_m \left(-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial R_m^2} + \frac{1}{2} M \omega_0^2 R_m^2 \right) + E(R_n) \right] a_n(\{R_1, \dots, R_N\}) - \sum_{m=n+1, n-1} t_{nm} a_m\{R_1, \dots, R_N\} = E a_n\{R_1, \dots, R_N\}, \quad (2.44)$$

where $t_{nm} = -\langle \phi(\mathbf{r} - n\mathbf{a}) | U(\mathbf{r} - n\mathbf{a}) | \phi(\mathbf{r} - m\mathbf{a}) \rangle$ is the hopping matrix element, and we ignored the perturbation $W_n = \sum_{m \neq n} \langle \phi(\mathbf{r} - n\mathbf{a}) | U(\mathbf{r} - m\mathbf{a}) | \phi(\mathbf{r} - n\mathbf{a}) \rangle$ on $E(R_n)$ from molecules at the other sites. Since the small polaron wavefunction is unit cell periodic, it is convenient to label the displacements for all sites by a constant $R_n = R$. By imposing the translational invariance of the system, the expansion coefficient should have a trivial form $a_n = A e^{ikR_n}$. Using this choice, the energy eigenvalue E_k can be obtained as:

$$E_k = E(R) + \frac{N}{2} M \omega_0^2 R^2 - 2t \cos(ka), \quad (2.45)$$

where $t = t_{n, n\pm 1}$ is the hopping parameter. By linearising $E(R)$ with respect to the displacement as $E(R) = -KR$, we can adapt Eq. (2.44) using second quantisation. The Holstein Hamiltonian in a one-dimensional molecular crystal reads:

$$H = -t \sum_n (c_n^\dagger c_{n+1} + c_{n+1}^\dagger c_n) + \hbar \omega_0 \sum_n \left(b_n^\dagger b_n + \frac{1}{2} \right) + g_0 \sum_n c_n^\dagger c_n (b_n + b_n^\dagger), \quad (2.46)$$

where $c_n^\dagger(c_n)$ is the fermionic creation (annihilation) operator for an electron at n -th site, $b_n^\dagger(b_n)$ is the bosonic creation (annihilation) operator for a phonon at n -th site and g_0 is the constant electron-phonon coupling which describes only the short-range interaction. Conceptually, the site representation of the Holstein Hamiltonian Eq. (2.46) is mathematically similar to the Fröhlich Hamiltonian Eq. (2.39) that we met earlier. Apart from the motion of electrons through different mechanisms between the two models, the constant e-ph coupling that appears in the Holstein model also

provides a modified scenario, which is complementary to the Fröhlich-type long-range interaction.

It is also worthwhile to discuss the ground state energies of self-trapped polarons described in the Holstein model, which happens when $g_0/\hbar\omega_0 \gg 1$. In this limit, the Holstein model can admit two distinctive solutions: namely a small radius polaron when the hopping $t/\hbar\omega_0 \ll 1$, or a large radius polaron when $t/\hbar\omega_0 \gg 1$. The solution for the small-radius polaron takes the translational-invariant form [78]:

$$\psi_k(x) = \frac{1}{\sqrt{N}} \sum_n e^{ikna} \langle x | c_n^\dagger e^{\frac{g_0}{\hbar\omega_0}(b_n^\dagger - b_n)} | 0 \rangle, \quad (2.47)$$

and the corresponding energy spectrum is:

$$\frac{E_k}{\hbar\omega_0} = -2 \left(\frac{t}{\hbar\omega_0} \right) e^{-\frac{g_0^2}{\hbar^2\omega_0^2}} \cos(ka) - \frac{g_0^2}{\hbar^2\omega_0^2}. \quad (2.48)$$

When the EP coupling is turned off by setting $g_0 = 0$, the conventional tight-binding solution is recovered. More interestingly, the hopping parameter carries the factor $e^{-\frac{g_0^2}{\hbar^2\omega_0^2}}$ which is known as the Holstein reduction factor [79]. This implies that the electronic bandwidth contracts due to the polaron formation. At the same time, the ground state energy is offset by $-\frac{g_0^2}{\hbar\omega_0}$. Such polaronic features have previously been found in the Landau-Pekar/Fröhlich model, in which the polaron energy vs momentum dispersion relation in the strong-coupling limit is given by $E(P) = -0.1085\alpha^2\hbar\omega_0 + \frac{P^2}{2M_{\text{eff}}}$, where M_{eff} is the polaron effective mass [80]. Since the self-trapped polaron is heavier than the band electron, the polaron band in the Fröhlich model turns out to be narrower than the Bloch bands. It is useful to note that the role of the e-ph coupling has a consistent renormalisation effect on the electronic bands in both models. Furthermore, the Holstein model can even describe the large polaron limit. To this end, we can write the fermionic operators by brackets and

Fourier-transform the bosonic operators as:

$$\begin{aligned} c_n &= |0\rangle \langle n|, \\ b_n &= \frac{1}{\sqrt{N}} \sum_k b_k e^{ikna}. \end{aligned} \quad (2.49)$$

Replacing Eq. (2.49) into Eq. (2.46) and taking $|\Psi\rangle = \sum_n a_n |n\rangle$, the second-quantised electronic term in the Holstein Hamiltonian is degraded by performing $\hat{H} = \langle\Psi|H|\Psi\rangle$:

$$\hat{H} = -t \sum_n (a_n^* a_{n+1} + a_n a_{n-1}^*) + \hbar\omega_0 \sum_k b_k^\dagger b_k + \frac{g_0}{\sqrt{N}} \sum_{n,k} |a_n|^2 (b_k e^{ikna} + b_k^\dagger e^{-ikna}). \quad (2.50)$$

If we consider the large polaron limit with a wide electronic bandwidth $t/\hbar\omega_0 \gg 1$ and using the finite difference scheme $a_{n\pm 1} = a_n \pm \frac{\partial a_n}{\partial(na)}a + \frac{1}{2} \frac{\partial^2 a_n}{\partial(na)^2}a^2$, the equivalent 1-D Fröhlich model is recovered by taking the continuum limit ($x = na$) [78]:

$$\hat{H} = -\frac{\hbar^2}{2m_x^*} \frac{\partial^2}{\partial x^2} + \hbar\omega_0 \sum_k b_k^\dagger b_k + \frac{g_0}{\sqrt{N}} \sum_k (b_k e^{ikx} + b_k^\dagger e^{-ikx}), \quad (2.51)$$

where the one-dimensional effective mass is related to the tight-binding parameter by

$$m_x^* = \frac{\hbar^2}{2ta^2}.$$

2.2.4 Toward model studies into *ab initio* formalism

Although the Holstein model goes somewhat beyond the Fröhlich model, it is rather difficult to make it widely applicable. Unlike the Fröhlich electron-phonon (e-ph) matrix element, which can be expressed using intrinsic properties of the crystal such as phonon frequencies and dielectric constants, however, there is a missing connection between the the constant e-ph coupling g_0 in the Holstein model and the microscopic detail of the crystal. As a result, studies of the Holstein model mostly focus on toy systems, in which the solution is either analysed in a qualitative manner or is obtained by fitting the model parameters to *ab initio* calculations.

In the subsequent years, more sophisticated models were proposed to generalise the Fröhlich and Holstein models to cover more realistic systems [42, 81–86], such as small-bandwidth polar materials and high-temperature cuprate superconductors. These works involved introducing the periodic lattice potential in the Fröhlich model to describe the discrete nature of lattice, incorporating the Fröhlich interaction and multiple bands into the discrete Holstein model, investigating dispersive phonons and studying charge transport properties by using these models [87–89]. Furthermore, numerous numerical approaches based on the exact diagonalisation method [90], continuous-time quantum Monte Carlo [91], and diagrammatic Monte Carlo algorithms [92, 93] moved successful for solving various polaron models with arbitrary coupling strengths, and served as a complement to the earlier variational methods.

More importantly, the variational and numerical studies clarified that the Fröhlich and dispersive Holstein models are asymptotically reduced to the same limit in both the weak and strong coupling regimes [69]. This conclusion based on the model studies implies that there is indeed some overlap between the two models as we have already seen once in Eq. (2.51). This motivates us to develop an unified *ab initio* formalism to study polarons, with precise accounts of electronic band structures, dispersive phonons and EPIs.

2.3 Experimental signatures of polarons

The interest in searching polarons in the last century were certainly not retracted to theoretical developments, but also lead to experimental measurements of the properties of polarons. The early experimental observation of polarons was discovered in alkali halides using spin resonance techniques in 1955 [8], followed by the electrical conductivity measurements on uranium dioxide in 1963 [94]. Subsequently, several experimental techniques such as electron paramagnetic resonance (EPR), scanning tun-

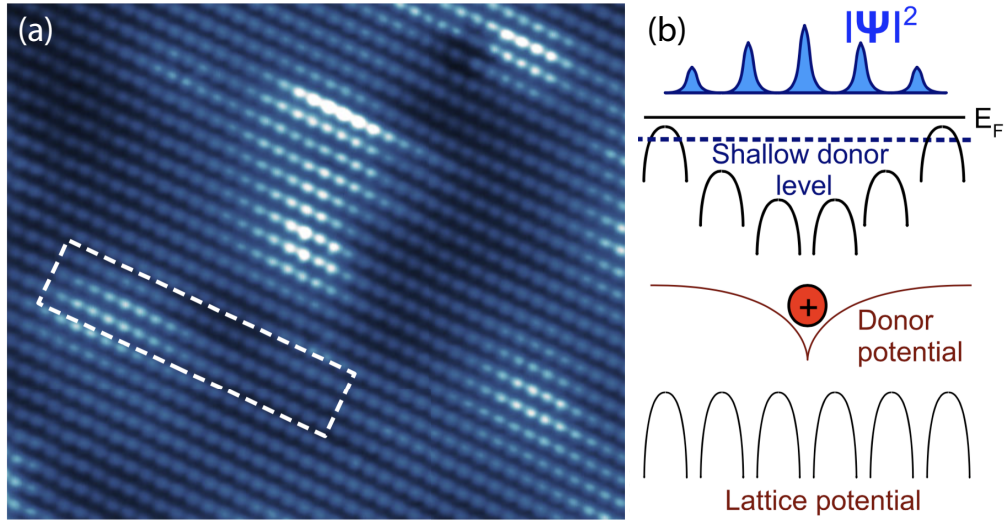


Figure 2.3: (a) The shallow donor state on anatase (101) obtained by constant-height STM measurement at $T = 6$ K and $V = +0.3$ V; and (b) a schematic view of a shallow donor state in semiconductors. (a) and (b) are adapted with permission from Ref. [51]. Copyright 2014 American Physical Society.

neling microscopy (STM) and angle-resolved photoemission spectroscopy (ARPES) have been used to investigate polaron properties from different viewpoints. For examples, EPR has been used to study trapped electrons in nanoparticles and films [66]; this technique enables small paramagnetic defects to be detected by hyperfine interactions with external magnetic fields [95]. On the other hand, STM has been employed to study defects in material surfaces; this technique enables a direct viewing of the spatial distribution of electrons and the localised defect states [51, 96]. In the following, we shall introduce the STM and the ARPES technique used to study the many-body spectral property of polarons.

2.3.1 Scanning Tunneling Microscopy

Scanning tunneling microscopy (STM) is an imaging technique for probing electron density of states at the atomistic scale. It allows for the study of pure crystalline samples with flat surfaces [97], through a precisely scanning over the crystal surface using a metal tip with angstrom-level control. The STM works by the effect of quan-

tum tunnelling, in which a bias voltage is applied between the the tip and the sample surface to measure the tunnelling current. This current is exponentially dependent of the tip-surface distance d , while the applied bias voltage V and the work function Φ are fixed. It can be related to the local electron density of states (LDOS) ρ_s near the Fermi level of the sample E_F at the probing surface via [98]:

$$I \propto V \rho_s(0, E_F) e^{-2 \frac{\sqrt{2m_e\Phi}}{\hbar} d}. \quad (2.52)$$

Since the current and the tip-surface distance are both free variables, it is possible to fix one and use the other to represent the contour of the LDOS (the surface topography). This means that there are two operation modes, namely the constant-height mode and the constant-current mode. The constant-height mode works straightly by fixing the tip-surface distance, followed by moving the tip and scan the sample surface line-by-line so that measured tunnelling currents at different positions correspond to the surface topography. In contrast, the constant-current mode works by maintaining a constant tunnelling current via adjusting the tip-surface distance. This requires an extra feedback circuit to monitor the current, and change the distance of the scanning tip using the piezoelectric effect [99]. In this mode, the tip-surface distances at difference surface points contribute to the surface topography. Overall, the constant-height mode is faster in comparison to the constant-current mode but it generally requires a smooth surface. The constant-current mode performs better for irregular sample surfaces but takes longer for the scanning [100].

Recently, the STM has been applied to access images of the polaronic states in anatase TiO_2 [51]. In that work, extra electrons are introduced to the anatase sample via the extrinsic Nb-dopant. An in-gap shallow donor state was identified at the Nb-doped anatase sample by the constant-height STM. Figure 2.3(a) presents a large-scale ($13 \times 11 \text{ nm}^2$) STM image of the anatase (101) surface, couples of shallow

donor states are clearly seen and found in different shapes. This was interpreted as the electron density is distributing over several unit cells as shown in Fig. 2.3(b), but the Nb donors reside at different depths and lead to an inhomogeneous net potential in which the electron is bound to. Interestingly the spatial extent of this STM electron distribution is characterised as $\Delta r[010] = 12 - 25 \text{ \AA}$ and $\Delta r[10\bar{1}] = 4 - 8 \text{ \AA}$, which are in close agreement with the Fröhlich model. The methods infer that this shallow donor state exhibits a polaron-like feature but physically cannot travel freely inside the crystal as the real polaron.

2.3.2 The ARPES technique

Angle-resolved photoemission spectroscopy (ARPES) is one of the most popular tools for studying the excited state properties of quantum matter. It has had a significant impact on the development of condensed matter physics because the single particle Green's function becomes experimentally measurable with this technique, which establishes a direct connection with theory calculations starting from the microscopic Hamiltonians and field-theoretic techniques. What makes ARPES powerful is that it can access the many-body effects beyond the mean-field level such as electron-hole, electron-phonon and electron-magnon interactions. This allows secondary excitations of electrons such as excitons, polarons [51], BCS-like Bogoliubov quasiparticles [101] and even the recently discovered chiral topological fermions [102] to be investigated.

ARPES probes the distribution of electrons in the energy-momentum space of the crystal. This distribution results from all possible interactions experienced by the electrons in the crystal, including the interaction of electrons with periodic crystal potential, lattice defects, impurities and thermal vibrations [103]. Overall, the photoemission process can be schematically explained by using the phenomenological three-step model [104]. In this model, the photoemission process involves three sequential steps: namely (1) the optical excitation of the bound electron by the incident

photon; (2) the excited electron travels to the sample surface and (3) extraction of the photoelectrons into vacuum by overcoming the surface confinement, so that photocurrents can be measured by the analyser. The photoemission intensity can be expressed by a product of the three independent events [105], i.e. the total probability for the optical transition approximated by Fermi's golden rule, the scattering probability for the travelling electrons and the transmission probability through the surface potential barrier. By invoking the sudden approximation to simplify the optical transition probability in step (1), one may express the ARPES photocurrent:

$$I(\mathbf{k}, \omega) \propto M_{\mathbf{k}}(h\nu) f(\omega) A(\mathbf{k}, \omega), \quad (2.53)$$

where $M_{\mathbf{k}}(h\nu)$ is proportional to the square modulus of the one-electron dipole matrix element $|M_{f,i}^{\mathbf{k}}|^2$ which is related to the probability of photon absorption. $f(\omega)$ is the Fermi-Dirac distribution, $A(\mathbf{k}, \omega)$ is the spectral function, which contains the Dirac delta functions from Fermi's golden rule by:

$$A(\mathbf{k}, \omega = h\nu - E_{kin}) = \sum_{\sigma=\pm 1} \sum_m |\langle \Psi_m^{N+\sigma} | c_{\mathbf{k}}^\sigma | \Psi_i^N \rangle|^2 \delta \left[h\nu - E_{kin} - \sigma (E_m^{N+\sigma} - E_i^N) \right], \quad (2.54)$$

where $\Psi_m^{N\pm 1}$ is the $(N \pm 1)$ -particle wavefunction with an excited state m , Ψ_i^N is the (N) -particle initial state, $c_{\mathbf{k}}$ is the annihilation operator for an electron with momentum $\mathbf{k}$, $h\nu$ is the energy of an incident photon and E_{kin} is the kinetic energy of the photoelectron. The distribution of the kinetic energies for photoelectrons are obtained by the analyser. The parallel component of the crystal momenta $\hbar \mathbf{k}_{||}$ to the surface can be deduced by measuring the polar emission angle θ and employing momentum conservation:

$$\begin{aligned} E_{kin} &= h\nu - \phi - |E_B|, \\ \hbar \mathbf{k}_{||} &= \sqrt{2mE_{kin}} \sin \theta, \end{aligned} \quad (2.55)$$

with ϕ denoting the material work function and E_B denoting the binding energy. Note that most ARPES experiments are typically performed at low photon energies where $h\nu < 100$ eV, in order to obtain better resolutions [105]. Fig. 2.4 (a)–(c) illustrates an example of the experimental ARPES intensity for TiO₂ anatase with different doping densities. We shall try to understand where these quasiparticle peaks and satellites are originated from by analysing further the spectral function in Eq. (2.54). The amplitude $|\langle \Psi_m^{N-1} | c_{\mathbf{k}} | \Psi_i^N \rangle|^2$ represents the probability of removing an electron from a N -particle system in state i , leaving the $(N - 1)$ -particles in a new state m . Suppose we have a non-interacting system in which the initial state i belongs to one of the eigenstate of the N -body Hamiltonian m_0 . The removal of an electron in such system will not induce a substantial change to the rest part of the electrons and therefore we expect that the amplitude $|\langle \Psi_m^{N-1} | c_{\mathbf{k}} | \Psi_{m_0}^N \rangle|^2 = \delta_{m,m_0}$. The spectral function for this system will collapse to a Dirac delta function $A(\mathbf{k}, \omega) = \delta(\omega - \epsilon_{\mathbf{k}})$, where $\epsilon_{\mathbf{k}} = E_{m_0}^N - E_{m_0}^{N-1}$. However, this is not the case in the interacting system in which more than one excited states will be generated, because the system will be largely perturbed while an electron is being removed. The spectral function for the interacting system will show a quasiparticle peak and satellites due to a number of excitation events.

2.3.3 Spectral feature of polarons in ARPES

In order to characterise the ARPES photocurrent for a system with polarons, for example the one reported in Fig. 2.4 (a)–(c), one has to calculate the spectral function for an electron-phonon interacting system. This requires the many-body Green's function formalism. Briefly, one define the time-ordered one-particle Green's function for a N -body system as $G(\mathbf{k}, t) = G^+(\mathbf{k}, t) + G^-(\mathbf{k}, t)$: where G^+ and G^- are given

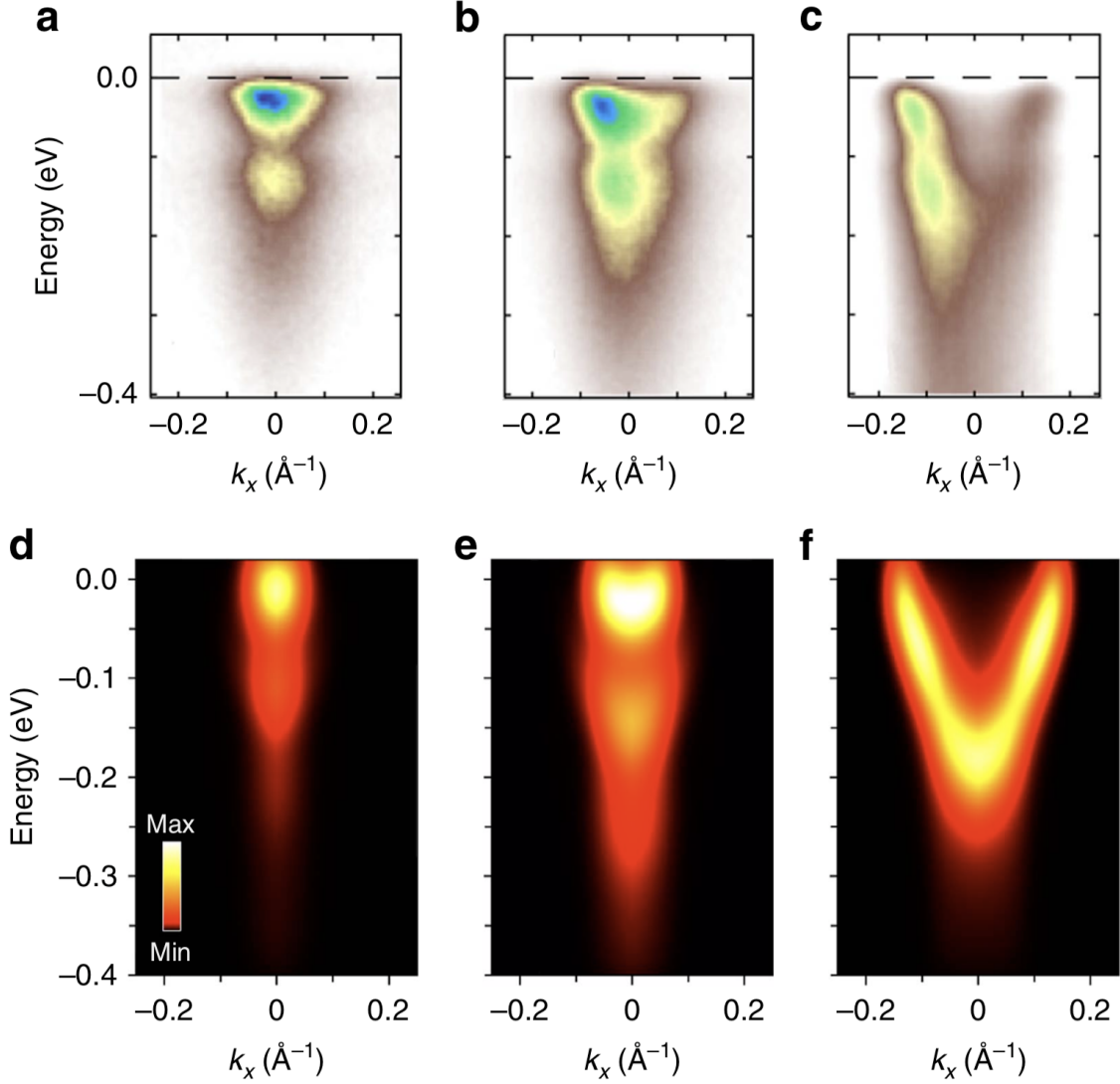


Figure 2.4: (a)-(c) Experimental ARPES spectra of anatase TiO_2 produced by Moser *et al.* with doping densities of (a) $5.0 \times 10^{18} \text{ cm}^{-3}$, (b) $3.0 \times 10^{19} \text{ cm}^{-3}$, and (c) $3.5 \times 10^{20} \text{ cm}^{-3}$ [11]. The wavevector k_x is along the path $\Gamma\Sigma$ in the Brillouin zone. Plots (a)-(c) are adapted with permission from Ref. [11]. Copyright 2013 American Physical Society. (d)-(f) Calculated spectral functions with the cumulant method using the same doping densities as in (a)-(c) produced by Verdi *et al.* [40]. Plots (d)-(f) are adapted with permission from Ref. [40]. Copyright 2017 Nature Publishing Group.

by:

$$\begin{aligned} G^+(\mathbf{k}, t) &= -i\theta(t) \langle \Psi_i^N | c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger | \Psi_i^N \rangle \\ G^-(\mathbf{k}, t) &= i\theta(-t) \langle \Psi_i^N | c_{\mathbf{k}}^\dagger c_{\mathbf{k}}(t) | \Psi_i^N \rangle \end{aligned} \quad (2.56)$$

Suppose that Ψ_i^N is the eigenstate of the N -body Hamiltonian, i.e. $\hat{H}\Psi_i^N = E_i^N \Psi_i^N$.

The Lehmann representation of the one-particle Green's function reads [106]:

$$G(\mathbf{k}, \omega) = \sum_m \frac{|\langle \Psi_i^{N+1} | c_{\mathbf{k}}^\dagger | \Psi_i^N \rangle|^2}{\omega - [E_m^{N+1} - E_i^N] + i\eta} + \sum_m \frac{|\langle \Psi_i^{N-1} | c_{\mathbf{k}} | \Psi_i^N \rangle|^2}{\omega + [E_m^{N-1} - E_i^N] - i\eta}. \quad (2.57)$$

It is important to point out that the spectral function Eq. (2.54) is indeed related to the time-order Green's function through $A(\mathbf{k}, \omega) = -\frac{1}{\pi} \text{Im } G(\mathbf{k}, \omega)$. Although the Lehmann representation Eq. (2.57) is exact, the eigenvalues and states of the interacting Hamiltonian are generally inaccessible. To overcome this difficulty, one can borrow the field-theoretic diagrammatic technique by expanding the time-ordered Green's function in the interacting picture as [106]:

$$G(\mathbf{k}, t) = \sum_{n=0}^{\infty} \frac{(-i)^{n+1}}{n!} \int_{-\infty}^{\infty} dt_1 \dots \int_{-\infty}^{\infty} dt_n \langle | \hat{T} [\hat{V}_I(t_1) \dots \hat{V}_I(t_n) c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger] | | \rangle, \quad (2.58)$$

where $\hat{V}_I(t) = e^{i\hat{H}_0 t} \hat{V} e^{-i\hat{H}_0 t}$ is the interacting potential $\hat{V}$ in the Heisenberg picture, which evolves according to the non-interacting Hamiltonian $\hat{H}_0$. The Dyson series can be obtained by the Fourier transform of Eq. (2.58):

$$G(\mathbf{k}, \omega) = G^0(\mathbf{k}, \omega) + G^0(\mathbf{k}, \omega) \Sigma(\mathbf{k}, \omega) G(\mathbf{k}, \omega). \quad (2.59)$$

The Σ is known as the self-energy which encodes the information about all the many-body interactions in the system. The calculations of self-energies are based on the diagrammatic perturbative series for a specific type of interaction, which requires

specific Feynman's rules. For example, the lowest order self-energy for a system with EPIs is the Fan-Migdal (FM) self-energy which has the form [107]:

$$\Sigma_{nn'}^{(1),\text{FM}}(\mathbf{k}, \omega) = \frac{1}{\hbar} \sum_{m\nu} \int \frac{d\mathbf{q}}{\Omega_{BZ}} g_{mn\nu}^*(\mathbf{k}, \mathbf{q}) g_{mn'\nu}(\mathbf{k}, \mathbf{q}) \times \left[\frac{n_{\mathbf{q}\nu} + 1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \frac{\epsilon_{m\mathbf{k}+\mathbf{q}}}{\hbar} - \omega_{\mathbf{q}\nu} + i\eta} + \frac{n_{\mathbf{q}\nu} + f_{m\mathbf{k}+\mathbf{q}}}{\omega - \frac{\epsilon_{m\mathbf{k}+\mathbf{q}}}{\hbar} + \omega_{\mathbf{q}\nu} - i\eta} \right], \quad (2.60)$$

where g is the e-ph matrix element from *ab initio*, $\mathbf{q}$ is the phonon wavevector, Ω_{BZ} is the volume of the first Brillouin zone, $f_{m\mathbf{k}}$ is the Fermi-Dirac occupation factor for an electronic state $|n\mathbf{k}\rangle$, $n_{\mathbf{q}\nu}$ is the Bose-Einstein occupation factor for a phonon state $\mathbf{q}\nu$, $\omega_{\mathbf{q}\nu}$ is the phonon frequency and η is the infinitesimal that defined in the propagator. In practice, expanding the self-energies to high diagrammatic orders is challenging because this requires the evaluation of many high-dimensional integrations. In model calculations of polarons, it was possible to access higher order diagrams using stochastic sampling methods, for example the diagrammatic Monte Carlo [20] which takes advantages of simple analytical forms of propagators and electron-phonon vertex. So far, many *ab initio* studies of EPI in materials are still rely on the Fan-Migdal self-energy. Taking the form of the non-interacting Green's function $G^0(\mathbf{k}, \omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}} + i\eta}$ for an insulator, the Green's function in Eq.(2.59) can be expressed using the complete form of self-energies:

$$G(\mathbf{k}, \omega) = \frac{1}{\omega - \epsilon_{\mathbf{k}} + \Sigma(\mathbf{k}, \omega) + i\eta}, \quad (2.61)$$

The spectral function is $A = -\frac{1}{\pi} \text{Im}G$, therefore:

$$A(\mathbf{k}, \omega) = -\frac{1}{\pi} \frac{\Sigma''(\mathbf{k}, \omega)}{[\omega - \epsilon_{\mathbf{k}} - \Sigma'(\mathbf{k}, \omega)]^2 + [\Sigma''(\mathbf{k}, \omega)]^2}. \quad (2.62)$$

In the non-interacting system where the self-energy is zero, the spectral function is simply a delta function $A(\mathbf{k}, \omega) = \delta(\omega - \epsilon_{\mathbf{k}})$. In contrast, the spectral function (2.62) behaves as a Lorentzian lineshape function in the interacting system where

the interaction is present. This is known as the quasiparticle peak, and its width is controlled by the magnitude of the imaginary part of the self-energy Σ'' , whereas the bands associated to the non-interacting system are renormalised by the real part Σ' . The imaginary part of the self-energy also carries additional information about the excited particles. For instance, the quasiparticle lifetime $\tau_{n\mathbf{k}}$ can be determined from the width of the quasiparticle peak by $\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \Sigma''_{n\mathbf{k}}$, and additional excitations (also known as satellites) beyond the quasiparticle peak that are expected in ARPES measurements can also be predicted by this many-body Green's function formalism.

Although the many-body formalism in Eqs. (2.60)–(2.62) used to investigate the spectral feature of polarons has been established for a while, it was still a big challenge to study this from the *ab initio* point of view because the Fan-Migdal self-energy Eq. (2.60) could only be obtained in a good accuracy by a heavy $\mathbf{q}$ -grid sampling. The efficiency of calculating e-ph matrix element and Fan-Migdal self-energy for polar materials was largely improved only after 2015 when the closed form of the *ab initio* Fröhlich matrix element was proposed by Sjakste *et al.* [108]. Nevertheless, the Fan-Migdal term was not sufficient to predict the satellite structure correctly, as mentioned in few studies of electron-plasmon interactions [109–111]. In order to recover certain amounts of many-body effects in the spectral function, in 2017 Verdi *et al* [40] borrowed the cumulant expansion technique which has been reported to successfully reproduce the plasmon satellites [112]. This extension works by including the so called cumulant factor $e^{C_{n\mathbf{k}}(t)}$ in the non-interacting Green's function, for example:

$$G_n(\mathbf{k}, t) = ie^{i\epsilon_{n\mathbf{k}}t/\hbar + C_{n\mathbf{k}}(t)}, \quad (2.63)$$

and the spectral function adapts as:

$$A(\mathbf{k}, \omega) = \frac{1}{\pi} \sum_n \text{Re} \int_{-\infty}^{\infty} dt e^{i(\omega - \epsilon_{n\mathbf{k}}\hbar)t} e^{C_{n\mathbf{k}}(t)}. \quad (2.64)$$

Here the cumulant function $C_{n\mathbf{k}}(t)$ is related to the Fan-Migdal self-energy by [40, 111]:

$$C_{n\mathbf{k}}(t) = \frac{1}{\pi\hbar} \int_{-\infty}^{\infty} d\omega \operatorname{Im}\Sigma_{n\mathbf{k}}(\epsilon_{n\mathbf{k}}/\hbar - \omega) \frac{e^{i\omega t} - i\omega t - 1}{\omega^2} \theta(\hbar\omega - \epsilon_{n\mathbf{k}}). \quad (2.65)$$

The form of cumulant function (2.65) guarantees the lowest order self-energy to be coincident with the Fan-Migdal term if one expresses the Green's function (2.63) in the form of (2.61) [111]. The introduction of cumulant factor $e^{C_{n\mathbf{k}}(t)}$ enables some additional diagrams beyond the Fan-Migdal order to be captured, but it should be noted that many higher order diagrams are still missing in this method. Although it is difficult to establish the accuracy of the cumulant method, it still works exceptionally well in predicting the satellite structure in comparison to the Fan-Migdal approximation, and thus has become a mainstream method for calculating spectral properties of many quantum systems.

Now we restrict ourselves to examine the spectral function for a prototypical polar material, for example TiO_2 anatase in which the Fröhlich coupling dominates. Figures 2.4(d)–(f) illustrate the calculated ARPES spectral function at the doping densities 5×10^{18} , 3×10^{19} and $3.5 \times 10^{20} \text{ cm}^{-3}$, which were directly compared with the experimental spectra in Figs. 2.4(a)–(c) obtained by the same doping densities as reported in Ref. [11]. In general, the calculated ARPES spectra by the cumulant method Eqs. (2.63)–(2.65) agrees well with the experimental ones. For the low doping densities such as the first two, a satellite appears clearly at the energy about 0.1 eV below the quasiparticle peak; the second satellite appears at again 0.1 eV below the first satellite and so on. This particular spectral feature in which multiple satellites exist and separated by a constant phonon energy is recognised as the polaron footprint [113]. It suggests polarons are generally present in many polar materials [11, 114–116], and therefore understanding its properties is important.

Since these experimental and theoretical ARPES spectra agree exceptionally well and calling for the existence of polarons in many polar materials, it is desirable to have an outstanding approach which can offer not only the excitation energies for this quasiparticle, but also the fascinating microscopic structure which illustrates how an electron is trapped in different materials. As the broader concept of polarons appeals to the general interests from many areas, this thesis is devoted to develop a theory of polarons in crystal lattices from first principle.

Chapter 3

First principle methods for extended systems

This chapter outlines the basic electronic structure methods for calculating the ground state electronic and vibrational properties of solids. The basis of such calculations are density functional theory (DFT) and density functional perturbation theory (DFPT). The central theme of this chapter is to introduce some practical concepts and useful tools for the understanding of *ab initio* calculations in computational materials science. At the beginning, we review the theoretical foundations of Kohn-Sham DFT and explain its capability in modelling many electrons system. Next, we describe how to formulate a lattice dynamics problem by using ground state DFT as a basic step, leading to the so-called DFPT approach. Finally, we lay out the key steps toward practical calculations of electron-phonon interactions by introducing the concept of maximally localized Wannier functions (MLWFs). At the end of the section, we also discuss a method that enable efficient calculations of electron-phonon interactions in polar materials.

3.1 Electronic structure in solids

3.1.1 The electronic structure problem

The fundamental building blocks of crystalline solids at the atomic scale are unit cells that are repeated periodically in space, where each unit cell hosts a collection of

interacting electrons and atomic nuclei. Before getting to describe the extended solid-state system which requires periodic boundary conditions, it is useful to begin with a simpler molecular-like system without periodicity. The many-body Hamiltonian that describes such an interacting system with N electrons and M nuclei maybe written as [117]:

$$\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i^N \nabla_i^2 - \sum_I^M \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i,I}' \frac{1}{4\pi\epsilon_0} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{i,j}' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{I,J}' \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (3.1)$$

where indices i and I label the i -th electron with mass m_e and the I -th nuclei with mass M_I respectively, Z_I is the atomic number and $\sum'$ is the primed sum. $\{\mathbf{r}_i|i = 1, 2, \dots, n\}$ and $\{\mathbf{R}_I|I = 1, 2, \dots, M\}$ denote the sets of electron and nuclear coordinates respectively. The time evolution of the many-body wavefunction $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}, t)$ follows the Schödinger's equation $i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi$. Obtaining the exact stationary states and energies for more than few atoms is extremely difficult, because these variables are mutually correlated and form a high-dimensional wave function. In order to obtain the solutions, it is useful to simplify the complexity of the many-body wave function by considering that the instantaneous effects on the electrons due to the nuclear motion are negligible. This is known as the Born-Oppenheimer (BO) approximation, which has been applied in many electronic structure calculations, although it is known the BO approximation fails to describe the nuclear quantum effects [118]. Once the wave function is decoupled in the adiabatic form $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = \psi_e(\{\mathbf{r}_i\})\chi(\{\mathbf{R}_I\})$, the many-body Schödinger's equation as described in Eq. (3.1) maybe disentangled as the coupled equations:

$$\left[-\sum_i^N \frac{\hbar^2 \nabla_i^2}{2m_e} - \sum_{i,I}' \frac{1}{4\pi\epsilon_0} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{i,j}' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi_e(\{\mathbf{r}_i\}) = E_e(\{\mathbf{R}_I\}) \psi_e(\{\mathbf{r}_i\}), \quad (3.2)$$

$$\left[-\sum_I^M \frac{\hbar^2 \nabla_I^2}{2M_I} + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{I,J}' \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + E_e(\{\mathbf{R}_I\}) \right] \chi(\{\mathbf{R}_I\}) = E^{tot} \chi(\{\mathbf{R}_I\}). \quad (3.3)$$

In these equations, we see that the kinetic energy (KE) for the nuclei is negligible in comparison to the electrons, because of their heavier mass. Therefore, the nuclei can be effectively viewed as classical point charges which do not exhibit any quantum effects. Since the KE term in Eq. (3.3) is dropped and the equation simply reduces to the classical potential energy surface, the problem is reduced to determining the electronic wave function and the eigenenergies for a given nuclear configuration $\{\mathbf{R}_I\}$. The eigenenergies for the electrons $E_e(\{\mathbf{R}_I\})$ together with the Coulomb repulsion of the nuclei form the adiabatic potential energy surface, which can be used to determine the Hellmann-Feynman force acting on the nuclei:

$$\mathbf{F}_I = -\nabla_I \left[E_e(\{\mathbf{R}_I\}) + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{I,J}' \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \right]. \quad (3.4)$$

Once these equations describing the adiabatic electronic structure are written down, the challenge is solving Eq. (3.2). This turns out to be as challenging as the original problem because Eq. (3.2) still contains the repulsive potential between the electrons. In 1927, Hartree employed the variational principle and a product ansatz to obtain an approximated ground state solution for the many electrons problem but it was found that this ansatz misses the important antisymmetric property of fermions. Later it was proposed to use the Slater determinant instead of the Hartree ansatz to recover the antisymmetry in the trial ansatz. This can be written as

$$\psi_e(\{\mathbf{r}_i\}) = \sqrt{n!} \hat{A}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\dots\psi_n(\mathbf{r}_n)], \quad (3.5)$$

where n denotes the number of electrons, $\psi_i(\mathbf{r}_i)$ denotes the wavefunction for the i th electron with coordinate $\mathbf{r}_i$, and

$$\hat{A} = \frac{1}{n!} \sum_{p \in S_n} (-1)^p \hat{P}(p) \quad (3.6)$$

is known as the antisymmetrising operator, and contains a collection of permutation operators used for implementing the Pauli principle for the fermions, with $(-1)^p \hat{P}(p)$ denoting the exchange operator and S_n denoting the symmetric group on n letters. Taking the expectation value for the electronic Hamiltonian (3.2) by using the Slater determinant (3.5), the ground state energy $E_e = \langle \psi_e | \hat{H} | \psi_e \rangle$ may be expressed as:

$$E_e = \sum_i \int d\mathbf{r} \psi_i^*(\mathbf{r}) \left[-\frac{\hbar^2 \nabla^2}{2m_e} + V_n(\mathbf{r}) \right] \psi_i(\mathbf{r}) + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{i,j} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{|\psi_i(\mathbf{r}_1)|^2 |\psi_j(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ - \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{i,j} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\psi_i^*(\mathbf{r}_1) \psi_j^*(\mathbf{r}_2) \psi_j(\mathbf{r}_1) \psi_i(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (3.7)$$

where $V_n(\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0} \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}$ is the bare nuclear potential, and the indices i, j run over the occupied states. The first term on the right hand side of Eq. (3.7) describes the non-interacting energy, the second term describes the classic electrostatic energy also known as the Hartree energy, and the third term is known as the Fock exchange energy, which results from the antisymmetrisation. The variation of this energy functional $E_e[\{\psi_k^*, \psi_k | 1 \leq k \leq n_{\text{occ}}\}]$ with respect to ψ_i^* leads to the Hartree-Fock equations. Moreover, if we take the expectation value on the number operator by using the Slater determinant (3.5), the total electron density of the system can be expressed as the sum of the individual densities contributed by the occupied states:

$$n(\mathbf{r}) = \langle \psi_e | \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) | \psi_e \rangle = \sum_{i \in \text{occ}} |\psi_i(\mathbf{r})|^2. \quad (3.8)$$

The development of the Hartree-Fock method led to first *ab initio* calculations, which are accurate for many systems even though the correlation energy is missing in the formalism. However, the non-local nature of the Fock exchange energy makes it computationally unfavourable for extended systems.

3.1.2 Density functional theory

In two landmark papers in *ab initio* electronic structure theory, Hohenberg and Kohn and Kohn and Sham in the 1960s formulated an alternative approach to the many electrons problem, which is nowadays called density-functional theory (DFT). Unlike the equations in Hartree-Fock theory which keep track of the orbitals for all the occupied electrons, DFT only needs to handle the electron density $n(\mathbf{r})$. This simplifies considerably the structure of the problem. To understand DFT we can express the energy functional using the many-body wave function $\psi_e(\{\mathbf{r}_i\})$ in the most general way:

$$E_e = \int d\mathbf{r}_1 \dots d\mathbf{r}_N \psi_e^*(\{\mathbf{r}_i\}) \left[- \sum_i^N \frac{\hbar^2 \nabla_i^2}{2m_e} + \sum_i V_n(\mathbf{r}_i) + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{i,j}' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \psi_e(\{\mathbf{r}_i\}). \quad (3.9)$$

In this form, it is clear that the E_e is a functional of the many-body wavefunction ψ_e , i.e. $E_e = F[\psi_e]$. Since the argument of the energy functional is a high-dimensional object that is difficult to compute, it would be desirable to express this functional in an alternative form with an argument of lower dimensionality.

The Hohenberg-Kohn theorem

Hohenberg and Kohn in 1964 [119], showed that the energy functional (3.9) depends solely on the electron density of the system, i.e. $E = F[n(\mathbf{r})]$. This is known as the Hohenberg-Kohn (HK) theorem, which forms the foundation of modern DFT. Now if we consider a system where its electrons experience an external potential say the

bare nuclear potential $V_n(\mathbf{r})$, the HK theorem states the following:

- (I). In the ground state, there exists an unique one-to-one mapping between the external potential and the electron density $n \leftarrow V_n$. Hence the total energy is also an unique functional of the electron density, i.e. $E = F[n(\mathbf{r})]$.
- (II). The total energy of the system E can be expressed as:

$$E_e[n(\mathbf{r})] = \int d\mathbf{r} n(\mathbf{r})V_n(\mathbf{r}) + \langle \psi_e | \hat{T} + \hat{W} | \psi_e \rangle [n]. \quad (3.10)$$

The second term involves both the kinetic energies $\hat{T} = -\sum_i \frac{\hbar^2 \nabla_i^2}{2m_e}$ and the electron repulsion $\hat{W} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum'_{i,j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$. This functional defines a variational problem to the energy, whose minimum is associated to the ground state electron density $n_0(\mathbf{r})$.

The Kohn-Sham equation

Since the HK-theorem states that the energy of the system depends solely on the ground state electron density, it is more convenient to separate the correlated part of the electronic motions from the total kinetic energy term. By grouping together this contribution with the exchange energy, Eq. (3.10) may be expressed as [120]:

$$E_e[n(\mathbf{r})] = \int d\mathbf{r} n(\mathbf{r})V_n(\mathbf{r}) + T_0[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})]. \quad (3.11)$$

where T_0 is the kinetic energy for the non-interaction electrons, E_H is the Hartree energy, and is the counterpart of the classical electrostatics energy for the electrons:

$$E_H[n(\mathbf{r})] = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (3.12)$$

and E_{xc} is called the exchange-correlation (xc) energy. Since the exact form of the exchange-correlation functional is still unknown, modern DFT calculations of solids mainly rely on semi-core apporimations, such as for example the local-density ap-

proximation (LDA) as a starting point. In LDA, the exchange energy E_x is based on the analytical form derived by considering the system as a homogeneous electron gas. The correlation energy E_c is, instead, parameterised by using the results from quantum Monte Carlo calculations [121]. If we variate the energy functional (3.11) with respect to the electron density $n(\mathbf{r})$, and followed by orthonormalising all the electronic states by performing a unitary transformation, we obtain the so-called Kohn-Sham (KS) equations as:

$$\left[-\frac{\hbar^2 \nabla^2}{2m_e} + V_{\text{KS}}[n](\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (3.13)$$

where the Kohn-Sham potential V_{KS} is given by:

$$V_{\text{KS}}[n](\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0} \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} + \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \quad (3.14)$$

The ground state electron density

$$n_0(\mathbf{r}) = \sum_{i \in \text{occ}} |\psi_i(\mathbf{r})|^2, \quad (3.15)$$

together with the KS eigenstates $\{\psi_i\}$ and energies $\{\epsilon_i\}$ can generally be found either by using the self-consistent field approach [117], or by performing a direct minimisation on the functional [122]. Finally, the total energy of the system can be expressed in terms of the solutions n_0 and $\{\epsilon_i\}$ [117]:

$$E_e[n_0] = \sum_{i \in \text{occ}} \epsilon_i - \left(E_H[n_0] + \int d\mathbf{r} \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \Big|_{n_0} n_0(\mathbf{r}) - E_{xc}[n_0] \right). \quad (3.16)$$

3.1.3 Periodicities in solids

So far we have focused on the electronic structure of an isolated system. We have not yet mentioned how to incorporate DFT for an extended solid-state system, which

contains infinite number of interacting electrons and nuclei. The problem is simplified by considering crystalline solids with translational symmetry.

The plane wave basis set

The atoms in a perfect crystalline solid form a periodic array of identical units, the primitive unit cell. The periodicity of the crystal imposes the same periodicity to the external potential experienced by the electrons, as well as on the electron density. Based on the Bloch theorem [123] and the BvK periodic boundary condition, the Bloch wavefunction is given by:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_p}} u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (3.17)$$

where N_p is the number of unit cell within the BvK supercell, $\mathbf{k}$ is the wavevector within the first Brillouin zone, which is used as an additional quantum number to incorporate the translational symmetry in the potential, $u_{n\mathbf{k}}(\mathbf{r})$ is a periodic function which has the same periodicity as the crystal, i.e. $u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R})$, and $\mathbf{R}$ is a lattice vector. Given its periodicity, the function $u_{n\mathbf{k}}$ can generally be expressed as a Fourier series:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_p\Omega}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \quad (3.18)$$

where $\mathbf{G}$ is the reciprocal lattice vector, Ω is the unit cell volume and $c_{n\mathbf{k}}(\mathbf{G})$ are the plane wave expansion coefficient for the KS state $|\psi_{n\mathbf{k}}\rangle$. The periodic function $u_{n\mathbf{k}}$ is normalised within the primitive unit cell, but the Bloch wavefunction is normalised in the BvK supercell. The latter choice is to ensure the normalisation of the electron density in this formalism. We note that the wavevector $\mathbf{k}$ is associated with the periodic boundary condition that defines the BvK supercell in real space, in which the Bloch state should satisfy $\psi_{n\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{T})$, where $\mathbf{T}$ is a superlattice vector. Since the Fourier series is defined on a complete metric space, it is naturally a complete

orthonormal basis set if all the $\mathbf{G}$ vectors are used. However, the number of $\mathbf{G}$ vectors used in practical calculations are truncated at the cutoff given by:

$$\frac{\hbar^2}{2m_e}|\mathbf{k} + \mathbf{G}|^2 \leq E_{\text{cut}}. \quad (3.19)$$

For each $\mathbf{k}$ -point, only $\mathbf{G}$ vectors with kinetic energy below this cutoff are included in the series. The truncation of the plane waves is reasonable because the Fourier coefficients of the KS potential $V_{\text{KS}}(\mathbf{G})$ decrease relatively fast with $\mathbf{G}$. Since the parameter E_{cut} controls the basis set errors, convergence test needs to be done as a priority step in any plane wave code. The KS equations (3.13) can also be expressed in the reciprocal space by using the plane wave expansion (3.18):

$$\sum_{\mathbf{G}'} \left[\frac{\hbar^2|\mathbf{k} + \mathbf{G}|^2}{2m_e} \delta_{\mathbf{G},\mathbf{G}'} + V_{\text{KS}}[n](\mathbf{G} - \mathbf{G}') \right] c_{n\mathbf{k}}(\mathbf{G}') = \epsilon_{n\mathbf{k}} c_{n\mathbf{k}}(\mathbf{G}). \quad (3.20)$$

In this form, the kinetic energy term contributes only to the diagonal entries, and the off-diagonal entries depend on the Fourier components of the KS potential. In modern DFT implementation Eq. (3.20) is never stored in the matrix form $H_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'}$, and the KS potential terms in Eq. (3.20) are evaluated on a real-space grid as $V_{\text{KS}}(\mathbf{r})\psi(\mathbf{r})$ and then rely on extensive use of the Fast-Fourier transform algorithm to optimise the calculations and the balance of memories.

Since the KS potential is formed by the electron density, it is convenient to relate the electron density in the reciprocal space $n(\mathbf{G})$ to the plane wave coefficients via:

$$n(\mathbf{r}) = \sum_{\mathbf{G}} n(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} = \sum_{\mathbf{G}} \left[\frac{1}{N_p\Omega} \sum_{n\mathbf{k},\mathbf{G}'} c_{n\mathbf{k}}^*(\mathbf{G}') c_{n\mathbf{k}}(\mathbf{G} + \mathbf{G}') \right] e^{i\mathbf{G}\cdot\mathbf{r}}. \quad (3.21)$$

A self-consistent system is formed by (3.20) and (3.21), in which the plane wave coefficients obtained by iterative diagonalisations are used to generate the electron density via (3.21) which will update the KS potential. Since the external and the

Hartree potential are local in $\mathbf{G}$ space, they can be expressed as a Ewald sum. The exchange-correlation potential is neither local in $\mathbf{G}$ space nor explicit in $n(\mathbf{G})$ [124], and is better to be evaluated in real-space first, and then followed by the Fast Fourier Transforms.

At last, we discuss the meaning of the $\mathbf{k}$ -point sampling in the first Brillouin zone. If we take the Bloch wavefunction defined in (3.17), the electron density can be expressed as $n(\mathbf{r}) = \frac{1}{N_p} \sum_{n\mathbf{k}} |u_{n\mathbf{k}}(\mathbf{r})|^2$. When we perform a real space integration on the electron density $n(\mathbf{r})$ over the BvK supercell, it can be shown that the average number of electrons in the unit cell is given by:

$$\frac{1}{N_p} \int_{SC} n(\mathbf{r}) d\mathbf{r} = \frac{1}{N_p} \sum_{\mathbf{k}} [n_{\text{occ}}] = n_{\text{occ}}, \quad (3.22)$$

where n_{occ} is the number of occupied states per unit cell, and the the summation domain of these discrete $\mathbf{k}$ -points is defined in the BvK boundary condition $e^{i\mathbf{k}\cdot\mathbf{T}} = 1$. The role of performing $\frac{1}{N_p} \sum_{\mathbf{k}} [\dots]$ is to average out the total number of electrons contributed individually from each unit cell. Therefore, the physical meaning to converge the $\mathbf{k}$ -point sampling in the self-consistently field calculations is that, ideally we would like to have a sufficiently large supercell to accommodate many different electron wavefunctions in order to simulate an extended crystal. This will give a good averaged electron density with a dense $\mathbf{k}$ -point sampling as:

$$n(\mathbf{r}) = \frac{1}{N_p} \sum_{\mathbf{k}} \left[\sum_{n \in \text{occ}} |u_{n\mathbf{k}}(\mathbf{r})|^2 \right]. \quad (3.23)$$

Since the shape of the electron wavefunctions should almost stay the same at two very closed $\mathbf{k}$ -points, we would expect only a small subset of different $\mathbf{k}$ -points will be sufficient to give a well-converged density. The Monkhorst-Pack method is widely used in modern plane wave DFT code, which generates special points in the Brillouin zone to achieve faster convergence than the uniform grid [125].

Pseudopotentials

Using plane wave basis carries several practical advantages, however, it also carries a weakness when describing the core electrons in the region of atomic nucleus. In order to resolve the fine structure of wavefunctions around the nuclei, a large set of Fourier components is required to describe the nodes; this can become very expensive computationally. Most of the physical and chemical properties of materials are determined by the valence electrons which participate in bonding. In contrast, core electrons are tightly bounded to the atomic nuclei and are chemically inert. Therefore, it is reasonable to introduce the concept of pseudopotentials [126, 127], which describes together the bare nucleus screened by the frozen core electrons as a single rigid ionic core.

There are several ways to construct a pseudopotential, from early empirical parameterisation to fully *ab initio* methods with improved efficiency and accuracy [128]. Apart from eliminating the undesirable core states, pseudopotentials are designed for replacing the nodal structure of the valence states by a smooth function to reduce the planewaves cutoff. This will produce a pseudo wavefunction that matches exactly the all-electron wavefunctions outside a set ionic radius r_c , but is smooth inside the ionic core. The other important feature for a pseudopotential is the transferability, which means that the same pseudopotential should perform reliably in broad class of environments, ranging from isolated molecules to bulk crystals [129].

A classic pseudopotential widely used in modern plane wave DFT code is called the norm-conserving pseudopotentials [130], which imposes the norm-conservation constrain between the (radial) pseudo wavefunctions u_l^{ps} and the (radial) all-electron wavefunctions u_{nl} [131]:

$$\int_0^{r_c} dr |u_l^{ps}(r; \epsilon_l^{ps})|^2 = \int_0^{r_c} dr |u_{nl}(r; \epsilon_{nl})|^2, \quad (3.24)$$

where n denotes the principal quantum number, l denotes the angular momentum quantum number, ϵ_{nl} is the all-electron eigenvalue and ϵ_l^{ps} is the pseudo eigenvalue. To construct the norm-conserving pseudopotentials *ab initio*, one has to obtain the all-electron solutions for an isolated atom using DFT. Since the screened KS potential for one atom exhibits a spherical symmetry, the usual KS equation can be simplified to the radial central equation:

$$\left[-\frac{\hbar^2}{2m_e} \frac{d^2}{dr^2} + \frac{l(l+1)\hbar^2}{2m_e r^2} + V(r) - \epsilon_{nl} \right] u_{nl}(r; \epsilon_{nl}) = 0, \quad (3.25)$$

where the screened all electron potential and the (radial) electron density ρ are given by:

$$V(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{Z}{r} + V_H[\rho](r) + V_{xc}[\rho](r), \quad (3.26)$$

$$\rho(r) = \frac{1}{4\pi} \sum_{n \in \text{occ}} \sum_{l=0}^{n-1} f_{nl} \left| \frac{u_{nl}(r; \epsilon_{nl})}{r} \right|^2. \quad (3.27)$$

Once obtained the all-electron solutions $u_{nl}(r; \epsilon_{nl})$ by solving Eqs. (3.25)–(3.27), one can map the outer core part of the all electron valence state $u_{nl}(r > r_c)$ to the pseudo wavefunction $u_l^{ps}(r > r_c)$. Both wavefunctions should have the same eigenvalue $\epsilon_l^{ps} = \epsilon_{nl}$, and the logarithmic derivatives of the wavefunctions should be consistent outside the ionic core to conserve the scattering properties. The next step is to construct the inner core part $u_{nl}(r < r_c)$. For example in the Troullier-Martin scheme one uses a 10-th order polynomial to optimise the smoothness [132]. The constructed pseudo wavefunction is nodeless, and one can calculate the screened pseudo potential by inverting Eq. (3.25) [131]:

$$V_l^{ps,screen}(r) = \epsilon_l^{ps} - \frac{l(l+1)\hbar^2}{2m_e r^2} + \frac{\hbar^2}{2m_e u_l^{ps}(r)} \frac{d^2}{dr^2} u_l^{ps}(r). \quad (3.28)$$

The last step is to obtain the bare ionic pseudopotential by removing the electronic

screening from the pseudo-density ρ_0 :

$$V_l^{ps}(r) = V_l^{ps,screen}(r) - V_H[\rho_0^{ps}](r) - V_{xc}[\rho_0^{ps}](r). \quad (3.29)$$

We note that this pseudopotential has been resolved in angular momentum, and importantly the ionic potential resolved at any l should behave the same outside the ionic radius, i.e. $V_l^{ps}(r > r_c) \equiv V_{loc}^{ps}(r) = -\frac{e^2}{4\pi\epsilon_0} \frac{Z_v}{|r|}$ with Z_v the number of valence electrons. Since norm-conserving pseudopotentials are derived in terms of a projection in the space of angular dependence, i.e. $\hat{V}^{ps} = \sum_{lm} |lm\rangle \hat{V}_l \langle lm|$, this leads to a semi-local pseudopotential in real-space with the kernel form:

$$V^{ps}(\mathbf{r}, \mathbf{r}') = V_{loc}^{ps}(r) \delta^{(3)}(\mathbf{r} - \mathbf{r}') + \frac{\delta(r - r')}{r^2} \sum_l \delta V_l^{ps}(r) \sum_m Y_{lm}^*(\Omega') Y_{lm}(\Omega), \quad (3.30)$$

where $Y_{lm}(\Omega)$ is the solid angle spherical harmonics, $\delta V^{ps}(r) = V_l^{ps}(r) - V_{loc}^{ps}(r)$ behaves as a short-range potential which vanishes beyond r_c . In the norm-conserving pseudopotential formalism, the bare nuclear potential (3.14) needs to be replaced as $V_{en}(\mathbf{r})\psi_i(\mathbf{r}) = \int d\mathbf{r}' V^{ps}(\mathbf{r}, \mathbf{r}')\psi_i(\mathbf{r}')$ in order to account for non-locality. More technical aspects on pseudopotentials can be found in the reviews [133, 134].

3.2 Vibrational properties of solids

3.2.1 Adiabatic lattice potentials

In Sec. 3.1, we only dealt with the basics of DFT and its practical usage for finding the ground-states energy and density. This involved only self-consistent field calculations based on the clamped-nuclei Kohn-Sham equation, where the ions are fixed at their equilibrium positions. However, it is also interesting to see how the electronic properties are changed by the motion of these ions. To do so, we need to understand

what is the underlying principle that determines the motions of ions in solids. Recall from Eq. (3.3) we have defined the potential energy surface (PES) under the adiabatic approximation for isolated systems. Now we proceed to the description of solids and consider a Born-von Karman (BvK) supercell of the crystal, containing N_p unit cells of volume Ω . We follow the same notation of Ref. [107] and use $\boldsymbol{\tau}_{\kappa p}$ and $\tau_{\kappa p \alpha}$ to indicate the position and Cartesian coordinates of the ion κ in the unit cell p along the Cartesian direction α , respectively. The ion κ has a charge $e Z_\kappa$. We adapt the pseudopotential formalism and the PES can be written as:

$$E(\{\boldsymbol{\tau}_{\kappa p}\}) = E_e(\{\boldsymbol{\tau}_{\kappa p}^0\}) + \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa p, \kappa' p', \mathbf{T}} \frac{Z_\kappa Z_{\kappa'}}{|\boldsymbol{\tau}_{\kappa p} - \boldsymbol{\tau}_{\kappa' p'} - \mathbf{T}|}, \quad (3.31)$$

where $E_e(\{\boldsymbol{\tau}_{\kappa p}^0\})$ is the ground state electron energy at the fixed ionic configuration $\{\boldsymbol{\tau}_{\kappa p}^0\}$ which can be obtained by DFT using Eq. (3.16), and the second term is the pairwise electrostatic repulsion between the ion, with $\mathbf{T}$ denoting the superlattice vector. The stabilisation of the lattice is achieved when the Hellmann-Feynman forces (3.4) on all ions are less than a set of numerical thresholds. In practice, direct minimisation of the PES can be performed by using conjugate-gradient method in which the Hellmann-Feynman forces are employed as the gradient functions.

Now let's consider the ions are slightly displaced from their equilibrium positions $\{\boldsymbol{\tau}_{\kappa p}^0\}$ and assume the electrons stay with density n . We can expand the PES about the equilibrium positions up to second order in the displacements:

$$E(\{\boldsymbol{\tau}_{\kappa p}\}) = E(\{\boldsymbol{\tau}_{\kappa p}^0\}) + \frac{1}{2} \sum_{\kappa p \alpha, \kappa' p' \alpha'} \frac{\partial^2 E}{\partial \tau_{\kappa p \alpha} \partial \tau_{\kappa' p' \alpha'}|_0} \Delta \tau_{\kappa p \alpha} \Delta \tau_{\kappa' p' \alpha'}. \quad (3.32)$$

Here the gradient term for the force at the equilibrium configuration vanished. We note that the second derivative of the PES with respect to the ionic coordinates is

the matrix of interatomic force constants (IFC), which is defined by:

$$C_{\kappa p\alpha, \kappa' p'\alpha'} = \frac{\partial^2 E[n; \{\boldsymbol{\tau}_{\kappa p}\}]}{\partial \tau_{\kappa p\alpha} \partial \tau_{\kappa' p'\alpha'} \big|_0}. \quad (3.33)$$

Once the force constants are determined, the equation of motion for the ions can be shown as:

$$M_\kappa \frac{\partial^2 \tau_{\kappa p\alpha}}{\partial t^2} = - \sum_{\kappa' p'\alpha'} C_{\kappa p\alpha, \kappa' p'\alpha'} \Delta \tau_{\kappa' p'\alpha'}. \quad (3.34)$$

Since the equation describes the harmonic motions of ions, it is reasonable to seek a normal mode solution of the type: $\Delta \tau_{\kappa p\alpha}(t) = \sum_{\mathbf{q}\nu} \chi_{\mathbf{q}\nu} e_{\kappa\alpha, \nu}(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{R}_p} e^{-i\omega_{\mathbf{q}\nu} t}$ where $\chi_{\mathbf{q}\nu}$ is the displacement amplitude for the phonon eigenmodes, and $e_{\kappa\alpha, \nu}(\mathbf{q})$ is a normal mode. By substituting the normal mode solution in Eq. (3.34) we obtain the eigenvalue problem:

$$\sum_{\kappa' \alpha'} D_{\kappa\alpha, \kappa' \alpha'}(\mathbf{q}) e_{\kappa' \alpha', \nu}(\mathbf{q}) = M_\kappa \omega_{\mathbf{q}\nu}^2 e_{\kappa\alpha, \nu}(\mathbf{q}), \quad (3.35)$$

where $D_{\kappa\alpha, \kappa' \alpha'}(\mathbf{q})$ is known as the dynamical matrix and is given by the Fourier transform of the matrix of interatomic force constants [107]:

$$D_{\kappa\alpha, \kappa' \alpha'}(\mathbf{q}) = \sum_{p'} C_{\kappa 0\alpha, \kappa' p'\alpha'} e^{i\mathbf{q} \cdot \mathbf{R}_{p'}}. \quad (3.36)$$

3.2.2 Density-functional perturbation theory

Since the dynamical properties in a crystal are governed by the matrix of interatomic force constants, there is a question on how to efficiently evaluate Eq. (3.33) without using finite differences. The so-called density-functional perturbation theory (DFPT) was developed to perform this task. A useful starting point is the alternative definition of IFC based on the derivative of the forces, i.e. $C_{\kappa p\alpha, \kappa' p'\alpha'} = -\partial_{\tau_{\kappa' p'\alpha'}} F_{\kappa p\alpha}$. The calculations of the Hellmann-Feynman forces involve only the electron-nuclei interactions

and the nuclei repulsion in the BO Hamiltonian (3.1). By evaluating the derivatives on the Hellmann-Feynman force using the chain rule, the complete expression which takes account of the electron responses to the displacements is [135]:

$$C_{\kappa p \alpha, \kappa' p' \alpha'} = \int d\mathbf{r} \frac{\partial n(\mathbf{r})}{\partial \tau_{\kappa' p' \alpha'}} \frac{\partial V_n(\mathbf{r}; \{\boldsymbol{\tau}_{\kappa p}\})}{\partial \tau_{\kappa p \alpha}} + \int d\mathbf{r} n(\mathbf{r}) \frac{\partial^2 V_n(\mathbf{r}; \{\boldsymbol{\tau}_{\kappa p}\})}{\partial \tau_{\kappa p \alpha} \partial \tau_{\kappa' p' \alpha'}} + \frac{\partial^2 E_N(\{\boldsymbol{\tau}_{\kappa p}\})}{\partial \tau_{\kappa p \alpha} \partial \tau_{\kappa' p' \alpha'}}. \quad (3.37)$$

In this expression, we almost have all the ingredients from the ground-state DFT calculations except for the electronic response $\frac{\partial n(\mathbf{r})}{\partial \tau_{\kappa' p' \alpha'}}$ to the ionic displacements. The calculation of the density variation can be performed using DFPT [135]. If we denote Δn as the first order correction of the electron density due to the ionic displacements, in the regime of linear response, Δn can be expressed in terms of the unperturbed Kohn-Sham (KS) states $\psi_i^0(\mathbf{r})$ and perturbed KS states $\Delta\psi_i(\mathbf{r})$:

$$\Delta n(\mathbf{r}) = 2\text{Re} \sum_{i \in \text{occ}} [\psi_i^0(\mathbf{r})]^* \Delta\psi_i(\mathbf{r}) \quad (3.38)$$

The perturbed KS states $\Delta\psi_i(\mathbf{r})$ can be obtained by expanding the KS equation (3.13) with a small perturbation λ . By collecting the terms in linear order of λ , the Sternheimer equation can be obtained as:

$$(\hat{H}_{\text{KS}}^0 - \epsilon_i^0) \Delta\psi_i(\mathbf{r}) = -[\Delta V_{\text{KS}}(\mathbf{r}) - \Delta\epsilon_i] \psi_i^0(\mathbf{r}), \quad (3.39)$$

where $\hat{H}_{\text{KS}}^0$ and ϵ_i^0 are the unperturbed KS Hamiltonian and energy respectively, $\Delta\epsilon_i = \langle \psi_i^0 | \Delta V_{\text{KS}} | \psi_i^0 \rangle$ is the first-order correction to the KS eigenenergy, the first-order correction to the unperturbed KS potential ΔV_{KS} can be written in the form of linear response [135, 136]:

$$\Delta V_{\text{KS}}(\mathbf{r}) = \Delta V_n(\mathbf{r}) + \int d\mathbf{r}' \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}(\mathbf{r}, \mathbf{r}') \right] \Delta n(\mathbf{r}'), \quad (3.40)$$

where $f_{xc}(\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E_{xc}[n]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')}$ is the exchange-correlation kernel. Eqs. (3.38)–(3.40) form a system of self-consistent equations which is known as the self-consistent Sternheimer system, similar to the KS system in DFT. The calculations procedure involves constructing a perturbed density and use it in Eq. (3.40) to form a perturbed KS potential ΔV_{KS} , followed by solving Eq. (3.39) for the perturbed KS states. Finally these perturbed states are used to update the perturbed electron density using Eq. (3.38). These steps are iterated self-consistently until the perturbed KS potential no longer varies. In this formulation, only the occupied electronic states obtained from the *prior* KS-DFT calculations enter the equation for $\Delta\psi_i(\mathbf{r})$.

In the practical implementation of DFPT for a periodic lattice, one uses the Bloch theorem and all quantities are computed in the primitive unit cells. Let us define:

$$\partial_{\tau_{\kappa\mathbf{q}\alpha}} = \sum_p e^{i\mathbf{q}\cdot\mathbf{R}_p} \partial_{\tau_{\kappa p\alpha}}, \quad (3.41)$$

where $\mathbf{R}_p$ is the p -th site lattice vector in real space and $\mathbf{q}$ is the phonon wavevector. Using Eq. (3.41), the dynamical matrix (3.36) and perturbed electron density (3.38) can be expressed as:

$$D_{\kappa\alpha, \kappa'\alpha'}(\mathbf{q}) = \frac{1}{N_p} \frac{\partial^2 E}{\partial \tau_{\kappa\mathbf{q}\alpha} \partial \tau_{\kappa'\mathbf{q}\alpha'}^*}, \quad (3.42)$$

$$\partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta n(\mathbf{q} + \mathbf{G}) = -\frac{2}{N_p \Omega} \sum_{v\mathbf{k}} \langle \psi_{v\mathbf{k}}^0 | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \Delta_{\kappa\alpha}^{\mathbf{q}} \psi_{v\mathbf{k}} \rangle, \quad (3.43)$$

where N_p is the total number of unit cells in the superlattice.

The derivative of the perturbed state $\Delta_{\kappa\alpha}^{\mathbf{q}} \psi_{v\mathbf{k}}$ is obtained by solving the Sternheimer equation in reciprocal space:

$$(\hat{H}_{\text{KS}}^{\mathbf{k}+\mathbf{q}} - \epsilon_{v\mathbf{k}}^0) |\Delta_{\kappa\alpha}^{\mathbf{q}} \psi_{v\mathbf{k}}\rangle = (P_v^{\mathbf{k}+\mathbf{q}} - 1) \partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta V_{\text{KS}} |\psi_{v\mathbf{k}}^0\rangle, \quad (3.44)$$

where the coordinate representation of the operator $\hat{H}_{\text{KS}}^{\mathbf{k}+\mathbf{q}}$ is defined in terms of the

self-consistent field Hamiltonian $\hat{H}_{\text{KS}}^0$ by the relation [135]:

$$\langle \mathbf{r} | \hat{H}_{\text{KS}}^{\mathbf{k}+\mathbf{q}} | \mathbf{r}' \rangle = e^{-i(\mathbf{k}+\mathbf{q}) \cdot \mathbf{r}} \langle \mathbf{r} | \hat{H}_{\text{KS}}^0 | \mathbf{r}' \rangle e^{i(\mathbf{k}+\mathbf{q}) \cdot \mathbf{r}'}, \quad (3.45)$$

where $\varepsilon_{v\mathbf{k}}^0$ is the unperturbed occupied bands, and $\hat{P}_v^{\mathbf{k}+\mathbf{q}} = \sum_v |\psi_{v\mathbf{k}+\mathbf{q}}^0\rangle \langle \psi_{v\mathbf{k}+\mathbf{q}}^0|$ is the projection operator for the occupied states of wave vector $\mathbf{k} + \mathbf{q}$ [135].

Equations (3.43), (3.44) together with (3.40) should be solved self-consistently for a fixed $\mathbf{q}$ -point to determine $\partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta n(\mathbf{q} + \mathbf{G})$ and similarly for $\partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta V_{\text{KS}}$. Finally, the electronic contribution to the dynamical matrix can be shown as [137]:

$$D_{\kappa\alpha, \kappa'\alpha'}^{\text{ele}}(\mathbf{q}) = \Omega \sum_{\mathbf{G}} \left[\partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta n(\mathbf{q} + \mathbf{G}) \partial_{\tau_{\kappa'\mathbf{q}\alpha'}}^* V_n(\mathbf{q} + \mathbf{G}) + n(\mathbf{G}) \partial_{\tau_{\kappa\mathbf{q}\alpha}} \partial_{\tau_{\kappa'\mathbf{q}\alpha'}}^* V_n(\mathbf{G}) \right]. \quad (3.46)$$

This expression is very similar to the electronic contribution that appears in the first two terms of Eq. (3.37) but more useful in practice. Since the perturbed quantities formulated in the Fourier space are uncoupled between the wavevectors $\mathbf{k}$ and $\mathbf{q}$, it carries some practical advantages in the calculations. For example, these calculations can be parallelised on a grid of $\mathbf{q}$ -points and symmetry reduction can be performed directly for the equivalent $\mathbf{q}$ -points. Furthermore, the dynamical matrices for a grid of $\mathbf{q}$ -points can be interpolated using discrete Fourier transforms.

3.2.3 Longitudinal optical phonons

As we mentioned for the Fröhlich model in Sec. 2.2, the formation of large polarons happens when an electron is coupled to a polarisation field induced by itself. Such description using the macroscopic polarisation and the electric field is associated with the so-called longitudinal optical (LO) phonons in the long-wavelength limit, in which the long-range nature of the Coulomb interactions plays a key role. In polar crystals, the electric field generated by the displacement of an atom is nonperiodic;

this corresponds to a singularity in reciprocal space when $\mathbf{q} \rightarrow 0$. DFPT has been refined to explicitly treat the long-wavelength LO phonons by adapting the theory of the electric polarisation [138]. In the long-wavelength limit, the PES depending on the atomic displacements $\Delta\boldsymbol{\tau}_{\kappa\mathbf{q}\rightarrow\mathbf{0}}$ and the macroscopic electric field $\mathbf{E}$ maybe expressed as [139]:

$$U = \frac{1}{2} \sum_{\kappa\alpha, \kappa'\alpha'} \tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{an} \Delta\tau_{\kappa\alpha} \Delta\tau_{\kappa'\alpha'} - \frac{\epsilon_0\Omega}{2} \sum_{\alpha\beta} E_\alpha \epsilon_{\alpha\beta}^\infty E_\beta - e \sum_{\kappa\alpha\beta} E_\alpha Z_{\kappa, \alpha\beta}^* \Delta\tau_{\kappa\beta}, \quad (3.47)$$

where the first term describes the usual analytical part of the dynamical matrix, the last two terms come from the electrostatic energy associated with the polarisation and the macroscopic field. $Z_{\kappa, \alpha\beta}^*$ is known as the Born effective charge tensor for the κ -th atom, and $\tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{an} = \tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{an}(\mathbf{q} = 0)$ denotes the long-wavelength dynamical matrix of the system in the absence of the macroscopic electric field. As we know from electrostatics, the macroscopic electric field created by the induced polarisation from the long-wavelength LO phonons exerts forces on the atoms. Therefore, we expect that the LO phonon frequencies will be enhanced. This effect is called the LO-TO splitting, and is unique signature of polar materials. If we minimise Eq. (3.47) with respect to the macroscopic field at a fixed displacement, we obtain a relation for the displacement field $\mathbf{D}$ that depends on both the macroscopic field and the polarisation density:

$$\mathbf{D} = -\frac{1}{\Omega} \frac{\partial U}{\partial \mathbf{E}} = \epsilon_0 \underline{\underline{\epsilon}}^\infty \cdot \mathbf{E} + \frac{e}{\Omega} \sum_{\kappa} \underline{\underline{Z}}_{\kappa}^* \cdot \Delta\boldsymbol{\tau}_{\kappa}, \quad (3.48)$$

where $\underline{\underline{\epsilon}}^\infty$ is the high-frequency dielectric tensor, and $\underline{\underline{Z}}_{\kappa}^*$ is the Born effective charge tensor for the κ -th atom. In the absence of free charges, the displacement field satisfies Maxwell's equation $\nabla \cdot \mathbf{D}(\mathbf{r}) = 0$. If we expand Eq. (3.48) as a Fourier series, it can be shown that:

$$\sum_{\mathbf{q}} \mathbf{q} \cdot \left[\epsilon_0 \underline{\underline{\epsilon}}^\infty \cdot \mathbf{E}_{\mathbf{q}\rightarrow\mathbf{0}} + \frac{e}{\Omega} \sum_{\kappa} \underline{\underline{Z}}_{\kappa}^* \cdot \Delta\boldsymbol{\tau}_{\kappa\mathbf{q}\rightarrow\mathbf{0}} \right] e^{i\mathbf{q}\cdot\mathbf{r}} = 0. \quad (3.49)$$

Since the Born effective charge and the high-frequency dielectric tensor are both block diagonal, the displacement and the macroscopic field induced by the LO phonon are thus in the same direction $\hat{\mathbf{q}}$, where $\hat{\mathbf{q}}$ is the unit vector of the phonon wavevector.

If we express the PES in terms of displacements only, i.e. replacing the macroscopic field by the displacement using Eq. (3.49), we can show that the PES can be written via an effective dynamical matrix which contains the usual analytical part plus an additional non-analytical contribution:

$$U = \frac{1}{2} \sum_{\kappa\alpha, \kappa'\alpha'} \left[\tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{an}(\mathbf{q} = 0) + \tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{na}(\mathbf{q} \rightarrow 0) \right] \Delta\tau_{\kappa\alpha} \Delta\tau_{\kappa'\alpha'}. \quad (3.50)$$

Here the non-analytical contribution $\tilde{D}^{na}$ is given by:

$$\tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{na}(\mathbf{q} \rightarrow 0) = \frac{e^2}{\epsilon_0 \Omega} \frac{[\mathbf{q} \cdot \underline{\underline{Z}}_{\kappa}^*]_{\alpha} [\mathbf{q} \cdot \underline{\underline{Z}}_{\kappa'}^*]_{\alpha'}}{\mathbf{q} \cdot \underline{\underline{\epsilon}}^{\infty} \cdot \mathbf{q}}. \quad (3.51)$$

The non-analytical contribution is responsible for the LO-TO splitting, and the enhancement of phonon frequency in the LO phonon can be obtained as:

$$\omega_{\nu}^2(\mathbf{q} \rightarrow 0) = \omega_{\nu}^2(\mathbf{q} \rightarrow 0) + \sum_{\kappa\alpha, \kappa'\alpha'} e_{\kappa\alpha, \nu}^*(\mathbf{q}) \frac{\tilde{D}_{\kappa\alpha, \kappa'\alpha'}^{na}(\mathbf{q} \rightarrow 0)}{\sqrt{M_{\kappa} M_{\kappa'}}} e_{\kappa'\alpha', \nu}^*(\mathbf{q}). \quad (3.52)$$

Although we have seen how the nonanalytical contribution $\tilde{D}^{na}$ depends on the phenomenological parameters $\underline{\underline{Z}}_{\kappa}^*$ and $\underline{\underline{\epsilon}}^{\infty}$, we have not yet discussed how they are obtained from *ab initio* calculations. To this aim it is useful to obtain the polarisation $\mathbf{P}$ via the identity (2.20):

$$\mathbf{P} = \epsilon_0(\underline{\underline{\epsilon}}^{\infty} - \underline{\underline{I}})\mathbf{E} + \frac{e}{\Omega} \sum_{\kappa} \underline{\underline{Z}}_{\kappa}^* \cdot \Delta\tau_{\kappa}. \quad (3.53)$$

This relation enables us to define the Born effective charge as:

$$Z_{\kappa, \alpha\beta}^* = \frac{\Omega}{e} \left(\frac{\partial P_{\alpha}}{\partial \Delta\tau_{\kappa\beta} \mathbf{q}=\mathbf{0}} \right)_{\mathbf{E}=\mathbf{0}}, \quad (3.54)$$

and the high-frequency dielectric tensor can be calculated by the derivative of the polarisation w.r.t. the macroscopic field [139]:

$$\epsilon_{\alpha\beta}^{\infty} = \delta_{\alpha\beta} + \frac{1}{\epsilon_0} \left(\frac{\partial P_{\alpha}}{\partial E_{\beta}} \right)_{\Delta\tau_{\kappa\beta\mathbf{q}=\mathbf{0}}=\mathbf{0}}. \quad (3.55)$$

As we see in Eq. (2.19) that the total polarisation is comprised of two contributions, i.e. the displacements of the bare nuclei plus the polarisation of the charge density response to the motions of nuclei. Therefore, the displacement derivative for the polarisation can be calculated by:

$$\frac{\partial P_{\alpha}}{\partial \Delta\tau_{\kappa\beta\mathbf{q}=\mathbf{0}}} = -\frac{e}{N_p\Omega} \int d\mathbf{r} r_{\alpha} \frac{\partial n(\mathbf{r})}{\partial \Delta\tau_{\kappa\beta\mathbf{q}=\mathbf{0}}} + \frac{e}{\Omega} Z_{\kappa} \delta_{\alpha\beta}. \quad (3.56)$$

Similarly, the derivative w.r.t the macroscopic field for the polarisation is given by:

$$\frac{\partial P_{\alpha}}{\partial E_{\beta}} = -\frac{e}{N_p\Omega} \int d\mathbf{r} r_{\alpha} \frac{\partial n(\mathbf{r})}{\partial E_{\beta}}. \quad (3.57)$$

The electronic contribution in the expressions (3.56) and (3.57) are mathematically ill-defined because the dipole moment operator is non-periodic, and thus contradict the BvK boundary conditions. To circumvent this problem, it is useful to consider the following identity for the off-diagonal matrix element of the position operator $\hat{r}_{\alpha}$:

$$\langle \psi_j^0 | \hat{r}_{\alpha} | \psi_i^0 \rangle = \frac{\langle \psi_j^0 | [\hat{H}_{\text{KS}}^0, \hat{r}_{\alpha}] | \psi_i^0 \rangle}{\epsilon_j^0 - \epsilon_i^0}, \quad \text{where } j \neq i. \quad (3.58)$$

Since Eq. (3.58) can be rearranged in the form of first-order perturbation problem:

$$|\bar{\psi}_i\rangle = \sum_{j \neq i} -\frac{\langle \psi_j^0 | [\hat{H}_{\text{KS}}^0, \hat{r}_{\alpha}] | \psi_i^0 \rangle}{\epsilon_i^0 - \epsilon_j^0} |\psi_j^0\rangle, \quad (3.59)$$

where $|\bar{\psi}_i\rangle \equiv \hat{r}_{\alpha} |\psi_i^0\rangle$ can be seen as an effective perturbed state, this enables the

solution to be determined by solving the Sternheimer equation [139]:

$$(\hat{H}_{\text{KS}}^0 - \epsilon_i^0) |\bar{\psi}_i\rangle = \hat{P}_c [\hat{H}_{\text{KS}}, \hat{r}_\alpha] |\psi_i^0\rangle, \quad (3.60)$$

where the commutator is given by:

$$[\hat{H}_{\text{KS}}^0, \hat{r}_\alpha] = -\frac{\hbar^2}{m} \frac{\partial}{\partial r_\alpha}. \quad (3.61)$$

The above commutation relation holds because the KS potential depends on the position operator, i.e. $V_{\text{KS}} = V_{\text{KS}}(\hat{\mathbf{r}})$. By combining Eqs. (3.54) and (3.56), the Born effective charge can be obtained as [139]:

$$Z_{\kappa, \alpha\beta}^* = Z_\kappa \delta_{\alpha\beta} - \frac{2}{N_p} \sum_{i \in \text{occ}} \langle \bar{\psi}_i | \Delta_{\kappa\beta}^{\mathbf{q}=0} \psi_i \rangle. \quad (3.62)$$

In the calculations of the dielectric tensor, however, the crystal has to be exposed to the applied field in order to measure how the polarisation density responds to the field, in which scenarios the electrons will feel additional potentials $V_{\text{ext}}(\mathbf{r}) = -e\mathbf{r} \cdot \mathbf{E}$. In linear response, Eq. (3.40) can be rather adapted to describe the response to the field instead of the perturbation in the bare nuclei potential:

$$\partial_{E_\alpha} V_{\text{KS}}(\mathbf{r}) = \partial_{E_\alpha} V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}(\mathbf{r}, \mathbf{r}') \right] \partial_{E_\alpha} n(\mathbf{r}'). \quad (3.63)$$

By the mean of solving the the Sternheimer equation (3.39), one can similarly determine the response of the KS states to the field $\partial_{E_\alpha} \psi_i$ self-consistently. Hence, the dielectric tensor can be obtained as [139]:

$$\epsilon_{\alpha\beta}^\infty = \delta_{\alpha\beta} - \frac{2e}{N_p \Omega \epsilon_0} \sum_{i \in \text{occ}} \langle \bar{\psi}_i | \partial_{E_\beta} \psi_i \rangle. \quad (3.64)$$

3.3 Electron-phonon interaction

3.3.1 *Ab initio* calculations of electron-phonon interactions

The electron-phonon matrix element

In the previous section, we demonstrated how the lattice dynamics can be studied by a linear response approach on the basis of DFT. Conceptually, this relies on the adiabatic approximation in which the nuclei are clamped in the adiabatic potential energy surface, where the electronic contribution is related to the nuclear motions via the static electronic screening [107]. Until now, we have only seen how the harmonic properties are influenced by the electronic states but not yet how vibrations affect the electrons. It is expected that the ions can also play an important role on the electronic states, via what is known as the electron-phonon coupling. In order to understand the origin of this coupling, let us first consider the bare electron-nuclei interaction in a periodic lattice:

$$V_{en}(\mathbf{r}, \{\boldsymbol{\tau}_{\kappa p}\}) = \sum_{\kappa p, \mathbf{T}} V_{\kappa}(|\mathbf{r} - \boldsymbol{\tau}_{\kappa p} - \mathbf{T}|), \quad (3.65)$$

where $V_{\kappa} = -\frac{e^2}{4\pi\epsilon_0} \frac{Z_{\kappa}}{|\mathbf{r}|}$ is the Coulomb interaction. When the ions are displaced from their equilibrium positions, we can define the electron-phonon (e-ph) matrix element as the scattering probability for an electron from the initial KS state $|n\mathbf{k}\rangle$ to the final state $|m\mathbf{k} + \mathbf{q}\rangle$, induced by the perturbed potential in response to the ionic displacements:

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = \sum_{\kappa\alpha} \sqrt{\frac{M_0}{M_{\kappa}}} l_{\mathbf{q}\nu} e_{\kappa\alpha, \nu}(\mathbf{q}) \langle \psi_{m\mathbf{k}+\mathbf{q}} | \partial_{\tau_{\kappa\mathbf{q}\alpha}} V_{en} | \psi_{n\mathbf{k}} \rangle, \quad (3.66)$$

where $l_{\mathbf{q}\nu} = \sqrt{\frac{\hbar}{2M_0\omega_{\mathbf{q}\nu}}}$ is the zero-point displacement amplitude [107]. This bare electron-phonon matrix element describes only the potential change due to the rigid ionic motions, which may only be valid for a perfect ionic crystal. As we have seen earlier in the DFPT formulation, in general the electron density also responds to the ionic displacements and screens the perturbation of the bare potential. As a result, the *ab initio* electron-phonon matrix element must take into account the screening effect, and thus is given by:

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}} | \Delta_{\mathbf{q}\nu} V_{\text{KS}} | \psi_{n\mathbf{k}} \rangle. \quad (3.67)$$

Here we introduced a new notation $\Delta_{\mathbf{q}\nu} \equiv \sum_{\kappa\alpha} \sqrt{\frac{M_0}{M_\kappa}} l_{\mathbf{q}\nu} e_{\kappa\alpha,\nu}(\mathbf{q}) \partial_{\tau_{\kappa\mathbf{q}\alpha}}$ for convenience. For the completeness of these definitions, we can define the dielectric response function ϵ by using the linear response relation (3.40), in order to describe how much the bare electron-nuclei interaction is screened by the electronic states:

$$\Delta V_{\text{KS}}(\mathbf{r}) = \int \epsilon_{\text{KS}}^{-1}(\mathbf{r}, \mathbf{r}'') \Delta V_{en}(\mathbf{r}'') d\mathbf{r}'', \quad (3.68)$$

where the dielectric response function can be represented by the density susceptibility of the KS system χ_0 and the Hartree-exchange-correlation kernel K_{Hxc} as a symbolic matrix equation $\epsilon = 1 - K_{Hxc}\chi_0$. As seen in Sec. 3.2.2 the perturbed KS potential $\partial_{\tau_{\kappa\mathbf{q}\alpha}} \Delta V_{\text{KS}}$ can be calculated in DFPT by the self-consistent scheme. Since the e-ph matrix element is formed by the overlapping between the perturbed KS potential and the unperturbed KS states, there is no extra computational effort to obtain this quantity once the self-consistency in the Sternheimer equation is achieved.

3.3.2 Fourier-Wannier interpolation

The calculation of electron-phonon effects, such as self energy effects on electrons and phonons, mobilities, polaronic effects, phonon-mediated pairing interaction and superconductivity, requires the numerical evaluation of the Brillouin-zone integrals, which usually involves a dense sampling of the phonon momenta. In practice, direct calculations of e-ph matrix elements using DFPT are still demanding even for simple systems, because the self-consistent Sternheimer equation has to be solved for every $\mathbf{q}$ -point. Therefore, efficient strategies for calculating e-ph matrix elements from first principle are imperative [140]. In the following, we will layout a useful computational method which is based on the Fourier interpolation and maximally localised Wannier functions [141, 142]. This allows e-ph matrix elements to be obtained in an efficient and reliable manner, without performing many brute-force calculations using DFPT.

Fourier interpolation

Before formally introducing the method of Ref. [140], it is worthwhile to discuss the concept of Fourier interpolation. The method is used to determine a function which passes through a given set of discrete data points, and is a given combination of trigonometric functions. It is most suitable for the interpolation of periodic functions. When the discrete points are equally spaced to form a regular lattice in reciprocal space, the function can efficiently be obtained by using the Fast Fourier Transform algorithm. Considering a periodic function in the reciprocal space $f(\mathbf{k}) = f(\mathbf{k} + \mathbf{G})$ that contains $N_p = N_i \times N_j \times N_k$ data points $\{(\mathbf{k}, f(\mathbf{k}))\}$, we can express the function as a discrete Fourier series:

$$f(\mathbf{k}) = \frac{1}{N_p} \sum_p \hat{f}(\mathbf{R}_p) e^{i\mathbf{R}_p \cdot \mathbf{k}}, \quad (3.69)$$

where $\hat{f}(\mathbf{R}_p)$ denotes the expansion coefficients of the series, and $\mathbf{R}_p$ is a lattice vector in real space. The orthogonality of plane waves property enables the series to be inverted as:

$$\hat{f}(\mathbf{R}_p) = \sum_{\mathbf{k}} f(\mathbf{k}) e^{-i\mathbf{R}_p \cdot \mathbf{k}}. \quad (3.70)$$

Once the expansion coefficients in the real space domain are computed, we can apply a backward Fourier transform (3.69) to obtain the function at arbitrary $\mathbf{k}$ in the reciprocal space. The interpolation quality generally depends on the functional form of $f(\mathbf{k})$ and the density of the given data set. If the function $f(\mathbf{k})$ is sufficiently smooth, the required number of data points to maintain good interpolations are small. This is equivalent to say that $\hat{f}(\mathbf{R}_p)$ is localised in the space $\mathbf{R}$. On the other hand, it is difficult to achieve good interpolation without a dense grid if the function contains singularities. This situation was met in an attempt to interpolate the e-ph matrix elements for the Fröhlich coupling in polar materials, in which the matrix element diverges in small $|\mathbf{q}|$ closed to the gamma point [143]. We will come back to this problem and discuss how to tackle it at the end of this section.

The Wannier function

The study of electronic properties of periodic crystals involves obtaining the extended Bloch wavefunctions $\{\psi_{n\mathbf{k}}(\mathbf{r}) | n, \mathbf{k} \in \text{BZ}\}$ by diagonalising the KS Hamiltonian, the periodicity in the effective potential assigns the crystal momentum $\mathbf{k}$ as an additional quantum number to the band index n . These Bloch states exhibit a smooth wave nature in the momentum space, therefore they should be accessible to Fourier interpolation. By reforming a generalised Fourier transform of the crystal momenta in the Bloch wavefunctions one obtains the so-called Wannier functions [141].

The Wannier functions $\{w_{\mathbf{R}}(\mathbf{r}) | \mathbf{R}\}$ are a complete orthonormal basis set used to alternatively represent electronic states in a period system. These states are solid-state counterpart of localised molecular orbitals for crystals [142]. The original form

of the Wannier function for a single isolated band is defined by:

$$w_{\mathbf{R}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \psi_{\mathbf{k}}(\mathbf{r}), \quad (3.71)$$

where N is the number of discrete $\mathbf{k}$ points in the uniform BZ grid, and $\mathbf{R}$ is the lattice vector in real space. When a group of Bloch bands $\{\psi_{n\mathbf{k}}(\mathbf{r}) | n = 1, 2, \dots, J\}$ with J denoting the number of entangled bands is connected among themselves by degeneracies but separated from other lower and higher bands, a more generalised Wannier functions $\{w_{m\mathbf{R}}(\mathbf{r}) | m = 1, 2, \dots, J\}$ that spans the same space as the group of Bloch bands were introduced [144], because Eq. (3.71) does not work in the case of multiple bands. There were very limited applications of Wannier functions in the past because of their nonuniqueness. This is because the Bloch states $|\psi_{n\mathbf{k}}\rangle$ carry an arbitrary phase for the corresponding state, such that $|\psi_{n\mathbf{k}}\rangle \rightarrow e^{i\phi_n(\mathbf{r})} |\psi_{n\mathbf{k}}\rangle$ is still a legitimate KS state. To resolve this arbitrariness problem Marzari *et al.* [141] introduced a method of constructing maximally localized Wannier functions (MLWF), which works by transforming the composited Bloch states into a set of Wannier functions via an arbitrary unitary matrix $U_{mn,\mathbf{k}}$ at every $\mathbf{k}$ as:

$$|w_{m\mathbf{R}}\rangle = \frac{1}{\sqrt{N_p}} \sum_{\mathbf{k}} \left[\sum_n U_{nm,\mathbf{k}} |\psi_{n\mathbf{k}}\rangle \right] e^{-i\mathbf{k}\cdot\mathbf{R}}, \quad (3.72)$$

where the band index n denotes the entangled manifold which contains a set of Bloch bands, the index m is used to denote the band index in the Wannier space. The arbitrary degree of freedom in the Wannier functions is fixed by minimising the spread functional Ω_s which is defined as the sum of the quadratic spreads of the Wannier functions about their centres (the name “maximal localized” comes from the fact that the Wannier functions become maximally localised while the total spreads are

subjected to be minimised):

$$\Omega_s = \sum_m \langle w_{m\mathbf{0}} | (\mathbf{r} - \bar{\mathbf{r}}_m)^2 | w_{m\mathbf{0}} \rangle, \quad (3.73)$$

where $\bar{\mathbf{r}}_m \equiv \langle w_{m\mathbf{0}} | \hat{\mathbf{r}} | w_{m\mathbf{0}} \rangle$ denotes the centre of the Wannier function $w_{m\mathbf{0}}(\mathbf{r})$. In practical implementation of finding MLWFs [145], the spread functional is expressed explicitly in terms of the overlap matrix:

$$M_{mn}^{(\mathbf{k}, \mathbf{b})} = \langle u_{m\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle, \quad (3.74)$$

where $u_{n\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{n\mathbf{k}}$ is the periodic part of the Bloch function. $U_{nm, \mathbf{k}}$ is updated on-the-fly in conjunction with $M_{mn}^{(\mathbf{k}, \mathbf{b})}$ by the iterative minimisation of the spread functional in Eq. (3.73).

By projecting the position ket onto Eq. (3.72), the Wannier function can be expressed as:

$$\langle \mathbf{r} | w_{m\mathbf{R}_p} \rangle = \frac{1}{\sqrt{N_p}} \sum_{n\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_p} U_{nm, \mathbf{k}} \langle \mathbf{r} | \psi_{n\mathbf{k}} \rangle, \quad (3.75)$$

The relation (3.75) is invertible because of the unitarity, which enables the Bloch states to be expressed as a linear superposition of MLWFs:

$$\langle \mathbf{r} | \psi_{n\mathbf{k}} \rangle = \frac{1}{\sqrt{N_p}} \sum_{mp} e^{i\mathbf{k}\cdot\mathbf{R}_p} U_{mn, \mathbf{k}}^\dagger \langle \mathbf{r} | w_{m\mathbf{R}_p} \rangle. \quad (3.76)$$

It can be shown by taking Eq. (3.75) and translate the position vector $\mathbf{r}$ by $\mathbf{r} \rightarrow \mathbf{r} - \mathbf{R}_p$, resulting in

$$w_{m\mathbf{0}}(\mathbf{r} - \mathbf{R}_p) = w_{m\mathbf{R}_p}(\mathbf{r}). \quad (3.77)$$

This relation is useful for reducing the memory usage in practical implementation, as it can be used to generate arbitrary Wannier functions at any site $\mathbf{R}_p$ but using only the MLWF at the primary site $w_{m\mathbf{0}}(\mathbf{r})$.

Interpolations of KS Hamiltonian and dynamical matrices

As a first example of Fourier-Wannier interpolation, we discuss the technique for interpolating the matrix elements formed by the KS Hamiltonian operator. The Hamiltonian matrix elements are given by:

$$H_{mn,\mathbf{k}} = \langle \psi_{m\mathbf{k}} | \hat{H}_{\text{KS}} | \psi_{n\mathbf{k}} \rangle = \delta_{mn} \epsilon_{n\mathbf{k}}. \quad (3.78)$$

By expressing the Bloch states in terms of the MLWF using Eq. (3.76), we can now establish the Fourier-Wannier transform of the Hamiltonian matrix elements:

$$H_{mn,\mathbf{k}} = \sum_{n',m',\mathbf{R}_e} e^{i\mathbf{k}\cdot\mathbf{R}_e} U_{mm',\mathbf{k}} \langle m'\mathbf{0} | \hat{H}_{\text{KS}} | n'\mathbf{R}_e \rangle U_{n'n,\mathbf{k}}^\dagger. \quad (3.79)$$

The matrix elements in the Wannier space can be expressed as a forward transform:

$$H_{mn,\mathbf{R}_e} \equiv \langle m\mathbf{0} | \hat{H}_{\text{KS}} | n\mathbf{R}_e \rangle = \frac{1}{N_e} \sum_{m',n',\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_e} U_{mm',\mathbf{k}}^\dagger H_{m'n',\mathbf{k}} U_{n'n,\mathbf{k}}. \quad (3.80)$$

In practical implementations, the KS matrix elements (3.78) are firstly obtained via DFT on a coarse $\mathbf{k}$ -grid, and then transformed into the Wannier representation $\{\mathbf{R}_e\}$ using Eq. (3.80). If the matrix elements are sufficiently localised in real space, i.e. $H_{mn,\mathbf{R}_e}$ decays fast with $|\mathbf{R}_e|$, then the matrix elements in the Bloch space $H_{mn,\mathbf{k}'}$ can be accurately obtained by the interpolation without using a dense coarse grid. The final step involves bringing back the matrix elements to the Bloch representation by using Eq. (3.79). However, the unitary matrix at arbitrary $\mathbf{k}'$ is still missing. Thanks to the diagonality of the KS matrix elements (3.78), we can obtain the interpolated matrix element from a standard diagonalisation, i.e. $\underline{\tilde{H}} = \underline{V} \cdot \underline{D} \cdot \underline{V}^\dagger$ by equating

Eqs. (3.78) and (3.79) [140]:

$$\delta_{mn}\epsilon_{n\mathbf{k}'} = \sum_{n',m'} U_{mm',\mathbf{k}'} \left[\sum_{\mathbf{R}_e} e^{i\mathbf{k}' \cdot \mathbf{R}_e} H_{m'n',\mathbf{R}_e} \right] U_{n'n,\mathbf{k}'}^\dagger, \quad (3.81)$$

where we defined the interpolated matrix element awaiting for being diagonalised as $\tilde{H}_{m'n',\mathbf{k}'} = \sum_{\mathbf{R}_e} e^{i\mathbf{k}' \cdot \mathbf{R}_e} H_{m'n',\mathbf{R}_e}$ and the unitary matrix obtained from the diagonalisation is given by $V_{mn} = U_{mn}^\dagger$. The relation (3.81) is powerful as it enables both the electronic bands $\epsilon_{n\mathbf{k}'}$ and the unitary matrix $U_{mm',\mathbf{k}'}$ to be interpolated on the same footing. The interpolation of the unitary matrix is a primary step for obtaining other matrix elements based on the method of FW interpolation.

The method layed out in this section is seamlessly applicable to phonons as well. In the case of phonons, the dynamical matrix elements can be interpolated by using the phonon Wannier representation [140]. Given that the mathematical structure for the FW interpolation between the electrons and phonons is similar, we will skip the phonons and proceed to describe how to interpolate e-ph matrix elements using this formalism.

Interpolation of electron-phonon matrix elements

Another remarkable application of the FW interpolation method is to efficiently obtain the e-ph matrix elements, because the brute-force calculation of them using DFPT is computationally very demanding. Consider the definition of e-ph matrix element in Eq. (3.67) where the perturbed screened potential $\Delta_{\mathbf{q}\nu} V_{\text{KS}}$ is sandwiched between the two Bloch states before and after scattering. This particular structure allows the interpolation scheme for the EP interaction to be formulated using MLWFs. If we introduce the FW transform of the Bloch state (3.76) into the definition (3.67) and also write the displacement derivative in real space using Eq. (3.41), then the e-ph matrix element can be expressed in the Wannier representation via the backward

FW transform:

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = \sum_{\mathbf{R}_e, \mathbf{R}_p} e^{i(\mathbf{q} \cdot \mathbf{R}_p + \mathbf{k} \cdot \mathbf{R}_e)} \sum_{m'n', \kappa\alpha} \left(\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} e_{\kappa\alpha, \nu}(\mathbf{q}) U_{mm', \mathbf{k}+\mathbf{q}} \times \langle m' \mathbf{0} | \partial_{\tau_{\kappa \mathbf{R}_p \alpha}} V_{\text{KS}} | n' \mathbf{R}_e \rangle U_{n'n, \mathbf{k}}^\dagger, \quad (3.82)$$

and similarly the forward transform $g_{mn, \kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p) \equiv \langle m' \mathbf{0} | \partial_{\tau_{\kappa \mathbf{R}_p \alpha}} V_{\text{KS}} | n' \mathbf{R}_e \rangle$ is given by:

$$g_{mn, \kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p) = \frac{1}{N_e} \sum_{\mathbf{k}, \mathbf{q}} e^{-i(\mathbf{q} \cdot \mathbf{R}_p + \mathbf{k} \cdot \mathbf{R}_e)} \sum_{m'n', \nu} \left(\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}} \right)^{-\frac{1}{2}} e_{\kappa\alpha, \nu}^*(\mathbf{q}) \times U_{n'n, \mathbf{k}+\mathbf{q}}^\dagger g_{m'n'\nu}(\mathbf{k}, \mathbf{q}) U_{n'n, \mathbf{k}}, \quad (3.83)$$

where we have used the following translational invariance property in deriving the transformations:

$$\langle m \mathbf{R}'_e | \partial_{\tau_{\kappa \mathbf{R}_p \alpha}} V_{\text{KS}} | n \mathbf{R}_e \rangle = \langle m \mathbf{0} | \partial_{\tau_{\kappa (\mathbf{R}_p - \mathbf{R}'_e) \alpha}} V_{\text{KS}} | n (\mathbf{R}_e - \mathbf{R}'_e) \rangle. \quad (3.84)$$

In order to perform the FW interpolation for the EPI, the initial step is to obtain the e-ph matrix elements in Eq. (3.67) on two coarse uniform grids using DFPT, with N_e and N_p denoting the dimensions of the $\mathbf{k}$ -grid and $\mathbf{q}$ -grid respectively. This is followed by computing the matrix elements in the Wannier representation $g_{mn, \kappa\alpha}(\mathbf{R}_e, \mathbf{R}_p)$ using the forward FW transform (3.83), where the required unitary matrices $U_{mn, \mathbf{k}}$ and the phonon eigenvectors $e_{\kappa\alpha, \nu}(\mathbf{q})$ were pre-computed from the prior calculations of MLWFs, and from the diagonalisation of the dynamical matrix. In a spirit similar to all other families of Fourier interpolation schemes, the e-ph matrix elements in the Wannier representation need to be well localised. This can be checked by inspecting the quantity $|g(|\mathbf{R}_e|, \mathbf{0})|$ and $|g(\mathbf{0}, |\mathbf{R}_p|)|$, to make sure that they are negligible outside of their Wigner-Seitz supercells [140]. The final step is to interpolate the matrix

elements at arbitrary $\mathbf{k}$ and $\mathbf{q}$ by performing the backward FW transform using Eq. (3.82). The required unitary matrix $U_{mn,\mathbf{k}}$, phonon eigenvectors $e_{\kappa\alpha,\nu}(\mathbf{q})$ and dispersions $\omega_{\mathbf{q}\nu}$ in this step are calculated on-the-fly by diagonalising the interpolated KS Hamiltonian (3.81) and the dynamical matrix.

Fourier-Wannier Interpolation for Fröhlich couplings

As we have seen earlier in Sec. 3.2.3 the long-wavelength LO phonons in polar materials induce a long-range electric field. This effect yields a discontinuity in the phonon dispersions, and results in the so-called LO-TO splitting. This behaviour is also inherited by the e-ph matrix elements in all the LO modes. The terms “Fröhlich coupling” or “polar coupling” are related to the same long-range electric field.

The form of the Fröhlich interaction was originally derived on the basis that a free electron interacts with a dispersionless LO phonon in a three-dimensional isotropic and homogeneous system, in which the coupling matrix elements could be expressed in a form [see Eq. (2.39)] that depends on simple parameters, such as the isotropic dielectric constants ϵ^∞ , ϵ^0 and the dispersionless phonon frequency $\hbar\omega_{\text{LO}}$. However, such simple expression is not enough to describe real materials, which may have more complicated structure such as entangled bands, dispersive phonons, and generally being inhomogeneous. While in principle DFPT in combination with the FW interpolation method is very general, the method has been limited for several years to the study of non-polar systems since the interpolation method was firstly proposed in 2007 [140]. Indeed the original method did not take into account the singularity nature $|\mathbf{q}|^{-1}$ of the Fröhlich couplings in the electron-phonon matrix elements.

Although the Fröhlich coupling has been used in many model calculations for a long period [70, 146], in the context of *ab initio* calculations this problem received attention only from 2015, with two independent works [108, 143]. Unlike the classic model proposed by Fröhlich, these works started from a microscopic approach, and

successfully recover the Fröhlich coupling using microscopic quantities such as Born effective charge $Z_{\kappa,\alpha\beta}^*$, normal mode displacements $\sqrt{\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}\nu}}}e_{\kappa\alpha,\nu}(\mathbf{q})$ and high-frequency dielectric tensor $\epsilon_{\alpha\beta}^\infty$. The most general form of the *ab initio* Fröhlich matrix element which takes into account the periodicity in the BvK supercell is given by [143]:

$$g_{mn}^L(\mathbf{k}, \mathbf{q}) = i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa} \left(\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \times \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{Z_\kappa^*}} \cdot \mathbf{e}_{\kappa,\nu}(\mathbf{q})}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon^\infty}} \cdot (\mathbf{q} + \mathbf{G})} \langle \psi_{m\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G}) \cdot (\mathbf{r}-\boldsymbol{\tau}_\kappa^0)} | \psi_{n\mathbf{k}} \rangle, \quad (3.85)$$

where $\boldsymbol{\tau}_\kappa^0$ denotes the equilibrium position of the κ -th atom. This expression precisely capture the nonanalytical part of the e-ph matrix element at $|\mathbf{q}| \rightarrow 0$ that arises from the long-range interaction induced by the Born charges of the ions. In contrast to the simple Fröhlich formula, this expression not only enables direct *ab initio* calculations of the Fröhlich e-ph matrix elements without relying on DFPT-FW interpolations, but also takes into account naturally the atomistic details and the dielectric anisotropy.

When using Eq. (3.85) to extract the polar coupling for model calculations, it is safe to impose the continuum approximation so that the scattering factor $\langle \psi_{m\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G}) \cdot (\mathbf{r}-\boldsymbol{\tau}_\kappa^0)} | \psi_{n\mathbf{k}} \rangle$ can be simplified to δ_{mn} . This is because this expression only describes the coupling of the long wavelength LO phonons at small $|\mathbf{q}|$, in this limit only the electrons near the conduction band bottom are responsible for the scatterings. Such electrons within the parabolic bands still behave as plane waves, which makes the continuum approximation to be valid.

The other practical advantage of Eq. (3.85) is to improve the *ab initio* FW interpolations for polar materials. Since the FW interpolation method suffers in the regime $|\mathbf{q}| \rightarrow 0$ that exhibits the polar singularity, it is convenient to partition the e-ph matrix element into the long-range g^L and the short-range g^S components, in

analogy with the phonon case in Eq. (3.50):

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = g_{mn\nu}^L(\mathbf{k}, \mathbf{q}) + g_{mn\nu}^S(\mathbf{k}, \mathbf{q}). \quad (3.86)$$

The choice of partitioning is arbitrary, and the general Fröhlich matrix element (3.85) is best suited for representing the long-range component g^L . The above partitioning is very useful because the e-ph matrix elements can be regularised by removing the g^L that contains the singularity, so that the FW interpolation method can effectively be performed on the smoother short-range part $g - g^L$. The new scheme designed for improving the original FW interpolation method in polar materials proceeds as follows: 1) compute the e-ph matrix element g on the coarse grids using DFPT; 2) regularise the e-ph matrix elements by calculating $g - g^L$ on the coarse $\mathbf{k}, \mathbf{q}$ grids, which leaves the short-range part g^S ; 3) perform the forward FW transform on g^S using Eq. (3.83); 4) interpolate the short-range part $g_{mn\nu}^S(\mathbf{k}, \mathbf{q})$ using the backward FW transform (3.82); 5) recover the complete matrix element on-the-fly by adding back the missing long-range component $g_{mn\nu}^S(\mathbf{k}, \mathbf{q}) + g_{mn\nu}^L(\mathbf{k}, \mathbf{q}) \rightarrow g_{mn\nu}(\mathbf{k}, \mathbf{q})$.

The method of FW interpolation discussed in this section offers an efficient pathway to calculate KS states and energies, phonon eigenvectors and dispersions, and the first-order e-ph matrix elements at arbitrary $\mathbf{k}$ and $\mathbf{q}$ points in reciprocal space. In this way, the computational cost for sampling the e-ph matrix elements with N_p $\mathbf{q}$ -points in the Brillouin zone can significantly be reduced to the cost of the pre-processing DFPT calculations on a small $\mathbf{q}$ -grid with n_p points. The interpolation method has been implemented as the main engine in the EPW code [147, 148], which is an integral part of the QUANTUM ESPRESSO suite [149, 150]. The code also uses the library mode of the WANNIER90 code for calculating MLWFs. The code has obviously been used for studies of phonon-assisted optical absorption [151], spectral properties of plasmons and polarons [40, 110], superconducting properties and charge

density waves [152–154], charge transport and mobilities [155–157]. In the next chapter, we will demonstrate how the computational techniques developed in this chapter can be applied on the *ab initio* studies of polarons but without supercells [45, 158].

Chapter 4

Ab initio theory of polarons

In this chapter, we present a novel theory and a computational method to study polarons in insulators and semiconductors from first principles. After emphasising the practical limitations in modelling polarons using the conventional DFT supercell approach, we develop a polaron functional by considering exclusively an extra electron deposited on the BvK supercell and identifying the self-interaction terms that hinders the formation of polarons. Starting from the variational minimisation to the functional, we derive the general form of polaron equations in real space that correctly describes the self-trapping of an electron by the displaced ions. Accordingly, we define what is called the formation energy of a polaron and the associated eigenvalue. We explain how to transform these equations in a practical form that can directly access the quantities obtained from *ab initio* calculations. We establish a formal connection between the polaron theory and the Landau-Pekar model. Lastly, we propose a DFT polaron functional with SIC in conjunction with the polaron formalism developed earlier in this chapter, which will be extensively used to cross validate the polaron theory.

4.1 Polarons in density-functional theory

In order to develop an *ab initio* theory of polarons, we begin with standard DFT that contains most of the essential physics, and can be used as a starting point. The modification to remove the undesirable self-interaction error in DFT will be presented in Sec. 4.2.

4.1.1 Derivation of the polaron equations

The physics of the LP model has already been incorporated by standard DFT formalism: if we introduce an electron in an otherwise empty conduction band of a semiconductor or insulator, the ions driven by an additional force are displaced to screen the extra charge. This physical picture is well established, and has been elucidated in several investigations of *small* polarons, i.e. polarons with a spatial extension corresponding to one or few atomic orbitals [53, 54, 159].

The main limitation of such direct calculations supercell is that only small polarons can be explored, because intermediate-size and large polarons would require prohibitively time-consuming calculations with supercells containing at least several thousands of atoms. Another limitation is that with direct calculations it is impossible to analyse the individual contributions to the polaron formation, for example which phonon modes are excited and contribute to the self-trapping, and which electrons participate in the polaron wavefunction. Lastly, direct calculations are particularly sensitive to the choice of the DFT exchange-correlation functional because of the self interaction errors, making it very challenging to obtain reliable polaron formation energies. To overcome these difficulties it is desirable to formulate an *ab initio* theory of polarons which does not resort to large supercell calculations, and where the individual contributions to the polaron energies and wavefunctions become easily accessible. In the following we discuss a new framework to address these challenges.

The starting point here is to express the DFT total energy of a semiconducting or insulating crystal, with the fully occupied valence bands and the emptied conduction bands. We consider a Born-von Karman supercell of the crystal, containing N_p unit cells of volume Ω . We write the equations with an all-electron implementation in mind; the transposition to a pseudopotential formalism is default in conjunction with modern DFT implementation of periodic systems. The KS states have wavefunctions

$\psi_{n\mathbf{k}}(\mathbf{r})$ with energies $\varepsilon_{n\mathbf{k}}$, where n is the band index and $\mathbf{k}$ the wavevector. The KS wavefunctions $\psi_{n\mathbf{k}}(\mathbf{r})$ are normalised within the supercell, and we have N_p $\mathbf{k}$ -points on a uniform grid. With this notation the electron density reads $n = n^\uparrow + n^\downarrow$, with $n^\uparrow(\mathbf{r}) = n^\downarrow(\mathbf{r}) = \sum_{v\mathbf{k}} |\psi_{v\mathbf{k}}(\mathbf{r})|^2$ including the spin degeneracy in the ground state and the subscript v running over all occupied states. The DFT total energy of the entire supercell maybe written as:

$$\begin{aligned} \frac{E[\{\psi_{v\mathbf{k}}\}, \{\tau_{\kappa\alpha}\}]}{E_{\text{Ha}}} = & -2 \sum_{v\mathbf{k}} \int d\mathbf{r} \psi_{v\mathbf{k}}^* \frac{a_0^2 \nabla^2}{2} \psi_{v\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{T}} \int d\mathbf{r} d\mathbf{r}' \frac{a_0 n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}' - \mathbf{T}|} \\ & + \frac{E_{xc}[n^\uparrow, n^\downarrow]}{E_{\text{Ha}}} - \sum_{\kappa p \mathbf{T}} \int d\mathbf{r} \frac{a_0 Z_\kappa n(\mathbf{r})}{|\mathbf{r} - \tau_{\kappa p} - \mathbf{T}|} + \frac{1}{2} \sum_{\substack{\kappa p \mathbf{T} \\ \kappa' p'}} \frac{a_0 Z_\kappa Z_{\kappa'}}{|\tau_{\kappa p} - \tau_{\kappa' p'} - \mathbf{T}|}, \quad (4.1) \end{aligned}$$

where a_0 is the Bohr radius, E_{Ha} is the Hartree energy, and all integration domains are restricted to the supercell. In the last term the contribution from $\kappa p = \kappa' p'$ is omitted when $\mathbf{T} = 0$. We now denote $\tau_{\kappa p}^0$ as the atomic positions at equilibrium in the ground state, so that a general ionic coordinate reads $\tau_{\kappa p} = \tau_{\kappa p}^0 + \Delta\tau_{\kappa p}$. Similarly, we denote $\psi_{v\mathbf{k}}^0$ as the wavefunctions obtained with the ions in the equilibrium positions, and n^0 the corresponding density. To second order in the displacements $\Delta\tau_{\kappa p}$, the total energy in Eq. (4.1) is expressed as:

$$E[\{\psi_{v\mathbf{k}}\}, \{\tau_{\kappa p}\}] = E[\{\psi_{v\mathbf{k}}^0\}, \{\tau_{\kappa p}^0\}] + \frac{1}{2} \sum_{\substack{\kappa\alpha p \\ \kappa'\alpha'p'}} C_{\kappa\alpha p, \kappa'\alpha'p'} \Delta\tau_{\kappa\alpha p} \Delta\tau_{\kappa'\alpha'p'} + \mathcal{O}(\Delta\tau^3), \quad (4.2)$$

where $C_{\kappa\alpha p, \kappa'\alpha'p'}$ is the interatomic force constant matrix [107, 135, 160], evaluated at the ground state. Upon adding an extra electron to the ground state, one conduction state is occupied and the system becomes spin polarised. Before proceeding we emphasise that the same reasoning can be generalised to the case of a hole at the top of the valence bands; the formalism is entirely symmetric in this respect. Let us denote the wavefunction of the excess electron ψ , and the corresponding density $\Delta n = |\psi|^2$.

For definiteness we assume this extra electron carries a spin up. A compensating jellium background, $-1/N_p\Omega$, is also added to avoid the Coulomb divergence of the total energy. The total energy from Eq. (4.1) is modified as follows:

$$\begin{aligned} \frac{E[\psi, \{\psi_{v\mathbf{k}}\}, \{\tau_{\kappa p}\}]}{E_{\text{Ha}}} &= -2 \sum_{v\mathbf{k}} \int d\mathbf{r} \psi_{v\mathbf{k}}^* \frac{a_0^2 \nabla^2}{2} \psi_{v\mathbf{k}} - \int d\mathbf{r} \psi^* \frac{a_0^2 \nabla^2}{2} \psi + \frac{E_{xc}[n^\uparrow + \Delta n, n^\downarrow]}{E_{\text{Ha}}} \\ &+ \frac{1}{2} \sum_{\mathbf{T}} \int d\mathbf{r} d\mathbf{r}' \frac{a_0}{|\mathbf{r} - \mathbf{r}' - \mathbf{T}|} [n(\mathbf{r}) + \Delta n(\mathbf{r}) - 1/N_p\Omega][n(\mathbf{r}') + \Delta n(\mathbf{r}') - 1/N_p\Omega] \\ &- \sum_{\kappa p \mathbf{T}} \int d\mathbf{r} \frac{a_0 Z_\kappa [n(\mathbf{r}) + \Delta n(\mathbf{r}) - 1/N_p\Omega]}{|\mathbf{r} - \tau_{\kappa p} - \mathbf{T}|} + \frac{1}{2} \sum_{\kappa p, \kappa' p' \mathbf{T}} \frac{a_0 Z_\kappa Z_{\kappa'}}{|\tau_{\kappa p} - \tau_{\kappa' p'} - \mathbf{T}|}. \quad (4.3) \end{aligned}$$

In order to continue we note the key observation as follows: the introduction of a single electron to a system of many electrons will only modify the electron density gently. Indeed, in the limit of very large polaron, the extra electron density at any point will be in the order of $(N_p\Omega)^{-1} \ll n$; in the limit of very small polaron the density will be in the order of Ω^{-1} in one unit cell, and negligible in the others. Following this argument, we make the approximations that, upon adding one electron, $\Delta n \ll n$ almost everywhere, and as a result the valence wavefunctions $\psi_{v\mathbf{k}}$ remain unaltered. The latter approximation allows us to expand the exchange-correlation energy as follows:

$$\begin{aligned} E_{xc}[n^\uparrow + \Delta n, n^\downarrow] &= E_{xc}[n^\uparrow, n^\downarrow] + \int d\mathbf{r} \frac{\delta E_{xc}}{\delta n^\uparrow} \Delta n(\mathbf{r}) + \int d\mathbf{r} d\mathbf{r}' \frac{1}{2} \frac{\delta^2 E_{xc}}{\delta n^\uparrow \delta n^\uparrow} \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') \\ &+ \mathcal{O}(\Delta n^3). \quad (4.4) \end{aligned}$$

We note that this expansion is only valid when the DFT functional has continuous derivatives. In exact DFT this is not the case due to the discontinuity of the exchange-

correlation potential. [161]. By combining Eqs. (4.1)–(4.4) and recasting we find:

$$\begin{aligned}
E[\psi, \{\psi_{v\mathbf{k}}\}, \{\boldsymbol{\tau}_{\kappa p}\}] &= E[\{\psi_{v\mathbf{k}}^0\}, \{\boldsymbol{\tau}_{\kappa p}^0\}] + \frac{1}{2} \sum_{\substack{\kappa\alpha p \\ \kappa'\alpha'p'}} C_{\kappa\alpha p, \kappa'\alpha'p'}^0 \Delta\tau_{\kappa\alpha p} \Delta\tau_{\kappa'\alpha'p'} \\
&+ E_{\text{Ha}} \int d\mathbf{r} \psi^*(\mathbf{r}) \left[-\frac{a_0^2}{2} \nabla^2 + \sum_{\mathbf{T}} \int d\mathbf{r}' \frac{a_0 n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}' - \mathbf{T}|} - \sum_{\kappa p \mathbf{T}} \frac{Z_{\kappa} a_0}{|\mathbf{r} - \boldsymbol{\tau}_{\kappa p} - \mathbf{T}|} + \frac{1}{E_{\text{Ha}}} \frac{\delta E_{xc}}{\delta n^{\uparrow}} \right] \psi(\mathbf{r}) \\
&+ \frac{1}{2} \left[\int d\mathbf{r} d\mathbf{r}' \frac{\delta^2 E_{xc}}{\delta n^{\uparrow} \delta n^{\uparrow}} \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') + E_{\text{Ha}} \sum_{\mathbf{T}} \int d\mathbf{r} d\mathbf{r}' \frac{[\Delta n(\mathbf{r}) - 1/N_p \Omega][\Delta n(\mathbf{r}') - 1/N_p \Omega]}{|\mathbf{r} - \mathbf{r}' - \mathbf{T}|/a_0} \right] \\
&+ E_{\text{B}} + \mathcal{O}(\Delta\tau^3) + \mathcal{O}(\Delta n^3), \quad (4.5)
\end{aligned}$$

where E_{B} is a constant term arising from the jellium background. The square brackets in the second line of this equation is recognised as the KS Hamiltonian $\hat{H}_{\text{KS}}[n(\mathbf{r}), \{\boldsymbol{\tau}_{\kappa p}\}]$ associated with the occupied manifold in *absence* of the excess electron. In analogy with Eq. (4.2), we can rewrite this term by performing a Taylor expansion around the equilibrium ionic coordinates:

$$\hat{H}_{\text{KS}}[n(\mathbf{r}), \{\boldsymbol{\tau}_{\kappa p}\}] = \hat{H}_{\text{KS}}[n^0(\mathbf{r}), \{\boldsymbol{\tau}_{\kappa p}^0\}] + \sum_{\kappa\alpha p} \frac{\partial V_{\text{KS}}^0}{\partial \tau_{\kappa\alpha p}} \Delta\tau_{\kappa\alpha p} + \mathcal{O}(\Delta\tau^2), \quad (4.6)$$

where we use V_{KS}^0 to indicate the KS self-consistent potential at equilibrium, in the absence of the excess electron. To keep the formalism in a simple form, we truncate the expansion to first order in $\Delta\tau_{\kappa\alpha p}$. This is the lowest order that admits non-trivial localised solutions, that is self-trapped polarons. By expanding to second order in the displacement we would obtain terms relating to the so-called Debye-Waller self-energy [162]. For simplicity we neglect these terms here.

The third line of Eq. (4.5) contains the Hartree and exchange-correlation *self-interaction* of the excess electron. These are spurious contributions that artificially increase the energy needed to form a polaron, and which tend to delocalise the polaron wavefunctions. Therefore, these terms should be removed from the functional for the

time being. In Sec. 4.2 we discuss how the correct procedure to handle these terms is to modify the exchange-correlation functional E_{xc} by including suitable self-interaction corrections.

Now we can combine Eqs. (4.5) and (4.6) to obtain the expression for the DFT functional of a polaron. At this point we omit the third and fourth lines of Eq. (4.5), and we employ the short-hand notation $\hat{H}_{\text{KS}}^0$ for $\hat{H}_{\text{KS}}[n^0(\mathbf{r}), \{\boldsymbol{\tau}_{\kappa p}^0\}]$:

$$E_{\text{p}}[\psi, \{\Delta\tau_{\kappa\alpha p}\}] = E[\{\psi_{v\mathbf{k}}^0\}, \{\boldsymbol{\tau}_{\kappa p}^0\}] + \frac{1}{2} \sum_{\substack{\kappa\alpha p \\ \kappa'\alpha'p'}} C_{\kappa\alpha p, \kappa'\alpha'p'}^0 \Delta\tau_{\kappa\alpha p} \Delta\tau_{\kappa'\alpha'p'} \\ + \int d\mathbf{r} \psi^*(\mathbf{r}) \left[\hat{H}_{\text{KS}}^0 + \sum_{\kappa\alpha p} \frac{\partial V_{\text{KS}}^0}{\partial \tau_{\kappa\alpha p}} \Delta\tau_{\kappa\alpha p} \right] \psi(\mathbf{r}). \quad (4.7)$$

The functional $E_{\text{p}}[\psi, \{\Delta\tau_{\kappa\alpha p}\}]$ defined by this equation constitutes as the DFT counterpart of the Landau-Pekar functional in Eq. (2.9). Also in this case the normalisation constraint on the wavefunction is taken account by introducing the Lagrange multiplier ε . By setting the derivatives with respect to ψ^* and $\Delta\tau_{\kappa\alpha p}$ to zero, we finally arrive the coupled system of equations which describes the self-trapped polaron in a practical consideration:

$$\frac{\delta}{\delta \psi^*} \left[E_{\text{p}} - \varepsilon \left(\int d\mathbf{r} |\psi(\mathbf{r})|^2 - 1 \right) \right] = 0 : \\ \hat{H}_{\text{KS}}^0 \psi(\mathbf{r}) + \sum_{\kappa\alpha p} \frac{\partial V_{\text{KS}}^0}{\partial \tau_{\kappa\alpha p}} \Delta\tau_{\kappa\alpha p} \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}), \quad (4.8)$$

$$\frac{\delta E_{\text{p}}}{\delta \Delta\tau_{\kappa\alpha p}} = 0 : \\ \Delta\tau_{\kappa\alpha p} = - \sum_{\kappa'\alpha'p'} (C^0)^{-1}_{\kappa\alpha p, \kappa'\alpha'p'} \int d\mathbf{r} \frac{\partial V_{\text{KS}}^0}{\partial \tau_{\kappa'\alpha'p'}} |\psi(\mathbf{r})|^2. \quad (4.9)$$

This coupled system of equations defines a self-consistent problem in ψ and $\Delta\tau_{\kappa\alpha p}$. The solution can yield a localised wavefunction and the associated pattern of atomic displacements for a polaron. In order to emphasise the analogy with the Landau-

Pekar polaron discussed in Sec. 2.1, it is convenient to replace Eq. (4.9) inside (4.8).

The result is given by:

$$\hat{H}_{\text{KS}}^0 \psi(\mathbf{r}) - \int d\mathbf{r}' K^0(\mathbf{r}, \mathbf{r}') |\psi(\mathbf{r}')|^2 \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}), \quad (4.10)$$

having defined the ‘polaron kernel’ $K^0(\mathbf{r}, \mathbf{r}')$ as:

$$K^0(\mathbf{r}, \mathbf{r}') = \sum_{\kappa\alpha p} \sum_{\kappa'\alpha'p'} \frac{\partial V_{\text{KS}}^0(\mathbf{r})}{\partial \tau_{\kappa\alpha p}} (C^0)^{-1}_{\kappa\alpha p, \kappa'\alpha'p'} \frac{\partial V_{\text{KS}}^0(\mathbf{r}')}{\partial \tau_{\kappa'\alpha'p'}}. \quad (4.11)$$

In this form the similarity with Eq. (2.7) is manifest: the KS Hamiltonian in Eq. (4.10) is the counterpart of the kinetic energy associated with the band effective mass in the LP model, while the kernel is the counterpart of the self-trapping potential. We will discuss this analogy in more detail in Sec. 4.1.7.

4.1.2 The formation energy of a polaron and the meaning of the polaron eigenvalue

As we have seen in the LP model, the eigenvalue ε appearing in Eq. (4.10) does not correspond to the energy of the polaron. To see this it is convenient to define the polaron *formation energy* ΔE_f as the energy required to trap a conduction band state into a localised polaronic state:

$$\Delta E_f = \min E_{\text{p}}[\psi, \{\Delta \tau_{\kappa\alpha p}\}] - \min E_{\text{p}}[\psi, \{\Delta \tau_{\kappa\alpha p} = 0\}]. \quad (4.12)$$

Here E_{p} is the functional defined by Eq. (4.7). This definition describes the energy gained by the system when a delocalised conduction electron becomes self-trapped in the superlattice, and allows us to separate the energetics of the polaron formation from that of the electron addition into the conduction band of the insulator/semiconductor with the ions in the equilibrium positions. By using Eqs. (4.7), (4.9) and (4.11) in

this expression we find:

$$\Delta E_f = \int d\mathbf{r} \psi^*(\mathbf{r}) \left(\hat{H}_{\text{KS}}^0 - \varepsilon_{\text{CBM}} \right) \psi(\mathbf{r}) - \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' |\psi(\mathbf{r})|^2 K^0(\mathbf{r}, \mathbf{r}') |\psi(\mathbf{r}')|^2, \quad (4.13)$$

where ε_{CBM} is the KS eigenenergy in the conduction band bottom. Similarly, an expression for the Lagrange multiplier ε in Eq. (4.10) is obtained by projecting onto ψ^* :

$$\varepsilon - \varepsilon_{\text{CBM}} = \int d\mathbf{r} \psi^*(\mathbf{r}) \left(\hat{H}_{\text{KS}}^0 - \varepsilon_{\text{CBM}} \right) \psi(\mathbf{r}) - \int d\mathbf{r} d\mathbf{r}' |\psi(\mathbf{r})|^2 K^0(\mathbf{r}, \mathbf{r}') |\psi(\mathbf{r}')|^2. \quad (4.14)$$

By subtracting the last two equations we obtain a relation between the formation energy ΔE_f and the eigenvalue ε :

$$\Delta E_f = \varepsilon - \varepsilon_{\text{CBM}} + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' |\psi(\mathbf{r})|^2 K^0(\mathbf{r}, \mathbf{r}') |\psi(\mathbf{r}')|^2. \quad (4.15)$$

This result indicates that the Lagrange multiplier carries a double counting of the Coulomb energy, which has to be removed in order to obtain the formation energy. This is analogous to the relation between the DFT total energy and the sum of the band eigenvalues [117].

By using Eqs. (4.9) and (4.11) we can rewrite Eq. (4.15) as follows:

$$\Delta E_f = \varepsilon - \varepsilon_{\text{CBM}} + \frac{1}{2} \sum_{\substack{\kappa\alpha p \\ \kappa'\alpha'p'}} C_{\kappa\alpha p, \kappa'\alpha'p'}^0 \Delta\tau_{\kappa\alpha p} \Delta\tau_{\kappa'\alpha'p'}. \quad (4.16)$$

This expression for the formation energy can be interpreted in the context of Franck-Condon principle: the difference $\varepsilon_{\text{CBM}} - \varepsilon$ can be viewed as the energy required for an ultrafast (adiabatic) excitation to promote the electron from the polaron state to a band state at the bottom of the conduction manifold, while the ions are still staying in the distorted polaron state; the sum on the r.h.s. then corresponds to energy released

by the distorted lattice upon relaxation. The same interpretation is often discussed in relation to the LP model [69].

4.1.3 Polaron equations in the basis of Kohn-Sham states and phonon modes

For the purpose of practical *ab initio* calculations, it is convenient to express the equations derived in Sec. 4.1.2 using the reciprocal space formulation. Since the KS states obtained in the ground state DFT calculation form a complete orthonormal basis, we can expand the polaron wavefunction as a linear superposition of these KS states:

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{N_p}} \sum_{n\mathbf{k}} A_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (4.17)$$

where the summation is restricted to the unoccupied (conduction) manifold because we are assuming that the valence band manifold remains unaltered. From the normalisations of the KS states $\psi_{n\mathbf{k}}$ and the polaron wavefunction ψ , it follows:

$$\frac{1}{N_p} \sum_{n\mathbf{k}} |A_{n\mathbf{k}}|^2 = 1. \quad (4.18)$$

Now we substitute Eq. (4.17) inside Eq. (4.10) and project both sides on a KS state. To proceed with the algebra one should keep in mind the standard relations between the electron-phonon matrix elements, the interatomic force constant tensor and the vibrational eigenmodes [107]:

$$g_{mn\nu}(\mathbf{k}, \mathbf{q}) = \sum_{\kappa\alpha p} \left(\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}\nu}} \right)^{1/2} e_{\kappa\alpha, \nu}(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{R}_p} \int d\mathbf{r} \psi_{m\mathbf{k}+\mathbf{q}}^*(\mathbf{r}) \frac{\partial V_{\text{KS}}^0(\mathbf{r})}{\partial \tau_{\kappa\alpha p}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (4.19)$$

$$(C^0)^{-1}_{\kappa\alpha p, \kappa'\alpha' p'} = \frac{1}{N_p} \sum_{\mathbf{q}\nu} \frac{e_{\kappa\alpha, \nu}(\mathbf{q}) e_{\kappa'\alpha', \nu}^*(\mathbf{q})}{\sqrt{M_\kappa M_{\kappa'} \omega_{\mathbf{q}\nu}^2}} e^{i\mathbf{q} \cdot (\mathbf{R}_p - \mathbf{R}_{p'})}. \quad (4.20)$$

By combining Eqs. (4.10)–(4.11) and (4.17)–(4.20) we arrive at the self-consistent eigenvalue problem again but in the reciprocal space:

$$\begin{cases} \frac{2}{N_p} \sum_{\mathbf{q}m\nu} B_{\mathbf{q}\nu} g_{mn\nu}^*(\mathbf{k}, \mathbf{q}) A_{m\mathbf{k}+\mathbf{q}} = (\varepsilon_{n\mathbf{k}} - \varepsilon) A_{n\mathbf{k}}, \\ B_{\mathbf{q}\nu} = \frac{1}{N_p} \sum_{m\mathbf{n}\mathbf{k}} A_{m\mathbf{k}+\mathbf{q}}^* \frac{g_{mn\nu}(\mathbf{k}, \mathbf{q})}{\hbar\omega_{\mathbf{q}\nu}} A_{n\mathbf{k}}. \end{cases} \quad (4.21)$$

$$\begin{cases} B_{\mathbf{q}\nu} = \frac{1}{N_p} \sum_{m\mathbf{n}\mathbf{k}} A_{m\mathbf{k}+\mathbf{q}}^* \frac{g_{mn\nu}(\mathbf{k}, \mathbf{q})}{\hbar\omega_{\mathbf{q}\nu}} A_{n\mathbf{k}}. \end{cases} \quad (4.22)$$

The term that multiplies $A_{m\mathbf{k}+\mathbf{q}}$ on the left-hand side of Eq. (4.21) is Hermitian. This can be verified after noting that from Eq. (4.22) we have $B_{\mathbf{q}\nu}^* = B_{-\mathbf{q}+\mathbf{G},\nu}$, where $\mathbf{G}$ is a reciprocal lattice vector that folds $-\mathbf{q}$ back into the first Brillouin zone (with $\mathbf{G} = 0$). The periodicity of $B_{\mathbf{q}\nu}$ is inherited from choosing the periodic gauge for both the KS states and phonon modes.

$$\begin{cases} \psi_{\mathbf{k}+\mathbf{G}} = \psi_{\mathbf{k}} \\ e_{\mathbf{q}+\mathbf{G}} = e_{\mathbf{q}} \end{cases} \quad (4.23)$$

An other useful property is that if $A_{n\mathbf{k}}$ is a solution vector, then also $A_{n,-\mathbf{k}+\mathbf{G}}^*$ is a solution for the same eigenvalue. This can be shown by taking the complex conjugate of Eq. (4.21) and apply $g_{mn\nu}^*(\mathbf{k}, \mathbf{q}) = g_{mn\nu}(-\mathbf{k} + \mathbf{G}, -\mathbf{q})$ from time-reversal symmetries:

$$\frac{2}{N_p} \sum_{\mathbf{q}m\nu} B_{\mathbf{q}\nu}^* g_{mn\nu}^*(-\mathbf{k} + \mathbf{G}, -\mathbf{q}) A_{m\mathbf{k}+\mathbf{q}}^* = (\varepsilon_{n\mathbf{k}} - \varepsilon) A_{n\mathbf{k}}^*. \quad (4.24)$$

We proceed further by inverting the wavevector $\mathbf{k} \rightarrow -\mathbf{k}$, and then followed by $\mathbf{q} \rightarrow -\mathbf{q}$. Finally, we apply the property $B_{\mathbf{q}\nu}^* = B_{-\mathbf{q}+\mathbf{G},\nu}$ to obtain:

$$\frac{2}{N_p} \sum_{\mathbf{q}m\nu} B_{\mathbf{q}\nu} g_{mn\nu}^*(\mathbf{k}, \mathbf{q}) A_{m,-(\mathbf{k}+\mathbf{q})}^* = (\varepsilon_{n\mathbf{k}} - \varepsilon) A_{n,-\mathbf{k}}^*. \quad (4.25)$$

This implies that, apart from a non-essential phases:

$$A_{n,-\mathbf{k}+\mathbf{G}} = A_{n\mathbf{k}}^*. \quad (4.26)$$

By using this property in the expansion Eq. (4.17) we infer that the polaron wavefunction ψ must be real-valued.

Equations (4.21) and (4.22) represent as the crucial result of this thesis. They allow one to calculate the polaron wavefunction without resorting to supercell calculations, but only relying on standard ingredients of DFT calculations in the primitive unit cell, such as KS states, phonon eigenmodes, and electron-phonon matrix elements [107, 135]. These equations have been discussed in Ref. [45, 158].

4.1.4 Lattice distortion in the polaronic ground state

The polaron eigenvector $A_{n\mathbf{k}}$ obtained from the solution of Eqs. (4.21), (4.22) can be used to determine the atomic displacements in the polaron ground state. To this aim we substitute Eqs. (4.17)–(4.20) and (4.22) inside Eq. (4.9). After few operations we obtain:

$$\Delta\tau_{\kappa\alpha p} = -\frac{2}{N_p} \sum_{\mathbf{q}\nu} B_{\mathbf{q}\nu}^* \left(\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}\nu}} \right)^{1/2} e_{\kappa\alpha,\nu}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{R}_p}. \quad (4.27)$$

It is clear that the quantity $B_{\mathbf{q}\nu}$ has the physical meaning of the amplitude of the excited phonon mode $\mathbf{q}\nu$ which contributes to the ionic displacement $\Delta\tau_{\kappa\alpha p}$. As in the case of the electron wavefunction discussed in the previous section, we can easily verify that the atomic displacements $\Delta\tau_{\kappa\alpha p}$ are real-valued as a result of time-reversal symmetry, $B_{\mathbf{q}\nu}^* = B_{-\mathbf{q}+\mathbf{G},\nu}$. By inverting Eq. (4.27) we also find that $B_{\mathbf{q}\nu}$ satisfies the sum rule:

$$\frac{1}{N_p} \sum_{\mathbf{q}\nu} \frac{|B_{\mathbf{q}\nu}|^2}{\omega_{\mathbf{q}\nu}} = \sum_{\kappa\alpha p} \frac{M_\kappa}{2\hbar} |\Delta\tau_{\kappa\alpha p}|^2, \quad (4.28)$$

where the r.h.s. can be interpreted as a measure of the lattice distortion.

4.1.5 Formation energy in the basis of Kohn-Sham states and phonon modes

As a closed analogue to the Eq. (4.27), the formation energy can also be similarly expressed in terms of the eigenvector $A_{n\mathbf{k}}$. To this aim we combine Eq. (4.13) with Eqs. (4.17)–(4.20) and (4.22). The result is given by:

$$\Delta E_f = \frac{1}{N_p} \sum_{n\mathbf{k}} |A_{n\mathbf{k}}|^2 (\varepsilon_{n\mathbf{k}} - \varepsilon_{\text{CBM}}) - \frac{1}{N_p} \sum_{\mathbf{q}\nu} |B_{\mathbf{q}\nu}|^2 \hbar\omega_{\mathbf{q}\nu}, \quad (4.29)$$

or equivalently, using Eq. (4.15):

$$\Delta E_f = \varepsilon - \varepsilon_{\text{CBM}} + \frac{1}{N_p} \sum_{\mathbf{q}\nu} |B_{\mathbf{q}\nu}|^2 \hbar\omega_{\mathbf{q}\nu}. \quad (4.30)$$

The formation energy in Eq. (4.29) consists of one term associated with the electron part of the polaron, described by $A_{n\mathbf{k}}$, and one term associated with the phonon part, described by $B_{\mathbf{q}\nu}$. By comparing Eqs. (4.30) and (4.16), we infer $|B_{\mathbf{q}\nu}|^2 \hbar\omega_{\mathbf{q}\nu}$ represents the contribution of every vibrational mode to the elastic energy of the polaron. Therefore it is natural to interpret $|B_{\mathbf{q}\nu}|^2$ as the number of excited phonons in each mode participating to the polaron.

This heuristic interpretation can be placed on more rigorous ground by moving from a classical to a quantum-mechanical description of the ionic coordinates, and by performing a Bogoliubov transformation [15]. For now we restrict ourselves to emphasise that in DFT calculations the nuclei are described in the adiabatic and classical approximation, therefore we do not strictly have phonon quanta in our formalism.

By introducing the spectral functions:

$$A^2(E) = \frac{1}{N_p} \sum_{n\mathbf{k}} |A_{n\mathbf{k}}|^2 \delta(E - \varepsilon_{n\mathbf{k}} + \varepsilon_{\text{CBM}}), \quad (4.31)$$

$$B^2(E) = \frac{1}{N_p} \sum_{\mathbf{q}\nu} |B_{\mathbf{q}\nu}|^2 \delta(E - \hbar\omega_{\mathbf{q}\nu}), \quad (4.32)$$

Eq. (4.29) can be expressed as:

$$\Delta E_f = \int_0^\infty A^2(E) E dE - \int_0^\infty B^2(E) E dE. \quad (4.33)$$

From these relations we see that the spectral functions $A^2(E)$ and $B^2(E)$ play a similar role in the polaron problem as the Eliashberg spectral function in the theory of superconductors [163]. In Eliashberg theory of superconductivity, the gap function partially depends on the phonon contribution V_n^{ph} . The phonon contribution can be related to the Eliashberg spectral function $\alpha^2 F(\omega)$ [164], which is defined by [148]:

$$\alpha^2 F(\omega) = \frac{1}{2} \sum_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu} \lambda_{\mathbf{q}\nu} \delta(\omega - \omega_{\mathbf{q}\nu}), \quad (4.34)$$

where the electron-phonon coupling strength $\lambda_{\mathbf{q}\nu}$ associated with a specific phonon mode ν and wavevector $\mathbf{q}$ is related to the electron-phonon matrix element by:

$$\lambda_{\mathbf{q}\nu} = \frac{1}{N(\epsilon_F) \omega_{\mathbf{q}\nu}} \sum_{nm} \sum_{\mathbf{k}} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2 \delta(\epsilon_{n\mathbf{k}} - \epsilon_F) \delta(\epsilon_{m\mathbf{k}+\mathbf{q}} - \epsilon_F), \quad (4.35)$$

and $N(\epsilon_F)$ is the density of state per spin at the Fermi level ϵ_F . Given these similarities, in Sec. 6.2.3 we will demonstrate that these polaron spectral functions can be used to identify the EPI mechanisms leading to the formation of polarons.

4.1.6 Visualisation of the polaron wavefunction

In order to visualise the polaron wavefunction ψ in Eq. (4.17), it is useful to utilise the Wannier function representation. Using the standard notation introduced in Ref. [142], each KS state can be expanded in a basis of maximally-localized Wannier functions as shown in Eq. (3.76). By combining Eq. (4.17) and Eq. (3.76) we obtain:

$$\psi(\mathbf{r}) = \sum_{mp} A_m(\mathbf{R}_p) w_m(\mathbf{r} - \mathbf{R}_p), \quad (4.36)$$

having defined:

$$A_m(\mathbf{R}_p) = \frac{1}{N_p} \sum_{n\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_p} U_{mn\mathbf{k}}^\dagger A_{n\mathbf{k}}. \quad (4.37)$$

Equation (4.36) naturally defines $A_m(\mathbf{R}_p)$ as the envelope function of the polaron, but starting from an *ab initio* perspective. It is interesting to realise that Eq. (4.37) for the electron part of the polaron is entirely analogous to Eq. (4.27) for the phonon part. Equation (4.37) is also useful for practical calculations, especially in combination with the Fourier-Wannier interpolation of the electron-phonon matrix elements, as we discuss in Sec. 6.2.2.

However, we should note that the use of Eq. (4.36) should receive a special care: the KS wavefunctions employed to determine $A_{n\mathbf{k}}$ from Eqs. (4.21) and (4.22) must be the *same* as those employed to construct maximally-localized Wannier functions, i.e. the matrix $U_{mn\mathbf{k}}^\dagger$ required in Eq. (4.37). Failure to do so would result in the introduction of spurious phases and the calculations of incorrect envelope function and polaron energies.

If the Wannier functions are real, and the wavefunctions $\psi_{n\mathbf{k}}$ fulfil time-reversal symmetry, i.e. $\psi_{n,-\mathbf{k}} = \psi_{n\mathbf{k}}^*$, that is not automatically guaranteed in *ab initio* calculations, then it follows that $U_{mn,-\mathbf{k}} = U_{mn\mathbf{k}}^*$. Combined with Eq. (4.37), these properties imply that also the envelope functions $A_m(\mathbf{R}_p)$ will also be real-valued.

By combining Eqs. (4.37) and (4.18), we obtain the normalisation condition for the envelope function:

$$\sum_{mp} |A_m(\mathbf{R}_p)|^2 = 1, \quad (4.38)$$

where we used the property that $U_{mn\mathbf{k}}$ is an unitary matrix.

4.1.7 Link with the Landau-Pekar model

Under suitable limits, the *ab initio* polaron equations Eqs. (4.21)–(4.22) recover exactly to the LP model presented in Sec. 2.1. To this aim we consider a model system with only one conduction band with effective mass m^* , one dispersionless phonon mode with frequency ω_{LO} , and electron-phonon coupling given by the simple Fröhlich interaction. The electron-phonon matrix element $g(q)$ is given by [107, 143, 165]:

$$|g(q)|^2 = \frac{e^2}{4\pi\epsilon_0} \frac{4\pi}{\Omega} \frac{\hbar\omega_{\text{LO}}}{2} \frac{1}{\kappa q^2}. \quad (4.39)$$

This expression is valid only for an isotropic bulk crystal with a single infrared-active phonon. By substituting Eqs. (4.37) and (4.39) inside Eqs. (4.21)–(4.22), after some manipulations we obtain:

$$-\frac{\hbar^2 \nabla^2}{2m^*} A(\mathbf{R}) - \sum_{\mathbf{R}'} \frac{1}{N_p} \left[\sum_{\mathbf{q}} e^{i\mathbf{q} \cdot (\mathbf{R}' - \mathbf{R})} \frac{2}{\hbar\omega} |g(q)|^2 |A(\mathbf{R}')|^2 \right] A(\mathbf{R}) = \varepsilon A(\mathbf{R}), \quad (4.40)$$

where we omitted the subscript p from $\mathbf{R}_p$ for the purpose of notational simplicities, and the gradient is with respect to $\mathbf{R}$.

In the limit of dense Brillouin-zone sampling, i.e. supercell of infinite size, we can replace the summation over $\mathbf{q}$ by an integral using $N_p^{-1} \sum_{\mathbf{q}} = \Omega_{\text{BZ}}^{-1} \int_{\text{BZ}} d\mathbf{q}$. Using this replacement and evaluate the integral, Eq. (4.40) becomes:

$$-\frac{\hbar^2 \nabla^2}{2m^*} A(\mathbf{R}) - \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \sum_{\mathbf{R}'} \frac{|A(\mathbf{R}')|^2}{|\mathbf{R}' - \mathbf{R}|} A(\mathbf{R}) = \varepsilon A(\mathbf{R}). \quad (4.41)$$

We can now replace the summation over the lattice vectors by an integral, but noting $\mathbf{R}$ as a continuous variable and using the substitution $\Omega \sum_{\mathbf{R}} = \int d\mathbf{R}$:

$$-\frac{\hbar^2 \nabla^2}{2m^*} A(\mathbf{R}) - \frac{e^2}{4\pi\epsilon_0} \frac{1}{\kappa} \frac{1}{\Omega} \int d\mathbf{R}' \frac{|A(\mathbf{R}')|^2}{|\mathbf{R}' - \mathbf{R}|} A(\mathbf{R}') = \varepsilon A(\mathbf{R}). \quad (4.42)$$

By comparing this result with Eq. (2.7), we see that the envelope function $\Omega^{-1/2} A(\mathbf{r})$ coincides with the solution $\psi(\mathbf{r})$ of the LP model. Therefore, in the case of single-band and single-phonon isotropic systems with Fröhlich electron-phonon coupling, there exists a direct and unambiguous connection between the LP model and first-principles calculations of polarons.

4.2 Self-interaction corrections to polarons

As we have seen in Sec. 4.1, the third line of Eq. (4.5) contains Hartree and exchange-correlation self-interaction energy of the polaron state. These terms are regarded as an artefact in DFT since the excess electron in a more accurate many-body picture should not interact with itself. The practical consequence of including these terms is that they tend to favour electron delocalisation. For example, if we consider the Hartree energy evaluated by the exponential ansatz used in the Landau-Pekar model (2.11):

$$V_H^{\text{LP}} = \frac{5}{16} \frac{e^2}{4\pi\epsilon_0 r_p}, \quad (4.43)$$

it is immediate to see that the Hartree term always decreases the formation energy of the polaron. As we discuss in Sec. 5.4, we confirmed by direct calculations that we are unable to obtain stable self-trapped polarons in the presence of these spurious self-interactions. This behaviour is also well documented in the literature [46, 159].

In order to remove the polaron self-interaction terms in Eq. (4.5), we introduce a

modified DFT functional with SIC as follows:

$$E^{\text{SIC}}[n_{\uparrow} + \Delta n, n_{\downarrow}] = E[n_{\uparrow} + \Delta n, n_{\downarrow}] - E_{\text{H}}[\Delta n - \Delta n_{\text{B}}] - \frac{1}{2} \left(E_{\text{xc}}[n_{\uparrow} + \Delta n, n_{\downarrow}] - 2E_{\text{xc}}[n_{\uparrow}, n_{\downarrow}] + E_{\text{xc}}[n_{\uparrow} - \Delta n, n_{\downarrow}] \right), \quad (4.44)$$

where $E[n_{\uparrow} + \Delta n, n_{\downarrow}]$ is a standard DFT functional, as in Eq. (4.5). The term E_{H} in this equation indicates the Hartree energy functional, and $\Delta n_{\text{B}} = (N_p \Omega)^{-1}$ is the compensating jellium background. The form of the functional E^{SIC} is designed in such a way that the Hartree self-interaction of the polaron is *exactly* removed, and the exchange-correlation self-interaction is cancelled up to *third order* in the polaron density $\Delta n = |\psi|^2$. In fact, upon Taylor expansions of the exchange-correlation terms in Eq. (4.44) we find:

$$E^{\text{SIC}}[n_{\uparrow} + \Delta n, n_{\downarrow}] = E[n_{\uparrow} + \Delta n, n_{\downarrow}] - \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \sum_{\mathbf{T}} \int d\mathbf{r} d\mathbf{r}' \frac{[\Delta n(\mathbf{r}) - \Delta n_{\text{B}}][\Delta n(\mathbf{r}') - \Delta n_{\text{B}}]}{|\mathbf{r} - \mathbf{r}' - \mathbf{T}|} - \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\delta^2 E_{\text{xc}}}{\delta n_{\uparrow} \delta n_{\uparrow}} \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') + \mathcal{O}(\Delta n^4), \quad (4.45)$$

which are precisely equivalent to the functional in Eq. (4.5) with the third line being removed. The present analysis demonstrates that, not only our starting functional [as defined by the first two lines of Eq. (4.5)] is physically motivated, but also it can be derived from a simple self-interaction-free DFT functional, as given by Eq. (4.44). This is particularly useful for benchmarking our formalism against direct calculations in large supercells.

In order to derive the KS Hamiltonians starting from Eq. (4.44), we evaluate the

functional derivatives with respect to $\psi_{v\mathbf{k}\uparrow}$, $\psi_{v\mathbf{k}\downarrow}$, and ψ . As a reminder we have:

$$\begin{aligned} n^\uparrow &= \sum_{v\mathbf{k}} |\psi_{v\mathbf{k}\uparrow}|^2, \\ n^\downarrow &= \sum_{v\mathbf{k}} |\psi_{v\mathbf{k}\downarrow}|^2, \\ \Delta n &= |\psi|^2. \end{aligned} \tag{4.46}$$

The total density is:

$$n = n^\uparrow + \Delta n + n^\downarrow, \tag{4.47}$$

where $n^\uparrow + \Delta n$ is the spin-up density, and $n^\downarrow$ is the spin-down density. We obtain the following modified KS Hamiltonians for spin-up valence electrons ($\hat{H}_{v\uparrow}^{\text{SIC}}$), spin-down valence electrons ($\hat{H}_{v\downarrow}^{\text{SIC}}$), and the polaron ($\hat{H}_{\text{pol}}^{\text{SIC}}$):

$$\hat{H}_{v\uparrow}^{\text{SIC}} = \hat{H}_{\uparrow}^{\text{KS}}[n_\uparrow + \Delta n, n_\downarrow] + V_{xc}^\uparrow[n_\uparrow, n_\downarrow] - \frac{1}{2}V_{xc}^\uparrow[n_\uparrow + \Delta n, n_\downarrow] - \frac{1}{2}V_{xc}^\uparrow[n_\uparrow - \Delta n, n_\downarrow], \tag{4.48}$$

$$\hat{H}_{v\downarrow}^{\text{SIC}} = \hat{H}_{\downarrow}^{\text{KS}}[n_\uparrow + \Delta n, n_\downarrow] + V_{xc}^\downarrow[n_\uparrow, n_\downarrow] - \frac{1}{2}V_{xc}^\downarrow[n_\uparrow + \Delta n, n_\downarrow] - \frac{1}{2}V_{xc}^\downarrow[n_\uparrow - \Delta n, n_\downarrow], \tag{4.49}$$

$$\hat{H}_{\text{pol}}^{\text{SIC}} = \hat{H}_{\uparrow}^{\text{KS}}[n_\uparrow + \Delta n, n_\downarrow] - V_{\text{H}}[\Delta n - \Delta n_{\text{B}}] - \frac{1}{2}V_{xc}^\uparrow[n_\uparrow + \Delta n, n_\downarrow] + \frac{1}{2}V_{xc}^\uparrow[n_\uparrow - \Delta n, n_\downarrow], \tag{4.50}$$

where the Hartree potential V_{H} and the exchange-correlation potentials $V_{xc}^{\uparrow,\downarrow}$ are defined as:

$$V_{\text{H}}[n](\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{4.51}$$

and

$$V_{xc}^{\uparrow,\downarrow}[n](\mathbf{r}) = \frac{\delta E[n]}{\delta n^{\uparrow,\downarrow}(\mathbf{r})}. \tag{4.52}$$

In order to avoid false minima, which are typically encountered in self-interaction corrected DFT [159, 166], we follow the method of Ref. [159] and choose to perform a constrained total energy minimisation with the constraint $\psi_{v\mathbf{k}\uparrow} = \psi_{v\mathbf{k}\downarrow}$. The ad-

vantage of this choice is that it is fully consistent with the assumptions that we used in Sec. 4.1.1 to derive the polaron equations, i.e. the spin-degenerate ground state with $n^\uparrow = n^\downarrow$. This choice is also favourable for implementation because we only need to store the Hamiltonians (4.50) and the average of (4.48)–(4.49) instead of three, and the minimisation only involves two arguments consisting of the spin up density $n^\uparrow + \Delta n$ and the spin down density $n^\downarrow$.

The functional E^{SIC} in Eq. (4.44) is similar but not identical to the SIC proposed in Ref. [159]. In that work the authors studied the self-trapping of a hole defect in α -quartz, by using the damped Car-Parrinello minimisation of the total energy. Using the present notation, their functional reads:

$$E^{\text{SIC, Ref. [159]}}[n_\uparrow + \Delta n, n_\downarrow] = E[n_\uparrow + \Delta n, n_\downarrow] - E_{\text{H}}[\Delta n] - E_{\text{xc}}[n_\uparrow + \Delta n, n_\downarrow] + E_{\text{xc}}[n_\uparrow, n_\downarrow]. \quad (4.53)$$

By comparing this expression with Eq. (4.44), we see that the Hartree self-interaction is removed in a similar way in both approaches. The difference lies in the exchange-correlation self-interaction: if we expand $E_{\text{xc}}[n_\uparrow + \Delta n, n_\downarrow]$ in Eq. (4.53) using the functional derivative, it is clear that the polaron does not experience the exchange-correlation interaction with the valence electrons. This can lead to artificially large band gaps. By including the SIC to quadratic order in Δn both for the Hartree and for the exchange-correlation contributions via Eq. (4.45), the polaron experiences the usual exchange-correlation interaction with the valence electrons, and band gaps remain unaffected.

The correction provided by Eq. (4.44) is rather easy to implement in DFT schemes which perform a direct minimisation on the energy functional, and requires minimal modifications to existing codes [150]. In the present thesis, we modified the `cp.x` code of the Quantum ESPRESSO distribution by altering the functional of Ref. [159] which was already implemented.

As we show in Sec. 6.4.1, the functional E^{SIC} defined by Eq. (4.44) overcomes the delocalisation problem of DFT, and correctly yields localised polaron wavefunctions in polar materials. Importantly, this method does not require the tuning of Hubbard corrections in DFT+ U [167, 168] or the mixing parameter α in hybrid functional calculations [169, 170], since the self-interaction error is removed from the outset without introducing additional parameters.

Chapter 5

Practical implementation

In this chapter we outline the technical aspects of the implementation of polaron theory in details. We start from describing the computational setup for our calculations on lithium fluoride (LiF), lithium peroxide (Li_2O_2) and α -quartz (SiO_2). After reporting the convergence parameters for these compounds, we explain how to numerically solve the polaron equations in practice. In particular, we introduce an effective scheme to reduce the size of the polaron Hamiltonian, in order to avoid memory overload and achieve faster convergence for calculations involving large polarons. We then explain why the gauge freedom for KS states and vibrational eigenmodes needs to be fixed, and how to obtain polaron energies in the dilute limit. Next, we describe the implementation of the polaron routine in the EPW code, and provide several model calculations and comparisons using model parameters from LiF for the purpose of validation. Lastly, we present a calculation for α -quartz to validate our implementation of the modified self-interaction corrected functional. The use of this new functional will be presented for Li_2O_2 in Chapter 6.

5.1 Density-functional theory calculations

In order to demonstrate the theory developed in Sec. 4.1 we perform DFT calculations based on planewaves and pseudopotentials, as implemented in the QUANTUM ESPRESSO materials simulation suite [149], together with the WANNIER90 [145] and EPW[148] codes. The polaron equations described in Sec. 4.1.3 are implemented in a modified version of the EPW code, and the visualisation of the polaron wave-

functions as described in Sec. 4.1.6 is performed using a modified version of the WANNIER90 code and VESTA for the visualisation [171]. We use the generalised gradient approximation to DFT of Perdew, Burke, and Ernzerhof (PBE) [172], and optimised norm-conserving Vanderbilt (ONCV) pseudopotentials [173], with planewaves kinetic energy cutoffs of 150 Ry, 105 Ry, and 70 Ry for LiF, Li₂O₂, and α -SiO₂, respectively. In the ground-state calculations we sample the Brillouin zone with Γ -centered uniform meshes of size $12 \times 12 \times 12$ and $8 \times 8 \times 8$ for LiF and Li₂O₂ respectively, while direct calculations of α -SiO₂ and Li₂O₂ with supercells are sampled at Γ . Lattice vectors and internal coordinates are optimised using this setup before proceeding to calculate polarons. Equations (4.21) and (4.22) require the evaluation of KS energies, phonon energies, and electron-phonon matrix elements on dense uniform grids. To this aim, we employ the Fourier-Wannier interpolation technique as introduced in Chapter 3.

In addition, we validate our approach against explicit supercell calculations by considering two systems, Al-doped α -SiO₂ and Li₂O₂. We perform Car-Parrinello calculations with self-interaction corrections using the CP code of QUANTUM ESPRESSO [174]. The SIC scheme implemented in CP was originally developed in Ref. [159], and corresponds to the functional in Eq. (4.53). To implement the functional in Eq. (4.45) we adapt the existing code with minor modifications.

5.2 Solution of the polaron equations

In order to solve Eqs. (4.21)–(4.22) we express Eq. (4.21) in an explicit matrix form as follows:

$$\sum_{m\mathbf{k}'} H_{n\mathbf{k},m\mathbf{k}'} A_{m\mathbf{k}'} = \varepsilon A_{n\mathbf{k}}, \quad (5.1)$$

with

$$H_{n\mathbf{k},m\mathbf{k}'} = \delta_{n\mathbf{k},m\mathbf{k}'} \varepsilon_{m\mathbf{k}'} - \frac{2}{N_p} \sum_{\nu} B_{\mathbf{k}-\mathbf{k}',\nu}^* g_{nm\nu}(\mathbf{k}', \mathbf{k} - \mathbf{k}'). \quad (5.2)$$

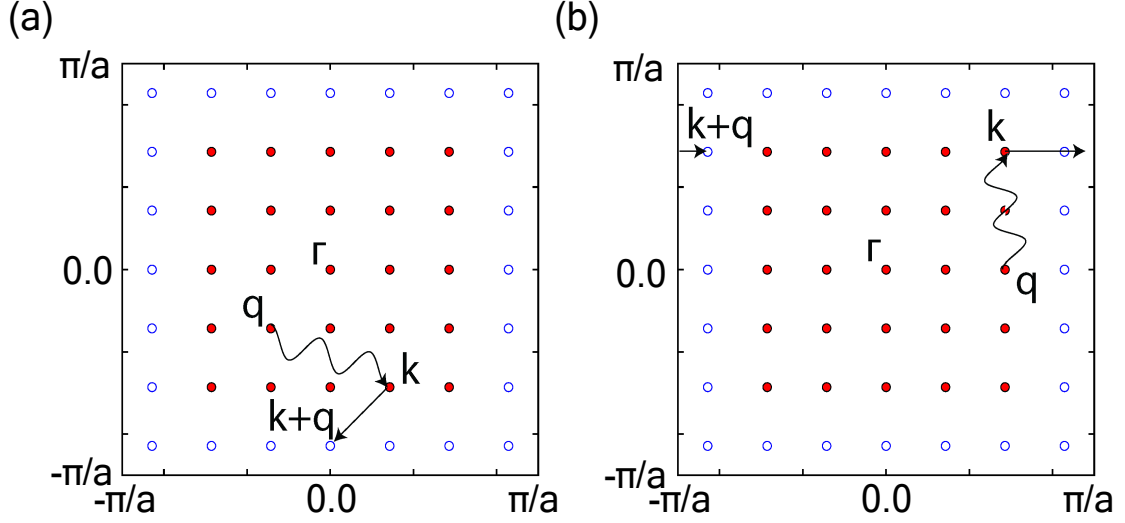


Figure 5.1: A two-dimensional schematic representation of the wavevector basis (uniform grid) in the first Brillouin zone which defines a 7×7 supercell. The red solid circle denotes an inner grid of 5×5 . The initial electronic state is labelled by the momentum $\mathbf{k}$, the incoming phonon is labelled by the momentum $\mathbf{q}$ with wiggly line, and the solid line indicates the scattering of the electron, where the arrow points to the final electronic state with momentum $\mathbf{k} + \mathbf{q}$. The panels (a) illustrates the scattering within the first Brillouin zone; and (b) illustrates the cross zone folding with the periodic Gauge.

In this form it is clear that the solution of Eq. (5.1) can be obtained using standard numerical eigensolvers. To this aim we initialise the vector of coefficients $A_{n\mathbf{k}}$ using a Gaussian lineshape centred at the band minimum. From this initial guess we proceed to construct the vector of coefficients $B_{\mathbf{q}\nu}$ using Eq. (4.22). At this point we can set up the Hamiltonian matrix of Eq. (5.2) and proceed to the solution of the eigenvalue problem in Eq. (5.1). The lowest-energy eigenvector $A_{n\mathbf{k}}$ is used to update the coefficients $B_{\mathbf{q}\nu}$ again in Eq. (4.22) and the whole procedure is repeated until convergence in the polaron formation energy as given by Eq. (4.30). In all calculations the energy convergence threshold of 0.1 meV is used.

The $\mathbf{k}$ -point grid employed in Eq. (5.1) defines the equivalent Born-von Kármán (BvK) supercell hosting the polaron. For example a $\mathbf{k}$ -point grid $7 \times 7 \times 1$ corresponds to calculating the polaron wavefunction, the corresponding atomic displacements, and

the energetics in an equivalent $7 \times 7 \times 1$ supercell. Since we need information on both $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$, we use the same uniform and Γ -centered grid for both the $\mathbf{k}$ -points and $\mathbf{q}$ -points. Such a uniform grid is shown in Fig. 5.1. When $\mathbf{k} + \mathbf{q}$ falls outside of the initial grid, we use the periodic gauge and set $A_{n\mathbf{k}+\mathbf{q}} = A_{n\mathbf{k}+\mathbf{q}+\mathbf{G}}$, with $\mathbf{G}$ a reciprocal lattice vector that folds $\mathbf{k} + \mathbf{q}$ inside the original grid [see Fig. 5.1(b)]. This procedure is necessary to guarantee that the solution vector $A_{n\mathbf{k}}$ fulfils time-reversal symmetry [see discussion after Eq. (4.22)].

In the case of large polarons dominated by the Fröhlich coupling, the electron-phonon matrix elements exhibit a singularity at $\mathbf{q} = 0$ [40]. As a result the solution vectors $A_{n\mathbf{k}}$ tend to have significant weight only in the vicinity of the band extrema. This is the case of the electron polaron in LiF, for example, as discussed in Sec. 6.2.2. In these situations one needs relatively fine $\mathbf{k}$ - and $\mathbf{q}$ -point meshes, but most grid points do not contribute to the calculations; to reduce computational cost we use fine grids but we restrict the Hamiltonian $H_{n\mathbf{k},n'\mathbf{k}'}$ to an inner grid of $\mathbf{k}, \mathbf{k}'$ -points near the band edges. We then increase the size of the inner grid to check for convergence.

Since in the present formalism we study a localised charge distribution in a supercell, the solutions of the eigenvalue problem in Eq. (5.1) contain a spurious interaction energy between the polaron and its periodic images. The same situation is also found in the study of charged defects in periodic supercells. In order to eliminate this spurious energy we employ the standard Makov-Payne correction [175]. To this aim we perform calculations for increasing size of the equivalent BvK supercell, and then extrapolate the formation energy and the polaron eigenvalue using the asymptotic trend L^{-1} , where L is the linear size of the equivalent supercell. For example, in the case of the large electron polaron in LiF, we use $\mathbf{k}$ -point grids up to $33 \times 33 \times 33$. In order to cope with such large grids we use a distributed-memory eigensolver from the ScaLAPACK library.[176]

One last aspect that requires some care is the gauge arbitrariness of the electron-

phonon matrix elements $g_{mn\nu}(\mathbf{k}, \mathbf{q})$ that one obtains from Wannier-Fourier interpolation. The arbitrariness relates to the facts that (i) the unitary rotation $U_{mn\mathbf{k}}^\dagger$ used in Eq. (3.76) to go from the smooth Bloch basis to the basis of KS states is determined from a separate diagonalisation at each $\mathbf{k}$ -point; (ii) the analogous rotation required for the atomic displacements in Eq. (4.19), that is the matrix of vibrational eigenvectors $e_{\kappa\alpha,\nu}(\mathbf{q})$, is also obtained by a separate diagonalisation at each $\mathbf{q}$. These diagonalisations have two drawbacks: 1) they do not satisfy the time-reversal symmetry requirements; 2) they may lead to different results on different architectures, and even on the same architecture but in different runs. This issue is particularly delicate because, in order to save memory, we recompute the matrix elements $g_{mn\nu}(\mathbf{k}, \mathbf{q})$ at each self-consistent iteration. Our benchmarks indicate that this issue can lead to (relatively small) numerical noise in the calculated formation energies, that shows up as small oscillations in plots of ΔE_f vs. L . In order to eliminate these fluctuations we enforce a predetermined choice for the gauge of eigenmodes and wavefunctions, in the same spirit as in Sec. VC of Ref. [140]. First we rotate $U_{mn\mathbf{k}}^\dagger$ and $e_{\kappa\alpha,\nu}(\mathbf{q})$ so that the first nonzero component is real and positive. Then we check for degeneracies in the electron or phonon energies, and we break these degeneracies using a fictitious perturbation. To this aim we set up a Hermitian perturbation $P_{mn\mathbf{k}}$ that spans the Bloch subspace. We fill this matrix by using a sequence of small prime numbers as matrix elements. Then, we diagonalise $P'_{ij\mathbf{k}} = \sum_{mn} U_{im\mathbf{k}} P_{mn\mathbf{k}} U_{nj\mathbf{k}}^\dagger$, where the indices i, j are restricted to the degenerate subspaces. By denoting with $V_{ij\mathbf{k}}$ the unitary matrix that diagonalises $P'_{ij\mathbf{k}}$, we construct $U'_{jn\mathbf{k}} = \sum_p V_{jp\mathbf{k}}^\dagger U_{pn\mathbf{k}}$. Finally we obtain interpolated KS states and energies from $U'_{mn\mathbf{k}}$ instead of $U_{mn\mathbf{k}}$. If the energies are all non-degenerate, then we are done. If there are still degeneracies we repeat the operation by filling the perturbation matrix using the next prime numbers in the sequence. We note that in the subsequent polaron calculation the KS energies remain unaffected by this fictitious perturbation, as from this procedure we

only retain the unitary rotation $V_{mn\mathbf{k}}$; formally this is equivalent to taking the limit of a vanishingly small perturbation. We operate similarly for the vibrational eigenmodes. This procedure guarantees that all KS states and phonon eigenmodes carry a unique gauge across successive iterations in the same calculation, or across different machines. We note that the present procedure is simpler and more efficient than the one used in Ref. [140], since here we only perform operations on very small matrices and we do not calculate explicitly the matrix elements of the fictitious perturbation using planewaves, unlike in Ref. [140]. Finally we enforce time-reversal symmetry by making sure that only half of the $\mathbf{k}$ -points are effectively employed in Eq. (5.1), using a simple mapping.

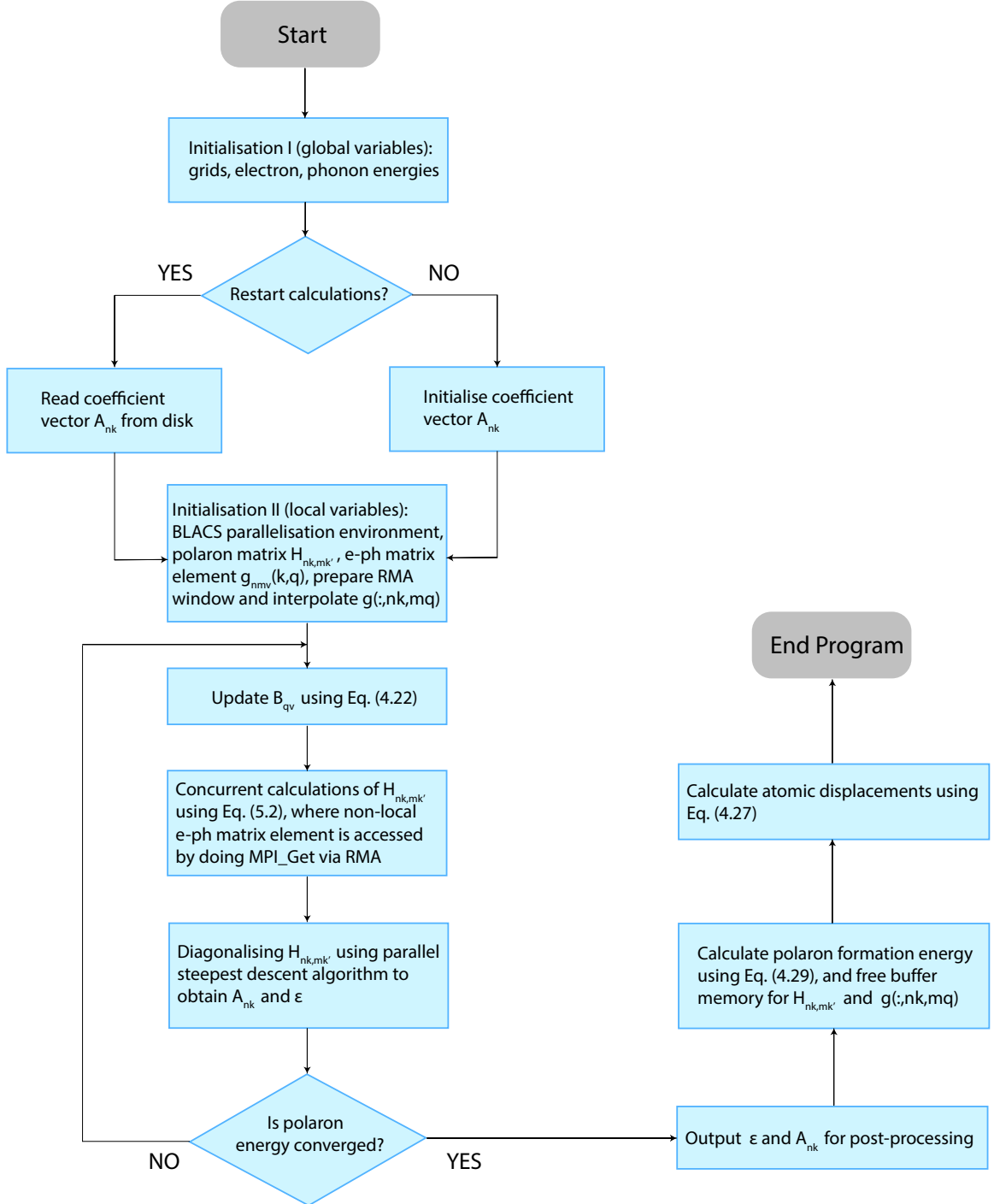


Figure 5.2: The flowchart illustrates the skeleton of our new implemented routine in the EPW code for *ab initio* calculations of polarons.

5.3 The implementation

5.3.1 Code development

In Fig. (5.2) we present a concise summary of our implementation for solving the polaron equations [Eqs. (5.1)–(5.2)]. The procedure was implemented as a Fortran90/MPI-2 code [177], which forms an integral part of the EPW software. Before describing the implementation, it is worth to discuss the practical challenges. If we consider the mathematical structure of the polaron equations, it is formulated as a standard self-consistent matrix eigenvalue problem. This can be relatively simple to deal with when we simulate a polaron in small supercells, in which small grids are sufficient to converge the calculation. However, the size of the Hamiltonian matrix grows quadratically with number of bands and the size of the Brillouin zone grid. Sequential computing is not sufficient to diagonalise very large matrices. In order to tackle this challenge, one might seek a distributed memory computing model which distributes the large matrix into different processors. In this way, each processor stores part of the matrix and performs local calculations synchronously. The final outcome is obtained by gathering all the pieces from different processors. In the following, we explain how to implement this in practice.

Stage I: Initialisations

The code starts from the initialisation of all the global variables which can be accessed by all the processors. This includes the grids, the electron band $\varepsilon_{n\mathbf{k}}$, the phonon dispersion $\omega_{\mathbf{q}\nu}$, the coefficient vectors $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$. These quantities are stored as one-dimensional arrays, running over the grid indices in multiples of band indices, as illustrated in Fig. 5.3(a). If we estimate the memory usage in storing these quantities, for the case of LiF with a $33 \times 33 \times 33$ grid each array will require $\approx 5 \times 10^{-4}$ GB for

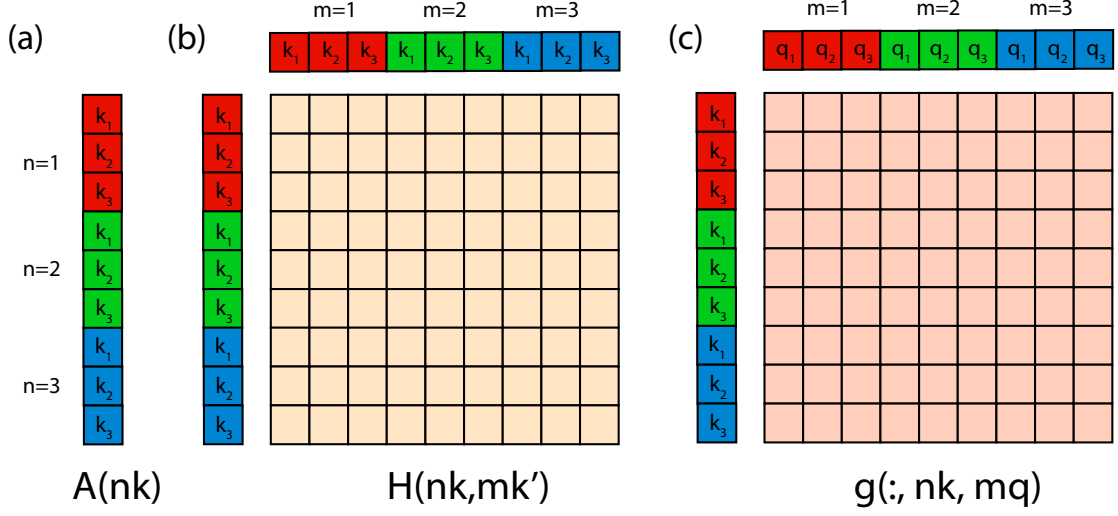


Figure 5.3: The global data structure for (a) the coefficient vector $A_{n\mathbf{k}}$; (b) the Hamiltonian matrix $H_{n\mathbf{k},m\mathbf{k}'}$ and (c) the e-ph matrix element $g_{nm\nu}(\mathbf{k}, \mathbf{q})$.

each band index. This is affordable even if each processor carries the entire part of these arrays, given that the average memory shared by each processor in a typical high performance computing node is ≈ 3 GB/ core. However, we immediately see that storing a two-dimensional array, such as for example the Hamiltonian matrix $H_{n\mathbf{k},m\mathbf{k}'}$ or the e-ph matrix element $g_{nm\nu}(\mathbf{k}, \mathbf{q})$ as illustrated in Fig. 5.3(b)(c), will require ≈ 19.2 GB per core. This set up is clearly not scalable.

Stage II: Block-cyclic distributions and parallelisation algorithms

In order to overcome this limitation, we initialise the parallelisation environment by using the BLACS library [178]. Both the $H_{n\mathbf{k},m\mathbf{k}'}$ and $g_{nm\nu}(\mathbf{k}, \mathbf{q})$ are allocated in 2D block-cyclic distributions [179], where each processor carries different sections of the matrix. This data structure not only resolves the memory scaling issue encountered in the global structure, but also is particularly useful for parallelising linear algebra operations. Fig. 5.4 demonstrates the concept of the block-cyclic distribution and the algorithm used to perform matrix-vector product in parallel, by using a small 5×5

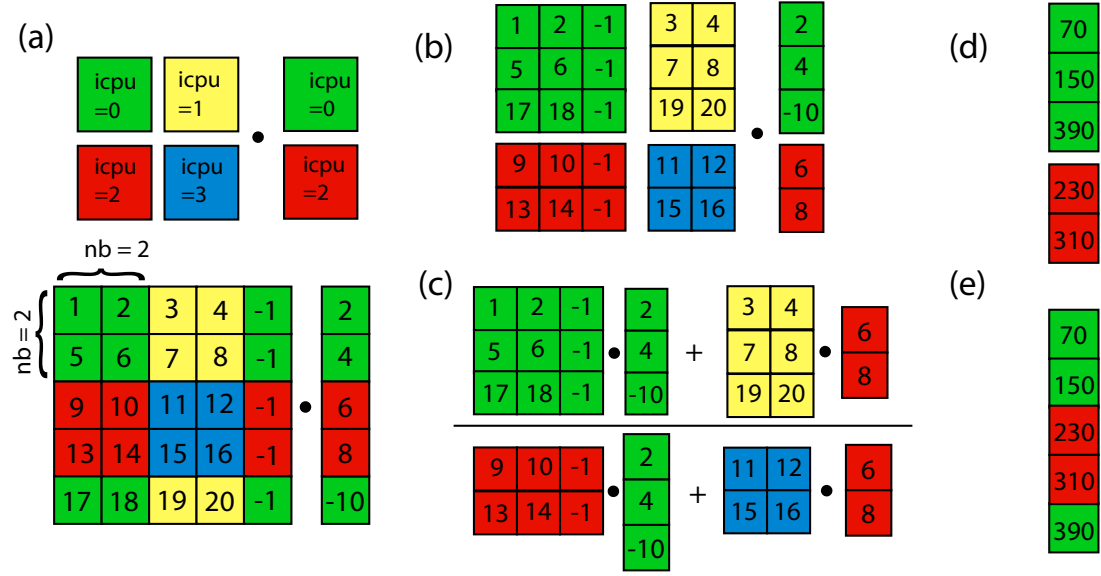


Figure 5.4: An outline of parallel implementation of a matrix-vector product using the block-cyclic distribution.

matrix and a 5×1 vector with four processors as an illustration.

In Fig. 5.4(a) we illustrate a 2×2 processor grid formed by the four processors, each of them is assigned a processor ID (we also label by colour for convenience). The block size $nb = 2$ is used to partition the global matrix into many $nb \times nb$ blocks and the remainder blocks, where each processor stores these blocks only in alternative as a submatrix, i.e. the processor only sees the submatrix labelled by the same colour. In the case of distributing a vector, we see that only processors in the first column of the processor grid are responsible for the subvectors. In Fig. 5.4(b) we show that these submatrices and subvectors can be arranged in the block form, so that the usual multiplication rule can be applied directly on the block matrices and vectors. In Fig. 5.4(c) we demonstrate how the multiplication task is broken down to smaller scales but processing in parallel. This requires some communications between processors. In Fig. 5.4(d)–(e) we show the final outcome in the block form and the global form respectively.

In general, this algorithm is also applicable to other matrix algebra operations,

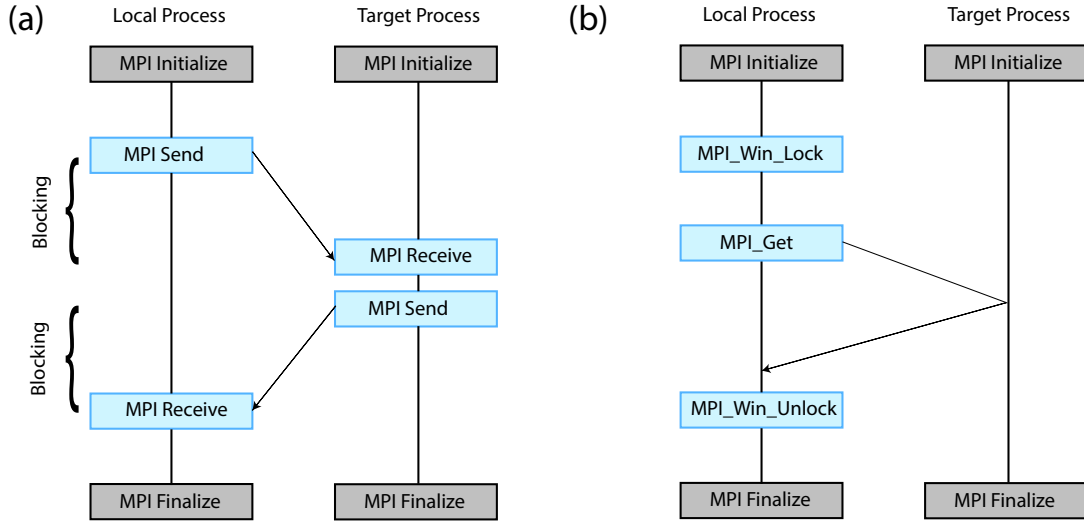


Figure 5.5: Program flow for (a) the MPI two-sided communication model; (b) the MPI one-sided communication model with the passive target synchronisation mode.

such as vector additions, dot products and matrix-matrix multiplications. We note that the step in Fig. 5.4(c) requires some hard coding from scratch to pass the submatrices across different processors via some message-passing techniques, for example the Message Passing Interface (MPI). This step is indeed hidden in our implementation by using the SCALAPACK library [176], which provides explicitly built in routines for parallelising matrix algebra operations. Our code offers two types of eigenvalue solvers: (I) a written parallel routine based on steepest descent, which can only calculate the lowest eigenvalues and eigenvectors; the time complexity scales between quadratically and cubically with the matrix size; (II) a default SCALAPACK routine PZHEEVD based on the Cholesky decomposition which computes the full eigenvalue spectra, but the time complexity scales cubically with the matrix size and requires more memory.

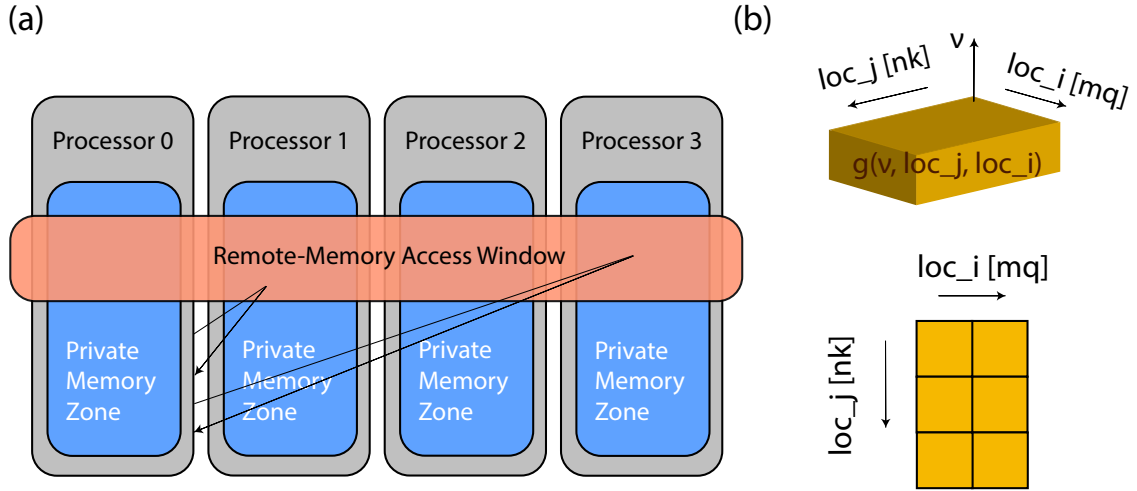


Figure 5.6: Schematic representations of (a) the remote-memory access window in a distributed computing architecture; (b) the local data structure for an e-ph submatrix $g(v, loc_j, loc_i)$, in which the symbol v , loc_j and loc_i indicate the phonon mode, the row and the column indices respectively. We indicate the corresponding global variables inside the square bracket.

Stage III: MPI one-sided communications and RMA

Since the Hamiltonian matrix defined in Eq. 5.2 is never stored globally, the ideal case is that the matrix entries can concurrently be computed locally, without any communications. Unfortunately, the off-diagonal elements in our Hamiltonian matrix depend on the electron-phonon matrix elements at different $\mathbf{q}$ -points, and it is not guaranteed that a processor can directly access this from its local memory. In order to access non-local matrix elements, an obvious choice is to establish the conventional point-to-point (two-sided) communication between the local processor and the target processor as illustrated in Fig. 5.5(a). In this model, whenever a local processor needs a matrix element not found in its own memory, it has to send a request to the target processor, which will need to know which processor is making the request in order to receive the buffer (or otherwise the blocking will last forever and cause a deadlock to the code). In response, the same MPI send-receive operation is performed again to

transfer the data from the target processor to the local processor as requested. We note that implementing such model is relatively demanding, because the send-receive communications need to be completed in pair to avoid deadlock. It turns out that a large startup overhead is also required to prepare for such communications.

Thanks to the development of communication libraries, the MPI-2 and the recent developed MPI-3 standard provide a powerful shared memory model which is known as the MPI one-sided communication [180]. We note that such a shared memory model is indeed tailored for our data structure. In Fig. 5.5(b) we show the one-sided communication model with the passive target synchronisation mode. In contrast to the two-sided communications, the key advantage of the passive one-sided communication is that the target processor no longer participates in the communication. Because of this non-blocking feature, multiple data transfers can process concurrently without blocking the targets. Recent benchmark shows that the one-sided communication could even scale up to ten times faster than the two-sided communications [181]. Given that the one-sided communication model has better performance and is easier to implement, we employ this model to transfer the non-local matrix elements between processors.

To establish the one-sided communication in our implementation, we need to create a remote-memory access (RMA) window as illustrated in Fig. 5.6(a). After allocating the distributed e-ph submatrices $g(nmodes, grows, gcols)$, where $nmodes$ is the total number of phonon modes, $grows$ and $gcols$ are the local dimension of the distributed e-ph submatrix, we use the call `MPI_Win_Create` to generate the RMA window for exposing $g(:, loc_j, loc_i)$ to all the processors. Once this set up is done, we can fill in these distributed e-ph submatrices by performing Fourier-Wannier interpolations. After the startup preparation for the communication is finalised, we can follow the MPI calls as shown in 5.5(b). To access a given non-local e-ph matrix element from a local processor, we need to supply the index of the target processor

that stores the desired matrix element, and its array index in the distributed e-ph submatrices to the call `MPI_Get`. As we see in Fig. 5.6(b) the distributed e-ph submatrices are allocated as 3-D arrays, the cyclic entries correspond to our variables are given by: $i \rightarrow \nu$, $j \rightarrow loc_j$ and $k \rightarrow loc_i$. We can now locate the array index *disp_aint* using the following mapping:

$$disp_aint = (\nu - 1) + nmodes \times (loc_j - 1) + nmodes \times grows \times (loc_i - 1), \quad (5.3)$$

where the first array index is offset to zero as required by the call `MPI_Get`.

Stage IV: Self-consistent cycles and post-processing

The concurrent calculations of the distributed Hamiltonian matrix and the corresponding eigenvalue solvers are embedded in the self-consistent solutions as shown in Fig. 5.2. The cycle begins with an update of the amplitude vector $B_{\mathbf{q}\nu}$, which corresponds to altering the ionic displacements in real space. This is followed by recalculating the Hamiltonian matrix based on the updated $B_{\mathbf{q}\nu}$. Finally, we diagonalise the Hamiltonian matrix to obtain the polaron eigenvalue ε and the coefficient vector $A_{n\mathbf{k}}$ by using the distributed-memory eigenvalue solver. We write both $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$ in plain text files which can be used to restart the calculations. At the end of the self-consistent cycle, we evaluate a condition to exit the cycle, by deciding whether the polaron energy is converged within a threshold.

After exiting the self-consistent cycle, the code writes the converged polaron eigenvalue, formation energy and the coefficient vector to the standard EPW output file. At this stage there is no need to use the Hamiltonian matrix and the distributed e-ph submatrices, we release the memory by deallocating the distributed objects. Finally, the code proceeds to calculate the ionic displacement for each atom and write to the output on the fly. This avoids resorting to distributing memory technique for storing

a large set of displacements. We note that there are several side routines used to support the post-processing calculations, for instances, the polaron spectral functions in Eqs. (4.31)–(4.32) and the polaron wavefunction in Eqs. (4.36)–(4.37). We leave these routine separated from the main polaron engine in Fig. 5.2, because they require more variables to describe specific features; for example, the broadening parameters, the domains of the 1D energy grid, or even the pre-computed Wannier functions by WANNIER90.

5.3.2 Benchmark

Before proceeding to study polarons using this new development, it is important to perform extensive benchmarks. As we seen in Sec. 4.1.7 the *ab initio* polaron equations reduce to the Landau-Pekar equation in the continuum limit; in this limit one assumes that characteristic length scale of the polaron is large so that the discrete atomistic details are unimportant. In this limit we can use a parabolic band structure, a dispersionless phonon and the simple Fröhlich matrix element in the equations. Following this reasoning, we can now test our implementation by comparing our numerical solutions, to the analytical result for the Landau-Pekar model. We employ LiF as the testing platform because it exhibits the strongest Fröhlich coupling among all the halide salts [182], and is expected to yield the smallest possible Fröhlich polaron. This choice is computationally advantages because it allows us to use reasonable small Brillouin-zone grids.

Since the most delicate aspect of the code is the distributed memory self-consistent diagonalisation, the main focus for the validation should be on testing if the converged polaron eigenvalue ε , the coefficients $A_{n\mathbf{k}}$, the polaron wavefunction ψ and the atomic displacements $\Delta\tau_{\kappa p\alpha}$ will reproduce the Landau-Pekar solutions. To this aim we use the ground state solutions of the Landau-Pekar equation Eq. (2.7) obtained varia-

Quantities	Values
a	7.6703 (Bohr)
Ω	112.8177 (Bohr ³)
ϵ^∞	2.0435 (ϵ_0)
ϵ^0	10.6146 (ϵ_0)
m^*	0.8784 (m_e)
Z_{Li}^*	1.0544 (e)
Z_{F}^*	-1.0507 (e)
$\hbar\omega_{\text{LO}}$	77.3020 (meV)
M_{Li}	6.9410 (m_p)
M_{F}	18.9980 (m_p)

Table 5.1: Calculated physical quantities of rocksalt LiF from *ab initio* using functional within the GGA-PBE approximation [172] and the ONCV pseudopotential [173]. We use these parameters to construct the Landau-Pekar model for testing our implementation.

tionally. The polaron wavefunction takes the form [23]:

$$\psi(\mathbf{r}) = 0.1231 r_p^{-1.5} f(|\mathbf{r}|), \quad (5.4)$$

where

$$r_p = \frac{1.51}{m^*} \left[\frac{1}{\epsilon^\infty} - \frac{1}{\epsilon^0} \right]^{-1} \quad (\text{Bohr}), \quad (5.5)$$

$$f(|\mathbf{r}|) = \left[1 + \left(\frac{|\mathbf{r}|}{r_p} \right) + 0.45 \left(\frac{|\mathbf{r}|}{r_p} \right)^2 \right] e^{-\frac{|\mathbf{r}|}{r_p}}, \quad (5.6)$$

and m^* is the isotropic electron effective mass in atomic units. The ground-state polaron eigenvalue reads [23, 30]:

$$\varepsilon = -4.435 \left(\frac{1}{\epsilon^\infty} - \frac{1}{\epsilon^0} \right)^2 m^* \quad (\text{eV}). \quad (5.7)$$

By taking the continuum limit in Eq. (4.9) and combining it with Eqs. (5.4)–(5.5), we obtain an analytical expression for the atomic displacements using the Landau-Pekar

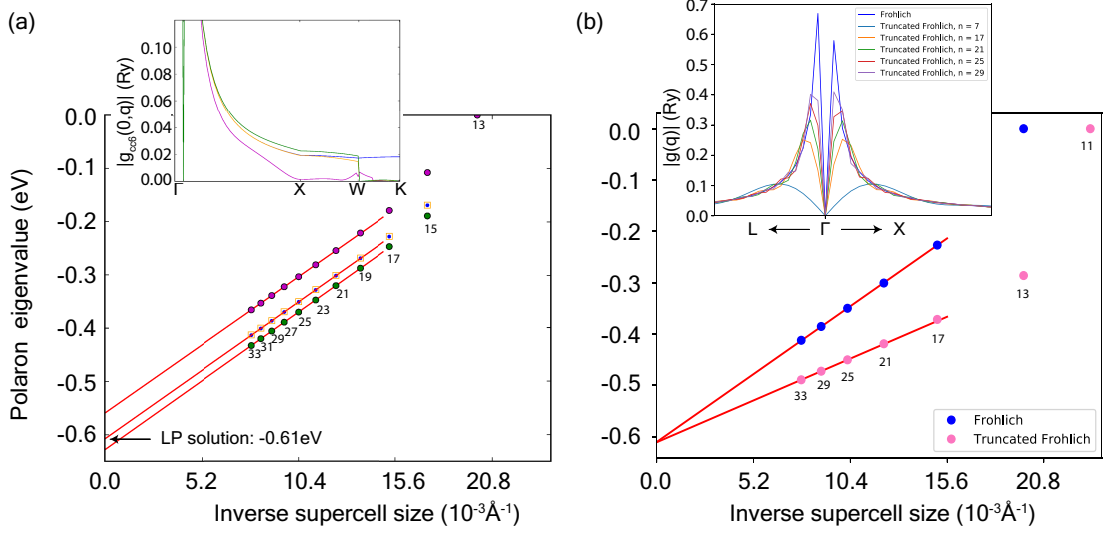


Figure 5.7: (a) Calculated polaron eigenvalues in LiF vs supercell size, where the size is given in $(na)^{-1}$. The numbers next to the data points indicate the supercell size, for example, $N_p^{1/3} \equiv n = 33$ means a supercell of size $33 \times 33 \times 33$. We consider various form of the e-ph matrix elements: (i) isotropic Fröhlich matrix element [Eq. (4.39)] with Einstein phonon ω_{LO} (blue dots); general Fröhlich matrix elements [Eq. (3.85)] (ii) with Einstein phonon ω_{LO} (orange squares); (iii) with dispersive phonons ω_{qLO} (green solid circles); and (iv) *ab initio* matrix elements with dispersive phonons ω_{qLO} (purple solid circles). The red solid lines indicate the Makov-Payne extrapolations. The inset shows the corresponding matrix elements along a $\mathbf{q}$ -path. (b) a comparison between the isotropic Fröhlich matrix element [Eq. (4.39)] and the general Fröhlich matrix element with Coulomb truncation [Eq. (5.10)] and Einstein phonon ω_{LO} .

model:

$$\Delta\tau(\tau_{\kappa p}^0) = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{M_\kappa\omega_{LO}^2} \frac{Z_\kappa^*}{\epsilon^\infty} \cdot \frac{4\pi(0.1231r_p^{-1.5})^2}{|\tau_{\kappa p}^0 - \mathbf{R}_{pol,c}|^3} \left[\int_0^{|\tau_{\kappa p}^0 - \mathbf{R}_{pol,c}|} dr r^2 f(r) \right] (\tau_{\kappa p}^0 - \mathbf{R}_{pol,c}), \quad (5.8)$$

where $\mathbf{R}_{pol,c}$ is the centre of the polaron and Z_κ^* is the isotropic Born-effective charge. Eqs. (5.4)–(5.8) can be evaluated directly using the parameters given in Table 5.1 and the results are presented in Figs. 5.7(a) and 5.8(a)(b) as “LP solution”. In order to compare our numerical solutions to the Landau-Pekar solutions, we consider only a single conduction band, and replace the *ab initio* ingredients in Eq. (5.2) by effective quantities: the parabolic band $\epsilon_{n\mathbf{k}} = \frac{\hbar^2|\mathbf{k}|^2}{2m^*}$, the Einstein phonon $\omega_{\mathbf{q}\nu} = \omega_{LO}$

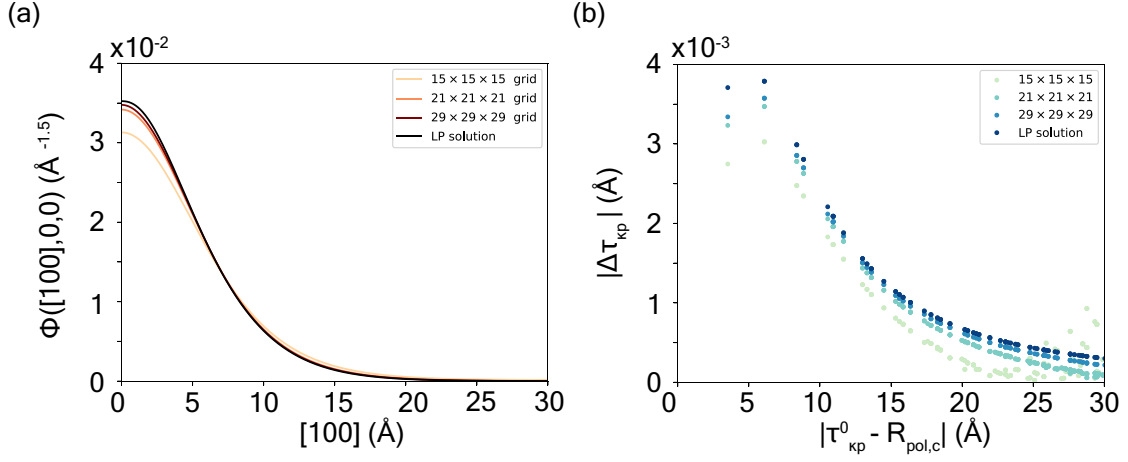


Figure 5.8: (a) The polaron wavefunctions resolved in the the x-axis $[100]$ and (b) The displacement in magnitude for the F atoms vs distance from the polaron centre. The ground state solutions of Landau-Pekar model [Eqs. (5.4) and (5.8)] are used for comparison.

and the isotropic Fröhlich matrix element in Eq. (2.40). The polaron eigenvalue ε , the coefficient vector $A_{\mathbf{k}}$ and the atomic displacements $\Delta \tau_{kp}$ are computed once the self-consistent calculations are converged. For the comparison of the polaron wavefunctions, we transform the coefficient vector $A_{\mathbf{k}}$ to obtain ψ via Eq. (4.17) while keeping only the plane wave factor in the Bloch wavefunctions:

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{N_p \Omega}} \sum_{\mathbf{k}} A_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (5.9)$$

In order to approach the thermodynamic limit of infinite supercell, we repeat the self-consistent calculations with increasingly dense grids starting from 5^3 up to 33^3 . In Fig. 5.7(a), we show the calculated polaron eigenvalues obtained via Eqs. (5.1)–(5.2) with various combinations of e-ph matrix elements and phonons (the same parabolic band with a fixed effective mass is employed). These energies all scale as the inverse of the linear dimension of the supercell, as expected in a charged defect system (see discussion in Sec. 5.2). We verified that both the polaron eigenvalues ε calculated using the isotropic Fröhlich matrix element and the general Fröhlich matrix element

coincides, and that the extrapolated energy is the same as the Landau-Pekar solution obtained by Eq. (5.7). This is well understood from the inset where the two matrix elements are exactly on top of each other, except at the zone edge. When a dispersive phonon is considered, we immediately see that polaron eigenvalue is enhanced by 3 % due to the influence on the Fröhlich matrix element by the phonon dispersion. In contrast to the model calculations, the polaron eigenvalue is lowered by 10 % when we use the *ab initio* matrix element. This is because the analytical expression of Fröhlich only captures correctly the long-range coupling in the small $|\mathbf{q}|$ regime, but it tends to overestimate the coupling at short wavelength, as shown in the inset of Fig. 5.7(a).

On the other hand, we note that the calculated of polaron energies using the isotropic Fröhlich matrix element do not correspond to the finite size Landau-Pekar model with the periodic boundary condition. The reason comes from the fact that the Fröhlich matrix element is derived from a dipole potential in infinite space, unlike in our formalism where we use periodic boundary conditions. To correct this artefact, one needs to introduce the Coulomb truncation into the e-ph matrix element as follows (see Appendix A for derivation):

$$g_{mn\nu}^{L,T}(\mathbf{k}, \mathbf{q}) = i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa} \left(\frac{\hbar}{2M_{\kappa}\omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{Z}}_{\kappa}^* \cdot \underline{\underline{e}}_{\kappa\nu}(\mathbf{q})}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon}}^{\infty} \cdot (\mathbf{q} + \mathbf{G})} \\ \times \left[1 - \text{sinc}(R'_c \sqrt{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon}}^{\infty} \cdot (\mathbf{q} + \mathbf{G})}) \right] \langle \psi_{m\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G}) \cdot (\mathbf{r}-\boldsymbol{\tau}_{\kappa})} | \psi_{n\mathbf{k}} \rangle, \quad (5.10)$$

where the interaction radius is given by:

$$R'_c = \left(\frac{3N_p\Omega}{4\pi} \right)^{\frac{1}{3}} \frac{1}{\sqrt{\epsilon^{\infty}}}. \quad (5.11)$$

In Fig. 5.7(b) we show the comparison between polaron eigenvalues calculated with the isotropic Fröhlich matrix element and the *ab initio* matrix element with Coulomb

truncation. We see that the polaron eigenvalues calculated with the Coulomb truncated matrix element are enhanced significantly for small supercells near the critical size for hosting a localised polaron. The enhancement becomes weaker until intercept at which supercell size approaches to infinite. This is expected because the Coulomb truncated matrix element must recover to the Fröhlich matrix element in the limit of infinite supercell size (i.e. the periodic boundary condition no longer exists). In the inset we illustrate how the Coulomb truncated matrix element depends on the supercell size.

Figures 5.8(a) and 5.8(b) show the calculated polaron wavefunction and displacements in magnitude for the F atoms that corresponds to the numerical calculations of polaron eigenvalues in Fig. 5.7(a)(i). These results are directly compared to the Landau-Pekar solutions. In Fig. 5.8(a) we see that the tail of the calculated polaron wavefunction is already similar to the LP solution for a 15^3 supercell, even though the supercell size is just beyond the critical limit for having a localised polaron. This is very different from the convergence of the polaron eigenvalues, which converge much more slowly. In Fig. 5.8(b) we also see a slow convergence in the atomic displacements for atoms closed to the polaron centre. We note that only few atoms near the polaron centre are significantly influenced by the polaron in comparison to the atoms far away from the polaron, but still the displacements in the latter case asymptotically approach to zero because of the long-range nature of the Fröhlich interaction.

5.4 Validation of the SIC functional

5.4.1 α -Quartz

Since in Sec. 4.2 we introduced a modified version of the SIC for polarons of Ref. [159], it is important to check that our functional yields results in line with previous work [159, 183, 184]. To this aim we repeat previous calculations on Al-doped α -quartz,

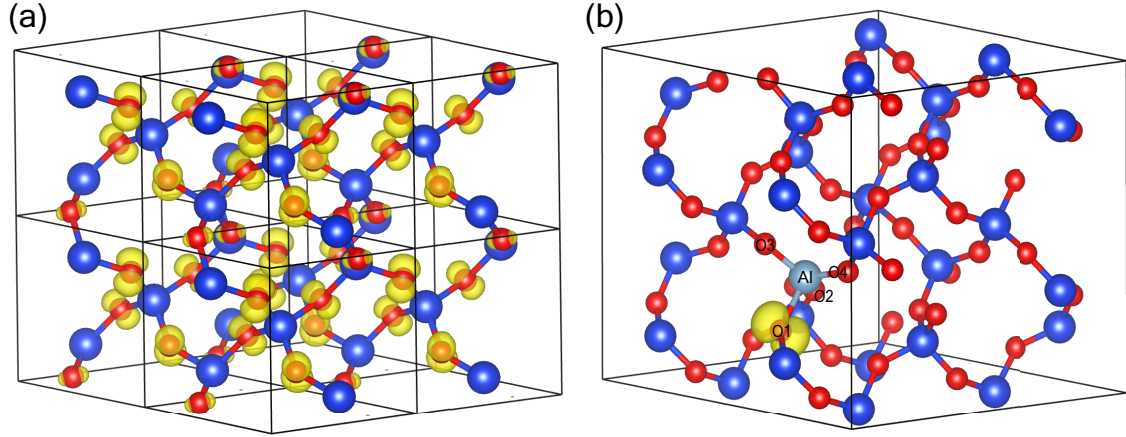


Figure 5.9: Ball and stick models of the compounds considered in this chapter. (a) $2 \times 2 \times 2$ Supercell of α -SiO₂, with Si and O atoms in blue and red, respectively. The isosurface plot represents the delocalised conduction band bottom in the undistorted structure. (b) $2 \times 2 \times 2$ supercell of α -SiO₂ with one Al atom (cyan) replacing Si. In this case the lowest unoccupied state is localised near the defect.

and we compare the localisation of the trapped hole with the existing results [159, 183, 184].

Figure 5.9(a) illustrates the optimised structure of the primitive unit cell of α -SiO₂, in the absence of the Al defect. Our optimised lattice parameters are $a = 4.913$ Å and $c/a = 1.100$, in agreement with the experimental values $a = 4.904$ Å and $c/a = 1.100$ [185]. We model the defect-induced localised hole using a supercell with 72 atoms, with one Si atom replaced by Al. The lattice parameters of the supercell are not re-optimised after this substitution. The defective structure is shown in Fig. 5.9(b).

5.4.2 Benchmark

In order to validate the SIC functional proposed in Eq. (4.44), we consider an Al defect in α -quartz, following previous work [159, 183, 184]. A calculation without SIC yields a delocalised electronic state and no lattice distortion, as shown in Fig. 5.9(a). However, when we include the SIC of Eq. (4.44), we obtain a localised solution, as seen in Fig. 5.9(b). This result is in agreement with previous work based on the

Bond	UHF	US	PW91	pSIC
Al-O(1)	1.924	1.957	1.880	1.946
Al-O(2)	1.688	1.710	1.700	1.696
Al-O(3)	1.703	1.705	1.710	1.708
Al-O(4)	1.689	1.715	1.700	1.699

Table 5.2: Bond lengths (in Å) for the localised geometry around the Al impurity using unrestricted Hartree-Fock [183], density functional theory with US-pSIC scheme [159], PW91-SIC scheme [184] and our polaron SIC functional (4.44)

unrestricted Hartree-Fock method [183] and other SIC schemes [159, 184].

To be more quantitative we also calculate the bond lengths around the defect site. Using the labeling convention set out in Fig. 5.9(b), we show in Table 5.2 our SIC functional yields the bond lengths 1.946/1.696/1.708/1.699 Å for the bonds Al-O(1) to Al-O(4), respectively. These values compare well with previous findings, with r.m.s. deviations of only 0.006 Å. We can conclude that our modified SIC functional yields the same geometry as in previous work. We also confirmed that the isosurface of the hole density [Fig. 5.9(b)] looks similar to what previously reported [159, 183, 184].

To avoid possible ambiguity, we emphasise that the localised hole in Al-doped α -SiO₂ does not constitute a polaron strictly speaking. In fact the localisation and self-trapping are driven by the crystal potential of Al, and do not reflect a spontaneous breaking of translational symmetry as in the cases of Li₂O₂ and LiF discussed below. Accordingly, in this case we do not compare with our linear-response polaron formalism, which addresses spontaneous symmetry breaking in perfect crystals.

Chapter 6

Application I: bulk materials

In this chapter we demonstrate the capability of our developed methodology by studying polarons in bulk materials using LiF and Li₂O₂ as prototypical examples; the study of a 2D layer will be presented in Chapter 7. We first show that our method can describe both large electron polaron and small hole polaron in LiF, and seamlessly captures Fröhlich-type polar electron-phonon coupling and non-Fröhlich coupling to acoustic and optical phonons. For each of the case we report calculations of polaron energies, wavefunctions, atomic displacements and the spectral decompositions. Using these results we calculate the spatial extent of these polarons hosted by LiF, which indicates the type of localisation. To explore further which phonon modes are responsible for the polaron formation, we quantify the fractional contributions of different phonon modes. Lastly, we perform the same analysis on the small electron polaron in Li₂O₂. In particular, we validate our theory using both analytical results and direct calculations on large supercells.

6.1 Test systems

6.1.1 Lithium fluoride

The first test system that we consider is a prototypical ionic insulator, lithium fluoride. LiF crystallises in a simple rock-salt structure and is known to be a wide-gap insulator. As the other members of the alkali halides family, LiF hosts color centers with interesting optoelectronic properties [186]. In particular, the V_K center

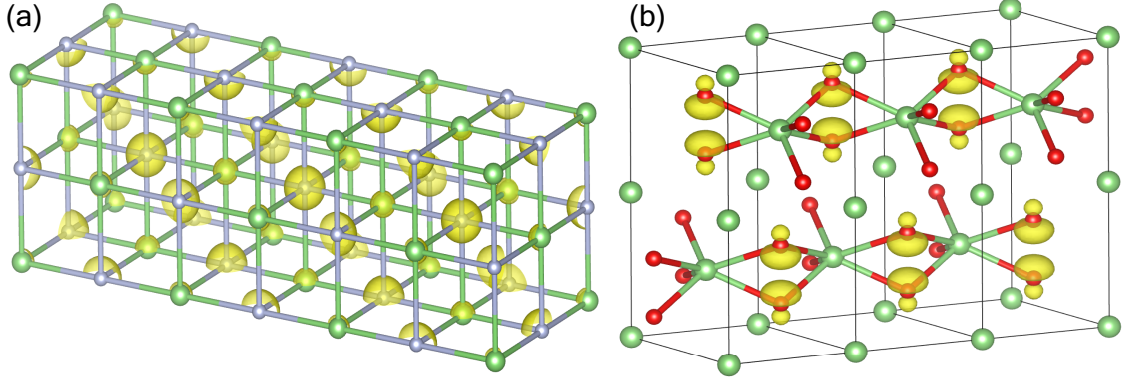


Figure 6.1: Ball and stick models of the compounds considered in this chapter. (a) $3 \times 1 \times 1$ supercell of Li_2O_2 , with Li and O atoms in green and red, respectively. Also in this case we show an isosurface plot of the density at the conduction band bottom, in the undistorted structure. The electron is completely delocalised. (b) $3 \times 1 \times 1$ supercell of LiF , with Li and F atoms in green and silver, respectively. We also show an isosurface plot of the density at the conduction band bottom. In the undistorted structure this state is completely delocalised.

is a self-trapped hole polaron which has been studied in a number of investigations [46, 53, 169, 187–193]. On the other hand, the electron polaron is expected to be a large polaron and has been investigated only by means of model Hamiltonians [194]. Here we perform calculations for both the large electron polaron and the small hole polaron of LiF , and we show that our formalism correctly describes both limits on the same footing.

Fig. 6.1(a) shows a supercell of LiF (the unit cell consists of only two atoms). Our optimised lattice parameter is $a = 4.058 \text{ \AA}$, in agreement with the experimental value $a = 4.02 \text{ \AA}$ [195]. Our calculated KS band gap is $E_g = 8.9 \text{ eV}$, and underestimates the experimental optical gap of 14.2 eV as expected [196]. We find isotropic electron and hole effective masses of $0.88 m_e$ and $3.73 m_e$, respectively. The electron mass is in good agreement with the reported values $0.78\text{--}1.2 m_e$ [182, 197], but we could not find previous values for the hole mass. The calculated relative dielectric constants are $\epsilon^0 = 10.62$ and $\epsilon^\infty = 2.04$, to be compared with the measured values $\epsilon^0 = 9.04$ and $\epsilon^\infty = 1.92$ [198, 199]. The highest computed phonon energy is $\hbar\omega_{\text{max}} = 77.0 \text{ meV}$,

close to the experimental value $\hbar\omega_{\max} = 80$ meV [200], and the Fröhlich coupling constant for the electrons is $\alpha = 4.92$.

6.1.2 Lithium peroxide

The second test system that we consider is lithium peroxide, Li_2O_2 . This compound crystallises in a layered hexagonal structure, with space group $\text{P6}_3/\text{mmc}$. The structure can be thought of as consisting of LiO_2 layers intercalated by Li planes as seen on Fig. 6.1(b). Li_2O_2 forms in battery cathodes during the operation of lithium-air batteries, and can degrade the battery performance through its low electrical conductivity [54, 201, 202]. It has been proposed that the low conductivity of this compound originates from a strong electron-phonon coupling, and several studies reported the calculation of small electron polarons using a supercell approach [54, 167, 203]. In Ref. [54] it was shown that a small electron polaron can form in a $3 \times 3 \times 2$ supercell, without the use of Hubbard corrections or hybrid functionals. This finding suggests that Li_2O_2 supports strongly bound small polarons. Furthermore Li_2O_2 is highly anisotropic. These properties make lithium peroxide an ideal candidate for testing the limits of our approach.

Figure 6.1(b) illustrates a supercell of this compound: in each unit cell we have four Li and four O atoms, and the optimised lattice parameters are $a = 3.153$ Å and $c/a = 2.433$. Using these parameters, we calculate a band gap of $E_g = 2.05$ eV, electron effective masses in- and out-of-plane of $2.19 m_e$ and $0.42 m_e$, respectively, and in-/out-of-plane relative dielectric constants $\epsilon^\infty = 2.73/3.94$ and $\epsilon^0 = 8.36/14.20$. The highest phonon energy that we calculate is $\hbar\omega_{\max} = 98.2$ meV, and the in-/out-of-plane Fröhlich coupling constants[40] are $\alpha = 4.74/1.54$. Our calculations are in good agreement with previous ones yielding $a = 3.17$ Å [204], $c/a = 2.43$ [204], $E_g = 3.6\text{-}4.8$ eV [205], and $\hbar\omega_{\max} = 99.3$ meV [206].

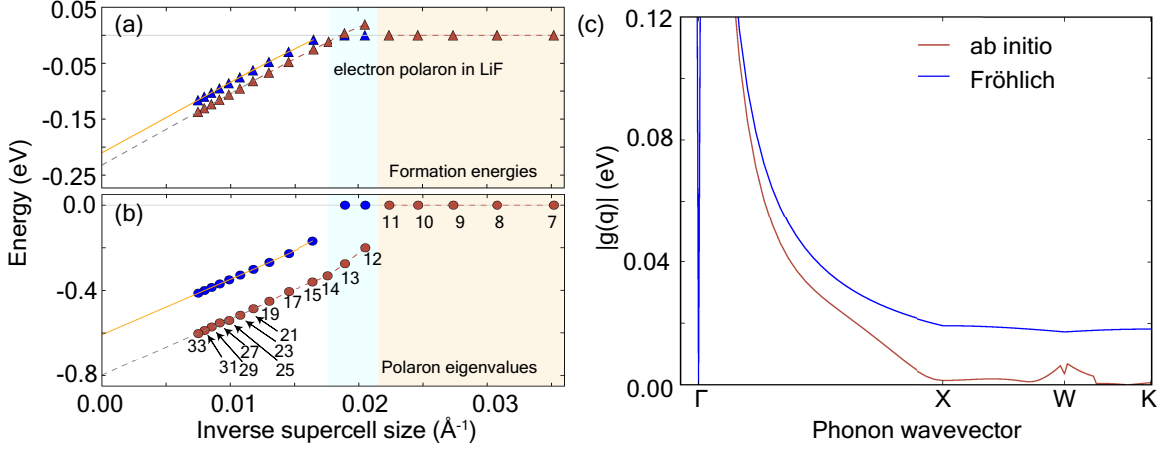


Figure 6.2: (a) Formation energy ΔE_f of the electron polaron in LiF vs. supercell size. We give the size as L^{-1} , where L^3 is the supercell volume. Brown symbols are our calculations using the polaron equations of this work. The dashed grey line is the Makov-Payne extrapolation to infinite supercell size. The orange line is the result of the LP model. The blue symbols are our calculations after considering a parabolic band and a dispersionless LO mode. The shaded regions (blue for parabolic and dispersionless and brown for *ab initio*) indicate supercells for which we did not find self-trapped polarons. (b) Same as in (a), but this time for the polaron eigenvalue ε . The numbers next to the data points indicate the supercell size, for example 33 means a supercell of size $33 \times 33 \times 33$. (c) Electron-phonon matrix element for an electron at the conduction band bottom of LiF, as a function of the phonon wavevector $|\mathbf{q}|$. The brown line is the *ab initio* matrix element, the blue line is the Fröhlich approximation, which retains only the long-range component.

6.2 Electron polaron in LiF

6.2.1 Polaron energy vs. supercell size and Mott transition

In Fig. 6.2(a),(b) we compare the formation energy and polaron eigenvalue obtained via our Eqs. (4.21)–(4.22) (brown symbols) with the results of the continuous LP model described in Sec. 2.1 (orange lines). We focus on the large electron polaron in LiF for definiteness, and for calculations using the LP model we take $\kappa = 2.53$ and $m^*/m_e = 0.88$ from Sec. 6.1.1. Fig. 6.2(a) shows that the polaron formation energy scales with the supercell size as L^{-1} , as expected. In the LP model, the formation energy extrapolated at infinite supercell size is $\Delta E_f = -210$ meV. In contrast, when

we solve the *ab initio* polaron equations, we find the extrapolated energy $\Delta E_f = -231$ meV. The difference between the LP model and our method relates to the fact that in our *ab initio* calculations the bands, phonons, and electron-phonon matrix elements are not as simple as in the LP model. To demonstrate this point, we show in the same figure a calculation carried out using our method, but after replacing the band structure by a parabolic band with the same effective mass as in the LP mode, the phonon dispersion relations by a single, non-dispersive longitudinal-optical (LO) mode, and retaining only the long-range component of the electron-phonon matrix element. This “trimmed” version of the calculation reproduces the LP model exactly, as shown by the blue symbols in Fig. 6.2(a). A comparison of the *ab initio* electron-phonon matrix element for this mode and the long-range Fröhlich component used in the LP model is shown in Fig. 6.2(c); here we see that the LP model overestimates the strength of the coupling at short range. Besides validating our method, the present comparison highlights the fact that even in a compound as simple as LiF the electron-phonon coupling is more complex than a simple Fröhlich interaction, and that details of band structures, phonon dispersions, and matrix elements are to be taken into account for predictive calculations.

In Fig. 6.2(b) we report the polaron eigenvalue ε measured from the conduction-band bottom as a function of supercell size L . In this case the Makov-Payne extrapolation to infinite supercell size yields $\varepsilon = -800$ meV with our method (brown symbols), and $\varepsilon = -609$ meV with the LP model (orange line). As for the formation energy, also in the case of the polaron eigenvalue we fully recover the LP result when we consider a parabolic band and a non-dispersive LO phonon [blue symbols in Fig. 6.2(b)]. It is interesting to note that in LiF the ratio between the polaron eigenvalue and its formation energy is 3.46; this ratio is close to the prediction of the LP model in Sec. 2.1, which yields $\varepsilon/\Delta E_f = 3$ using the exponential ansatz in Eqs. (2.13) and (2.14); note that in the LP model ΔE_f coincides with the energy

E_{LP} .

In Fig. 6.2(a) we also see that when the LiF supercell is smaller than $12 \times 12 \times 12$ unit cells there is no localised polaron solution, i.e. $\Delta E_f = 0$. The existence of a critical supercell size for polaron formation can be explained in terms of the Mott transition: when the periodic replicas of the polaron are too close, they form an extended wavefunction, and the corresponding lattice deformation is too shallow to trap an electron. In this case the excess electron becomes fully delocalised. Therefore a localised polaron can only form when the overlap between nearest-neighbor replicas is negligible. This is the same criterion used by Mott to identify the metal-insulator transition [207]. Using the Mott criterion in the standard form $r_p n_c^{1/3} = 0.26$ [207], with n_c being the critical density and r_p from Eq. (2.12), we can estimate a critical density:

$$n_c \simeq 3.6 \left(\frac{m^*/m_e}{\kappa} \right)^3 \cdot 10^{21} \text{cm}^{-3}. \quad (6.1)$$

We note that this is only a *crude* estimate since it is based on a simplified solution to the Landau-Pekar model. Using $\kappa = 2.53$ and $m^*/m_e = 0.88$ from Sec. 6.1.1 inside Eq. (6.1), we obtain $n_c = 15 \cdot 10^{19} \text{cm}^{-3}$. This estimate is of the same order of magnitude as our calculation in Fig. 6.2(a), which places the transition between supercells of size 11^3 and 12^3 , that is $n_c = 4 \cdot 10^{19} \text{cm}^{-3}$.

6.2.2 Polaron wavefunctions

Figure 6.3 shows the electron polaron in LiF, obtained by solving Eqs. (4.21) and (4.22). The electron wavefunction is computed using Eqs. (4.36)–(4.37), and the atomic displacements are obtained via Eq. (4.27). Figures 6.3(a),(b) show the electron wavefunction as an isosurface and as a contour plot in a plane cutting through the center, respectively. When we compare with the delocalised electronic state shown in Fig. 6.1(a) we see that now we are in the presence of a localised, but large, polaron.

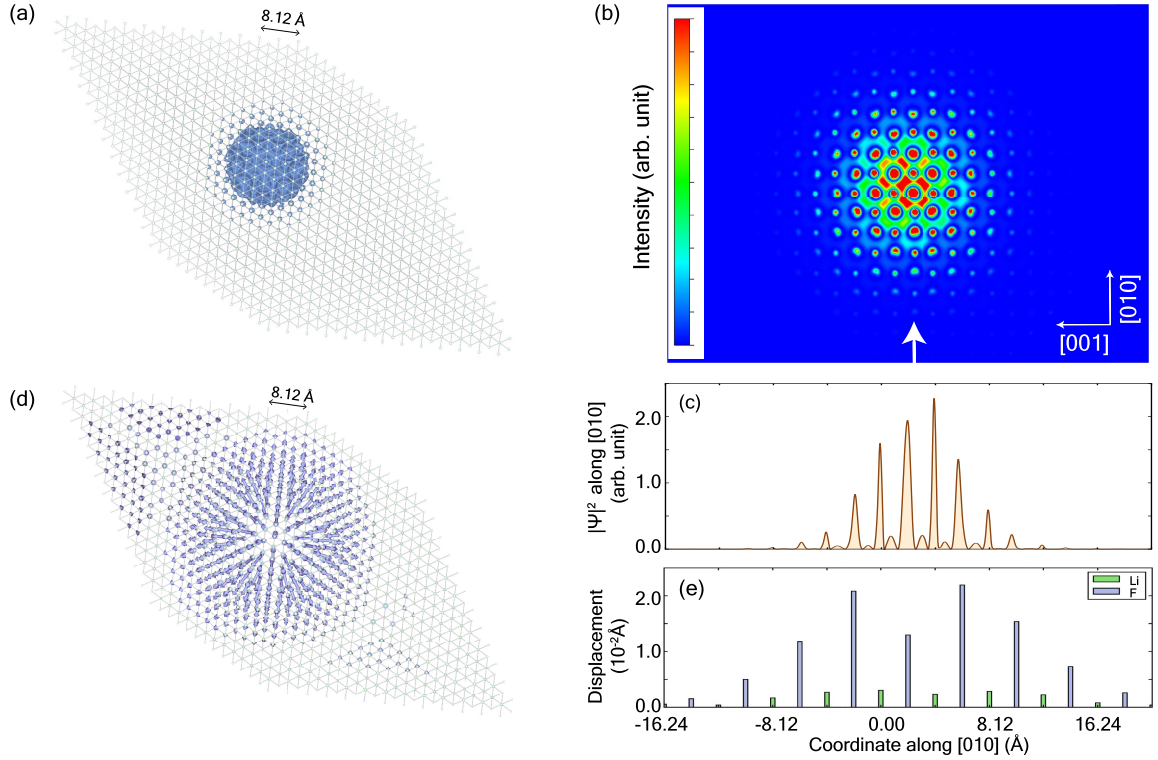


Figure 6.3: Electron polaron in LiF. (a) Isosurface plot of the polaron wavefunction ψ computed with our method for an extra electron in LiF. We use a $12 \times 12 \times 12$ supercell, as it can be seen from the underlying ball-stick model (Li and F are in green and silver, respectively). (b) Same wavefunction as in (a), but as a contour plot in a plane passing through the center and perpendicular to the $[100]$ direction. (c) One-dimensional profile of the polaron density $|\psi|^2$ along the line indicated by the arrow in (b). (d) Displacements of F atoms in this polaron state. The length of the arrows has been scaled $\times 150$ for visualisation purposes. (e) Absolute values of the Li and F displacements along a line passing near the polaron center.

To quantify the spatial extension of the polaron we plot the electron density along a line going through the polaron center, see Fig. 6.3(c). The envelope of the resulting function resembles a Gaussian; if we define the polaron size as the full width at half maximum we obtain $2r_p = 9.0 \text{ \AA}$. Therefore this polaron extends approximately over two unit cells of the LiF lattice. In Figs. 6.3(d),(e) we report the atomic displacements associated with this polaron, using a vectorial representation and a one-dimensional cut, respectively. We note that the displacements of the F anions are consistently larger than those of the Li cations. This may appear counterintuitive because the F anions are heavier, but it is consistent with the fact that the electron charge is mostly concentrated around the Li cations due to the character of the conduction band bottom, therefore the F atoms experience the strongest electrostatic force. The largest atomic displacement is $0.02 \text{ \AA}$, and this value is only 1% of the Li-F bond length. Therefore we are well within the remit of the harmonic approximation.

6.2.3 Spectral decomposition of the polaron

In Figs. 6.4 we present the spectral decomposition of the polaron wavefunctions and atomic displacements in terms of the underlying band states. Figure 6.4(a) shows the electronic weights $|A_{n\mathbf{k}}|^2$ plotted on top of the band structure for the case of the large electron polaron in LiF. The corresponding electronic density of states and spectral function $A^2(E)$ are shown in Fig. 6.4(b). These plots are meant to mimic similar representations of the excitons calculated via the Bethe-Salpeter method [208–210].

The large electron polaron is dominated by states at the bottom of the conduction band; as expected the localisation in reciprocal space mirrors the delocalisation in real space. The corresponding atomic displacements are resolved using the weights $|B_{\mathbf{q}\nu}|^2$ in Fig. 6.4(c), and the density of vibrational states and phonon spectral function $B^2(E)$ are given in Fig. 6.4(d). We see that the polaron is dominated by the LO mode at 76 meV, as expected from earlier work on Fröhlich polarons in halide salts

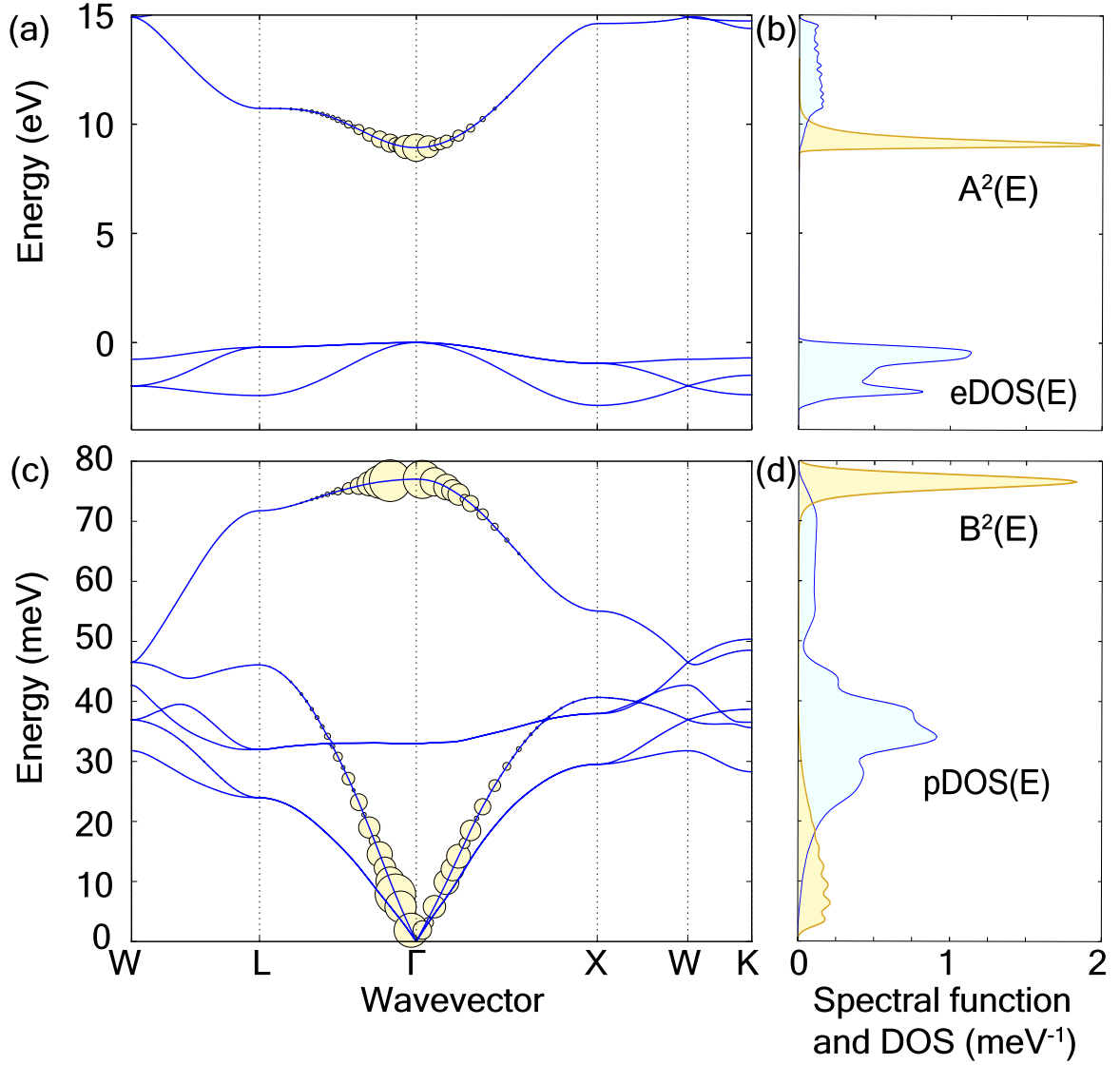


Figure 6.4: Spectral decomposition of the electron polaron in LiF. (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of LiF. The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. The zero of the energy is aligned with the top of the valence bands. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{q\nu}$ plotted on top of the phonon dispersion relations of LiF. The radius of each circle is proportional to $|B_{q\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

[194], but we also have smaller contributions coming from the acoustic branches. By integrating $B^2(E)$ in Fig. 6.4(d) we can quantify the roles of these phonons: we find that the LO mode accounts for 62% of the polaronic distortion, while the the transverse acoustic (TA) mode is responsible for the remaining 38%. To the best of our knowledge this is the first study where the importance of TA phonons in the polaron physics of LiF has been identified. This is precisely the kind of new insight that our method can offer.

6.3 Hole polaron in LiF

6.3.1 Polaron energy vs. supercell size and Mott transition

In Figs. 6.5(a),(b) we show our calculated formation energy and eigenvalue for the hole polaron in LiF, respectively. In this case we obtain self-trapped polarons already for supercells as small as $2 \times 2 \times 2$ unit cells. This result is consistent with Eq. (6.1) and the heavy effective mass of the valence bands. In fact if we use $m^*/m_e = 3.73$ from Sec. 6.1.1 we obtain $n_c = 1.15 \cdot 10^{22} \text{ cm}^{-3}$, which corresponds approximately to one electron in a $2 \times 2 \times 2$ supercell. The Makov-Payne extrapolation yields $\Delta E_f = -1.98 \text{ eV}$ and $\varepsilon = +4.76 \text{ eV}$ (measured from the valence-band top), therefore we are in the presence of a strongly bound polaron. We note that the polaron eigenvalue is positive because the localised hole state lies above the valence band, but the energy is still within the KS gap of this system ($E_g = 8.9 \text{ eV}$ from Sec. 6.1.1). For comparison with explicit DFT calculations, in Fig. 6.5(a) we also report the formation energies calculated in Ref. [46] using SIC or hybrid functionals (filled squares). These calculations correspond to $5 \times 5 \times 5$ supercells and are in very good agreement with our results.

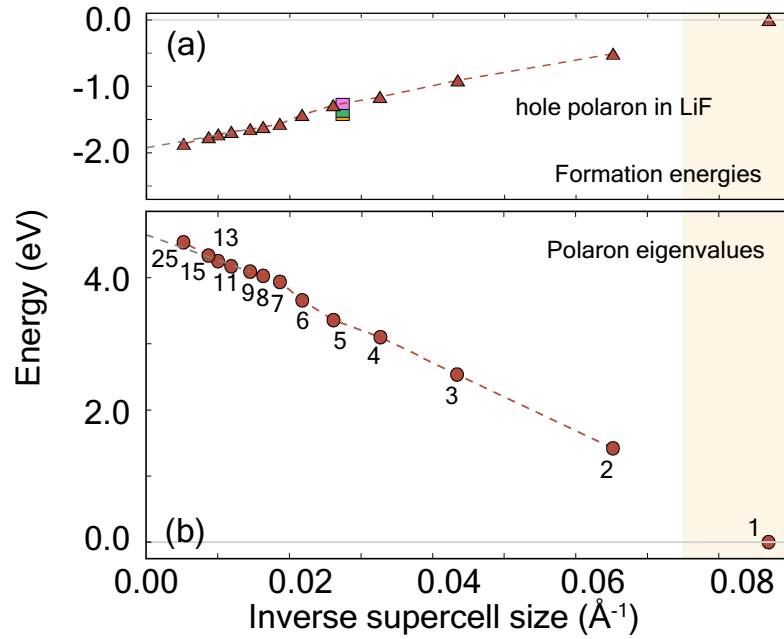


Figure 6.5: (a) Formation energy of the hole polaron in LiF vs. supercell size (brown symbols). The dashed line is the Makov-Payne extrapolation. The filled squares are the formation energies calculated in Ref. [46]. (b) Same as in (a), but for the eigenvalue of the hole polaron in LiF (brown symbols). The numbers represent the supercell size as in Fig. 6.2(b).

6.3.2 Polaron wavefunction

Figure 6.6 shows the hole polaron in LiF, namely (a) the wavefunction isosurface, (b) the same function as a contour plot, (c) a line-cut of the wavefunction, (d) the atomic displacements as arrows, and (e) the size of the displacements along a line passing through the center. Here we are in the presence of a small hole polaron, which is expected given the much heavier masses of the holes as compared to electrons in this system and the much narrowed valence band width [Fig. 6.4(a)]. As it will be discussed in Sec. 6.3.3 the hole polaron in LiF is much closer to a Holstein polaron than a Fröhlich polaron. In this case the wavefunction extends over approximately two atomic orbitals, and from Fig. 6.6(c) we obtain the full width at half maximum $2r_p = 0.97 \text{ \AA}$. Accordingly only a few atoms undergo significant displacements, as shown in Fig. 6.6(e). The largest displacements are found for Li cations, in line with

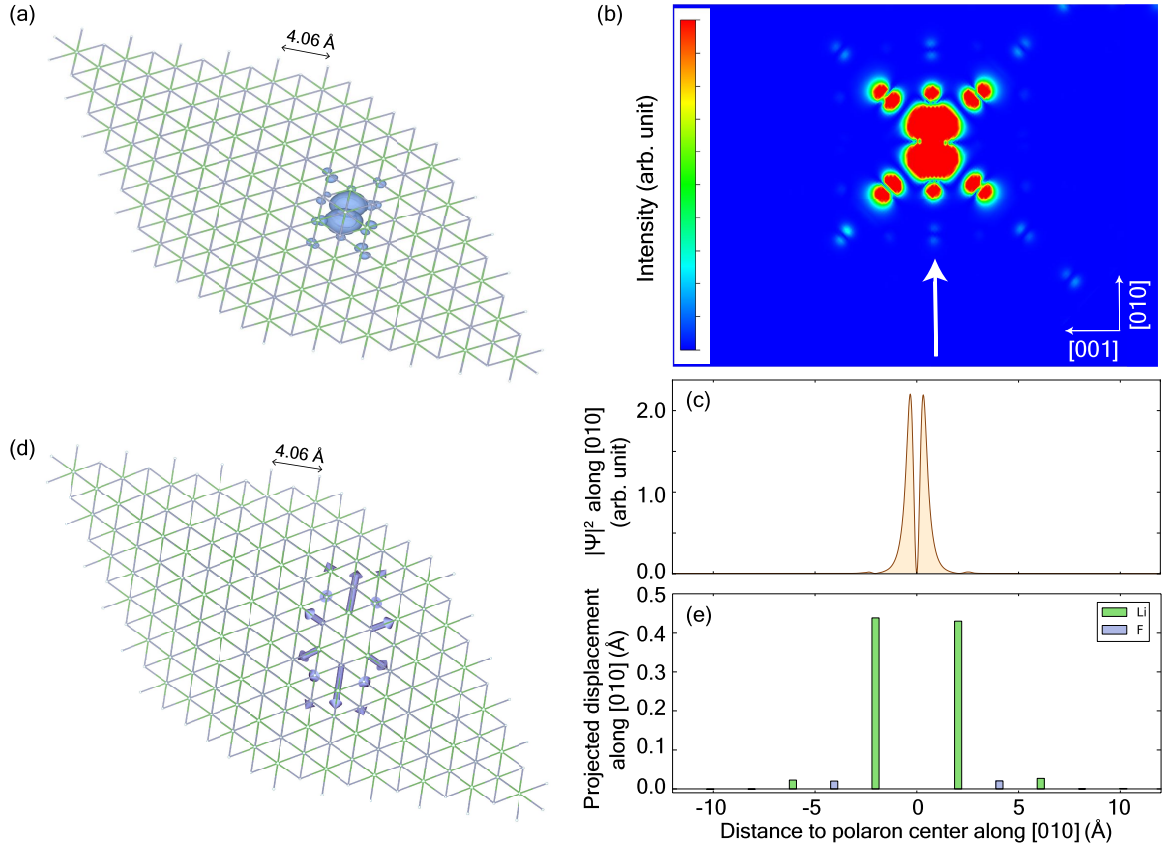


Figure 6.6: Hole polaron in LiF. (a) Isosurface plot of the polaron wavefunction ψ computed with our method for an extra electron in LiF. We use a $5 \times 5 \times 5$ supercell, as shown by the underlying ball-stick model (Li and F are in green and silver, respectively). (b) Same wavefunction as in (a), but as a contour plot in a plane passing through the center and perpendicular to the $[100]$ direction. (c) One-dimensional profile of the polaron density $|\psi|^2$ along the line indicated by the arrow in (b). (d) Displacements of Li atoms in this polaron state. The length of the arrows has been scaled $\times 8$ for clarity. (e) Absolute values of the Li and F displacements along the same line used in (c).

the fact that the wavefunctions at the top of the valence band are localised around the anions, which therefore experience a weaker force. We obtain a maximum displacement of 0.44 Å, which is approximately 20% of the bond length of Li-F (2.03 Å). It is remarkable that our formalism is able to capture this limit of very small polaron, even when the atomic displacements are definitely beyond the harmonic regime. We believe that the reason why the formalism works in this extreme case is that the distortion caused by the small polaron affects only a small portion of the crystal; therefore the use of bands, phonons, and electron-phonon matrix elements calculated for the undistorted unit cell does not lead to significant inaccuracies.

6.3.3 Spectral decomposition of the polaron

Figure 6.7 shows the spectral decomposition of the small hole polaron in LiF. Here the main observation is that the entire highest valence band contributes to the polaronic wavefunction, with smaller contributions from lower-lying bands. This behavior suggests that the small hole polaron of LiF is closer to the Holstein limit [14] than the Fröhlich limit [12]. We emphasise that, at variance with model Hamiltonians, our approach is parameter-free, therefore it captures seamlessly both limits. Also in this case the LO phonon branch dominates the coupling, however now it is the entire branch that contributes, as shown in Fig. 6.7(d). This observation is in line with the fact that the small polaron requires short-range electron-phonon coupling, therefore the range of important phonon wavevectors must extend away from the zone center. By integrating the spectral function $B^2(E)$ we find that the LO branch contributes 78% of the coupling in this case.

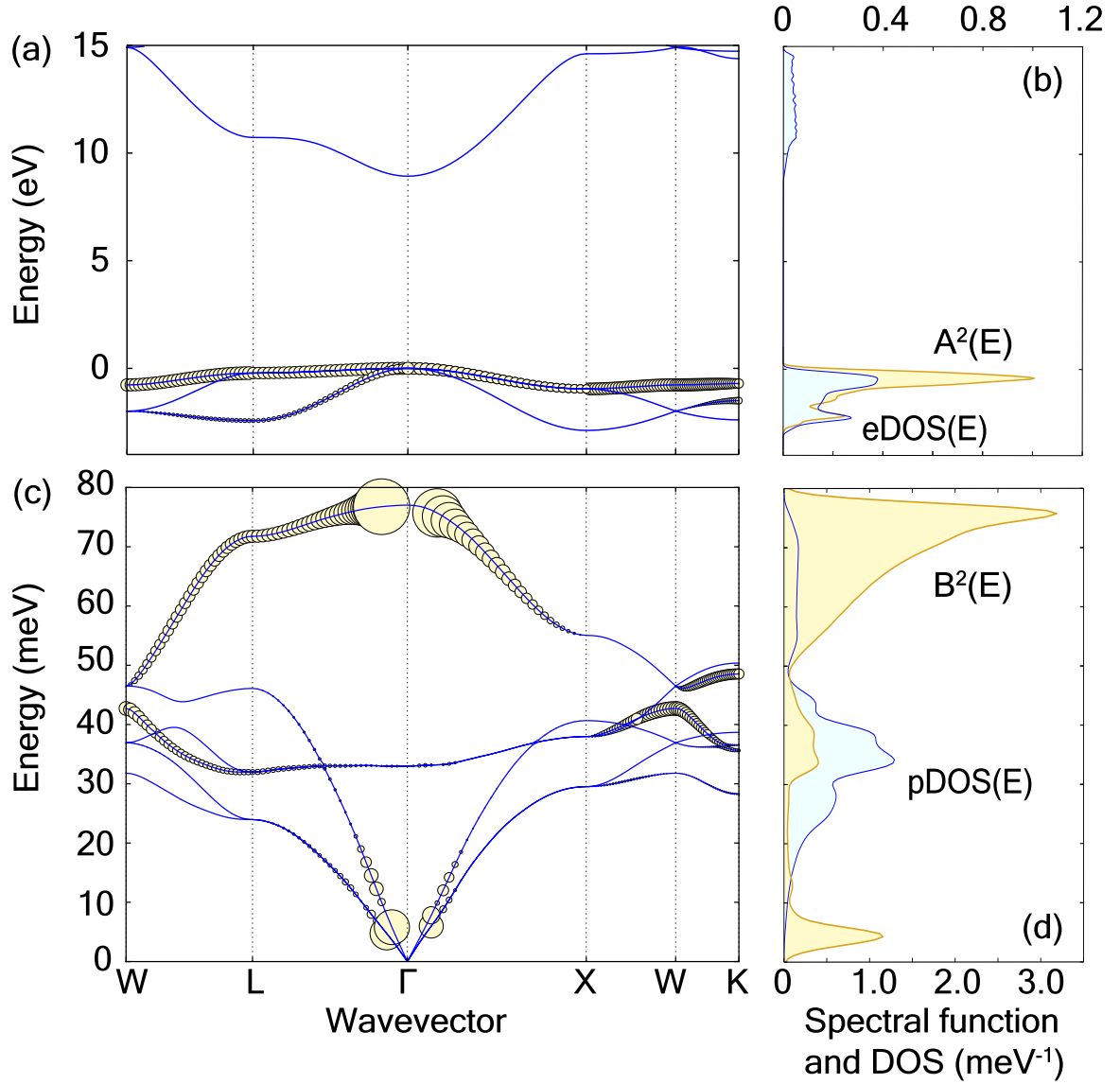


Figure 6.7: Spectral decomposition of the hole polaron in LiF. (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of LiF. The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{\mathbf{q}\nu}$ plotted on top of the phonon dispersion relations of LiF. The radius of each circle is proportional to $|B_{\mathbf{q}\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

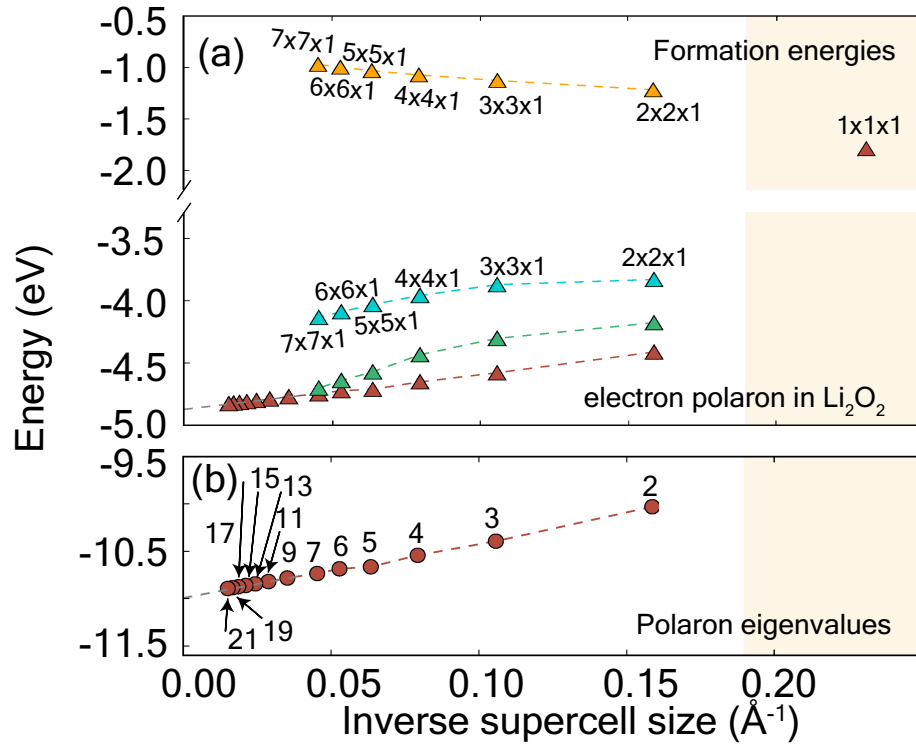


Figure 6.8: (a) Formation energy of the electron polaron in Li_2O_2 vs. supercell size. $N \times N \times 1$ indicates a non-uniform supercell (used for computational convenience). We compare the results of our polaron equations (green symbols) and our explicit DFT calculations with the SIC functional of Eq. (4.44) (cyan symbols), both on non-uniform $N \times N \times 1$ supercells. We also include DFT calculations without self-interaction correction (orange symbols), and our polaron equations on uniform supercells (brown symbols). (b) The eigenvalue of the electron polaron in Li_2O_2 vs. supercell size (brown symbols). The notation is the same as in Fig. 6.2(b).

6.4 Electron polaron in Li_2O_2

6.4.1 Polaron energy vs. supercell size and Mott transition

In Figs. 6.8(a),(b) we show the eigenvalues and formation energies of the electron polaron in Li_2O_2 , respectively, as a function of supercell size. Using $\kappa = 4.05$ and $m^*/m_e = 2.19$ from Sec. 6.1.2 inside Eq. (6.1), we obtain the estimate $n_c = 6 \cdot 10^{20} \text{ cm}^{-3}$; therefore we expect to see localised solutions already for supercells as small as $5 \times 5 \times 1$ unit cells. Our calculations indeed find polarons already at $2 \times 2 \times 1$, see Fig. 6.8(a). In this case the formation energy and eigenvalue extrapolated at infinite supercell size (brown symbols) are $\Delta E_f = -4.87 \text{ eV}$ and $\varepsilon = -10.98 \text{ eV}$, respectively. The polaron eigenvalue falls within a band gap in the valence manifold. In this figure we also compare to direct calculations using the SIC functional of Sec. 4.2. The formation energy in our explicit DFT SIC calculation (cyan symbols) is close to the results of our linear response polaron equations (green symbols), and exhibits the same trend as a function of supercell size; the DFT SIC calculation for the largest supercell considered here ($7 \times 7 \times 1$) yields -4.13 eV , to be compared to our linear-response result -4.70 eV . The deviation of $\sim 14\%$ can be attributed to the fact that our formalism neglects the response of the valence electrons to the localised lattice distortion caused by this strongly bound polaron, or to the fact that the approximation of linear electron-phonon coupling becomes inaccurate for such large atomic displacements. In the same figure we also show that a DFT calculation without SIC fails to predict the correct formation energy, and tends to delocalise the polaron when increasing the supercell size (orange symbols). We note that a study of the scaling of the polaron energy vs. supercell size has not yet been reported in the literature.

6.4.2 Polaron wavefunction

Figure 6.9 shows the electron polaron in Li_2O_2 . In this case we compare three calculations: in (a)-(d) we show the small electron polaron obtained with our formalism; in (e)-(h) we show an explicit supercell calculation using the SIC functional of Sec. 4.2; in (i)-(l) we show the results of a standard DFT calculation without SIC. In each column we report, from top to bottom: the electron wavefunction, its one-dimensional cut across the polaron center, the atomic displacements as arrows, and the one-dimensional cut of these displacements. The first observation to be made is that standard DFT yields a two-dimensional electronic state that is localised along the c axis [Fig. 6.9(i)] but delocalised in the ab plane. The SIC leads to electron localisation also in the plane, and this is observed both in the explicit supercell calculation [Fig. 6.9(e)] and using our method [Fig. 6.9(a)]. The explicit supercell calculation yields a slightly asymmetric wavefunction, while our method gives a perfectly symmetric polaron. This is an artifact of the constraint $\psi_{v\mathbf{k}\uparrow} = \psi_{v\mathbf{k}\downarrow}$ used in the SIC calculation, in fact previous work using hybrid functionals and supercells also found a symmetric polaron [203], as in our method. In this case the polaron is also very small, and extends over two adjacent O- p orbitals. From Fig. 6.9(b) we determine $2r_p = 0.63 \text{ \AA}$, and from Fig. 6.9(d) we find the largest displacement to be $0.38 \text{ \AA}$. Also in this case the atomic displacements are large ($\sim 25\%$ of the O-O distance, $1.51 \text{ \AA}$), but our method correctly predicts the distorted structure as the explicit DFT SIC calculation. This success is remarkable if we consider that our theory is based on small displacements and linear electron-phonon interactions.

6.4.3 Cross validation of the SIC functional

In addition to the first test on a defect induced hole polaron in the α -quartz, we perform a second test to check the geometry of the small electron polaron in Li_2O_2 .

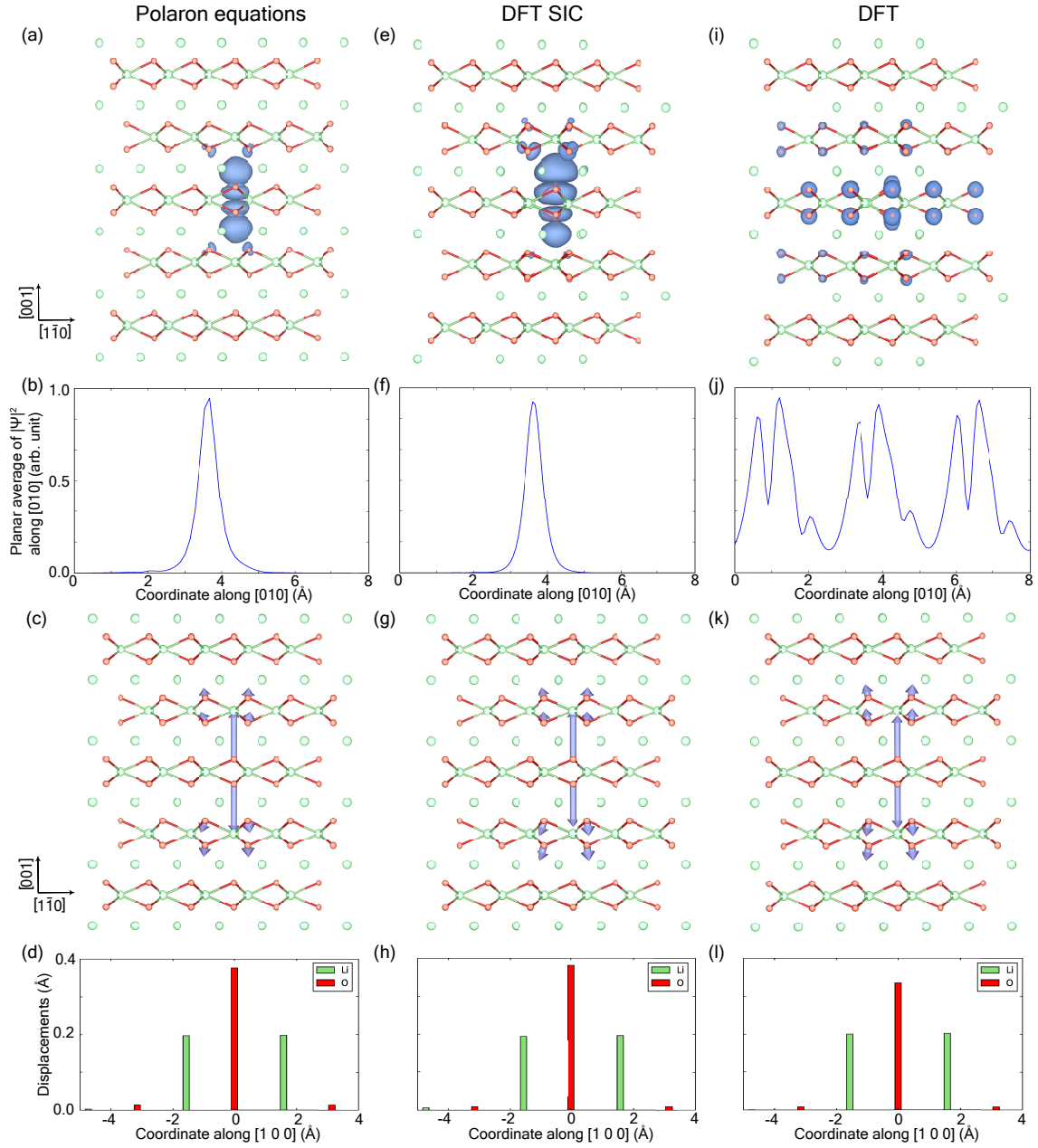


Figure 6.9: Electron polaron in Li_2O_2 . All calculations were performed in a $3 \times 3 \times 3$ supercell. (a) Isosurface plot of the polaron wavefunction ψ , computed using our method. The green and red spheres are Li and O atoms, respectively. (b) Planar average of $|\psi|^2$ along a $[010]$ line passing through the polaron center. (c) Displacements of the O atoms in this polaron, amplified $\times 8$ for clarity. (d) Absolute value of the displacements along a $[100]$ line passing near the polaron center. In (e)-(h) we repropose the same set of data, this time using the DFT SIC functional of Eq. (4.44). In (i)-(l) we repropose the same set of data, this time using standard DFT calculations without SIC.

In this case previous work finds an electron localised around two nearest-neighbor O atoms in the LiO_2 plane, see for example Fig. 6.9(a). The O-O distance in the pristine lattice is 1.54 Å (1.51 Å in Ref. [203]). Using hybrid functional calculations, Ref. [203] reported that this distance increases to 2.20 Å upon adding one excess electron in a supercell with 192 atoms. Our calculations using the SIC functional of Eq. (4.44) also yield an electron localised around the same pair of oxygen atoms, as shown in Fig. 6.9(e). The resulting O-O distance is 2.25 Å, only 2 % larger than in Ref. [203].

6.4.4 Spectral decomposition of the polaron

Finally Fig. 6.10 shows the spectral decomposition for the small electron polaron in Li_2O_2 . In this case the two lowest conduction bands contribute equally to the polaron wavefunctions. This is a case where one-band model Hamiltonians such as the models of Fröhlich and Feynman would not be sufficient to capture the essential features of the problem. We also point out that the higher-lying conduction bands do not contribute appreciably to the polaron wavefunction, as it can be seen from the spectral density $A^2(E)$ in Fig. 6.10(b). The largest contribution to the lattice distortion comes from TO modes around 96 meV, which account for 64% of the coupling.

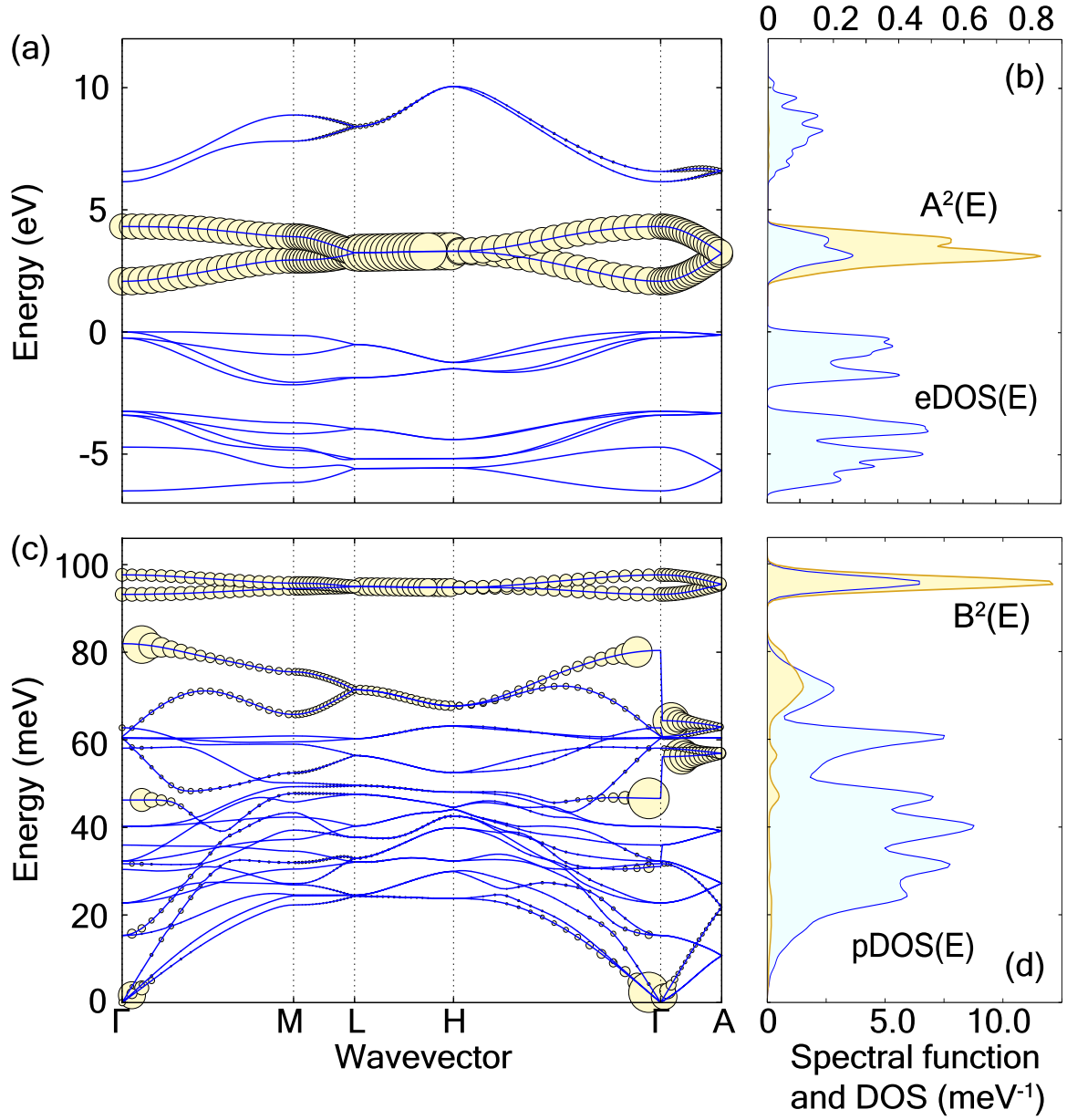


Figure 6.10: Spectral decomposition of the electron polaron in Li_2O_2 . (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of Li_2O_2 . The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. In this case we scale all radii by a large factor in order to show the tiny contribution arising from the topmost unoccupied bands. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{\mathbf{q}\nu}$ plotted on top of the phonon dispersion relations of Li_2O_2 . The radius of each circle is proportional to $|B_{\mathbf{q}\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

Chapter 7

Application II: 2D materials

In this chapter, we apply our methodology to study polarons in 2D materials. The main purpose is to examine the influence of the dimensionality effects on the polaron properties, by using hexagonal boron nitride as an example. This allows us to observe for the first time a 2D polaron wavefunction with atomistic details. After identifying the key contributions from KS states and phonons to the formation of the large polarons, we describe how to converge polaron energies in monolayer calculations. Lastly, we elucidate the origin of dimensionality effects by comparing the polaron energies, wavefunctions and spectral decompositions between the bulk and the monolayer limit.

7.1 Hexagonal boron nitride

As a rapidly growing field in nanoscience and technology, graphene [56] and the related classes of two-dimensional (2D) materials, including transition metal dichalcogenides (TMDCs) [57, 211], phosphorenes [58, 212], chalcogenides [59] and boron nitrides [60], have received substantial attention in recent years [213–216]. 2D materials are interesting due to their electronic structure and the potential to study novel van der Waals heterostructures [61, 62], 2D materials are considered as a potential platform for next generation optoelectronic devices [217–219]. One of the applications of these 2D materials is as transparent layers in optoelectronic devices. For instances, graphene layers can serve as a high-performance transparent conducting electrode in solar cells because of their tunable conductivity and transparency with increasing number of

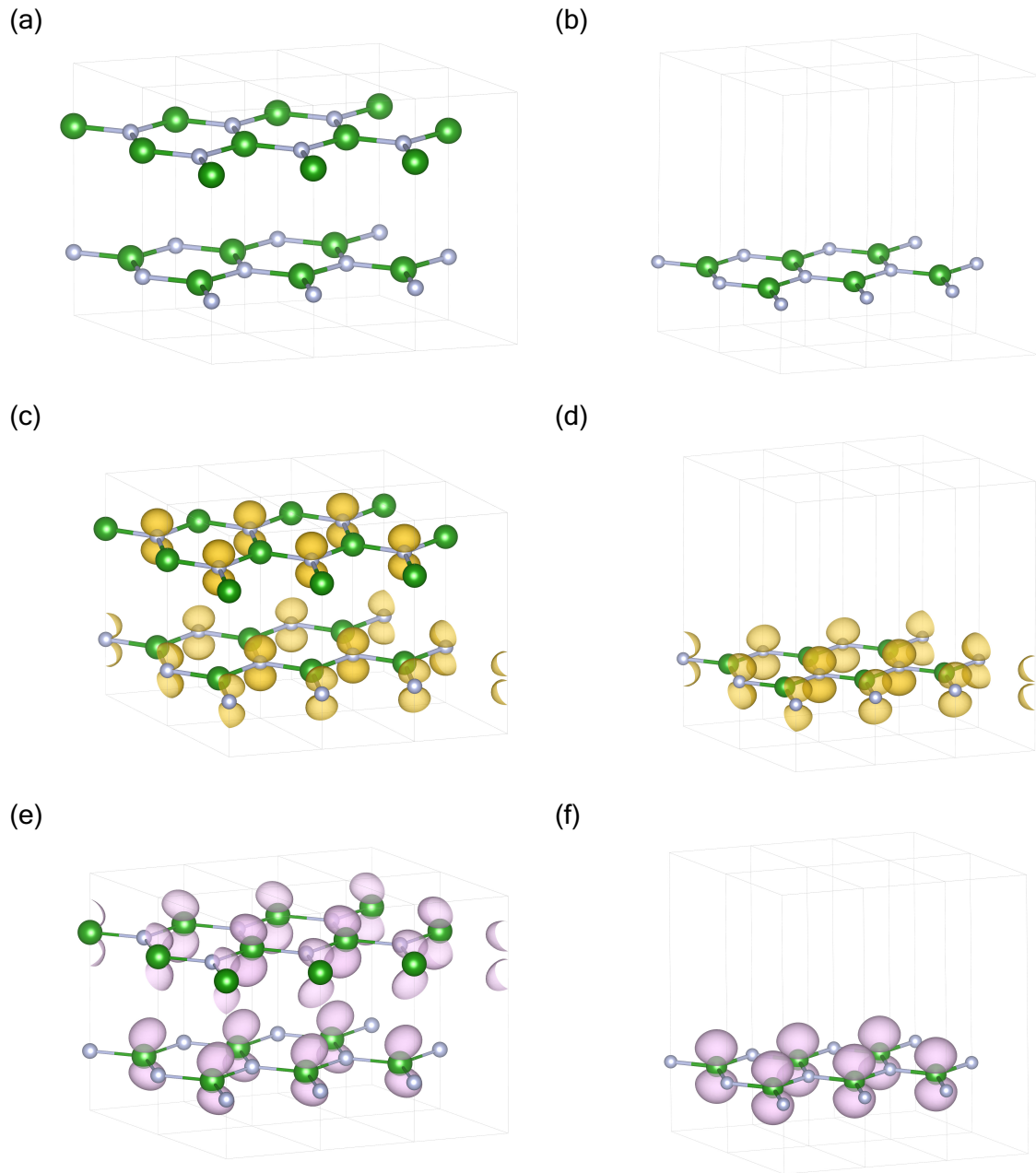


Figure 7.1: Ball and stick models of (a) bulk h-BN which exhibits AA'-type stacking with alternating B and N atoms along the $[001]$ direction; (b) h-BN monolayer in a $3 \times 2 \times 1$ supercell, with B and N atoms in green and silver respectively. We show in (c) an isosurface plot of the density at the valence band top, with h-BN in bulk ; (d) an isosurface plot of the density at the valence band top, with h-BN in monolayer; (e) same as (c) but this time at the conduction band bottom; and (f) same as (d) but this time at the conduction band bottom.

layers [220]; TMDCs such as MoS₂ and WS₂ can be used as hole-extraction layers in low-cost perovskite solar cells due to their desirable work functions [221, 222].

Hexagonal boron nitride (h-BN) is an isomorph of graphite with a similar layered structure but exhibits a wide-bandgap of ~ 6 eV in the ultraviolet (UV) spectral range. Due to its insulating nature, thermal and chemical stability, it has initially been used as an engineering ceramic for a long period of time [223], and as a substrate for graphene and other heterostructures since the past decade [224, 225]. The large bandgap and high internal quantum efficiency in h-BN make it a promising candidate for deep-UV lighting devices [226, 227]; the capability of hosting colour centers make it promising to engineer for a single-photon emitters [228, 229]. These novel applications motivate more in-depth theoretical studies of exciton-phonon couplings of by h-BN [230–232].

To date, a detailed picture of pure polaronic effects in h-BN and how they scale with dimensionality is still missing. Since the excitonic effects are more pronounced in fewer layers as compared to the bulk, due to the reduced electronic screening in lower dimensions, it is important to characterise whether the polaronic effect would be enhanced and become comparable to the excitonic effects in monolayers. If this is the case, a more subtle question is how to incorporate large polaronic effects within the excitonic picture.

In this work, we calculate the polaron excitation energies, wavefunctions and spectral decomposition for both hole and electron polarons in h-BN monolayer and bulk, by employing our recent developed *ab initio* formalism of Chapter 4. Our calculations show that both polarons exhibit a very similar long-range Fröhlich coupling nature. In order to correctly extrapolate the polaron energies, we follow the method in Ref. [233] to eliminate the spurious electrostatic interactions in the out-of-plane direction. Finally, we show that the hole polaron in the monolayer which yield $\sim 50\%$ more energetic than in the bulk crystal.

7.2 Computational details

We perform ground state electronic structure and phonon calculations for h-BN using the same electronic structure code as outlined in Sec. 5.1. To describe better the interlayer interaction for the bulk h-BN as shown in Fig. 7.1(a), we use the local density approximation (LDA) to DFT [234], and ONCV pseudopotentials [173], with planewaves kinetic energy cutoffs of 125 Ry. In the ground-state calculations of electrons and phonons, we sample the Brillouin zone with a Γ -centred uniform grid of $14 \times 14 \times 6$ points. The lattice vectors and internal coordinates are optimised using this setup before proceeding to the polaron calculations for the h-BN monolayer. To construct a monolayer as shown in Fig. 7.1(b), we remove the upper layer in bulk h-BN and increase the out-of-plane lattice constant to 20 Å. The latter choice is essential to prevent spurious interactions between periodic replicas [235]. In Sec. 7.3.1 we show that the direct gap of h-BN monolayer can only be captured correctly when the interlayer distance increases beyond 10 Å; we also confirmed that the polaron energies are converged with this choice (see Sec. 7.4.2). In the ground-state calculations of electrons and phonons for the monolayer, we sample the Brillouin zone with a Γ -centred uniform grid of $16 \times 16 \times 2$ points.

We note that our computed electron (conduction band) effective mass of h-BN monolayer using LDA differs by more than 20 % from more accurate hybrid functional calculations [236]. In order to obtain a better in-plane electron mass, which is important to completely describe polarons, as we seen in Sec. 2.1, we follow exactly the same procedures as outlined above, but this time using the generalised gradient approximation to DFT of PBE [172], and ONCV pseudopotentials [173], with the same planewaves kinetic energy cutoffs of 125 Ry. In the ground-state calculations of electrons and phonons for the monolayer, we sample the Brillouin zone with a Γ -centred uniform grid of $16 \times 16 \times 2$ points.

Figure 7.1(a) illustrates the supercells used for bulk h-BN. Since the interlayer interactions in bulk layers are well described by LDA but not PBE, here we only report calculations with LDA for bulk h-BN. Our optimised lattice parameters for the bulk unit cell are $a = 2.49$ Å (LDA) and $c/a = 2.60$, in agreement with the experimental values $a = 2.50$ Å and $c/a = 2.66$ [237]. Our calculated KS band gap is $E_g = 4.06$ eV, and underestimates the experimental optical gap of 5.97 eV as expected [237]. We find in-plane hole effective mass of $0.54 m_e$ (LDA), which is in good agreement with the reported value of $0.49 m_e$ [238]. The in (out-of)-plane relative dielectric constants $\epsilon_{||}^{\infty}$ ($\epsilon_{\perp}^{\infty}$) = 4.85 (2.85) and $\epsilon_{||}^0$ ($\epsilon_{\perp}^0$) = 6.61 (3.39). The highest computed phonon energy is $\hbar\omega_{\max} = 201$ meV, close to the reported value $\hbar\omega_{\max} = 200$ meV [239], and the Fröhlich couplings constant for the hole is $\alpha_{||}$ ($\alpha_{\perp}$) = 0.331 (0.337).

Figure 7.1(b) illustrates the supercells of h-BN in monolayer. In the monolayer calculations we use the same lattice constants optimised in bulk $a = 2.49/2.51$ Å (LDA/PBE), but adjusting the aspect ratio to $c/a = 8.04/7.97$ (LDA/PBE) which corresponds to the vacuum size of 20 Å. For the monolayer we find the in-plane hole and electron effective massess of $0.62/0.65 m_e$ (LDA/PBE) and $0.89/0.94 m_e$ (LDA/PBE) respectively. Both the hole and electron masses are in a good agreement with the reported values $0.61\text{-}0.63 m_e$ and $0.83\text{-}1.07 m_e$ [216, 236], respectively. The in (out-of)-plane relative dielectric constants $\epsilon_{||}^{\infty}$ ($\epsilon_{\perp}^{\infty}$) = 1.63 (1.12)/1.62 (1.12) (LDA/PBE) and $\epsilon_{||}^0$ ($\epsilon_{\perp}^0$) = 1.91 (1.12)/1.92 (1.12) (LDA/PBE).

7.3 Preliminary analyses

We perform calculations for both the hole polaron and the electron polaron in a h-BN monolayer, and we show that both polarons exhibit the Fröhlich long-range character as anticipated in Ref. [240] and should be classified as large polarons. When we solve the *ab initio* polaron equations, we sample the grid only along the in-plane

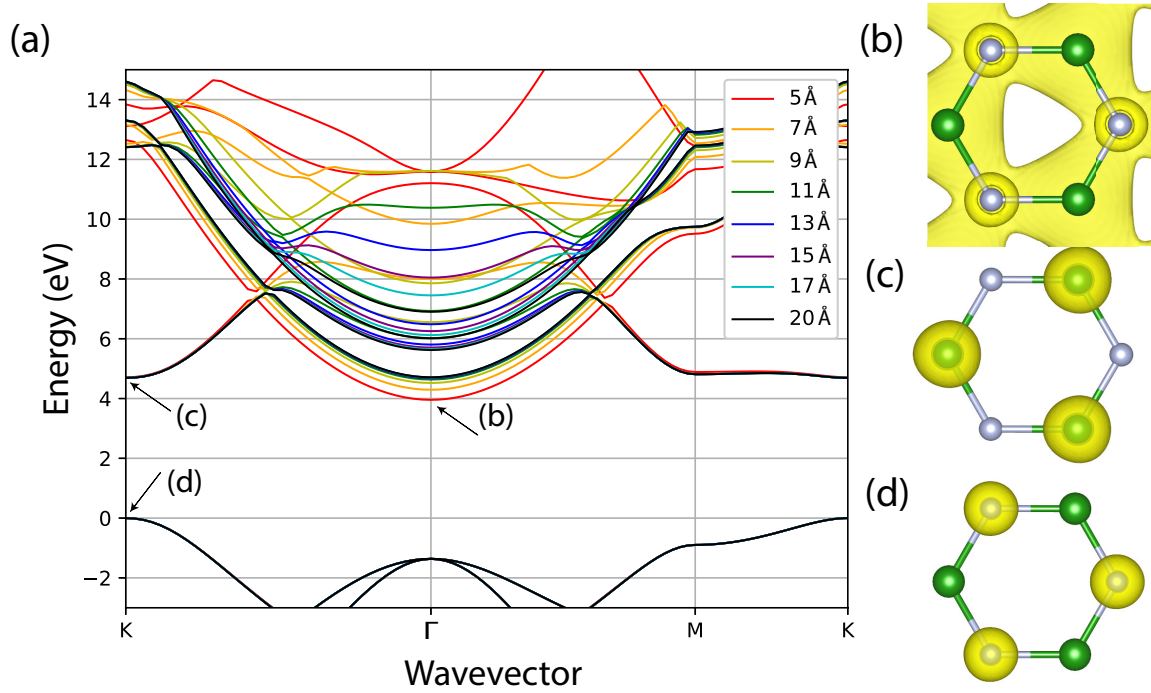


Figure 7.2: (a) Band structures for the h-BN monolayer calculated by PBE functional. Plots of the electron density of the KS state at (b) the conduction band bottom at Γ ; (c) the conduction band bottom at K; and (d) the valence band top at K, respectively.

direction while keeping $k_z = 0$. This choice leads to incorrect energetics but it is computationally very affordable and allows us to perform a preliminary analysis. The equivalent physical picture is that the constructed supercells is still unit cell periodic in the z -direction, each of them hosts a localised polaron but with the same unit cell periodicity along z . Since 3D Fröhlich interactions are long-range, we expect that the polaron energies would diverge while such in-plane expansions of the supercell approach to infinite. If we refer to the polaron interaction term in Eq. (4.13), we note that this divergence is indeed closely related to the Coulomb divergence in the Ewald summation, which happens when a charge is presented under a periodic boundary condition. However, we still consider this set up as a preliminary calculation which allows us to analyse which electron bands and excited phonon modes are involved in the 2D polaron formation, and we will present the Makov-Payne extrapolation of polaron energies in Sec. 7.4.1.

7.3.1 Convergence of band structures against vacuum size

In order to describe the direct bandgap in monolayer of BN [227], we conduct a convergence test to ensure that the interlayer interactions are reduced to a suitable level that the character of the band gap does not change. We perform band structure calculations on unit cells of the h-BN monolayer with various c -axis lengths, from 5 Å up to 20 Å, while the lattice constant is fixed at the one optimised with the PBE functional. Figure 7.2(a) shows the band structures as colour contours, each of them is labelled by the corresponding vacuum size, and the energy zero is aligned with the valence band top. We see that the occupied bands do not depend on the vacuum size, but the unoccupied bands change rapidly when the vacuum size is below 11 Å. To compromise between the interlayer interactions and the number of planewaves, we choose a vacuum size of 20 Å for our monolayer calculations of polarons.

Since polarons in h-BN are constructed from the low-lying states near the band edge, it is useful to explore how these states behave. Figures 7.2(b)-(d) illustrate the electron densities correspond to three band edge states of our interest. In Fig. 7.2(b) we see that a delocalised electron density is present on top of the h-BN layer, which behaves as two-dimensional electron gas and contributes to the parabolic-like band structure in near Γ . In Sec. 7.3.4 we will show that these nearly-free electron states lead to additional couplings in the case of the large electron polaron. For the conduction/valence band edge states at $\mathbf{k}=\text{K}$ as shown in Figs. 7.2(c)-(d), $2p_z$ orbitals are found to localise around the boron and the nitrogen atoms, respectively. These orbitals contribute to the tight-binding bands as recognised in many layered systems [241, 242].

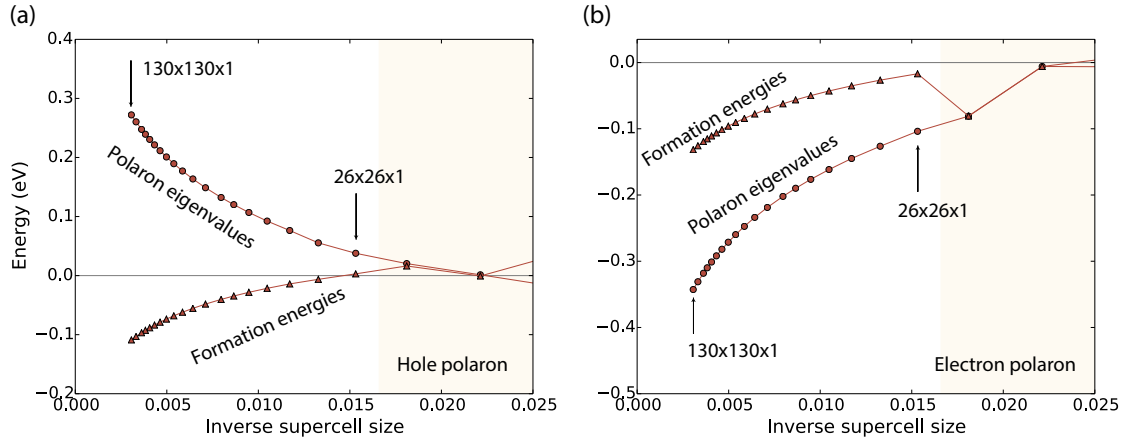


Figure 7.3: Preliminary calculations of polaron formation energies and eigenvalues sampled by the grid size of $L \times L \times 1$ points for (a) the hole polaron; and (b) the electron polaron, respectively.

7.3.2 Polaron energies

Figure 7.3 shows the formation energies and polaron eigenvalues for the hole and electron polarons in the h-BN monolayer. We see that these energies tend to diverge when the Brillouin zone grid is expanded from $18 \times 18 \times 1$ to $130 \times 130 \times 1$, as expected from our previous discussion. In Fig. 7.3(a) we report the calculated formation energies and eigenvalue ε of the hole polaron as a function of the in-plane supercell size L . In this case, we obtain self-trapped polarons beginning at the supercell size corresponding to $22 \times 22 \times 1$ unit cells. In Fig. 7.3(b) we show the formation energies and eigenvalues of the electron polaron. Our calculations find localised electron polarons starting at a supercell size corresponding to $22 \times 22 \times 1$ unit cells. We discuss the interesting feature of these localised polarons further in the following section.

7.3.3 Polaron wavefunctions

Figure 7.4 shows the hole polaron and electron polaron in h-BN monolayer, obtained by solving the *ab initio* polaron equations at the grid size $26 \times 26 \times 1$. The polaron wavefunction is calculated via Eq. (4.17). Figs. 7.4(a) and (b) show the hole polaron

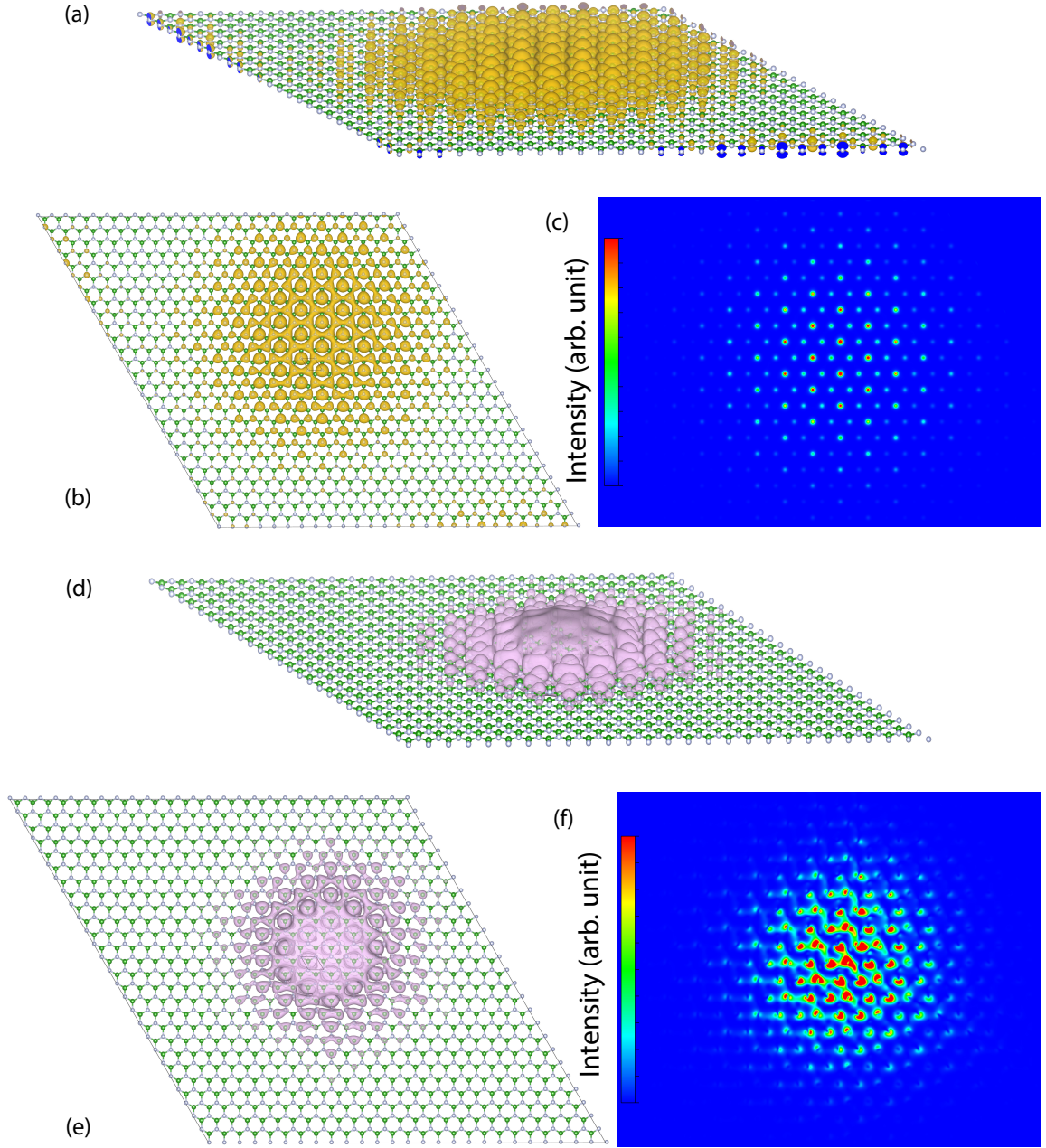


Figure 7.4: Isosurface plots of the polaron wavefunction ψ for the hole polaron [(a)–(c)] and the electron polaron [(d)–(f)] in monolayer h-BN, computed using the *ab initio* polaron formalism. We used the same isosurface values in the plots of (a),(b),(d) and (e) for the comparison of their spacial extent.

wavefunction as an isosurface viewed from the side and from the top, respectively. If we compare these plots with the delocalised hole state shown in Fig. 7.1(d), we see that the polaron is localised but extends over many unit cells. To characterise the orbital nature of this large hole polaron, we show in Fig. 7.4(c) the same wavefunction but as a contour plot in a horizontal plane parallel to the monolayer. Here we recognise the polaron wavefunction is comprised of $2p_z$ orbitals of nitrogen atoms, modulated by a localised envelope function.

Figs. 7.4(d) and (e) show the electron polaron wavefunction as an isosurface, viewed from side and from top, respectively. If we compare these plots with the delocalised electron state shown in Fig. 7.1(f), we also see the electron polaron is localised but extends over many unit cells. The localised radius is smaller than for the hole polaron. This is expected given that the electron polaron possess larger formation energy (if we consider the same supercell size, see Fig. 7.3). Moreover, we also see that the polaron centre is covered by an additional layer of electrons, which is missing in the hole polaron. To perform a similar orbital characterisation for the large electron polaron, we show in Fig. 7.4(f) the same wavefunction but as a contour plot in a horizontal plane parallel to the monolayer. In order to understand the distinctive shape between the electron polaron wavefunction and the hole polaron wavefunction, it is useful to analyse the spectral decomposition of each polaron.

7.3.4 Spectral decompositions of the polaron

Figures 7.5–7.6 illustrate the spectral decomposition of the polaron wavefunctions and atomic displacements in terms of KS states and phonon modes. In Fig. 7.5(a), we visualise the electronic Fourier amplitudes $|A_{n\mathbf{k}}|^2$ by plotting it on top of the band structure for the hole polaron in monolayer h-BN. The corresponding electronic density of states and spectral function $A^2(E)$ are shown in Fig. 7.5(b). We see that the large hole polaron is formed by the states near the valence band maximum, which

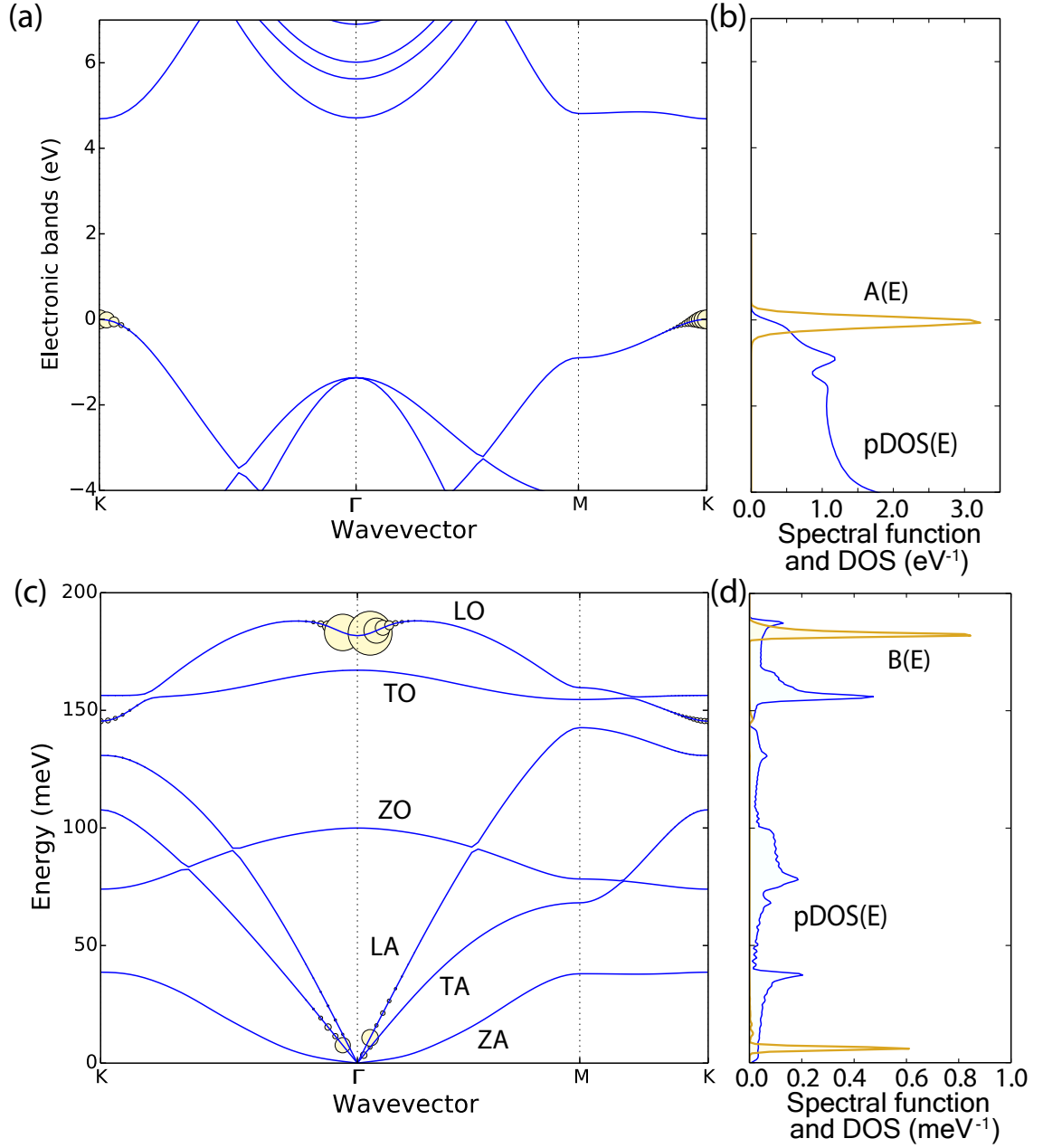


Figure 7.5: Spectral decomposition of the hole polaron in h-BN monolayer. (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of h-BN monolayer. The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. The zero of the energy is aligned with the top of the valence bands. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{\mathbf{q}\nu}$ plotted on top of the phonon dispersion relations of h-BN monolayer. The radius of each circle is proportional to $|B_{\mathbf{q}\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

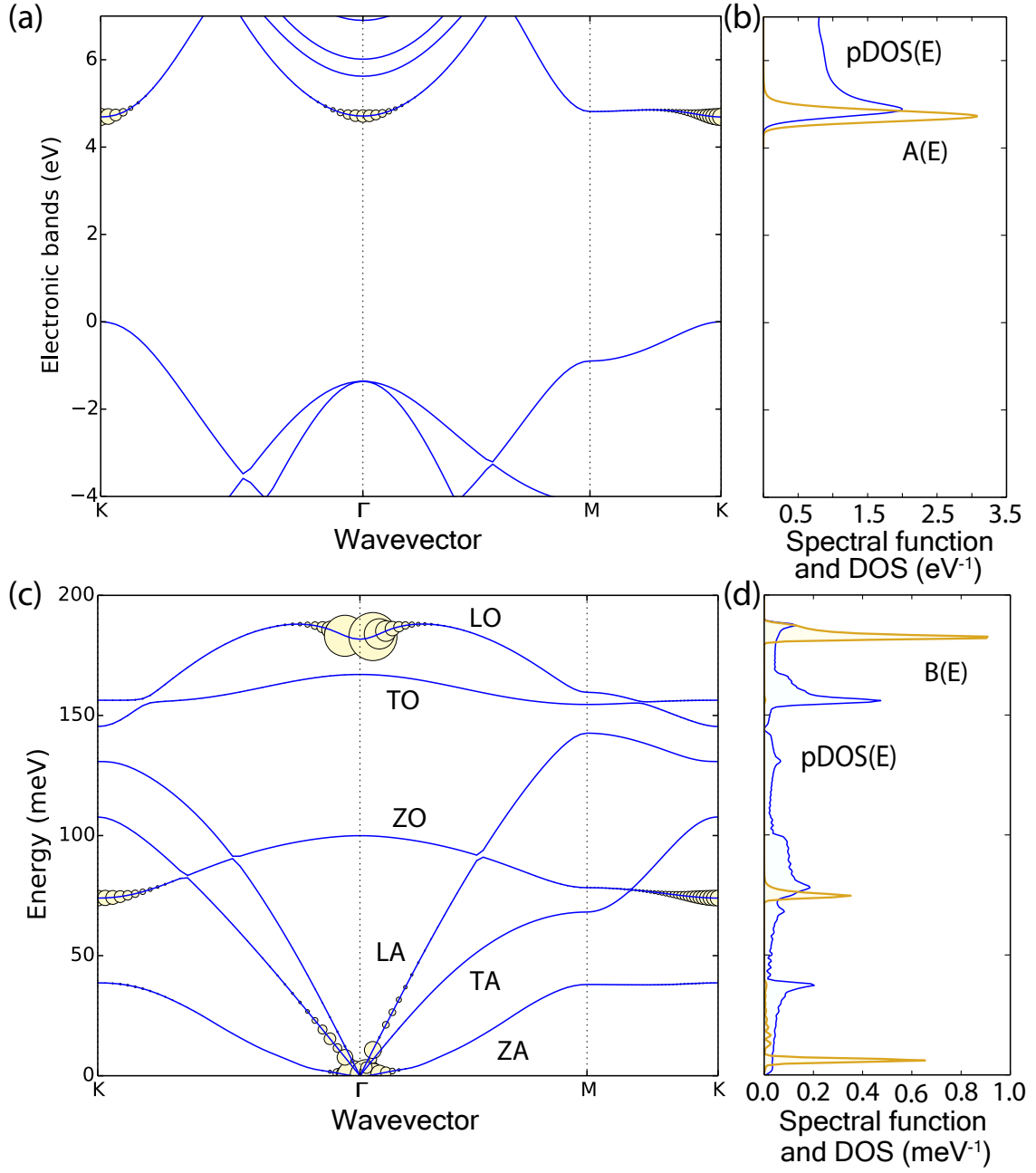


Figure 7.6: Spectral decomposition of the electron polaron in h-BN monolayer. (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of h-BN monolayer. The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. The zero of the energy is aligned with the top of the valence bands. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{\mathbf{q}\nu}$ plotted on top of the phonon dispersion relations of h-BN monolayer. The radius of each circle is proportional to $|B_{\mathbf{q}\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

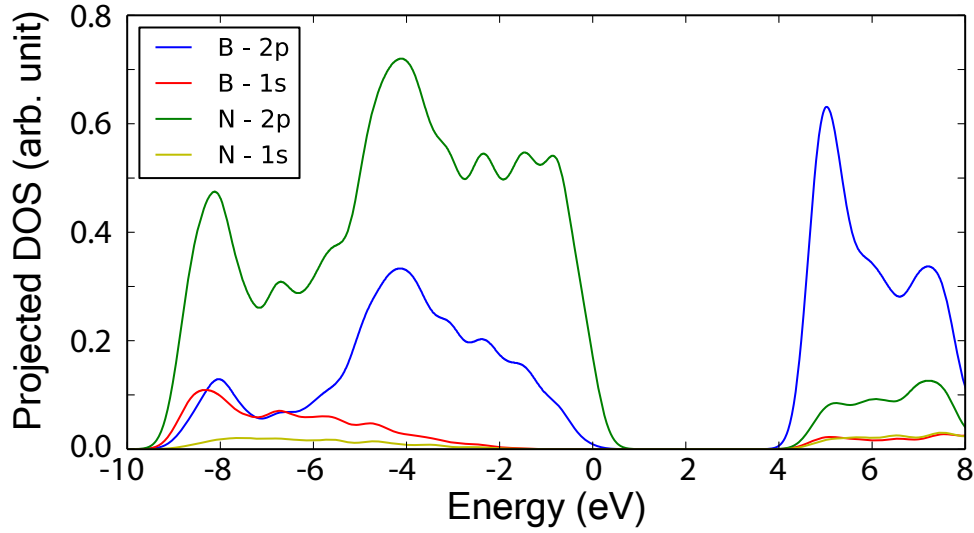


Figure 7.7: Projected density of states against energy for monolayer h-BN.

are dominated by the $2p_z$ orbital of the nitrogen atoms as suggested by the projected density of states in Fig. 7.7. This agrees to the orbital character of the hole polaron wavefunction. In Fig. 7.5(c), we present the displacements Fourier amplitudes $|B_{\mathbf{q}\nu}|^2$, and the density of vibrational states and phonon spectral function $B^2(E)$ are shown in Fig. 7.5(d). We see that the polaron is dominated by the long-wavelength LO E_{2g} phonon at 182 meV, where the two atoms exhibit out-of-phase vibrations along the same direction of the wavevector as shown in Fig. 7.8(f). As in the phonon spectrum of the large polaron in LiF, we also find small contributions from the LA and TA modes here. If we integrate $B^2(E)$ in Fig. 7.5(d), we can quantify the fractional contributions from these phonon modes. For example, we find that the LO mode accounts for 66 % of the total displacement, while the acoustic modes contribute the remaining 34 %.

Figure 7.6 shows the spectral decomposition of the large electron polaron in monolayer h-BN. We see that the main contribution to the large electron polaron comes from the two different conduction bands. In addition to the contribution from the usual band with conduction band minima at the K-point, where the low-lying states

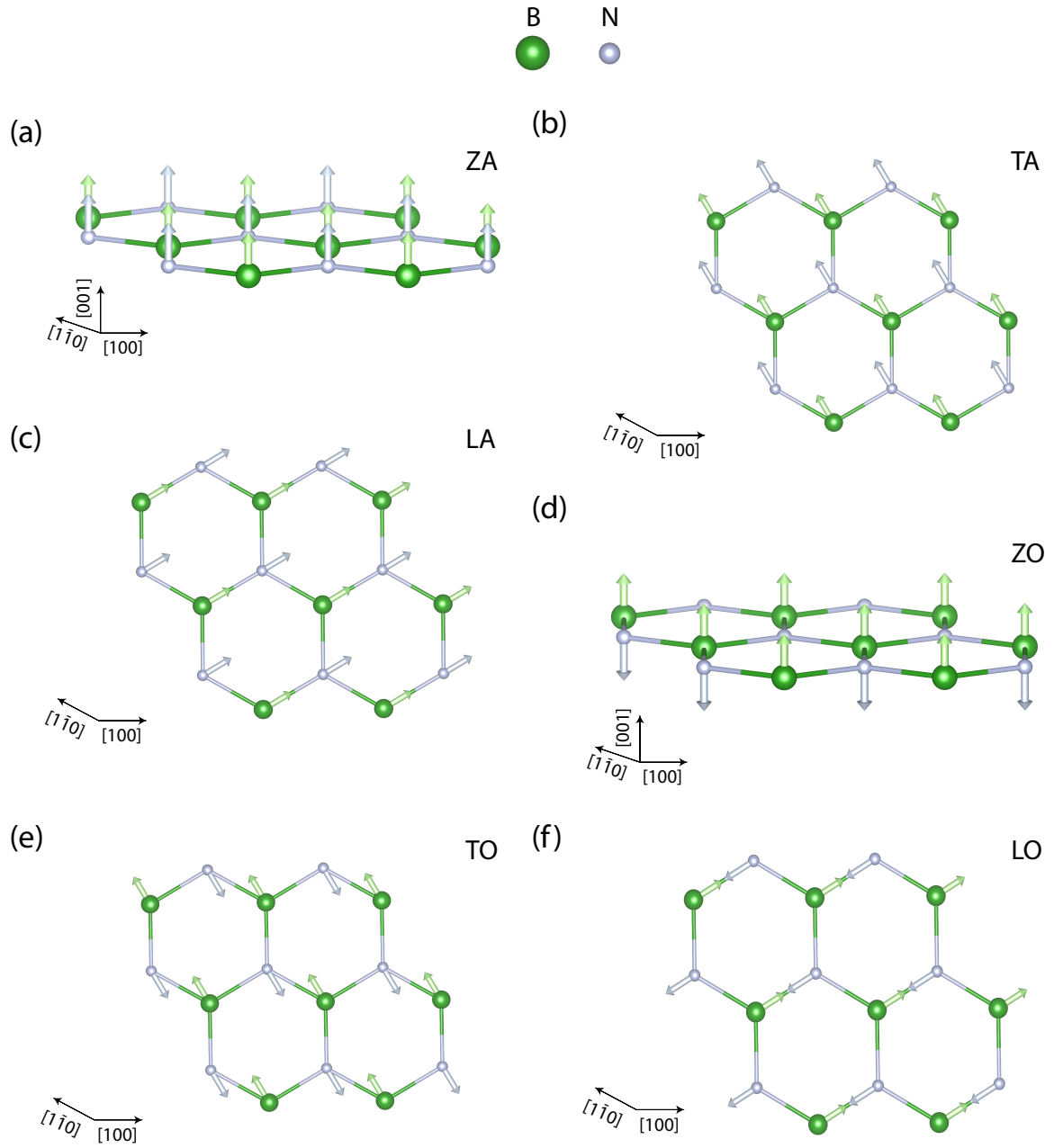


Figure 7.8: Ball-and-stick representation of (a)-(c) acoustic phonons ($\nu = 1 - 3$); and (d)-(f) optical phonons ($\nu = 4 - 6$) at $|\mathbf{q}| \rightarrow 0$ from the Γ -M path, showing for $3 \times 3 \times 1$ unit cells of the monolayer h-BN.

are comprised of the $2p_z$ orbitals of the boron atoms as suggested by the projected density of states in Fig. 7.7, we note that the parabolic band centred at the Γ -point also plays a role. Since the conduction band minima of the two bands are very close, the inclusion of these low energy parabolic states can lead to the coupling to additional phonons. To emphasise this point we compare the plot of $|B_{\mathbf{q}\nu}|^2$ in Fig. 7.6(c) to the case of the hole polaron as shown in Fig. 7.5(c). We see that additional optical phonons are excited near the zone boundary. Again we can quantify the fractional contributions from these phonon modes by integrating the spectral function. We find that the LO modes contribute 55 % of the total excited phonons, the acoustic modes contribute 27 % and the additional ZO phonons [see Fig. 7.8(d)] contribute the remaining 18 %.

Based on this analysis, we see that the measure of the low-lying nearly-free electron band [243, 244] in the case of electron polaron rescales the fractional contributions from the LO and acoustic phonons, by coupling to additional modes. In addition, we note that the low-lying parabolic band is formed by the nearly-free electron states. This observation indeed explains why we see an additional layer-like electron density in the electron polaron as shown in Figs. 7.4(d)–(e), in which these low-lying parabolic states form a localised continuum electron density on top of the large polaron.

7.4 Formation energy of polarons in monolayer h-BN

In order to extrapolate the polaron energies corresponding to the correct monolayer limit, we follow previous works in Refs. [233, 245] by considering a 3D sampling of the Brillouin zone. This approach allows us to perform a Madelung correction to obtain the formation energy of an isolated charge in an infinite supercell. We note that this approach is equivalent to constructing a supercell that contains many

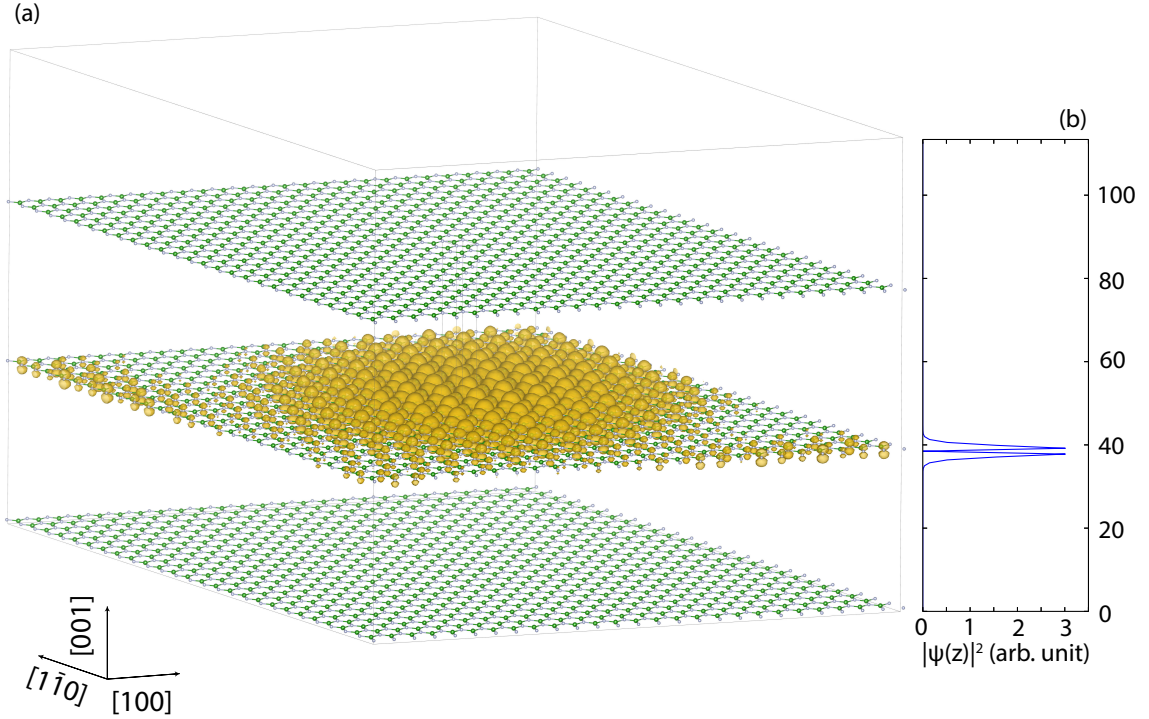


Figure 7.9: (a) An example supercell comprised of $26 \times 26 \times 3$ primitive unit cells of h-BN, where each of them contains a layer of atoms. The hole polaron wavefunction shown here is calculated at the uniform grid of $26 \times 26 \times 3$ points. (b) The horizontal planar average of the polaron wavefunction against the z -axis in Bohr.

isolated monolayers, as shown in Fig. 7.9(a), where each of the layer is separated from its neighbours by a large distance of 20 Å. To guarantee the calculated energies correspond to the monolayer limit, we confirm in Fig. 7.9(b) the calculated polaron wavefunction at the uniform grid of $26 \times 26 \times 3$ points is indeed tightly bounded to one of the layer, and the shape of the wavefunction remains similar to those in Fig. 7.4.

7.4.1 Polaron energies and metal-insulator transitions

Figure 7.10 shows the formation energies and polaron eigenvalues for the hole polaron in the h-BN monolayer as a function of grid size. We see that these energies scale with the supercell size as L^{-1} , when the grid is expanded in 3D from $10 \times 10 \times 1$ points up to $88 \times 88 \times 11$ points. Our calculations find localised hole polarons starting at the

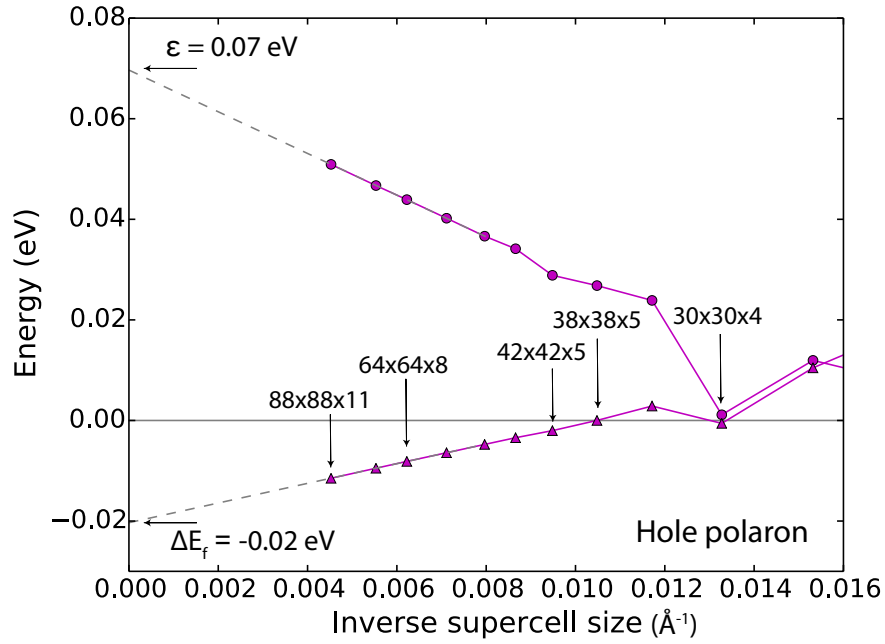


Figure 7.10: Polaron formation energies and eigenvalues for the hole polaron.

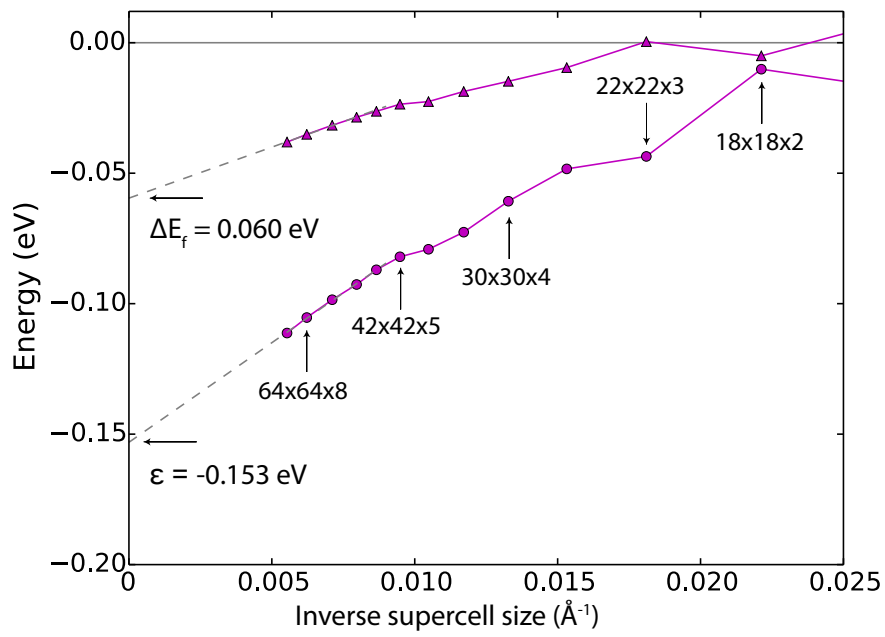


Figure 7.11: Polaron formation energies and eigenvalues for the electron polaron

supercell size of $22 \times 22 \times 3$ unit cells. In this case, the formation energy and eigenvalue extrapolated to the infinite supercell are $\Delta E_f = -0.02$ eV and $\varepsilon = 0.07$ eV (measured from the valence band maximum). Now we can apply the Mott criterion (6.1), to estimate at which supercell size the metal-insulator transition happens. Using $\kappa_{||} = 10.48$ and $m^*/m_e = 0.65$ for the hole, we obtain the critical density $n_c = 8.6 \times 10^{17} \text{ cm}^{-3}$. Taking the unit cell volume of the monolayer h-BN $\Omega = 736.89 \text{ Bohr}^3$, we obtain that the critical supercell size of 22^3 marks the MIT. We note that this estimate is very close to our calculation, which also confirms that the hole polaron is indeed a large one, and is well described by the Landau-Pekar model.

In Fig. 7.11 we present the formation energies and polaron eigenvalues in the same system but for the electron polaron. Our calculations find localised polarons starting at the supercell size of $14 \times 14 \times 2$ unit cells. In this case, the extrapolated formation energy and eigenvalue are $\Delta E_f = -0.06$ eV and $\varepsilon = 0.153$ eV (measured from the conduction band minimum), thus the electron polaron determined here is a large polaron. We find that the formation energy and polaron eigenvalue of the electron polaron are 200 % and 119 % larger than for the hole polaron, respectively. This is due to the additional coupling arising from the free electron states. Using $\kappa_{||} = 10.48$ and $m^*/m_e = 0.94$ from the electron, we obtain the critical density $n_c = 2.6 \times 10^{17} \text{ cm}^{-3}$. Taking the unit cell volume of the monolayer h-BN $\Omega = 736.89 \text{ Bohr}^3$, we obtain that the MIT corresponds to the critical supercell size of 15^3 , not too far from the hole polaron case yielded by 22^3 .

7.4.2 Convergence of polaron energies with vacuum size

In order to check that our calculated polaron energies are indeed converged at our chosen vacuum size of $20 \text{ \AA}$, we repeat the calculation of polaron formation energies and eigenvalues by systematically increasing the vacuum sizes from $10 \text{ \AA}$ to $20 \text{ \AA}$. Since we would like to compare the monolayer calculations to the bulk in the next

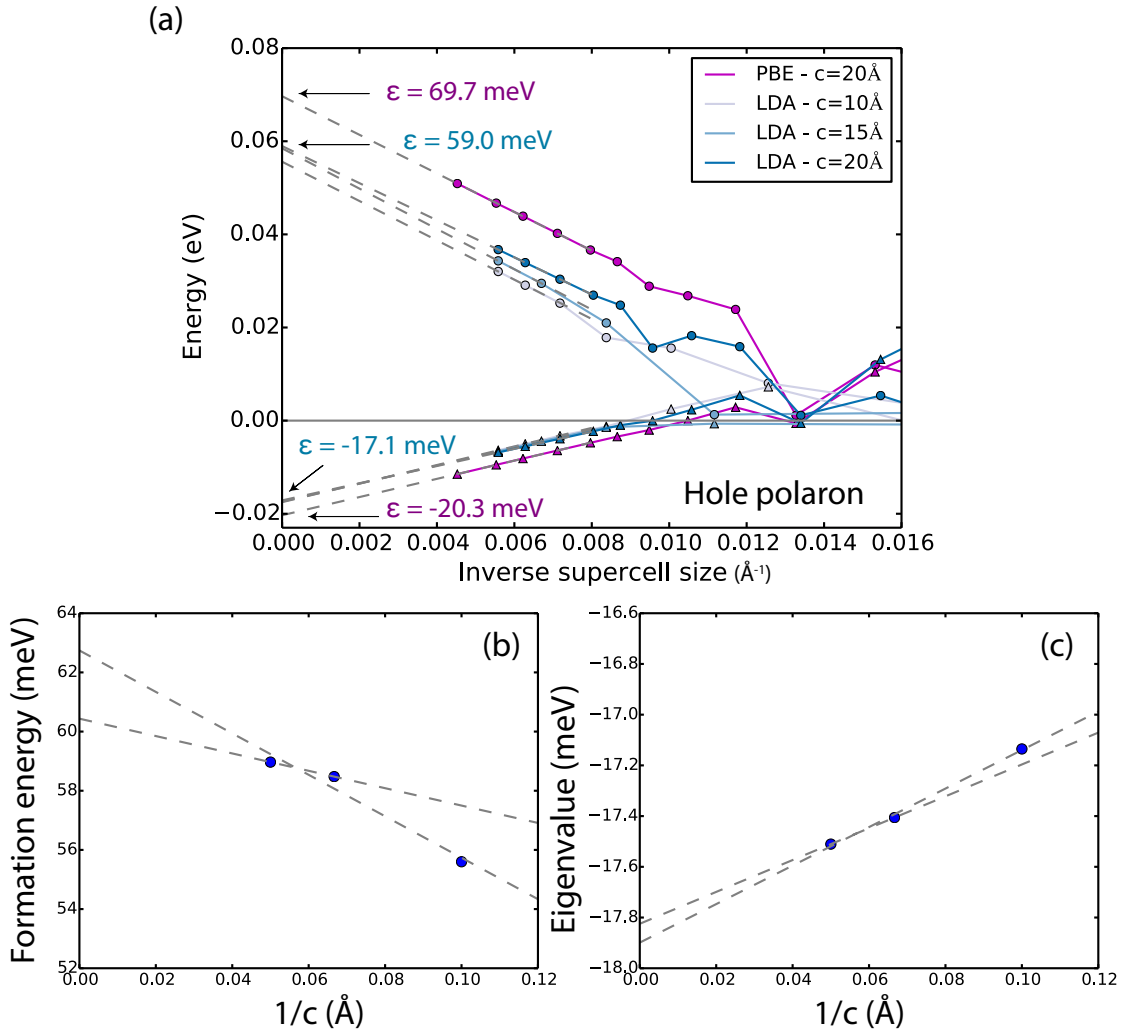


Figure 7.12: (a) Hole polaron energies with increasingly vacuum sizes from 10 $\text{\AA}$ to 20 $\text{\AA}$ in the *h*-BN monolayer. Previous calculations based on the PBE functional is shown in purple as a reference. Extrapolated polaron formation energies (b); and eigenvalues (c) from (a) against the inverse of lattice constant c . The two grey dash lines represent the best fit of all the data points (10-20 $\text{\AA}$) and the points (15-20 $\text{\AA}$), respectively.

section, LDA functional was employed for the repeated calculations of monolayers to ensure the consistency.

Figure 7.12(a) shows the calculated polaron energies with increasingly vacuum sizes from 10 Å to 20 Å in the h-BN monolayer. We confirm that both formation energies and eigenvalues extrapolated at infinite supercell size are converged at the vacuum size of 15 Å. In order to obtain the polaron energies correspond to an isolated monolayer, in principle one needs to push the polaron calculations by using a very large vacuum. We note that this can be very expensive to achieve. To this aim, instead we perform a further extrapolation on our energies with several vacuum sizes to estimate the energies in the monolayer limit. In Figs. 7.12(b)-(c) we present the extrapolated polaron energies against different vacuum sizes. Using these data, we determine the polaron formation energy and the eigenvalue in the monolayer limit are about 17.8–17.9 meV and 60.4–62.7 meV, respectively.

7.5 Comparisons between bulk and monolayer

To explore how the polaron energies change from monolayer to bulk, we repeat the polaron calculations but this time using bulk h-BN as shown in Fig. 7.1(a). Since the c-axis parameter is better described by the LDA functional, we consider the LDA functional for both calculations. In addition, we focus on the hole polaron which exhibits simple spectral decompositions.

Figure 7.13 shows both the formation energies and polaron eigenvalues for the h-BN in bulk and in monolayer, respectively. In the bulk calculations, we find localised polarons starting with a supercell of $24 \times 24 \times 8$ unit cells. The formation energy and eigenvalue extrapolated at infinite supercell size (black symbols) are $\Delta E_f = -13.6$ meV and $\varepsilon = 38.2$ meV, respectively. In the same figure we also compare to the monolayer calculations using the same LDA functional. Both the formation

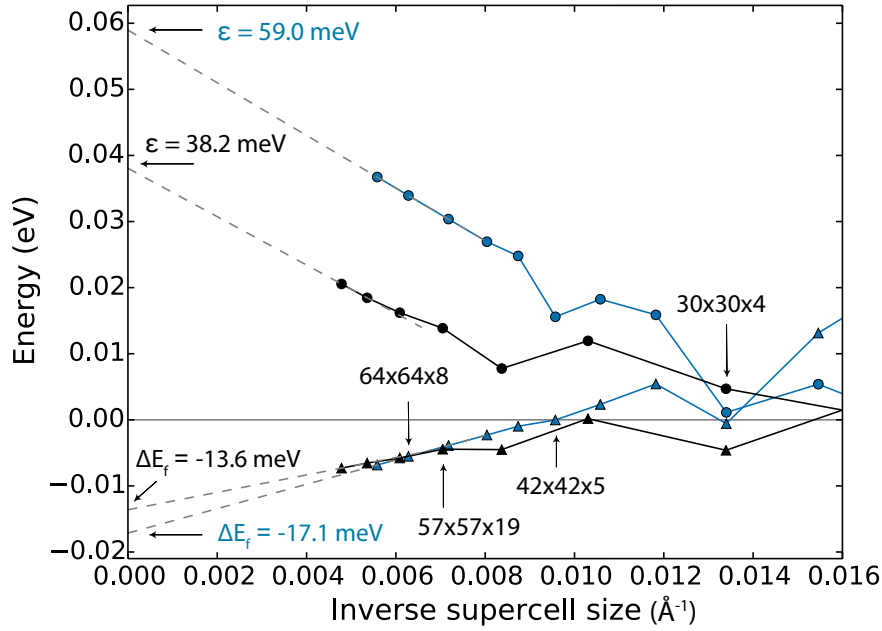


Figure 7.13: Comparisons of hole polaron formation energies and eigenvalues in h-BN between the bulk (black) and the monolayer (blue).

energies and eigenvalues exhibit a similar trend but different slopes as a function of L^{-1} , and the extrapolated energies at infinite supercell size (blue symbols) as $\Delta E_f = -17.1$ meV and $\varepsilon = 59.0$ meV, respectively.

The formation energy is large in the monolayer due to the reduced screening. This is similar to what is found in the study of excitons in the h-BN bulk and monolayer [246, 247]. Using $\kappa_{||} = 18.22$ and $m^*/m_e = 0.54$ from the hole, we obtain the critical density $n_c = 9.55 \times 10^{16} \text{ cm}^{-3}$. Taking the unit cell volume of the bulk h-BN $\Omega = 234.65 \text{ Bohr}^3$, we obtain the critical supercell size of 67^3 . The estimate of the critical size is few times larger than in our calculation, which implies that the polaron in bulk may not exhibit purely Fröhlich interactions but also other phonon modes may play a role. To understand this mechanism further, we analyse the spectral decomposition in the following paragraph.

Figure 7.14 shows the 3D hole polaron in bulk h-BN, obtained by solving the *ab initio* polaron equations at the grid size of $24 \times 24 \times 8$ points. The polaron wavefunc-

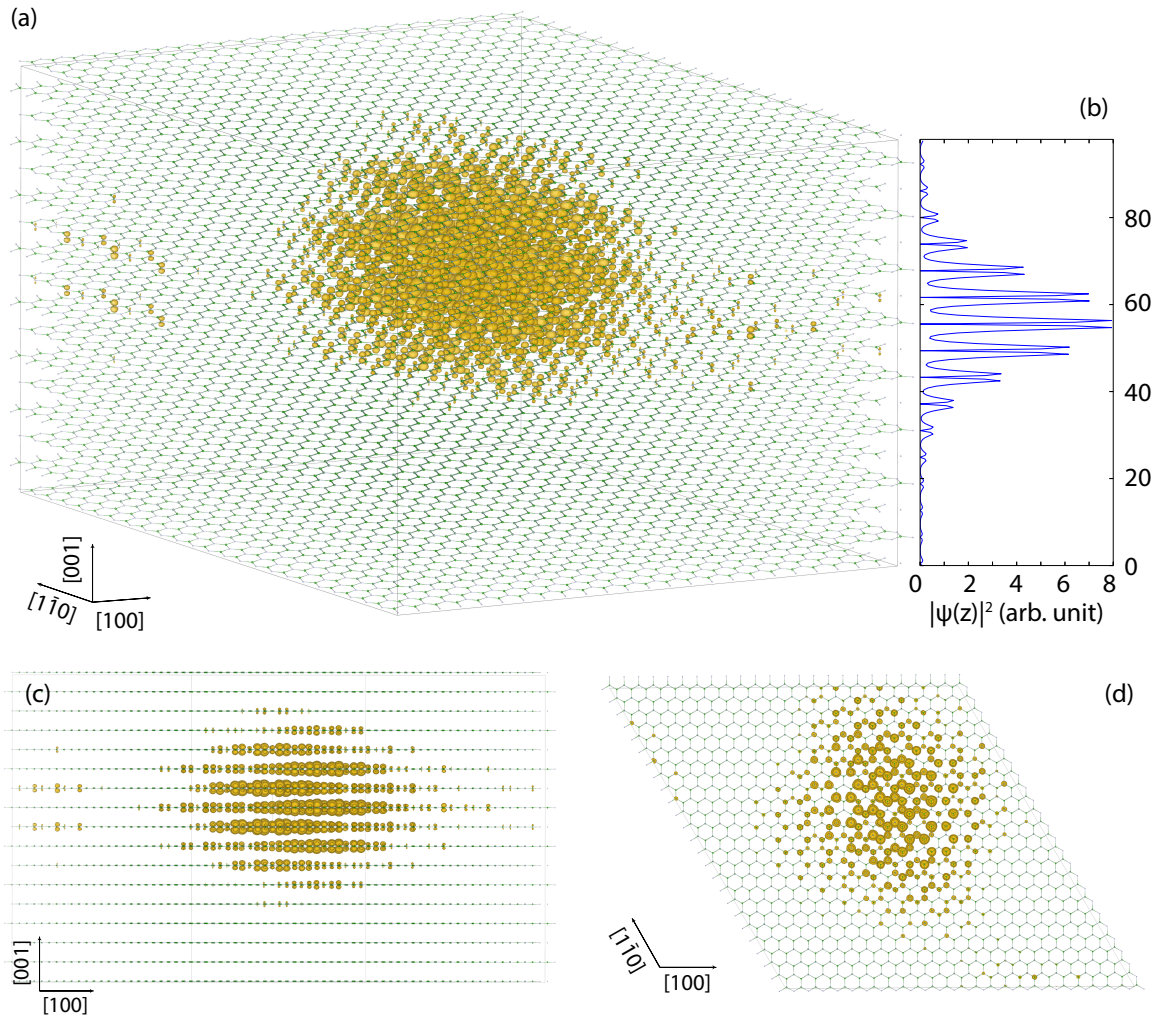


Figure 7.14: Isosurface plots of the polaron wavefunction ψ for the hole polaron in bulk h-BN, viewed from sides (a),(c); and from top (d) respectively. (b) The horizontal planar average of the polaron wavefunction against the z -axis in Bohr.

tion is calculated via Eq. (4.17). Figs. 7.14(a),(c) and (d) illustrate the hole polaron wavefunction as an isosurface viewed from different sides and from the top, respectively. Overall, the 3D hole polaron looks very similar to the 2D hole polaron shown in Fig. 7.4(a)–(c), where the polaron wavefunction is still comprised of $2p_z$ orbitals of nitrogen atoms, modulated by a localised envelop function. If we compare these plots with the 2D hole polaron shown in Fig. 7.9, we see that the polaron in bulk h-BN is not only bound to one of the layer, but also it extends over several unit cells before it decays. In Fig. 7.14(b) we show the 1D-profile of the polaron wavefunction along the out-of-plane direction. We determine that the out-of-plane width (full-width at half maximum) of this 3D hole polaron is of 13 Å, which is 9 times larger than for the 2D hole polaron in monolayer h-BN. This large width is also expected from the smaller formation energy in the bulk case. In Figs. 7.14(c)–(d) we see that the 3D hole polaron possesses a very symmetric localisation in both in-plane and out-of-plane directions. In particular, the alternating pattern of the polaron wavefunction along the [001] is induced by the alternative arrangement of the atoms in the setup of the AA' type stacking. Interestingly, such alternance has been previously observed in the exciton wavefunctions for few-layer h-BN [248].

Figure 7.15 shows the spectral decomposition of the polaron wavefunctions and atomic displacements in terms of the KS states and the vibrational normal modes. In Fig. 7.15(a) we visualise the electronic Fourier amplitudes $|A_{n\mathbf{k}}|^2$ by plotting it on top of the band structure for the hole polaron in bulk h-BN. The corresponding electronic density of states and spectral function $A^2(E)$ are shown in Fig. 7.15(b). We see that the main contribution to the large hole polaron comes from the high-lying valence states near the valence band maximum. We note that the valence band maxima (T_1) for the bulk h-BN is off from the symmetric points [249], which introduces extra complication on the couplings. For example, the states near the local maxima in the path M-K also contribute to the polaron formation, in addition to the principal

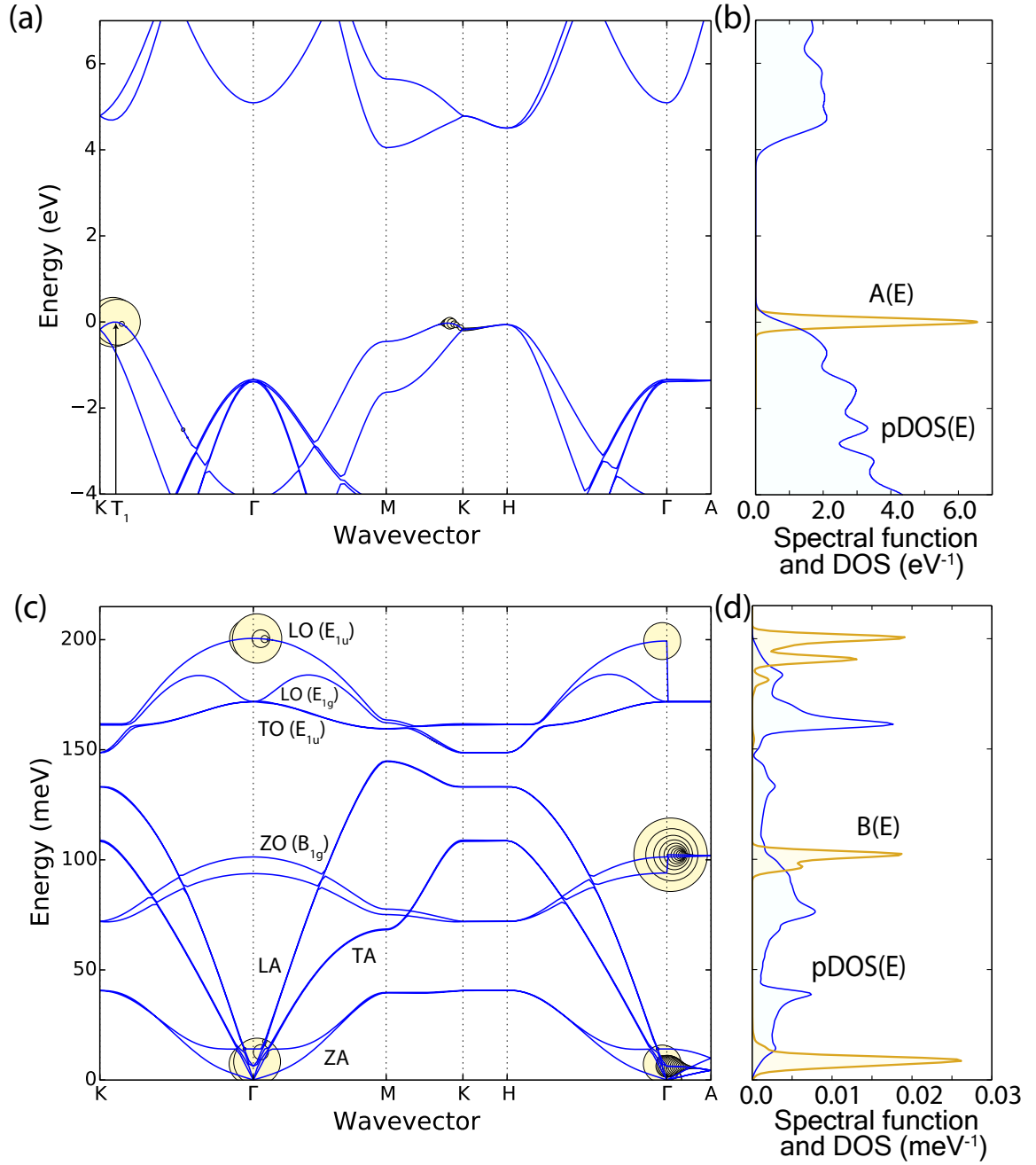


Figure 7.15: Spectral decomposition of the hole polaron in bulk h-BN. (a) Generalised Fourier amplitudes $A_{n\mathbf{k}}$ plotted on top of the band structure of bulk h-BN. The radius of each circle is proportional to $|A_{n\mathbf{k}}|^2$. The zero of the energy is aligned with the top of the valence bands. (b) Electronic density of states (blue, arb. unit) and spectral function $A^2(E)$ (yellow), aligned with the bands in (a). (c) Generalised Fourier amplitudes $B_{\mathbf{q}\nu}$ plotted on top of the phonon dispersion relations of h-BN monolayer. The radius of each circle is proportional to $|B_{\mathbf{q}\nu}|^2$. (d) Phonon density of states (blue, arb. unit) and spectral function $B^2(E)$, aligned with the dispersions in (c).

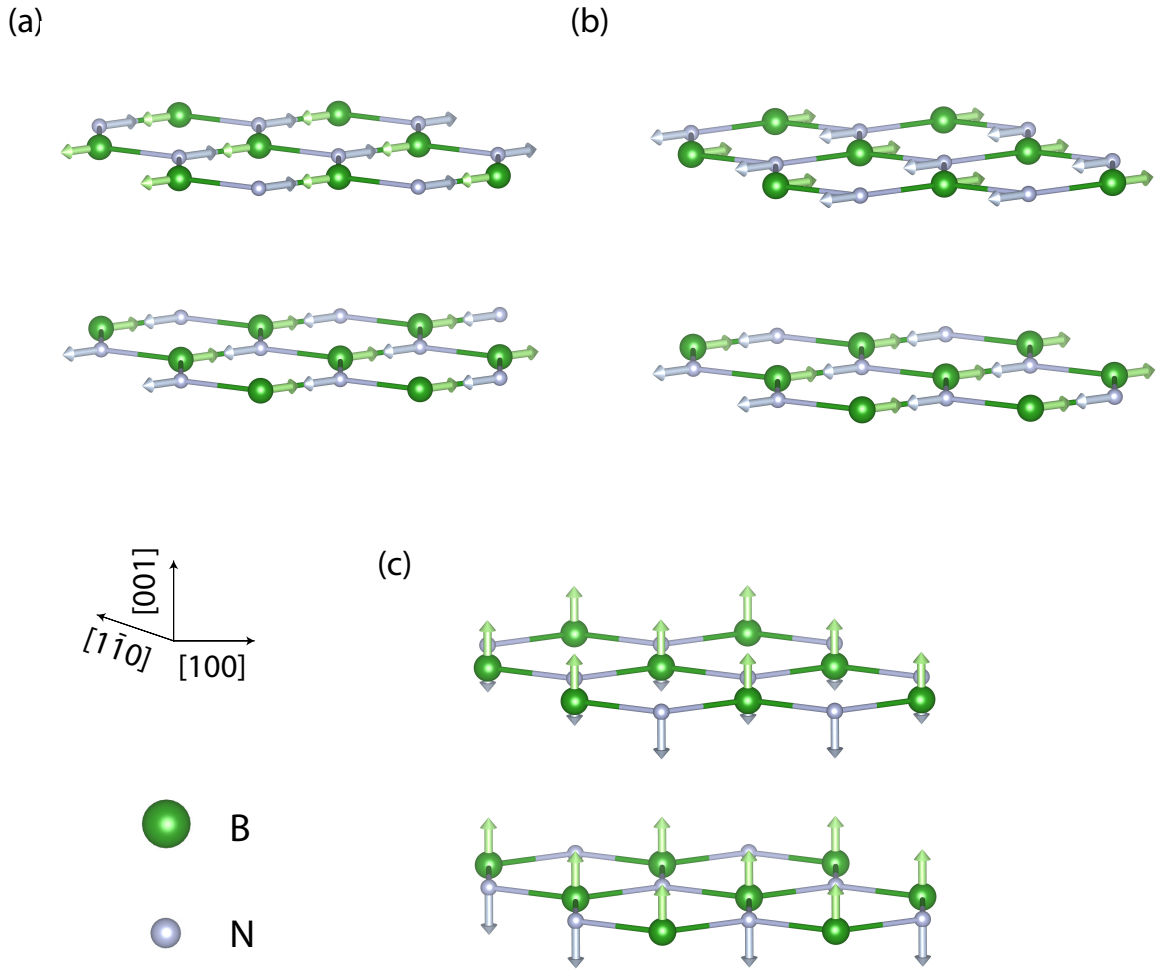


Figure 7.16: Ball-and-stick representation of (a) LO E_{1g} phonon; (b) LO E_{1u} phonon at $|\mathbf{q}| \rightarrow 0$ from the Γ -M path; and (c) optical phonons ($\nu = 8$) at $|\mathbf{q}| \rightarrow 0$ from the Γ -A path, showing for $3 \times 3 \times 1$ unit cells of the bulk h-BN.

contributions at T_1 .

In Fig. 7.15(c) we show the displacements Fourier amplitudes $|B_{\mathbf{q}\nu}|^2$, and the density of vibrational states and phonon spectral function $B^2(E)$ are presented in Fig. 7.15(d). We find that the polaron is dominated by the LO E_{1u} phonon at 201 meV [Fig. 7.16(a)], but also the ZO mode [Fig. 7.16(c)] and acoustic phonons have significant contributions. If we compare the phonon contributions to the case of the 2D hole polaron in monolayer as shown in Fig. 7.5(c), we see that the polaron in bulk h-BN is indeed not induced by the same LO mode with E_{1g} symmetry but a higher frequency one with E_{1u} symmetry. Fig. 7.16(a)–(b) illustrates the difference between these two modes. Since the LO E_{1u} mode represents the atoms in the same position but with different layer oscillating out-of-phase, this explains why the side view of the 3D hole polaron exhibits the alternating pattern as shown in Fig. 7.14(c). Another useful remark is that the ZO phonons are substantially excited for the bulk case in comparison to monolayer. This kind of coupling comes from the inhomogeneous, multilayer nature of the system, which cannot be described accurately by the Landau-Pekar model. This additional coupling indeed affects our estimate of the critical size in which the metal-insulator transition happens, based on Eq. (6.1). Again we can quantify the fractional contributions from these phonon modes by integrating the spectral function in Fig. 7.15(d). We find that the LO modes contribute 38 % of the total excited phonons, the acoustic modes contribute 36 % and the additional ZO phonons contribute the rest of 26 %.

7.6 Further work

Having reported calculations of 2D polarons by using the newly developed method in Chapters 4–5, it is desirable to understand the underlying mechanism by using models. For example, we identified that the long-wavelength LO phonons are dominant in h-

BN. One could devise out an effective analytical model to describe polarons in the monolayer limit, but resorting only to simple parameters such as in-plane effective mass, dielectric constants and phonon frequencies.

To understand the dependence of electron-phonon couplings on the screening effects, it would be desirable to establish a model for the monolayer which includes an inhomogeneous dielectric profile. From our point of view, these model studies are important in the sense that they will provide a first step towards the *ab initio* studies of polarons in more complicated heterostructure and interfacial systems [250].

A possible potential work can also be integrating the established polaron framework with the exciton theory to provide a complete picture of these many-body effects, although this might not be trivial.

Chapter 8

Conclusion

8.1 Summary

In this thesis we developed an *ab initio* theory and a computational methodology that enables calculations of self-trapped polarons in semiconductors and insulators, without resorting to the conventional supercell approach. Before formally applying the method to real materials, we established numerical tests and benchmarks for validating our implementation. The application starts from the investigation of polarons in two bulk crystals, a halide salt, LiF and a layered metal oxide, Li₂O₂. Furthermore, we applied our developed method for calculating two-dimensional (2D) polarons in h-BN from first-principle, in order to explore the dimensionality effect of the self-trapped state in layered materials.

We presented the theoretical formalism in Chapter 4, which employs electronic band structures, phonon dispersion relations, and electron-phonon matrix elements calculated in the crystal unit cell by density-functional theory and density-functional perturbation theory. In the theory we formulated the polaron problem as a variational minimisation of a self-interaction corrected DFT functional with respect to the polaron wavefunction, which led to a non-linear system of two coupled equations for the polaron wavefunction and the associated atomic displacements. We showed that this approach is mathematically similar to the classic Landau-Pekar polaron problem, but in our case the electron-phonon couplings are generalised to all phonon modes, including acoustic and optical, and both short- and long-range interactions,

are all considered in our formalism. We also established an explicit form for the self-interaction corrected DFT functional, which was used to cross validate our polaron formalism.

In Chapter 5 we outlined the computational setup used for these studied compounds, and explained how to solve the coupled polaron equations in practice. Furthermore, we proposed an efficient concurrent computing model allows for boosting self-consistent diagonalisations of the giant polaron matrix with capable power of modern high-performance computer, which was successfully tested on LiF using model parameters.

In Chapter 6 we applied this new approach to three cases, namely the large electron polaron in LiF, the small hole polaron in the same material, and the small electron polaron in Li_2O_2 . In the case of the large polaron we validated our calculation using the Landau-Pekar model; in the case of the small polaron we compared our results with explicit supercell calculations. We noted that our developed method describes correctly both large and small polarons, therefore this method reveals general validity across the length scales. We also introduced a spectral analysis of the polaron wavefunction and atomic displacements in order to quantify which electron bands, phonon modes, and electron-phonon couplings play the most important role in the formation of the polaron. This analysis allowed us to identify Fröhlich-type electron polarons in LiF, and Holstein-type polarons in LiF (holes) and Li_2O_2 (electrons). We anticipated that this type of analysis will be useful to devise model polaron Hamiltonian starting from realistic materials parameters computed from first principles.

In Chapter 7 we employed this new approach to investigate polarons in h-BN, we presented a first-principle calculations of polaron formation energies, eigenvalues, wavefunctions and spectral decompositions for both the hole and electron polarons in h-BN monolayer. Our calculations suggested that both polarons exhibit a long-range spatial extent, and thus should be classified as large polarons. In particular, we show

that our approach can lead to the formation of a 2D polaron, with the correct choice. Lastly, we demonstrated by calculations the 2D polaron in h-BN monolayer exhibits a stronger polaronic effect than the 3D one found in its bulk form.

8.2 Future work

We believe that the work introduced in this thesis provided a solid foundation for blueprinting predictive calculations of lattice polarons in real materials, and will possibly open up several potential directions for future developments.

One immediate development would be to explore excited polaron states beyond the ground state. This involves solving Eq. (4.21) for higher-lying electronic eigenstates instead of keeping only the ground state in the self-consistent calculations. The study of electronic excitations at fixed lattice distortion could be useful to understand the response of polarons to ultrafast optical excitations.

Another important work would be to go beyond the adiabatic approximations and the assumption of classical ions. Indeed, the main limitation of the present formalism is that we started from the DFT energy functional in Eq. (4.1). In this functional the electronic structure is described as a parametric function of classical ionic coordinates, therefore both DFT calculations of polarons and our formalism are both similar in spirit to the Landau-Pekar polaron model. In principle, this problem should be formulated by using a field-theoretic approach, as provided for example by the self-consistent Hedin-Baym equations for the coupled electron-phonon system [107]. While it may be possible to proceed along this direction, we speculate that it may be easier to begin with the present formulation, but upgrading the polaron theory by reintroducing non-adiabatic effects and quantum nuclear fluctuations from the outset. For example, the kinetic energy of the nuclei could be introduced into the Eq. (4.7), which allows us to express the problem in terms of the correlated

electron-ion wavefunction $\Psi(\mathbf{r}, \{\Delta\tau_{\kappa\alpha p}\})$:

$$E'_p[\Psi] = \int d\mathbf{r} d\{\Delta\tau_{\kappa\alpha p}\} \Psi^*(\mathbf{r}, \{\Delta\tau_{\kappa\alpha p}\}) \times \left[-\frac{1}{2} \sum_{\kappa\alpha p} \frac{\hbar^2}{2M_\kappa} \frac{\partial^2}{\partial \Delta\tau_{\kappa\alpha p}^2} + \frac{1}{2} \sum_{\substack{\kappa\alpha p \\ \kappa'\alpha'p'}} C_{\kappa\alpha p, \kappa'\alpha'p'}^0 \Delta\tau_{\kappa\alpha p} \Delta\tau_{\kappa'\alpha'p'} + \hat{H}_{\text{KS}}^0 + \sum_{\kappa\alpha p} \frac{\partial V_{\text{KS}}^0}{\partial \tau_{\kappa\alpha p}} \Delta\tau_{\kappa\alpha p} \right] \Psi(\mathbf{r}, \{\Delta\tau_{\kappa\alpha p}\}). \quad (8.1)$$

This formulation takes an advantage by focusing only on a single electron interacting with a phonon bath, because the electron-electron interaction is already captured by the KS-DFT Hamiltonian. Equation (8.1) can be reformulated in terms of phonon ladder operators and electron-phonon matrix elements, following steps similar to Sec. 4.1.4. Using the same notation as in Ref. [107], the Hamiltonian inside the square brackets becomes (apart from a constant provided by the zero-point energy):

$$\begin{aligned} \hat{H}'_p = & \sum_{n\mathbf{k}} \varepsilon_{n\mathbf{k}} |n\mathbf{k}\rangle \langle n\mathbf{k}| + \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu} \hat{a}_{\mathbf{q}\nu}^\dagger \hat{a}_{\mathbf{q}\nu} \\ & + N_p^{-\frac{1}{2}} \sum_{mn\nu, \mathbf{k}, \mathbf{q}} g_{mn\nu}(\mathbf{k}, \mathbf{q}) (\hat{a}_{\mathbf{q}\nu} + \hat{a}_{-\mathbf{q}\nu}^\dagger) |m\mathbf{k} + \mathbf{q}\rangle \langle n\mathbf{k}|, \end{aligned} \quad (8.2)$$

where the summations over bands are restricted to conduction or to valence states for electron or hole polarons, respectively. In this equation we do not employ the usual electron field operator because we have only one electron, therefore second quantisation only applies to phonons.

Equation (8.2) can be considered as the *ab initio* counterpart of the Fröhlich electron-phonon Hamiltonian.[70] In fact the standard Fröhlich Hamiltonian is recovered by retaining only one parabolic band and considering only one LO phonon branch. This equation suggests a potential way to connect the present approach with many-body calculations of model polaron Hamiltonians: (i) for a given system we could identify the most important electronic bands, phonons, and electron-phonon couplings using the spectral decompositions of $A_{n\mathbf{k}}$ and $B_{\mathbf{q}\nu}$; (ii) we could accord-

ingly truncate Eq. (8.2) to retain only the most important contributions; (iii) at this point we could employ advanced many-body techniques for polaron Hamiltonians, for example the diagrammatic Monte Carlo (DMC) approach [20]. In this way one could envision complete first-principles calculations of polarons, where the atomistic details and predictive power of DFT approaches are combined with the wealth of many-body physics and numerical techniques of Monte Carlo.

In conclusion, we would like to stress that investigating polarons from first-principle is not only of a theoretical interest, but also equivalently important for envisioning meaningful applications in which polarons may play a unique role. We hope that the present work will serve as a basis for enabling predictive calculations of polarons in more complex materials, and it will ultimately help the *ab initio* design of many optoelectronic materials that host polarons for a broad range of applications.

Appendix A

Truncation of the long-range Fröhlich matrix element

In this appendix, we introduce a spherical Coulomb truncation to speed up the convergence of polaron calculations. We derive a closed-form expression for the truncated Fröhlich matrix element. First, considering the Coulomb potential generated by a point charge $+Q$ immersed by anisotropic medium as in Ref. [251]:

$$\phi(\mathbf{r}; \boldsymbol{\tau}) = \frac{Q}{4\pi\epsilon_0\sqrt{\det \underline{\underline{\epsilon}}_r}} \frac{1}{\sqrt{(\mathbf{r} - \boldsymbol{\tau}) \cdot \underline{\underline{\epsilon}}_r^{-1} \cdot (\mathbf{r} - \boldsymbol{\tau})}}.$$

The dipole contribution may be written as:

$$\phi_{dip}(\mathbf{r}; \boldsymbol{\tau}) = \sum_{\alpha} \frac{\partial \phi}{\partial \tau_{\alpha}|_{\boldsymbol{\tau}=0}} \tau_{\alpha} = \frac{1}{4\pi\epsilon_0\sqrt{\det \underline{\underline{\epsilon}}_r}} \frac{\mathbf{p} \cdot \underline{\underline{\epsilon}}_r^{-1} \cdot \mathbf{r}}{[\mathbf{r} \cdot \underline{\underline{\epsilon}}_r^{-1} \cdot \mathbf{r}]^{\frac{3}{2}}}. \quad (\text{A.1})$$

The Fourier coefficient of a periodic dipole potential with a spherical cutoff R_c maybe obtained by $\phi_{\mathbf{q}+\mathbf{G}}^{dip} = \frac{1}{N_p\Omega} \int_{SC} \phi^{dip}(\mathbf{r}) \theta(R_c - |\mathbf{r}|) e^{-i(\mathbf{q}+\mathbf{G}) \cdot \mathbf{r}} d\mathbf{r}$:

$$\phi_{\mathbf{q}}^{dip} = \frac{1}{N_p\Omega} \frac{1}{4\pi\epsilon_0\sqrt{\det \underline{\underline{\epsilon}}}} \mathbf{p} \cdot \underline{\underline{\epsilon}}_r^{-1} \cdot i \underline{\underline{\nabla}}_{\mathbf{q}} \underbrace{\int_{SC} d\mathbf{r} \frac{e^{-i\mathbf{q} \cdot \mathbf{r}} \theta(R_c - |\mathbf{r}|)}{[\mathbf{r} \cdot \underline{\underline{\epsilon}}_r^{-1} \cdot \mathbf{r}]^{\frac{3}{2}}}}_J. \quad (\text{A.2})$$

Now we are in the position of computing the integral J . This part involves a particular choice of coordinate transformation. First, we diagonalise the full dielectric tensor by

an orthogonal matrix $\underline{\underline{C}}$:

$$\underline{\underline{\epsilon}}_r = \underline{\underline{C}} \cdot \underline{\underline{\epsilon}}'_r \cdot \underline{\underline{C}}^T, \quad (\text{A.3})$$

where $\underline{\underline{\epsilon}}'_r = \text{diag}(\epsilon_1, \epsilon_2, \epsilon_3)$ stores the three eigenvalues of $\underline{\underline{\epsilon}}_r$. Second, we define a scaling matrix $\underline{\underline{D}} = \text{diag}(\sqrt{\epsilon_1}, \sqrt{\epsilon_2}, \sqrt{\epsilon_3})$ and such that:

$$\begin{aligned} \underline{\underline{\epsilon}}_r &= \underline{\underline{C}} \cdot \underline{\underline{D}}^2 \cdot \underline{\underline{C}}^T, \\ \underline{\underline{\epsilon}}_r^{-1} &= \underline{\underline{C}} \cdot \underline{\underline{D}}^{-2} \cdot \underline{\underline{C}}^T. \end{aligned} \quad (\text{A.4})$$

While evaluating the J integral, we seek a transformation $\mathbf{r} = \underline{\underline{C}} \cdot \underline{\underline{D}} \cdot \mathbf{u}$ and the volume element $d^3\mathbf{r} = \det(\underline{\underline{C}} \cdot \underline{\underline{D}}) d^3\mathbf{u} = \sqrt{\det \underline{\underline{\epsilon}}_r} d^3\mathbf{u}$:

$$J = \sqrt{\det \underline{\underline{\epsilon}}_r} \int d^3\mathbf{u} \frac{e^{-i\mathbf{q}^T \cdot \underline{\underline{C}} \cdot \underline{\underline{D}} \cdot \mathbf{u}} \theta(R_c - \sqrt{\mathbf{u} \cdot \underline{\underline{D}}^2 \cdot \mathbf{u}})}{[\mathbf{r}^T \cdot \underline{\underline{C}} \cdot (\underline{\underline{D}}^{-1})^T \cdot \underbrace{\underline{\underline{D}}^{-1} \cdot \underline{\underline{C}}^T \cdot \mathbf{r}}_{\mathbf{u}}]^{\frac{3}{2}}}. \quad (\text{A.5})$$

Note that this integral may have to be evaluated using ellipsoidal polar coordinates since the integration domain is defined within $(\frac{u_x}{R_c/\sqrt{\epsilon_1}})^2 + (\frac{u_y}{R_c/\sqrt{\epsilon_2}})^2 + (\frac{u_z}{R_c/\sqrt{\epsilon_3}})^2 < 1$. To simply the problem, we may consider a weak polarised system for which it is valid to approximate the ellipsoidal volume by an effective sphere (i.e. $\frac{4\pi R_c^3}{3} = \frac{4\pi}{3} \frac{R_c^3}{\sqrt{\epsilon_1 \epsilon_2 \epsilon_3}}$). But we do not hope to lose too much accuracy in an extreme case. More generally, it is reasonable to consider the radius of the effective sphere that covers up to the largest axis of the ellipsoid although this may introduce a slightly extra cost. Hence

$$R'_c = \frac{R_c}{\sqrt{\min(\epsilon_1, \epsilon_2, \epsilon_3)}}$$

$$\begin{aligned} J &= \sqrt{\det \underline{\underline{\epsilon}}_r} \int_{|\mathbf{u}| < R'_c} d^3\mathbf{u} \frac{e^{-i(\underline{\underline{D}} \cdot \underline{\underline{C}}^T \cdot \mathbf{q})^T \cdot \mathbf{u}}}{|\mathbf{u}|^3} \\ &= \sqrt{\det \underline{\underline{\epsilon}}_r} \frac{4\pi}{|\mathbf{q}'|} \int_0^{R'_c} \frac{\sin(|\mathbf{q}'|u)}{u^2} du \quad \text{with} \quad \mathbf{q}' \equiv \underline{\underline{D}} \cdot \underline{\underline{C}}^T \cdot \mathbf{q}. \end{aligned} \quad (\text{A.6})$$

Since 1D integral in J is improper at $u = 0$, one has to also take into account the curl operator $\nabla_{\mathbf{q}}$ together with J in order to regularise the integral:

$$\underline{\nabla}_{\mathbf{q}} J = \sqrt{\det \underline{\epsilon}_r} \left[\underline{\nabla}_{\mathbf{q}} \frac{4\pi}{|\mathbf{q}'|} \int_0^{R'_c} \frac{\sin(|\mathbf{q}'|u)}{u^2} du \right]. \quad (\text{A.7})$$

Remark: a small subtle point here is the partial derivative $\nabla_{\mathbf{q}}$ is operated on a function with transformed coordinate $\mathbf{q}'$, one has to recall some tensor algebra. Considering a contravariant vector $q'^i = D_j^i (C^T)^j_k q^k$ and differentiate a scalar function f by:

$$\frac{\partial f}{\partial q^i} = \frac{\partial q'^a}{\partial q^i} \cdot \frac{\partial f}{\partial q'^a}.$$

We note that:

$$\frac{\partial q'^a}{\partial q^i} = D_j^a (C^T)^j_k \delta_i^k = D_j^a (C^T)^j_i.$$

Hence,

$$\begin{aligned} \frac{\partial}{\partial q^i} &= D_j^a (C^T)^j_i \frac{\partial}{\partial q'^a} \\ &= (\underline{D} \cdot \underline{C}^T)^a_i \frac{\partial}{\partial q'^a}. \end{aligned}$$

This shows that partial derivative is a covector with rank (0,1) and we can transpose the above relation to obtain the corresponding contravariant curl operator:

$$\begin{aligned} \underline{\nabla}_{\mathbf{q}} &= \underline{C} \cdot \underline{D}^T \cdot \underline{\nabla}_{\mathbf{q}'} \\ &= \underline{C} \cdot \underline{D} \cdot \underline{\nabla}_{\mathbf{q}'}. \end{aligned} \quad (\text{A.8})$$

Now substituting this result into Eq. (A.7):

$$\begin{aligned} \underline{\nabla}_{\mathbf{q}} J &= \left(\sqrt{\det \underline{\epsilon}_r} \right) \underline{C} \cdot \underline{D} \cdot \underline{\nabla}_{\mathbf{q}'} \left[\frac{4\pi}{|\mathbf{q}'|} \int_0^{R'_c} \frac{\sin(|\mathbf{q}'|u)}{u^2} du \right] \\ &= -4\pi \left(\sqrt{\det \underline{\epsilon}_r} \right) \underline{C} \cdot \underline{D} \cdot \frac{\hat{\mathbf{q}}'}{|\mathbf{q}'|} \left[1 - \frac{\sin |\mathbf{q}'| R'_c}{|\mathbf{q}'| R'_c} \right]. \end{aligned} \quad (\text{A.9})$$

Replacing this result into Eq. (A.2):

$$\begin{aligned}\phi_{\mathbf{q}}^{dip} &= \frac{1}{N_p \Omega} \frac{1}{4\pi\epsilon_0 \sqrt{\det \underline{\underline{\epsilon_r}}}} \mathbf{p} \cdot \underline{\underline{\epsilon_r}}^{-1} \cdot i \left[-4\pi \left(\sqrt{\det \underline{\underline{\epsilon_r}}} \right) \underline{\underline{C}} \cdot \underline{\underline{D}} \cdot \frac{\mathbf{q}'}{|\mathbf{q}'|^2} \left[1 - \frac{\sin |\mathbf{q}'| R'_c}{|\mathbf{q}'| R'_c} \right] \right] \\ &= -\frac{i}{N_p \Omega \epsilon_0} \frac{\mathbf{p} \cdot \mathbf{q}}{\mathbf{q} \cdot \underline{\underline{\epsilon_r}} \cdot \mathbf{q}} \left[1 - \frac{\sin(R'_c \sqrt{\mathbf{q} \cdot \underline{\underline{\epsilon_r}} \cdot \mathbf{q}})}{R'_c \sqrt{\mathbf{q} \cdot \underline{\underline{\epsilon_r}} \cdot \mathbf{q}}} \right].\end{aligned}\tag{A.10}$$

At this point, we re-arrived at the Fourier series of a periodic dipole potential with a cutoff R'_c :

$$\begin{aligned}\phi^{dip}(\mathbf{r}; \mathbf{p}) &= -\frac{i}{N_p \Omega \epsilon_0} \sum_{\mathbf{q}, \mathbf{G} \neq -\mathbf{q}} \frac{\mathbf{p} \cdot (\mathbf{q} + \mathbf{G})}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon_r}} \cdot (\mathbf{q} + \mathbf{G})} \\ &\quad \times \left[1 - \frac{\sin(R'_c \sqrt{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon_r}} \cdot (\mathbf{q} + \mathbf{G})})}{R'_c \sqrt{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon_r}} \cdot (\mathbf{q} + \mathbf{G})}} \right] e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}},\end{aligned}\tag{A.11}$$

where R'_c is defined as

$$R'_c = \frac{R_c}{\sqrt{\min(\epsilon_1, \epsilon_2, \epsilon_3)}}.\tag{A.12}$$

The dipole potential in $\mathbf{q}$ -space and the electron-phonon matrix element are therefore:

$$\begin{aligned}\Delta_{\mathbf{q}\nu} V^L(\mathbf{r}) &= i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa'} \left(\frac{\hbar}{2M_{\kappa'} \omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \\ &\quad \times \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{Z}}_{\kappa'}^* \cdot \underline{\underline{e}}_{\kappa'\nu}(\mathbf{q}) [1 - \text{sinc}(R'_c \sqrt{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon_r}} \cdot (\mathbf{q} + \mathbf{G})})]}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon_r}} \cdot (\mathbf{q} + \mathbf{G})} e^{i(\mathbf{q} + \mathbf{G}) \cdot (\mathbf{r} - \boldsymbol{\tau}_{\kappa'})},\end{aligned}\tag{A.13}$$

$$\begin{aligned}
g_{mn\nu}^{L,T}(\mathbf{k}, \mathbf{q}) = & i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa'} \left(\frac{\hbar}{2M_{\kappa'}\omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \\
& \times \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{Z}}_{\kappa'}^* \cdot \underline{\underline{e}}_{\kappa'\nu}(\mathbf{q}) [1 - \text{sinc}(R'_c \sqrt{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon}}_r \cdot (\mathbf{q} + \mathbf{G})})]}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon}}_r \cdot (\mathbf{q} + \mathbf{G})} \\
& \times \langle \psi_{m\mathbf{k}+\mathbf{q}} | e^{i(\mathbf{q}+\mathbf{G}) \cdot (\mathbf{r}-\boldsymbol{\tau}_{\kappa'})} | \psi_{n\mathbf{k}} \rangle . \quad (\text{A.14})
\end{aligned}$$

This expression has been used in Chapter 5, with the comparison between the usual Fröhlich matrix element and the truncated one presenting in Fig. 5.7(b).

Appendix B

Average phonon number

In this section, we derive an analytical expression for calculating the average phonon number in a strong-coupling polaron system. Considering the second quantised formalism in Ref. [15], the average phonon number is given by:

$$\langle \hat{N} \rangle = \sum_{\mathbf{q}\nu} |c_{\mathbf{q}\nu}|^2, \quad (\text{B.1})$$

where $c_{\mathbf{q}\nu}$ is a c-number from the linear Bogoliubov transformation. It is possible to relate the polaron ground state wavefunction to this c-number via [15]:

$$c_{\mathbf{q}\nu} = -\frac{1}{\sqrt{N_p} \hbar \omega_{\mathbf{q}\nu}} \int d\mathbf{r}' |\Psi(\mathbf{r}')|^2 [\Delta_{\mathbf{q}\nu} V^{\text{KS}}(\mathbf{r}')]^*. \quad (\text{B.2})$$

Here $\Delta_{\mathbf{q}\nu} V^{\text{KS}}(\mathbf{r}')$ denotes the perturbed KS potential defined in Chapter 3. In the most general case, the long range contribution can be expressed in terms of *ab initio* variables, namely Born effective charge, dielectric tensor and phonon eigenvectors as described in Eq. (A.13):

$$\Delta_{\mathbf{q}\nu} V^L(\mathbf{r}) = i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \sum_{\kappa'} \left(\frac{\hbar}{2M_{\kappa'} \omega_{\mathbf{q}\nu}} \right)^{\frac{1}{2}} \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{Z}}_{\kappa'}^* \cdot \underline{\underline{e}}_{\kappa'\nu}(\mathbf{q}) e^{i(\mathbf{q}+\mathbf{G}) \cdot (\mathbf{r}-\boldsymbol{\tau}_{\kappa'})}}{(\mathbf{q} + \mathbf{G}) \cdot \underline{\underline{\epsilon}}^\infty \cdot (\mathbf{q} + \mathbf{G})}. \quad (\text{B.3})$$

In order to relate this expression to the isotropic Landau-Pekar model, one has to consider only longitudinal optical mode and the continuum limit, i.e. the unit cell size effectively goes to zero, the type of atomic label κ and phonon mode ν are omitted.

The expression now reads

$$\Delta_{\mathbf{q}} V^{\text{LP}}(\mathbf{r}) = i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \left(\frac{\hbar}{2M\omega_{\text{LO}}} \right)^{\frac{1}{2}} \frac{Z^*}{\epsilon^\infty} e^{i\mathbf{q}\cdot\mathbf{r}} \mathbf{e}(\mathbf{q}) \cdot \sum_{\mathbf{G} \neq -\mathbf{q}} \frac{(\mathbf{q} + \mathbf{G})}{|\mathbf{q} + \mathbf{G}|^2} e^{i\mathbf{G}\cdot\mathbf{r}}. \quad (\text{B.4})$$

As in the continuum limit the first Brillouin zone becomes infinite, the discrete sum $\mathbf{G}$ should naturally pick up the primary one $\mathbf{G} = \mathbf{0}$. Hence,

$$\Delta_{\mathbf{q}} V^{\text{LP}}(\mathbf{r}) = i \frac{4\pi}{\Omega} \frac{e^2}{4\pi\epsilon_0} \left(\frac{\hbar}{2M\omega_{\text{LO}}} \right)^{\frac{1}{2}} \frac{Z^*}{\epsilon^\infty} \frac{e^{i\mathbf{q}\cdot\mathbf{r}}}{|\mathbf{q}|}. \quad (\text{B.5})$$

Now we express the Born effective charge using high / low frequency dielectric constants $\epsilon_\infty/\epsilon_0$ and phonon frequency $\hbar\omega_{\text{LO}}$,

$$\frac{Z^*}{\epsilon^\infty} = \left[\frac{\epsilon_0 \Omega M \omega_{\text{LO}}^2}{e^2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) \right]^{\frac{1}{2}}. \quad (\text{B.6})$$

Therefore, the perturbed potential now reads

$$\begin{aligned} \Delta_{\mathbf{q}} V^{\text{LP}}(\mathbf{r}) &= \frac{i}{|\mathbf{q}|} \left[\frac{e^2}{4\pi\epsilon_0} \left(\frac{4\pi}{\Omega} \right) \frac{\hbar\omega_{\text{LO}}}{2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) \right]^{\frac{1}{2}} e^{i\mathbf{q}\cdot\mathbf{r}}. \\ &\equiv \frac{i}{|\mathbf{q}|} F e^{i\mathbf{q}\cdot\mathbf{r}}. \end{aligned} \quad (\text{B.7})$$

Inserting this into Eq. (B.2), the phonon c-number with Fröhlich type interaction becomes

$$c_{\mathbf{q}} = \frac{i}{\sqrt{N_p} \hbar \omega_{\text{LO}}} \frac{F}{|\mathbf{q}|} \int d\mathbf{r}' |\Psi(\mathbf{r}')|^2 e^{i\mathbf{q}\cdot\mathbf{r}}. \quad (\text{B.8})$$

Combining Eqs. (B.1) and (B.8), the average phonon number is expressed as

$$\begin{aligned} \langle \hat{N} \rangle &= \frac{\Omega F^2}{\hbar^2 \omega_{\text{LO}}^2} \frac{1}{(2\pi)^3} \int d\mathbf{r} d\mathbf{r}' |\Psi(\mathbf{r})|^2 |\Psi(\mathbf{r}')|^2 \int_{\infty^3} \frac{d\mathbf{q}}{|\mathbf{q}|^2} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \frac{\Omega F^2}{\hbar^2 \omega_{\text{LO}}^2} \frac{1}{4\pi} \int d\mathbf{r} d\mathbf{r}' \frac{|\Psi(\mathbf{r})|^2 |\Psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \end{aligned} \quad (\text{B.9})$$

By taking the variational polaron wavefunction Eqs. (5.4)–(5.6), the Coulomb integral can be evaluated as:

$$\int d\mathbf{r}d\mathbf{r}' \frac{|\Psi(\mathbf{r})|^2 |\Psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} = 0.2183 \frac{m^*}{\kappa} \left[\frac{m_e}{\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right) \right]. \quad (\text{B.10})$$

Replacing this result into Eq. (B.9), the phonon average number is therefore:

$$\begin{aligned} \langle \hat{N} \rangle &= 0.1092 \frac{m^*}{\kappa^2} \frac{1}{\hbar\omega_{LO}} m_e \left(\frac{e^2}{4\pi\epsilon_0\hbar} \right)^2 \\ &= 0.2183\alpha^2. \end{aligned} \quad (\text{B.11})$$

As anticipated in the strong-coupling limit, this result is different to the one derived from the perturbation theory which is given by [106]:

$$\langle \hat{N} \rangle = \frac{\alpha}{2}. \quad (\text{B.12})$$

We note that Eq. (B.11) should be compared to the measure of phonon number by integrating the phonon spectral function $B^2(E)$ as presented in Ref. [158]. For example, it was found that the electron polaron in LiF involves the excitation of 5 LO phonons. By taking $\alpha = 4.92$ in Sec. 6.1.1 we determine the average number of phonon is of 5.3 from Eq. (B.11), in contrast perturbation theory only gives 2.5 phonons which is clearly underestimated. To understand why the integral of the phonon spectral function corresponds to the number of excited phonons in the system, one needs to relate Eqs. (B.1)–(B.2) to show:

$$\langle \hat{N} \rangle = \frac{1}{N_p} \sum_{\mathbf{q}\nu} |B_{\mathbf{q}\nu}|^2, \quad (\text{B.13})$$

where $B_{\mathbf{q}\nu}$ is the phonon Fourier amplitude defined in Eq. (4.22). This indeed can be expressed as an integral with the phonon spectral function $B^2(E)$:

$$\langle \hat{N} \rangle = \int_0^\infty B^2(E) dE, \quad (\text{B.14})$$

where $B^2(E)$ is defined in Eq. (4.32).

Bibliography

- [1] G. Lindemann, R. Lassnig, W. Seidenbusch, and E. Gornik, “Cyclotron resonance study of polarons in GaAs,” *Phys. Rev. B* **28**, 4693 (1983).
- [2] R. Popp and R. Murray, “Diffusion of the V_k -polaron in alkali halides: Experiments in NaI and RbI,” *J. Phys. Chem. Solids* **33**, 601 (1972).
- [3] P. M. Chaikin, A. F. Garito, and A. J. Heeger, “Excitonic polarons in molecular crystals,” *Phys. Rev. B* **5**, 4966 (1972).
- [4] D. Emin, “On the existence of free and self-trapped carriers in insulators: an abrupt temperature-dependent conductivity transition,” *Adv. Phys.* **22**, 57 (1973).
- [5] M. Umehara, “Importance of the site-off-diagonal electron-phonon interaction on the self-trapped small magnetic polaron,” *Phys. Rev. B* **27**, 5669 (1983).
- [6] A. Stoneham, J. Gavartin, A. Shluger, A. Kimmel, D. M. Ramo, H. Rønnow, G. Aepli, and C. Renner, “Trapping, self-trapping and the polaron family,” *J. Phys. Condens. Matter* **19**, 255208 (2007).
- [7] L. D. Landau, “Electron motion in crystal lattices,” *Phys. Z. Sowjet.* **3**, 664 (1933).
- [8] W. Känzig, “Electron spin resonance of V_1 -centers,” *Phys. Rev.* **99**, 1890 (1955).
- [9] T. D. Lee, F. E. Low, and D. Pines, “The motion of slow electrons in a polar crystal,” *Phys. Rev.* **90**, 297 (1953).
- [10] S. Pekar, “Autolocalization of the electron in a dielectric inertially polarizing medium,” *J. Exp. Theor. Phys.* **16**, 335 (1946).

-
- [11] S. Moser *et al.*, “Tunable polaronic conduction in anatase TiO_2 ,” *Phys. Rev. Lett.* **110**, 196403 (2013).
 - [12] H. Fröhlich, H. Pelzer, and S. Zienau, “Xx. properties of slow electrons in polar materials,” *Lond. Edinb. Dubl. Phil. Mag. J. Sci.* **41**, 221 (1950).
 - [13] R. P. Feynman, “Slow electrons in a polar crystal,” *Phys. Rev.* **97**, 660 (1955).
 - [14] T. Holstein, “Studies of polaron motion: Part I. the molecular-crystal model,” *Ann. Phys.* **8**, 325 (1959).
 - [15] N. Bogolubov, *Polaron Theory: model problems* (CRC Press, Amsterdam, 2014).
 - [16] E. P. Gross, “Small oscillation theory of the interaction of a particle and scalar field,” *Phys. Rev.* **100**, 1571 (1955).
 - [17] J. T. Marshall and L. R. Mills, “Second-order correction to Feynman’s path-integral calculation of the polaron self-energy,” *Phys. Rev. B* **2**, 3143 (1970).
 - [18] J. M. Luttinger and C.-Y. Lu, “Generalized path-integral formalism of the polaron problem and its second-order semi-invariant correction to the ground-state energy,” *Phys. Rev. B* **21**, 4251 (1980).
 - [19] A. L. Kholodenko and K. F. Freed, “Direct path-integral treatment of the polaron problem,” *Phys. Rev. B* **27**, 4586 (1983).
 - [20] N. V. Prokof’ev and B. V. Svistunov, “Polaron problem by diagrammatic quantum Monte Carlo,” *Phys. Rev. Lett.* **81**, 2514 (1998).
 - [21] J. T. Titantah, C. Pierleoni, and S. Ciuchi, “Free energy of the Fröhlich polaron in two and three dimensions,” *Phys. Rev. Lett.* **87**, 206406 (2001).

- [22] F. Grusdt, “All-coupling theory for the Fröhlich polaron,” *Phys. Rev. B* **93**, 144302 (2016).
- [23] A. S. Alexandrov and J. T. Devreese, *Advances in polaron physics* (Springer, Berlin, 2010).
- [24] F. Grusdt and E. Demler, *Quantum Matter at Ultralow Temperatures*, edited by S. S. M. Inguscio, W. Ketterle and G. Roati (IOS Press, Amsterdam, 2016).
- [25] J. Devreese, “Fröhlich polarons. Lecture course including detailed theoretical derivations,” arXiv:1611.06122 .
- [26] M. Bahrami, A. Großardt, S. Donadi, and A. Bassi, “The Schrödinger–Newton equation and its foundations,” *New J. Phys.* **16**, 115007 (2014).
- [27] R. Penrose, “On the gravitization of quantum mechanics 1: Quantum state reduction,” *Found. Phys.* **44**, 557 (2014).
- [28] I. M. Moroz, R. Penrose, and P. Tod, “Spherically-symmetric solutions of the Schrödinger–Newton equations,” *Class. Quantum Grav.* **15**, 2733 (1998).
- [29] P. Tod and I. M. Moroz, “An analytical approach to the schrödinger-newton equations,” *Nonlinearity* **12**, 201 (1999).
- [30] R. Harrison, I. Moroz, and K. Tod, “A numerical study of the Schrödinger–Newton equations,” *Nonlinearity* **16**, 101 (2002).
- [31] R. L. Frank and G. Zhou, “Derivation of an effective evolution equation for a strongly coupled polaron,” *Anal. PDE* **10**, 379 (2017).
- [32] J. Tempere, W. Casteels, M. K. Oberthaler, S. Knoop, E. Timmermans, and J. T. Devreese, “Feynman path-integral treatment of the BEC-impurity polaron,” *Phys. Rev. B* **80**, 184504 (2009).

- [33] S. P. Rath and R. Schmidt, “Field-theoretical study of the Bose polaron,” *Phys. Rev. A* **88**, 053632 (2013).
- [34] J. H. Denschlag, J. Simsarian, H. Häffner, C. McKenzie, A. Browaeys, D. Cho, K. Helmerson, S. Rolston, and W. D. Phillips, “A Bose-Einstein condensate in an optical lattice,” *J. Phys. B. Atom. Molec.* **35**, 3095 (2002).
- [35] M. Bruderer, A. Klein, S. R. Clark, and D. Jaksch, “Polaron physics in optical lattices,” *Phys. Rev. A* **76**, 011605 (2007).
- [36] M. Cetina, M. Jag, R. S. Lous, I. Fritsche, J. T. Walraven, R. Grimm, J. Levinsen, M. M. Parish, R. Schmidt, M. Knap, *et al.*, “Ultrafast many-body interferometry of impurities coupled to a Fermi sea,” *Science* **354**, 96 (2016).
- [37] F. Scazza, G. Valtolina, P. Massignan, A. Recati, A. Amico, A. Burchianti, C. Fort, M. Inguscio, M. Zaccanti, and G. Roati, “Repulsive Fermi polarons in a resonant mixture of ultracold ^6Li atoms,” *Phys. Rev. Lett.* **118**, 083602 (2017).
- [38] M. Sendner, P. K. Nayak, D. A. Egger, S. Beck, C. Müller, B. Epding, W. Kowalsky, L. Kronik, H. J. Snaith, A. Pucci, *et al.*, “Optical phonons in methylammonium lead halide perovskites and implications for charge transport,” *Mater. Horiz.* **3**, 613 (2016).
- [39] M. Schlipf, S. Poncé, and F. Giustino, “Carrier lifetimes and polaronic mass enhancement in the hybrid halide perovskite $\text{CH}_3\text{NH}_3\text{PbI}_3$ from multiphonon Fröhlich coupling,” *Phys. Rev. Lett.* **121**, 086402 (2018).
- [40] C. Verdi, F. Caruso, and F. Giustino, “Origin of the crossover from polarons to Fermi liquids in transition metal oxides,” *Nat. Commun.* **8**, 15769 (2017).

-
- [41] J. Vlietinck, W. Casteels, K. Van Houcke, J. Tempere, J. Ryckebusch, and J. T. Devreese, “Diagrammatic Monte Carlo study of the acoustic and the Bose–Einstein condensate polaron,” *New J. Phys.* **17**, 033023 (2015).
- [42] Y. Lépine and Y. Frongillo, “Ground-state energy and effective mass of the intermediate-radius polaron in small-bandwidth polar materials,” *Phys. Rev. B* **46**, 14510 (1992).
- [43] C. Spreefico and J. VandeVondele, “The nature of excess electrons in anatase and rutile from hybrid DFT and RPA,” *Phys. Chem. Chem. Phys.* **16**, 26144 (2014).
- [44] F. Zheng and L.-w. Wang, “Large polaron formation and its effect on electron transport in hybrid perovskites,” *Energy Environ. Sci.* **12**, 1219 (2019).
- [45] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, “*Ab initio* theory of polarons: Formalism and applications,” *Phys. Rev. B* **99**, 235139 (2019).
- [46] B. Sadigh, P. Erhart, and D. Åberg, “Variational polaron self-interaction-corrected total-energy functional for charge excitations in insulators,” *Phys. Rev. B* **92**, 075202 (2015).
- [47] T. Xu, T. Shimada, Y. Araki, M. Mori, G. Fujimoto, J. Wang, T.-Y. Zhang, and T. Kitamura, “Electron engineering of metallic multiferroic polarons in epitaxial BaTiO₃,” *Npj Comput. Mater.* **5**, 23 (2019).
- [48] M. Reticcioli, U. Diebold, G. Kresse, and C. Franchini, “Small polarons in transition metal oxides,” *Handbook of Materials Modeling: Applications: Current and Emerging Materials*, 1 (2019).
- [49] J. Nightingale, J. Wade, D. Moia, J. Nelson, and J.-S. Kim, “Impact of molec-

- ular order on polaron formation in conjugated polymers,” *J. Phys. Chem. C* **122**, 29129 (2018).
- [50] J. P. Perdew and A. Zunger, “Self-interaction correction to density-functional approximations for many-electron systems,” *Phys. Rev. B* **23**, 5048 (1981).
- [51] M. Setvin, C. Franchini, X. Hao, M. Schmid, A. Janotti, M. Kaltak, C. G. Van de Walle, G. Kresse, and U. Diebold, “Direct view at excess electrons in TiO₂ rutile and anatase,” *Phys. Rev. Lett.* **113**, 086402 (2014).
- [52] B. Himmetoglu, A. Janotti, L. Bjaalie, and C. G. Van de Walle, “Interband and polaronic excitations in YTiO₃ from first principles,” *Phys. Rev. B* **90**, 161102 (2014).
- [53] S. Kokott, S. V. Levchenko, P. Rinke, and M. Scheffler, “First-principles supercell calculations of small polarons with proper account for long-range polarization effects,” *New J. Phys.* **20**, 033023 (2018).
- [54] Z. Feng, V. Timoshevskii, A. Mauger, C. M. Julien, K. H. Bevan, and K. Zaghib, “Dynamics of polaron formation in Li₂O₂ from density functional perturbation theory,” *Phys. Rev. B* **88**, 184302 (2013).
- [55] A. Janotti, J. B. Varley, M. Choi, and C. G. Van de Walle, “Vacancies and small polarons in SrTiO₃,” *Phys. Rev. B* **90**, 085202 (2014).
- [56] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, “The electronic properties of graphene,” *Rev. Mod. Phys.* **81**, 109 (2009).
- [57] M. Xu, T. Liang, M. Shi, and H. Chen, “Graphene-like two-dimensional materials,” *Chem. Rev.* **113**, 3766 (2013).
- [58] A. Favron, E. Gaufrès, F. Fossard, A.-L. Phaneuf-LHeureux, N. Y. Tang, P. L. Lévesque, A. Loiseau, R. Leonelli, S. Francoeur, and R. Martel, “Photooxi-

- dation and quantum confinement effects in exfoliated black phosphorus,” *Nat. Mater.* **14**, 826 (2015).
- [59] L. C. Gomes and A. Carvalho, “Phosphorene analogues: Isoelectronic two-dimensional group-IV monochalcogenides with orthorhombic structure,” *Phys. Rev. B* **92**, 085406 (2015).
- [60] Y. Kubota, K. Watanabe, O. Tsuda, and T. Taniguchi, “Deep ultraviolet light-emitting hexagonal boron nitride synthesized at atmospheric pressure,” *Science* **317**, 932 (2007).
- [61] A. K. Geim and I. V. Grigorieva, “Van der Waals heterostructures,” *Nature* **499**, 419 (2013).
- [62] K. Novoselov, A. Mishchenko, A. Carvalho, and A. C. Neto, “2D materials and van der Waals heterostructures,” *Science* **353**, aac9439 (2016).
- [63] M. Reticcioli, M. Setvin, M. Schmid, U. Diebold, and C. Franchini, “Formation and dynamics of small polarons on the rutile $\text{TiO}_2(110)$ surface,” *Phys. Rev. B* **98**, 045306 (2018).
- [64] J. Wiktor and A. Pasquarello, “Electron and hole polarons at the BiVO_4 -water interface,” *ACS Appl. Mater. Interfaces*. (2019).
- [65] W. Kong, J. Zhou, Y. Z. Luo, T. Yang, S. Wang, J. Chen, A. Rusydi, Y. P. Feng, and M. Yang, “Formation of two-dimensional small polarons at the conducting $\text{LaAlO}_3/\text{SrTiO}_3$ interface,” *Phys. Rev. B* **100**, 085413 (2019).
- [66] S. Yang, A. T. Brant, N. C. Giles, and L. E. Halliburton, “Intrinsic small polarons in rutile TiO_2 ,” *Phys. Rev. B* **87**, 125201 (2013).
- [67] J. D. Jackson, *Classical electrodynamics* (Wiley, New York, 1998).

- [68] S. J. Miyake, “Strong-coupling limit of the polaron ground state,” J. Phys. Soc. Jpn. **38**, 181 (1975).
- [69] J. T. Devreese and A. S. Alexandrov, “Fröhlich polaron and bipolaron: recent developments,” Rep. Prog. Phys. **72**, 066501 (2009).
- [70] H. Fröhlich, “Electrons in lattice fields,” Adv. Phys. **3**, 325 (1954).
- [71] B. Gerlach and H. Löwen, “Analytical properties of polaron systems or: Do polaronic phase transitions exist or not?” Rev. Mod. Phys. **63**, 63 (1991).
- [72] N. Bogolyubov, “On a new form of the adiabatic theory of disturbances in the problem of interaction of particles with a quantum field,” Ukr. Mat. Zh. **2**, 3 (1950).
- [73] G. R. Allcock, “On the polaron rest energy and effective mass,” Adv. Phys. **5**, 412 (1956).
- [74] Y. Ōsaka, “Polaron state at a finite temperature,” Prog. Theor. Exp. Phys. **22**, 437 (1959).
- [75] G. Sewell, “Electrons in polar crystals,” Philos. Mag. **3**, 1361 (1958).
- [76] D. M. Eagles, “Optical absorption in ionic crystals involving small polarons,” Phys. Rev. **130**, 1381 (1963).
- [77] I. Lang and Y. A. Firsov, “Kinetic theory of semiconductors with low mobility,” Sov. Phys. JETP **16**, 1301 (1963).
- [78] V. Lakhno, “Large-radius Holstein polaron and the problem of spontaneous symmetry breaking,” Prog. Theor. Exp. Phys. **2014** (2014).
- [79] A. Chatterjee, S. Mukhopadhyay, L. Landau, L. Landau, S. Pekar, S. Pekar, S. Pekar, H. Fröhlich, C. Kuper, G. Whitfield, *et al.*, “Issledovaniya po elec-

- tronnoy teovii kristallov,” in *Polarons and Bipolarons: An Introduction*, Vol. 3 (Plenum Moscow, Russia, 2019) pp. xi–xiv.
- [80] L. Landau and S. Pekar, “Effective mass of a polaron,” *J. Exp. Theor. Phys* **18**, 419 (1948).
- [81] Y. Lépine, “The phase diagram of the tight-binding frohlich polaron,” *J. Phys. Condens. Matter* **6**, 6611 (1994).
- [82] A. S. Alexandrov and P. E. Kornilovitch, “Mobile small polaron,” *Phys. Rev. Lett.* **82**, 807 (1999).
- [83] A. S. Alexandrov, “Polaron dynamics and bipolaron condensation in cuprates,” *Phys. Rev. B* **61**, 12315 (2000).
- [84] A. S. Alexandrov and B. Y. Yavidov, “Small adiabatic polaron with a long-range electron-phonon interaction,” *Phys. Rev. B* **69**, 073101 (2004).
- [85] C. Perroni, V. Cataudella, and G. De Filippis, “Polaron features for long-range electron–phonon interaction,” *J. Phys. Condens. Matter* **16**, 1593 (2004).
- [86] O. Rösch and O. Gunnarsson, “Electron-phonon interaction in the t – J model,” *Phys. Rev. Lett.* **92**, 146403 (2004).
- [87] L. Cruzeiro-Hansson, J. Eilbeck, J. Marin, and F. Russell, “Interplay between dispersive and non-dispersive modes in the polaron problem,” *Phys. Lett. A* **266**, 160 (2000).
- [88] A. S. Alexandrov, *Theory of superconductivity: from weak to strong coupling* (IOP, Bristol, 2003).
- [89] F. Ortmann, F. Bechstedt, and K. Hannewald, “Theory of charge transport in organic crystals: Beyond Holstein’s small-polaron model,” *Phys. Rev. B* **79**, 235206 (2009).

- [90] H. Fehske and S. Trugman, “Numerical solution of the holstein polaron problem,” in *Polarons in Advanced Materials*, Springer Series in Material Sciences 103, edited by A. S. Alexandrov (Springer Verlag, Dordrecht, 2007) pp. 393–461.
- [91] P. Kornilovitch, “Path integrals in the physics of lattice polarons,” in *Polarons in Advanced Materials*, Springer Series in Material Sciences 103, edited by A. S. Alexandrov (Springer Verlag, Dordrecht, 2007) pp. 192–230.
- [92] A. S. Mishchenko, N. V. Prokof’ev, A. Sakamoto, and B. V. Svistunov, “Diagrammatic quantum Monte Carlo study of the Fröhlich polaron,” *Phys. Rev. B* **62**, 6317 (2000).
- [93] T. Hahn, S. Klimin, J. Tempere, J. T. Devreese, and C. Franchini, “Diagrammatic Monte Carlo study of Fröhlich polaron dispersion in two and three dimensions,” *Phys. Rev. B* **97**, 134305 (2018).
- [94] P. Nagels, M. Denayer, and J. Devreese, “Electrical properties of single crystals of uranium dioxide,” *Solid State Commun.* **1**, 35 (1963).
- [95] L. Kispert, J. Joseph, G. Miller, and R. Baughman, “EPR study of polarons in a conducting polymer with nondegenerate ground states: Alkali metal complexes of poly (p-phenylene) and phenylene oligomers,” *J. Chem. Phys.* **81**, 2119 (1984).
- [96] M. Reticcioli, I. Sokolović, M. Schmid, U. Diebold, M. Setvin, and C. Franchini, “Interplay between adsorbates and polarons: Co on rutile $\text{TiO}_2(110)$,” *Phys. Rev. Lett.* **122**, 016805 (2019).
- [97] M. Setvín, U. Aschauer, P. Scheiber, Y.-F. Li, W. Hou, M. Schmid, A. Selloni, and U. Diebold, “Reaction of O_2 with subsurface oxygen vacancies on TiO_2 anatase (101),” *Science* **341**, 988 (2013).

-
- [98] C. J. Chen, *Introduction to scanning tunneling microscopy*, Vol. 4 (Oxford University Press on Demand, 1993).
- [99] F. Besenbacher, “Scanning tunnelling microscopy studies of metal surfaces,” Rep. Prog. Phys. **59**, 1737 (1996).
- [100] H. Zhang, X. Li, Y. Chen, J. Park, A.-P. Li, and X.-G. Zhang, “Postprocessing algorithm for driving conventional scanning tunneling microscope at fast scan rates,” Scanning **2017** (2017).
- [101] H. Matsui, T. Sato, T. Takahashi, S.-C. Wang, H.-B. Yang, H. Ding, T. Fujii, T. Watanabe, and A. Matsuda, “BCS-like Bogoliubov quasiparticles in high- T_c superconductors observed by angle-resolved photoemission spectroscopy,” Phys. Rev. Lett. **90**, 217002 (2003).
- [102] D. Takane, Z. Wang, S. Souma, K. Nakayama, T. Nakamura, H. Oinuma, Y. Nakata, H. Iwasawa, C. Cacho, T. Kim, K. Horiba, H. Kumigashira, T. Takahashi, Y. Ando, and T. Sato, “Observation of chiral fermions with a large topological charge and associated Fermi-arc surface states in CoSi,” Phys. Rev. Lett. **122**, 076402 (2019).
- [103] A. Kordyuk, “ARPES experiment in fermiology of quasi-2D metals,” Low Temp. Phys. **40**, 286 (2014).
- [104] S. Hüfner, *Photoelectron spectroscopy: principles and applications* (Springer, 2013).
- [105] A. Damascelli, Z. Hussain, and Z.-X. Shen, “Angle-resolved photoemission studies of the cuprate superconductors,” Rev. Mod. Phys. **75**, 473 (2003).
- [106] G. D. Mahan, *Many-particle physics* (Springer Science & Business Media, 2013).

- [107] F. Giustino, “Electron-phonon interactions from first principles,” *Rev. Mod. Phys.* **89**, 015003 (2017).
- [108] J. Sjakste, N. Vast, M. Calandra, and F. Mauri, “Wannier interpolation of the electron-phonon matrix elements in polar semiconductors: Polar-optical coupling in GaAs,” *Phys. Rev. B* **92**, 054307 (2015).
- [109] J. Lischner, D. Vigil-Fowler, and S. G. Louie, “Physical origin of satellites in photoemission of doped graphene: An *Ab initio* GW plus cumulant study,” *Phys. Rev. Lett.* **110**, 146801 (2013).
- [110] F. Caruso, H. Lambert, and F. Giustino, “Band structures of plasmonic polarons,” *Phys. Rev. Lett.* **114**, 146404 (2015).
- [111] B. Gumhalter, V. Kovač, F. Caruso, H. Lambert, and F. Giustino, “On the combined use of GW approximation and cumulant expansion in the calculations of quasiparticle spectra: The paradigm of Si valence bands,” *Phys. Rev. B* **94**, 035103 (2016).
- [112] J. Inglesfield, “Plasmon satellites in core-level photoemission,” *J. Phys. C: Solid State Phys.* **16**, 403 (1983).
- [113] M. Kang, S. W. Jung, W. J. Shin, Y. Sohn, S. H. Ryu, T. K. Kim, M. Hoesch, and K. S. Kim, “Holstein polaron in a valley-degenerate two-dimensional semiconductor,” *Nat. Mater.* **17**, 1 (2018).
- [114] C. Chen, J. Avila, E. Frantzeskakis, A. Levy, and M. C. Asensio, “Observation of a two-dimensional liquid of Fröhlich polarons at the bare SrTiO₃ surface,” *Nat. Commun.* **6**, 8585 (2015).
- [115] J. M. Riley, F. Caruso, C. Verdi, L. Duffy, M. D. Watson, L. Bawden, K. Volkert, G. van der Laan, T. Hesjedal, M. Hoesch, *et al.*, “Crossover from lattice

- to plasmonic polarons of a spin-polarised electron gas in ferromagnetic EuO,” Nat. Commun. **9**, 2305 (2018).
- [116] E. Pastor, J.-S. Park, L. Steier, S. Kim, M. Grätzel, J. R. Durrant, A. Walsh, and A. A. Bakulin, “In situ observation of picosecond polaron self-localisation in α -Fe₂O₃ photoelectrochemical cells,” Nat. Commun. **10**, 3962 (2019).
- [117] F. Giustino, *Materials modelling using density functional theory: properties and predictions* (Oxford University Press, 2014).
- [118] L. Lin, *Density functional theory and nuclear quantum effects* (Princeton University, 2011).
- [119] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” Phys. Rev. **136**, B864 (1964).
- [120] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects,” Phys. Rev. **140**, A1133 (1965).
- [121] D. M. Ceperley and B. J. Alder, “Ground state of the electron gas by a stochastic method,” Phys. Rev. Lett. **45**, 566 (1980).
- [122] D. Marx and J. Hutter, “Ab initio molecular dynamics: Theory and implementation,” Modern methods and algorithms of quantum chemistry **1**, 141 (2000).
- [123] N. Ashcroft and N. Mermin, “Solid state physics int. ed,” Holt-Saunders, Philadelphia **16** (1976).
- [124] Y. Yao and Y. Kanai, “Plane-wave pseudopotential implementation and performance of scan meta-gga exchange-correlation functional for extended systems,” J. Chem. Phys. **146**, 224105 (2017).
- [125] H. J. Monkhorst and J. D. Pack, “Special points for brillouin-zone integrations,” Phys. Rev. B **13**, 5188 (1976).

-
- [126] W. C. Topp and J. J. Hopfield, “Chemically motivated pseudopotential for sodium,” *Phys. Rev. B* **7**, 1295 (1973).
- [127] T. Starkloff and J. D. Joannopoulos, “Local pseudopotential theory for transition metals,” *Phys. Rev. B* **16**, 5212 (1977).
- [128] J. R. Chelikowsky, “The origin of the pseudopotential density functional method. perspective on microscopic theory of phase transformation and lattice dynamics of Si,” in *Theoretical Chemistry Accounts* (Springer, 2000) pp. 340–342.
- [129] S. Goedecker and K. Maschke, “Transferability of pseudopotentials,” *Phys. Rev. A* **45**, 88 (1992).
- [130] D. R. Hamann, M. Schlüter, and C. Chiang, “Norm-conserving pseudopotentials,” *Phys. Rev. Lett.* **43**, 1494 (1979).
- [131] M. Fuchs and M. Scheffler, “Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory,” *Comput. Phys. Commun.* **119**, 67 (1999).
- [132] N. Troullier and J. L. Martins, “Efficient pseudopotentials for plane-wave calculations,” *Phys. Rev. B* **43**, 1993 (1991).
- [133] W. E. Pickett, “Pseudopotential methods in condensed matter applications,” *Computer Physics Reports* **9**, 115 (1989).
- [134] D. J. Singh and L. Nordstrom, *Planewaves, Pseudopotentials, and the LAPW method* (Springer Science & Business Media, 2006).
- [135] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, “Phonons and related crystal properties from density-functional perturbation theory,” *Rev. Mod. Phys.* **73**, 515 (2001).

-
- [136] S. Baroni, P. Giannozzi, and A. Testa, “Green’s-function approach to linear response in solids,” *Phys. Rev. Lett.* **58**, 1861 (1987).
- [137] R. Heid, “12 density functional perturbation theory and electron phonon coupling,” Emergent Phenomena in Correlated Matter: Autumn School Organized by the Forschungszentrum Jülich and the German Research School for Simulation Sciences at Forschungszentrum Jülich 23-27 September 2013; Lecture Notes of the Autumn School Correlated Electrons 2013 **3** (2013).
- [138] D. Vanderbilt and R. D. King-Smith, “Electric polarization as a bulk quantity and its relation to surface charge,” *Phys. Rev. B* **48**, 4442 (1993).
- [139] P. Giannozzi and S. Baroni, “Density-functional perturbation theory,” in *Handbook of Materials Modeling* (Springer, 2005) pp. 195–214.
- [140] F. Giustino, M. L. Cohen, and S. G. Louie, “Electron-phonon interaction using Wannier functions,” *Phys. Rev. B* **76**, 165108 (2007).
- [141] N. Marzari and D. Vanderbilt, “Maximally localized generalized Wannier functions for composite energy bands,” *Phys. Rev. B* **56**, 12847 (1997).
- [142] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, “Maximally localized Wannier functions: Theory and applications,” *Rev. Mod. Phys.* **84**, 1419 (2012).
- [143] C. Verdi and F. Giustino, “Fröhlich electron-phonon vertex from first principles,” *Phys. Rev. Lett.* **115**, 176401 (2015).
- [144] N. Marzari, I. Souza, and D. Vanderbilt, “An introduction to maximally-localized Wannier functions,” *Psi-K newsletter* **57**, 129 (2003).
- [145] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and

- N. Marzari, “An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions,” *Comput. Phys. Commun.* **185**, 2309 (2014).
- [146] P. Vogl, “Microscopic theory of electron-phonon interaction in insulators or semiconductors,” *Phys. Rev. B* **13**, 694 (1976).
- [147] J. Noffsinger, F. Giustino, B. D. Malone, C.-H. Park, S. G. Louie, and M. L. Cohen, “EPW: A program for calculating the electron–phonon coupling using maximally localized Wannier functions,” *Comput. Phys. Commun.* **181**, 2140 (2010).
- [148] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, “EPW: Electronphonon coupling, transport and superconducting properties using maximally localized Wannier functions,” *Comput. Phys. Commun.* **209**, 116 (2016).
- [149] P. Giannozzi *et al.*, “QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials,” *J. Phys. Condens. Matter* **21**, 395502 (2009).
- [150] P. Giannozzi *et al.*, “Advanced capabilities for materials modelling with Quantum ESPRESSO,” *J. Phys. Condens. Matter* **29**, 465901 (2017).
- [151] J. Noffsinger, E. Kioupakis, C. G. Van de Walle, S. G. Louie, and M. L. Cohen, “Phonon-assisted optical absorption in silicon from first principles,” *Phys. Rev. Lett.* **108**, 167402 (2012).
- [152] E. R. Margine and F. Giustino, “Anisotropic Migdal-Eliashberg theory using Wannier functions,” *Phys. Rev. B* **87**, 024505 (2013).
- [153] E. R. Margine and F. Giustino, “Two-gap superconductivity in heavily n -doped graphene: *Ab initio* Migdal-Eliashberg theory,” *Phys. Rev. B* **90**, 014518 (2014).

-
- [154] C. Heil, S. Poncé, H. Lambert, M. Schlipf, E. R. Margine, and F. Giustino, “Origin of superconductivity and latent charge density wave in NbS₂,” *Phys. Rev. Lett.* **119**, 087003 (2017).
- [155] S. Poncé, E. R. Margine, and F. Giustino, “Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors,” *Phys. Rev. B* **97**, 121201 (2018).
- [156] S. Poncé, M. Schlipf, and F. Giustino, “Origin of low carrier mobilities in halide perovskites,” *ACS Energy Lett.* **4**, 456 (2019).
- [157] W. Li, S. Poncé, and F. Giustino, “Dimensional crossover in the carrier mobility of two-dimensional semiconductors: The case of InSe,” *Nano Lett.* **19**, 1774 (2019).
- [158] W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, “Polarons from first principles, without supercells,” *Phys. Rev. Lett.* **122**, 246403 (2019).
- [159] M. d’Avezac, M. Calandra, and F. Mauri, “Density functional theory description of hole-trapping in SiO₂: A self-interaction-corrected approach,” *Phys. Rev. B* **71**, 205210 (2005).
- [160] X. Gonze and C. Lee, “Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory,” *Phys. Rev. B* **55**, 10355 (1997).
- [161] L. Sham and M. Schlüter, “Density-functional theory of the band gap,” *Phys. Rev. B* **32**, 3883 (1985).
- [162] P. B. Allen and V. Heine, “Theory of the temperature dependence of electronic band structures,” *J. Phys. C: Solid State Phys.* **9**, 2305 (1976).

-
- [163] P. B. Allen and B. Mitrović, “Theory of superconducting T_c ,” *Solid State Phys.* **37**, 1 (1982).
- [164] A. Sanna, J. A. Flores-Livas, A. Davydov, G. Profeta, K. Dewhurst, S. Sharma, and E. Gross, “Ab initio Eliashberg theory: making genuine predictions of superconducting features,” *J. Phys. Soc. Jpn.* **87**, 041012 (2018).
- [165] F. Giustino, “Erratum: Electron-phonon interactions from first principles [Rev. Mod. Phys. 89, 15003 (2017)],” *Rev. Mod. Phys.* **91**, 019901 (2019).
- [166] S. Goedecker and C. J. Umrigar, “Critical assessment of the self-interaction-corrected–local-density-functional method and its algorithmic implementation,” *Phys. Rev. A* **55**, 1765 (1997).
- [167] J. M. Garcia-Lastra, J. S. G. Myrdal, R. Christensen, K. S. Thygesen, and T. Vegge, “DFT + U study of polaronic conduction in Li_2O_2 and Li_2CO_3 : Implications for Li – Air batteries,” *J. Phys. Chem. C* **117**, 5568 (2013).
- [168] P. Erhart, A. Klein, D. Åberg, and B. Sadigh, “Efficacy of the DFT + U formalism for modeling hole polarons in perovskite oxides,” *Phys. Rev. B* **90**, 035204 (2014).
- [169] D. M. Ramo, A. Shluger, J. Gavartin, and G. Bersuker, “Theoretical prediction of intrinsic self-trapping of electrons and holes in monoclinic HfO_2 ,” *Phys. Rev. Lett.* **99**, 155504 (2007).
- [170] C. Franchini, G. Kresse, and R. Podloucky, “Polaronic hole trapping in doped BaBiO_3 ,” *Phys. Rev. Lett.* **102**, 256402 (2009).
- [171] K. Momma and F. Izumi, “*VESTA3* for three-dimensional visualization of crystal, volumetric and morphology data,” *J. Appl. Crystallogr.* **44**, 1272 (2011).

- [172] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple,” *Phys. Rev. Lett.* **77**, 3865 (1996).
- [173] D. R. Hamann, “Optimized norm-conserving Vanderbilt pseudopotentials,” *Phys. Rev. B* **88**, 085117 (2013).
- [174] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, “Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials,” *Phys. Rev. B* **47**, 10142 (1993).
- [175] G. Makov and M. C. Payne, “Periodic boundary conditions in *ab initio* calculations,” *Phys. Rev. B* **51**, 4014 (1995).
- [176] J. Choi, J. Demmel, I. Dhillon, J. Dongarra, S. Ostrouchov, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, “ScaLAPACK: A portable linear algebra library for distributed memory computers design issues and performance,” *Comput. Phys. Commun.* **97**, 1 (1996).
- [177] B. Barker, “Message passing interface (mpi),” in *Workshop: High Performance Computing on Stampede*, Vol. 262 (2015).
- [178] C. L. Lawson, R. J. Hanson, D. R. Kincaid, and F. T. Krogh, “Basic linear algebra subprograms for fortran usage,” *ACM Trans. Math. Software* **5**, 308 (1977).
- [179] L. Prylli and B. Tourancheau, “Efficient block cyclic data redistribution,” in *European Conference on Parallel Processing* (Springer, 1996) pp. 155–164.
- [180] W. D. Gropp and R. Thakur, “Revealing the performance of MPI RMA implementations,” in *European Parallel Virtual Machine/Message Passing Interface Users Group Meeting* (Springer, 2007) pp. 272–280.

-
- [181] J. Squar, *MPI-3 Algorithms for 3D Radiative Transfer on Intel Xeon Phi Co-processors*, Ph.D. thesis, Universität Hamburg, Fachbereich Informatik (2016).
- [182] G. Iadonisi, “Electron-phonon interaction: Effects on the excitation spectrum of solids,” *Riv. Nuovo Cimento* **7**, 1 (1984).
- [183] J. Lægsgaard and K. Stokbro, “Hole trapping at Al impurities in silica: A challenge for density functional theories,” *Phys. Rev. Lett.* **86**, 2834 (2001).
- [184] H. Gudmundsdottir *et al.*, “Calculations of Al dopant in α -quartz using a variational implementation of the Perdew–Zunger self-interaction correction,” *New J. Phys.* **17**, 083006 (2015).
- [185] G. Smith and L. Alexander, “Refinement of the atomic parameters of α -quartz,” *Acta Crystallogr.* **16**, 462 (1963).
- [186] G. Baldacchini and R. M. Montereali, “New perspectives of coloured LiF for optoelectronic devices,” *Opt. Mater.* **16**, 53 (2001).
- [187] F. Karsai, P. Tiwald, R. Laskowski, F. Tran, D. Koller, S. Gräfe, J. Burgdörfer, L. Wirtz, and P. Blaha, “F center in lithium fluoride revisited: Comparison of solid-state physics and quantum-chemistry approaches,” *Phys. Rev. B* **89**, 125429 (2014).
- [188] M. R. Pederson and B. M. Klein, “Improved theoretical methods for studies of defects in insulators: Application to the F center in LiF,” *Phys. Rev. B* **37**, 10319 (1988).
- [189] G. Mallia, R. Orlando, C. Roetti, P. Ugliengo, and R. Dovesi, “F center in LiF: A quantum mechanical *ab initio* investigation of the hyperfine interaction between the unpaired electron at the vacancy and its first seven neighbors,” *Phys. Rev. B* **63**, 235102 (2001).

-
- [190] R. Williams and K. Song, “The self-trapped exciton,” *J. Phys. Chem. Solids* **51**, 679 (1990).
- [191] A. Shluger and A. Stoneham, “Small polarons in real crystals: concepts and problems,” *J. Phys. Condens. Matter* **5**, 3049 (1993).
- [192] J. L. Gavartin, P. V. Sushko, and A. L. Shluger, “Modeling charge self-trapping in wide-gap dielectrics: Localization problem in local density functionals,” *Phys. Rev. B* **67**, 035108 (2003).
- [193] O. Schirmer, “O- bound small polarons in oxide materials,” *J. Phys. Condens. Matter* **18**, R667 (2006).
- [194] M. Inoue, C. K. Mahutte, and S. Wang, “Electronic polarons in alkali halides,” *Phys. Rev. B* **2**, 539 (1970).
- [195] L. Dressler, U. Griebner, and R. Kittner, “Precision measurement of lattice parameters in LiF monocrystals,” *Cryst. Res. Technol.* **22**, 1431 (1987).
- [196] M. Piacentini, D. W. Lynch, and C. G. Olson, “Thermoreflectance of LiF between 12 and 30 eV,” *Phys. Rev. B* **13**, 5530 (1976).
- [197] L. J. Page and E. H. Hygh, “Calculation of energy bands in alkali halides,” *Phys. Rev. B* **1**, 3472 (1970).
- [198] C. Andeen, J. Fontanella, and D. Schuele, “Low-frequency dielectric constant of LiF, NaF, NaCl, NaBr, KCl, and KBr by the method of substitution,” *Phys. Rev. B* **2**, 5068 (1970).
- [199] E. R. Levin and E. L. Offenbacher, “Theory of dielectric constants of LiF,” *Phys. Rev.* **118**, 1142 (1960).
- [200] G. Dolling, H. G. Smith, R. M. Nicklow, P. R. Vijayaraghavan, and M. K. Wilkinson, “Lattice dynamics of lithium fluoride,” *Phys. Rev.* **168**, 970 (1968).

- [201] S. P. Ong, Y. Mo, and G. Ceder, “Low hole polaron migration barrier in lithium peroxide,” *Phys. Rev. B* **85**, 081105 (2012).
- [202] V. Viswanathan, K. S. Thygesen, J. Hummelshøj, J. K. Nørskov, G. Girishkumar, B. McCloskey, and A. Luntz, “Electrical conductivity in Li_2O_2 and its role in determining capacity limitations in non-aqueous Li- O_2 batteries,” *J. Chem. Phys.* **135**, 214704 (2011).
- [203] J. Kang, Y. S. Jung, S.-H. Wei, and A. C. Dillon, “Implications of the formation of small polarons in Li_2O_2 for Li-air batteries,” *Phys. Rev. B* **85**, 035210 (2012).
- [204] M. K. Chan, E. L. Shirley, N. K. Karan, M. Balasubramanian, Y. Ren, J. P. Greeley, and T. T. Fister, “Structure of lithium peroxide,” *J. Phys. Chem. Lett.* **2**, 2483 (2011).
- [205] J. M. G. Lastra, J. Bass, and K. S. Thygesen, “Communication: Strong excitonic and vibronic effects determine the optical properties of Li_2O_2 ,” *J. Chem. Phys.* **135**, 121101 (2011).
- [206] K. C. Lau, D. Qiu, X. Luo, J. Greeley, L. A. Curtiss, J. Lu, and K. Amine, “Theoretical exploration of various lithium peroxide crystal structures in a Li-air battery,” *Energies* **8**, 529 (2015).
- [207] N. F. Mott, “Metal-insulator transition,” *Rev. Mod. Phys.* **40**, 677 (1968).
- [208] M. Rohlfing and S. G. Louie, “Electron-hole excitations and optical spectra from first principles,” *Phys. Rev. B* **62**, 4927 (2000).
- [209] M. Bokdam, T. Sander, A. Stroppa, S. Picozzi, D. Sarma, C. Franchini, and G. Kresse, “Role of polar phonons in the photo excited state of metal halide perovskites,” *Sci. Rep.* **6**, 28618 (2016).

- [210] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, “*Ab Initio* calculations of the quasiparticle and absorption spectra of clusters: The sodium tetramer,” *Phys. Rev. Lett.* **75**, 818 (1995).
- [211] A. Molina-Sanchez, K. Hummer, and L. Wirtz, “Vibrational and optical properties of MoS₂: From monolayer to bulk,” *Surf. Sci. Rep.* **70**, 554 (2015).
- [212] X. Ling, H. Wang, S. Huang, F. Xia, and M. S. Dresselhaus, “The renaissance of black phosphorus,” *Proc. Natl. Acad. Sci.* **112**, 4523 (2015).
- [213] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, “Single-layer MoS₂ transistors,” *Nat. Nanotechnol.* **6**, 147 (2011).
- [214] J. Pei, X. Gai, J. Yang, X. Wang, Z. Yu, D.-Y. Choi, B. Luther-Davies, and Y. Lu, “Producing air-stable monolayers of phosphorene and their defect engineering,” *Nat. Commun.* **7**, 10450 (2016).
- [215] J. Zhou, J. Lin, X. Huang, Y. Zhou, Y. Chen, J. Xia, H. Wang, Y. Xie, H. Yu, J. Lei, *et al.*, “A library of atomically thin metal chalcogenides,” *Nature* **556**, 355 (2018).
- [216] F. Ferreira, A. Chaves, N. Peres, and R. Ribeiro, “Excitons in hexagonal boron nitride single-layer: a new platform for polaritonics in the ultraviolet,” *J. Opt. Soc. Am. B* **36**, 674 (2019).
- [217] R. Mas-Balleste, C. Gomez-Navarro, J. Gomez-Herrero, and F. Zamora, “2D materials: to graphene and beyond,” *Nanoscale* **3**, 20 (2011).
- [218] J. S. Ponraj, Z.-Q. Xu, S. C. Dhanabalan, H. Mu, Y. Wang, J. Yuan, P. Li, S. Thakur, M. Ashrafi, K. McCoubrey, *et al.*, “Photonics and optoelectronics of two-dimensional materials beyond graphene,” *Nanotechnology* **27**, 462001 (2016).

-
- [219] J. Cheng, C. Wang, X. Zou, and L. Liao, “Recent advances in optoelectronic devices based on 2D materials and their heterostructures,” *Adv. Opt. Mater.* **7**, 1800441 (2019).
- [220] S. Das, D. Pandey, J. Thomas, and T. Roy, “The role of graphene and other 2D materials in solar photovoltaics,” *Adv. Mater.* **31**, 1802722 (2019).
- [221] Y. G. Kim, K. C. Kwon, Q. Van Le, K. Hong, H. W. Jang, and S. Y. Kim, “Atomically thin two-dimensional materials as hole extraction layers in organolead halide perovskite photovoltaic cells,” *J. Power Sources* **319**, 1 (2016).
- [222] S. Kohnehpoushi, P. Nazari, B. A. Nejand, and M. Eskandari, “MoS₂: a two-dimensional hole-transporting material for high-efficiency, low-cost perovskite solar cells,” *Nanotechnology* **29**, 205201 (2018).
- [223] H. Yang, H. Fang, H. Yu, Y. Chen, L. Wang, W. Jiang, Y. Wu, and J. Li, “Low temperature self-densification of high strength bulk hexagonal boron nitride,” *Nat. Commun.* **10** (2019).
- [224] C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, *et al.*, “Boron nitride substrates for high-quality graphene electronics,” *Nat. Nanotechnol.* **5**, 722 (2010).
- [225] G.-H. Lee, Y.-J. Yu, X. Cui, N. Petrone, C.-H. Lee, M. S. Choi, D.-Y. Lee, C. Lee, W. J. Yoo, K. Watanabe, *et al.*, “Flexible and transparent MoS₂ field-effect transistors on hexagonal boron nitride-graphene heterostructures,” *ACS Nano* **7**, 7931 (2013).
- [226] H. Jiang and J. Lin, “Hexagonal boron nitride for deep ultraviolet photonic devices,” *Semicond. Sci. Technol.* **29**, 084003 (2014).

- [227] C. Elias, P. Valvin, T. Pelini, A. Summerfield, C. Mellor, T. Cheng, L. Eaves, C. Foxon, P. Beton, S. Novikov, *et al.*, “Direct band-gap crossover in epitaxial monolayer boron nitride,” *Nat. Commun.* **10**, 2639 (2019).
- [228] S. A. Tawfik, S. Ali, M. Fronzi, M. Kianinia, T. T. Tran, C. Stampfl, I. Aharonovich, M. Toth, and M. J. Ford, “First-principles investigation of quantum emission from hBN defects,” *Nanoscale* **9**, 13575 (2017).
- [229] M. Abdi, J.-P. Chou, A. Gali, and M. B. Plenio, “Color centers in hexagonal boron nitride monolayers: A group theory and ab initio analysis,” *ACS Photonics* **5**, 1967 (2018).
- [230] F. Paleari, H. P. C. Miranda, A. Molina-Sánchez, and L. Wirtz, “Exciton-phonon coupling in the ultraviolet absorption and emission spectra of bulk hexagonal boron nitride,” *Phys. Rev. Lett.* **122**, 187401 (2019).
- [231] E. Cannuccia, B. Monserrat, and C. Attaccalite, “Theory of phonon-assisted luminescence in solids: Application to hexagonal boron nitride,” *Phys. Rev. B* **99**, 081109 (2019).
- [232] D. Wigger, R. Schmidt, O. Del Pozo-Zamudio, J. A. Preuß, P. Tonndorf, R. Schneider, P. Steeger, J. Kern, Y. Khodaei, J. Sperling, *et al.*, “Phonon-assisted emission and absorption of individual color centers in hexagonal boron nitride,” *2D Materials* **6**, 035006 (2019).
- [233] H.-P. Komsa and A. Pasquarello, “Finite-size supercell correction for charged defects at surfaces and interfaces,” *Phys. Rev. Lett.* **110**, 095505 (2013).
- [234] J. P. Perdew and Y. Wang, “Accurate and simple analytic representation of the electron-gas correlation energy,” *Phys. Rev. B* **45**, 13244 (1992).

- [235] A. V. Lebedev, I. V. Lebedeva, A. A. Knizhnik, and A. M. Popov, “Interlayer interaction and related properties of bilayer hexagonal boron nitride: *ab initio* study,” RSC Adv. **6**, 6423 (2016).
- [236] N. Drummond, V. Zólyomi, and V. Falko, “Electronic structure of two-dimensional crystals of hexagonal boron nitride,” .
- [237] K. Zhang, Y. Feng, F. Wang, Z. Yang, and J. Wang, “Two dimensional hexagonal boron nitride (2D-hBN): synthesis, properties and applications,” J. Mater. Chem. C **5**, 11992 (2017).
- [238] H. Henck, D. Pierucci, G. Fugallo, J. Avila, G. Cassabois, Y. J. Dappe, M. G. Silly, C. Chen, B. Gil, M. Gatti, F. Sottile, F. Sirotti, M. C. Asensio, and A. Ouerghi, “Direct observation of the band structure in bulk hexagonal boron nitride,” Phys. Rev. B **95**, 085410 (2017).
- [239] J. Serrano, A. Bosak, R. Arenal, M. Krisch, K. Watanabe, T. Taniguchi, H. Kanda, A. Rubio, and L. Wirtz, “Vibrational properties of hexagonal boron nitride: Inelastic x-ray scattering and *Ab Initio* calculations,” Phys. Rev. Lett. **98**, 095503 (2007).
- [240] T. Sohler, M. Gibertini, M. Calandra, F. Mauri, and N. Marzari, “Breakdown of optical phonons splitting in two-dimensional materials,” Nano lett. **17**, 3758 (2017).
- [241] S. Fang, R. Kuate Defo, S. N. Shirodkar, S. Lieu, G. A. Tritsarlis, and E. Kaxiras, “*Ab initio* tight-binding Hamiltonian for transition metal dichalcogenides,” Phys. Rev. B **92**, 205108 (2015).
- [242] Y. Wang, C. Huang, D. Li, P. Li, J. Yu, Y. Zhang, and J. Xu, “Tight-binding model for electronic structure of hexagonal boron phosphide monolayer and bilayer,” J. Phys. Condens. Matter **31**, 285501 (2019).

-
- [243] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, “Quasiparticle band structure of bulk hexagonal boron nitride and related systems,” *Phys. Rev. B* **51**, 6868 (1995).
- [244] B. Huang and H. Lee, “Defect and impurity properties of hexagonal boron nitride: A first-principles calculation,” *Phys. Rev. B* **86**, 245406 (2012).
- [245] H.-P. Komsa, N. Berseneva, A. V. Krashenninnikov, and R. M. Nieminen, “Charged point defects in the flatland: Accurate formation energy calculations in two-dimensional materials,” *Phys. Rev. X* **4**, 031044 (2014).
- [246] A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, and T. F. Heinz, “Exciton binding energy and nonhydrogenic Rydberg series in monolayer WS_2 ,” *Phys. Rev. Lett.* **113**, 076802 (2014).
- [247] A. Hanbicki, M. Currie, G. Kioseoglou, A. Friedman, and B. Jonker, “Measurement of high exciton binding energy in the monolayer transition-metal dichalcogenides WS_2 and WSe_2 ,” *Solid State Commun.* **203**, 16 (2015).
- [248] F. Paleari, T. Galvani, H. Amara, F. Ducastelle, A. Molina-Sánchez, and L. Wirtz, “Excitons in few-layer hexagonal boron nitride: Davydov splitting and surface localization,” *2D Materials* **5**, 045017 (2018).
- [249] D. Wickramaratne, L. Weston, and C. G. Van de Walle, “Monolayer to bulk properties of hexagonal boron nitride,” *J. Phys. Chem. C* **122**, 25524 (2018).
- [250] C. Chen, J. Avila, S. Wang, Y. Wang, M. Mucha-Kruczynski, C. Shen, R. Yang, B. Nosarzewski, T. P. Devereaux, G. Zhang, *et al.*, “Emergence of interfacial polarons from electron–phonon coupling in graphene/h-BN van der Waals heterostructures,” *Nano Lett.* **18**, 1082 (2018).

-
- [251] G. Fischerauer, “Comments on real-space Green’s function of an isolated point-charge in an unbounded anisotropic medium,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* **44**, 1179 (1997).

Acknowledgement

I would like to express my sincere gratitude to my supervisor Professor Feliciano Giustino for the inspiring discussions, helpful suggestions, and continual supporting throughout this DPhil. I appreciate his wisdom and strong intuition towards into physics questions, which motivates a lot of my researches. I believe that his rigorous attitudes had not only advanced my research quality and output, but also made indispensable contributions to the field. His passion and dedication to physics is particularly worth for one to learn from.

I am indebted to the Department of Chemistry and Materials of the University of Oxford for providing a good environment for my study and for providing kind support on academic activities. It is thoroughly enjoyable to work in such a place with a friendly atmosphere. Especially, I would like to thank Professor David E Logan and the TMCS doctoral training centre for their keen support and the extensive training of transferable skills.

I would like to give special thanks to Marina, Carla and Samuel for patiently guiding me at the beginning of this DPhil, and also I have to mention everyone in the Materials Modelling Laboratory, including Christoph, Fabio, Felix, George, Kun, Marios, Miguel, Martin, Nikolaus, Nourdine, Sven, Tim, Wenbin and Xinlei for countless lunches, cakes, and pubs.

I would also thank all the wonderful friends who I met during my student life here in Oxford and in the United Kingdom, and all my old friends in Macau and Hong Kong – although I cannot list all of you here. A special remark goes to Becky, Kana, Senio, and especially to Louise – thank you very much for all the great moments we had in the past few years. Life would never be complete without them.

Finally, I have to say I am so lucky to have a great family for supporting me in making any big decisions in my life. My parents Kuai-Hoi, Chi-Mui and my brother Eric Weng-Hou, have provided me unrelenting love and selfless support throughout my study abroad. My gratitude to them is beyond description.

List of publications

1. G. Volonakis, A. A. Haghighirad, R. L. Milot, W. H. Sio, M. R. Filip, B. Wenger, M. B. Johnston, L. M. Herz, H. J. Snaith, and F. Giustino, Cs₂InAgCl₆: A New Lead-Free Halide Double Perovskite with Direct Band Gap. *J. Phys. Chem. Lett.* 8, 4, 772-778 (2017).
2. W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, Polarons from First Principles, without Supercells. *Phys. Rev. Lett.* 122, 246403 (2019).
3. W. H. Sio, C. Verdi, S. Poncé, and F. Giustino, *Ab initio* theory of polarons: Formalism and applications. *Phys. Rev. B* 99, 235139 (2019).
4. W. H. Sio and F. Giustino, Two-dimensional polarons: the case of monolayer BN. *Paper in preparation.*