

## Density Functional Theory in the Solid State

Philip J. Hasnip<sup>1</sup>, Keith Refson<sup>2</sup>, Matt I.J. Probert<sup>1</sup>, Jonathan R. Yates<sup>3</sup>,  
Stewart J. Clark<sup>4</sup> and Chris J. Pickard<sup>5</sup>

<sup>1</sup>Department of Physics, University of York, York YO10 5DD, UK

<sup>2</sup>Scientific Computing Department, STFC Rutherford Appleton Laboratory,  
Chilton, Didcot OX11 0QX, UK

<sup>3</sup>Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, UK

<sup>4</sup>Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

<sup>5</sup>Department of Physics and Astronomy, University College London,  
London WC1E 6BT, UK

### Abstract

Density functional theory (DFT) has been used in many fields of the physical sciences, but none so successfully as in the solid-state. From its origins in condensed matter physics it has expanded into materials science, high-pressure physics and mineralogy, solid-state chemistry and more, powering entire computational sub-disciplines. Modern DFT simulation codes can calculate a vast range of structural, chemical, optical, spectroscopic, elastic, vibrational and thermodynamic phenomena. The ability to predict structure-property relationships has revolutionised experimental fields such as vibrational and solid-state NMR spectroscopy, where it is the primary method to analyse and interpret experimental spectra. In semiconductor physics great progress has been made in the electronic structure of bulk and defect states despite the severe challenges presented by the description of excited states. Studies are no longer restricted to known crystallographic structures. DFT is increasingly used as an exploratory tool for materials discovery and computational experiments, culminating in *ex nihilo* crystal structure prediction, which addresses the long-standing difficult problem of how to predict crystal structure polymorphs from nothing but a specified chemical composition. We present an overview of the capabilities of solid-state DFT simulations in all of these topics, illustrated with recent examples using the CASTEP computer program.

**Keywords:** Density functional theory, condensed matter theory, computational materials science, electronic structure theory, computational chemistry

### 1. Introduction

It has long been recognised that many materials properties are governed by their electronic structure, yet only relatively recently has it become possible to simulate this behaviour with predictive accuracy. Advances in electronic structure theory, improved software and substantial year-on-year improvement in computing power have combined to make accurate materials simulations not only possible, but almost routine. Such is the speed and accuracy of these *ab initio* simulations that they have become an integral tool in most materials investigations, aiding the interpretation of experimental data and guiding experimental design.

The unparalleled success of density functional theory (DFT) in explaining materials phenomena has led to the concept of “materials by design” - the idea that for any given set of properties we can predict novel materials that outperform any of the known “usual suspects”. Whilst achieving this lofty goal fully may be some time away, DFT calculations can now predict a significant portion of the properties for a given material. Given a modest set of elements, it is already possible to predict the thermodynamically stable crystal structures; not only under ambient conditions, but spanning a vast pressure range to terapascals and beyond.

Given a crystal structure, it is straightforward to compute its mechanical properties, including acoustic wave propagation and elastic constants, the band-structure and density of states. It is also possible to compute a wide range of spectral properties including optical and electron-energy loss spectra (EELS), nuclear magnetic resonance (NMR), infrared (IR), Raman and inelastic neutron scattering (INS). By computing the forces on each atom and the stress tensor, and propagating the atoms and lattice vectors accordingly, it is also possible to simulate the evolution of the material structure [over a few hundred picoseconds](#). This enables the study of material dynamics and kinetics, as well as the calculation of thermodynamic properties and ensemble averages.

## 2. Theoretical Background

[A full quantum-mechanical treatment of a material would require the calculation of the system's many-nuclei, many-electron wavefunction; however the relatively large mass of the nuclei means that for the vast majority of simulations their behaviour is decoupled from the electrons and they may be treated as classical point-like particles \(the Born-Oppenheimer approximation\). In contrast to the nuclei, the low mass of electrons means that a full quantum mechanical treatment is required to understand their behaviour, but the computational complexity of the many-body Schrödinger equation means that a solution is beyond current or foreseeable technologies for almost all materials problems. However, density functional theory allows us to sidestep that computational difficulty by focusing on the electron density, instead of the many-body wavefunction. The underlying principle of DFT is that the total energy of the system is a unique functional of the electron density \[1\], hence it is unnecessary to compute the full many-body wavefunction of the system. However the precise functional dependence of the energy on the density is not known.](#)

Kohn and Sham [2] transformed this DFT problem of computing the ground-state energy and particle density of an  $N$ -electron system to that of solving a set of independent-particle equations. These Kohn-Sham equations consist of  $N$  single-particle (3-dimensional) Schrödinger-like equations with a modified effective potential, and are much easier to solve than the original ( $3N$ -dimensional) many-body problem. The modified potential is itself a functional of the total particle density, and contains a contribution from the quantum-mechanical exchange and correlation of the particles. No expression for this exchange-correlation (XC) potential is known, but as it is relatively small compared to the single-particle kinetic and Coulomb terms the fractional error arising from its approximation is rather small. [Nevertheless, the importance of this XC term cannot be underestimated, as it is central to the success of DFT. The energy differences of chemical reactions, bonding, etc. are small, so finding XC approximations which yield chemical accuracy presents an ongoing challenge \(see section 10\).](#)

The first general approximation for exchange and correlation was the local density approximation (LDA) [2]. In the LDA the XC energy density at position  $\mathbf{r}$  depends only on the particle density at that point,  $n(\mathbf{r})$ . This density-dependence must be identical to that of a *homogeneous* electron gas of the same density  $n^{\text{HEG}}=n(\mathbf{r})$  and since this has been calculated accurately [3,4], it provides a usable approximate density functional. An important property of any density functional is the exchange-correlation hole, the region around any particle in which the probability of finding another identical particle is reduced. This hole arises from the Pauli exclusion principle, and the total reduction of the density of the other particles should equal  $-1$ ; i.e. exactly one particle should be excluded. The LDA satisfies this exclusion principle exactly, meaning that whilst the shape of the LDA exchange-correlation hole is incorrect it has the correct spherical average [5].

Most extensions to the LDA include a dependence on the local density gradient as in the

generalised-gradient approximations (GGAs) including the PW91 [6] and PBE functionals [7]. These extensions preserve the spherical average of the exchange-correlation hole, often based on the effect of weak perturbations on the homogeneous electron gas. GGAs systematically improve the atomisation or cohesive energies of a wide range of molecules and solids [8] and correct the LDA's severe overbinding of hydrogen-bonded solids [9]. However they are not a universal improvement over the LDA as in both cases their accuracy relies on some cancellation of error, which is system dependent. There are several approximations beyond the GGA; some commonly used approximations include (i) meta-GGAs [10] where the laplacian of the density can be included (in practical terms, this is expressed in terms of the laplacian of the wavefunction, ie. the kinetic energy); (ii) hybrid functionals where an empirical fraction (often around 20%-25%) of Hartree-Fock exchange is included to obviate the band gap problem; and (iii) DFT+U where an on-site Hubbard-U potential is included to enhance localisation of electrons, usually applied to the d or f shells, which improve on magnetic properties of materials. The latter two of these methods will be detailed in Section 3.

Once an exchange-correlation potential has been chosen, what remains is the choice of basis for the single-particle wavefunctions. The inherent periodicity of crystals may be exploited by invoking Bloch's theorem to express the wavefunction as a periodic Bloch function multiplied by a complex phase factor, whose wavevector is drawn from the first Brillouin zone of the reciprocal lattice. Integrals over the Brillouin zone are approximated numerically, with Bloch functions sampled on a discrete mesh of wavevectors often referred to as k-points. The smoothly-varying nature of Kohn-Sham states means that for insulators and semiconductors a well-converged sampling density can usually be achieved using a modest number of wavevectors. For metals however, the abrupt change in the *occupancy* of each state with wavevector means that much denser grids are required.

It is convenient to represent the Bloch functions as an expansion in terms of some set of mathematical *basis functions*, where the coefficients of the functions in this *basis set* are the primary values used to build a computational representation. Several different families have been developed for use in periodic solid-state calculations, including plane-waves, augmented plane-waves, muffin-tin orbitals, numerical and mixed basis sets and a variant of the Gaussian basis sets commonly encountered in quantum chemical codes with open boundary conditions. A detailed description of all of these can be found in reference [11]. While many of these have been implemented in the widely-used solid-state codes referenced in section 2.1, the most popular is the plane-wave basis set [11,14]. Each Bloch state is expressed as a Fourier series whose basis states are plane-waves whose wavevector is a reciprocal lattice vector. This has several key advantages over other basis sets, whose functions are centred on atomic positions (referred to as *localised* basis sets): it reflects the periodicity of an ordered material; it allows matrix elements to be computed efficiently and straightforwardly; it is an orthonormal basis set; and perhaps most importantly its size (and therefore accuracy) is controlled by a single parameter, the cut-off energy  $E_{\text{cut}}$  or equivalently wavevector  $G_{\text{cut}}$ . The ground-state energy is variational with respect to  $G_{\text{cut}}$ , which allows the accuracy to be systematically improved simply by increasing  $G_{\text{cut}}$ . The main disadvantages are, first, that the number of basis functions required is large compared to localised basis sets, which increases computational cost, a particularly serious problem for the evaluation of the Fock exchange term as required by hybrid exchange-correlation functionals. The second is the difficulty of representing sharp peaks in the Kohn-Sham states, such as occur in the core regions near nuclei.

In the core regions the electron-nuclear Coulomb attraction is singular, and it is this singularity which gives rise to the high wavevector, sharp peaks in the Kohn-Sham states. The Coulomb interaction dominates the potential in these core regions, and the form of the Kohn-Sham states near a nucleus

is almost completely independent of the chemical environment. This environmental independence means the states in the core region have a negligible contribution to the chemical or electronic properties of a material, and it is not necessary to represent them or the Coulomb potential exactly. This freedom may be exploited to improve the convergence of the computed Kohn-Sham states with  $G_{\text{cut}}$ , without loss of accuracy, by replacing the Coulomb potential in the core regions with a smoother pseudopotential. This pseudopotential is constructed to reproduce the atomic scattering properties and Coulombic form outside the core region, but is weaker and smoother inside. Figure 1 shows the Coulombic potential for a carbon atom along with an example pseudopotential, together with the original s-orbital and the pseudo-s-orbital. Furthermore any core states (those localised entirely within a core region) may be precomputed (the frozen core approximation), avoiding the need to include them in the materials simulation explicitly. Further information on pseudopotential theory may be found in reference [11].

Determining the plane-wave coefficients of the single-particle wavefunctions requires the solution of the Kohn-Sham equations. These Schrödinger-like equations take the form of eigenvalue equations, and could be solved in principle by constructing the Hamiltonian explicitly and diagonalising it or by iterative methods. However the Hamiltonian is a density functional, and the density depends on the wavefunctions. The Hamiltonian's implicit wavefunction-dependence transforms the Kohn-Sham equations into non-linear eigenvalue equations. Instead of solving these non-linear eigenvalue equations directly, the solution may be found efficiently using iterative diagonalisation techniques [11,13,14]. One common method of solution is the self-consistent field (SCF) method, in which the non-linearity of the Kohn-Sham eigenvalue equations is removed by solving them for a given, approximate input Hamiltonian. The resultant Kohn-Sham states yield a particle density which is not, in general, consistent with the input Hamiltonian but this density may be used to compute an improved approximate Hamiltonian, and the states recalculated. This procedure may be followed iteratively until the change to the Hamiltonian is negligible, the computed states and Hamiltonian are (approximately) self-consistent, and the algorithm has converged to the ground state solution.

## 2.1 Simulation Software

The widespread use of DFT for solid-state simulations has only been made possible by the development of robust, high-performance density functional computer programs. For a materials simulation to have predictive power, the simulation software itself must be robust enough to provide accurate results even when the modelled system is under extreme or unusual conditions, and fast enough to perform simulations with a reasonable amount of computer time. With reasonable computer power (~500 compute cores) DFT simulations may be performed routinely on systems of hundreds of atoms, with calculations for thousands of atoms being demanding, but possible. It should be noted that the methods described in this work require computational time that scales as the cube of the simulation size for large systems (though see section 10.3).

The success of DFT in the solid-state has led naturally to a proliferation of computer programs able to perform such simulations. The illustrations in this paper have been drawn from the authors' own research with the CASTEP computer program[15], but there exist many other capable, well-used DFT programs for periodic solid-state simulations, including: ABINIT [16,17], CP2K [18], CPMD [19], CRYSTAL [20], DMOL [21], FHI-AIMS [22], GPAW [23], Quantum Espresso [24], SIESTA [25], VASP [26] and WIEN2K [27].

## 3. Band-gaps, defect levels and quasiparticles

Whilst the LDA and GGAs have been applied successfully to a wide range of materials for structural and chemical properties, certain electronic properties, in particular the band-gap (or equivalently the HOMO-LUMO gap) are not predicted reliably. Computing the band-gap from the Kohn-Sham valence- and conduction-band eigenvalues underestimates band-gaps for most semiconductors, insulators, and strongly correlated systems, notoriously predicting a metallic ground state for many transition-metal oxides.

The electronic band-gap of a material may be defined as the difference between the electron affinity (the energy of adding an electron to the system) and the first ionisation energy (the energy of removing an electron from the system); each of these may be calculated using conventional DFT, leading to the so-called delta-SCF method for band-gap calculation. However this is only directly applicable to finite systems; for extended solids it would be necessary to calculate the effect of addition or removal of a single electron from the infinite total. It might be thought that an equivalent approach would be to compute the energy of removal or addition of an infinitesimal fraction of an electron from a single periodic cell; but this yields exactly the same gap as the difference between Kohn-Sham eigenvalues for LDA or GGA approximations. One of the deep results of DFT is that when extended to fractional occupation numbers, the exchange-correlation potential as a function of the electron count is *discontinuous* at the Fermi level [28]. This fundamental discontinuity in the exact  $V_{xc}$  is precisely the difference between the Kohn-Sham and true band gaps, but is not reproduced in LDA or GGA.

A major contribution to the band-gap error arises from the electrostatic electron-electron contribution to the Hamiltonian. This energy is usually computed as the Hartree energy  $E_H$ , the classical interaction of the electronic valence charge density  $n$  at a point  $\mathbf{r}$  with that at  $\mathbf{r}'$ , integrated over all  $\mathbf{r}$  and  $\mathbf{r}'$ .

$$E_H = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}$$

Whilst this includes the correct Coulomb repulsion, by using the total density it also includes a Coulomb repulsion between an electron and its own charge density. **This spurious *self-interaction* is exactly cancelled by the exchange term in some non-DFT methods such as Hartree-Fock theory (discussed below), but it is only partially cancelled by LDA (or GGA) exchange. Residual self-interaction is one of the most significant causes of the underestimation of the band-gap in LDA (or GGA) based DFT calculations.**

In many systems the self-interaction error has only a small effect on properties other than the band-gap, but it can be a significant problem when simulating **systems with localised electrons**. Because the self-interaction energy is always positive, it raises the energy of localised states and favours delocalisation instead. This erroneous delocalisation is common in simulations with d- and f-block elements, and can cause spurious metallisation or incorrect magnetic ground states. Many methods have been proposed to mitigate against the self-interaction problem. The self-interaction can be corrected directly by ensuring each particle only interacts with the density of the *other* particles, usually using the so-called SIC method proposed by Perdew and Zunger [3]. Unfortunately the computational expense and technical difficulties inherent in this form of self-interaction correction mean that it is used only rarely.

A second approach to removing self-interaction error is to use a method beyond DFT, for example

the GW method [29]. The name “GW” comes from the mathematical formulation of the problem where “G” is the Greens function  $G(\mathbf{r},\mathbf{r}',E)$ ,  $W(\mathbf{r},\mathbf{r}',E)$  is the screened Coulomb interaction, and  $E$  is energy; integrating over the product  $GW$  gives the XC interaction. The square of  $G$  can be thought of as the probability of finding a pair of electrons at  $\mathbf{r}$  and  $\mathbf{r}'$  with energy  $E$  whilst  $W$  is often given by an approximation of the inverse dielectric function. In this method the quasiparticle energies of the material are obtained from a perturbative expansion of the electron self-energies and the dielectric function. Not only is the GW method free from self-interactions, it also captures the true non-analytic behaviour of  $V_{xc}$  upon addition (or removal) of an electron. GW gives generally reliable values for the band gaps of semiconductors and insulators, but suffers from major disadvantages. In principle the GW equations ought to be solved self-consistently, but the calculations are so computationally demanding that it is often only practical to perform a single iteration, which is known as the  $G_0W_0$  method. Even this single iteration requires considerable computer power (typically two to three orders of magnitude more than a DFT calculation), effectively restricting its application to modest simulation sizes of a few dozen atoms, and as  $G_0W_0$  only uses a single iteration the result is dependent on the starting configuration. Furthermore neither forces nor stresses can be computed straightforwardly within the standard GW approach, so it is not possible in practice to obtain the atomic structure or crystal lattice of a material.

A third approach to correcting for self-interaction is the LDA+ $U$  or DFT+ $U$  method [30], which introduces a repulsion between the localised electrons on a given atom. This repulsion can cause a breaking of symmetry and lead to the opening of an insulating gap. The repulsive term is akin to that in a Hubbard model and is usually referred to as the Hubbard  $U$ . Computationally, it is a low-cost method for open shell systems that can also be used self-consistently for structural relaxations. The inclusion of the Hubbard  $U$  term does not fix the self-interaction problem itself, but it can correct for the resultant under-localisation of particles and incorrect magnetic structure.

DFT+ $U$  is a popular method for treating the effects of self-interaction, and it has been applied to a wide range of materials. One recent application is in the Heusler alloy  $\text{CoFe}_{0.5}\text{Mn}_{0.5}\text{Si}$  [31], which is a half-metal, meaning that the states at the Fermi level are completely spin-polarised. Half-metals are of great interest for spintronics applications because an applied voltage causes a spin current to flow. Furthermore half-metal-based heterostructures can be used to create spin valves, devices which can switch to allow or forbid the transport of a single spin component. The underestimation of the band-gap with LDA or GGAs causes a complete closing of the computed minority-spin band-gap in  $\text{CoFe}_{0.5}\text{Mn}_{0.5}\text{Si}$ , predicting the material to be an ordinary metal and consequently of no use in spintronics applications. However the inclusion of a small Hubbard  $U$  term (typically around 2 eV) on the Fe and Mn d-states recovers the localisation, and opens up the minority-spin band-gap whilst preserving, correctly, the metallic character of the majority spins (see Figure 2).

Many practitioners regard the strength of the Hubbard  $U$  term as a parameter of the simulation, and DFT+ $U$  has often been used incorrectly as a method for widening the band-gap for general semiconductors. Empirically fixing the band-gap for closed shell systems usually requires an unphysically large Hubbard  $U$ , causing over-localisation and a flattening of the valence and conduction bands. In this sense it might be thought that DFT+ $U$  calculations are semi-empirical, rather than *ab initio* in nature; however Cococcioni and de Gironcoli [32] showed that the correct value of the Hubbard  $U$  terms could in fact be computed using density functional perturbation theory. This technique removes the empiricism, and ensures that the values used are physically reasonable, though it is not straightforward to apply to materials with many different elements.

The final approach we shall discuss is Hartree-Fock and related methods. In Hartree-Fock (HF) the



nonlocal Fock (exact) exchange energy is not a density functional, but depends on the single-particle states. This exchange exactly cancels the spurious self-interaction, but takes no account of electronic correlation. The exchange itself is very long ranged, decaying only as  $1/r$ , and is not screened leading to unrealistically high excitation energies and a large over-estimation of the band-gap. A hybrid XC functional is a combination of HF and LDA (or GGA) functionals, motivated by the observation that HF over-estimates the fundamental gap and LDA or GGA functionals underestimate it. These hybrid functionals yields a result intermediate between HF and DFT and consequently a more realistic gap. More recent functionals such as the HSE class [33] contain a partitioning between long- and short-ranged contributions, introducing screening into the HF term itself. All hybrid functionals introduce at least one empirical parameter, the relative proportions of HF and LDA etc. which is usually fitted to the properties of a molecular test set. Whilst these mixing parameters may give reasonable results within the range of physical parameters to which they are fitted (often including band gaps), the corresponding XC functionals are no longer fully *ab initio* and have diminishing predictive power outside this range. Furthermore the optimum value of the mixing parameter depends strongly on which property is being computed [34].

Whilst the key shortcomings of HF, namely the lack of screening or any correlation, can be addressed with an *ad hoc* mixture of HF and LDA (or GGA), there is an alternative approach. In screened exchange (SX) the screening is included by introducing an exponential damping to the Fock exchange, with a characteristic screening length. The screening length could be considered a free parameter, but as it has a physical meaning its value is usually approximated by a simple model (e.g. the Thomas-Fermi screening length). The lack of correlation in HF may be addressed by adding in the correlation contribution from an XC functional such as LDA, leading to the SX-LDA approximation [35]. **It should be noted that the computational cost of such HF-like functionals for plane waves is significantly larger than for semi-local functionals (see section 2). Whilst the computational time for large simulations scales approximately the same as an ordinary (semi-)local DFT calculation, the absolute time required is an order of magnitude greater.**

Many results of calculations using hybrid functionals have been published recently for solid state systems. These have been used primarily in areas where excitation energies are important, for example, in semiconductor science. In the following section we briefly outline a few examples.

#### **4. Semiconductor physics/device technologies**

DFT simulations are increasingly important in the field of device technologies, and used to study everything from conventional CMOS devices to future devices based on half-metals, thermoelectrics or multiferroics. Obtaining the correct band-gap for the constituent (doped/defect) materials is essential to describe the functionality of these devices. Currently, within the DFT formalism, **(sometimes empirical)** hybrid functionals have become the normal way of obtaining this. There are various pros and cons involved in using the hybrid functionals, however here we will give some successful examples and give more details on hybrid functionals, below.

Increasing device performance usually requires decreasing the size of the components, and for transistors the size has now reached the quantum limit of the constituent materials, in that the standard  $\text{SiO}_2$  gate dielectric is so small that the electrons can tunnel. To limit this excessive leakage current higher dielectric constant materials such as  $\text{HfO}_2$  are now being used, but these high-k materials have a higher defect concentration leading to large electron/hole trapping, lower mobility, changes in threshold voltages and other sources of instability. Therefore it is crucial to understand the electronic structure of the defects and the processing conditions required to minimise their concentration. This requires accurate, predictive electronic structure calculations, including excited

states and semiconductor defect energy levels, for which screened exchange is well-suited [36].

In the case of  $\text{HfO}_2$ , screened-exchange calculations showed the oxygen vacancy gives rise to defect levels near the conduction band edge of silicon (see Figure 3) so it was determined that this defect was the main charge trap in  $\text{HfO}_2$ :Si interfaces. For  $\text{HfO}_2$  to be used successfully in new devices, such defects must be passivated using an oxygen rich processing technique.

The accuracy of screened exchange is not limited to high-k dielectrics like  $\text{HfO}_2$ , indeed SX gives reasonable band gap energies across a range of semiconductors and insulators (see Figure 4). This is particularly important for determining defect levels, which are crucial to describing electrical transmission in device materials. The LDA or GGA band-gaps are too small, and defect levels are incorrectly situated in the gap; this is disastrous when attempting to model defect semiconductors for device technologies. Furthermore, this incorrect electronic structure can lead to incorrect geometries of defects. Hybrid XC functionals can correct such errors, for example in the charged defect states of ZnO (see Figure 5). Whilst hybrid functionals are generally improvements over LDA and GGAs, there are significant differences between the different variants with SX and HSE out-performing other hybrids (particularly B3LYP) for some semiconductor defect states and structures [37].

## 5. Computational Experiments

With the power to compute the total energy of a particular atomic configuration comes the ability to experiment computationally with the system. This experimentation allows the exploration not only of known structures and mechanisms, but also of possible alternatives. For example it was long thought impossible to grow thick films of polar materials such as magnesium oxide ( $\text{MgO}$ ) in their polar direction; each atomic bilayer adds another dipole, raising the electrostatic energy which becomes arbitrarily large as the film grows. When such growth is attempted experimentally in ultra-high vacuum, the films are of poor quality with surface reconstructions, dislocations, islands and other 3D growth phenomena. In contrast, Lazarov *et al.* [38] used DFT simulations to show that perfect film growth is possible in the presence of hydrogen. The hydrogen passivates the oxygen-terminated  $\text{MgO}$  (111) surface, leading to a hydroxylated surface, but when another Mg layer is added the hydroxyl group splits and the hydrogen moves up to the new surface. Furthermore the hydrogen does not remain over the oxygen atom, but moves across to the next site in the crystallographic stacking, exactly where it is needed to terminate the next oxygen layer (see Figure 6). Thus a perfect bilayer of  $\text{MgO}$  can be added to a seed crystal without any problems with the electrostatic energies. This finding has been confirmed experimentally by the growth of thick, ordered  $\text{MgO}$  films whose X-ray photoelectron spectra (XPS) show a significant OH-peak throughout the growth process.

When modelling the polar oxide growth, the structure of the existing  $\text{MgO}$  seed layers was well known and the study focussed on the growth mechanism itself. However such computational experiments need not be restricted to variations of known structures and growth mechanisms, and this becomes crucial when studying new materials or even well-known materials under unusual conditions. The challenge of *predicting* the atomic structure from first principles is the topic of the next section.

## 6. Atomic structure

Understanding how atoms are arranged in a material, and the implications for its properties of that arrangement, is central to modern science and technology. The discovery and development of new - stronger, cheaper, lighter, more functional - materials is important to maintain the steady technological progress to which we have become accustomed. It is increasingly recognised that theory and computation plays a key role in materials discovery.



The challenge of computational materials discovery is fundamentally combinatorial in nature - there is no hope of the complete enumeration of the entire chemical, compositional and configurational materials space. Despite this, the computational exploration of materials space offers advantages over a purely experimental attack.

Finding the most stable (lowest in energy or free energy) structure of a large assembly of atoms is a very difficult problem. Whilst many well-established methods exist for optimising a given structure to find the local energy (or enthalpy, at pressure) minimum, predicting the most stable structure of all possible structures requires a global optimisation method. Furthermore, only fully quantum mechanical calculations are sufficiently reliable to deliver the level of accuracy required because of the wide range of inter-atomic bonding that may be encountered during the searches. First-principles DFT methods offer a high-level description of the electronic structure at a cost that is affordable for the many thousands of structures that must be considered in the course of a reliable search.

Many global optimisation strategies for performing first principles crystal structure prediction have been advocated, ranging from simulated annealing [39] and genetic/evolutionary algorithms [40-43] to minima hopping [44] and particle swarms [45]. A particularly straightforward, and powerful, approach is *Ab initio* random structure searching (AIRSS) [46,47]. In the AIRSS approach random structures are generated - the type of atoms, and the number of them, may be specified, or themselves selected at random. These atoms are then placed randomly in a unit cell (if a crystal is desired) of random shape, but with volume chosen to give approximately the correct starting density. A large ensemble of such random initial structures is generated, and each one is structurally optimised (using first-principles DFT) to minimise the atomic forces and cell stresses. Each initially random structure is therefore moved to a nearby local energy (or enthalpy) minimum. The calculations may be performed one after each other, but this problem is very well suited to high-throughput computation - where each random structure is independently optimised on a subset of the total compute cores. The greatest throughput can be obtained if a single core is used for each structure, as there is no parallel communication cost. However, this communication cost is small between compute cores within a single compute node or workstation, so parallelising the optimisation of each structure over a multicore compute node is reasonable and returns initial results speedily. Because each initial random structure is uncorrelated with all the others, if the few most stable (lowest energy/enthalpy) structures are found repeatedly in the searches a degree of confidence that they represent true global minima is obtained. The number of times a given structure is found is proportional to the volume its basin of attraction occupies in the configurational hyperspace. The initial structures need not be entirely random - rejecting candidates where the atoms are too close to each other (or far apart) improves the stability of the electronic energy minimisation and encourages the formation of connected structures rather than isolated fragments, and can space the different species out appropriately. These “random sensible structures” may be constructed so as to be clusters, defects, surfaces or built from molecular units. The use of symmetry allows more complex materials to be explored, and experimental constraints can be applied in combined experimental/computational structure determinations. As shown in Figure 7, if AIRSS is applied to a range of compositions the thermodynamically stable ones can be predicted through the convex-hull plot, or Maxwell construction.

AIRSS has been applied extensively to the study of matter at high pressure, where the ability to predict structure in an unbiased manner is particularly valuable given the challenges that experiments face, and the deficiency of standard chemical intuition. Early predictions of the structure of silane ( $\text{SiH}_4$ ) [46] and alane ( $\text{AlH}_3$ ) [51] were rapidly confirmed by experiment. The mixed phases of

dense hydrogen (computationally) discovered as part of a study of the phase III of hydrogen [52] have also been recognised recently [53-55] to be an excellent structural model of the room temperature phase IV of hydrogen [56]. Further predictions remain to be confirmed, or otherwise, such as an ionic form of ammonia ( $\text{NH}_3$ ) at megabar pressures [57]. There is no reason to expect that DFT is inaccurate at extremely high pressures, and so first principles structural searches can be used to explore the nature of matter at densities well beyond those routinely studied experimentally, but relevant to planetary science and within range of the latest generation of shock experiments. Through the application of AIRSS, unexpected phase transitions have been predicted in iron [58], carbon [59] and oxygen [60] at pressures up to 1 PPa. Aluminium was predicted to undergo a transition to a complex, incommensurate, host guest phase at around 3 TPa [61]. This simultaneously led to a significant revision of the equation of state of aluminium (which is an important reference material for shock experiments) and revealed an unexpected richness of structure at extreme densities, stimulating considerable experimental activity. The complex structure of aluminium was rationalised in terms of a transition to an electride phase which, as described in Figure 8, led to the successful search for new kinds of elemental magnetism [61,62].

## 7. Phonons and Vibrational Spectroscopy

The study of the dynamical properties of atoms in the solid state is founded on the theory of lattice dynamics. In contrast with the liquid state, atomic motion is primarily vibrational in character and well described as collective oscillations about the equilibrium crystal structure. The theory of lattice dynamics in various refinements has been amply demonstrated to quantitatively describe the thermal properties of solids including specific heat, elasticity, thermal expansion and other thermoelastic properties. It extends to the description of solid-solid phase transitions, for the thermal entropy term in calculations of phase equilibrium at first-order transitions, and it forms the core of the theory of soft mode behaviour in second-order, displacive phase transitions. The lattice dynamical description captures much of the essential thermal physics across a very broad range of crystalline solid-state materials.

The excitations which appear in the quantum theory of lattice vibrations - phonons - typically have an energy scale of 1-500 meV (8-4000  $\text{cm}^{-1}$ ), and a variety of low-energy spectroscopy techniques have been developed and refined for their study. This includes not only the optical probes of infrared and Raman spectroscopies, but also inelastic neutron and more recently inelastic X-Ray techniques which couple to short-wavelength phonons and are thereby able to reveal the dispersion within the Brillouin zone characteristic of crystalline solids.

Lattice Dynamics theory is an ideal match to DFT methods developed for the solid state. Both formulate their respective Hamiltonians using Born von Karman periodic boundary conditions based on the crystal unit cell. In its quantum formulation, lattice dynamics uses the Born-Oppenheimer approximation as does DFT, and the quantum nuclear problem elegantly maps onto a classical description of nuclear vibration (at least within the harmonic approximation). Most significantly, a system-specific set of force constants can be determined using DFT calculations of the response of that system to a perturbation from its equilibrium structure. Solid-state DFT calculations can thereby be used to parametrise the lattice dynamics description of a specific crystalline material by calculating the matrix of force constants from an LDA or other DFT Hamiltonian.

*Ab initio* lattice dynamics (AILD) was conceived and implemented early in the history of solid-state DFT [63,64] and has subsequently developed into a powerful, general and mature computational technique. Since the 1990s, AILD has been used in many thousands of studies across the solid state sciences. It is not the purpose of this article to review either the technical details of the methods or

the wide scientific literature of applications. For that we refer the reader to the book of Richard Martin [11] and the review article by Stefano Baroni *et al.* [65]

More recently, AILD has become an indispensable adjunct of experimental studies using inelastic neutron scattering (INS), inelastic x-ray scattering (IXS), infra-red (IR) spectroscopy and Raman spectroscopy to probe vibrational spectroscopy. Most spectroscopy experiments yield a cross section as a function of frequency, such as its absorptivity, with peaks corresponding to the fundamentals and overtones (one and higher-order phonon processes). The magnitude of this cross section is a property of the response of the electrons in the system to the incoming radiation (except for neutrons where the interaction is entirely nuclear) and can also be calculated using DFT-based methods[65]. But experiments with powder samples (by far the most common case) yield no information on the atomic displacements associated with a peak, making identification and assignment of the modes a difficult and error-prone process. Even with single crystal samples it is only in rare cases that any information on the mode eigenvectors is obtained. By matching a calculated spectrum against the experimental one, the atomic displacements and symmetry associated with each peak can be identified. While non-periodic DFT is also used with gas- and liquid- state spectroscopy, it has become particularly indispensable for solid-state studies due to the greater difficulty of mode assignment. The variety of chemical bonding and co-ordination environments found in many crystal structures, and the more comprehensive sampling of the periodic table in systems of interest to solid-state physicists, mean that simple models for assignment based on known structural motifs are simply not available or are unreliable. Indeed there are many instances in the literature of incorrect assignment which has only been corrected with the aid of an AILD calculation [66].

Buckminsterfullerene -  $C_{60}$  - was the first of the "new" polymorphs of carbon to be discovered, before nanotubes and graphene and its rich set of physical and chemical properties have inspired a multitude of experimental and theoretical investigations. However a definitive analysis of the vibrational spectrum of either the molecule or the solid was lacking until very recently, with uncertainties remaining over the mode assignment even for the molecule. One of the main reasons for the incomplete success of the previous studies was the attempt to compare gas-phase DFT calculations with experiments on solid-state samples. Using the DFPT implementation of AILD in the CASTEP code [67] [a series of calculations was performed](#) in close collaboration with an experimental programme of infrared, Raman and inelastic neutron scattering (INS) experiments resulting in a definitive account of the entire vibrational spectrum of the low-temperature Pa3 structure and almost all of the observed spectrum of the high-temperature phase [68]. The most powerful of these spectroscopic experiments was inelastic neutron scattering, because there are no silent modes due to selection rules unlike infrared and Raman scattering.

The low-temperature ordered Pa3 phase ( $T < 260\text{K}$ ) contains 4  $C_{60}$  units - 240 atoms - in the unit cell which means that a complete AILD calculation was only feasible by using massive parallelism in order to exploit the large amount of CPU power available on national high-performance computing facilities. A librational mode (ie a molecular rotation) of frequency  $31.7\text{ cm}^{-1}$  at a wavevector of  $(1/3, 1/3, 1/3)$  is illustrated in Figure 9. The content of the periodic solid-state calculation is richer than the gas-phase calculation not only by the inclusion of dispersion and external molecular modes, but also the shifts and factor-group splitting of the molecular-like modes by the crystalline environment. These advantages resulted in a prediction of the INS spectrum in remarkable agreement with experiment (Figure 10) including a quantitative match of the peak intensities as well as positions. Consequently it was possible to definitively assign the entire spectrum, resolving all ambiguities in previous studies.

Above 260 K buckminsterfullerene undergoes a phase transition to a Fm3m structure containing just one C<sub>60</sub> per unit cell. The higher symmetry arises from complete rotational disorder of the molecules, which is dynamic in nature. While this disorder can not be modelled directly in the periodic AILD framework, it proved an excellent approximation to impose an artificial orientational order and to model the system in a Fm-3 ordered structure containing one C<sub>60</sub> unit. Consequently it was possible to compute the Raman spectrum in the internal molecular energy range including the electronic contribution to the scattering cross section. The comparison with the experiment is shown in Figure 11, which demonstrates excellent agreement of frequencies and rather good agreement of peak intensities. This calculation correctly predicted a triply degenerate librational mode of imaginary frequency, indicating a structural instability which is clearly the origin of the nearly-free molecular rotation observed above 260K. The limitations of the artificial Fm-3 model were more evident in the comparison with the infrared absorption spectrum where additional, small, peaks not predicted by the calculation were present. It is likely that their origin lies in the orientational disorder, which breaks the periodic lattice symmetry, lifting the selection-rule absence of what are zone-boundary modes in the Fm-3 model.

## 8. Molecular Dynamics

While many of the thermodynamic and thermoelastic properties of crystalline solids are very well described by *ab initio* lattice dynamics, this approach is limited to ordered systems with predominantly vibrational motion. In particular glasses and high-temperature behaviour involving transport such as fast-ion conductors and orientationally disordered molecular crystals lie beyond its scope. Nor is the treatment of anharmonic effects straightforward where *quantum* nuclear dynamics is manifest. In these circumstances the natural route is to exploit the power of molecular dynamics (MD) and the ergodic theorem.

Within the CASTEP computer program, there are two fundamentally different approaches to MD: Born-Oppenheimer MD using classical equation of motion and Path-Integral MD using the Feynman path-integral approach to quantum mechanics. A full explanation of these and other related ideas, such as Car-Parrinello MD, can be found in the textbook by Marx and Hutter [69].

In the Born-Oppenheimer approach the electronic energy is minimised at each new MD configuration, so that the system evolves along the Born-Oppenheimer potential energy surface with *ab initio* forces derived within DFT. It is an advantage of this approach over Car-Parrinello MD, that the same approach works equally well for metals and insulators.

The Feynman path-integral approach (PIMD) is used to include a quantum treatment of the nuclei in addition to the DFT treatment of the electrons. This approach is based upon the isomorphism between finite temperature quantum mechanics and statistical mechanics where each quantum particle is replaced by a classical ring polymer. We can then use PIMD to sample configurations of this system of ring polymers and hence calculate finite temperature quantum expectation values, so that effects such as quantum delocalisation, zero point motion and tunnelling are accurately included. Recent developments such as centroid-PIMD and ring-polymer PIMD go beyond this to do true quantum dynamics.

With either MD approach, we need to choose which statistical mechanical ensemble to work in, as that dictates the equation of motion to use, and the range of physical observables that can be measured over the trajectory. Within CASTEP, the available ensembles are: NVE (the microcanonical ensemble); NVT (the canonical ensemble); NPH (iso-enthalpic calculation - not a

true ensemble but useful for doing 2-phase coexistence studies); and NPT (isothermal-isobaric ensemble).

In general, PIMD requires considerably more computer power than conventional *ab initio* MD, and so its major uses so far have been on systems for which quantum properties of the ions (not just the electrons) are expected to be important, e.g. those containing hydrogen. Recent studies have looked at the wetting of metal surfaces [70] and the phase diagram of cold solid hydrogen [71].

In the case of cold solid hydrogen, earlier work using the AIRSS approach with CASTEP [52] had suggested some interesting new candidate structures for the high-pressure phases. Obviously, the quantum effects of the hydrogen nuclei are expected to be important, so in this early work the basic enthalpy of each structure was calculated for each geometry-optimised structure and then a quasi-harmonic zero-point energy was calculated, and the combined energy was then used to evaluate the stability of the different structures with pressure. In the more recent work of Li *et al.* [71], the PIMD technique was used with the AIRSS candidate structures to include the thermal and quantum effects rigorously. In addition to calculating the structure and energies, GW theory was used to calculate the optical response, in order to determine at which pressure the different candidates went transparent (300 GPa according to experiment). The variation of IR and phonon modes with pressure was also calculated and compared to experiment. The conclusion was that the  $P2_1/c-24$  structure, which is a rotationally restricted phase based upon a large hcp-type unit cell (with 24 atoms in the primitive cell) is the best candidate so far for Phase II, and a layered structure with C2/c symmetry is most plausible for Phase III. Figure 12 shows the representative centroid configurations of the  $P2_1/c-24$  structure at 80 GPa (phase II) using both conventional *ab initio* MD and also PIMD. The comparison of hydrogen and deuterium illustrates the relative importance of the quantum and thermal effects. Deuterium at 50 K (either with or without PIMD) has restricted rotational motion, whereas hydrogen at 50 K (with PIMD) and deuterium at 150 K (with PIMD) rotate freely.

## 9. Other spectroscopies

Solid-state Nuclear-Magnetic Resonance (NMR) is a powerful probe of structure and dynamics on the atomic scale. The past decade has seen significant advances in this technique both in terms of hardware and methodology, for example the availability of high-field magnets with field strengths up to 23.5 T, rotors for performing magic-angle spinning up to 70 kHz, and sophisticated radio-frequency pulse sequences to extract correlated spectra. While NMR is a spectroscopy of the nuclear spins, the observed spectrum is influenced by a number of interactions mediated by the electronic structure of the material, i.e. the chemical shift, indirect spin-spin (or J) coupling and, for quadrupolar nuclei (eg  $^2\text{H}$ ,  $^{17}\text{O}$ ), the electric field gradient. All of these interactions can be computed for solid materials using DFT and such calculations are now widely used in the solid-state NMR community for the assignment and interpretation of experimental spectra [72]. Calculations play a further role in guiding the design and optimisation of new experiments.

NMR often provides complementary information to diffraction based studies; NMR is a probe of short-range order whilst diffraction relies on longer range order. NMR is hence of particular importance when a material is amorphous, or exhibits some degree of disorder. NMR can also be a sensitive probe of dynamics in a material. For these reasons there is increasing use of the combination of both solid-state NMR and diffraction experiments together with first-principles calculations for materials characterisation.

An interesting illustration is given by the simple sugar galactose [73]. Three crystal structures can be



found in the literature, each one has the same crystal symmetry but a different arrangement of atoms. After an initial optimisation of the atomic positions in which the total energy and forces on the atoms are minimised, it was found that two of the structures had relaxed to the same final structure and were thus equivalent. This equivalence leaves only two distinct structures each with a different hydrogen-bonding network. By comparing the computed proton NMR spectra with that measured experimentally it is clear that only one of the two possible structure is consistent with experiment. Hence it was only with combination of diffraction, NMR spectroscopy and first-principles simulations that the correct crystal structure could be identified (Figure 13). One area of application is in the area of pharmaceuticals in which there are commercial and regulatory requirements to identify all possible polymorphs, salts, and hydrated forms of a given drug molecule. Often such compounds do not form sufficiently large crystals for single-crystal diffraction to provide a complete characterisation of the crystal structure and the combined approach is vital.

Simulations have an important role to play in the interpretation of NMR experiments on materials that are not fully ordered. A recent example is the study of layered double hydroxides (LDH), a family of anionic clays of considerable interest as host structures for drug delivery systems, nanocomposite materials, or for catalysis [74]. The physicochemical properties of a particular LDH are believed to be determined by the nature of its cationic ordering. Cadars *et al.* [74] studied the nature of the Al/Mg ordering in LDHs. In these materials X-ray diffraction provides no information on the local ordering of the cations. However, using high-field magnetics and very fast magic angle spinning it was possible to record  $^1\text{H}$  and  $^{27}\text{Al}$  NMR spectra with clear spectral features. First-principles calculations were then used to assign these features and hence provide a quantitative description of the various defect moieties present in the samples (Figure 14).

Turning briefly to other spectroscopic techniques, first-principles calculations have been applied to the study of electron paramagnetic resonance (EPR) for example phosphorus defects at the Si/SiO<sub>2</sub> interface [75] and hydrogen vacancies on a silicon surface.[76]

First-principles calculations have also been extensively applied to interpret electron energy loss (EELS) measurements - many (scanning) transmission electron microscopes are fitted with an EELS spectrometer. EELS provides a method of probing the bonding within a material. Depending on the energy range examined the processes involved include excitations of core electrons into the conduction states, and plasma excitations of the valence electrons. DFT has been shown to be sufficient to interpret EELS spectra of a wide range of materials. It is a particularly important technique for the study of nanomaterials as the latest generation of STEMs can obtain high-energy resolution EELS at atomic-level spatial resolution. As an illustration, CASTEP was recently used to predict the change in the carbon K-edge spectrum (an excitation of the 1s electron) due to the presence of a substitutional nitrogen defect in graphene (Figure 15). The calculations predicted a pronounced shoulder to the sigma\* peak for the carbon atom directly bonded to the nitrogen. Carbon atoms greater than three bonds from the defect were shown to have spectra essentially identical to pristine graphene. These subtle changes were subsequently observed [77] in STEM EELS measurements on N-doped graphene, providing direct evidence of the N-C bond.

## 10. Current challenges

The plane-wave method described in this paper provides an excellent balance between accuracy and computational cost, and has been applied successfully to model a wide range of experimental phenomena. Nevertheless there remain several challenges for DFT and its applications in the future, some scientific and others technical in nature. Here we highlight three specific issues: exchange and correlation; dispersion interactions; and the application of DFT to large solid-state systems.

## 10.1 Exchange-Correlation

There is now a large *zoo* of *ab initio* XC functionals to choose from when performing density functional calculations, including GGAs, meta-GGAs, hybrid functionals, Hubbard-U parameters, and many more. Although a DFT calculation with any particular choice of XC functional may be considered *ab initio*, this freedom in the choice of functional naturally leads to many practitioners *deciding* which is best for a particular simulation, and therefore introducing an element of empiricism. Since there is no known universal functional, nor even a framework in which to improve XC approximations systematically, the performance of an XC functional may only be tested by comparison to simple model systems, known experimental results or high quality, computationally intensive “post-DFT” quantum chemistry calculations.

The success of DFT with (semi-)local functionals such as LDA and GGAs is that it gives reasonable accuracy for a large set of material properties for reasonable compute costs (we intentionally leave “reasonable” ambiguously defined). The predictive power of such simulations arises precisely because they are *ab initio*, with no free parameters and therefore no way in which the computational scientist could have influenced the simulations’ outcomes. However these functionals are not universal, and where they fail to describe a material with sufficient accuracy there is little choice but to turn to the aforementioned functional zoo. Since it is not possible for the materials scientist to try every functional to see which suits their particular application best, these DFT simulations require some *a priori* knowledge of the XC functional’s range of validity and approximate accuracy for a specific property - essentially a choice based on knowledge, experience and comparison to experiment; this may yield useful insights into the system under study, but the *ab initio* philosophy is lost and some of the predictive power with it.

The challenges for XC approximations in the future thus come in two categories: (i) can the researchers working in the fundamental properties of DFT (the XC researchers) find a “more universal” functional [78]? Are there other properties of the exact XC functional that can be used to improve present approximations? This search is still ongoing, though with the occasional incremental improvement in the quality of some material property predictions; (ii) can specific physical or chemical mechanisms be identified which are not captured properly in standard XC approximations? The identification of these mechanisms sheds light on the validity of XC functionals, and also paves the way for the shortcomings to be addressed. Indeed it was just such an identification which led to the DFT+U methodology, and as will be seen in the next section there is at least one other mechanism lacking; one whose treatment has become an active research area in recent years.

## 10.2 Dispersion interactions

One area where DFT has traditionally struggled is in the treatment of dispersion interactions (such as van der Waals) as they are both non-local and dynamical in origin. This has become something of a “hot topic” of research in recent years, with a number of different approaches being promoted by different groups, including semi-empirical correction schemes (usually a functional of the form  $D(r)$ , where  $r$  represents spatial coordinate), non-local density functionals (of the form  $D(r,r')$ ), and many-body schemes such as the Random Phase Approximation (RPA) (of the form  $D(r,r,E)$ , where  $E$  is energy). These schemes differ dramatically in computational cost, with the semi-empirical schemes being effectively free at the one extreme (compared to the cost of the overall DFT calculation), and the RPA being very costly at the other; the computational cost rises quickly with the “dimensionality” of the functional. Unfortunately the increase in computational cost is not always rewarded by quantitatively superior results, and it is often found that the semi-empirical schemes are good enough for practical purposes. A recent review article by Björkman *et al.* [79] summarises the

performance of a range of different approaches to dispersion interactions.

### 10.3 Large simulations

The descriptive power of the plane-wave DFT method, coupled with the dramatic increase in available computational resources, has led to DFT being applied to ever-larger simulation systems. However for large simulations ( $\sim 10^3$  atoms), the computer memory required by a conventional DFT simulation scales as the *square* of the simulation size. Furthermore the computational time required scales as the *cube* of the system size. Whilst available computer power continues to increase, it is attained for the most part not with faster compute cores, but with *more* compute cores. To address large physical and chemical problems, or even move into the molecular biological sciences, is likely to require simulations to run on hundreds of thousands, if not millions of cores. Not only does this present an extremely difficult challenge for the parallel scaling of the DFT software, particularly any parallel Fourier transforms, it also presents new problems. As the number of cores involved in a calculation increases, the probability of a core failing increases combinatorially; in fact in a calculation involving a million cores for many hours, it is likely that one or more cores will fail during the course of the simulation. This requires a new consideration in the design of DFT programs: the capability to detect and recover from severe hardware faults such as core, memory or network connection failure. Such fault tolerant computing is already crucial for key computer server software and is an active area of research in Computer Science, but its impact has not yet spread widely in computational science simulations. It is vital that such fault tolerant techniques and algorithms are used in DFT software for the future if it is ever to make full use of exascale computers and beyond.

An alternative approach to the problem of large simulations is that of linear-scaling DFT. The Kohn-Sham equations can be reformulated in terms of the density matrix,  $n(\mathbf{r}, \mathbf{r}')$ , and for systems with a band-gap this density matrix decays exponentially with  $|\mathbf{r} - \mathbf{r}'|$ . This exponential decay with distance means the density matrix is diagonally dominant, and can be truncated safely beyond a specified cut-off radius. By only storing the non-zero elements of the density matrix and careful use of iterative matrix methods, both the computer memory and time required for a large DFT calculation can be made to scale linearly with simulation size.

For extremely large calculations, this improvement of the scaling by two orders reduces the computational time dramatically. However linear-scaling methods are relatively new compared to the established plane-wave methods, and they are inherently more complex. At present conventional, cubically-scaling methods out-perform linear-scaling computations for the vast majority of solid-state simulations; the notable exception is for systems with large vacuum regions (e.g. isolated nanowires), where linear-scaling methods can already surpass the performance of standard DFT simulations for relatively modest simulation sizes.

Linear-scaling DFT is an area of active research in computational physics, with the performance of the simulations improving steadily, especially on parallel high-performance machines. Historically, linear-scaling implementations were restricted to basic ground state energy and density calculations, but this has also improved in recent years with geometry optimisation and molecular dynamics (see Section 8) becoming available. The user-friendliness of the programs has also been improved, with the developers of linear-scaling codes such as CONQUEST [80,81] and ONETEP [82] seeking to increase and broaden their user base.

### Acknowledgements

The authors acknowledge financial support from EPSRC (PJH, MIJP, SJC and CJP), STFC (KR) and The Royal Society (JRY) with gratitude. The authors are grateful for computational support

from the UK national high performance computing service, HECToR, for which access was obtained via the UKCP consortium and funded by EPSRC grant ref. EP/K013564/1.

## References

- [1] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas", *Phys. Rev.* **136** (3B) B864-B871 (1964), doi: 10.1103/PhysRev.136.B864.
- [2] W. Kohn and L.J. Sham, "Self-consistent equations including exchange and correlation effects", *Phys. Rev.* **140** (4A) A1133-A1138 (1965), doi: 10.1103/PhysRev.140.A1133.
- [3] J.P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", *Phys. Rev. B* **23** 5048-5079 (1981), doi: 10.1103/PhysRevB.23.5048.
- [4] D.M. Ceperley and B.J. Alder, "Ground state of the electron gas by a stochastic method", *Phys. Rev. Lett.* **45** 566-569 (1980), doi: 10.1103/PhysRevLett.45.566.
- [5] O. Gunnarsson and B.I. Lundqvist, "Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism", *Phys. Rev. B* **13** 4274-4298 (1976), doi: 10.1103/PhysRevB.13.4274.
- [6] J.P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", *Phys. Rev. B* **45** 13244-13249 (1992), doi:10.1103/PhysRevB.45.13244.
- [7] J.P. Perdew, K. Burke and M. Ernzerhof, "Generalized gradient approximation made simple", *Phys. Rev. Lett.* **77** 3865-3868 (1995), doi: 10.1103/PhysRevLett.77.3865.
- [8] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh and C. Fiolhais, "Atoms, molecules, solids and surfaces - applications of the generalized gradient approximation for exchange and correlation", *Phys. Rev. B* **46** 6671-6687 (1992), doi: 10.1103/PhysRevB.46.6671.
- [9] B. Winkler, V. Milman, B. Hennion, M.C. Payne, M.-H. Lee and J.S. Lin, "Ab Initio Total Energy Study of Brucite, Diaspore and Hypothetical Hydrous Wadsleyite", *Physics and Chemistry of Minerals* **22** (1995), doi:10.1007/BF00200324.
- [10] J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, "Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids", *Phys. Rev. Lett.* **91** 146401 (2003) doi: 10.1103/PhysRevLett.91.146401.
- [11] Richard M. Martin, "*Electronic Structure: Basic theory and practical methods*," Cambridge University Press (2004). ISBN-10: 0521534402; ISBN-13: 978-0521534406.
- [12] R. Ditchfield, W. J. Hehre and J. A. Pople, "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules", *J. Chem. Phys.* **54**, 724 (1971), doi: 10.1063/1.1674902.
- [13] M.J. Gillan, "Calculation of the vacancy formation energy in aluminium", *J. Phys.: Condens. Matter* **1** 689 (1989), doi:10.1088/0953-8984/1/4/005.

- [14] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias and J.D. Joannopoulos, "Iterative minimization techniques for *ab initio* total-energy calculations - molecular-dynamics and conjugate gradients", *Rev. Mod. Phys.* **64** 1045-1097 (1992), doi: 10.1103/RevModPhys.64.1045.
- [15] S.J. Clark, M.D. Segall, C.J. Pickard, P.J. Hasnip, M.J. Probert, K. Refson and M.C. Payne, "First principles methods using CASTEP", *Z. Kristallogr.* **220**(5-6), 567-570 (2005), doi: 10.1524/zkri.220.5.567.65075.
- [16] X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, Ph. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D.R. Hamann, and D.C. Allan, "A brief introduction to the ABINIT software package", *Z. Kristallogr.* **220**(5-6), 558-562 (2005), doi: 10.1524/zkri.220.5.558.65066.
- [17] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Cote, T. Deutsch, L. Genovese, Ph. Ghosez, M. Giantomassi, S. Goedecker, D.R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M.J.T. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M.J. Verstraete, G. Zerah, J.W. Zwanziger, "ABINIT : First-principles approach of materials and nanosystem properties", *Comp. Phys. Comm.* **180**, 2582-2615 (2009), doi: 10.1016/j.cpc.2009.07.007.
- [18] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, "Quickstep: fast and accurate density functional calculations using a mixed Gaussian and plane waves approach", *Comp. Phys. Comm.* **167**, 103 (2005), doi: 10.1016/j.cpc.2004.12.014.
- [19] J. Hutter and M. Iannuzzi, "CPMD: Car-Parrinello molecular dynamics". *Z. Kristallogr.* **220** 549-551 (2005), doi: 10.1524/zkri.220.5.549.65080.
- [20] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V.R. Saunders and C.M. Zicovich-Wilson, "CRYSTAL: a computational tool for the *ab initio* study of the electronic properties of crystals", *Z. Kristallogr.* **220** 571-573 (2005), doi: 10.1524/zkri.220.5.571.65065.
- [21] B. Delley, "From molecules to solids with the DMol approach." *J. Chem. Phys. B* **113** 7756-7764 (2000), doi 10.1063/1.1316015.
- [22] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X.G. Ren, K. Reuter and M. Scheffler, "Ab initio molecular simulations with numeric atom-centered orbitals". *Comp. Phys. Comm.* **180** 2175-2196 (2009), doi: 10.1016/j.cpc.2009.06.022.
- [23] J. Enkovaara, C. Rostgaard, J.J. Mortensen, J. Chen, M. Dulak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H.A. Hansen, H.H. Kristoffersen, M. Kuisma, A.H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P.G. Moses, J. Ojanen, T. Olsen, V. Petzold, N.A. Romero, J. Stausholm-Moller, M. Strange, G.A. Tritsarlis, M. Vanin, M. Walter, B. Hammer, H. Hakkinen, G.K.H. Madsen, R.M. Nieminen, J.K. Nørskov, M. Puska, T.T. Rantala, J. Schiotz, K.S. Thygesen and K.W. Jacobsen, "Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method", *J. Phys.: Condens. Matter* **22**, 253202 (2010), doi: 10.1088/0953-8984/22/25/253202.



- [24] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari and R.M. Wentzcovitch, “QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials.” *J. Phys.: Condens. Matter* **21** 395502 (2009), doi: 10.1088/0953-8984/21/39/395502.
- [25] J.M. Soler, E. Artacho, J.D. Gale, A. Garcia, J. Junquera, P. Ordejon and D Sanchez-Portal, “The SIESTA method for ab initio order-N materials simulation”, *J. Phys.: Condens. Matter* **14** 2745–2780 (2002), doi: 10.1088/0953-8984/14/11/302.
- [26] G. Kresse and J. Furthmüller, “Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set”, *Phys. Rev. B* **54**(16) 11169–11186 (1996), doi: 10.1103/PhysRevB.54.11169.
- [27] K. Schwarz and P. Blaha, “Solid state calculations using WIEN2k”, *Computational Materials Science*, **28** 259–273 (2003), doi: 10.1016/S0927-0256(03)00112-5.
- [28] J. P. Perdew, R. G. Parr, M. Levy and J.L. Balduz, “Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy”, *Phys. Rev. Lett.* **49**(23), 1691–1694 (1982), doi:10.1103/PhysRevLett.49.1691.
- [29] L. Hedin, “New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem”, *Phys. Rev.* **139** (3A) A796–A823 (1965), doi: 10.1103/PhysRev.139.A796.
- [30] A.I. Lichtenstein, V.I. Anisimov and J. Zaanen, “Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators”, *Phys. Rev. B* **52**, 5467 (1995), doi: 10.1103/PhysRevB.52.R5467.
- [31] P.J. Hasnip, J.H. Smith and V.K. Lazarov, “*Ab initio* studies of disorder in the full Heusler alloy  $\text{Co}_2\text{Fe}_x\text{Mn}_{1-x}\text{Si}$ ”, *J. Appl. Phys.* **113** 17B106 (2013), doi:10.1063/1.4801745.
- [32] M. Cococcioni and S. de Gironcoli, “Linear response approach to the calculation of the effective interaction parameters in the LDA+U method”, *Phys. Rev. B* **71** 035105 (2005), doi: 10.1103/PhysRevB.71.035105.
- [33] J. Heyd, G.E. Scuseria and M. Ernzerhof, “Hybrid functionals based on a screened Coulomb potential”, *J. Chem. Phys.* **118** 8207 (2003), doi: 10.1063/1.1564060.
- [34] F. Cora., M. Alfredsson, G. Mallia, D. S. Middlemiss, W. C. Mackrodt and R. Dovesi, “The performance of hybrid density functionals in solid state chemistry”, *Structure and Bonding* **113** 171–232 (2004), doi:10.1007/b97944.
- [35] S.J. Clark and J. Robertson, “Screened exchange density functional applied to solids”, *Phys. Rev. B* **82** 085208 (2010), doi: 10.1103/PhysRevB.82.085208.

- [36] K. Xiong, J. Robertson, M.C. Gibson, S.J. Clark, "Defect energy levels in HfO<sub>2</sub> high-dielectric-constant gate oxide", *Appl. Phys. Lett.* **87** 183505 (2005), doi: 10.1063/1.2119425.
- [37] S.J. Clark, J. Robertson, S. Lany and A. Zunger, "Intrinsic defects in ZnO calculated by screened exchange and hybrid functionals", *Phys. Rev. B* **81**, 115311 (2010), doi: 10.1103/PhysRevB.81.115311.
- [38] V.K. Lazarov, Z.H. Cai, K. Yoshida, K.H.L.L. Zhang, M. Weinert, K.S. Ziemer and P.J. Hasnip, "Dynamically Stabilized Growth of Polar Oxides: The Case of MgO(111)", *Phys. Rev. Lett.* **107** 056101 (2011), doi: 10.1103/PhysRevLett.107.056101.
- [39] J.C. Schon and M. Jansen, "Determination, prediction, and understanding of structures, using the energy landscapes of chemical systems - Part I", *Z. Kristallgr.* **216** 307-325 (2001), doi: 10.1524/zkri.216.6.307.20339.
- [40] G.L.W. Hart, V. Blum, M.J. Walorski and A. Zunger, "Evolutionary approach for determining first-principles hamiltonians", *Nat. Mater.* **4** 391-394 (2005), doi: 10.1038/nmat1374.
- [41] N.L. Abraham and M.I.J. Probert, "A periodic genetic algorithm with real-space representation for crystal structure and polymorph prediction", *Phys. Rev. B* **73** 224104 (2006), doi: 10.1103/PhysRevB.73.224104.
- [42] N.L. Abraham and M.I.J. Probert, "An improved real-space genetic algorithm for crystal structure and polymorph prediction", *Phys. Rev. B* **77** 134117 (2008), doi: 10.1103/PhysRevB.77.134117.
- [43] C.W. Glass, A.R. Oganov and N. Hansen, "USPEX - Evolutionary crystal structure prediction", *Comp. Phys. Comm.* **175** 713-720 (2006), doi:10.1016/j.cpc.2006.07.020.
- [44] S. Goedecker, "Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems", *J. Chem. Phys.* **120** 9911-9917 (2004), doi:10.1063/1.1724816.
- [45] Y.C. Wang, J.A. Lv, L. Zhu and Y.M. Ma, "Crystal structure prediction via particle-swarm optimization", *Phys. Rev. B* **82** 094116 (2010), doi: 10.1103/PhysRevB.82.094116.
- [46] C.J. Pickard and R.J. Needs, "High-pressure phases of silane", *Phys. Rev. Lett.* **97** 045504 (2006), doi: 10.1103/PhysRevLett.97.045504.
- [47] C.J. Pickard and R.J. Needs, "*Ab initio* random structure searching", *J. Phys.: Condens. Matter* **23** 053201 (2011), doi: 10.1088/0953-8984/23/5/053201.
- [48] M.H.F. Sluiter and Y. Kawazoe, "Cluster expansion method for adsorption: Application to hydrogen chemisorption on graphene", *Phys. Rev. B* **68** 085410 (2003), doi: 10.1103/PhysRevB.68.085410.
- [49] J.O. Sofo, A.S. Chaudhari and G.D. Barber, "Graphane: A two-dimensional hydrocarbon", *Phys. Rev. B* **75** 153401 (2007), doi: 10.1103/PhysRevB.75.153401.

- [50] D.C. Elias, R.R. Nair, T.M.G. Mohiuddin, S.V. Morosov, P. Blake, M.P. Halsall, A.C. Ferrari, D.W. Boukhvalov, M.I. Katsnelson, A.K. Geim and K.S. Novoselov, "Control of Graphene's Properties by Reversible Hydrogenation: Evidence for Graphane", *Science* **323** 610-613 (2009), doi: 10.1126/science.1167130.
- [51] C.J. Pickard and R.J. Needs, "Metallization of aluminum hydride at high pressures: A first-principles study", *Phys. Rev. B* **76** 024401 (2007), doi: 10.1103/PhysRevB.76.024401.
- [52] C.J. Pickard and R.J. Needs, "Structure of phase III of solid hydrogen", *Nat. Phys.* **3** 473-476 (2007), doi: 10.1038/nphys.6258.
- [53] R.T. Howie, C.L. Guillaume, T. Scheler, A.F. Goncharov and E. Gregoryanz, "Mixed Molecular and Atomic Phase of Dense Hydrogen", *Phys. Rev. Lett.* **108** 125501 (2012), doi: 10.1103/PhysRevLett.108.125501.
- [54] C.J. Pickard, M. Martinez-Canales and R.J. Needs, "Density functional theory study of phase IV of solid hydrogen", *Phys. Rev. B* **85** 214114 (2012), doi:10.1103/PhysRevB.85.214114.
- [55] H.Y. Liu, L. Zhu, W.W. Cui and Y.M. Ma, "Room-temperature structures of solid hydrogen at high pressures", *J. Chem. Phys.* **137** 074501 (2012), doi: 10.1063/1.4745186.
- [56] M.I. Eremets and I.A. Troyan, "Conductive dense hydrogen", *Nat. Mater.* **10** 927-931 (2011), doi: 10.1038/nmat3175.
- [57] C.J. Pickard and R.J. Needs, "Highly compressed ammonia forms an ionic crystal", *Nat. Mater.* **7** 775-779, doi: 10.1038/nmat2261.
- [58] C.J. Pickard and R.J. Needs, "Stable phases of iron at terapascal pressures", *J. Phys.: Condens. Matter* **21** 452205 (2009), doi: 10.1088/0953-8984/21/45/452205.
- [59] M. Martinez-Canales, C.J. Pickard and R.J. Needs, "Thermodynamically Stable Phases of Carbon at Multiterapascal Pressures", *Phys. Rev. Lett.* **108** 045704 (2012), doi: 10.1103/PhysRevLett.108.045704.
- [60] J. Sun, M. Martinez-Canales, D.D. Klug, C.J. Pickard and R.J. Needs, "Persistence and Eventual Demise of Oxygen Molecules at Terapascal Pressures", *Phys. Rev. Lett.* **108** 045503 (2012), doi: 10.1103/PhysRevLett.108.045503.
- [61] C.J. Pickard and R.J. Needs, "Aluminium at terapascal pressures", *Nat. Mater.* **9** 624-627 (2010), doi: 10.1038/nmat2796.
- [62] C.J. Pickard and R.J. Needs, "Predicted Pressure-Induced s-Band Ferromagnetism in Alkali Metals", *Phys. Rev. Lett.* **107** 087201 (2011), doi: 10.1103/PhysRevLett.107.087201.
- [63] K. Kunc and R.M. Martin, "Ab initio force constants of GaAs: A new approach to calculation of phonons and dielectric properties", *Phys. Rev. Lett.* **48**(6), 406-409 (1982), doi: 10.1103/PhysRevLett.48.406.

- [64] S. Baroni, P. Giannozzi and A. Testa, “Green’s-function approach to linear response in solids”, *Phys. Rev. Lett.* **58** 1861–1864 (1987), doi:10.1103/PhysRevLett.58.1861.
- [65] S. Baroni, S. de Gironcoli and A. Dal Corso, “Phonons and related crystal properties from density-functional perturbation theory”, *Rev. Mod. Phys.* **73** 515–562 (2001), doi:10.1103/RevModPhys.73.515.
- [66] S.F. Parker, K. Refson, R.I. Bewley, and G. Dent, “Assignment of the vibrational spectra of lithium hydroxide monohydrate,  $\text{LiOH}\cdot\text{H}_2\text{O}$ ”, *J. Chem. Phys.* **134**, 084503 (2011), doi:10.1063/1.3553812.
- [67] K. Refson, P.R. Tulip, and S.J. Clark, “Variational density-functional perturbation theory for dielectrics and lattice dynamics”, *Phys. Rev. B* **73**, 155114 (2006), doi:10.1103/PhysRevB.73.155114.
- [68] S.F. Parker, S.M. Bennington, J.W. Taylor, H. Herman, I. Silverwood, P. Albers, and K. Refson, “Complete assignment of the vibrational modes of  $\text{C}_{60}$  by inelastic neutron scattering spectroscopy and periodic-DFT”, *PCCP* **13**, 7789 (2011), doi:10.1039/c0cp02956d.
- [69] D. Marx and J. Hutter, “*Ab initio* molecular dynamics: basic theory and advanced methods”, Cambridge University Press (2009). ISBN-10: 0521898633; ISBN-13: 978-0521898638.
- [70] X-Z Li, M.I.J. Probert, A. Alavi and A. Michaelides, “Quantum nature of the proton in water-hydroxyl overlayers on metal surfaces”, *Phys. Rev. Lett.* **104** 066102 (2010), doi: 10.1103/PhysRevLett.104.066102.
- [71] X-Z. Li, B. Walker, M.I.J. Probert, C.J. Pickard, R.J. Needs and A. Michaelides, “Classical and quantum ordering of protons in cold solid hydrogen under megabar pressures”, *J. Phys.: Condens. Matter* **25** 085402 (2013), doi: 10.1088/0953-8984/25/8/085402.
- [72] C. Bonhomme, C. Gervais, F. Babonneau, C Coelho, F Pourpoint, T Azais, S.E. Ashbrook, J.M. Griffin, J.R. Yates, F. Mauri, C.J. Pickard, “First-Principle Calculation of NMR Parameters Using the GIPAW (Gauge Including Projector Augmented Wave) Method: a Chemist's Point of View.”, *Chem. Rev.* **112** 5733-5779 (2012), doi: 10.1021/cr300108a.
- [73] M. Kibalchenko, D. Lee, L. Shao, M.C. Payne, J.J. Titman, J.R. Yates, “Distinguishing hydrogen bonding networks in [alpha]-d-galactose using NMR experiments and first principles calculations”, *Chem. Phys. Lett.* **498** 270-276 (2010), doi: 10.1016/j.cplett.2010.08.077.
- [74] S. Cadars, G. Layrac, C. Gérardin, M. Deschamps, J.R. Yates, D. Tichit, D. Massiot, “Identification and Quantification of Defects in the Cation Ordering in Mg/Al Layered Double Hydroxides”, *Chemistry of Materials* **23** 2821–2831 (2011), doi: 10.1021/cm200029q.
- [75] F. Hoehne, J. Lu, A.R. Stegner, M. Stutzmann, M.S. Brandt, Martin Rohrmüller, Wolf Gero Schmidt, and Uwe Gerstmann, “Electrically Detected Electron-Spin-Echo Envelope Modulation: A Highly Sensitive Technique for Resolving Complex Interface Structures”, *Phys. Rev. Lett.* **106**, 196101 (2011), doi: 10.1103/PhysRevLett.106.196101.
- [76] U. Gerstmann, M. Rohrmüller, F. Mauri and W.G. Schmidt, “*Ab initio* g-tensor calculation for

paramagnetic surface states: hydrogen adsorption at Si surfaces”, *Phys. Status Solidi (c)* **7** 157 (2010), doi: 10.1002/pssc.200982462.

[77] R.J. Nicholls, A.T. Murdock, J. Tsang, J. Britton, T.J. Pennycook, A. Koós, P.D. Nellist, N. Grobert and J.R. Yates, “Probing the Bonding in Nitrogen-Doped Graphene Using Electron Energy Loss Spectroscopy”, *ACS Nano* **7(8)** 7145–7150 (2013), doi: 10.1021/nm402489v.

[78] R. Peverati and D. Truhlar, “The quest for a universal density functional: The accuracy of density functionals across a broad spectrum of databases in chemistry and physics”, *Phil. Trans. A* (2013) (in press).

[79] T. Björkman, A. Gulans, A. V. Krasheninnikov and R. M. Nieminen, “Are we van der Waals ready?”, *J. Phys.: Condens. Matter* **24** 424218 (2012), doi:10.1088/0953-8984/24/42/424218.

[80] D. R. Bowler, T. Miyazaki and M. J. Gillan, “Recent progress in linear-scaling ab initio electronic structure techniques”, *J. Phys. Condens. Matter* **14** 2781 (2002), doi: 10.1088/0953-8984/14/11/303.

[81] D. R. Bowler, R. Choudhury, M. J. Gillan and T. Miyazaki, “Recent progress with large-scale ab initio calculations: The CONQUEST code”, *Phys. Stat. Sol.* **243** 989 (2006), doi: 10.1002/pssb.200541386.

[82] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi and M. C. Payne, “Introducing ONETEP: Linear-scaling density functional simulations on parallel computers”, *J. Chem. Phys.* **122** (8) 084119 (2005), doi:10.1063/1.1839852.

## Figure captions

**Figure 1:** (Left) The Coulomb potential (dashed line) and an example pseudopotential (solid line) for a carbon atom, along with the corresponding radial components of the 2s-orbital. The potential and pseudopotential match beyond  $r_c$ , as do the computed wavefunctions (right).

**Figure 2:** The density of states around the Fermi level (dashed line) for the half-metal Heusler alloy  $\text{CoFe}_{0.5}\text{Mn}_{0.5}\text{Si}$  computed using PBE (top) and PBE+U (bottom). With the PBE exchange-correlation functional the system is predicted to be fully metallic, whereas the inclusion of a modest Hubbard U ( $U=2.1$  eV) opens up a band-gap for the minority spins. Adapted with permission from Hasnip *et al.* [31]. Copyright 2013, American Institute of Physics.

**Figure 3:** Defect levels in  $\text{HfO}_2$  arising from oxygen vacancies (V) and interstitials (I) in a variety of charge states. The top valence band (VB) and bottom conduction band (CB) are shown for reference. **The short lines show the energy level position within the gap and the dots show electron occupancies. On the x-axis, V indicates ‘vacancy’ and I ‘interstitial’ with the superscripts showing the defect charge state.** Reprinted with permission from Xiong *et al.* [36]. Copyright 2005, American Institute of Physics.

**Figure 4:** A comparison between the computed band-gap from a screened-exchange (SX) DFT calculation (blue), an ordinary PBE DFT calculation (green) and the experimental band gap for a variety of materials. Reprinted with permission from Clark and Robertson [35]. Copyright 2010 by the American Physical Society.



**Figure 5:** Gap eigenstate of the charged oxygen vacancy and zinc vacancy in ZnO. The coloured contours show the charge density associated with the defect state ranging from low electron density (blue) through to high electron density (red). Reprinted with permission from Clark *et al.* [37]. Copyright 2010 by the American Physical Society.

**Figure 6:** The potential energy surface experienced by a hydrogen atom on a MgO (111) surface during layer-by-layer growth. MgO has a rocksalt structure, with an A-B-C stacking sequence along (111); initially the surface is OH-terminated (site B, with the Mg sublayer at site A), but as the next Mg layer is deposited (at site C) the OH dissociates and the H atoms rise to the new surface and move across to the potential energy minimum. This energy minimum is at site A, the next site in the MgO stacking sequence. Adapted with permission from Lazarov *et al.* [38]. Copyright 2011 by the American Physical Society.

**Figure 7:** Thermodynamically stable compositions can be predicted using a convex-hull plot (or Maxwell construction more generally). In this case, A is hydrogen, and B is carbon. At ambient pressure the stable compositions are predicted to be H<sub>2</sub> (x=0), CH<sub>4</sub> (x=0.2) and C (x=1, graphite). At 10 GPa further compositions become thermodynamically stable, including polyethene and the predicted (and experimentally confirmed) graphane.

**Figure 8:** Discovery of novel magnetism in the elements. (Left) Ferromagnetism is confined to the elements highlighted in red, promoted by narrow 3d and 4f bands. Many elements, such as aluminum [61], become “electrides” at high pressures - the valence electrons are squeezed between the atomic cores, localising them and leading to band narrowing. Could this localisation lead to unexpected ferromagnetism? Extensive spin unrestricted searches over many elements in the periodic table, at a range of pressures, were performed. The alkali elements highlighted in blue were revealed to exhibit (weak) ferromagnetism. (Right) Potassium (purple) adopts the simple cubic structure at about 20 GPa, and is a ferromagnetic electride [62]. The spin-density isosurface is shown in green. Nearly one electron spin-polarised is localised in the centre of the cubic cell, leading to a CsCl crystal structure with the electron playing the role of the anion. Adapted with permission from Pickard and Needs [62]. Copyright 2011 by the American Physical Society.

**Figure 9:** Visualisation of the eigenvectors of a 31.7 cm<sup>-1</sup> phonon mode with a wavevector  $q=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$  from a DFPT phonon calculation of the Pa3 low-temperature phase of C<sub>60</sub>. “External” librational modes such as this one emerge from exactly the same crystalline periodic lattice dynamics formalism as “internal” molecular deformation modes. Reproduced with the permission of the PCCP owner societies from Parker *et al.* [68].

**Figure 10:** Inelastic neutron spectra of the internal modes of C<sub>60</sub> in the Pa3 phase recorded at the ISIS neutron scattering facility on the TOSCA instrument (blue), the MARI instrument (olive green and black) compared with the predicted spectrum from a DFPT calculation (red). Reproduced with the permission of the PCCP owner societies from Parker *et al.* [68].

**Figure 11:** Calculated Raman spectrum of C<sub>60</sub> in the Fm-3 model structure (blue) compared with the measured spectrum at room temperature (red) corrected for instrumental factors.

**Figure 12:** Trajectories of cold high-pressure structures of hydrogen obtained from simulations with classical and quantum nuclei at 80 GPa starting from the P2<sub>1</sub>/c-24 structure. Yellow balls show the representative configurations of the centroids throughout the course of the simulation. The red rods

show the static (geometry-optimised) structure. A conventional hexagonal cell containing 144 atoms was used. Panels (a), (c), (e), and (g) show the z–x plane and panels (b), (d), (f), and (h) show the x–y plane of the hcp lattice. The four simulations are: (1) MD with classical nuclei at 50 K (panels (a) and (b)), (2) PIMD for D at 50 K (panels (c) and (d)), (3) PIMD for H at 50 K (panels (e) and (f)), and (4) PIMD for D at 150 K (panels (g) and (h)). In the MD simulation, the anisotropic inter-molecular interaction outweighs the thermal and quantum nuclear fluctuations. Therefore, the molecular rotation is highly restricted. The thermal plus quantum nuclear fluctuations outweigh the anisotropic inter-molecular interactions in the PIMD simulations of H at 50 K and D at 150 K. Reprinted with permission from Li *et al.* [71]. Copyright 2013, the Institute of Physics.

**Figure 13:** The structure of galactose as determined by comparison between the computed and observed NMR chemical shifts. (Left) The galactose structure showing hydrogen (white), carbon (grey) and oxygen (red), with hydrogen bonds shown in blue. (Right) The experimental  $^1\text{H}$  NMR spectrum together with the spectra calculated using the two proposed hydrogen bond networks (see text). The atoms contributing to the diagnostic small downfield peak are highlighted. Reprinted from Kibalchenko *et al.* [73]. Copyright 2010, with permission from Elsevier.

**Figure 14:** Hypothetical layered double hydroxide (LDH) structure (left) Atom colours are hydrogen (white), oxygen (red), nitrogen (blue), aluminium (purple) and magnesium (green).  $^1\text{H}$  MAS NMR spectra of two Mg/Al LDH materials showing the proportions of the different hydroxyl groups present (right). Reprinted with permission from Cadars *et al.* [74]. Copyright 2011, the American Chemical Society.

**Figure 15:** (a) Medium angle annular dark field (MAADF) image of a nitrogen substitutional dopant atom in graphene; (b) Model used for simulations; (c) Experimental and calculated EELS data. Reprinted with permission from Nicholls *et al.* [77]. Copyright 2013, the American Chemical Society.

## Short title for page headings

DFT in the solid state