Applications of Electronic Structure Theory in Electron Microscopy



Tim Naginey
Corpus Christi College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2018

For Liam

Acknowledgements

The work contained in this thesis would have been absolutely impossible without the constant support and mentorship of my supervisors: Professor Pete Nellist, Professor Jonathan Yates, and Dr. Rebecca Nicholls.

The support of Professor David Logan and the entire cast and crew of the Theory and Modelling in the Chemical Sciences CDT were likewise instrumental in making this journey achievable–at times even enjoyable!

I am deeply indebted to Professor Colm Whelan for his wonderful mentorship over the years, and for his encouragement and aid in getting me into Oxford in the first place.

Thanks to my family who helped me at so many crucial points along the way, especially Mom (Grandma) and Tammy (Nana) for spending so much time taking care of (read: spoiling) Liam while I pursued my studies.

Also thanks to Hannah for being more than understanding and supportive of me during the chaos of these final months of writing up. Your love and your baked treats sustain me.

Thanks also to Brandi for putting your heart and soul into raising our little guy while I've been gone the better part of these past four years.

Lastly, thanks to Liam for putting up with conversations with me on computer screens in place of the real thing for too many evenings. Let's make up for lost time.

Abstract

Furthering the integration of theoretical and computational electronic structure methods with electron microscopy techniques is the overarching theme of this thesis. In this work, theoretical and computational tools are developed and applied to several specific experimental problems in electron microscopy.

Hartree-Fock theory is combined with methods for calculating damped dispersion and short-range correlation energies and used to calculate the blue shift in the $1s^2 \rightarrow 1s2p(^1P)$ excitation energy as a function of the density of confined helium atoms under extreme pressures. This is then compared to existing electron energy loss spectroscopy measurements on these systems. We find that, though the trends vary widely between different experiments, our theoretical predictions do not match very closely with any of them. We conclude that this is due either to the difficulty of the experimental technique or physical effects that are not accounted for in our model or the interpretation of the data.

A general approach for obtaining projected electrostatic potentials from pseudopotential plane wave (PPW) density functional theory (DFT) is presented in detail. Projected potentials are the crucial physical quantity of interest in the simulation of (scanning) transmission electron microscopy ((S)TEM) images, and the incorporation of more accurate electronic structure theory into the calculation of these potentials improves the accuracy of (S)TEM simulations. A novel solution to the problems arising in handling the singularities due to nuclei in the projected potentials is also presented and shown to improve the precision of image simulations while simultaneously reducing the complexity of obtaining projected potentials from DFT for the user. Through its implementation in CASTEP, a PPW code, this method of obtaining projected potentials is then used to demonstrate the sensitivity of the experimental technique of (S)TEM ptychography to the electronic structure of monolayer hexagonal boron nitride.

Contents

1 Introduction and outline **1**

2 Background and literature review **4**

2.1 Hartree-Fock theory 4

2.2 Density functional theory 6

2.3 EELS, electronic structure, and confined atoms 9

2.4 Modelling and simulation in (S)TEM 11

2.5 Modelling and measuring electronic structure in the electron microscope 17

3 Excited helium under high pressures in the bulk and in nanobubbles **25**

4 Calculation of projected potentials using pseudopotentials for use in electron microscopy simulations **36**

5 The treatment of nuclear singularities in electron microscopy simulations **48**

6 Direct imaging of charge redistribution in monolayer hexagonal boron nitride using electron ptychography in the scanning transmission electron microscope **57**

7 Conclusions **63**

Bibliography **66**

Chapter 1

Introduction and outline

Electron microscopy is a valuable experimental tool for gathering a wide range of information about many different types of samples, all the way down to the atomic scale. At this scale, comparison with reliable physical simulations is especially crucial for the interpretation of experimental data. Electronic structure theory has played an important role in bringing simulations in electron microscopy up to speed with the ever-improving instrumental capabilities in the field, but there is still much room for improvement.

In this thesis, we will identify several areas in which electronic structure theory may be applied to further the progress of the interface between theory and experiment. It is the aim of my project to evaluate how these applications of electronic structure theory may improve the agreement between simulations and experiment and aid in the interpretation of the data, through the development and application of electronic structure software. The outline of the thesis is as follows:

- We begin with a **literature review** in chapter 2, which will include brief summaries of the concepts from electronic structure theory and electron microscopy that will be used throughout this work. The literature review will go on to provide an in-depth look at what has been done so far in applying electronic structure theory to the interpretation of two types of microscopy experiments: electron energy loss spectroscopy (EELS) measurements of excitation energies of confined helium under extreme pressure; and the quantitative interpretation of experimental (scanning) transmission electron microscopy ((S)TEM) images via their comparison with simulated ones. The literature review will also highlight the gaps in the literature that we intend to fill with the methods and results presented in the subsequent papers that constitute the body of this thesis.

- The **first paper** [1] to be presented constitutes chapter 3 and concerns the modelling of clusters of helium atoms under extreme pressure using Hartree-Fock theory. The integrity of the metal coating of the fuel rods used in present-day nuclear reactors may be jeopardized by the formation of high-pressure helium bubbles inside of the coating. It is of interest to be able to measure the pressure of these helium bubbles in order to evaluate whether they pose a serious threat to the integrity of the coating. This can be deduced from the blue shift in the $1s^2 \rightarrow 1s2p(^1P)$ excitation energy of the helium atoms in the bubble, which can now be measured very accurately using EELS. We utilize Hartree-Fock theory in combination with two corrections related to correlation effects to predict the dependence of the excitation energy and scattering cross section on the density of the helium.

The rest of the papers [2–4], which make up chapters 4, 5, and 6, deal with improving the accuracy of (S)TEM image simulations using density functional theory (DFT). Most of the software used to simulate (S)TEM images completely neglects interatomic interactions in the sample and thus leaves out the effects of charge transfer and chemical bonding. The effect of these interactions can be essential to understanding many of the properties of materials. These interactions are especially important for fully quantitative interpretations of (S)TEM images of bulk oxides because of the strong ionic bonds, and 2D materials composed of light elements since most of the electrons are involved in bonding. In order to interpret electron microscopy data in terms of charge transfer, we employ DFT to calculate the projected electrostatic potential of the sample, which is the main physical quantity needed to model (S)TEM images.

- The **second paper** [2] is presented in chapter 4, in which we introduce a general approach for calculating projected potentials in pseudopotential plane wave (PPW) DFT so that these quantities can be directly obtained from an electronic structure calculation. The implementation of this method will enable the routine simulation of microscopy images with charge transfer and bonding effects included in a systematic way.

- In the **third paper** [3], which makes up chapter 5, we introduce a new way to minimize the artifacts that arise in (S)TEM image simulations as a result of discrete sampling of the projected potentials. This method is particularly useful for projected potentials stored on the regular Cartesian grids used in many PPW DFT codes.

- In the fourth and **final paper** [4], given in chapter 6, we bring together the methods presented in the preceding two papers and apply them–through their implementation in the PPW DFT code CASTEP–to the study of ptychographic (S)TEM measurements of monolayer hexagonal boron nitride.

- Finally, we will summarize all of our results and outline directions for future investigations in the **conclusions** section, chapter 7.

The contributions of the coauthors to the papers that constitute Chapters 3-6 of this work are as follows:

In Chapter 3, the dispersion and short range correlation corrections to the energy differences of the helium bubbles of varying densities were calculated using codes written by N. C. Pyper, who also performed the calculations of these quantities.

In Chapter 6, all of the experimental data and analyses thereof were performed by G. T. Martinez, P. D. Nellist, et. al. The simulated ptychographic reconstruction and histogram analysis of the resulting simulated image was performed by G. T. Martinez, whilst the raw data supplied as input for these simulations (SCF and IAM projected potentials) were calculated by myself.

In Chapters 3-6, the remaining coauthors contributed in a mainly supervisory fashion, offering advice and input via invaluable conversation.

Chapter 2

Background and literature review

In this literature review, relevant background information relating to electronic structure theory and electron microscopy will be given. The need for improving the quality of the electronic structure theory included in simulations of electron microscopy will be made apparent.

First we will point out the disparity in experimental results concerning electron energy loss spectroscopy (EELS) measurements of excitation energies and number densities of helium under extreme pressure and demonstrate an opportunity for theoretical calculations to shed light on this discrepancy. The theory that we found most useful for this purpose is Hartree-Fock theory; a section including a brief overview of this branch of electronic structure theory will also be presented as background.

Second, we will point out the complete lack of a streamlined technique for the incorporation of more advanced electronic structure theory into the simulation of (scanning) transmission electron microscopy ((S)TEM) images. To put this into context, we will briefly review the background of both image simulation techniques in the microscopy community and the use of density functional theory (DFT) in the electronic structure community, illustrating the strengths and weaknesses of the most widely used implementations of theory in both areas and pointing out some promising and exciting contributions that each field might make to the other.

2.1 Hartree-Fock theory

Hartree-Fock theory [5] is an approach to solving the many body Schrödinger equation, given in time-independent, non-relativistic form under both the Born-Oppenheimer

and clamped nuclei approximations by

$$\left[-\sum_{i}^{N}\frac{\nabla_i^2}{2}+\sum_{i}^{N}V_n(\boldsymbol{r}_i)+\frac{1}{2}\sum_{j}^{N}\sum_{i\neq j}^{N}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}\right]\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)=E\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N), \tag{2.1}$$

in which N is the number of electrons, E is the energy of the system *minus* the nuclear-nuclear interaction energy, and the electron-nuclear potential energy operator V_n is given by

$$V_n(\boldsymbol{r}_i)=-\sum_{I}^{M}\frac{Z_I}{|\boldsymbol{R}_I-\boldsymbol{r}_i|}, \tag{2.2}$$

where M is the total number of nuclei, and Z_I and $\boldsymbol{R}_I$ are the atomic number and position of the I-th nucleus, respectively [6, 7]. We have made use of Hartree atomic units, as is common in electronic structure theory. Equation (2.1) governs the electronic structure of many systems with more than one interacting electron.

Hartree-Fock theory is a mean-field theory; the interaction of each electron with all the others is approximated as the interaction with the mean electric field of all the others. As Slater pointed out in [8], this approximation can be derived by assuming that the wave function in (2.1) takes the form of a Slater determinant of orbital functions ψ,

$$\Psi(\boldsymbol{r}_1,...,\boldsymbol{r}_n)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_1(\boldsymbol{r}_1) & \psi_2(\boldsymbol{r}_1) & \cdots & \psi_n(\boldsymbol{r}_1)\\ \psi_1(\boldsymbol{r}_2) & \psi_2(\boldsymbol{r}_2) & \cdots & \psi_n(\boldsymbol{r}_2)\\ \vdots & \vdots & \ddots & \vdots\\ \psi_1(\boldsymbol{r}_n) & \psi_2(\boldsymbol{r}_n) & \cdots & \psi_n(\boldsymbol{r}_n)\end{vmatrix}. \tag{2.3}$$

Slater determinants take into account the fact that the wave function for the electrons must be antisymmetric. After some mathematical manipulation, substituting (2.3) into (2.1) yields the following eigenvalue equation for each orbital ψ_i:

$$\begin{aligned}&\left[-\frac{\nabla^2}{2}+V_n(\boldsymbol{r})+\frac{1}{2}\sum_{j}^{N}\int d\boldsymbol{r}'\frac{|\psi_j(\boldsymbol{r}')|^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]\psi_i(\boldsymbol{r})\\ &-\frac{1}{2}\sum_{j}^{N}\int d\boldsymbol{r}'\frac{\psi_j(\boldsymbol{r})\psi_j^*(\boldsymbol{r}')\psi_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}=\epsilon_i\psi_i(\boldsymbol{r}).\end{aligned} \tag{2.4}$$

By invoking the variational principle and subjecting the set of eigenvalue equations to the appropriate constraints, Ψ can be solved for in a self-consistent manner. This entails, conceptually, starting with a guess for the functional form of the set of orbitals $\{\psi_i\}$, solving the set of eigenvalue equations–or Hartree-Fock equations–for that guess, constructing the wave function Ψ from the new solutions, and then solving (2.1) for

Ψ. If the new value of the energy E differs from the previous value by more than a prescribed amount, the process is carried out again. Each cycle through this process is called a self-consistent field (SCF) loop, and the whole program carries out as many SCF loops as necessary until the value of E has converged.

The exchange interaction due to Pauli repulsion of the electrons, which is embodied in the last term on the LHS of equation (2.4), is accounted for exactly by this theory. However correlation interactions–the dependence of the wave function of an electron on the probability of finding another electron in its vicinity–are completely neglected. A way of including correlation effects while still retaining a mathematical form similar to the Hartree-Fock approach was developed in the early 1960s. We turn now to a brief outline of this theory.

2.2 Density functional theory

DFT is another theoretical approach to solving the many-body problem. In contrast to Hartree-Fock theory, it is a way of mapping the actual fully-interacting system of electrons to a fictitious, mean-field system of non-interacting electrons with a different potential but an identical charge density. The electrons of the fictitious system are treated in a mean-field fashion, but the error of this approximation is accounted for– at least in principle–by the use of an exchange-correlation energy functional, $E_{xc}[\rho]$, in the Hamiltonian.

DFT has its roots in Thomas-Fermi theory [9] in which the electron charge density,

$$\rho(\boldsymbol{r}) = \int d\boldsymbol{r}_1 ... \int d\boldsymbol{r}_{N-1} |\Psi(\boldsymbol{r}_1, ..., \boldsymbol{r}_N)|^2, \tag{2.5}$$

is the principle quantity of interest. An advantage of solving for $\rho(\boldsymbol{r})$ rather than the full wave function is that it depends on $\boldsymbol{r}$ only and so is a function of three variables whereas the wave function is a function of $3N$ variables, N being the number of electrons. It is apparent why it would be computationally advantageous to deal with $\rho(\boldsymbol{r})$. Moreover, Hohenberg and Kohn [10] proved in 1964 that the ground state wave function of the many-body Schrödinger equation is a unique functional of $\rho(\boldsymbol{r})$. Consequently, since all other observables of the ground state system are functionals of the wave function, they are functionals of the ground state density. The issue then becomes accurate determination of $\rho(\boldsymbol{r})$. In 1965 Kohn and Sham [11] introduced a set of self-consistent equations and a scheme for obtaining a solution for $\rho(\boldsymbol{r})$ from them, relying on the variational principle. This set of equations and the method

for solving them resemble those of Hartree-Fock theory. The eigenvalue equations of DFT–the Kohn-Sham equations–take the form [6, 7]

$$\left[-\frac{\nabla^2}{2} + V_n(\boldsymbol{r}) + \frac{1}{2}\sum_j^N \int d\boldsymbol{r}' \frac{|\psi_j(\boldsymbol{r}')|^2}{|\boldsymbol{r}-\boldsymbol{r}'|} + V_{xc}(\boldsymbol{r})\right]\psi_i(\boldsymbol{r}) = \epsilon_i\psi_i(\boldsymbol{r}), \tag{2.6}$$

where V_{xc} is the exchange-correlation potential given by the functional derivative

$$V_{xc}(\boldsymbol{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho}\Big|_{\rho(\boldsymbol{r})}. \tag{2.7}$$

It was shown in [10] that there exists an exact energy functional such that the ground state density obtained by solving the Kohn-Sham equations is identical to the true ground state density of the exact fully interacting many body system. This universal functional has not been found, but has to be approximated.

The first functional to be introduced was the local density approximation (LDA) [10], which is derived from properties of the homogeneous electron gas and gives the exact answer for this system. Generalized gradient approximations (GGAs) were later developed [12], which can be thought of as somewhat akin to adding the next term in a Taylor expansion to the LDA. These functionals are derived from the properties of the slowly varying electron gas. One of the most popular GGAs is the PBE functional [13]. Generally speaking these functionals better accommodate systems with more rapidly varying electron densities. There are many other functionals that have been developed, far too many to list here, some of which are also semi-empirical. The functional of choice for the purposes of this thesis was PBE.

The main theoretical approximation involved in DFT is the use of an approximation to the exchange-correlation functional. However, there are additional approximations that must be made from a computational standpoint, to which we now turn our attention.

Many DFT codes developed for applications in solid-state physics and materials science represent the orbital functions ψ as superpositions of plane waves, summed over wave vectors $\boldsymbol{G}$. This representation of the wave functions stems from Bloch's theorem [14], which states that the wave function of an electron in a periodic potential can be expressed as the product of a function with the same periodicity as the potential and a function which simply modulates the phase. The upper limit of the sum over the wave vectors $\boldsymbol{G}_{cut}$ is given in terms of a cutoff energy $E_{cut} = \frac{1}{2}|\boldsymbol{G}_{cut}|^2$, and defines the size of the basis set used in this type of DFT calculation.

The periodicity of the wave functions in the interstitial regions of a solid are represented quite accurately by a linear combination of relatively few plane wave basis

functions, i.e. a very low cutoff energy. However, the rapid variation and nodal behaviour of the wave functions near the atomic cores demands a much larger plane wave basis set for accurate representation, greatly increasing the computational cost of a DFT calculation. In order to overcome this problem, pseudopotential theory was developed. Originally introduced by Fermi [15] as a theoretical tool in neutron scattering, the basic idea in the application of pseudopotentials to plane wave DFT is to replace the divergent atomic core potential within some cutoff radius $\boldsymbol{r}_c$ with a fictitious, smoother potential when solving for the valence states. This gets rid of the undesirable steep gradients and nodal behaviour of the valence wave functions within the cutoff radius, but reproduces the exact wave functions and scattering properties outside of it, for each angular momentum state separately. In these types of calculations, the core electrons are effectively frozen in their atomic state for the duration of the calculation. Since most properties of interest depend on the behavior of the electrons in the interstitial regions, and the core electrons are assumed to be relatively insensitive to the chemical environment, this is an acceptable approximation in many cases. DFT programs which use a plane wave basis together with pseudopotentials are referred to as pseudopotential plane wave (PPW) codes, with the class of codes that does not use the pseudopotential approximation being referred to as all-electron (AE).

The two most popular types of pseudopotentials are norm-conserving [16] and ultrasoft [17]. One of the constraints imposed on the generation of pseudopotentials in the norm-conserving scheme is that the norm of the pseudo-wave functions within the cutoff radius is kept identical to the norm of the all-electron wave functions. If we denote the pseudized charge density of the i-th state as $\tilde{\rho}_i(\boldsymbol{r})$ and the all-electron charge density as $\rho_i(\boldsymbol{r})$, then this constraint is equivalent to enforcing $\int_0^{\boldsymbol{r}_c} \tilde{\rho}_i(\boldsymbol{r})d\boldsymbol{r} = \int_0^{\boldsymbol{r}_c} \rho_i(\boldsymbol{r})d\boldsymbol{r}$, and so can be thought of as a redistribution of the charge within the cutoff radius. In the ultrasoft scheme, this constraint is relaxed.

DFT has an advantage over other theoretically exact methods, such as coupled cluster theory which approaches the exact wave function as the basis is expanded and is limited only by the size of the basis set [6], in that it is much less computationally expensive. The computational cost is such that methods like coupled cluster theory can only be used on systems with a small number of electrons. They can offer useful theoretical benchmarks in smaller systems for different variations of DFT, and then DFT can be applied to much larger systems at much lower cost.

In addition to being computationally efficient, DFT is precise and accurate over a range of materials. With regard to accuracy, charge densities calculated using

DFT have compared well with extremely precise measurements of electronic charge densities obtained from electron and X-ray diffraction experiments for some test cases. For example, R-factors (which are quantities commonly used in crystallography as measures of the disagreement of structure factors) of less than 1% were found for silicon and germanium in [18]. In terms of precision, a recent comprehensive study [19] demonstrated the reproducibility of results across a very wide range of completely independent DFT software. Its accuracy, precision, and computational efficiency make DFT a solid candidate for application to the problem of including electronic structure theory in electron microscopy image simulations.

Having outlined the basic ideas behind DFT and Hartree-Fock theory, we move on to a discussion of how these might be applied in new ways within the field of electron microscopy.

2.3 EELS, electronic structure, and confined atoms

One of the many areas of research in which electronic structure theory has been successfully applied is the modelling and interpretation of electron energy loss spectroscopy (EELS) experiments [20]. In transmission EELS, electrons of a known kinetic energy are fired into a sample and the intensity of transmitted electrons as a function of the amount of energy lost is obtained. This technique can be used alongside CTEM or STEM methods. The range of energies of the incident electrons must be kept very narrow–this spread is what determines the energy resolution of the EELS experiment. The transmitted electrons are passed through a spectrometer which uses a magnetic field to separate the electrons according to their final energies. The intensity is then measured as a function of final kinetic energy. The intensities of the energy loss over a range of energies is then known to the degree that the initial kinetic energy of the electrons is controlled.

Many processes contribute to the energy loss spectra. Here we consider only a few in describing the qualitative features of a typical energy loss spectrum. At 0 eV, there is a sharp maximum called the zero-loss peak which is the intensity of the portion of the electrons that were transmitted through the sample without losing any appreciable amount of energy. Excitations of phonon modes typically result in extremely small amounts of energy loss that used to be considered part of the zero-loss peak [20]. However, with recent improvements in both the spatial and energetic resolution [21–23], these can be measured and studied in detail; see for example [24]. In most materials, the next feature is the plasmon peak that results from the

collective excitation of valence electrons and typically extends over the 4-40 eV range. Finally, the core-to-valence excitations cause smaller, more narrow edge features in the spectra, corresponding to the series of core-to-valence excitations characteristic of each of the elements that compose the specimen. These are the features of interest in chapter 3 of this thesis. Some of these edges can appear superposed atop the broader plasmon peak, with subsequent edges appearing over the remaining range of energies.

The electrons in the core states are not typically sensitive to the chemical environment, but the valence states to which they are excited are more strongly influenced by it and thus information about the bonding of atoms in materials can be obtained from EELS. When used alongside high resolution STEM, precise spatial and energetic information can be obtained simultaneously. For these reasons EELS has been used in the study of confined excited atomic species [25, 26]. A particular case that has received attention due to its importance in the safety of nuclear fission reactors is the measurement of the excitation energies of helium confined under extreme pressures. EELS is used to measure the excitation energy of helium atoms trapped in bubbles within metals. This energy is always higher than that of a free atom. The difference between the excitation energy of the atom in the cluster and the free atom is defined as the energy shift, (ΔE). ΔE is a function of the number density (n) of the helium trapped in the cluster. There are two main components of these experiments, then: (i) the measurement of ΔE, and (ii) the calculation of n which involves using experimentally measured quantities in conjunction with some theoretical assumptions. The goal of these challenging experiments is to extract a simple relationship between the two quantities so that the more easily obtained quantity ΔE would be all that needed to be measured in the future, and values for n could just be read off.

By around the year 2000, several theoretical approaches had been used to calculate ΔE as a function of n from first principles [27–30]. At this time, the experimental data was limited to only a few data points at small values of n [27,31]. Given the difficulty of the experiments, the idea was to develop a theoretical technique that could match the existing data and then extend over a wide range of n values. DFT was found to be inadequate for this purpose; in [32] it was shown that DFT (in over 18 different variations) gave very poor results when used to predict the binding energies of noble gas dimers, even when including Van der Waals interactions. However, the method of [33], which employed the Hartree-Fock plus short-range correlation plus damped dispersion (HFSCDD) technique, was successful in matching the existing data and offered an advantage over the other theoretical approaches in that it provided more straightforward physical insight into the nature of the relationship between ΔE and n.

The HFSCDD method uses generalized valence-bond Hartree-Fock theory to obtain the Hartree-Fock energy and wave functions. These wave functions are then used to calculate two corrections to the Hartree-Fock energy: those arising from the damped inter-atomic dispersion potential V^{disp} [34] and the short-range correlation potential V^{SC} calculated using the theory of Gordon and Kim [35]. Thus the potential of two atoms a and b separated by a distance R in the HFSCDD method is

$$V_{ab}(R) = V_{ab}^{HF}(R) + V_{ab}^{SC}(R) + V_{ab}^{disp}(R). \tag{2.8}$$

Since the publication of [33] in 2001, many more experiments have been performed on helium bubble systems [25, 36–40]. However, no consistent agreement has been found. Many of these experiments attempt to find a linear fit to ΔE vs. n data, and the values obtained for the slope C of the fitted line $\Delta E = C\rho + D$ vary across the literature from 0.016 to 0.044 eV·nm^3 [25, 37–39]. In [1] (which forms chapter (3) of this thesis) we apply the same HFSCDD method of [33] to a wider range of systems and compare this with the larger body of available experimental data with the aim of shedding light on the nature of the disagreement in the experimental literature. We also discovered in this analysis that the method used in nearly all of the experimental literature to calculate n from measured quantities, taken from [31], neglected the effects of the overlapping wave functions of the atoms. We show in [1] that this is not accurate for the high densities n in question, and present a theoretical method to correct for this.

2.4 Modelling and simulation in (S)TEM

The first conventional transmission electron microscopy (CTEM) image was obtained in 1931 by Max Knoll and his doctoral student Ernst Ruska [41]. The basic principle was to use fast electrons rather than visible light to visualize samples in order to achieve resolution much higher than that of conventional light microscopes. The first working CTEM machines had approximately the same resolution as light microscopes, but were an important proof of principle. By 1933 Ernst Ruska had improved the magnetic lens enough to push past the resolution limit of light microscopes. However, many more technical improvements would be necessary before anything like atomic resolution could be achieved. One of the most important improvements was the development of techniques for correcting for some of the effects of the aberrations due to the magnetic lenses of the microscope. After a significant amount of time and

effort, aberration correction finally pushed CTEM into the realm of atomic resolution in the 1990s (for a detailed account of this development, see [42]).

Scanning transmission electron microscopy (STEM) began in 1938 [43], soon after the development of the first CTEM apparatus, but was not fully developed until the work of Crewe, Wall, and Welter in [44]. In STEM, a focused convergent beam of electrons, called the probe, is passed through the sample at each of a number of probe positions as it is scanned across the specimen. STEM images are related to CTEM images through reciprocity [45]. There are several different experimental methods for performing STEM measurements, such as annular dark field (ADF) and electron ptychography. However, the basic physical principle behind all imaging modes of the microscope is the same: information about the specimen is gathered from the scattering of electrons from the specimen [46].

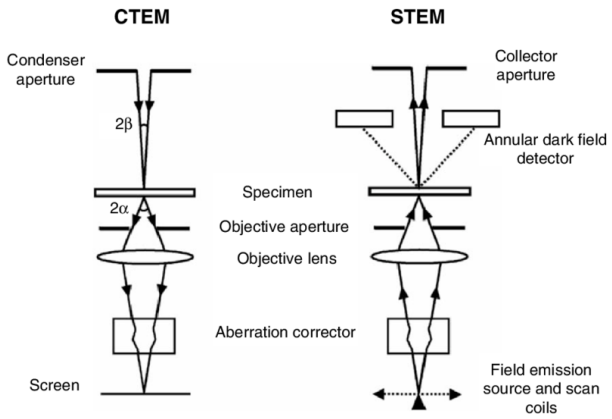


Figure 2.1: *Schematic showing ray diagrams of CTEM and STEM modes of the electron microscope. Image taken from [47].*

Projection assumption In order to accurately simulate images for comparison with (S)TEM experiments, we need to model the scattering of electrons interacting with the sample. The energies of the electrons in the beam, or "fast electrons," are typically on the order of 10^2 keV and are much higher than the energies of the electrons found in the specimen. Saldin has shown that the many-body exchange and correlation effects between the beam and specimen can be neglected for beam energies of over 10 keV [48]. In addition to neglecting the many body effects between the fast electrons and those within the sample, under certain conditions we can also simplify the problem by making the projection assumption [49]. This approximation, also known as the phase object approximation, is valid if (i) the sample is sufficiently thin

and (ii) the potential of the sample varies on length scales much larger than the wavelength of the fast electrons. To illustrate, consider a single plane wave representing the incident fast electron wave function Ψ_i in CTEM mode,

$$\Psi_i(x, y, z) = e^{ikz}, \tag{2.9}$$

where k is the wave number and the amplitude has been taken to be unity. The projection assumption states that the transmitted wave function Ψ_t after passing through the specimen will be

$$\Psi_t(x, y, z) = t(x, y)\Psi_i(x, y, z), \tag{2.10}$$

with t being the transmission function of the specimen

$$t(x, y) = \exp\{i\sigma V_z(x, y)\}. \tag{2.11}$$

In (2.11), the interaction parameter σ is given by

$$\sigma = \frac{2\pi me\lambda}{h^2}, \tag{2.12}$$

where m and e are the rest mass and charge of the electron, λ is the relativistic wavelength of the fast electrons, h is Planck's constant, and V_z is the potential of the specimen integrated, or "projected," through the z axis:

$$V_z(x, y) \equiv \int_{-\infty}^{\infty} V(x, y, z)dz. \tag{2.13}$$

Thus, under the projection assumption, the total effect of a thin specimen on an incident plane wave can be accounted for through the projected potential V_z. This carries over to STEM as well, since the probe can be thought of in terms of a superposition of plane waves. A more detailed derivation of the projection assumption will be given in chapter 4.

If the sample sufficiently thick, the multislice approach must be used [50], in which the electrostatic potential of the sample is partitioned into thin slices perpendicular to the incident beam, the projection assumption is made within each slice, and the transmitted wave function is propagated between slices with the Fresnel propagator [51]. Thus, in both single slice and multislice calculations, the primary quantity of interest is V_z. Next, we give a brief overview of how the three imaging modes that will be studied in this thesis are modelled.

CTEM In the simulation of CTEM experiments, the z dependence in (2.10) is typically dropped since it will have no effect on the final image intensity. Thus we have our transmitted wave given simply by

$$\Psi_t(x, y) = t(x, y). \tag{2.14}$$

The final image intensity g is given by

$$g(x, y) = |\Psi_t \otimes h_o|^2, \tag{2.15}$$

where the convolution with the point spread function $h_o(x, y)$ embodies the effects of the optics dictated by the experimental parameters of the microscope.

ADF STEM To model ADF STEM experiments, the first step is to construct the probe function $\Psi_p(\boldsymbol{r})$ which represents the narrow, focused electron beam formed by passing the electrons through the lens *before* they pass through the sample. Here $\boldsymbol{r}$ is a two dimensional vector lying in the specimen plane. For a given probe position with coordinates $\boldsymbol{r}_p$, the effect of the sample on the beam is represented mathematically as a simple product of the probe function with the transmission function,

$$\Psi_t(\boldsymbol{r}, \boldsymbol{r}_p) = \Psi_p(\boldsymbol{r} - \boldsymbol{r}_p)t(\boldsymbol{r}). \tag{2.16}$$

Next, the Fourier transform of Ψ_t is taken with respect to $\boldsymbol{r}$ to obtain the wave function in the diffraction plane, $\widetilde{\Psi}_t$. The square modulus of the portion of this wave function that falls on the annular detector is then integrated to obtain the intensity of the pixel for that probe position,

$$g(\boldsymbol{r}_p) = \int\int |D(\boldsymbol{k})\widetilde{\Psi}_t(\boldsymbol{r}_p, \boldsymbol{k})|^2 d\boldsymbol{k}, \tag{2.17}$$

where $\boldsymbol{k}$ is the 2D vector in reciprocal space and $D(\boldsymbol{k})$ is the detector function, equal to unity for values of $\boldsymbol{k}$ on the detector and zero otherwise. The total image is then constructed pixel by pixel by repeating this process for each probe position, with the dimensions of the pixels being equivalent to the step-size of the probe in the x and y directions. Since only those electrons that scatter out to higher angles are collected in ADF STEM, the image primarily contains information about the potentials of the nuclei. Thus it is sometimes referred to as Z contrast imaging.

Electron ptychography in the STEM Ptychography, originally proposed in 1969 [52], is a method of solving the phase problem, which arises in electron microscopy due to the fact that in recording the intensity–which is proportional to the modulus of the wave function squared–phase information is lost [53]. To recover the phase information encoded in the transmitted probe function, electron ptychography [54] uses a probe that is defocused–that is, focused in a plane either slightly above or below the specimen plane–so that, in the specimen plane, the probe function is a bit wider than the step-size of the probe so that there is overlap between neighboring probe positions. As a diffraction pattern is recorded for each probe position, the overlap ensures that there is a redundancy in the information stored in the set of diffraction patterns. This redundancy can be exploited, through deconvolution of the dataset, to extract the relative phase [54–58].

As in ADF STEM, the transmitted wave function in the diffraction plane is given by the Fourier transform of the transmitted wave function with respect to the spatial coordinates $\boldsymbol{r}$,

$$\widetilde{\Psi}_t(\boldsymbol{r}_p, \boldsymbol{k}), \tag{2.18}$$

where Ψ_t is the same as in (2.17). The intensity of the microdiffraction pattern for each probe position is

$$g(\boldsymbol{r}_p, \boldsymbol{k}) = |\widetilde{\Psi}_t(\boldsymbol{r}_p, \boldsymbol{k})|^2, \tag{2.19}$$

which leaves us with a complex, 4D dataset. Forward Fourier transforming g with respect to the probe coordinates and invoking the convolution theorem yields

$$\widetilde{g}(\boldsymbol{r}_p, \boldsymbol{k}) = \widetilde{\Psi}_p(\boldsymbol{k})\widetilde{\Psi}_p^*(\boldsymbol{k} + \boldsymbol{q}_p) \otimes_{\boldsymbol{k}} \widetilde{t}(\boldsymbol{k})\widetilde{t}^*(\boldsymbol{k} - \boldsymbol{q}_p), \tag{2.20}$$

where $\otimes_{\boldsymbol{k}}$ represents the convolution with respect to $\boldsymbol{k}$. Backward Fourier transformation of (2.20) with respect to $\boldsymbol{k}$ yields a product of two Wigner distribution functions [58]. From there, deconvolution in conjunction with knowledge of the probe function and some additional Fourier analysis leads to a full reconstruction of $\widetilde{t}(\boldsymbol{q}_p)$, the Fourier transform of the transmission function as a function of the spatial frequency of the probe. This method is sometimes referred to as ptychographic phase reconstruction.

One major advantage of ptychography over other phase contrast imaging modes is that ADF STEM can be performed simultaneously, offering a way to extract both the coherent phase and the incoherent Z-contrast information from a sample. This is especially useful when applied to samples containing a mixture of heavy and light elements. Additionally, ptychography is: relatively robust since aberrations can be

corrected after the data is acquired; efficient for low dose (low-intensity electron beam) imaging of beam-sensitive samples; and can be used to extract 3D information [58]. The utility of DFT in the interpretation of electron ptychography experiments in terms of charge transfer will be studied in chapter 6.

Though all three of the above (S)TEM techniques vary significantly, the key quantity relating the image to the specimen is V_z in each case, via the transmission function t given in equation (2.11). We look next at a description of the most common method of computing V_z in the electron microscopy community.

Independent atom model In order to easily generate projected potentials for any arbitrary specimen on the fly, most multislice software makes use of an independent atom model (IAM). In the IAM, the electrostatic potential of the specimen is constructed from the superposition of potentials of isolated atoms. Analytic functional forms with several free parameters are fitted to results of Hartree-Fock or multi-configuration calculations of the radial charge densities and potentials of an isolated atom for each species on the periodic table [59]. These parametrized quantities are easily stored in libraries in electron microscopy codes, so that no actual electronic structure calculation is carried out on the fly when constructing the potential of the sample. There are multiple popular techniques for parametrizing IAM quantities [59–61], with accuracy improving over the decades. Regardless of the accuracy of these parametrizations, projected potentials constructed from the IAM will never be able to take into account the interaction of the electrons of one atom with all the others in a sample. These effects–often classified as "charge transfer" or "bonding" or referred to collectively as "charge redistribution" effects–can only be taken into account by obtaining projected potentials which result from applying many body theory to the specimen as a whole, rather than each isolated atom individually. For this purpose, DFT is well-suited.

We now turn to a survey of cases in the literature in which DFT has already been used in conjunction with (S)TEM simulation software to interpret experimental images. As we will see, several different approaches have been used to simulate and measure the electronic structure of materials. However, each study seems to use a computational approach specifically tailored to the individual application and a general approach for integrating electronic structure calculations with electron microscopy image simulations is completely absent from the literature, thus hindering the modelling capabilities of the electron microscopy community at large.

2.5 Modelling and measuring electronic structure in the electron microscope

While bonding effects can be probed with EELS, our focus in this section is on the effects of electronic structure on diffraction and imaging techniques in (S)TEM. As early as 1973, the influence of atomic bonding on electron diffraction experiments was being considered [62]. However it was not until the mid 1990s that the question of how important it was to consider the charge redistribution in ionic crystals in modelling high resolution TEM images began to attract more attention. This was driven by technological improvements in electron microscopy which had recently led to sub-resolution [63, 64]. By this time DFT had become a well-established method in condensed matter communities and there were several reliable software packages available which implemented the theory [65–67].

$\Delta\rho$ method In their case study of sapphire (Al_2O_3) in 1998, Gemming *et al.* [68] performed both simulations and experiments to determine if the ionic nature of sapphire had a significant effect on HRTEM images. This appears to be the first publication which included a comparison of original experimental electron microscopy data with image simulations that used electronic structure theory to go beyond the IAM. The authors used CASTEP [69]–a pseudopotential plane wave DFT code–to calculate the charge density difference between the IAM and the full DFT models. This charge density difference was then incorporated into the EMS software package [70] which was used to generate the simulated HRTEM images for both independent atom and DFT models (the technique of utilizing the charge density difference between DFT and IAM calculations as input for simulating the TEM images will be referred to as the $\Delta\rho$ method and it will be discussed in further detail below). The authors found that the difference between the images resulting from the IAM and the DFT charge densities was smaller than the difference between the experimental image and the image generated from the IAM. Thus it was concluded that the effects of the ionic bonds on the electronic structure of sapphire was not detectable in the electron microscope and the use of the default IAM was justified in this case. The authors then went on to speculate that this particular case represented a trend, and that the IAM would be sufficient for modelling electron microscopy images for other materials that exhibit ionicity. While their conclusion was later confirmed for the case of sapphire [71], studies have since shown that the IAM is insufficient for the image simulation of most of these types of materials in general [71–73].

The $\Delta\rho$ method used in [68] was also used in [71] in 2004 to simulate TEM images for MgO as well as an Ag-MgO interface. Two different DFT codes were employed to obtain the charge densities in [71]: CASTEP and DACAPO [74], another pseudopotential plane wave code. They also used the EMS software package to generate the images. Thus the computational approaches in both [68] and [71] were very similar. In [71] the authors found that the difference between IAM and DFT image simulations should be significant enough to be observed in the electron microscope, and found good agreement between images derived from charge densities from both DFT programs, even though different functionals were used–LDA in CASTEP and GGA in DACAPO. It appears that the authors wanted to assess whether the use of different functionals had a noticeable effect on simulated TEM images. However, due to the pseudopotentials employed in each code, they could only use LDA in CASTEP and GGA in DACAPO, whereas in an "ideal world" the two different functionals would have been compared within the framework of a single code. This is now easily achieved in CASTEP for example. It is also stated in [71] that, due to a lack of highly accurate experimentally measured charge densities to compare against for the interfaces they studied, a determination of which functional offered a better description of the charge density was not possible. This is slightly puzzling, since at least one extensive study in 1997 [75] had found that the GGA gave charge densities that were twice as accurate as those coming from the LDA for silicon.

At this point it is worthwhile to take a closer look at the $\Delta\rho$ method used in [68] and [71]. EMS uses the multislice algorithm, and so V_z must be obtained for each material under consideration. Since V_z is a functional of the charge density ρ, this density is in principle all that's needed from the DFT calculation. However, ultrasoft pseudopotentials were used in both the CASTEP and DACAPO simulations of [68] and [71]. The use of ultrasoft pseudopotentials adds an offset to ρ, so the authors had to come up with a way to get rid of this. The technique they used was as follows: DFT calculations were run for each atom of each species of the material in which only that single atom was present in the unit cell. The individual charge densities obtained for each of these calculations were then added together to obtain a sort of IAM charge density from the DFT code, ρ^{DFT}_{atomic}. Then a DFT calculation was carried out for the fully populated unit cell, and a charge density ρ^{DFT} was obtained. Since the core electrons are treated the same in each atom in all of these calculations, the difference $\Delta\rho = \rho^{DFT}_{atomic} - \rho^{DFT}$ contains only the difference in the valence charge density, and the offset in ρ arising from the use of ultrasoft pseudopotentials is cancelled.

The next step is the calculation of V_z within EMS. EMS uses electron scattering factors parametrized by the Doyle-Turner method [60] to describe the charge densities in an IAM approach by default–these charge densities will be denoted ρ^{IAM}. To incorporate the valence charge density from the DFT calculation into EMS, an approximation must be made, namely that the charge densities described by the Doyle-Turner parametrizations and those computed from the DFT codes are identical, except for the additional component in ρ^{DFT} arising from the use of pseudopotentials. This assumption allows the projected potential within EMS to be calculated from the charge density difference $\rho^{IAM} - \Delta\rho$. Of course this approximation is not entirely accurate, but the differences are assumed by the authors to be small enough so as to be negligible. However, this assumption was not verified in either [68] or [71]. Since the Doyle-Turner scattering factors are parametrizations of Hartree-Fock calculations on atoms, differences between charge densities extracted from DFT and Doyle-Turner will arise from the differences in the underlying theories as well as inaccuracies arising from the process of parametrization. We also note that an additional drawback of the $\Delta\rho$ method is that the source code of EMS must be edited in order to include this charge density difference in the generation of projected potentials [71].

SF method Another method for incorporating DFT into HRTEM image simulations was introduced by Deng and Marks [72] in 2006. Their approach was to calculate the electron charge density of the material using an all-electron DFT program (Wien2k [76]), obtain the X-ray structure factors by Fourier transforming this charge density, obtain the electron structure factors by applying the Mott-Bethe formula to the X-ray structure factors, and from these calculate the projected potentials and apply the multislice method to simulate the images. This general approach will be referred to as the SF method, for "structure factor." In [72], the SF method was used to simulate HREM images for a range of oxides, and comparison of these with IAM images led the authors to conclude that charge redistribution effects should be detectable in aberration-corrected microscopes. In a related paper [77], the SF method was used to model the effects of local charge defects in MgO on HREM images, and a similar conclusion was reached–the difference between the images generated using the IAM and the SF method was significant enough to suggest that these effects should be observable. Later, in 2011, Ciston *et al.* [78] used the SF method (using Wien2k for the DFT calculations and NUMIS 2.0 [79] for implementing the multislice algorithm) to study the effects of charge transfer in HREM image simulations of Forsterite (Mg_2SiO_4). Among the many important observations made in the careful

study of [78], it was predicted that it was not the spatial resolution, but the signal to noise ratio of the microscope that mattered most when evaluating whether charge transfer effects could be observed in an experiment. Another important conclusion drawn by these authors was that the electrostatic potential was sensitive to the parameters of the DFT calculations. They stressed the need for a fully converged DFT calculation, reasoning that a small difference in the electronic charge density could affect the screening of the atomic cores which would have a large impact on the potential. This contrasts the results of Kurasch *et al.* [80] who found that the projected potentials they calculated from Wien2k were "not very sensitive to the DFT input parameters." However, in [80] the materials under consideration were 2D as opposed to the bulk oxides considered in [78]. Perhaps the dependence of potentials on the parameters of a DFT calculation needs more thorough evaluation, and the question of whether the differing claims in [78] and [80] are due to the dimensionality of the materials or not remains to be answered.

The SF method benefits from the fact that the projected potential is calculated in reciprocal space (from the Fourier transform of the charge density) rather than in real space, which would be more expensive computationally since a very fine sampling of the potential in the atomic cores is needed. However this advantage is much less significant if one is attempting to model defects in crystals because of the large number of Fourier components needed to accurately represent local defects. A disadvantage of the SF method is the error introduced at small angles in the electron structure factors by using the Mott-Bethe formula to calculate them. In [72] this error is accounted for by using an approximate limiting form, which in turn leads to small errors in the inner potential, the authors explain. However it is claimed that this only really leads to a uniform phase shift of the exit wave, and so is not taken to be significant. Another problem that arises from the use of the Mott-Bethe formula is that it makes it difficult to use different Debye-Waller factors for different atoms in the unit cell, so in [72, 77, 78] a uniform Debye-Waller factor is used. It is also noted that, while an all-electron code was used in all three publications listed here who employed the SF method, this does not appear to be absolutely necessary to the technique.

RSP method In contrast to the $\Delta\rho$ and SF methods, several other sources [80–82] have used real space electrostatic potentials obtained from DFT calculations as input for their TEM image simulations. This method will be referred to as the real space potential (RSP) method. In this technique, the RSP is either calculated from the charge density of a DFT calculation, or simply taken straight from the output of a

DFT calculation, projected through the z-axis, and used as the input for a multislice simulation.

The effects of charge transfer on annular dark field (ADF) STEM images of a wide range of crystals were examined in [82] and [73] in 2016. ADF experiments extract information from the electrons that are scattered through larger angles, and thus depend more heavily on the interactions of the electrons in the beam with the atomic cores in the sample. In [82] ADF-STEM image simulations were performed for C (diamond), w-BN, c-BN, AlN, BeO, MgO, LiF, and NaF. In general, it was found that polar bonding had a visible effect on the simulated ADF-STEM images, over a wide range of instrumental parameters. This was confirmed experimentally for AlN in [73]. For the DFT calculations, Quantum ESPRESSO–a pseudopotential plane wave code–was used [83]. Charge densities were calculated for 3 different models of each material: IAM, a bonded crystal model (BCM), and a fully ionized model (FIM). The IAM was calculated as the superposition of isolated atom DFT calculations for each species. The BCM calculation was a superposition of atomic core densities from the IAM calculations and valence charge densities from a typical SCF DFT calculation for the full unit cell of the material. In the FIM, densities were calculated for independent full-valence-shell ions and then a linear superposition of these was taken. The IAM and FIM were limiting cases, with the more accurate BCM always lying somewhere in between. Projected potentials were calculated from the charge densities for each of these models and used as direct input for the TEMSIM [61] multislice code for the simulation of images. In order to incorporate thermal effects in their simulations, the frozen phonon approach was used. Thermal effects are very important in ADF-STEM. However, a complication arose in [82] when the inclusion of these effects was considered. In order to allow TEMSIM to calculate frozen phonon configurations, the projected potentials from Quantum ESPRESSO had to be parametrized in such a way that spherical symmetries of the core potentials were enforced. Thus some of the accurate physics of the bonds in the materials was lost when thermal effects were included. In addition to this parametrization, a truncation of the potentials near the nuclei was also necessary. Despite these complications, good agreement was found between experimental and BCM images for AlN [73], adding credibility to the use of pseudopotential plane wave DFT in the simulation of STEM images.

In [80] and [81], 2D materials are exclusively investigated, using the RSP approach. In [81], the authors compared simulation with experiment for HRTEM images in two cases: a nitrogen substitution point defect in graphene and monolayer hexagonal boron nitride (h-BN). It was found that the inclusion of bonding effects in the image

simulations was necessary for the correct interpretation of the experimental data. In h-BN, the experimentally measured contrast between boron and nitrogen atoms was much less than that predicted by the IAM approximation, but correctly modelled by DFT. The IAM also predicts that a nitrogen substitution in graphene should not be detectable in these experiments, but in actuality bonding effects do enable its detection, as is correctly predicted by DFT-based simulations.

The details of the methods used in [81] to simulate the TEM images using DFT potentials were elaborated on in [80], and a few image simulations on additional systems were also carried out, but this time not compared with experiment. The new simulations performed were for nitrogen and oxygen substitutions in graphene, and an oxygen adatom on graphene, and found that the inclusion of charge redistribution effects in the image simulations were crucial in these cases. The RSP method as implemented in [81] and [80] made use of two separate DFT software packages. The first, VASP [67], is a pseudopotential plane wave code and was used to perform the structural relaxation calculations. This type of calculation is essential for the accurate modelling of defects in materials, and is generally less computationally expensive in a plane wave pseudopotential code than in an all electron implementation of DFT. However, the authors were unable to extract the electrostatic potentials from VASP. They used the relaxed structure as input for WEIN2k, from which they were then able to extract the all-electron potentials. The use of two separate implementations of DFT–one for the structural relaxation and another for obtaining the potentials– is an obvious complication. The inability to access the total electrostatic potential is perhaps an issue specific to VASP rather than being general to pseudopotential plane wave implementations, since Odlyzko *et al.* were able to compute potentials from the total charge density in Quantum ESPRESSO without any reported issues in [82]. A further complication in the construction of DFT projected potentials for use in a multislice program explained in [80] has to do with sampling near the nuclei. Some typical IAM potentials such as Doyle-Turner are smoothed out near the nuclear singularity before they are interpolated onto a grid and used in a multislice simulation. However the authors were not able to smooth the divergent nuclear potentials from WIEN2k, as they did not have access to the necessary software components. They were thus forced to use a very fine grid spacing in the z-direction of the unit cell when computing DFT potentials to maintain accuracy in projecting the potentials. They also had to impose a cutoff for the maximum value of the potential in a pixel–a "very crude approach" in their own words. They justified this by pointing out that interaction with these pixels would be rare enough not to affect the overall images

in this type of study. Treatment of the potential near the nuclei would presumably become more important when considering ADF-STEM images, such as in [82].

Within the RSP approach then–which is arguably the simplest of the three general techniques outlined in this section–there appears to be room for improvement in DFT software. To model defects with minimal computational cost, a pseudopotential plane wave code would be desirable. Direct access to projected electrostatic potentials and a consistent and accurate handling of potentials close to the nucleus would also be needed. Software development along these lines would facilitate a greater number of users in the microscopy community to easily simulate images which include the effects of atomic bonding and charge transfer, and thus enable a great many more important studies into the electronic structure of materials through electron microscopy. For example, this type of software would be very useful coupled with a technique like ptychography which is particularly well-suited for observing charge transfer.

Diffraction DFT has also been used in interpreting experimental electron diffraction and X-ray diffraction (XRD) experiments [84–86]. In these studies, structure factors of different materials are measured using diffraction techniques and compared to the structure factors computed from IAM and DFT charge densities. The importance of accounting for the redistribution of charge (using DFT) has been demonstrated in these experiments. A disadvantage of diffraction experiments is that local defects cannot be studied since they rely on the periodicity of the sample [81]. One exception is the parallel recording of dark-field images (PARODI) technique [84, 87, 88] which combines diffraction and images, and is thus capable of accommodating local defects.

A Convergent beam electron diffraction (CBED) pattern is the square modulus of the Fourier transform of the transmitted wave function in a STEM experiment. A single 2D CBED pattern is obtained for each probe position. Likewise, a selected area electron diffraction (SAED) pattern is the analogous quantity in CTEM mode. CBED and SAED patterns can be directly calculated from projected potentials, such as those obtained from the method outlined in chapter 4 of this thesis. For XRD simulations, structure factors are needed. These can be obtained from a Fourier transformation of the quantities obtained from the method in chapter 4.

DFT under the microscope In addition to the many experiments which have compared structure factors measured with electron diffraction and XRD techniques to those calculated from DFT in order to better interpret experimental data, there

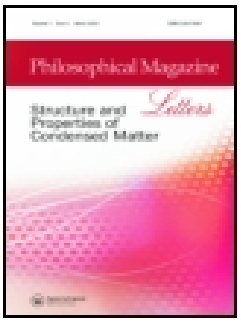
have also been diffraction studies the other way round which have compared the performance of different DFT functionals in reproducing highly accurate experimentally measured structure factors for a few materials [75, 86, 89].

This is a very important test of the theory, since the principle quantity dealt with in DFT is the electron density, and structure factors are essentially just Fourier transforms of this. How accurately this density is calculated is a direct measure of the reliability of various functionals and other approximations used within DFT.

Most new functionals are designed so that the total ground state energies of different systems are accurately reproduced. In a recent article entitled *Density functional theory is straying from the path toward the exact functional* [90], the need to test the charge densities–rather than just the energies–produced by the growing number of available functionals is stressed. The ground state charge densities calculated from 128 different functionals for several different types of single atoms were compared against nearly exact theoretical CCSD-full calculations. When testing the oldest functionals up through some that were developed around the early 2000s, it was found that agreement between the DFT and the CCSD charge densities generally improved over time. Since the early 2000s however, newer functionals generally compare worse with the CCSD densities than the older ones. Nearly exact theoretical charge density calculations obtained from other methods solving the many body problem–such as CCSD–are computationally intensive and thus are only available for rather simple systems. To test the accuracy of DFT when applied to more complicated systems, comparison with experiment is the only option. As has been previously discussed, most experimental measurements on larger systems involve electron diffraction which is not suitable for systems with defects. In addition to PARODI, perhaps highly accurate STEM ptychography experiments could one day be compared with DFT calculations to test how well different functionals reproduce charge densities of systems with defects.

Chapter 3

Excited helium under high pressures in the bulk and in nanobubbles



Philosophical Magazine Letters

ISSN: 0950-0839 (Print) 1362-3036 (Online) Journal homepage: http://www.tandfonline.com/loi/tphl20

Excited helium under high pressures in the bulk and in nanobubbles

N. C. Pyper, T. C. Naginey, P. D. Nellist & Colm T. Whelan

To cite this article: N. C. Pyper, T. C. Naginey, P. D. Nellist & Colm T. Whelan (2017): Excited helium under high pressures in the bulk and in nanobubbles, Philosophical Magazine Letters, DOI: 10.1080/09500839.2017.1347724

To link to this article: http://dx.doi.org/10.1080/09500839.2017.1347724

Published online: 05 Jul 2017.

Submit your article to this journal

Article views: 15

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=tphl20

Download by: [the Bodleian Libraries of the University of Oxford] **Date:** 31 July 2017, At: 05:50

PHILOSOPHICAL MAGAZINE LETTERS, 2017
https://doi.org/10.1080/09500839.2017.1347724

Taylor & Francis
Taylor & Francis Group

Check for updates

Excited helium under high pressures in the bulk and in nanobubbles

N. C. Pyper[a‡], T. C. Naginey[b‡], P. D. Nellist[b] and Colm T. Whelan[c]

[a]University Chemical Laboratories, University of Cambridge, Cambridge, UK; [b]Department of Materials, University of Oxford, Oxford, UK; [c]Physics Department, Old Dominion University, Norfolk, VA, USA

ABSTRACT

We systematically investigate the effects of intense pressures on the excitation energies of helium trapped in bubbles in order to deepen our understanding of the fundamental physics of atoms in extreme conditions. The $1s^2 \rightarrow 1s\,2p$ excitation energy of a confined helium atom is known to differ from that of a free atom being greater in both the bulk liquid or solid or a bubble confined in a metallic matrix state. We compare calculations for the energy shift with both laboratory experiments for bulk systems and results derived from scanning transmission electron microscope (STEM) studies of helium nanobubbles embedded in different matrices. We find excellent agreement between our calculations and the latest extensive measurements in the bulk. However, we find significant discrepancies when we compare with results deduced using the 'standard' approach for analysing STEM data. Here, we show the scattering matrix element determining the intensity of this excitation in a STEM experiment is significantly affected by the same environmental factors that shift the excitation energy. Consequently, there is a serious theoretical inconsistency in the way the STEM results are calculated, in that the 'standard' approach depends on a supposedly known $1s^2 \rightarrow 1s\,2p$ scattering cross section, whereas we show here that this cross section is itself dependent on the environment. Correcting for this inconsistency does not, in itself, improve agreement.

ARTICLE HISTORY
Received 27 October 2016
Accepted 22 June 2017

KEYWORDS
Electronic structure; ab initio; computer modelling; EELS

1. Motivation

The interaction of electrons with matter provides a powerful way to characterize materials. Electron energy-loss spectroscopy (EELS) in the scanning transmission electron microscope (STEM) allows electronic excitations of materials to be explored at high spatial resolution. Both the energy and strength of an excitation can be measured [1]. In most cases, quantitative interpretation of such data requires comparison to calculations. The excitation energies and cross-section values are often calculated making the assumption of isolated atoms (see for example [2]), whereas for most materials the atoms are in a different condition to the isolated state leading to substantially different excitation energies and cross

CONTACT T. C. Naginey timothy.naginey@materials.ox.ac.uk
‡Both authors contributed equally to this work.

© 2017 Informa UK Limited, trading as Taylor & Francis Group

sections. To improve the precision to which EELS data can be quantified, it is therefore important to include the environment of the atoms in the calculation.

The study of confined atoms is known to be of both fundamental and technological significance [3]. An important example of the latter is the formation of helium bubbles in the metallic components of either existing fission or proposed fusion nuclear reactors. In a reactor, the in-service materials will experience bombardment from radiation produced by nuclear transformations. Some of these radiations result in the formation of helium atoms within the metal. Continual bombardment causes these atoms to aggregate into bubbles trapped inside the metal where they exert very high pressures. The occurrence of such high-pressure bubbles weakens the structure of the metal in these components and would, at a sufficiently high pressure, cause the material to crack [4]. Knowledge of the formation of such helium bubbles, especially of the pressures they exert, is needed to ensure that reactors are still safe after several years of operation. Since it is not possible to measure the bubble pressures directly, these must be deduced from a reliably known and tested equation of state which requires as input the bubble density or equivalently the masses and volumes of the bubbles. The pressure is highly sensitive to the density [5,6], so it is important to determine the density to the highest possible accuracy. Although reliable equations of state are available and information about bubble dimensions has been derived by (S)TEM experiments, it is not possible to measure directly either the mass or the density of helium gas in the bubbles.

It is known that the $1s^2$ to $1s2p$ helium atom excitation energy depends on its environment, this being revealed through STEM experiments. We show here that such dependency not only extends to the scattering cross section but also that the latter is very significantly more environmentally sensitive than the excitation energy. Consequently, it should be expected that there are significant errors in previous interpretations of STEM/EELS experiments based on the assumption that the scattering cross sections are the same as that of an isolated helium atom.

2. Properties of helium atoms in bubbles and bulk condensed phases

The ground ($1s^2$) to $1s2p$ excitation energy in a bubble differs from that of 21.218 eV in a free helium atom being greater in both, a bubble as well as in the bulk liquid or solid states. In previous work, both the excitation energy and the strength of the inelastic scattering due to that excitation have been measured with the object of deriving the helium density in a bubble [5,7–10].

The excitation energy shifts (ΔE) can only be trusted to provide a reliable determination of a bubble density and hence pressure after the quantum mechanical predictions have been validated against experimental measurements for systems in which the bubble densities are known from entirely independent sources. Until recently, only two such experimental data points were available: one for the bulk liquid [11] and the other for the 'solid' [12]. In [13], excellent agreement with these results was found by using ab initio electronic structure computations, each consisting of a self-consistent field (SCF) calculation augmented by dispersion and short-range correlation contributions. The energy shift was explained by considering one of the helium atoms to be surrounded by a cluster of ground-state helium atoms. When the central helium atom is excited into the $1s\ 2p$ state, the $2p$ orbital on the central atom overlaps the electron densities of the surrounding atoms. The Pauli

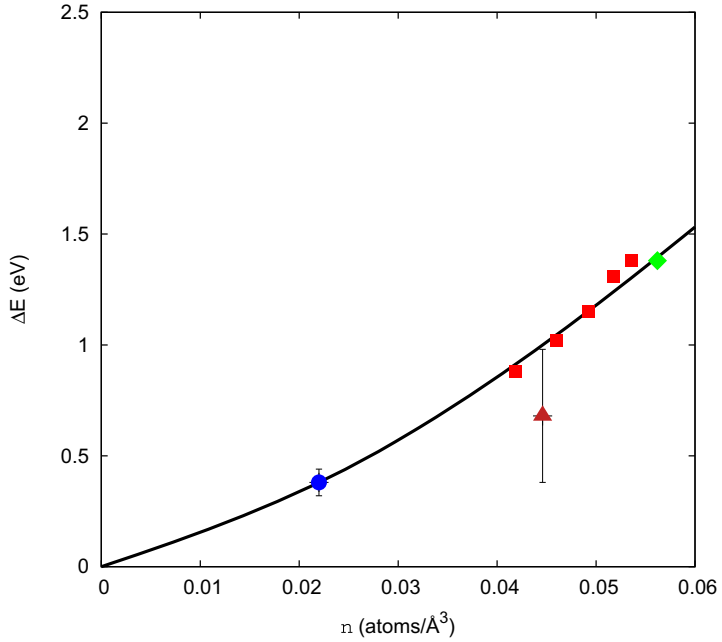


Figure 1. Comparison of our ab initio energy shift calculations using the 'extended' model as described in [13] (solid line) with experimental data for the bulk liquid from [11] (circle) and [15] (squares) as well as for the bulk solid from [12] (triangle) and [14] (diamond).

principle requires that the wavefunction describing all the electrons in the system should be antisymmetrized. When the 2*p* orbital is orthogonalized to the 1*s* orbitals on the surrounding atoms, extra nodes are introduced into the 2*p* orbital, which increases its energy. The energy of the cluster can be reduced by using a more contracted 2*p* orbital but the energy of the cluster is still greater than before the Pauli principle was taken into account.

In [13], the SCF wave functions were calculated in two models: a 'basic' model in which the excited 2*p* electron remains on the central helium atom with all neighbouring atoms in the free-atom ground state, and an extended model in which the possibility of delocalization of the 2*p* electron onto the surrounding atoms in the cluster was included. For the full details of the models used in our computations, the reader is referred to [13]. More extensive and accurate experimental data have been reported [14,15] since the publication of [13]. We have repeated and confirmed the calculations in our earlier work and extended their range to allow for comparison with the more recent experimental results, and have found excellent agreement between theory and experiment, shown in Figure 1.

At the time, 2001, of the publication of [13], there was only one experimental investigation which measured the properties of individual bubbles. The three quantities directly measured were the energy shift, the bubble diameter d and the ratio $\frac{I_p}{I_Z}$ where I_z is the unscattered beam intensity and I_p is the intensity in the He 1s $\rightarrow$ 2p peak. They presented values of the number density n using the relation [16]

$$n = \frac{I_p}{\sigma_{2p} I_z d} \tag{1}$$

by assuming a scattering cross section σ_{2p} that is derived from the free helium atom wavefunctions [17]. More recent measurements of ΔE for an extensive range of bubbles [7–9] differ very significantly from the bulk helium ΔE if the bubble densities are derived using the free atom sigma in (1). For example, shifts of over 2 eV are reported [8] for bubbles with number densities derived using the free helium cross sections to be around 0.04 atoms/Å^3 whilst for this density in the bulk ΔE is only found to be of the order of 0.9 eV.

The results from our previous calculations showed that the mean squared radius of the $2p$ orbital was significantly decreased with increasing density. Thus the free atom $\langle r^2 \rangle$ of 31.56 a.u. is lowered to 15.05 a.u. for $n = 0.05$ atoms/Å^3. These results indicate σ_{2p} might be significantly modified by its environment in both the bulk and the bubbles. In the first Born approximation, the cross section σ_{2p} for excitation of the target atom from the ground state to the $2p$ excited state by a fast incident electron is given in [17,18] by the equation:

$$\sigma_{2p} = \frac{2\pi}{E_i \Delta U_n} \int_{q_{\min}}^{q_{\max}} \frac{f_{2p}(q)}{q} dq. \tag{2}$$

here E_i is the energy of the incident electron in the STEM experiment, ΔU_n is the energy required to excite a ground state helium atom to its $1s\,2p\ ^1P$ state and $q_{\min}$ and $q_{\max}$ are the miniumum and maximum values of the momentum transfer as given in [1,17,18]. In (2) $f_{2p}(q)$ is the generalized oscillator strength (GOS)

$$f_{2p}(q) = 2\Delta U_n \frac{|\epsilon_{2p}(q)|^2}{q^2}, \tag{3}$$

and $\epsilon_{2p}(q)$ is the atomic form factor, given by the integral

$$\epsilon_{2p}(q) = \int \Psi^*_{2p}(\mathbf{r}_1, \mathbf{r}_2) \sum_{j=1}^{2} e^{i\mathbf{q}\cdot\mathbf{r}_j} \Psi_0(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \tag{4}$$

Here $\Psi_{2p}(\mathbf{r}_1, \mathbf{r}_2)$ and $\Psi_0(\mathbf{r}_1, \mathbf{r}_2)$ are the two-electron wavefunctions for the $1s\,2p\ ^1P$ and $1s^2$ states, respectively. After expressing these two states in terms of single-electron orbitals, (4) becomes

$$\epsilon_{2p}(q) = \sqrt{2} \langle \psi_{1s}(\mathbf{r}) | \psi_{1s0}(\mathbf{r}) \rangle \langle \psi_{2p}(\mathbf{r}) | e^{i\mathbf{q}\cdot\mathbf{r}} | \psi_{1s0}(\mathbf{r}) \rangle, \tag{5}$$

where $\psi_{1s}(\mathbf{r})$ and $\psi_{1s0}(\mathbf{r})$ are the $1s$ wavefunctions of the excited and ground states, respectively. Although these two functions are not identical, their overlap $\langle \psi_{1s}(\mathbf{r}) | \psi_{1s0}(\mathbf{r}) \rangle$, which is explicitly evaluated in all our computations, is always very close to unity. This shows that the overwhelming majority of the environmental dependence of $\epsilon_{2p}(q)$ arises from the compression of the $2p$ orbital acting to increase its value closer to the nucleus where $\psi_{1s0}(\mathbf{r})$ is more significant. The term $\langle \psi_{2p}(\mathbf{r}) | e^{i\mathbf{q}\cdot\mathbf{r}} | \psi_{1s0}(\mathbf{r}) \rangle$ in (5) is thereby enhanced.

We have calculated $\epsilon_{2p}(q)$ for a range of densities from wavefunctions computed with the GAMESS-UK program [19] for helium atom clusters as described previously [13]. Figure 2 shows how $|\epsilon_{2p}(q)|$ is significantly enhanced on passing from the free atom

through the n value of 0.0286 atoms/Å^3 presented in [5] to a density of 0.05 atoms/Å^3. The accuracy of our free atom $\epsilon_n(q)$ computed from the Hartree Fock wave functions is shown by its very close agreement with the result derived from very accurate correlated wave functions [17].

The density dependence of the computed cross section is presented in Figure 3. It is clear that the environmentally induced enhancements of the $\epsilon_{2p}(q)$ are carried over into σ_{2p}. It is notable that, even for moderate densities around 0.05 atoms/Å^3, the cross section is double that of the free helium atom.

Although the overwhelming majority of the variation in the calculated σ_{2p} values arises from the compression of the $2p$ orbital, one other factor can in principle contribute to a density dependence of σ_{2p}, namely the effect of ΔU_n on the integration limits $q_{\min}$ and $q_{\max}$ [5]. It should be noted that the ΔU_n in the dominator of (2) is cancelled after substitution of (3) for $f_{2p}(q)$. However in [5], a parametrized form for the free atom f_{2p} was used which implicitly contains a factor of $\frac{1}{\Delta U_0}$ (the free atom ΔU_n value) while the density-dependent ΔU_n was used in (4). The n dependence resulting from both these factors of σ_{2p} is seen from Figure 3 to be very small. The n dependence of this calculated σ_{2p} would be even smaller if the inconsistent treatment of the ΔU_n factors was removed.

The large density dependence of σ_{2p} presented in Figure 3 clearly shows that any value of n deduced from (1) using the free atom value for σ_{2p} as performed in [5,7–10,20] cannot be trusted.

It is clear that

$$\frac{I_p}{\sigma_{2p} I_z d}$$

cannot be taken to be an explicit function of n. It is, rather, an implicit function of the density since σ_{2p} is itself a function of density.

The quantities directly measured in the STEM experiments are I_p, I_z, and d. This shows that the purely experimental quantity related to (1) is the product $n \cdot \sigma_{2p}$ given by

$$\Sigma = \frac{I_p}{I_z d}. \tag{6}$$

3. Comparison of theory and experiment via the Σ values

The quantities Σ that were directly measured in STEM experiments can be reconstructed from the reported densities n by multiplying the latter by the σ_{2p} value, derived from the free atom GOS, that was used to calculate n. However both I_p and I_z depend on the details of the experiments, namely on the collection angles determining the $q_{\min}$ and $q_{\max}$ entering (2) as well as on the energy and convergence angle of the incident beam. The latter parameter can be described in terms of an effective maximum collection angle using the CONCOR2 code of Egerton [1]. The relation (1) then shows, consitent with equation (2) for σ_{2p}, that the value of the latter to be used to reconstruct the experimental Σ depends on the collection angles. For each of the experimental data considered below, we derived the Σ value from the free atom σ_{2p} computed using the collection angles which were either reported in the literature or communicated to us from the authors of the original

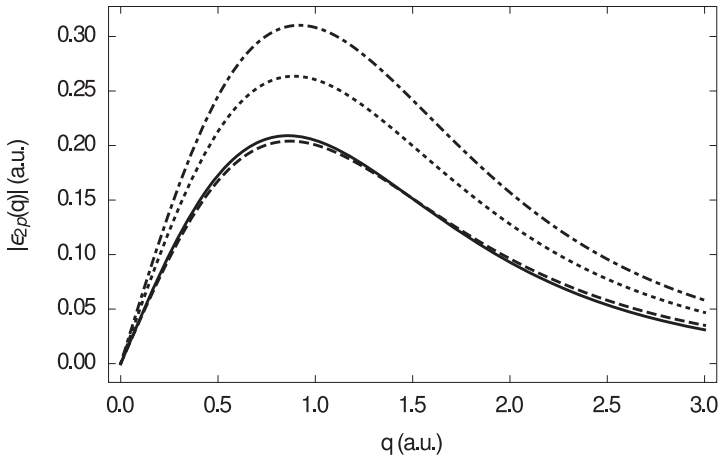


Figure 2. A comparison of the atomic form factor for free helium (solid) with our calculation for free helium (dashed), our delocalized calculation for BCC helium at density $n = 0.0286$ atoms/Å^3 (dotted), and our delocalized calculation for BCC helium at density $n = 0.05$ atoms/Å^3 (dot-dashed). For free helium, the analysis in [5] is entirely correct.

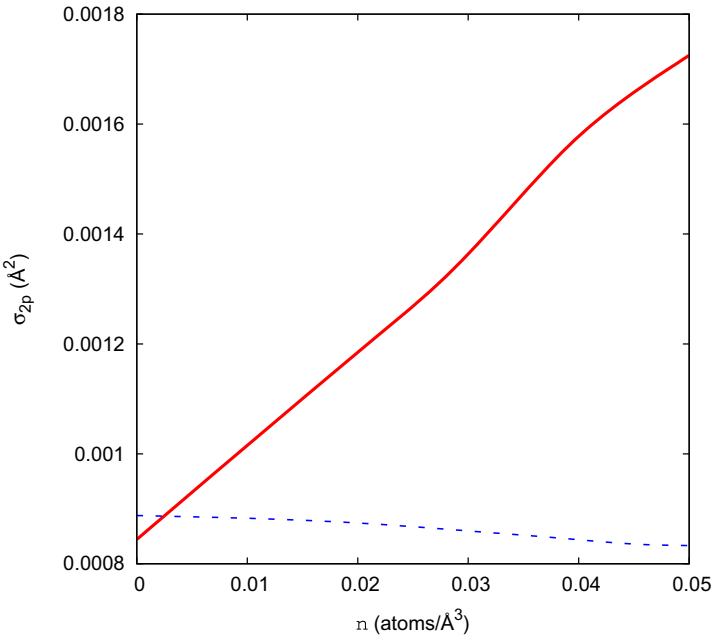


Figure 3. Cross section plotted as a function of number density n using the free atom wave function (dashed line) and the delocalized wave functions of the present work (solid line).

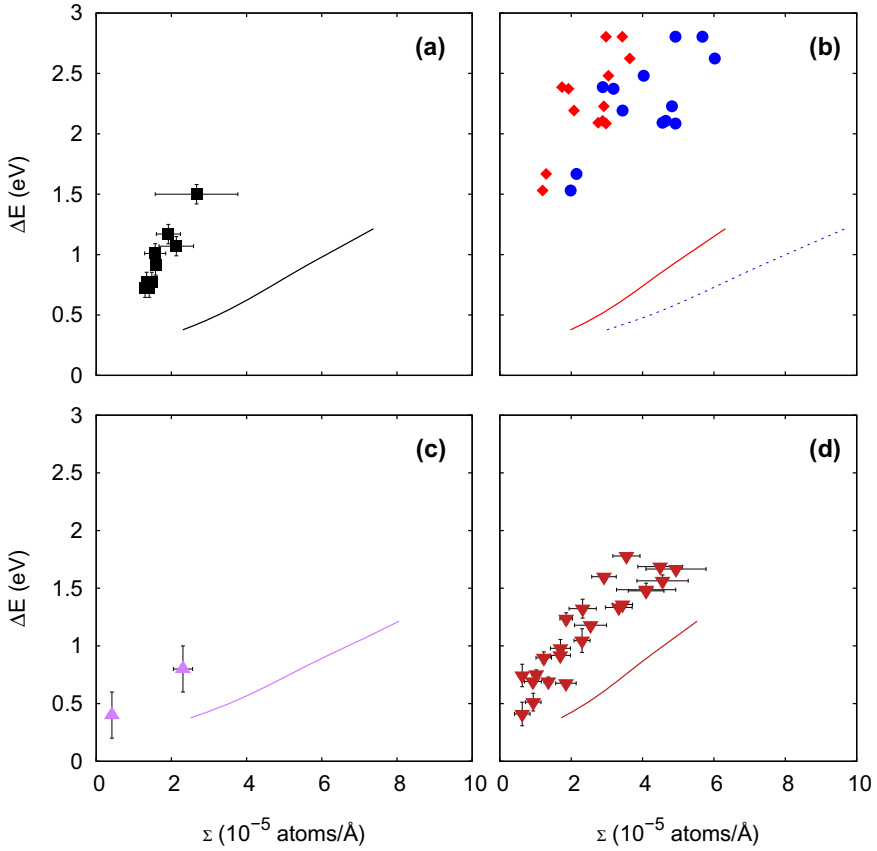


Figure 4. Comparison of our theoretical values of ΔE as a function of Σ (solid and dashed lines) with experimental data. Plotted in part (a) are the (corrected) experimental values of [7] (squares). In part (b) the upper (discs) and lower (diamonds) limiting values of [8] are both shown since only a range of values of σ are given; the solid and dashed theoretical curves correspond to the diamonds and discs, respectively. In (c) the data from [5] is plotted (triangles), and in (d) the data from [9] (inverted triangles) are shown. The axes on all four plots shown are identical.

paper. Σ values directly comparable to each of four sets of experiments were derived from the GOS generated by our computations by calculating the σ_{2p} (2) using the appropriately corresponding instrumental parameters. The differences in the latter necessitate presenting separately the four comparisons between theory and experiment (Figure 4).

The report [7] of the study of bubbles in a PdPt matrix presents two sets of data points which differ in that two densities are reported for the same ΔE measured at the centre of the same bubble. However, the derivation of the densities involves some averaging over the scattering from different regions in the bubble. The first set of densities, called uncorrected, was derived using the method of [5] without modification. The second set of densities, labelled corrected, was produced because it was considered that the scattering properties of helium atoms in proximity to the metallic wall are modified. The Σ values for corrected densities are plotted in Figure 4(a), the uncorrected values (not shown in Figure 4(a)) would appear to the left of those shown, with this leftward move increasing from about

8% for the smallest ΔE to 16% for the largest. Thus, application of the correction in [7] does not remove the large differences between the bubble and bulk $\Sigma/\Delta E$ relationships. It can be assumed that even though the application of this correction to the rest of the experimental data shown in Figure 4 would lead to a smaller disagreement between theory and experiment, the disagreement would still be far too significant to be explained solely by this effect.

In the study [8] of bubbles in Martensitic steel, variations between the instrumental parameters in the different experiments produced a range of the σ_{2p} values used in (1) to deduce the densities. The limits of this range were 5.5×10^{-4} Å^2 and 9.1×10^{-4} Å^2. For both these σ_{2p} values, we have calculated Σ and plotted in Figure 4(b) their measured ΔE_n against Σ. The two theoretical curves shown in Figure 4(b) correspond to the two sets of experimental parameters which give rise to the reported limits of σ_{2p} in [8]. The differences between the theoretical and experimental Σ presented in Figure 4(b) are at least as large, if not larger, than those in Figure 4(a).

The paper [5] of the study of bubbles in the PE16 alloy reports all the experimental details for the two bubbles studied. Our calculations showed that the effect of the differences in the collection angles of the two experiments is extremely small. Consequently, the comparison in Figure 4(c) between theory and experiment reports only the theoretical curve which is closer to the experiments. Again the differences between theory and experiment lie well outside the error bars reported for the latter.

For bubbles in a silicon matrix, theory is compared with the experiment [9] in Figure 4(d). The clear differences between theory and experiment would only be reduced if a constant energy offset was included in the latter. The difficulty of accurately measuring small energy shifts has been discussed [10].

4. Conclusions

There are three main conclusions resulting from the present work. Firstly, we have demonstrated that the values of the energy shifts predicted using the extended model of [13] are in excellent agreement with the laboratory measurements of the bulk. This encourages us to believe that the contractions of the $2p$ orbital with increasing density have been correctly described and as a consequence to deduce that the quantities directly measured in STEM experiments are implicitly rather than explicitly dependent on the density. Secondly, we have shown that the contraction of the $2p$ orbital causes σ_{2p} to increase with density very markedly.

Thirdly, it has been shown that most existing STEM measurements for the bubbles are inconsistent with our theoretical predictions, with the possible exception of the bubbles in silicon. This suggests either some systematic experimental problem with this type of measurement or that there may be physics present in the bubbles which is absent from the bulk. If the experimentally measured energy shifts in silicon needed to be corrected with a constant offset bringing theory and experiment into near agreement, it would follow that bubbles in silicon behaved differently from those in the three metals. This would provide further evidence for the influence on bubble properties from the encapsulating matrix.

Additionally, we note that there have been attempts to find a linear relationship between ΔE and n [5,7–10]; however, this would not be consistent with our theory and indeed leads to a large error at higher densities.

Acknowledgements

TCN would like to thank Prof. Martyn Guest of Cardiff University for his invaluable assistance in getting the GAMESS-UK calculations to run smoothly and also Profs. Marie-Laure David and Laurent Pizzagalli of Université Poitiers and Prof. Mathieu Kociak of the Laboratory for Solid State Physics in Orsay for helpful correspondence regarding experimental details. NCP and TCN thank Alex Thom for his help in running some of the GAMESS calculations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by a grant from the Royal Societynder the International Exchange Scheme. TCN gratefully acknowledges funding from EPSRC [grant number EP/LO15722/1].

References

[1] R.F. Egerton, *Electron Energy Loss Spectroscopy in the Electron Microscope*, Plenum, New York, 1996.
[2] P. Rez, Ultramicroscopy 9 (1982) p.283.
[3] J.P. Connerade, *Structure and Properties of Atomic Clusters*, Imperial College Press, London, 2008. p. 53.
[4] H. Trinkaus and B.N. Singh, J. Nucl. Mater. 323 (2003) p.229.
[5] C.A. Walsh, J. Yuan and L.M. Brown, Phil. Mag. A 80 (2000) p.1507.
[6] H. Trinkaus, Radiat. Eff. 78 (1983) p.189.
[7] D. Taverna, M. Kociak, O. Stéphan, A. Fabre, E. Finot, B. Décamps and C. Colliex, Phys. Rev. Lett. 100 (2008) p.035301.
[8] S. Fréchard, M. Walls, M. Kociak, J.P. Chevalier, J. Henry and D. Gorse, J. Nucl. Mater. 393 (2009) p.102.
[9] M.L. David, F. Pailloux, V. Mauchamp and L. Pizzagalli, Appl. Phys. Lett. 98 (2011) p.171903.
[10] K. Alix, M.L. David, G. Lucas, D.T.L. Alexander, F. Pailloux and L. Pizzagalli, Micron 77 (2015) p.57.
[11] C.M. Surko, G.J. Dick, F. Reif and W.C. Walker, Phys. Rev. Lett. 23 (1969) p.842.
[12] N. Schell, R.O. Simmons, A. Kaprolat, W. Schulke and E. Burkel, Phys. Rev. Lett. 74 (1995) p.2535.
[13] N.C. Pyper, D.W. Essex and C.T. Whelan, Phil. Mag. B 81 (2001) p.91.
[14] D.A. Arms, R.O. Simmons, M. Schwöerer-Bohning, A.T. Macrander and T.J. Graber, Phys. Rev. Lett. 87 (2001) p.156402.
[15] D.A. Arms, T.J. Graber, A.T. Macrander, R.O. Simmons, M. Schwöerer-Bohning and Y. Zhong, Phys. Rev. B 71 (2005) p.233107.
[16] A.J. McGibbon and L.M. Brown, Trans. R. Microsc. Soc. 1 (1990) p.23.
[17] M. Inokuti, Rev. Mod. Phys 43 (1971) p.297.
[18] A. Burgess and J.A. Tully, J. Phys. B 11 (1978) p.4271.
[19] M.F. Guest, I.J. Bush, H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A. Havenith and J. Kendrick, Mol. Phys 103 (2005) p.719–747. GAMESS-UK is a package of ab initio programs. See: "http://www.cfs.dl.ac.uk/gamess-uk/index.shtml".
[20] R. Schierholz, B. Lacroix, V. Godinho, J. Caballero-Hernández, M. Duchamp and A. Fernández, Nanotechnology 26 (2015) p.1–10.

Chapter 4

Calculation of projected potentials using pseudopotentials for use in electron microscopy simulations

Calculation of projected potentials using pseudopotentials for use in electron microscopy simulations

T. C. Naginey, Rebecca J. Nicholls, Peter D. Nellist, and Jonathan R. Yates
Department of Materials, University of Oxford, Parks Road, Oxford, OX1 3PH, U. K.

A general procedure for obtaining projected electrostatic potentials from pseudopotential plane wave implementations of density functional theory (DFT) is given. The projected potential is the key specimen-related quantity needed in simulations of electron microscopy experiments. Results from an implementation of this technique in CASTEP, a pseudopotential plane wave DFT code, are used in the simulation of conventional transmission electron microscopy images of hexagonal boron nitride. Projected potentials and images obtained from CASTEP are compared with the widely used independent atom model and with results from all-electron DFT. Substantial disagreement with results from the independent atom model is found, while nearly exact agreement is found with all-electron DFT.

Keywords: electron microscopy, dft, *ab-initio*

I. INTRODUCTION

Modern electron microscopes are now capable of routinely achieving sub-Ångstrom resolution [1–4], due to fairly recent technological advancements such as aberration correction. The quantitative interpretation of the data relies heavily on comparison with simulation [5]. One crucial deficiency in commonly used microscopy simulation software is the neglect of inter-atomic interactions between electrons of neighboring atoms in the specimen. These interactions have been accounted for in isolated studies through the application of density functional theory (DFT) [6–11]. However, a general procedure for including these effects with DFT in a systematic way is absent from the literature. It is this gap in the literature that we attempt to close with the approach outlined in this work for obtaining projected electrostatic potentials from pseudopotential plane wave (PPW) DFT.

In section II, a brief outline of some of the principles of simulations in electron microscopy and calculations using DFT are presented. We move on to demonstrate a general procedure for extracting the projected potential from any PPW DFT implementation in III. Next, we show results from an implementation of this method in CASTEP in IV, where we compare against results of both the independent atom model (IAM) and all-electron (AE) DFT, before moving on to our conclusions in V.

II. BACKGROUND

Here we present a brief discussion of some key concepts related to modelling in both electron microscopy and electronic structure theory. We begin with a derivation of the projection assumption to demonstrate the importance of the projected electrostatic potential in electron microscopy simulations. Next we present a discussion of some of the other basic aspects of modelling in electron microscopy, followed by a brief overview of the concepts from DFT that will be used in this paper.

A. The projection assumption

Many approximations are commonly used in the simulation of electron microscopy images [5, 12]. The theoretical and computational techniques that will be outlined in this work have been developed with the purpose of improving the approximations made concerning the electronic structure of the sample, but we will make most of the other common approximations. We begin by outlining the origin of the major assumptions and approximations in the context of modelling (scanning) transmission electron microscopy ((S)TEM) experiments.

Many of the underlying principles of electron microscopy image simulations can be understood from simply considering a plane wave passing through a potential. The high energy (usually between 60-300 keV) electrons that comprise the beam of the microscope are referred to as the fast electrons and can be thought of in terms of plane waves. In conventional TEM the fast electron is modelled as a single plane wave. In the case of STEM, in which a focused electron beam is scanned across the sample, the beam can be thought of as a superposition of plane waves. Thus it suffices to consider the behavior of a single plane incident wave $\Psi_i(x)$ whose initial form before passing through the sample is expressed as

$$\Psi_i(x) = Ae^{ikx} \tag{1}$$

in one dimension, with amplitude A and wave number $k = 2\pi/\lambda$. The simplest case to consider is that of a very thin specimen, which is where we begin. If we assume that the potential of the specimen $V(x)$ varies slowly over distances that are large compared to the wavelength of the fast electrons λ, then we can use the WKB approximation [13] which gives the form of the transmitted plane wave after passing through the sample as

$$\Psi_t(x) = \frac{c}{\sqrt{p(x)}} e^{\frac{i}{\hbar}\int p(x)dx} \tag{2}$$

where c is some complex constant and the momentum $p(x)$ is given by

$$p(x) = \sqrt{2m_e(E - eV(x))}, \tag{3}$$

with E the energy of the fast electron, and m_e and e the mass and charge of the electron, respectively. Now we seek to simplify the form of Ψ_t by making some additional approximations, going beyond those of the WKB model. We begin by expanding p in terms of the function V:

$$p = \sqrt{2m_eE(1 - \frac{V}{E})} \approx \sqrt{2m_eE}(1 - \frac{V}{2E}). \tag{4}$$

Likewise we approximate $1/\sqrt{p}$ as

$$\frac{1}{\sqrt{p}} \approx (2m_eE)^{-1/4}. \tag{5}$$

The integral in the exponent of Eq. (2) can be made definite by assuming that our potential $V(x')$ is nonzero only over a small region of space between spatial coordinates 0 and x. Applying Eq. (4) and Eq. (5) to Eq. (2) and carrying out the integral in the exponent from 0 to x, we arrive at

$$\Psi_t(x) \approx A \exp\left\{ \frac{i}{\hbar}\sqrt{2m_eE}\left(x - \frac{e}{2E}\int_0^x V(x')dx' \right) \right\} \tag{6}$$

where we have absorbed all constants into A which we set equal to the amplitude in Eq. (1). Since the sample is thin, $V(x')$ is only nonzero over a small region on the interval $(0, x)$, so $\int_0^x V(x')dx' = \int_{-\infty}^{\infty} V(x')dx'$. Defining an "interaction parameter," σ, as

$$\sigma \equiv \frac{e}{\hbar}\sqrt{\frac{m_e}{2E}} = \frac{2\pi m_e e\lambda}{h^2}, \tag{7}$$

we can recast Eq. (6) as

$$\Psi_t(x) \approx Ae^{ikx} \exp\left\{ i\sigma \int_{-\infty}^{\infty} V(x')dx' \right\} \tag{8}$$

$$= \Psi_i(x) \exp\left\{ i\sigma \int_{-\infty}^{\infty} V(x')dx' \right\}. \tag{9}$$

Thus we see that, under the above approximations, the incident plane wave merely picks up a phase shift after passing through our thin sample.

Generalizing to three dimensions and choosing our coordinate system so that the sample lies in the x-y plane and the plane wave travels from the $+z$ direction and is orthogonal to the z-axis we have

$$\Psi_t(x, y) \approx \Psi_i(x, y) \exp\{i\sigma V_z(x, y)\} \tag{10}$$

where we have defined the "projected potential" as

$$V_z(x, y) \equiv \int_{-\infty}^{\infty} V(x, y, z)dz. \tag{11}$$

The exponential factor in Eq. (10) is commonly referred to as the transmission function and denoted as

$$t(x, y) \equiv \exp\{i\sigma V_z(x, y)\}. \tag{12}$$

From Eq. (10) it can be seen that V_z is the physical quantity that relates the sample to the transmitted wave function in this model. The set of assumptions that gives rise to Eq. (10) is known as the projection assumption and the scope of its validity is discussed in detail in Ref. [**?**]. For thicker samples, or samples containing heavy nuclei, it is necessary to slice the potential into subsections along the x-y plane and then project each subsection separately. Between each slice, the plane wave emerging from the previous slice is propagated to the next using a Fresnel propagator [5, 12]. This technique is called the multi-slice method, whereas taking only one integral as in Eq. (11) is referred to as the single slice technique. Unless otherwise noted, we will explain the concepts in this paper in the context of the single-slice model for simplicity, since generalizing to the multi-slice model is usually straightforward.

B. The independent atom model (IAM)

In most of the software used to simulate (S)TEM images and cross-sections, the projected potential of the sample is constructed on the fly from the projected potentials of isolated atoms. Fully relativistic electrostatic potentials are taken from Hartree-Fock calculations [14] and then parametrized with analytic basis functions. For example, I. Lobato and D. Van Dyck fit the projected potentials using a set of modified Bessel functions of the second kind with ten free parameters for each element [15]. Since the basis functions remain the same for each atom, storing projected potentials for all elements on the periodic table requires only that a library of parameters be stored within the code. These independent atom projected potentials are then simply superposed on the lattice of the sample being modelled. The accurate parametrization of these independent atom potentials has been the subject of serious research in and of itself [5, 15, 16]. However, no matter how accurate the parametrization, the use of this independent atom model (IAM) leaves out all inter-atomic interactions of electrons in the solid. It is this complete neglect of basic electronic structure that we intend to systematically overcome with the techniques presented in this work by making it easy and user-friendly to use DFT to generate the projected potential of the material as a whole to be used in electron microscopy simulations.

C. Other aspects of electron microscopy simulation

The focus of this work is on the generation of more accurate projected potentials for use in the transmission function in Eq. (12), but a brief discussion of the basic principles of generating the images in standard (S)TEM simulations will be given first. We will consider only conventional TEM and annular dark field (ADF) STEM simulations here, and neglect the effects that thermal diffuse scattering due to phonons has on the image.

1. Conventional TEM

In conventional TEM, the first step of the simulation is to obtain the transmitted wave function, given by

$$\Psi_t(x, y) = t(x, y), \tag{13}$$

where the initial plane wave from Eq. (10) has been taken to be unity. After passing through the sample, the wave function then passes through the aperture and magnetic lenses of the microscope. Mathematically this is approximated as the convolution of Ψ_t with the point spread function $h_o(x, y)$ which represents the effects of the aperture, defocus, and aberrations of the objective lens. Approximating the interaction of Ψ_t with the lens as a convolution is known as the linear image approximation [5, 17]. Once this convolution has been performed, the final step is to take the square modulus of the wave function. The result is the two dimensional image intensity,

$$g(x, y) = |\Psi_t \otimes h_o|^2. \tag{14}$$

2. ADF STEM

To model ADF STEM experiments, the first step is to construct the probe function $\Psi_p(x, y)$ which represents the narrow, focused electron beam of the microscope formed by passing the electrons through the lens *before* they pass through the sample. For a given probe position with coordinates (x_p, y_p), the effect of the sample on the beam is represented mathematically as a simple product of the probe function with the transmission function,

$$\Psi_t(x, y, x_p, y_p) = \Psi_p(x - x_p, y - y_p)t(x, y). \tag{15}$$

Next, the Fourier transform of Ψ_t is taken with respect to x and y to obtain the wave function in the diffraction plane, $\widetilde{\Psi}_t$. The square modulus of the portion of this wave function that falls on the annular detector is then integrated to obtain the intensity of the pixel for that probe position,

$$g(x_p, y_p) = \int\int |D(\boldsymbol{k})\widetilde{\Psi}_t(x_p, y_p, \boldsymbol{k})|^2 d\boldsymbol{k}, \tag{16}$$

where $\boldsymbol{k}$ is the 2D vector in reciprocal space and $D(\boldsymbol{k})$ is the detector function, equal to unity for values of $\boldsymbol{k}$ on the detector and zero otherwise. The total image is then constructed pixel by pixel by repeating this process for each probe position, with the dimensions of the pixels being equivalent to the step-size of the probe in the x and y directions.

We see that in both conventional TEM and ADF STEM simulations, all dependence of the final image on the sample comes from V_z, via the transmission function t.

3. Additional general remarks

Further approximations are common in electron microscopy simulations. One such approximation is the use of non-relativistic quantum mechanics to model the fast-electronic wave functions, for example in the derivation of Eq. (10). However, relativity is partially taken into account by using the relativistic wavelength of the electron, given by

$$\lambda = \frac{hc}{\sqrt{eV(eV + 2m_e c^2)}} \tag{17}$$

where c is the speed of light in vacuum and V the accelerating potential of the microscope. In addition, the many-body interactions of the fast electron with the electrons in the sample material are neglected [18]. Finally, the effects of phonons in electron microscopy experiments are frequently approximated with the Debye-Waller factor or an absorptive potential approach [12]. The limits of these approximations and their effects on simulated images have been studied and can be found in [5, 12, 19]. The only approximation under examination in the present work, however, will be the use of the IAM.

D. Density functional theory (DFT)

DFT has been the theoretical framework of choice in previous studies seeking to go beyond the IAM in electron microscopy simulations [8, 9, 20–22] due to its balance between accuracy and computational efficiency in its applications in solid state physics [23]. Within DFT, there are several types of implementations to choose from [24–27]. Two main categories are all-electron (AE) and pseudopotential plane wave (PPW) codes. The AE approach is considered a gold standard withing DFT, as all of the electrons in the solid are treated explicitly and allowed to react to the chemical environment of the solid. If however the frozen core approximation is made, which

rests on the assumption that the core electronic states are insensitive to the chemical environment and remain identical to their atomic state, computational cost can be greatly reduced. This is because we can then use pseudopotentials to model nuclear and core potentials [28–31]. These pseudopotentials are chosen so that the wave functions representing the valence states are smooth and nodeless near the nuclei. Thus a plane wave basis can be used for all valence electrons involved in the calculation, simplifying the mathematics and reducing computational cost.

In some previous studies [6, 7] in which (S)TEM images were generated from DFT projected potentials, two different DFT codes were used to model a single system. This was because the PPW software used for the more expensive structural relaxation calculations did not allow access to the full projected potentials needed for the microscopy simulations. The procedure was then to perform the structural relaxation in the PPW code first, and subsequently use the structure output by this calculation as input to an AE DFT program from which the extraction of the projected potentials was more straightforward. One of the main advantages, then, of obtaining projected potentials directly from PPW implementations as opposed to AE codes is that the user is able to perform more expensive calculations and obtain projected potentials within the framework of a single code.

In what follows, we will outline a general procedure for obtaining projected potentials from PPW density functional theory. We will then use our implementation of this procedure in CASTEP to demonstrate the differences between IAM- and full DFT-generated microscopy images for a few simple cases to illustrate the basic physics that are left out of the IAM and easily incorporated by using PPW DFT to generate the projected potentials. In this work, the simulation of electron microscopy images was performed using software written in-house, following the general approach outlined in [5]. Unless otherwise stated, all quantities are in atomic units.

III. METHODOLOGY: OBTAINING THE PROJECTED POTENTIALS

The fundamental physical quantity of interest in DFT is the ground state electronic charge density, $\rho(\boldsymbol{r})$ [32, 33]. It is known from the Hohenberg-Kohn theorem that all other ground state physical observables are functionals of $\rho(\boldsymbol{r})$. The ground state electrostatic potential, $V(\boldsymbol{r})$, is a member of this class of observables. Once an accurate charge density has been obtained from a (carefully) converged self-consistent field calculation, the potential for a periodic solid can be obtained, in theory, simply from the solution to Poisson's equation,

$$V(\boldsymbol{r}) = V[\rho](\boldsymbol{r}) = \int_{\Omega} \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}', \tag{18}$$

where the integral has been taken over the volume of the unit cell, Ω. The rest of this section will be concerned with how this integral is handled in practice.

In PPW implementations of DFT, the core and valence electrons are treated separately. This is due to the frozen core approximation, which is central to the PPW approach. The electrons in the lowest lying energy states–the core electrons–are assumed to remain unchanged from their atomic states. Charge densities for these core electrons, $\rho_c(\boldsymbol{r})$, can be obtained from the single atom calculations that are necessary for the generation of the pseudopotentials. These core quantities exhibit discontinuities near the nucleus, and are thus typically stored numerically on logarithmic radial grids with very fine spacing close to the origin. The valence charge densities $\rho_v(\boldsymbol{r})$ are, by construction, much smoother and are able to be stored on a regular Cartesian grid. The different structure of the numerical grids used for the core and valence quantities necessitate that they be treated separately in the calculation of the potential. We divide the system into a valence portion, containing quantities related only to the valence electrons, and an ionic portion, containing quantities related to the combined system of nuclei and core electrons. The charge density of the latter will be denoted $\rho_i(\boldsymbol{r})$.

While the notation used in this work is general, one must be careful to understand that the quantity represented by ρ_v in this work depends on the details of the pseudopotentials used. Within the PPW approach there are several different types of pseudopotential schemes. In the norm conserving scheme [31], the total amount of charge contained within the cutoff radius r_c around the nucleus–called the augmentation region–is constrained to be equal to the AE value, and so can be thought of as a redistribution of the charge within the augmentation region. In the ultrasoft scheme [29, 30], the valence charge density is divided into a soft portion, represented on the plane wave basis of the calculation, and a hard portion which is essentially the difference between the AE and the soft valence charge densities and is thus localized within the augmentation regions. The soft portion is not subject to the norm-conserving constraint within the augmentation region. The charge missing from the soft portion is contained within the hard charge. However, this hard charge is itself pseudized within a smaller cutoff radius, r_{cc}. All of the results shown in Section IV of this work were calculated by using the sum of the soft and pseudized hard charges for ρ_v. This is justified by the fact that, within the smaller augmentation region bounded by r_{cc}, the Coulomb potential of the nucleus dominates the resulting projected potential, and so the charge missing from ρ_v has little effect on the final image. For reference, $r_{cc} = 0.84$ for boron and 0.77 for nitrogen, in atomic units, in the hexagonal boron nitride system examined in Section IV.

The electrostatic potential of an infinitely periodic system is only well-defined if the net charge of the unit cell is zero. To keep the separate valence and ionic systems neu-

tral, we apply the general procedure used for calculating the local electrostatic energy in [28] to the calculation of the local electrostatic potential. We define an auxiliary charge $\rho_a(\boldsymbol{r})$ which obeys

$$\int_\Omega (\rho_v(\boldsymbol{r}) + \rho_a(\boldsymbol{r}))\, d\boldsymbol{r} = \int_\Omega (\rho_i(\boldsymbol{r}) - \rho_a(\boldsymbol{r}))\, d\boldsymbol{r} = 0. \tag{19}$$

We can then separate the ionic and core potentials in the following way:

$$V(\boldsymbol{r}) = \underbrace{\int_\Omega \frac{\rho_v(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}'}_{V_v(\boldsymbol{r})} + \underbrace{\int_\Omega \frac{\rho_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}'}_{V_i(\boldsymbol{r})} \tag{20}$$

$$= \underbrace{\int_\Omega \frac{\rho_v(\boldsymbol{r}') + \rho_a(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}'}_{V_v'(\boldsymbol{r})} + \underbrace{\int_\Omega \frac{\rho_i(\boldsymbol{r}') - \rho_a(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}'}_{V_i'(\boldsymbol{r})}, \tag{21}$$

where the potential due to ρ_a has been added and subtracted in the last step. The final projected potential will be the sum

$$V_z(x,y) = \underbrace{\int V_v'(\boldsymbol{r})\, dz}_{V_z^v(\boldsymbol{r})} + \underbrace{\int V_i'(\boldsymbol{r})\, dz}_{V_z^i(\boldsymbol{r})} \tag{22}$$

with the integrals performed over the z axis of the unit cell. The two integrals in Eq. (21) will be considered separately, beginning with the valence term $V_v'(\boldsymbol{r})$.

A. The valence term

The potential due to the sum of the valence ρ_v and the auxiliary ρ_a charge densities,

$$V_v'(\boldsymbol{r}) = \int_\Omega \frac{\rho_v(\boldsymbol{r}') + \rho_a(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}', \tag{23}$$

is conveniently evaluated in Fourier space. The Fourier transform of a function $f(\boldsymbol{r})$ with the periodicity of the unit cell will be defined in the present work as

$$\tilde{f}(\boldsymbol{G}) \equiv \mathscr{F}[f(\boldsymbol{r})] \equiv \frac{1}{\Omega}\int_\Omega f(\boldsymbol{r}) e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} d\boldsymbol{r} \tag{24}$$

where $\boldsymbol{G}$ is a reciprocal lattice vector of magnitude G. It follows that the inverse Fourier transform is given by

$$\mathscr{F}^{-1}[\tilde{f}(\boldsymbol{G})] = \sum_{G=0} \tilde{f}(\boldsymbol{r}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}}. \tag{25}$$

In practice, there is a finite upper limit to the sum in Eq. (25), defined in terms of a cutoff energy which sets the size of the plane wave basis used in DFT calculations. From Eq. (23), it can be seen that V_v' is a convolution. We can then invoke the convolution theorem to obtain

$$V_v'(\boldsymbol{r}) = \mathscr{F}^{-1}[\mathscr{F}[V_v'(\boldsymbol{r})]] \tag{26}$$

$$= \sum_{G=0} (\tilde{\rho}_v(\boldsymbol{G}) + \tilde{\rho}_a(\boldsymbol{G})) \frac{4\pi}{G^2} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}. \tag{27}$$

The singularity at the $G = 0$ term is avoided since

$$\tilde{\rho}_v(0) + \tilde{\rho}_a(0) = \frac{1}{\Omega}\int_\Omega (\rho_v(\boldsymbol{r}) + \rho_a(\boldsymbol{r}))d\boldsymbol{r} = 0, \tag{28}$$

the last equality coming from Eq. (19). Using this result, equation (27) becomes

$$V_v'(\boldsymbol{r}) = \sum_{G=1} (\tilde{\rho}_v(\boldsymbol{G}) + \tilde{\rho}_a(\boldsymbol{G})) \frac{4\pi}{G^2} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}. \tag{29}$$

The derivation thus far is true of any form of $\rho_a(\boldsymbol{r})$ that satisfies Eq. (19). The calculation of V_v in Fourier space simplifies further if $\rho_a(\boldsymbol{r})$ is chosen to be a uniformly distributed charge over the volume of the unit cell,

$$\rho_a(\boldsymbol{r}) = \frac{-Z_v}{\Omega}, \tag{30}$$

with Z_v being the total valence charge contained in Ω,

$$Z_v = \int_\Omega \rho_a(\boldsymbol{r}) d\boldsymbol{r}. \tag{31}$$

The Fourier transform of Eq. (30) is

$$\tilde{\rho}_a(\boldsymbol{G}) = \frac{-Z_v}{\Omega}\delta_{G\,0}, \tag{32}$$

where δ is the Kronecker delta. On substitution of (32) into (29), V_v' becomes

$$V_v'(\boldsymbol{r}) = \sum_{G=1} \tilde{\rho}_v(\boldsymbol{G}) \frac{4\pi}{G^2} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}. \tag{33}$$

It is important to bear in mind that Eq. (33) is the potential due to the sum of the valence and uniform background auxiliary charge densities. The projected potential due to this charge distribution is

$$V_z^v(x,y) = \int V_v'(\boldsymbol{r}) dz. \tag{34}$$

Since the function V_v' is smooth and well-behaved, its numerical integration through the z axis in Eq. (34) is not complicated and basic numerical integration techniques can be employed. The projected potential V_z^v will be

stored on a 2D regular grid whose spacing will be determined by the real space grid used in the DFT calculation. The grid spacing in a standard DFT calculation is typically greater than what is needed in an electron microscopy simulation. The projected potentials used as input for microscopy simulations should be sampled at a rate that is three to four times the resolution of the microscope being modelled, which results in grid spacings on the order of 10^{-2} a.u. However, since $V_z^v(x,y)$ is smooth and well-behaved, it can be interpolated efficiently after it has been projected onto the default DFT grid using a simple 2D cubic spline. Having outlined the basic procedure for obtaining the projected potential related to the valence charge, we now turn our attention towards the ionic quantities.

B. The ionic term

The potential due to the difference between the ionic ρ_i and the auxiliary ρ_a charge densities given by

$$V_i'(\boldsymbol{r}) = \int_\Omega \frac{\rho_i(\boldsymbol{r}') - \rho_a(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}', \tag{35}$$

will be handled using the method of Ewald summation [28].

Core charge densities ρ_c in PPW DFT are taken from calculations on isolated atoms. In the isolated atomic state, the core charge density and its potential are spherically symmetric. Using the convention that $V = 0$ as $r \to \infty$, the potential due to the core charge density of an atom α centered at the origin is

$$V_c^\alpha(r) = 4\pi \left(\frac{1}{r} \int_0^r r'^2 \rho_c^\alpha(r') dr' + \int_r^\infty r' \rho_c^\alpha(r') dr' \right). \tag{36}$$

The radial ionic potential of an atom α centered at $\boldsymbol{r} = \boldsymbol{r}_\alpha$ due to the sum of the nuclear and core charge densities is given by

$$V_i^\alpha(r) = V_c^\alpha(|\boldsymbol{r} - \boldsymbol{r}_\alpha|) - Z^\alpha/|\boldsymbol{r} - \boldsymbol{r}_\alpha|, \tag{37}$$

where Z^α is the atomic number and the minus sign comes from the convention common in DFT to treat electronic charge as positive and nuclear charge as negative.

The exact ionic charge density corresponding to an atom α at position $\boldsymbol{r}_\alpha$ is the sum of the core and nuclear charge densities,

$$\rho_{i,exact}^\alpha(\boldsymbol{r}) = \rho_c^\alpha(\boldsymbol{r} - \boldsymbol{r}_\alpha) + Z_n^\alpha \delta(\boldsymbol{r} - \boldsymbol{r}_\alpha), \tag{38}$$

where Z_n^α is the nuclear charge of atom α (which is taken to be negative) and δ is the Dirac delta function.

Far away from each nucleus, the potential due to the ion can be approximated as that arising from a point particle with ionic charge Z_i^α and charge density

$$\rho_{i,approx}^\alpha(\boldsymbol{r}) = Z_i^\alpha \delta(\boldsymbol{r} - \boldsymbol{r}_\alpha), \tag{39}$$

where $Z_i^\alpha = Z_v^\alpha + Z_n^\alpha$ with Z_v^α being the total valence charge of atom α.

Within the innermost 27 unit cells of our calculation, we use the exact expression, while outside of this range we use the approximate charge density of Eq. (39). Thus our total ionic charge density is

$$\rho_i(\boldsymbol{r}) = \sum_{\boldsymbol{R}_{inner}} \sum_\alpha \rho_{i,exact}^\alpha(\boldsymbol{r} - \boldsymbol{R}) + \sum_{\boldsymbol{R}_{outer}} \sum_\alpha \rho_{i,approx}^\alpha(\boldsymbol{r} - \boldsymbol{R}) \tag{40}$$

where the sum over α runs over the atoms in the unit cell, and the sum over $\boldsymbol{R}_{inner}$ is the sum over the set of lattice translation vectors corresponding to the first 27 unit cells and the sum over $\boldsymbol{R}_{outer}$ is in practice only taken over the next shell of unit cells.

In order to solve for V_i', we need to take into account the uniform background auxiliary charge distribution. Thus we will need to find the potential due to the total charge distribution given by

$$\rho_i'(\boldsymbol{r}) = \rho_i(\boldsymbol{r}) - \rho_a(\boldsymbol{r}), \tag{41}$$

where $\rho_i(\boldsymbol{r})$ is given by Eq. (40).

We follow the Ewald potential strategy of adding and subtracting Gaussian functions g_α to the charge distribution at each atomic position to obtain

$$\rho_i'(\boldsymbol{r}) = \underbrace{\rho_i(\boldsymbol{r}) - \sum_\alpha \sum_{\boldsymbol{R}} g_\alpha(\boldsymbol{r} - \boldsymbol{R})}_{\rho_1} \underbrace{- \rho_a(\boldsymbol{r}) + \sum_\alpha \sum_{\boldsymbol{R}} g_\alpha(\boldsymbol{r} - \boldsymbol{R})}_{\rho_2}, \tag{42}$$

where the functions g_α are given by

$$g_\alpha(\boldsymbol{r}) = A_\alpha e^{-|\boldsymbol{r} - \boldsymbol{r}_\alpha|^2/\mu_\alpha^2} \qquad A_\alpha = \frac{Z_i^\alpha}{(\sqrt{\pi}\mu_\alpha)^3}, \tag{43}$$

and μ_α is a parameter that can be chosen arbitrarily. The value A_α ensures that the Gaussian functions are normalized such that $\int g_\alpha(\boldsymbol{r})\, d\boldsymbol{r} = Z_i^\alpha$.

We will solve for the potentials due to ρ_1 and ρ_2 separately and add them together at the end to find

$$V_i'(\boldsymbol{r}) = \phi_1(\boldsymbol{r}) + \phi_2(\boldsymbol{r}) \tag{44}$$

where ϕ_1 and ϕ_2 solve the equations

$$\nabla^2 \phi_1(\boldsymbol{r}) = -4\pi \rho_1(\boldsymbol{r}), \qquad \nabla^2 \phi_2(\boldsymbol{r}) = -4\pi \rho_2(\boldsymbol{r}). \tag{45}$$

We first solve for ϕ_1. It can be shown that the potential arising from the Gaussian charge distribution $g_\alpha(\boldsymbol{r})$, given by Eq. (43), is

$$\phi_g^\alpha(\boldsymbol{r}) = \frac{Z_i^\alpha}{|\boldsymbol{r}-\boldsymbol{r}_\alpha|}\,\mathrm{erf}\left(\frac{|\boldsymbol{r}-\boldsymbol{r}_\alpha|}{\mu_\alpha}\right), \tag{46}$$

where erf is the error function given by

$$\mathrm{erf}\,(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt. \tag{47}$$

From Eq. (40) and Eq. (37), we can see that the potential arising from ρ_i is

$$\phi_i(\boldsymbol{r}) = \sum_{\boldsymbol{R}_{inner}}\sum_\alpha V_i^\alpha(|\boldsymbol{r}-\boldsymbol{R}|) + \sum_{\boldsymbol{R}_{outer}}\sum_\alpha \frac{Z_i^\alpha}{|\boldsymbol{r}-\boldsymbol{r}_\alpha-\boldsymbol{R}|} \tag{48}$$

The potential due to ρ_1 is then given by:

$$\begin{aligned}\phi_1(\boldsymbol{r}) &= \phi_i(\boldsymbol{r}) + \sum_{\boldsymbol{R}}\sum_\alpha \phi_g^\alpha(\boldsymbol{r}-\boldsymbol{R}) \\ &= \sum_{\boldsymbol{R}_{inner}}\sum_\alpha \left[V_i^\alpha(|\boldsymbol{r}-\boldsymbol{R}|) + \phi_g^\alpha(\boldsymbol{r}-\boldsymbol{R})\right] \\ &+ \sum_{\boldsymbol{R}_{outer}}\sum_\alpha \frac{Z_i^\alpha}{|\boldsymbol{r}-\boldsymbol{r}_\alpha-\boldsymbol{R}|}\mathrm{erfc}\left(\frac{|\boldsymbol{r}-\boldsymbol{r}_\alpha-\boldsymbol{R}|}{\mu_\alpha}\right).\end{aligned} \tag{49}$$

Next, we evaluate the potential ϕ_2 by solving Poisson's equation Eq. (45) in reciprocal space,

$$\widetilde{\phi}_2(\boldsymbol{G}) = \frac{4\pi}{G^2}\widetilde{\rho}_2(\boldsymbol{G}). \tag{50}$$

To find $\widetilde{\rho}_2(\boldsymbol{G})$, we first apply the definition of the Fourier transform for periodic functions Eq. (24) to the Gaussian term to find

$$\mathscr{F}\left[\sum_R g_\alpha(\boldsymbol{r}-\boldsymbol{R})\right] = \frac{Z_i^\alpha}{\Omega}e^{-i\boldsymbol{G}\cdot\boldsymbol{r}_\alpha}e^{G^2\mu_\alpha^2/4}. \tag{51}$$

The Fourier transform of the uniform background auxiliary charge distribution is given by

$$\mathscr{F}\left[\rho_a(\boldsymbol{r})\right] = \frac{Z_i^\alpha}{\Omega}\delta(\boldsymbol{G}). \tag{52}$$

Thus we find

$$\begin{aligned}\widetilde{\rho}_2(\boldsymbol{G}) &= \sum_\alpha \mathscr{F}\left[\sum_R g_\alpha(\boldsymbol{r}-\boldsymbol{R})\right] - \mathscr{F}\left[\rho_a(\boldsymbol{r})\right] \\ &= \sum_\alpha \frac{Z_i^\alpha}{\Omega}e^{-i\boldsymbol{G}\cdot\boldsymbol{r}_\alpha}e^{G^2\mu_\alpha^2/4} - \frac{Z_i^\alpha}{\Omega}\delta(\boldsymbol{G}).\end{aligned} \tag{53}$$

Inserting Eq. (53) into Eq. (50) and taking the inverse Fourier transform of both sides gives us

$$\phi_2(\boldsymbol{r}) = \sum_{\boldsymbol{G}=1}\sum_\alpha \frac{4\pi Z_i^\alpha}{\Omega G^2}e^{i\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{r}_\alpha)}e^{-G^2\mu_\alpha^2/4}. \tag{54}$$

Inserting Eq. (49) and Eq. (54) into Eq. (44) completes the derivation of the ionic term of the potential V_i'. The projected potential due to this term, V_z^2, is then simply the integral through the z axis.

It should be noted that the calculation of ϕ_2 in Eq. (54) is the most expensive part of the entire process, since it involves a sum over all $\boldsymbol{G}$-vectors (except $\boldsymbol{G}=0$) for each real space grid point. Thus the calculation of projected potentials scales most strongly with the size of the plane wave basis used in the calculation. However this quantity can be evaluated on a fairly coarse grid, being smooth and well-behaved in real space. The quantity ϕ_2 can then be integrated through z numerically and then interpolated easily onto however fine a grid is desired for the microscopy calculation.

The same is true of the second term of ϕ_1 in Eq. (54),

$$\sum_{\boldsymbol{R}_{outer}}\sum_\alpha \frac{Z_i^\alpha}{|\boldsymbol{r}-\boldsymbol{r}_\alpha-\boldsymbol{R}|}\mathrm{erfc}\left(\frac{|\boldsymbol{r}-\boldsymbol{r}_\alpha-\boldsymbol{R}|}{\mu_\alpha}\right), \tag{55}$$

since this is also smooth and well-behaved over the central unit cell.

However, more care must be taken with the first term of ϕ_1 in Eq. (54),

$$\sum_{\boldsymbol{R}_{inner}}\sum_\alpha \left[V_i^\alpha(|\boldsymbol{r}-\boldsymbol{R}|) + \phi_g^\alpha(\boldsymbol{r}-\boldsymbol{R})\right], \tag{56}$$

because of the singularity arising from the nuclear portion of V_i^α. This quantity is readily available since it is necessary for the calculation of pseudopotentials, and is typically stored on a radial, logarithmic grid. Thus, careful numerical integration can be performed because of the finer grid spacing near the nuclei. The projection of Eq. (56) should be done explicitly for each point on the 2D grid fine enough for a microscopy calculation.

It is worthwhile to note here that the singularity arising from V_i^α poses a further computational challenge for the modelling of electron microscopy experiments. Depending on how close the nuclei of the atoms lie to the nearest grid point, the value of the projected potential can vary greatly. Large changes in the value of V_z at the single grid point nearest the nucleus can translate into noticeable artifacts in the final image, thus violating translational symmetry. Our method of preserving the translational invariance of the sample with respect to the grid in electron microscopy simulations is discussed in detail in [34], in which we demonstrate that approximating the projected potential near the nucleus as the projection of a Yukawa potential and taking the average value of this over the region surrounding that grid point

fixes the problem. This technique can be implemented as a post-processing step once the full projected potential has been obtained. All projected potentials used in the generation of images in this paper have been treated in this way.

IV. RESULTS

A. Comparison of projected potentials and microscopy image simulations between IAM and DFT

The necessity of using projected potentials obtained from DFT calculations rather than those made from the IAM is exemplified in the simple case of a monolayer of hexagonal boron nitride (h-BN). It has been shown through comparison with experiment that the accuracy of simulated electron microscopy images is greatly improved by using DFT rather than IAM projected potentials as input [6, 7]. Since most of the electrons of both boron and nitrogen are valence electrons, the effects of neglecting charge redistribution in the chemical environment in the IAM are readily visible in both the projected potentials and the microscopy images resulting from them when compared with DFT potentials and images.

The SCF quantities were obtained from our implementation of the method outlined in section III in the form of a software tool developed for use with the plane-wave pseudopotential DFT code CASTEP. The IAM quantities were also obtained from CASTEP. CASTEP generates its own (ultrasoft) pseudopotentials on the fly. To do this, the electronic structure of an isolated atom of each species in the unit cell is solved for using all-electron DFT. These isolated atomic quantities are easily accessed within CASTEP, and from them projected potentials are calculated. Since each atom is neutral in this case, there is no need for auxiliary background charge distributions and so the more complicated mathematical treatment of section III is avoided.

In comparing SCF projected potentials and the related images with similar IAM quantities generated within the same DFT implementation rather than the conventional parametrized Hartree-Fock IAM potentials, we avoid complicating our analysis with additional variables such as the accuracy of the parametrizations or differences between theoretical frameworks. Thus the differences shown in Figs. 1 and 2 are due entirely to changes in the electronic structure of the sample.

The CASTEP calculation was found to converge with a plane wave cutoff energy of 1100 eV and a 45 k-point grid, and the default on-the-fly ultrasoft pseudopotentials were used. Our choice of functional was PBE. An orthorhombic unit cell was used to make the calculation of the projected potentials straightforward, taking advantage of the regular Cartesian nature of the 3D numerical grid on which all valence quantities are stored. This

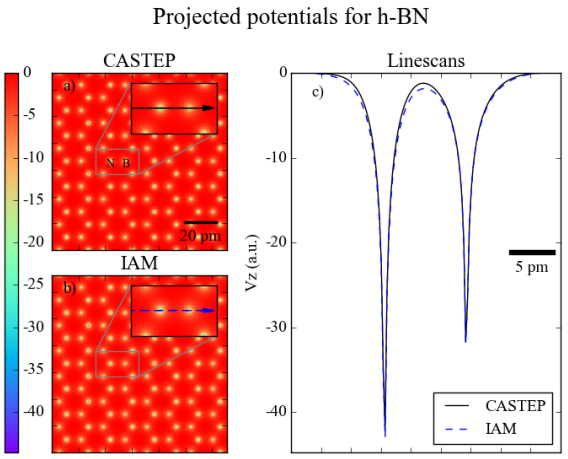


FIG. 1. *A comparison of the projected potentials of monolayer h-BN. In a) and b) we show the supercells of the projected potentials obtained from a fully converged SCF calculation using the PPW DFT code CASTEP and an IAM calculation, respectively. A nitrogen and a boron atom are labeled in a) for reference, and the scale bar in a) applies to b) as well. The color bar shows the strength of the projected potential in atomic units ($a_0 \cdot E_h/e$). The insets show the unit cell used in the DFT calculation in a) and the calculation of the IAM projected potentials in b), with the arrows indicating the position and direction of the linescans shown in c).*

orthorhombic unit cell contained four atoms and had a width and length of 4.7313 a.u. and 8.1948 a.u. The calculation was also converged with respect to the height of the unit cell, so as to be careful not to include the effects of periodic replicas of the monolayer. The converged value of the height was found to be 20 a.u.

In Fig. 1 we compare the projected potentials from a full SCF calculation with IAM projected potentials. From the comparison of the linescans in Fig. 1 a), the most obvious difference is that the potential around the SCF nitrogen atom is shallower than its IAM counterpart. This can be interpreted as a result of stronger screening of the nitrogen atom in the SCF case relative to the IAM case due to the ionic nature of the bonding in h-BN.

Our electron microscopy image simulations were calculated for a microscope in CTEM mode with a beam energy of 80 kV. We used a coefficient of spherical aberration value of 0.02 mm and the corresponding Scherzer defocus of 111.2 Å. A 4×7 supercell of the projected potential was used as input, so as to make the length and width of the supercell approximately equal in order to minimize artifacts in the images. The unit cell was 77×132 pixels, and the resulting grid spacing gives a scattering angle that extends out to 640 mrad.

The CTEM images resulting from the SCF and IAM projected potentials (from Fig. 1) are shown in Fig. 2.

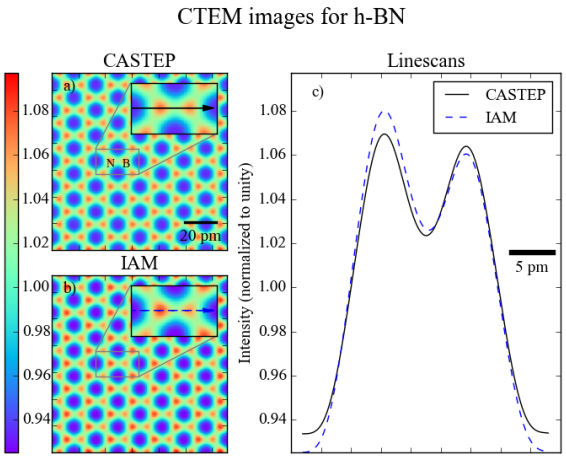


FIG. 2. *A comparison of the simulated 80 kV CTEM images of monolayer h-BN. In a) and b) we show the images from a CTEM simulation performed on the projected potentials shown in Fig. 1 a) and b), respectively. The color bar shows the intensity of the CTEM image, which is normalized over the supercell to unity. In c) is shown the comparison of the linescans indicated by the arrows in the insets of a) and b).*

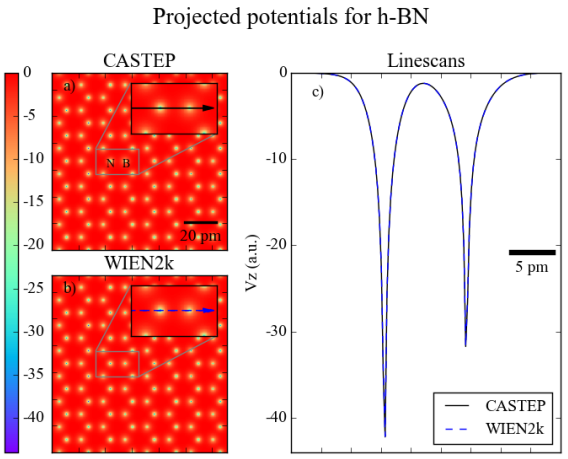


FIG. 3. *A comparison of the projected potentials of monolayer h-BN. In a) and b) we show the supercells of the projected potentials obtained from a fully converged SCF calculation using the PPW DFT code CASTEP and the all-electron DFT program WIEN2k, respectively. A nitrogen and a boron atom are labeled in a) for reference, and the scale bar in a) applies to b) as well. The color bar shows the strength of the projected potential in atomic units ($a_0 \cdot E_h/e$). The insets show the unit cell used in the DFT calculations in a) and b), with the arrows indicating the position and direction of the linescans shown in c).*

We can see that the "charge transfer" that results from an increase in charge density around the nitrogen atoms and a related decrease in the charge density around the boron atoms in h-BN, relative to their isolated atomic charge densities, results in corresponding changes in the intensities near these atomic sites in the CTEM image. In this way, DFT predicts that it is possible to observe directly the difference in electronic structure that results from the adjustment of the valence charge densities of the atoms due to their chemical environment.

Another observable difference between the IAM and SCF intensities is the small reduction of contrast in going from the former to the latter. This can be interpreted as the result of less electronic charge density occupying the interstitial regions in the centers of the hexagons, or equivalently the adherence of the electronic charge density to the symmetry of the system. In the IAM, the charge densities of each atom are artificially forced to be spherically symmetric, resulting in spurious electronic charge density occupying the interstitials of the hexagons. When the electrons are allowed to relax through the SCF calculation, they tend to redistribute themselves closer to the atomic sites and in the regions directly between nearest neighbors in the lattice. This results in greater screening of the nuclear potentials in the SCF case, so that the potential towards the middle of the hexagons is less negative.

B. Comparison of projected potentials and microscopy image simulations between all-electron and PPW DFT

While the lower computational cost and ease of use of PPW codes make them appealing for their widespread application in the microscopy community, it is important to know what effect the approximations made within PPW DFT have on microscopy images. To this end, we compare here the projected potentials and CTEM images of CASTEP with WIEN2k, an all-electron DFT code.

To obtain projected potentials from WIEN2k, we used a modified version of the wien2venus script [35], also used in [6, 7]. We found a converged value of 7 for $R_{MT} \cdot K_{max}$ and a grid of 209 k-points. We used the same unit cell dimensions as in our CASTEP calculation, and also used the PBE functional in both instances.

In Fig. 3, we compare the projected potentials of these two codes. The difference in the projected potential at most points is on the order of 10^{-2} a.u.

With such small numerical differences in the projected potentials, one would expect the resulting CTEM images to be indistinguishable. Indeed this is what is seen in Fig. 4. Close agreement between AE and PPW methods was also observed in the quantitative analysis of simulated ptychographic (S)TEM images in [36]. Thus the precision between these very different theoretical implemen-

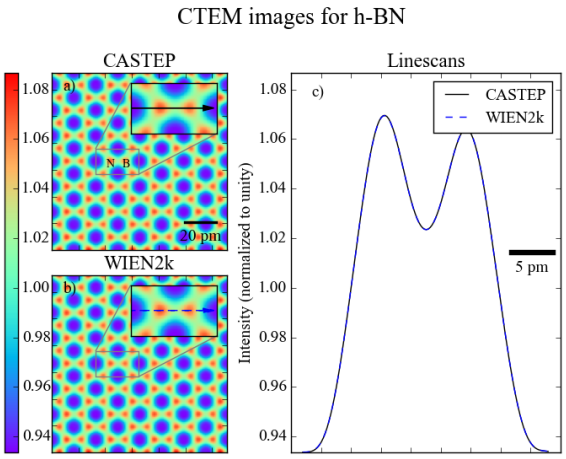


FIG. 4. *A comparison of the simulated 80 kV CTEM images of monolayer h-BN. In a) and b) we show the images from a CTEM simulation performed on the projected potentials shown in Fig. 3 a) and b), respectively. The color bar shows the intensity of the CTEM image, which is normalized over the supercell to unity. In c) is shown the comparison of the linescans indicated by the arrows in the insets of a) and b).*

tations of DFT is very high with respect to the projected potentials they produce, at least for the particular case of h-BN. The precision of (S)TEM images from all-electron and PPW codes over a wide range of materials is beyond the scope of the present work. However, a qualitative argument can be made in support of the general agreement between the two: The valence charge densities of PPW DFT agree exactly with those of AE DFT outside of a small augmentation region around the nuclei. Within this region, these charge densities generally differ observably. However, the nuclear potential dominates in this region, and so differences in the potential due to changes in the electronic charge density are largely washed out by the strong nuclear singularity. While one might expect the largest differences between AE and PPW projected potentials in materials where core polarization plays an important role, it seems reasonable to expect the close agreement seen in Figs. 3 and 4 to persist over a range of materials.

V. CONCLUSIONS

We have outlined a procedure for obtaining projected potentials from PPW DFT in a general way, for the single-slice approximation explicitly. The generalization of this technique to the multislice case is straightforward and would involve changing only the limits of integration in the last step in the calculation of the ionic and valence projected potentials.

Using fully converged DFT potentials as opposed to the traditional IAM in electron microscopy simulations offers increased accuracy with respect to experiment due to the incorporation of more realistic electronic structure theory. In addition to this, DFT also allows for a deeper physical interpretation of microscopy data. Whereas in the IAM, one is confined to interpreting images primarily in terms of the atomic numbers and positions of the atoms in a material, DFT adds to this the ability to observe the effects of the chemical environment on the electronic structure of the material.

One of the key benefits of using PPW codes specifically in the calculation of the projected potentials is that structural relaxations and projected potentials can be computed inexpensively using the same code, simplifying the process for the user. As we have seen, the use of the approximations underlying the PPW approach is justified in the particular example of h-BN, with reason to believe that this agreement will carry over to a large class of materials.

[1] B. C. McCallum P. D. Nellist and J. M. Rodenburg. Resolution beyond the 'information limit' in transmission electron microscopy. Nature, 374:630–632, 1995.

[2] P. E. Batson, N. Delby, and O. L. Krivanek. Sub-angstrom resolution using aberration corrected electron optics. Nature (London), 418:617–620, 2002.

[3] P. D. Nellist, M. F. Chisholm, N. Delby, O. L. Krivanek, M. F. Murfitt, A. S. Szilagyi, A. F. Lupini, A. Borisevich, W. H. Sides Jr., and S. J. Pennycook. Direct sub-angstrom imaging of a crystal lattice. Science, 305(5691):1741–1741, 2004.

[4] H. Sawada et al. Correction of higher order geometrical aberration by triple 3-fold astigmatism field. J. Electron Microsc., 58:341–347, 2009.

[5] Earl J. Kirkland. Advanced Computing in Electron Microscopy. Springer, New York, USA, 2010.

[6] S. Kurasch, J. C. Meyer, D. Künzel, A. Groß, and U. Kaiser. Beilstein J. Nanotechnol., 2:394–404, 2011.

[7] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Groß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser. Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy and others. Nat. Mater., 10:209–215, 2011.

[8] Michael L. Odlyzko, Burak Himmetoglu, Matteo Cococcioni, and K. Andre Mkhoyan. Atomic bonding effects in annular dark field scanning transmission electron microscopy. I. Computational predictions. J. Vac. Sci. Technol. A, 34(041602):1–11, 2016.

[9] Michael L. Odlyzko, Burak Himmetoglu, Matteo Cococcioni, and K. Andre Mkhoyan. Atomic bonding effects in annular dark field scanning transmission electron microscopy. II. Experiments. J. Vac. Sci. Technol. A, 34(041603):1–8, 2016.

[10] J. M. Zuo. Measurements of electron densities in solids: a real-space view of electronic structure and bonding in inorganic crystals. Rep. Prog. Phys., 67:2053–2103, 2004.

[11] T. Gemming, G. Möbus, M. Exner, F. Ernst, and M. Rühle. Ab-initio HRTEM simulations of ionic crystals: a case study of sapphire. J. Microsc., 190:89–98, 1998.

[12] Budhika G. Mendis. Electron beam-specimen interactions and simulation methods in microscopy. John Wiley Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK, 2018.

[13] David J. Griffiths. Introduction to Quantum Mechanics, 2nd ed. Pearson Prentice Hall, Upper Saddle River, NJ, 2005.

[14] D. R. Hartree. The calculation of atomic structures. Rep. Prog. Phys., 11:114–143, 1947.

[15] I. Lobato and D. Van Dyck. An accurate parameterization for scattering factors, electron densities and electrostatic potentials for neutral atoms that obey all physical constraints. Acta Cryst., A70:636–649, 2014.

[16] P. A. Doyle and P. S. Turner. Relativistic Hartree-Fock X-ray and electron scattering factors. Acta Crystallogr. A., 24:390–397, 1968.

[17] Joseph W. Goodman. Introduction to Fourier Optics. McGraw-Hill, Inc., San Francisco, USA, 1968.

[18] D. K. Saldin and J. C. H. Spence. On the mean inner potential in high- and low-energy electron diffraction. Ultramicroscopy, 55(4):397–406, 1994.

[19] M. Vulović, L. M. Voortman, L. J. van Vliet, and B. Rieger. Ultramicroscopy, 136:61–66, 2014.

[20] Jin-Cheng Zheng, Lijun Wu, and James W. Davenport. On the sensitivity of electron and X-ray scattering factors to valence charge distributions. J. Appl. Cryst., 38:648–656, 2005.

[21] Stefan Löffler et al. arXiv:1606.08329v1 [cond-mat.mtrl-sci, 2016.

[22] M. P. Prange et al. Phys. Rev. Lett., 109(246010), 2012.

[23] Kurt Lejaeghere et al. Reproducibility in density functional calculations of solids. Science, 351(6280):1415, 2016.

[24] S. J. Clark et al. First principles methods using castep. Zeitschrift fuer Kristallographie, 220(5-6):567–570, 2005.

[25] P. Blaha, K. Schwarz, G. Madsen, D. Kvasnicka, and J. Luitz. WIEN2K: An augmented plane wave plus local orbitals program for calculating crystal properties. 2001.

[26] Paolo Giannozzi et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. J. Phys.: Condens. Matter, 21(39):395502, 2009.

[27] G. Kresse and J. Furtm:uller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Phys. Rev. B, 54:11169, 1996.

[28] Jorge Kohanoff. Electronic Structure calculations for solids and molecules. Cambridge University Press, Cambridge, UK, 2006.

[29] P. E. Blöchl. Phys. Rev. B, 50(24), 1994.

[30] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. Phys. Rev. B, 41:7892–7895, 1990.

[31] D. R. Hamann, M. Schlüter, and C. Chiang. Norm conserving pseudopotentials. Phys. Rev. Lett., 43:1494–1497, 1979.

[32] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. Phys. Rev., 136(3B):864–871, 1964.

[33] W. Kohn and L. J. Sham. Phys. Rev., 140(4A), 1965.

[34] T. C. Naginey, Jonathan R. Yates, and P. D. Nellist. The treatment of nuclear singularities in electron microscopy simulations. In preparation.

[35] M. Arai. wien2venus script. `http://www.nims.go.jp/cmsc/staff/arai/wien/venus.html` note =.

[36] G. T. Martinez, T. C. Naginey, C. O'Leary, R. J. Nicholls, J. R. Yates, P. D. Nellist, and T. J. Pennycook. Direct imaging of charge redistribution in monolayer hexagonal boron nitride using electron ptychography in the scanning transmission electron microscope. In preparation.

Chapter 5

The treatment of nuclear singularities in electron microscopy simulations

The treatment of nuclear singularities in electron microscopy simulations

T. C. Naginey, J. R. Yates, P. D. Nellist
Department of Materials, University of Oxford, Parks Road, Oxford, OX1 3PH, U. K.

A new way to handle the singularities due to the nuclei in electrostatic potentials used in the simulation of electron microscopy images is developed and shown to improve precision. This technique has been developed to preserve the translational invariance of the sample in (scanning) transmission electron microscopy (S)TEM image simulations. The singularities that arise in potentials generated from density functional theory (DFT) for use in the modelling of (S)TEM experiments pose a particular threat to this invariance, due to the discrete nature of the output of many DFT codes. We present here a novel method for handling these singularities, the theory of which is motivated by modelling the electrostatic potentials close to the nuclei of the the atoms in the sample with a Yukawa potential in addition to reproducing the correct transmitted wave functions to first order in the (projected) electrostatic potential of the sample. This method is trivial to implement computationally, improves the precision of (S)TEM simulations, and increases the usability of DFT in the workflow of modelling in microscopy.

Keywords: electron microscopy, dft, *ab-initio*

I. INTRODUCTION AND BACKGROUND

Transmission electron microscopy (TEM) images are often quantitatively interpreted through comparison with image simulations [1]. Recent technical developments in transmission electron microscopy have led to images with sufficient sensitivity that the conventional simulation methods that make use of scattering factors based on independent atoms [1] can no longer be used and the effects of charge transfer must be included. Examples of images where the effects of transfer and bonding can be seen include conventional TEM (CTEM) imaging of hexagonal boron nitride and doped graphene [2] and scanning TEM (STEM) imaging of $BaTiO_3$ using differential phase contrast imaging [3]. The enhanced imaging sensitivity has been enabled through the development of hardware correctors for the geometric aberrations of the imaging lenses and through the development of new detector technologies. It is clear that we are now entering an era where image simulations must be based on calculations of the electronic structure of materials, such as through density functional theory (DFT) calculations.

The electrostatic potential plays a similar role for electrons as the refractive index does for light. For a thin sample, the effect of the electrostatic potential is to shift the phase of the incident wavefront by an amount proportional to the projected electrostatic potential. For thicker samples where multiple scattering is important, the electrostatic potential is still used via either a Bloch-wave approach [1] or multislice approach [4].

The multislice method [4] is now dominant, and involves slicing the potential of the sample, V, into several sub-slices using a set of planes perpendicular to the velocity of the incoming electron beam along the optical axis, which is taken to be the z-axis by convention. Within each subslice, we make the assumption that V varies on length scales that are much longer than the wavelength of the fast electron [5]. We can then make the WKB approximation [6] within each slice, which lets us relate the transmitted electronic wave function ψ_t to the incident wave function ψ_i via the potential V:

$$\psi_t = \exp\left\{i\sigma \int_{slice} V dz\right\}\psi_i, \tag{1}$$

where σ is a scalar called the interaction parameter and $\int V dz$ is the integral of the potential along the optical axis over the subslice, *i.e.* the potential projected onto each plane partitioning the sample. This 'projected potential' is the key quantity relating the electronic structure of the sample to the experimentally measured quantities in (S)TEM imaging modes. Eq. (1) is invoked within each plane of the potential of the sample, and then between planes ψ is propagated with the Fresnel propagator [7]. If the sample is thin enough and the potential weak enough, the use of a single slice for the whole of the specimen potential is sufficiently accurate [5]. In this case we have the single two-dimensional scalar function

$$V_z \equiv \int_{-\infty}^{+\infty} V(x, y, z) dz, \tag{2}$$

coming from the sample which enters the (S)TEM simulation. For simplicity, this approximation is used for the derivations in the present work. Generalizing to the multislice case is straightforward.

In most (S)TEM modeling software, the projected potential of the sample is constructed from a simple superposition of projected potentials of isolated atoms. These independent atom model potentials come from Hartree-Fock calculations performed on individual atoms which are then parametrized using basis functions with fitted parameters [8, 9]. Thus in the independent atom model the projected potential is constructed from analytic functions which can be rapidly evaluated at any point. The price paid for this computational convenience is the loss in physical accuracy due to the complete absence of interactions between electrons of neighboring

atoms. DFT offers a good balance between accuracy and computational efficiency in approximating the solution to the many body problem [10, 11]. Obtaining V_z from a DFT code is thus a relatively straightforward way of improving upon the independent atom model by taking into account in a self-consistent fashion the interactions between electrons of neighboring atoms. Indeed, it has been shown that using DFT in place of the more conventional independent atom model leads to crucial improvements in the interpretation of (S)TEM images for certain materials [2, 12–15].

II. NUMERICAL TREATMENT OF THE SINGULARITY

If the nucleus is modelled as a point particle, then each atomic potential has a divergence due to the $Ze/4\pi\epsilon_0 r$ singularity of the nuclear potential, where Z, e, and ϵ_0 are the atomic number, electron charge and permittivity of free space, respectively. The nuclei present a numerical problem, since V_z is calculated pointwise on a regular grid for each simulation. The value of V_z at each grid point is assigned to a pixel with dimensions equal to the distance between grid points. The spacing of this grid is ideally equal in x and y and is determined by the maximum scattering angle detectable by the microscope. The value of V_z at a given grid point grows arbitrarily large the closer the nucleus happens to lie to that grid point. This can give rise to non-physical variations in the simulated images due to translations of the sample relative to the grid used in the simulation. It may naively be assumed that the value of V_z over a single pixel would not have a major effect on the overall image, but the effects are long range and can be quite significant due to the convolution with the point spread function in calculating the image. An example of the effect that changing the value of V_z at the single pixel containing the nucleus has on the final simulated image is shown in Fig. 1 for a single hydrogen atom.

There are several ways that this nuclear singularity is handled in (S)TEM codes that employ the independent atom model. One way is to bandwidth limit the potentials according to the length/width of the two-dimensional grid in reciprocal space [1]. Another way is to interpolate V with a cubic spline within a small volume around the nucleus, the length scale of the volume being determined by the grid spacing in real space (Ivan Lobato, private communication). The principle behind each approach is that there is a limit to the spatial frequencies that can be detected with the microscope due to the experimental set-up (limited collection angles, pixel size, beam width, etc.) and so, in simulations, V should not contain any spatial frequencies above this limit if aliasing is to be avoided [16].

Both the interpolation and bandwidth limiting techniques used to handle the violation of translational invariance in the independent atom model case depend on

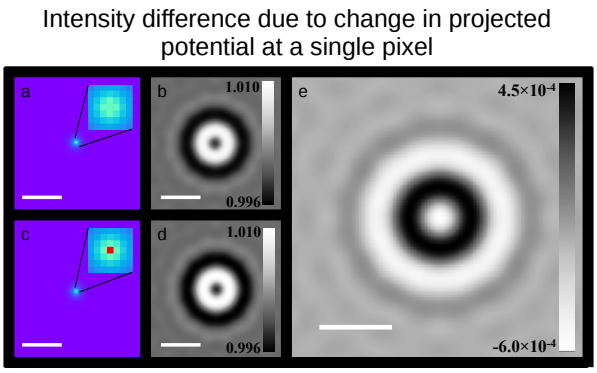


FIG. 1. *Analysis of the effect that the value of V_z at a single pixel has on a conventional TEM image. The scale bar is 2.5 Å. a) and c) show the analytic projected potential of a (non-relativistic) hydrogen atom, the only difference being that the value of the pixel containing the nucleus in a) was taken to be equal to the next-nearest grid point and in c) it was taken as the value of the analytic V_z at a distance 5×10^{-5} Åfrom the nucleus. The width of the pixels is about 5×10^{-2} Å. b) and d) show the simulated TEM phase contrast images resulting from a) and c), respectively. We used a beam energy of 80 keV, coefficient of spherical aberration of 0.02 mm, defocus of 180 Å, and a maximum spatial frequency k_0 of 0.81 $Å^{-1}$ in our simulation. e) shows the difference between b) and d). The contrast in e) was enhanced for visualisation. The range of contrast in e) is about 8% of that in b) or d).*

V_z having an analytic form. Thus these methods do not translate well to DFT, where either the basis functions used are far more complicated or V simply consists of numerical values on a discrete grid. This problem has been addressed previously. For example, in the studies on TEM images of defects and adatoms in various 2D materials in [2, 15], projected potentials obtained from the all-electron DFT code WIEN2k [17] were used. To deal with the nuclear singularities, the authors combined two techniques. First, they took care to ensure that no grid points coincided exactly with the position of a nucleus by carefully choosing the grid spacing and atomic coordinates. Second, they used a fixed cutoff value in the 3D electrostatic potential of the sample before integrating through the z-axis.

This problem is not limited to potentials obtained from WIEN2k. The authors of the current work have been developing a tool for CASTEP, a pseudopotential plane-wave DFT code [18]. Within CASTEP the potential close to the nucleus is represented by the potential of a finite nucleus. Thus outside of a small ($\sim 5 \times 10^{-5}$ Å) radius, the nuclear potential is $Ze/4\pi\epsilon_0 r$, and within this radius, the potential has a smooth analytic form corresponding to a conveniently chosen charge distribution within the small spherical nucleus. In this way, nuclear singularities are avoided in CASTEP calculations. However, the values of V_z close to the nucleus are still very

large, again giving rise to a lack of translational invariance.

It would be of benefit to the microscopy community in general to make the DFT calculations for obtaining V_z as simple as possible and treat the singularities in such a way that the proximity of an atom relative to the numerical grid does not need to be taken into consideration by the user. It would also be desirable to use a technique a bit more physically motivated than a universal cutoff value for the electrostatic potential. Thus we are motivated to come up with a new, general solution to the problem of nuclear singularities when obtaining projected potentials from DFT calculations for use in (S)TEM modelling.

III. A NEW WAY TO HANDLE THE SINGULARITY

The modelling of microscopy images involves convolutions of wave functions representing the beam with functions that embody the optics of the microscope [1, 7]. Numerically we work with a grid of values of these functions at discrete points in space. Pixelated versions of these functions are then made by assigning the value of each point to a pixel. Ideally, the value of each pixel should be the average of the function over that region. Handling the projected potentials in the former way is a good approximation at most points within the sample, where V_z varies on length scales comparable to the pixel size, but very close to nuclei the approximation breaks down.

The mathematics of the fast Fourier transform is based on discretely sampled functions [16]. When taking a Fourier transform of a discrete sampling of a function, aliasing can occur if there are any Fourier components of the function which have spatial frequencies higher than the Nyquist frequency. The result is that the *discrete* Fourier transform of a discretely sampled function is not equal to the discretely sampled Fourier transform of the continuous function that was originally sampled. We would like these two quantities to match as closely as possible. If we focus our attention only on the isolated pixel that contains the nucleus–assuming that the rest of the projected potential is smooth enough that a step-wise representation is a good approximation–and consider the value of V_z over this pixel to be a parameter, we can take the liberty of setting this value so that the discrete Fourier transform of the step-wise approximation matches more closely the discrete sampling of the Fourier transform of the analytic function.

To begin with, consider the simple case of the projected potential of an isolated atom, $V_z(\boldsymbol{x})$, centered at the origin, where $\boldsymbol{x}$ is the 2D vector in the x-y plane. We construct another function $V_z'(\boldsymbol{x})$, which is equal to $V_z(\boldsymbol{x})$ everywhere except over a small region of width Δ_x and length Δ_y around the nucleus, over which it has a constant value of V_z^0:

$$V_z'(\boldsymbol{x}) = \begin{cases} V_z(\boldsymbol{x}), & |x| \geq \frac{\Delta_x}{2} \vee |y| \geq \frac{\Delta_y}{2} \\ V_z^0, & |x| < \frac{\Delta_x}{2} \wedge |y| < \frac{\Delta_y}{2}. \end{cases} \tag{3}$$

The values of Δ_x and Δ_y are the pixel widths in each dimension. Working in the single slice approximation for clarity, we set out to minimize the difference in the annular dark field (ADF) STEM images produced by the two projected potentials in Eq. (3) by choosing an appropriate value for V_0.

The intensity g of a pixel in an ADF STEM image for probe position $\boldsymbol{x}_p$ is given by

$$g(\boldsymbol{x}_p) = \int D(\boldsymbol{k}) \left| \mathscr{F} \left[\psi_p(\boldsymbol{x}, \boldsymbol{x}_p) t(\boldsymbol{x}) \right] \right|^2 d\boldsymbol{k}, \tag{4}$$

where ψ_p is the probe function, $t(\boldsymbol{x}) = \exp\{i\sigma V_z(\boldsymbol{x})\}$ is the transmission function and $D(\boldsymbol{k})$ is the detector function, the latter being given by

$$D(\boldsymbol{k}) = \begin{cases} 1, & |\boldsymbol{k}| \geq |\boldsymbol{k}_{\min}| \ \wedge \ |\boldsymbol{k}| \leq |\boldsymbol{k}_{\max}| \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

For our derivation, it will be useful to put Eq. (4) into the following form:

$$g(\boldsymbol{x}_p) = \int \left| \mathscr{F} \left[d(\boldsymbol{x}) \otimes \{ \psi_p(\boldsymbol{x}, \boldsymbol{x}_p) t(\boldsymbol{x}) \} \right] \right|^2 d\boldsymbol{k}, \tag{6}$$

where we have used the convolution theorem and the fact that $|D(\boldsymbol{k})|^2 = D(\boldsymbol{k})$, with $d(\boldsymbol{x})$ being the inverse Fourier transform of $D(\boldsymbol{k})$. We wish to minimize the difference between the intensity resulting from V_z and that resulting from V_z':

$$\begin{aligned} g(\boldsymbol{x}_p) - g'(\boldsymbol{x}_p) = \int & \left| \mathscr{F} \left[d(\boldsymbol{x}) \otimes \{ \psi_p(\boldsymbol{x}, \boldsymbol{x}_p) t(\boldsymbol{x}) \} \right] \right|^2 \\ & - \left| \mathscr{F} \left[d(\boldsymbol{x}) \otimes \{ \psi_p(\boldsymbol{x}, \boldsymbol{x}_p) t'(\boldsymbol{x}) \} \right] \right|^2 d\boldsymbol{k}. \end{aligned} \tag{7}$$

Since ψ_p varies on a much larger length scale than t or t' over the region near the nucleus, we can take it outside of the convolution and Eq. (7) becomes

$$\begin{aligned} g(\boldsymbol{x}_p) - g'(\boldsymbol{x}_p) \approx \int & \left| \mathscr{F} \left[\psi_p(\boldsymbol{x}, \boldsymbol{x}_p) \{ d(\boldsymbol{x}) \otimes t(\boldsymbol{x}) \} \right] \right|^2 \\ & - \left| \mathscr{F} \left[\psi_p(\boldsymbol{x}, \boldsymbol{x}_p) \{ d(\boldsymbol{x}) \otimes t'(\boldsymbol{x}) \} \right] \right|^2 d\boldsymbol{k}. \end{aligned} \tag{8}$$

Thus it is evident that if we choose $t'(\boldsymbol{x})$ such that

$$d(\boldsymbol{x}) \otimes t(\boldsymbol{x}) = d(\boldsymbol{x}) \otimes t'(\boldsymbol{x}), \tag{9}$$

we would have $g(\boldsymbol{x}_p) - g'(\boldsymbol{x}_p) = 0$. Writing out the convolution explicitly, we can rewrite Eq. (9) as

$$\int_{-\infty}^{\infty} d(\boldsymbol{x} - \boldsymbol{x}') \left(t(\boldsymbol{x}') - t'(\boldsymbol{x}') \right) d\boldsymbol{x}' = 0. \tag{10}$$

Taylor expanding t and t' to second order, we write the LHS of Eq. (10) as

$$\int_{-\frac{\Delta_x}{2}}^{\frac{\Delta_x}{2}} \int_{-\frac{\Delta_y}{2}}^{\frac{\Delta_y}{2}} i\sigma d(\boldsymbol{x}-\boldsymbol{x}') \left(V_z(\boldsymbol{x}') - V_z^0\right) dx'dy', \quad (11)$$

using the definition of $V_z^{'}$ from Eq. (3). If we assume that d varies on a much longer length scale than $\Delta_{x,y}$, we can then retain only the first term in its Taylor expansion to arrive at

$$i\sigma d(\boldsymbol{x}) \int_{-\frac{\Delta_x}{2}}^{\frac{\Delta_x}{2}} \int_{-\frac{\Delta_y}{2}}^{\frac{\Delta_y}{2}} \left(V_z(\boldsymbol{x}') - V_z^0\right) dx'dy'. \quad (12)$$

The expression in Eq. (12) can be set to zero if we choose

$$V_z^0 = \frac{1}{\Delta_x \Delta_y} \int_{-\frac{\Delta_x}{2}}^{\frac{\Delta_x}{2}} \int_{-\frac{\Delta_y}{2}}^{\frac{\Delta_y}{2}} V_z(\boldsymbol{x}) d\boldsymbol{x}. \quad (13)$$

The condition in Eq. (13) then minimizes Eq. (7), subject to the assumptions and approximations made above. This condition Eq. (13) amounts to setting the value of the V_z at the grid point nearest the nucleus equal to the average value of V_z over the pixel containing the nucleus. Thus we will use

$$V_z^0 = \frac{1}{\Delta_x \Delta_y} \int_{x_i-\frac{\Delta_x}{2}}^{x_i+\frac{\Delta_x}{2}} \int_{y_i-\frac{\Delta_y}{2}}^{y_i+\frac{\Delta_y}{2}} V_z(x,y) dydx \quad (14)$$

for the value of the projected potential at the grid point closest to the nucleus of each atom in the sample, where (x_i, y_i) are the coordinates of the center of the pixel containing the i-th atom and $\Delta_{x,y}$ is the pixel width along the x or y axis. The condition given by Eq. (13) also minimizes the difference in intensities in the conventional TEM (CTEM) case by a similar mathematical argument. Qualitatively, we can argue that this must be the case, since ADF STEM is more sensitive to the nuclear potential than conventional bright field TEM.

However, Eq. (13) requires knowledge of the functional form of the projected potential of the sample close to the nuclei. For reasons stated above, we don't have access to this in a DFT code. To get a good approximation of this in the single slice case, we consider the behavior of the projected potential of an *isolated* atom, close to the nucleus.

IV. ANALYTIC APPROXIMATION OF THE PROJECTED POTENTIAL CLOSE TO THE NUCLEUS

We begin by modelling the potential of the atom simply as a Yukawa potential,

$$V^Y(r) = Z\frac{e^{-\alpha r}}{r}, \quad (15)$$

with Z the atomic number and α a parameter determining the screening length. The projected potential that results from Eq. (15) is

$$V_z^Y(\rho) = 2ZK_0(\alpha\rho), \quad (16)$$

where K_0 is the zeroth order modified Bessel function of the second kind and $\rho = \sqrt{x^2+y^2}$. The Yukawa potential was chosen because its projection takes a simple analytic form, which is not true of a bare ionic potential $\frac{Z}{r}$. The projection of $\frac{Z}{r}$ is infinite everywhere, as can be seen by setting $\alpha = 0$ in Eq. (16). For small values of the product $\alpha\rho$ a Taylor expansion of Eq. (16) gives

$$2ZK_0(\alpha\rho) \approx -2Z\left(\gamma + \ln\frac{\alpha}{2} + \ln\rho\right), \quad (17)$$

where γ is Euler's constant.

Let ρ_0 be the radius of the area close to the nucleus over which we will use the approximation Eq. (17). Then we use the value of the actual projected potential at this radius $V_z(\rho_0)$ as our boundary value to determine the screening constant α in Eq. (17):

$$2ZK_0(\alpha\rho_0) = V_z(\rho_0) \rightarrow \alpha = \frac{2}{\rho_0}\exp\left(\frac{V_z(\rho_0)}{2Q} - \gamma\right). \quad (18)$$

Substituting this value of α into Eq. (17) yields what we will use as the approximate projected potential for an atom at small distances from the nucleus:

$$V_z(\rho < \rho_0) \approx V_z(\rho_0) - 2Z\ln\frac{\rho}{\rho_0}. \quad (19)$$

This approximation is compared with several IAM projected potentials from the parametrization of [19] in Fig. 2, using $\rho_0 = 0.1\ a_0$, or about 0.0529 Å.

The agreement between Eq. (19) and the projected potentials of isolated atoms encourages the application of our method to projected potentials obtained from DFT calculations. In doing this, we are making the assumption that the behavior of V_z at very small distances from the nucleus for isolated atoms is similar for atoms in the sample. Since the form of V_z is dominated by the divergence in the nuclear potential at these small distances from the nuclei, this seems like a reasonable approximation. For projected potentials obtained from DFT calculations, we will use the value of V_z at one of the eight grid points surrounding the grid point closest to the nucleus for $V_z(\rho_0)$ in Eq. (19). We now have an asymptotic functional form for the projected potential that we can integrate to determine V_z^0 for the grid point nearest each atom in the sample. Putting Eq. (19) into Eq. (14) we have the expression

$$V_z^0 = [V_z(\rho_0) + 2Z\ln\rho_0] - \frac{2Z}{\Delta_x\Delta_y}\int \ln\sqrt{(x-x_0)^2+(y-y_0)^2}\, dA, \quad (20)$$

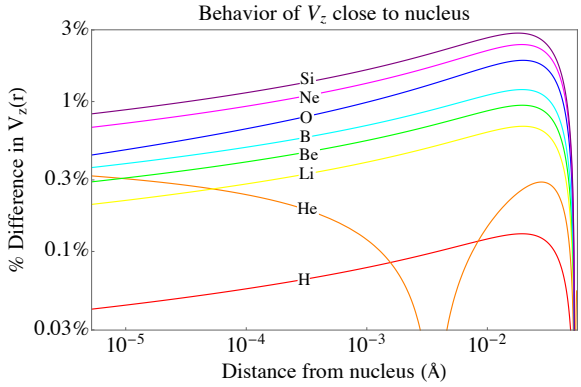


FIG. 2. *Log-log plot of the absolute value of the difference between the independent atomic projected potentials of [19] and our approximation in Eq. (19), as a percentage of the independent atomic potentials. Data is shown for the first five elements of the periodic table along with O, Ne and Si. The value of ρ_0 was fixed at 0.1 $a_0$for each element. There is a sign change in the difference between the two potentials for helium, but the percentage difference still does not exceed one percent.*

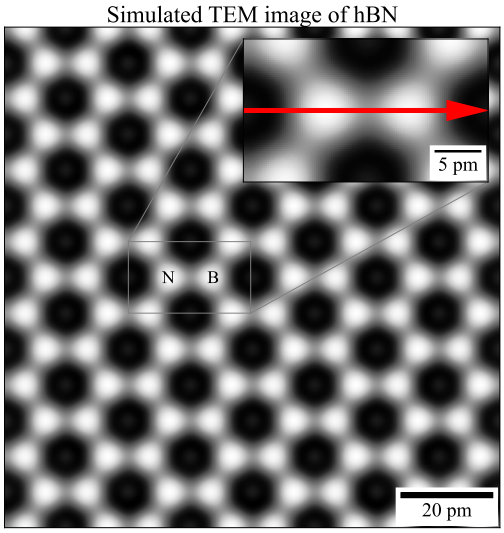


FIG. 3. *A representative supercell generated by our TEM software. This particular image is the intensity resulting from a corrected CASTEP calculation. The linescans plotted in the remaining images were all taken along the arrow shown. Atomic species are labelled in the unit cell shown outlined in the small rectangle, and its dimensions are* 4.34×2.50 *Å*^{2}.

for the value of V_z for the grid point nearest an atom of charge Z and coordinates (x_0, y_0), where the integral is over the area of the pixel containing the atom.

For our implementation, it was possible to evaluate only one dimension of the integral in Eq. (20) analytically so that we were left with a 1-D numerical integral. The removable singularity at $x = x_0$, $y = y_0$ is handled numerically by adding a small offset to the argument of the integrand. This does not change the value of the integral appreciably, since the singularity is removable and converges so long as the offset is small enough. We obtain the value of $V_z(\rho_0)$ from the furthest of the four second-nearest neighboring grid points from the position of the atom. This guarantees that the sign of the integrand in Eq. (20) does not change over the pixel.

V. RESULTS

We have tested how well our method improves the precision in basic conventional TEM image simulations generated from projected potentials from two different DFT codes, CASTEP and WIEN2k. The projected potentials from CASTEP were obtained using a software tool developed by the authors, and from WIEN2k they were obtained by using a modified version of the wien2venus script from [20], similar to the method used in [2, 15]. As in [15], only the defocus and spherical (first order) aberration effects of the microscope optics were accounted for in the generation of the TEM phase images. We have used a value of 0.02 mm for the coefficient of spherical aberration, a beam energy of 80 keV, and a value of -11 nm for defocus (extended Scherzer defocus). The TEM microscopy simulations were performed using a python script developed in-house, which omitted thermal diffuse scattering.

In the CASTEP calculations, we used a plane wave cutoff energy of 1100 eV and the default on-the-fly ultrasoft pseudopotentials. For the WIEN2k calculations, we used a value of 7 for $R_{MT} \cdot K_{max}$. For all DFT calculations, the height of the unit cell was fixed at 10.6 Å, the maximum spacing of the k-point grid was 0.1 Å^{-1}, and the PBE functional was used.

Hexagonal boron nitride (h-BN) was chosen as our test sample for this study. Our basic approach was to compare images generated using slightly different pixel sizes while keeping all other parameters equal. In this way, the differences in the images results almost entirely from the positions of the grid points relative to the atomic coordinates. The projected potentials were first calculated for an orthorhombic unit cell of h-BN containing four atoms (see Fig. 3) with dimensions 4.34×2.50 Å^2. We did not use the primitive hexagonal unit cell since orthorhombic unit cells are more convenient for evaluating the projected potential. The number of pixels in the shorter dimension was fixed at 77, while we varied the number of pixels in the longer dimension from 132 (in which case the nitrogen and boron nuclei on the linescan of the unit cell each lay directly on a grid point) to 133 (in which case the same nuclei lay in between grid points). Linescans of two CASTEP calculations with the number of pixels varied in the manner described above are shown in Fig. 4. Since the number of pixels changes only by

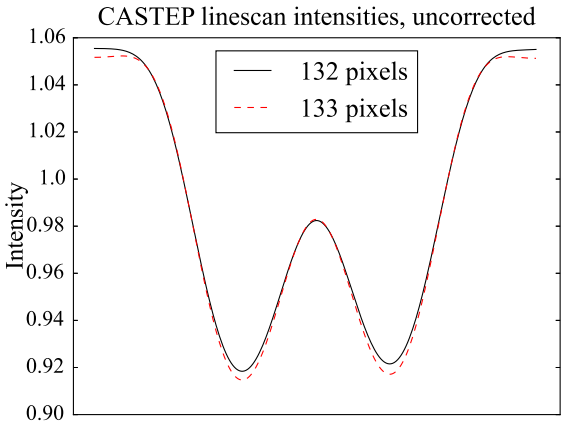


FIG. 4. *Linescans of simulated TEM images of hBN, using projected potentials from CASTEP calculations. The solid line is the intensity when the number of pixels/grid points along the longer edge of the unit cell is 132, while the dashed line corresponds to 133 pixels in this dimension. The nuclear singularities were not corrected for in the projected potentials in either case. The intensity was normalized to unity.*

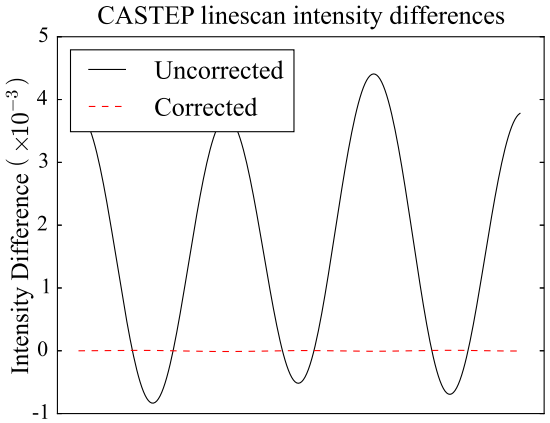


FIG. 5. *Differences in linescans of simulated TEM images of hBN, using projected potentials from CASTEP calculations. The solid line is the difference in intensity on changing the number of pixels/grid points along the longer edge of the unit cell from 132 to 133, using uncorrected projected potentials. The dashed line represents the same difference using corrected projected potentials.*

one, the spatial resolution was remains effectively constant between all calculations. With this grid spacing, the scattering angle extends to about 640 mrad.

The input for the TEM simulations was an approximately square 4×7 supercell, composed of 28 periodic replicas of the original unit cell, so as to make the dimensions of the supercell as close to square as reasonably possible thus minimizing artifacts in the final image. An example of an image from the calculations is shown in Fig. 3. To compare the intensities, we used linescans across the same region in each image, shown in the inset of Fig. 3. We will label the calculations resulting from our method of integrating over the nuclear singularities as "corrected." Images resulting from the projected potentials that have not been averaged over the pixels containing nuclei we will call "uncorrected." The images resulting from the cut-off method used in [15] applied to WIEN2k will be termed the "uncorrected WIEN2k" results. Images resulting from the projected potentials from CASTEP where nuclei have only been treated with the default finite nucleus approximation we will call "uncorrected CASTEP." We have used the same maximum value of the potential for the cutoff in our WIEN2k calculations as was used in [15]–50 kV. All intensities have been normalized so that the integral over the supercell is equal to unity.

In Figs. (5) and (6) the differences between the corrected and uncorrected images resulting from CASTEP and WIEN2k potentials is apparent. In both cases, the method of handling the nuclear singularity presented in the current work has essentially eliminated the differences due to slightly varying the size of the pixels. The largest differences in the intensities of the uncorrected images can be seen to occur in the center of the hexagons of the h-BN sample, for both CASTEP and WIEN2k. This is a consequence of the separation distance between pairs atoms on opposite sides of the hexagons.

It is of interest to compare the effects that the nuclear singularities can have on an image with the effects of bonding and charge transfer. To do this, we compare the difference between an independent atom model image and a DFT image, plotted as a single linescan, with the linescans in Figs. (5) and (6). It is also instructive to compare these with the differences between two different DFT codes. This will give a measure of how important the differences in DFT implementations are in the context of modelling (S)TEM experiments.

In Fig. (7), the linescans of the differences in intensity incurred by small changes in the position of the nuclei relative to the grid points in the uncorrected cases (those labelled "uncorrected" in Figs. (5) and (6)) are compared with the difference between images generated from corrected DFT and independent atom model potentials. This latter difference is essentially a measure of how important bonding effects are in this sample. Also plotted in Fig. (7) is the difference between images resulting from projected potentials from a CASTEP and a WIEN2k calculation, both corrected. It can be seen that while the differences in intensity incurred by small changes in the position of the nuclei relative to the grid points in the uncorrected cases are smaller than the difference between corrected independent atom model and DFT images, they are larger than the the difference between images generated using corrected potentials from the two different DFT codes examined here. This is significant since these two particular codes use very different implementations of DFT; WIEN2k is an all-electron code which treats core electrons explicitly while CASTEP is a plane-

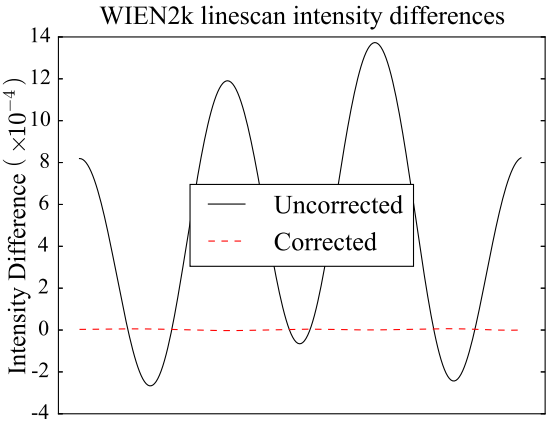


FIG. 6. *Differences in linescans of simulated TEM images of hBN, using projected potentials from WIEN2k calculations. The solid line is the difference in intensity on changing the number of pixels/grid points along the longer edge of the unit cell from 132 to 133, using uncorrected projected potentials. The dashed line represents the same difference using corrected projected potentials.*

wave pseudopotential code which uses the frozen core and pseudopotential approximations. Despite these fundamental differences, we see that the intensity differences between DFT implementations disappears when the singularities are handled consistently with our method.

We conclude that the choice of which type of DFT to use to generate the projected potential is actually less important than the treatment of the nuclear singularity, while the inclusion of bonding effects remains more important than both of these factors. The agreement between CASTEP and WIEN2k images is addressed in [21].

VI. CONCLUSIONS

The method presented here for the treatment of the nuclear singularities in projected potentials has been shown to preserve translational invariance of samples in single-slice electron microscopy image simulations with respect to the grids on which they are computed. While the derivation presented here was specific to conventional TEM, electron microscopy simulations usually involve the same types of integrals as Eq. (14) at some level of approximation–in STEM by reciprocity, in annular dark field (ADF) techniques through the integrals over the annulus, etc. Thus improving the consistency of the integral in Eq. (14) over the pixels containing singularities means improved consistency in the modelling of most imaging modes in microscopy simulations using DFT projected potentials.

The only input needed to handle the singularities in this way is a 2D grid of the projected potential and the

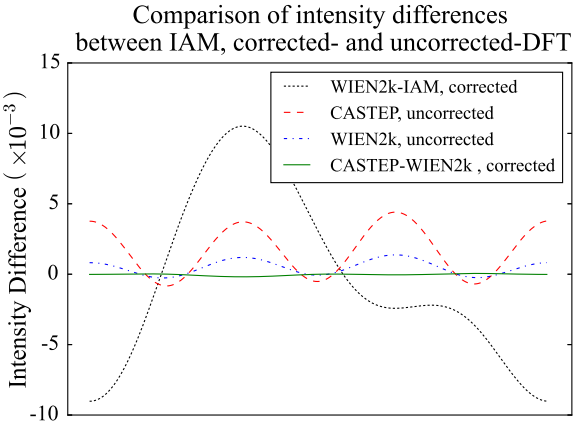


FIG. 7. *Comparison of the uncorrected differences from Figs. 5 (dashed) and 6 (dot-dashed) with the difference between intensities of corrected IAM and DFT (WIEN2k) images (dotted) and corrected calculations of two different DFT codes, WIEN2k and CASTEP (solid). The differences resulting from changing the grid in an uncorrected DFT calculation are smaller than the differences between IAM and DFT images, but larger than the difference between completely different corrected DFT codes.*

coordinates and nuclear charge for each atom in the sample, making the technique general and independent of the code or theory used to generate the projected potentials. The computational cost is minimal, resulting solely from the evaluation of the 2D integral in Eq. (20) for each non-equivalent atom in the sample.

Our method is an improvement over other techniques for handling the singularities because it effectively eliminates the difference between images resulting from small changes in the relative positions of atoms and grid points. The burden on the user is thus reduced since there is no need to take into account atomic positions when choosing exact pixel sizes, making it easier for microscopists to improve their simulations by making use of projected potentials from DFT.

[1] Earl J. Kirkland. Advanced Computing in Electron Microscopy. Springer, New York, USA, 2010.
[2] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Groß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser. Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy and others. Nat. Mater., 10:209–215,

2011.

[3] Y. Kohno H. Sawada Y. Kondo Y. Ikuhara N. Shibata, S. D. Findlay. Differential phase-contrast microscopy at atomic resolution. Nature Phys., 8:611–615, 2012.

[4] J. M. Cowley and A. F. Moodie. The scattering of electrons by atoms and crystals. I. a new theoretical approach. Acta Crystalogr., 10:609–619, 1957.

[5] M. Vulović, L. M. Voortman, L. J. van Vliet, and B. Rieger. Ultramicroscopy, 136:61–66, 2014.

[6] Jr. S. C. Miller and Jr. R. H. Good. A WKB-type approximation to the schrödinger equation. Phys. Rev., 91:174–179, 1953.

[7] Joseph W. Goodman. Introduction to Fourier Optics. McGraw-Hill, Inc., San Francisco, USA, 1968.

[8] P. A. Doyle and P. S. Turner. Relativistic Hartree-Fock X-ray and electron scattering factors. Acta Crystallogr. A., 24:390–397, 1968.

[9] I. Lobato and D. Van Dyck. Ultramicroscopy, 156:9–17, 2015.

[10] P. Hohenburg and W. Kohn. Phys. Rev., 136(3B), 1964.

[11] W. Kohn and L. J. Sham. Phys. Rev., 140(4A), 1965.

[12] S. Mogck, B. J. Kooi, J. Th. M. De Hosson, and M. W. Finnis. Ab-initio transmission electron microscopy image simulations of coherentAg-MgO interfaces. Phys. Rev. B, 70(245427):1–9, 2004.

[13] B. Deng and L. D. Marks. Theoretical strucure factors for selected oxides and their effects in high-resolution electron-microscope (HREM) images. Acta Crystalogr. Section A, 62:208–216, 2006.

[14] Michael L. Odlyzko, Burak Himmetoglu, Matteo Cococcioni, and K. Andre Mkhoyan. Atomic bonding effects in annular dark field scanning transmission electron microscopy. I. Computational predictions. J. Vac. Sci. Technol. A, 34(041602):1–11, 2016.

[15] S. Kurasch, J. C. Meyer, D. Künzel, A. Groß, and U. Kaiser. Beilstein J. Nanotechnol., 2:394–404, 2011.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical Recipes in Fortran, Second Edition. Cambridge University Press, Trumpington Street, Cambridge, CB2 1RP, UK, 1992.

[17] P. Blaha, K. Schwarz, G. Madsen, D. Kvasnicka, and J. Luitz. WIEN2K: An augmented plane wave plus local orbitals program for calculating crystal properties. 2001.

[18] S. J. Clark et al. First principles methods using castep. Zeitschrift fuer Kristallographie, 220(5-6):567–570, 2005.

[19] I. Lobato and D. Van Dyck. An accurate parameterization for scattering factors, electron densities and electrostatic potentials for neutral atoms that obey all physical constraints. Acta Cryst., A70:636–649, 2014.

[20] M. Arai. wien2venus script. `http://www.nims.go.jp/cmsc/staff/arai/wien/venus.html` note =.

[21] T. C. Naginey, R. J. Nicholls, P. D. Nellist, and Jonathan R. Yates. Calculation of projected potentials using pseudopotentials for use in electron microscopy simulations. In preparation.

Chapter 6

Direct imaging of charge redistribution in monolayer hexagonal boron nitride using electron ptychography in the scanning transmission electron microscope

Direct imaging of charge redistribution in monolayer hexagonal boron nitride using electron ptychography in the scanning transmission electron microscope

G T Martinez,[1] T C Naginey,[1] L Jones,[2] C O'Leary,[1] T J Pennycook,[3] R J Nicholls,[1] J R Yates,[1] and P D Nellist[1,*]

[1] *Department of Materials, University of Oxford, Parks Road, Oxford, OX1 3PH, U. K.*
[2] *School of Physics and CRANN, Trinity College Dublin, University of Dublin, Ireland*
[3] *Stuttgart Center for Electron Microscopy, Max Planck Institute for Solid State Research, Stuttgart, Germany*

(Dated: October 4, 2018)

Aberration-free phase imaging of hexagonal boron nitride using Wigner Distribution Deconvolution electron ptychography is presented. The measured phase shift in thin samples is a function of the atomic potential and thus phase images are sensitive to charge redistribution due to bonding. The experimental results demonstrate the direct imaging of charge redistribution in this 2D material. Image simulations that consider independent spherical atoms do not match the experimental data. Incorporating bonding effects is required to explain the experimental data, which is done here using density functional theory (DFT) calculations. We find that both all-electron and pseudopotential plane wave DFT codes deliver identical results and provide image simulations much closer in agreement with the experimental data, though a statistically significant mismatch between theory and experiment is still observed.

Scattering by high-energy electrons can be used not only to determine the location of atoms in structures, but also to detect the redistribution of charge in the sample due to bonding. Here we use electron ptychography to measure the phase shift of an electron wave transmitted through a monolayer of hexagonal boron nitride (hBN). The measurement is accurate and sufficiently precise to provide a local, real-space measurement of charge redistribution due to bonding.

Convergent-beam electron diffraction (CBED) has been the most widely used technique to measure the effects of bonding and charge redistribution, which has proven to be sensitive enough to detect charge density variations in pure metals [1]. CBED can be applied to crystalline samples to accurately measure structure factors, which are related to the variation in charge density in the crystal. The dynamical diffraction that arises in thicker samples leads to intensity variations within the diffracted discs and the structure factors can be fitted with high accuracy to these intensity variations through forward simulations of the diffraction patterns (see for example [2]). For thin crystals, such as the monolayer hBN considered here, the diffracted discs will not show intensity variations, making precise measurements of the structure factors challenging. Furthermore, diffraction does not offer a route to local measurements and so precludes the examination of defects.

Alternative imaging modes can overcome both of these limitations. It has been shown previously that high-resolution conventional transmission electron microscopy (CTEM) can detect the effects of charge redistribution in hBN [3]. Achieving this requires images that are sensitive to the small phase shifts induced in the high-energy electron beam as it transits a thin sample. The approach used in [3] was to deliberately inject lens aberrations, such as defocus, to form an approximation to a Zernicke phase plate, thereby generating a phase contrast image. The resulting image contrast is highly sensitive to the lens aberrations and other imaging parameters, and while qualitative agreement was found, a fully quantitative match to DFT calculations was not attempted in that work. Other methods–such as off-axis holography–can also be used to detect phase shifts, but have a similar sensitivity to residual aberrations.

To provide a local technique with sufficient phase accuracy, we make use of electron ptychography. First proposed as a solution to the phase problem in Hoppe [4], Hoppe and Strube [5], Hoppe [6], the implementation of ptychography in the present work makes use of the atomic-scale focused electron probe used in scanning transmission electron microscopy (STEM). The highly convergent nature of this illuminating beam leads to overlap and interference of waves scattered at different angles, and the interference can be used to solve the phase problem, allowing inversion to an image [7]. Ptychography therefore bridges the gap between diffraction and imaging techniques. Ptychography has previously been applied in overcoming the diffraction resolution limit imposed by the numerical aperture of the objective lens [8, 9]. In the present work, we exploit its ability to provide accurate and precise phase measurement, along with detection and correction of residual lens aberrations arising from small optical misalignments. Recent developments in fast pixelated detector technology [10–12] have enabled the recording of the STEM detector plane for each illuminating probe position, leading to the four-dimensional (4D) data-set that is used for a ptychographical reconstruction. High angle scattering, predominantly from the atomic nuclei, can simultaneously be recorded using an annular dark-field (ADF) detector, leading to a set of images that is sensitive to both atomic number (Z-contrast imaging) and charge redistribution effects [13].

In this work, we employ atomic resolution Z-contrast imaging to confirm that the sample was indeed a mono-

layer, and to provide the crucial unambiguous identification of the boron and nitrogen sublattices. We then utilize simultaneous electron ptychography to correct for small residual aberrations, providing an extremely accurate phase image. This experimental phase image is then compared with simulated images generated from independent spherical atoms–the independent atom model (IAM) which neglects any bonding effects–and two different DFT calculations.

For this study, a sample of monolayer hBN was imaged using a probe corrected JEOL ARM200F microscope. The accelerating voltage was 80 kV with a 31.5 mrad semi-angle of convergence. The 4D data-set was acquired using a JEOL 4DCanvas system incorporating a pnCCD device [10] with 512x512 pixels in the image plane and 66x264 pixels in the detector plane at a detection speed of 4000 frames per second.

The real space sampling in the image plane corresponds to a pixel size of 0.135Å. Non-rigid registration [14] of the entire 4D data set from each of the 16 scan frames was used to overcome scan instabilities, providing higher precision of phase measurements and improving the ADF image contrast. The calculated average dose per frame was found to be $7.75\text{x}10^5 \text{e}/\text{Å}^2$, which leads to a total dose of the non-rigid aligned 4D data-set of $1.24\text{x}10^7 \text{e}/\text{Å}^2$. The ADF image allows the identification of the N and B sub-lattices and the confirmation by comparison with simulations that the sample is indeed a monolayer (see supplemental information).

The non-rigid aligned ptychographic data-set was then reconstructed using the Wigner Distribution Deconvolution (WDD) method [7] and the probe aberrations were determined using the method described in [13]. The retrieved phase images were analyzed using a modified version of the Absolute Integrator software [15], which constructs Voronoi cells around atomic sites over which the phase signal is integrated. We will refer to this quantity as the Integrated Phase Cross-section (IPCS). This quantity was found to be more robust to residual aberrations than the peak phase value (see supplemental information). We subject the distribution of IPCS values for each of the B and N sites to statistical analysis, and we show in figure 2 f) the histogram of the IPCS for the two atom types (N - green and B - blue) for the ptychographic phase image reconstruction, prior to correction of the residual aberrations. These values are plotted with respect to the mean phase obtained in the center of the hexagonal atomic rings, which was assumed to be vacuum and assigned a zero phase. The IPCS values were fitted to a Gaussian function to determine the mean value for each atom type. All images in Figure 2 were analysed in a similar manner.

In hBN, there is a transfer of electronic charge from B to N sites, leading to increased screening of the potentials of the N nuclei [16]. This effect induces a smaller phase shift in the electron wave than the IAM potential. Figure 2 a) shows the reconstructed phase image without aberration correction and its corresponding phase difference histogram in f). It can be seen that the mean IPCS difference between N and B is -1.45 mradÅ^2. After aberration correction, this difference is reduced to 0.98 mradÅ^2 (Figures 2 b) and g)). It is therefore clear that the small residual aberrations, though not observable in the ADF image, are still significant at the level of precision offered by ptychography. In Figure 2 c) through e) we compare the experimental results with image simulations. An in-house code was employed to simulate the CBED patterns which were then used to construct the ptychographical phase images. An electron source probe broadening of 0.95 Å at FWHM was determined to match the mean experimental IPCS of all N and B atoms present in the image. The same source size effect was included in the ADF simulations, which similarly matched the experimental data, bolstering confidence in this value. The width of the histogram peaks for the simulated images is a result of the inclusion of thermal lattice vibrations in our simulations. Figures 2 c) and h) show the results of the WDD of the 4D data-set associated with the IAM. The IAM potentials were obtained from DFT calculations on isolated atoms using the PBE exchange-correlation functional [17]. In using DFT to obtain the IAM potentials, we avoid the need to use parametrized IAM potentials which are commonly used in electron microscopy simulation software [18]. From Figure 2 h), we see that the simulated mean difference between the IPCSs for the two species is 6.88 mradÅ^2, significantly different from the experimental result which is much smaller. We conclude that the experimental results indicate the direct observation of charge transfer, demonstrating the necessity of including bonding effects in our simulations.

To include the effects of bonding we used projected potentials obtained from two different implementations of DFT: the plane-wave pseudopotential code CASTEP [19] and the all-electron LAPW+lo code WIEN2K [20]. A modified version of the wien2venus script [21] was used to calculate projected potentials from WIEN2k. The method used to obtain projected potentials from CASTEP is reported elsewhere [22]. Calculations were performed using an orthorhombic unit cell of monolayer h-BN containing 4 atoms, with lattice parameters $a = 2.50$ Å, $b = 4.34$ Å, and $c = 10.58$ Å. The WIEN2k SCF calculation was performed using a value of 7 for the quantity $R_{MT} \cdot K_{max}$. For the SCF calculation in CASTEP, a plane wave cutoff energy of 1100 eV was used, and pseudopotentials were generated using the ultrasoft scheme [23]. All calculations used the PBE functional [17] and a maximum k-point spacing of 0.095 Å^{-1}. The ptychographic phase image simulations were performed using the same source broadening and phonon configurations used for the IAM simulations. The resulting images are shown in Figure 2 d) and e) with

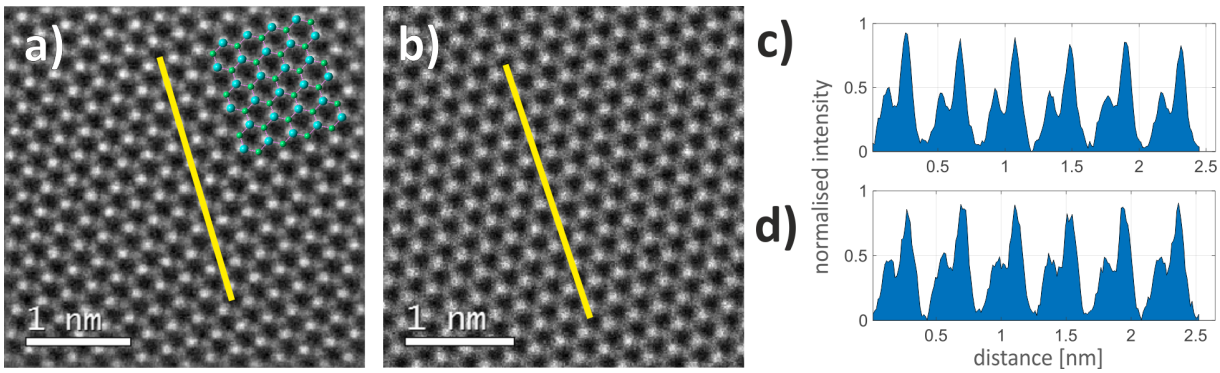


FIG. 1. Monolayer of hexagonal boron nitride. a) Template matching average of experimental image. b) Annular Dark Field image simulation considering experimental conditions. c) Profile of experimental image integrated over the width indicated with the yellow bar in a). d) Profile of simulated image indicated with the yellow line in b) normalised to the average peak pixel intensity for N. The ratio between B and N in experiment is 0.54 and for simulation is 0.56.

their corresponding histograms in i) and j). The analysis of these simulations shows that the results of these two modelling approaches are within statistical error of each other. The difference in the IPCS values for the DFT calculations is 2.5 mrad·Å^2, much closer to the 0.98 mrad·Å^2 observed experimentally than the 6.88 mrad·Å^2 value for the IAM. The DFT calculations have therefore accounted for about 75% of the observed difference between experimental IPCS and the results of the IAM, but there remains a statistically significant mismatch.

In this work, we have presented the experimental observation of charge transfer in hexagonal boron nitride using non-rigid aligned aberration-free phase imaging by means of WDD electron ptychography. Electron ptychography allows for the correction of probe aberrations and delivers accurate phase measurements that can be compared to image simulations. The IAM does not match our experimental results, since it does not take into account the bonding between atoms. We find much better agreement between experiment and simulations that utilize DFT. There remains a statistically significant mismatch between theory and experiment, however, and perhaps the method described here could be used to compare the results of different exchange correlation functionals.

We acknowledge funding from the EPSRC (grant numbers EP/M010708/1, EP/K032518/1 and EP/K040375/1). The authors are thankful to Y. Sasaki from JFCC for providing the sample for this study. Technical support and fruitful discussions with Y. Kondo and R. Sagawa from JEOL Ltd and M. Huth, M. Simsom and R. Ritz from PNDetector GmbH and H. Soltau from PNSensor GmbH in the implementation of 4D Canvas system are also greatly appreciated. TCN is grateful for support from the EPSRC Centre for Doctoral training, Theory and Modelling in Chemical Sciences, under grant EP/L015722/1.

* peter.nellist@materials.ox.ac.uk

[1] P. N. H. Nakashima, A. E. Smith, J. Etheridge, and B. C. Muddle, Science **331**, 1583 (2011).
[2] J. M. Zuo, Reports on Progress in Physics **67**, 2053 (2004).
[3] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Gro, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser, Nat Mater **10**, 209 (2011).
[4] W. Hoppe, Acta Crystallographica Section A **25**, 495 (1969).
[5] W. Hoppe and G. Strube, Acta Crystallographica Section A **25**, 502 (1969).
[6] W. Hoppe, Acta Crystallographica Section A **25**, 508 (1969).
[7] J. Rodenburg and R. Bates, Phil. Trans. R. Soc. Lond. A **339**, 521 (1992).
[8] P. Nellist, B. McCallum, and J. M. Rodenburg, Nature **374**, 630 (1995).
[9] Y. Jiang, Z. Chen, Y. Han, P. Deb, H. Gao, S. Xie, P. Purohit, M. W. Tate, J. Park, S. M. Gruner, *et al.*, Nature **559**, 343 (2018).
[10] H. Ryll, M. Simson, R. Hartmann, P. Holl, M. Huth, S. Ihle, Y. Kondo, P. Kotula, A. Liebel, K. Mller-Caspary, A. Rosenauer, R. Sagawa, J. Schmidt, H. Soltau, and L. Strder, Journal of Instrumentation **11**, P04006 (2016).
[11] M. W. Tate, P. Purohit, D. Chamberlain, K. X. Nguyen, R. Hovden, C. S. Chang, P. Deb, E. Turgut, J. T. Heron, D. G. Schlom, and et al., Microscopy and Microanalysis **22**, 237249 (2016).
[12] M. Krajnak, D. McGrouther, D. Maneuski, V. O. Shea, and S. McVitie, Ultramicroscopy **165**, 42 (2016).
[13] H. Yang, R. N. Rutte, L. Jones, M. Simson, R. Sagawa, H. Ryll, M. Huth, T. J. Pennycook, M. L. H. Green, H. Soltau, Y. Kondo, B. G. Davis, and P. D. Nellist, Nat Comm **7**, 12532 (2016).
[14] L. Jones, H. Yang, T. J. Pennycook, M. S. J. Marshall, S. Van Aert, N. D. Browning, M. R. Castell, and P. D. Nellist, Advanced Structural and Chemical Imaging **1**, 8

(2015).
[15] L. Jones, "Absolute integrator," (2014).
[16] L. Pauling, Proceedings of the National Academy of Sciences of the United States of America **56**, 1646 (1966).
[17] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[18] I. Lobato and D. Van Dyck, Acta Crystallographica A **70**, 636 (2014).
[19] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. Payne, Z. Kristall. **220**, 567 (2005).
[20] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, J. Luitz, R. Laskowski, F. Tran, and L. D. Marks, (Karlheinz Schwarz, Techn. Universitt Wien, Austria). ISBN 3-9501031-1-2 (2018).
[21] M. Arai, "wien2venus.py,".
[22] T. Naginey, R. Nicholls, P. Nellist, and J. Yates, "Obtaining projected potentials from plane wave pseudopotential density functional theory for use in electron microscopy simulations," (2018).
[23] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

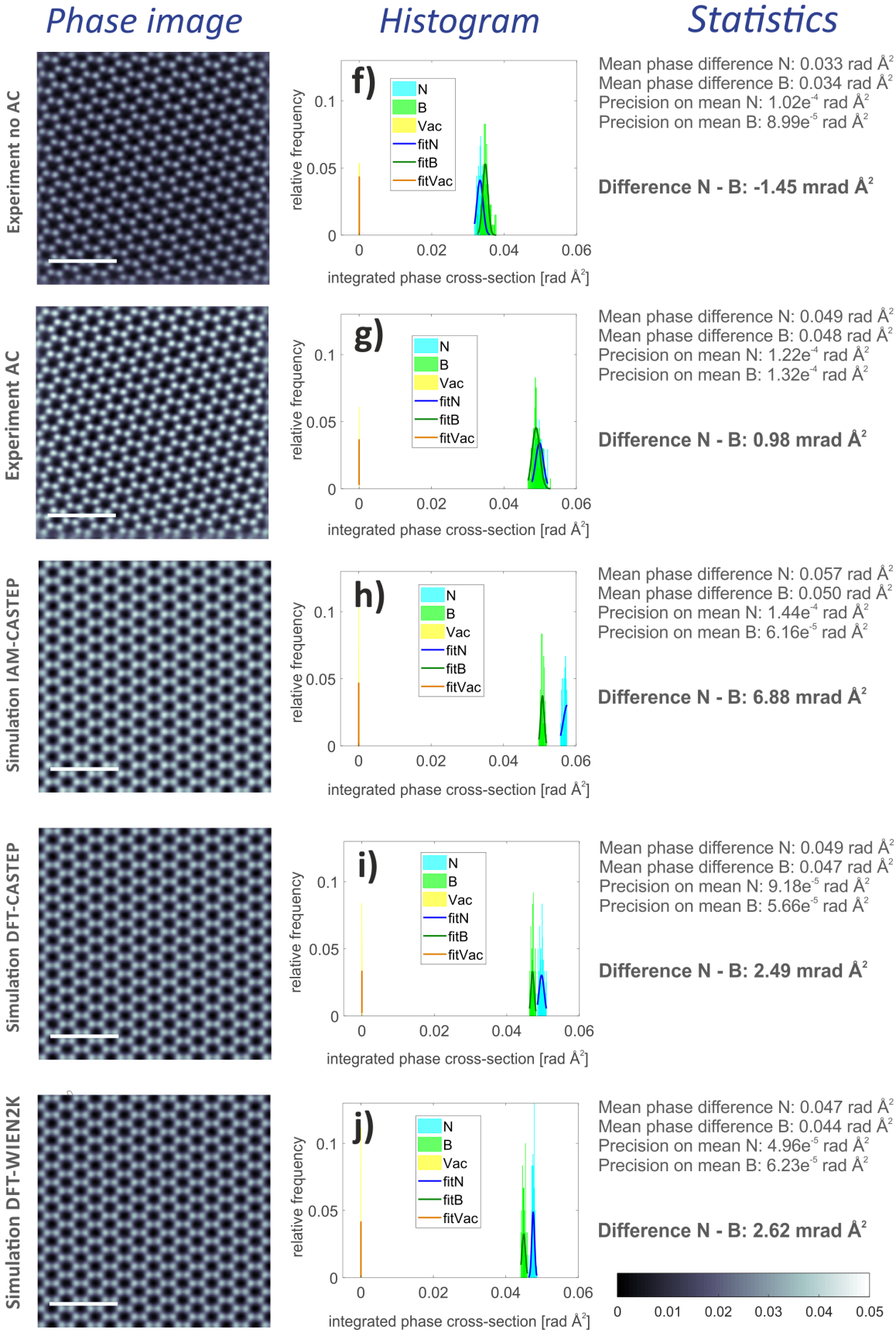


FIG. 2. a) Non-aberration corrected and b) aberration corrected WDD experimental phase images. c), d), e) WDD phase image of simulations considering IAM, DFT-CASTEP and DFT-WIEN2K respectively. f) through j) show the histograms of the IPCS of N and B, with respect to vacuum, obtained from the parametric model fitted to the images in a) to e). The values for the mean IPCS of each atomic species with their corresponding precision are displayed for each of the measurements. Color bar at bottom left is for the phase images, in rad, and the scale bar corresponds to 1 nm.

Chapter 7

Conclusions

We have seen several examples of applications of electronic structure theory to electron microscopy that have facilitated the physical interpretation of experimental data. We have also made the case that, through the development of a user-friendly code that provides access to projected potentials calculated from DFT, many more useful physical interpretations of (S)TEM data could be made by a wide audience of scientists.

In [1], we applied the method of [33] to compare to a much larger range of experimental EELS data relating the energy shift and the number density of high pressure helium bubbles. We discovered and corrected a deficiency in the calculation of excitation cross sections for the helium, but this increased the disagreement between our theoretical values and the experimental ones. The very good agreement of our theory with bulk helium led us to conclude that there is either a systematic problem with the experiments or some new physics present in the bubbles that is absent from the bulk.

In the future we would like to continue to look deeper into this problem by testing the convergence of not only the energies but also the cross sections and mean squared radii of the $2p$ orbitals with respect to cluster size to be more certain about our calculations. We will also attempt to compute energy shifts and cross sections at much higher values of the number density n and possibly collaborate with experimentalists to compare these with a new wider set of data. We have begun work on a second publication in which we will attempt to accomplish these goals.

We have shown in [2] how to obtain projected potentials from PPW DFT in a general way, for the single-slice approximation explicitly. We have attempted to make this method even more precise and robust by developing a technique for preventing artifacts in the image that arise from the singularities in the nuclear potentials [3]. All of the mathematics can be generalized to the case of multislice straightforwardly.

Since the projected potential is the only quantity related to a material that is needed to simulate the corresponding (S)TEM images, this offers a clear path for adding the simulation of (S)TEM images to the capabilities of any PPW DFT software implementation.

We've also explained the advantages of using DFT in general to obtain the projected potentials. Using fully converged DFT potentials as opposed to the traditional IAM offers increased accuracy with respect to experiment due to the incorporation of more realistic electronic structure theory. In addition to this, DFT also allows for a deeper physical interpretation of microscopy data. Whereas in the IAM one is confined to interpreting images primarily in terms of the atomic numbers and positions of the atoms in a material, DFT adds to this the ability to observe the effects of the chemical environment on the electronic structure of the material.

One of the key benefits of using PPW codes specifically in the calculation of the projected potentials is that structural relaxations and projected potentials can be computed inexpensively using the same code. This saves the user–in this case assumed to be a microscopist rather than an electronic structure theorist–the trouble of having to learn to use two different codes, as was the case in [80, 81]. We also demonstrated why it is important to make it easy and inexpensive for more members of the microscopy community to incorporate DFT into their simulations by citing instances beyond systems that exhibit charge transfer in which DFT offered improved accuracy over the IAM. The improved Stobbs factor of graphene is a particular example of this. As we have seen through comparison to AE results, the use of the approximations underlying the PPW approach is justified in the particular example of image simulations of h-BN, and this agreement is expected to carry over to a large class of materials. In the future, it would be instructive to carry out comparisons of images from PPW and AE potentials across a wide range of materials to determine the limits of the PPW approximations.

We saw promising results from implementing our method in CASTEP [4] and applying it to the analysis of ptychographic STEM measurements of hexagonal boron nitride. We were able to show that ptychography is indeed highly sensitive to the effects of charge redistribution. Given this, it would be interesting to use ptychography in conjunction with DFT to study the electronic structure of other systems, perhaps those containing vacancies and defects.

It may also be instructive to examine systems in which charge redistribution effects are not typically considered important. For example, the reduction in the overall contrast of the image seen in [2] when SCF projected potentials are used is not

unique to h-BN. This same effect was observed in [80] for graphene, in which charge transfer between atoms is not a factor. There is a known discrepancy between the absolute contrast (defined as the standard deviation of the intensity of all pixels in an image) of simulated and experimental (S)TEM images, the former sometimes being a factor of two or three larger than the latter. This is referred to as the Stobbs factor problem [91], and using projected potentials from SCF DFT calculations moves the absolute contrast of the simulated image closer to the experimental value, although it is certainly not able to account for all of it. However, the implication remains that the use of IAM projected potentials in (S)TEM image simulations appears to be a contributing factor to the Stobbs factor problem. For example, in [80] a 10% reduction was found in absolute contrast on going from the IAM to SCF DFT generated images of graphene, while in [92] the experimental Stobbs factor (between experimental and IAM-generated simulations) was found to be 20% for the same system, suggesting that up to half of this factor could be accounted for by using more accurate electronic structure theory in the simulations.

The implementation of the method for obtaining projected potentials outlined in [2] opens the door to further exciting possibilities in computational (S)TEM. As the basic computational approach to multislice (S)TEM image simulation is not challenging to implement, it may be beneficial to the microscopy community for the developers of DFT codes to implement image simulation directly within their own software. One could imagine, with the addition of relatively few lines of code, developing DFT software that would output (S)TEM images directly. This might encourage a more widespread use of healthier electronic structure theory in the microscopy community. Furthermore, the calculation of all parameters related to the specimen for input into the microscopy simulation could be done entirely from first principles. This would be a big step forward since, in the field of microscopy simulation, "Specimen parameters such as atom coordinates, Debye-Waller factors and thickness are treated as variables and are deemed accurate when there is close agreement between simulation and experiment," according to [93] (p. 87).

On the subject of Debye-Waller factors, another interesting application would be the inclusion of thermal effects in (S)TEM simulations in a novel way, from first principles. Commonly, Debye-Waller factors or absorptive potentials [93] are used to take thermal effects into account in microscopy simulations. But phonon modes can be calculated by CASTEP explicitly. By using the frozen phonon approach for each mode weighted by the Bose-Einstein factor, it would be possible to incorporate temperature dependence into the simulation for each material from first principles,

enabling the study of thermal diffuse scattering from irregular systems, such as those containing defects or adatoms.

In addition, with the ever-improving and presently impressive resolution of modern electron microscopes, it may be possible to evaluate the accuracy of different functionals or other approximations of DFT by comparing directly with (S)TEM experiments. The new meta-GGA functional SCAN [94] has shown promising results so far in improving the accuracy of DFT calculations. It would be interesting to compare the effects of SCAN-generated projected potentials with LDA- or PBE-generated ones and to evaluate which are closer to experiment for an array of materials.

Armed with the explicit road map to the computation of accurate projected potentials within the PPW DFT framework outlined in this work, it is our hope that developers of these codes may implement this technique for the possible benefit of both the microscopy and the electronic structure communities.

Through the theoretical development, computational implementation, and application to experiment demonstrated in this thesis, it is clear that electronic structure theory is an indispensable tool for electron microscopy. If history is a reliable guide, electronic structure theory software is likely to continue to advance in both computational efficiency and physical accuracy. Similarly, the spatial and energetic resolution of electron microscopes continue to increase with each new generation of machines. It seems likely then that the interface between these two disciplines will continue to increase in significance, with fruitful outcomes for both electronic structure theory and the electron microscopy alike.

Bibliography

[1] N. C. Pyper, T. C. Naginey, P. D. Nellist, and Colm T. Whelan. *Philos. Mag. Lett.*, 97(8), 2017.

[2] T. C. Naginey, R. J. Nicholls, P. D. Nellist, and Jonathan R. Yates. *In preparation.*

[3] T. C. Naginey, Jonathan R. Yates, and P. D. Nellist. *In preparation.*

[4] G. T. Martinez, T. C. Naginey, C. O'Leary, R. J. Nicholls, J. R. Yates, P. D. Nellist, and T. J. Pennycook. *In preparation.*

[5] D. R. Hartree. *Rep. Prog. Phys.*, 11, 1947.

[6] Jorge Kohanoff. *Electronic Structure calculations for solids and molecules.* Cambridge University Press, Cambridge, UK, 2006.

[7] Feliciano Giustino. *Materials Modelling using Density Functional Theory.* Oxford University Press, Great Clarendon Street, Oxford, OX2 6DP, UK, 2014.

[8] J. C. Slater. *Phys. Rev.*, 34(10), 1929.

[9] L. H. Thomas. *Math. Proc. Cambridge Philos. Soc.*, 23(5), 1927.

[10] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136(3B), 1964.

[11] W. Kohn and L. J. Sham. *Phys. Rev.*, 140(4A), 1965.

[12] David C. Langreth and M. J. Mehl. *Phys. Rev. B*, 28(4), 1983.

[13] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. *Phys. Rev. Lett.*, 77(18), 1996.

[14] F. Bloch. *Z. Phys.*, 52, 1928.

[15] E. Fermi. *Ricer. Scient.*, 7, 1936.

[16] D. R. Hamann, M. Schlüter, and C. Chiang. *Phys. Rev. Lett.*, 43, 1979.

[17] D. Vanderbilt. *Phys. Rev. B*, 41, 1990.

[18] Z. W. Lu, Alex Zunger, and Moshe Deutsch. *Phys. Rev. B*, 47(15), 1993.

[19] Kurt Lejaeghere et al. *Science*, 351(6280), 2016.

[20] R. F. Egerton. *Electron Energy Loss Spectroscopy in the Electron Microscope.* Plenum, New York, 1996.

[21] P. E. Batson, N. Delby, and O. L. Krivanek. *Nature (London)*, 418, 2002.

[22] P. D. Nellist, M. F. Chisholm, N. Delby, O. L. Krivanek, M. F. Murfitt, A. S. Szilagyi, A. F. Lupini, A. Borisevich, W. H. Sides Jr., and S. J. Pennycook. *Science*, 305(5691), 2004.

[23] H. Sawada et al. *Jpn. J. Appl. Phys.*, 46, 2007.

[24] Fredrik S. Hage1, Rebecca J. Nicholls, Jonathan R. Yates, Dougal G. McCulloch, Tracy C. Lovejoy, Niklas Dellby, Ondrej L. Krivanek, Keith Refson, and Quentin M. Ramasse. *Sci. Adv.*, 4(6), 2018.

[25] D. Taverna, M. Kociak, O. Stéphan, A. Fabre, E. Finot, B. Décamps, and C. Colliex. *Phys. Rev. Lett.*, 100(035301), 2008.

[26] Konstantin Iakoubovskii, Kazutaka Mitsuishi, and Kazuo Furuya. *Phys. Rev. B*, 78(6), 2008.

[27] W. Jager, R. Manzke, H. Trinkaus, R. Zeller, J. Fink, and G. Crecelius. *Radiat. Effects*, 78(315), 1983.

[28] A. A. Lucas, J. P. Vigneron, S. E. Donnelly, and J. C. Rife. *Phys. Rev. B*, 28(5), 1983.

[29] P. R. Taylor. *Chem. Phys. Lett.*, 121(205), 1985.

[30] C. M. Rohlfing. *J. nucl. Mater.*, 165(84), 1989.

[31] C. A. Walsh, J. Yuan, and L. M. Brown. *Phil. Mag. A*, 80(7), 2000.

[32] M. P. Housden and N. C. Pyper. *Mol. Phys.*, 105(17-18), 2007.

[33] N. C. Pyper, D. W. Essex, and Colm T. Whelan. *Phil. Mag. B*, 81(1), 2001.

[34] N. Jacobi and G. Csanak. *Chem. Phys. Lett.*, 30(367), 1975.

[35] R. G. Gordon and Y. S. Kim. *J. chem. Phys.*, 56(3122), 1972.

[36] H. Trinkaus and B. N. Singh. *J. Nucl. Mater.*, 323, 2003.

[37] S. Fréchard, M. Walls, M. Kociak, J. P. Chevalier, J. Henry, and D. Gorse. *J. Nucl. Mater.*, 393, 2009.

[38] M.-L. David, F. Pailloux, V Mauchamp, and L. Pizzagalli. *Appl. Phys. Lett.*, 98(171903), 2011.

[39] K. Alix, M-L. David, G. Lucas, D. T. L. Alexander, F. Pailloux, and L. Pizzagalli. *Micron*, 77, 2015.

[40] D. A. Arms, T. J. Graber, A. T. Macrander, R. O. Simmons, M. Schwöerer-Bohning, and Y. Zhong. *Phys. Rev. B*, 71(233107), 2005.

[41] M. Knoll and E. Ruska. *Z. Phys. A*, 78, 1932.

[42] Harald H. Rose. *J. Electron Microsc.*, 58(3), 2009.

[43] M. von Ardenne. *Z. Tech. Phys.*, 19, 1938.

[44] A. V. Crewe, J. Wall, and L. M. Welter. *J. Appl. Phys.*, 39, 1968.

[45] Florian F. Krause and Andreas Rosenauer. *Micron*, 92, 2017.

[46] S. J. Pennycook. *A scan through the history of STEM*, pages 1–90.

[47] S.J. Pennycook, Maria Varela, Crispin Hetherington, and Angus Kirkland. 31, 01 2006.

[48] D. K. Saldin and J. C. H. Spence. *Ultramicroscopy*, 55(4), 1994.

[49] M. Vulović, L. M. Voortman, L. J. van Vliet, and B. Rieger. *Ultramicroscopy*, 136, 2014.

[50] J. M. Cowley and A. F. Moodie. *Acta Crystalogr.*, 10, 1957.

[51] Joseph W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, Inc., San Francisco, USA, 1968.

[52] W. Hoppe. *Acta. Crystallagr. A*, 25, 1969.

[53] R. E. Burge, M. A. Fiddy, A. H. Greenaway, and G. Ross. *Proc. R. Soc. Lond.*, 350, 1976.

[54] J. M. Rodenburg and R. H. T. Bates. *Phil. Trans. R. Soc. Lond.* A, 339, 1992.

[55] P. C. Nellist and J. M. Rodenburg. *Acta Crystalogr. Section A*, 54, 1998.

[56] J. M. Rodenburg, B. C. McCallum, and P. D. Nellist. *Ultramicroscopy*, 48, 1993.

[57] T. J. Pennycook, A. R. Lupini, H. Yang, M. F. Murfitt, L. Jones, and P. D. Nellist. *Ultramicroscopy*, 151, 2015.

[58] Hao Yang, Ian MacLaren, Lewys Jones, Gerardo T. Martinez, Martin Simson, Martin Huth, Henning Ryll, Heike Soltau, Ryusuke Sagawa, Yukihito Kondo, Colin Ophus, Peter Ercius, Lei Jin, András Kovács, and Peter D. Nellist. *Ultramicroscopy*, 180, 2017.

[59] I. Lobato and D. Van Dyck. *Acta Cryst.*, A70, 2014.

[60] P. A. Doyle and P. S. Turner. *Acta Crystallogr. A.*, 24, 1968.

[61] Earl J. Kirkland. *Advanced Computing in Electron Microscopy.* Springer, New York, USA, 2010.

[62] G. R. Anstis, D. F. Lynch, A. F. Moodie, and M. A. O'Keefe. *Acta Crystalogr. Section A*, 29, 1973.

[63] F. Phillipp, R. Höschen, M. Osaki, G. Möbus, and M. Rühle. *Ultramicroscopy*, 56, 1994.

[64] G. Möbus, F. Phillipp, T. Gemming, R. Schweinfest, and M. Rühle. *J. Electron Microsc.*, 46, 1997.

[65] P. Blaha, K. Schwarz, and P. Sorantin. *Comput. Phys. Commun.*, 59, 1990.

[66] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. *Rev. Mod. Phys.*, 64, 1992.

[67] G. Kresse and J. Furtm:uller. *Phys. Rev. B*, 54, 1996.

[68] T. Gemming, G. Möbus, M. Exner, F. Ernst, and M. Rühle. *J. Microsc.*, 190, 1998.

[69] S. J. Clark et al. *Zeitschrift fuer Kristallographie*, 220(5-6), 2005.

[70] P. S. Stadelmann. *Ultramicroscopy*, 21(2), 1987.

[71] S. Mogck, B. J. Kooi, J. Th. M. De Hosson, and M. W. Finnis. *Phys. Rev. B*, 70(245427), 2004.

[72] B. Deng and L. D. Marks. *Acta Crystalogr. Section A*, 62, 2006.

[73] Michael L. Odlyzko, Burak Himmetoglu, Matteo Cococcioni, and K. Andre Mkhoyan. *J. Vac. Sci. Technol. A*, 34(041603), 2016.

[74] B. Hammer, L. Hansen, and J. K. Norskov. *Phys. Rev. B*, 59, 1999.

[75] J. M. Zuo, P. Blaha, and K. Schwarz. *J. Phys.: Condens. Matter*, 9, 1997.

[76] P. Blaha, K. Schwarz, G. Madsen, D. Kvasnicka, and J. Luitz. 2001.

[77] Bin Deng, Laurence D. Marks, and James M. Rodinelli. *Ultramicroscopy*, 107, 2006.

[78] J. Ciston, J. S. Kim, S. J. Haigh, A. I. Kirkland, and L. D. Marks. *Ultramicroscopy*, 111, 2011.

[79] L. D. Marks et al. NUMIS multislice imaging software 2.0. http://josephsmith.numis.northwestern.edu/cgi-bin/cvs/cvs.gi/Numis-2.0/.

[80] S. Kurasch, J. C. Meyer, D. Künzel, A. Groß, and U. Kaiser. *Beilstein J. Nanotechnol.*, 2, 2011.

[81] J. C. Meyer, S. Kurasch, H. J. Park, V. Skakalova, D. Künzel, A. Groß, A. Chuvilin, G. Algara-Siller, S. Roth, T. Iwasaki, U. Starke, J. H. Smet, and U. Kaiser. *Nat. Mater.*, 10, 2011.

[82] Michael L. Odlyzko, Burak Himmetoglu, Matteo Cococcioni, and K. Andre Mkhoyan. *J. Vac. Sci. Technol. A*, 34(041602), 2016.

[83] Paolo Giannozzi et al. *J. Phys.: Condens. Matter*, 21(39), 2009.

[84] Jin-Cheng Zheng, Lijun Wu, and James W. Davenport. *J. Appl. Cryst.*, 38, 2005.

[85] Lijun Wu, Yimei Zhu anbd T. Vogt, Haibin Su, and J. W. Davenport. *Phys. Rev. B*, 69, 2004.

[86] J. M. Zuo. *Rep. Prog. Phys.*, 67, 2004.

[87] Y. Zhu and J. Tafto. *Phil. Mag. B*, 75, 1997.

[88] Yimei Zhu, Lijun Wu, and J. Tafto. *Microsc. Microanal.*, 9, 2003.

[89] J. Friis, G. K. H. Madsen, F. K. Larsen, B. Jiang, K. Marthinsen, and R. Holmestad. *J. Chem. Phys.*, 119(21), 2003.

[90] Michael G. Medvedev, Ivan S. Bushmarinov, Jianwei Sun, John P. Perdew, and Konstantin A. Lyssenko. *Science*, 355, 2017.

[91] M. J. Hÿtch and W. M. Stobbs. *Ultramicroscopy*, 53(191), 1994.

[92] Liu-Gu Chen, Jamie Warner, Angus I. Kirkland, Fu-Rong Chen, and Dirk Van Dyck. *Sci Rep.*, 7(10839), 2017.

[93] Budhika G. Mendis. *Electron beam-specimen interactions and simulation methods in microscopy*. John Wiley Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK, 2018.

[94] Jianwei Sun, Adrienn Ruzsinszki, and John P. Perdew. *Phys. Rev. Lett.*, 115(036402), 2015.