Department of Physics and Astronomy
University College London

Ab initio surface energetics: Beyond chemical accuracy

Simon J. Binnie

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of the
University College London.

2011

I, Simon J. Binnie, confirm that the work presented in the thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Simon J. Binnie

Abstract

Density functional theory (DFT) is the work–horse of modern materials modeling techniques, but scattered evidence indicates it often fails for important surface properties. This thesis investigates how DFT estimates of the surface energy (σ) and molecular adsorption energies of ionic systems are affected by the choice of exchange–correlation (xc) functional. Accurate diffusion Monte–Carlo (DMC) and quantum chemistry (QC) calculations are presented for these quantities showing marked improvement over DFT and agreement of much better than chemical accuracy.

DFT estimates of σ are presented for the (001) surfaces of LiH, LiF, NaF and MgO. Five xc functionals, LDA, PBE, RPBE, Wu–Cohen and PW91 are used. A clear xc functional bias is demonstrated with $\sigma_{\mathrm{LDA}} > \sigma_{\mathrm{WC}} > \sigma_{\mathrm{PBE}} > \sigma_{\mathrm{PW91}} > \sigma_{\mathrm{RPBE}}$.

To improve the picture detailed pseudopotential DMC calculations are presented for LiH and LiF. The lattice parameters and cohesive energies agree with experiment to better than 0.2 % and 30 meV respectively. For LiH novel all–electron DMC calculations are also presented showing significant improvement over pseudopotential DMC.

Accurate all–electron Hartree–Fock calculations of σ for LiH(001) and LiF(001) are presented along with calculations of the LiF bulk using specially adapted Gaussian basis–sets. Combined with existing QC correlation estimates the bulk and surface properties of LiH and LiF show excellent agreement to both experiment and DMC and allow a longstanding disagreement between two experimental estimates for σ_{LiF} to be resolved.

Finally the adsorption energy curve for water on LiH(001) is obtained by both DMC and incremental QC techniques leading to agreement of better than 10 meV. DFT and dispersion corrected DFT estimates are also presented highlighting the large xc functional dependence.

Thus we demonstrate that is possible and necessary to obtain agreement between higher levels of theory and produce benchmark values beyond DFT.

List of Publications

Work from this thesis has appeared in the following publications:

C. Cazorla, S. Binnie, D. Alfè, and M. J. Gillan. Effect of the exchange-correlation energy and temperature on the generalized phase diagram of the 4d transition metals. *High Pressure Res.*, **28**: (2008) 449.

S. J. Binnie, E. Sola, D. Alfè, and M. J. Gillan. Benchmarking DFT surface energies with quantum Monte Carlo. *Mol. Sim.*, **35**: (2009) 609.

S. J. Binnie, S. J. Nolan, N. D. Drummond, D. Alfè, N. L. Allan, F. R. Manby, and M. J. Gillan. Bulk and surface energetics of crystalline lithium hydride: Benchmarks from quantum Monte Carlo and quantum chemistry. *Phys. Rev. B*, **82**: (2010) 165431.

S. J. Nolan, P. J. Bygrave, N. L. Allan, M. J. Gillan, S. J. Binnie, and F. R. Manby. Ch 4: "The hierarchical scheme for electron correlation in solids". in *Accurate Condensed–Phase Quantum Chemistry.* F. R. Manby (Editor) (CRC Press, Boca Raton, 2011).

S. J. Nolan, S. J. Binnie, D. Alfè, N. L. Allan, F. R. Manby, and M. J. Gillan. Bulk and surface energetics of crystalline lithium fluoride. In preparation.

S. J. Binnie, B. Paulus, D. Alfè, and M. J. Gillan. High accuracy benchmarking with quantum Monte Carlo and incremental coupled cluster calculations: Water adsorption on (001) lithium hydride. In preparation.

Presented Talks

"Benchmarks for surface formation energy from quantum Monte Carlo and quantum chemistry". Presented at Ψ_k 2010, Free University, Berlin, Germany

"*Ab initio* surface energetics: Beyond chemical accuracy". Presented at TYC Student Day 2010, Kings College London, UK

Acknowledgements

I would like to thank my supervisors Mike Gillan and Dario Alfè for the support and knowledgeable insight they have provided to me throughout the my time spent working on these calculations. Mike's enthusiasm for solving the great problems is infectious and Dario has offered much welcomed guidance through the often difficult world of quantum Monte–Carlo.

As a computationalist I have also been at the mercy of many bizarre and unfamiliar computational facilities so I would also like to thank the technical staff at the London Centre for Nanotechnology who's help has been invaluable.

As a physicist I am also indebted to the groups of Beate Paulus and Fred Manby of Berlin and Bristol respectively. Both groups have provided vital collaborations for the work done in this thesis and both have been instrumental in inducting me into the world of computational chemistry with great warmth and hospitality.

I also wish to thank the EPSRC who have funded this work and therefore given me the opportunity to perform it.

Finally I would like to thank my friends and family both in London and in the North for their support, good humor and most importantly their skilled belaying.

AMANTISSIMISSPONSIS

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The papers of Hohenberg and Kohn [1] and Kohn and Sham [2] that form the basis of Kohn–Sham density functional theory (DFT) are often said to be among the most cited papers of all time[1]. This is hardly surprising given that DFT is one of the most widely used techniques within the field of electronic structure theory, and one that is used extensively in much cross–discipline work. It is well known however that its main approximation, that of the exchange–correlation (xc) functional is its great weakness and in some cases the choice of xc functional can lead to wildly varying results.

One important area where this can happen, and the area which this thesis concentrates on is that of surface processes. In this thesis we will highlight that for one important surface property, the surface formation energy (σ), current DFT functionals are simply not accurate enough to make meaningful predictions and other more accurate methods must be used. In order to go beyond DFT for this case we will show that it is possible and important that the two leading candidates for higher accuracy, quantum Monte Carlo (QMC) and quantum chemistry agree in their predictions of this property.

We intend to show that for more complex properties, in this case the adsorption energy of a water monomer on the LiH (001) surface DFT is let down by the functionals' inability to accurately describe the surface. Whereas by using both QMC and quantum chemistry techniques these adsorption energies can be computed with perfect agreement and to much better than chemical accuracy.

This chapter is split into four sections, in the first an overview of the ideas behind electronic structure theory are presented along with the universal approximations upon which all of the methods in this thesis are based. Then follows a discussion of surface formation energy, its definition and previous work both experimental and computational. The third section is a similar discussion this time of surface adsorption and adsorption energies. The final section summarises the aims of the thesis and provides a full outline of later chapters.

1.1 Electronic structure theory

The motivation behind modern electronic structure theory is to be able to calculate and predict the properties of materials, the nature of chemical reactions and interactions, spectra and a multitude of other things. Agreement with experiment is the minimum we must expect from any such calculations. If

[1] They top the list of the most cited papers in articles appearing in the journals of the American Physical Society up to 2003.

the calculations cannot be relied upon to accurately reproduce experimental data applications where we are relying upon the predictive power of the calculations, where experimental data is unavailable or confused, would prove to be out of reach. It cannot be an understatement to say that accuracy is king.

As outlined by Gillan [3] these theoretical methods have come a long way since the original simulations of materials that dealt with atoms represented as hard spheres moving in empirical potentials. Modern day methods start with the assumption that the quantum nature of the electrons within a given material is pivotal to its properties.

1.1.1 The main approximations

The first approximation therefore is that we are attempting to solve the Schrödinger equation (SE). So for a system of N electrons and M nuclei with the nuclei having charge Z_μ and mass m_μ the SE is[2],

$$\left[-\sum_{i}^{N}\frac{1}{2}\nabla_i^2-\sum_{\mu}^{M}\frac{1}{2m_\mu}\nabla_\mu^2-\sum_{i}^{N}\sum_{\mu}^{M}\frac{Z_\mu}{r_{i\mu}}+\sum_{i<j}^{N}\frac{1}{r_{ij}}+\sum_{\mu<\nu}^{M}\frac{Z_\mu Z_\nu}{r_{\mu\nu}}\right]|\Psi\rangle=E\,|\Psi\rangle\,. \tag{1.1}$$

Here $|\Psi\rangle$ is an energy eigenstate of the Hamiltonian and E the associated energy eigenvalue.

The SE is not invariant under an arbitrary Lorentz transform hence does not include any relativistic effects. This approximation means that the effects of particle spin must be included via other constraints, something that results in a serious problem for one of the methods used in this thesis, diffusion Monte Carlo. Other relativistic effects such as particles at near relativistic speeds can be included through subsequent corrections.

Modern electronic structure theory usually operates within what is known as the Born–Oppenheimer (BO) approximation. This assumes that due to the nuclei having a much larger mass compared to the electrons, the nuclei can be considered stationary (or clamped in position) while the electrons around them instantaneously relax to their ground state. This reduces the SE to,

$$\left[-\sum_{i}^{N}\frac{1}{2}\nabla_i^2-\sum_{i}^{N}\sum_{\mu}^{M}\frac{Z_\mu}{r_{i\mu}}+\sum_{i<j}^{N}\frac{1}{r_{ij}}\right]|\Psi\rangle=E\,|\Psi\rangle\,. \tag{1.2}$$

The resultant force of this new electronic potential on the nuclei can then be resolved and the nuclear positions relaxed. The electronic ground state can then be recalculated.

It is usually enough to ignore the quantum nuclear effects and treat them just as classical point charges but in cases where the nuclei are light enough to make their nuclear motion important this can be re–introduced using path–integral methods [4, 5, 6].

The one exception to this is the zero–point motion of the nuclei which will be dealt with in due course.

1.1.2 Numerical calculations

In this work we are assuming that the errors introduced by using the SE and the BO approximation are negligible. Essentially we assume that if we had the exact solution to Equation 1.2 for a given system we should be able to calculate its properties with exact agreement to experiment.

[2] Atomic units are used here and throughout the chapter.

Exact solutions to this equation however are not forthcoming once we go beyond the complexity of the hydrogen atom. While Feynman [7] and Lloyd [8] have shown that it is possible to simulate an arbitrary Hamiltonian using the methods of quantum computation, a working quantum computer capable of such calculations is some way off. We are therefore left with more traditional computational techniques.

All of the current computational methods are based upon approximating the problem in some way. Some, such as the family of post Hartree–Fock methods, create a hierarchy where these approximations are slowly removed allowing complete accuracy to be approached. However the technique that is most widely used, and which has revolutionized the field is based upon a rather more uncontrolled approximation, this is density functional theory.

1.1.3 Density functional theory

A full discussion of DFT will come in Section 2.2 but for now it is enough to understand that its main approximation is in the form of the exchange–correlation (xc) functional.

DFT has allowed a large number of high–quality calculations to be performed on numerous physical applications, for instance descriptions of normal bulk materials through to the behavior of those materials under extreme conditions [9]. There are also instances where DFT appears to have failed, from radial distribution functions in water [10] to surface adsorption [11]. In this work we are interested in those situations where it may be necessary to go beyond DFT in order to get accurate results.

1.2 Surface formation energy

One such application comes when we look at systems involving surfaces or interfaces. Catalysis is one area where a good description of the electronic structure at a surface is vital. Also with the rise of nanotechnology surface effects play an important role. Indeed the working definition of something being in the realm of nanotechnology tends to be when the surface dominates the bulk. The range of real world applications of such systems whether it be hydrogen fuel storage [12, 13] to novel methods of drug delivery [14] are huge and provide an excellent area where theory can make a difference.

One particularly important quantity is the surface formation energy (σ) of a given crystal. This energy, associated with a given surface of crystal provides an important role in many processes that happen at the surface. It provides an important contribution to the 'wetting' of a surface and adhesion to it in general. The shape of small nano–particles can be determined partially by a balance of the surface energy vs. the bulk strain. Also the growth of crystal structures on surfaces, for instance the structure of germanium huts on silicon [15] is determined by a balancing of the various surface energies with other variables in the system.

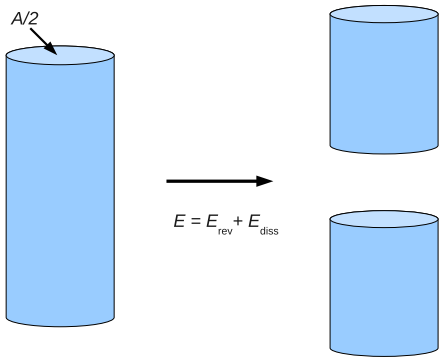
1.2.1 Defining surface formation energy

The surface energy is simply defined as the reversible work done in creating a unit area of a given crystal surface. If one considers a rod of crystal with a cross-sectional area $A/2$ the surface energy[3] would be

[3]Strictly speaking this is only true in the limit of $A \to \infty$ as otherwise there would be a small contribution from the edge formation energy.

E_{rev}/A where E_{rev} is the reversible part of the work done cleaving the rod in two as shown in Figure 1.1.

Figure 1.1: A rod of cross sectional area A is split in two with the work done being E, which itself is made up of a reversible component E_{rev} and a dissipative component E_{diss}.

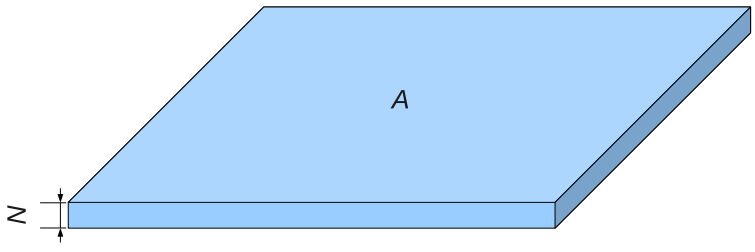


Another equivalent but slightly more useful way of defining it is in terms of a slab of atomic thickness N and surface area A as in Figure 1.2,

$$E_{\mathrm{slab}}(N) = 2A\sigma + NE_{\mathrm{bulk}}. \tag{1.3}$$

Here $E_{\mathrm{slab}}(N)$ is the total energy of the given slab and E_{bulk} represents the total energy per atomic layer as you would find in the *bulk* crystal. In this way σ is seen more as the energy in excess of the bulk obtained at the surface. In the limit of $N \to 0$ all that is left is the energy associated with the surface itself, which is σ.

Figure 1.2: A slab of thickness N with two surfaces both of area A.



1.2.2 Measuring surface formation energy

Direct measurement of the surface formation energy is a problematic area. For systems that are easily cleaved along a given plane such as MgO or LiF the method of Gilman [16] is the one that is frequently referred to in the literature. It is in essence a fairly crude method but one which has the reputation of being effective. A sample block of crystal is obtained, preferably in the form of a cuboidal rod with the cleavage plane of interest running longitudinally along it. A small longitudinal crack is initiated by

placing the block in a screw vice and pressing a blade against the face of the rod. The block is then placed in a cantilever system with the whole assembly placed in a low temperature environment such as liquid nitrogen. The crystal is then pulled apart by the cantilever and the peak force measured using a force balance. Form this is straightforward to estimate the reversible work done in cleaving the crystal and hence σ. The magnitude of the energy lost through dissipative effects such as plastic flow of the crystal can be estimated by etching the newly created surface and counting the number of steps which indicate the level of plastic flow.

One obvious drawback of the method is that the spread of estimates for σ is quite large when the experiment is repeated. In Gilman's work on LiF his estimates range from 0.34 $\mathrm{J\,m^{-2}}$ to 0.5 $\mathrm{J\,m^{-2}}$. In theory the lowest estimate produced by the method should be the most accurate. While Gilman's work and subsequent papers based on it [17, 18] continue to be cited as perhaps the only experimental contribution[4] to surface energies for these ionic systems [20, 21, 22] two papers by Burns & Webb [23, 24] that appear to have been ignored in the literature, provide a dissenting voice about the accuracy of the method.

Taking the case of LiF they obtain a surface formation energy of 0.48(5) $\mathrm{J\,m^{-2}}$, a significant increase on Gilman's estimate of 0.34 $\mathrm{J\,m^{-2}}$. Their method relies on the use of high speed photography to track the crack propagation in a transparent crystal. Applying their derived equations of motion for crack propagation they are able to recover the surface formation energy itself. They account for the difference in their estimate to their inbuilt treatment of plastic deformation effects which is lacking in the Gilman method.

1.2.3 Calculating surface formation energy

Being a quantity that can be obtained from the total energies of periodic systems the calculation of σ can be quite straight forward using electronic structure methods that employ periodic boundary conditions (PBC).

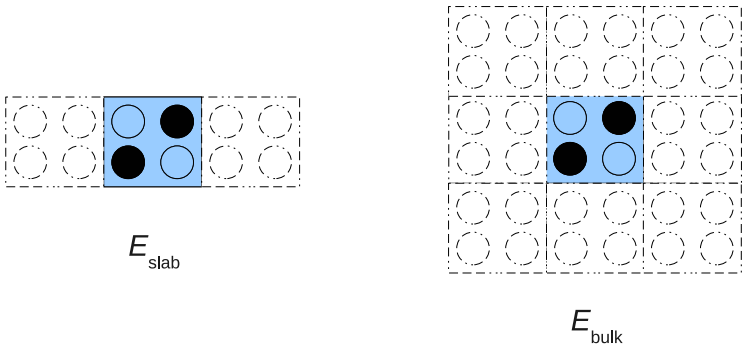
The most obvious way of calculating σ is taking the difference of the total energy between a surface unit cell in 2D PBCs and the corresponding unit cell in 3D PBCs as shown in Figure 1.3.

However Fiorentini and Methfessel [25] have shown that such an approach is hard to converge. This is mainly due to the differing errors that are present in the slab and bulk calculations. These errors tend not to cancel and hence require very highly converged calculations of both the bulk and the slab in order to achieve modest convergence of σ.

The method they propose which is the method adopted in this work is to calculate the total energies of a sequence of slabs of increasing atomic thickness. This gives a dataset that corresponds to $E_{\mathrm{slab}}(N)$. By fitting this to Equation 1.3 an estimate of σ can be obtained. Moreover this estimate takes into account the favorable cancellation of errors between successive slab calculations.

[4] For other systems such as lead other more indirect experimental methods such as that of Bombis *et al.* [19] have produced unquestionably accurate estimates of σ. These methods have not been applied to these simple ionic systems however.

Figure 1.3: An example of a slab and bulk calculation for calculating σ. The unit cells in blue show the cells over which the respective total energies are calculated. The dashed cells show the nearest periodic images of the primary cell in the directions where PBCs are enforced. In both cases periodicity is also enforced in the direction perpendicular to the plane of the paper.



1.2.4 Previous work

Being such an important property it is not surprising that investigations into σ have been performed in the past. As far back as the 1950's calculations of σ and related quantities for ionic systems were performed using the Born–Stern equations [26] based on treating the ions as fixed points in associated inter–ionic potentials with a short range "hard–shell" repulsion around each. It is interesting to note that for these ionic materials the results obtained by Gilman *et al.* matched up quite well [26]. This unfortunate coincidence may be the reason why little more experimental investigation was performed on their surface energy.

1.2.4.1 Jellium

One system that has been of long standing interest to the electronic structure community is jellium or the free electron gas. This is a cloud of electrons set against a neutralizing charged background is a popular system for study when new methods are being developed. The surface energy of jellium has been studied with a variety of methods including DFT, the Random Phase Approximation (RPA) and quantum Monte Carlo (QMC) [27, 28, 29]. Here there is broad agreement between the various different approximations in DFT (the worst agreement coming from PBE) and methods like the RPA. Agreement with QMC and other many–body approaches however was harder to obtain and remained a major problem in the field until a study by Wood *et al.* [30] indicated that error in the QMC calculations were to blame. Wood *et al.* suggested that the many–body finite size–errors had a large effect on $\sigma_{\mathrm{jellium}}$ and did not cancel as well as those in single–particle theory. By adopting an approach similar to that outlined in Section 1.2.3 QMC and single–particle results were reconciled. The fact that DFT could produce good results for a range of electron densities should be quite encouraging suggesting that applying it to real world systems should yield good results.

1.2.4.2 Exploring the periodic table

This optimism would be bolstered by the work of Methfessel *et al.* [31] which found that for a series of transition metal surfaces (from yttrium to sliver) that LDA-DFT agreed quite well with the existing high quality experimental data. Work by Da Silva *et al.* [32] reproduced similar results using a more accurate treatment of the nuclei than Methfessel. However they also performed the calculations using the PBE GGA. This in theory should be an improvement over the LDA used in Methfessel's work. In the cases Da Silva *et al.* studied they found that PBE consistently predicted values of σ that were 30% lower than the LDA, a somewhat amplified echo of the error in PBE from the work on jellium. Yu and Scheffler [33] tell a similar story with lead surfaces again with GGAs being 30% lower than the LDA. In this case strong experimental data by Bombis *et al.* [19] confirms that it is the LDA value that is correct.

In the cases of TiO_2 and SnO_2 Goniakowski *et al.* are less lucky and although they too find that GGAs predict values 30% less than the LDA [34] there are no experimental values available to decide which, if any, are correct.

Finally Alfè and Gillan [35] performed both DFT and QMC studies of the MgO (001) surface. This again confirmed that PBE was 30% lower than the LDA and in this case the LDA agreed with QMC and the experimental work using Gilman's method.

1.2.5 This work

This work sets out to add to this small but growing collection of data on the accuracy of DFT when applied to surfaces. Results of a DFT study of four different ionic surfaces using five different exchange–correlation functionals are presented in Section 4.1. Due to the lack of definitive experimental data the results for LiH and LiF are benchmarked against QMC studies. As the controversy over jellium has shown in the absence of experimental data one method that is theoretically more accurate should not be relied upon without other supporting evidence. To that end Hartree–Fock calculations on LiH and LiF are presented so as to allow comparison of the QMC results with the quantum chemistry correlation estimates of Nolan *et al.* [36]. Through this a clearer picture emerges on the accuracy of DFT and indeed QMC for ionic surfaces.

1.3 Surface adsorption

While surface formation energy is obviously an important quantity it is only one part of the picture. Most real world situations rarely consist of perfectly clean isotropic surfaces. One important consideration is matter that is adsorbed onto the surface.

This interaction between molecules and surfaces is of course central to processes ranging from catalysis to protein adsorption in biological systems. The fact that the interaction is determined by a balance of surface formation energy against other concerns makes it an ideal candidate for a process where the accuracy of DFT can be called into question, requiring methods beyond DFT.

1.3.1 Defining adsorption energy

Firstly it is necessary to fully define how we are going to measure adsorption. Most real world applications consist of one or more layers of adsorbate bound to the surface. For our purposes a certain level of simplification is needed, so we are concerned with a single molecule of adsorbate bound to an infinite perfect surface. While accuracy in this case is not sufficient to be able to obtain accuracy at higher coverage it is necessary and therefore offers a useful test.

So taking the example of a single water molecule adsorbed on a lithium hydride surface there exists a relaxed geometry for the H_2O–LiH complex. We shall denote its total energy as $E(\mathrm{H_2O{-}LiH})$. The isolated water molecule has a total energy $E(\mathrm{H_2O})$, and the surface $E(\mathrm{LiH})$. Hence the adsorption energy is given by,

$$E_{\mathrm{ads}} = E(\mathrm{H_2O{-}LiH}) - E(\mathrm{H_2O}) - E(\mathrm{LiH}). \tag{1.4}$$

Obviously if this quantity must be negative for adsorbed complex to be stable.

By varying the distance between the molecule and the surface in the complex E_{ads} can be plotted as an adsorption curve.

1.3.2 Previous work

As we might expect previous attempts to treat surface adsorption problems within DFT have not been met with unbridled success. The most famous example is the adsorption of CO on the Pt (111) surface. The situation is summed up in the paper of Feibelman *et al.* [11], despite numerous independent studies DFT is unable to predict the correct binding site. Failure of DFT to obtain even qualitative accuracy could be classed as very embarrassing and something which would call into question the suitability of DFT for these calculations.

The work of Hammer *et al.* [37] show that this is far from a single pathological case. In studying molecular adsorption on several metal surfaces they show variations of almost 1 eV in the adsorption energy predicted by different xc functionals for the same system with no functional consistently agreeing with experiment to better than 0.5 eV. Subsequent studies by Stroppa and Kresse [38] and by Schimka *et al.* [39] confirm such worries and also indicate that the use of so–called "hybrid" xc functionals do little to alleviate the problem. Indeed Stroppa and Kresse show that the only functional that gives reasonable adsorption energies, BLYP, also produces a particularly bad description of the bulk and the surface formation energy, suggesting at the least some very lucky error cancellation.

In terms of high accuracy work the main contributions are from quantum chemistry methods. The incremental method of Paulus and Stoll [40] has been applied to the problem of molecular adsorption on ceria with good results [41, 42]. Also Paulus and Rosciszewski [43] have applied the method to hydrogen adsorption on graphene. A full discussion of the incremental scheme will come later but these calculations have not yet demonstrated chemical accuracy for adsorption energies.

An alternative is the cluster approach which has been used for many studies on adsorption energies such as Cu on MgO(001)[44] and others [45, 46]. The most accurate calculation reported using this method is that of water on NaCl(001) by Li *et al.*[47]. This proof of concept calculation shows it is

possible to approach chemical accuracy.

1.3.3 This work

As with the case of surface formation energies we aim in this work to first examine DFT and then go beyond it. Our test system is water adsorption on the LiH (001) surface, a system that is small enough to allow us to fully examine the accuracy of all of the methods we adopt and pave the way for more demanding systems in the future.

So a DFT study examining several different xc functionals and the adsorption energy curves they produce will first be presented in an attempt to establish the shortcomings of DFT.

In terms of going beyond DFT studies on hydrogen dissociation on metal surfaces [48, 49] show that QMC methods are a clear contender. Hence we will present an extensive QMC study of the water–LiH adsorption energy curve allowing direct comparison to be made with the DFT results.

Also we shall present a quantum chemistry study of the system, using both the incremental scheme of Paulus, with whom we have a collaboration, in concert with the cluster approach of Li *et al.*. This allows us to obtain results using the current 'gold–standard' quantum chemistry technique of coupled cluster singles–doubles and perturbative triples, very near to the basis–set limit. Using these results the veracity of the QMC work to be cross–checked demonstrating that we have indeed gone beyond DFT accuracy and in this case achieved much better than chemical accuracy.

1.4 Aims and thesis outline

Having shown the importance of both surface formation energy and surface adsorption the aims of this thesis are two–fold.

Firstly we intend to show that it is indeed necessary to go beyond density functional theory for these calculations. In the case of surface formation energy we have already outlined some of the sparse evidence already available in the literature that this is the case. However to make a more convincing case we will present a detailed study of the surface formation energies of four different ionic systems using a variety of xc functionals. In the case of surface adsorption we will also present a detailed DFT study of the adsorption energy of a single water molecule on LiH (001) surface, showing the large exchange–correlation functional dependence of this quantity.

Secondly, and more importantly, we intend to show that it is possible to go beyond DFT for these two quantities. For both quantities the establishment of accuracy for any method is hampered by the lack of reliable experimental data. In this respect we will demonstrate the high accuracy of our calculations through the novel use of both quantum Monte Carlo and quantum chemistry methods on extended systems in order to produce results that agree to within chemical accuracy.

1.4.1 Outline

The thesis is organized as follows,

- Chapter 1 provides a basic outline of the initial assumptions in any electronic structure theory method, a discussion of surface formation energies and adsorption energies and why it is necessary to go beyond DFT.

- Chapter 2 is an in depth discussion of single–particle methods, including DFT. A discussion of Hartree–Fock theory and post Hartree–Fock methods is included as these are some of the methods we intend to use to provide our "beyond DFT" benchmarks.

- Chapter 3 is a discussion of quantum Monte Carlo (QMC) methods including specifically diffusion Monte Carlo which will be the main method used for going beyond DFT.

- Chapter 4 presents the results of the DFT and QMC investigations into surface formation energies. Also the "hierarchical method" for calculating surface formation energies using post–HF methods is outlined so our QMC results can be compared to existing quantum chemistry calculations.

- Chapter 5 is concerned with the adsorption energy of water on LiH (001) and both DFT and QMC results are presented. The "incremental method" is also explained and results of a post–HF study of the adsorption energy using the method are also presented showing the accuracy of the QMC and post–HF calculations is much better than 1 kcal mol^{-1}, chemical accuracy.

- A summary and full discussion of all of the thesis and its results is given in Chapter 6.

Chapter 2

Single–particle Theory

The most widely used computational methods in electronic structure theory are based on a single–particle type approach. This approach dates back to the approximate method of Thomas and Fermi and is central today in both density functional theory and Hartree–Fock theory.

This chapter aims to review first Hartree–Fock theory then DFT. A full review of DFT methods is obviously required if we are to be able to make any substantive judgments about its accuracy in both the surface energy and the surface adsorption cases. Hartree–Fock theory simultaneously offers useful insight into the advantages and drawbacks of DFT, highlighting the exchange–correlation partition while also leading to a collection of techniques that will allow us to go beyond DFT especially in the case of surface adsorption.

These "post–HF" techniques are also outlined in this chapter are later used in Chapter 5 in order to obtain highly accurate adsorption energies.

2.1 Hartree–Fock theory

The origins of the Hartree–Fock self consistent field method (HF) seem to owe a lot to the iterative ideas involved in the theory itself. Originally proposed by Hartree in 1927 the work was then re–derived by Slater in a more elegant fashion, before being extended by Fock to correctly respect the Pauli exclusion principle. Finally in 1935 Hartree re–visited the theory (including Fock's modifications) an re–derived it in more or less the form we see today. An iterative approach to a theory that relies on iteration towards self–consistency.

2.1.1 Spin–orbitals and the mean–field

The fundamental *ansatz* of Hartree–Fock theory is that the wavefunction solution to the Schrödinger equation can be represented as a Slater determinant of single–particle spin–orbitals,

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.1}$$

here a system of N electrons each with position and spin $\mathbf{x}_n$ and spin–orbital χ_n. The choice of such a wavefunction means that the anti-symmetry of the Pauli exclusion principle is automatically enforced[1]. The energy change that comes about due to this anti-symmetry is termed the *exchange* energy. The idea of these single–particle orbitals must also be explained. Part of the problem of the many–body Schrödinger equation is the electron–electron interaction which, necessarily, is a two–electron operator. This is the many–body nature that causes the difficulties in solving the equation. Hartree and later Hartree–Fock theory adopted a mean–field approach. Here an electron would move not according to explicit interactions with other electrons but would move according to a mean–field determined by all of the electrons in the system.

2.1.2 Fock operator

Firstly lets us re-call the Hamiltonian for a system of N electrons and M nuclei within the Born–Oppenheimer approximation,

$$\hat{H} = -\sum_{i}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i}^{N}\sum_{\mu}^{M} \frac{Z_\mu}{r_{i\mu}} + \sum_{i<j}^{N} \frac{1}{r_{ij}} + V_{\mathrm{nn}}. \tag{2.2}$$

Here the first term is just the kinetic energy of the i^{th} electron, the second term the Coulomb interaction between the i^{th} electron and the μ^{th} nucleus with charge Z_μ, the third term is the electron–electron interaction and V_{nn} is just the energy shift due to the Coulomb repulsion between the nuclei.

The kinetic energy operator and the electron–nuclear interaction are already single–particle operators so we can define a new operator as follows,

$$h_i = -\frac{1}{2}\nabla_i^2 - \sum_{\mu}^{M} \frac{Z_\mu}{r_{i\mu}}. \tag{2.3}$$

So now we can take our ansatz wavefunction $|\Psi\rangle$ and minimize the energy $E = \langle\Psi|\hat{H}|\Psi\rangle$ with the condition that the spin–orbitals are orthonormal. This gives us the one–electron equations,

$$\left[\hat{h} + \sum_{j}^{N}\left(\hat{\mathcal{J}}_j - \hat{\mathcal{K}}_j\right)\right]|\chi_i\rangle = \epsilon_i\,|\chi_i\rangle\,. \tag{2.4}$$

The operators $\hat{\mathcal{J}}_j$ and $\hat{\mathcal{K}}_j$ are the single–particle Coulomb and exchange operators respectively and are described below [50].

The single–particle Coulomb operator is defined as,

$$\hat{\mathcal{J}}_j(\mathbf{x}_1) = \int d\mathbf{x}_2\,|\chi_j(\mathbf{x}_2)|^2\,\frac{1}{r_{12}}. \tag{2.5}$$

It gives the classical Coulomb potential at point $\mathbf{x}_1$ due to the average position of the electron in the j^{th} spin–orbital, this is the mean–field approximation that was made by assuming the wavefunction was a product of single–particle orbitals.

The single–particle exchange operator is defined as,

$$\hat{\mathcal{K}}_j(\mathbf{x}_1)\chi_i(\mathbf{x}_1) = \int d\mathbf{x}_2\chi_j^*(\mathbf{x}_2)\frac{1}{r_{12}}\chi_i(\mathbf{x}_2)\chi_j(\mathbf{x}_1). \tag{2.6}$$

[1] This being the development to Hartree's original theory that Fock introduced.

This arises purely because of the anti–symmetry that was enforced in the wavefunction. It is the repulsive interaction between parallel spin electrons resulting from the Pauli exclusion principle.

It is usual to define the operators in Equation (2.4) collectively as the Fock operator,

$$\hat{f} = \hat{h} + \sum_{j}^{N} \left(\hat{\mathcal{J}}_j - \hat{\mathcal{K}}_j \right). \tag{2.7}$$

At this point it is interesting to note that due to the summation over all orbitals there is a spurious self–interaction term in the Coulomb operator. This is exactly balanced by the same term in the exchange operator. As we will see in Section 2.2 this is not always the case in mean–field approaches.

2.1.3 Self–consistent field

Now with the equations for the single–particle orbitals determined all that is left is to solve the equations themselves. The Fock equations are non-linear in the sense that the Coulomb and exchange operators depend on the orbitals themselves and hence cannot be solved directly.

The usual approach is to expand the orbitals as some linear combination of atomic orbitals,

$$\chi_i = \sum_{k}^{K} c_{ik} \phi_k. \tag{2.8}$$

This reduces the Fock equations to the following matrix equation known as the Roothaan–Hall equations,

$$\mathbf{FC} = \mathbf{SC}\epsilon. \tag{2.9}$$

With the matrix elements of the Fock matrix,

$$F_{ij} = \langle \phi_i | f | \phi_j \rangle, \tag{2.10}$$

and the overlap matrix,

$$S_{ij} = \langle \phi_i | \phi_j \rangle. \tag{2.11}$$

So an initial guess can be made at the orbitals (coefficient matrix $\mathbf{C}$) which can then be used to construct the Fock matrix. Standard diagonalisation techniques can then solve for a new coefficient matrix. The process is then iterated until $\mathbf{C}$ ceases to change outside of some convergence threshold. This is known as reaching self–consistency.

It is important to note that the Fock equation 2.9 is merely the equation that must be satisfied for a given Slater determinant to give the minimum energy when the full Hamiltonian is applied. Hence in order to obtain the groundstate energy it is not sufficient to just sum the eigenvalues of Equation 2.9, instead the final determinant must be substituted into the many–body Schrödinger equation .

2.1.4 Basis sets

The atomic orbitals of Equation 2.8 are more commonly referred to as the "basis–set" of a calculation. Initially the basis–set used was that of the atomic orbitals suggested by Slater [51]. These Slater type orbitals (STOs) take the following form,

$$\phi(\mathbf{r}) = N r^{(n-1)} e^{-\zeta r} Y_l^m(\theta, \phi). \tag{2.12}$$

Here r is the radial distance from the nucleus, n is the principle quantum number, N is some normalization constant and ζ a constant related to the nuclear charge. $Y_l^m(\theta, \phi)$ are the usual hydrogenic spherical harmonics.

STOs provide a very accurate description of the behavior of atomic orbitals both near to the nucleus and far away from it. Importantly they obey the Kato cusp conditions of which there will be further discussion in Section 3.2.1.2. This is a great benefit because in order to completely span Hilbert space the basis must be infinitely large. Hence in actual calculations the set must be truncated. With STOs their inherent qualitative accuracy means relatively few basis functions need to be included to produce high accuracy results. Unfortunately however from a computational point of view evaluating integral products of STOs is relatively expensive as the integrations must be performed numerically.

2.1.4.1 Gaussian type orbitals

The alternative that now forms the vast majority of Hartree–Fock calculations done today is to use Gaussian type orbitals (GTOs). Here the exponential in Equation 2.12 is of Gaussian form leading to,

$$\phi(\mathbf{r}) = Nr^{(n-1)}e^{-\zeta r^2}Y_l^m(\theta, \phi). \tag{2.13}$$

These orbitals now do not have the correct qualitative behavior and thus require far more functions when using truncated expansions. However integral products can be evaluated analytically in a fraction of the time it would take to evaluate the same integrals if they contained STOs. The smallest basis–set that can be used for a given problem contains the same number of basis functions as spin–degenerate orbitals in the system. This is referred to as a 'minimal' or 'single–zeta' basis. When using GTOs a single basis function is often a summation or 'contraction' of several Gaussians, termed primitives. The coefficients of this contractions are held constant and are usually determined by least squares fitting to an STO or by optimization using the atomic system.

Obviously the variational freedom of the basis–set can be increased by including twice as many basis functions as the minimal system. This is termed a double–zeta basis. This idea can be continued in order to produce a hierarchy of basis–sets of increasing accuracy. Also orbitals of higher angular momentum can be added to the basis to increase variational freedom. These are termed polarization functions, and they allow for the distortion of orbitals encountered in bonding situations.

2.1.4.2 The HF complete basis limit

Many such hierarchies of basis–sets have been produced, mainly for use with post Hartree–Fock methods. However Jensen has created a set of polarization consistent basis–sets (pcN) specifically for Hartree–Fock calculations [52]. This hierarchy is ordered with polarization functions of increasing angular momentum added to the basis. Hence the pc0 set for hydrogen consists of just two s orbitals, the pc1 set includes a p orbital etc. Jensen has shown that this hierarchy converges exponentially towards the complete basis–set limit (CBS) and a three–point extrapolation of the following form [53] can be used,

$$E(N) = E_{\mathrm{CBS}} + Be^{-CN}. \tag{2.14}$$

Here $E(N)$ is the energy obtained with the pc$N$ basis–set, E_{CBS} is the energy in the complete basis–set limit, C and B represent fitting parameters.

It is worth noting that post–HF methods with a correlation consistent basis–sets can also be extrapolated this time with,

$$E(N) = E_{\mathrm{CBS}} + BN^3. \tag{2.15}$$

2.1.4.3 Basis–set superposition

In cases where complete basis–set limit is not reached some calculations can suffer from what is known as basis–set superposition error.

This occurs when calculations involving energy differences are performed. Say we want the interaction energy of the water dimer, we would take the difference between the interacting water system and twice the isolated water molecule. However in the interacting system the basis on one water molecule is augmented by the diffuse tails from the basis on the other water molecule. This means that compared to the isolated system energy (where the basis is not augmented) the interacting system has its energy reduced thus reducing the interaction energy.

In order to mitigate this such calculations are often performed using whats known as "counterpoise correction". Here when the isolated system is calculated it is done with the basis functions of the second molecule in place as if in the interacting system. All of the calculations using Gaussian basis–sets in this work are done using this scheme where appropriate.

2.1.5 Virtual orbitals and excitations

If something other than a minimal basis–set is used then after solving the Roothaan–Hall equations using the SCF method a set of K orbitals are produced, where K basis–functions were used in the expansion. For a system of N electrons obviously only the first N orbitals are occupied in the ground–state. The remaining $K - N$ orbitals are usually termed *virtual orbitals*. It is of course possible to create a determinant from a set of Hartree–Fock orbitals where one or more of he electrons have been promoted into one or more virtual orbitals. In the second quantization such a system could be constructed in the following way,

$$|\Psi_i^a\rangle = \hat{a}_a^\dagger \hat{a}_i \, |\Psi\rangle \tag{2.16}$$

Here $|\Psi\rangle$ is the HF groundstate, $|\Psi_i^a\rangle$ is the single excitation where the electron in orbital i has been excited to the virtual orbital a, $\hat{a}_a^\dagger$ and $\hat{a}_i$ represent the creation operator for state a and annihilation operator for state i respectively. A double excitation could be constructed in a similar way,

$$|\Psi_{ij}^{ab}\rangle = \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j \, |\Psi\rangle \, . \tag{2.17}$$

In fact for an N electron system with K basis functions the total number of determinants that could be formed is given by,

$$\frac{K!}{N!\,(K-N)!}. \tag{2.18}$$

These excited wavefunctions are an important part of attempts to improve on the Hartree–Fock approximations and shall be revisited in Section 2.3

2.2 Density Functional Theory

2.2.1 Hohenberg–Kohn

With the full many–body Schrödinger equation looking to be intractable for all but the simplest systems the problem obviously needed to be reformulated. Hohenberg and Kohn managed to transform the problem in order to drastically reduce it dimensionality [1]. They dispensed with the need to solve the many–body Schrödinger equation in order to obtain the ground–state wavefunction of the system and from this the ground–state energy. Instead they showed it was only necessary to find the ground state electron density of the system. This they suggested would uniquely determine the ground state energy.

They did this through proving two fairly simple theorems.

The first states that for a system of particles operating in an external potential $V_{\rm ex}(\mathbf{r})$, the ground state particle density $n_0(\mathbf{r})$, is sufficient to fully–determine the external potential. Obviously it follows that the ground state particle–density fully determines the Hamiltonian of the system, since the particle–particle interaction and the kinetic energy operator for the particles are common to all systems containing those particles. Given that the Hamiltonian determines the wavefunctions of all of the states of the system we conclude that the wavefunctions and all of the properties of the system are fully–determined by $n_0(\mathbf{r})$.

Proving this is incredibly simple. Let us assume there are two Hamiltonians $\hat{H}_a$ and $\hat{H}_b$ which take the form,

$$\hat{H}_{a,b} = \hat{T} + \hat{V}_{\rm int} + V_{a,b}(\mathbf{r}) \tag{2.19}$$

where $\hat{T}$ and $\hat{V}_{\rm int}$ are the kinetic energy operator and particle–particle interaction operator respectively. $V_{a,b}(\mathbf{r})$ represents the non–equivalent external potentials for system a and system b. Systems a and b have groundstate wavefunctions Ψ_a and Ψ_b respectively. Now let us assume they both have the same ground state particle–density $n_0(\mathbf{r})$.

The variational principle states that for an arbitrary wavefunction $|\Psi\rangle$,

$$E_0 \leq \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{2.20}$$

that is that the expectation value of the Hamiltonian acting on an arbitrary wavefunction is always greater than or equal to the ground–state energy. This allows us to write,

$$\langle\Psi_a|\hat{H}_a|\Psi_a\rangle < \langle\Psi_b|\hat{H}_a|\Psi_b\rangle\,. \tag{2.21}$$

Using Equation (2.19) we can re–write this in the following way,

$$\langle\Psi_a|\hat{H}_a|\Psi_a\rangle < \langle\Psi_b|\hat{T}|\Psi_b\rangle + \langle\Psi_b|\hat{V}_{\rm int}|\Psi_b\rangle + \int d^3r V_a(\mathbf{r})n_0(\mathbf{r}) \tag{2.22}$$

$$< \langle\Psi_b|\hat{H}_b|\Psi_b\rangle + \int d^3r\,[V_a(\mathbf{r}) - V_b(\mathbf{r})]\,n_0(\mathbf{r}). \tag{2.23}$$

The arbitrary indices can be swapped leading to the following inequality,

$$\langle\Psi_a|\hat{H}_a|\Psi_a\rangle + \langle\Psi_b|\hat{H}_b|\Psi_b\rangle < \langle\Psi_b|\hat{H}_b|\Psi_b\rangle + \langle\Psi_a|\hat{H}_a|\Psi_a\rangle\,. \tag{2.24}$$

As this can clearly never be satisfied[2] we can see that the initial theorem must hold.

The second theorem states that there is a functional $E[n(\mathbf{r})]$ for the total energy of a system valid for any external potential $V_{\text{ex}}(\mathbf{r})$ and that the global minimum of this functional is the groundstate energy for that system. Also this minimum occurs at the ground state particle density; when $n(\mathbf{r}) \equiv n_0(\mathbf{r})$.

Since the first theorem has shown that the ground state particle–density determines the wavefunction and the Hamiltonian it is reasonable to write the following,

$$E = \langle \Psi | \hat{T} + \hat{V_{\text{int}}} | \Psi \rangle + \int d^3 r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \tag{2.25}$$

$$= T[n(\mathbf{r})] + V_{\text{int}}[n(\mathbf{r})] + \int d^3 r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) = E_{\text{HK}}[n(\mathbf{r})] \tag{2.26}$$

We can then define a functional $F[n(\mathbf{r})]$ as,

$$F[n(\mathbf{r})] = T[n(\mathbf{r})] + V_{\text{int}}[n(\mathbf{r})]. \tag{2.27}$$

So the total energy functional is given by,

$$E_{\text{HK}}[n(\mathbf{r})] = F[n(\mathbf{r})] + \int d^3 r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \tag{2.28}$$

Now if we take a Hamiltonian $\hat{H}$, its ground state wavefunction Ψ_0 and associated ground state particle–density n_0, we find its ground state energy E_0 is,

$$E_0 = E[n_0] = \langle \Psi | \hat{H} | \Psi \rangle \tag{2.29}$$

Again appealing to the variational principle we can see that for any other arbitrary wavefunction Ψ_x with its associated density n_x,

$$E_0 < \langle \Psi_x | \hat{H} | \Psi_x \rangle \tag{2.30}$$

$$E_0 < E[n_x]. \tag{2.31}$$

So by minimizing $E[n]$ with respect to n we necessarily obtain the ground state energy and the ground state particle–density. In the expression for $E[n(\mathbf{r})]$ the functional $F[n(\mathbf{r})]$ known as the Hohenberg–Kohn functional, is universal for a given particle type within the system. Hence if $F[n(\mathbf{r})]$ was known for a given particle type any system consisting of those particles in an external potential could be solved by direct minimization of $E[n(\mathbf{r})]$. Needless to say for systems of interacting electrons $F[n(\mathbf{r})]$ remains unknown and the Hohenberg–Kohn theorems offer no guidance in how to find it. In order to make progress with calculations of actual systems we must use the method proposed by Kohn and Sham in 1965 [2].

2.2.2 Kohn–Sham

Where Hohenberg and Kohn re–framed the problem of solving the many–body Schrödinger equation into a problem of minimizing a functional of an electron density Kohn and Sham re–framed the problem of an interacting many–body system into that of a non–interacting one. This mean–field approach is

[2]The implicit assumption here is that the ground state is non–degenerate. Further derivation can be done to extend the Hohenberg–Kohn theorem to include degenerate ground states. [54]

similar to Hartree–Fock theory in that the particles are only interacting with a mean–field of the effects of all of the particles in the system. This substitution of systems can be justified if we consider that the total energy functional of the interacting system can be partitioned as follows,

$$E[n(\mathbf{r})] = T_{\mathrm{S}}[n(\mathbf{r})] + \int V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'d\mathbf{r} + E_{\mathrm{NN}} + E_{\mathrm{xc}}[n(\mathbf{r})]. \tag{2.32}$$

Here we have $T_{\mathrm{S}}[n(\mathbf{r})]$ being the kinetic energy for the non–interacting system of density $n(\mathbf{r})$, the second term is due to the external potential, the third term the classical Hartree interaction between an element of the density at a given point in space and the rest of the density, finally $E_{\mathrm{xc}}[n(\mathbf{r})]$ is defined as the "exchange–correlation" functional for the system. This expression however can be seen as the energy functional for an non–interacting system subject to a potential defined as,

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{\mathrm{xc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}. \tag{2.33}$$

Such a system would consist of N single particle orbitals $|\psi_i\rangle$ which would obey,

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r})\right]|\psi_i\rangle = \epsilon_i\,|\psi_i\rangle\,. \tag{2.34}$$

These orbitals are usually termed Kohn–Sham orbitals, and the density is of course given by,

$$n(\mathbf{r}) = \sum_i \langle\psi_i|\mathbf{r}\rangle\,\langle\mathbf{r}|\psi_i\rangle\,. \tag{2.35}$$

They must be found self–consistently given the circular dependency of effective potential on the density. The introduction of an orbital representation does of course increase the complexity of the problem. This is necessary however due the lack of the exact form of $T_{\mathrm{S}}[n]$. Orbital–free DFT which harks back to the earlier Thomas–Fermi theory and attempts to solve exclusively with the density is an area of on–going research but as of yet fails to live up to the accuracy of Kohn–Sham DFT.

It is clear that the crucial element to all of this is E_{xc}. What Kohn and Sham did was to concentrate all of our ignorance about the full many–body system into this functional. It contains in it the difference between the interacting and the non–interacting kinetic energy and crucially the difference between the classical Hartree interaction and the true many–body Coulomb interaction. Furnished *deus ex machina* with the exact form of E_{XC} the only limit to simulation would be computer power. Every system could in principle be solved, within of course the BO approximation and in the non–relativistic limit. It is unfortunately the case that the form of E_{XC} remains an approximation. The following section deals with this issue and outlines some of the approximations used.

2.2.3 Exchange–correlation functionals

2.2.3.1 Local density approximation

The first attempts at an exchange–correlation functional came with the idea that the electrons in the system could be thought of in the limit of the homogeneous electron gas (HEG). As already described this is a system of interacting electrons in a neutralizing background positive field. The exchange part can be written down analytically but the correlation part must be calculated for many different electron densities[3] and then parameterized in some way. This gives you a function for $\epsilon_{\rm xc}(n)$.

One then makes the approximation that the exchange–correlation functional $E_{\rm xc}[n]$, only depends on the local value of $\epsilon_{\rm xc}(n)$. This is called the local density approximation (LDA) and leads to the following expression,

$$E_{\rm xc}^{\rm GGA}[n] = \int n(\boldsymbol{r})\epsilon_{\rm xc}(n(\boldsymbol{r}))d^3\boldsymbol{r}. \tag{2.36}$$

Amazingly this naïve assumption works quite well and is still used for many calculations, however it can be expanded into the generalized gradient approximation (GGA) which functionals such as PW91 and PBE are based on.

2.2.3.2 Generalized gradient approximation

Here one assumes a dependence not only on the local value of the electron density but its local gradient as well, giving us a form of,

$$E_{\rm xc}^{\rm LDA}[n] = \int n(\boldsymbol{r})\epsilon_x^{\rm HEG}(n)F_{\rm xc}(n, |\nabla n|)d^3\boldsymbol{r}. \tag{2.37}$$

here $\epsilon_x^{\rm HEG}(n)$ is the exchange part of the HEG energy and $F_{\rm xc}$ is a new enhancement factor that depends on n and $|\nabla n|$.

It is in this enhancement factor that different GGA functionals differ (two examples can be seen in Section 4.1.1.1). These enhancement factors do have to be constructed in a very careful way as $|\nabla n|$ for real systems is often quite large and would lead naïve choices of $F_{\rm xc}$ to breakdown.

Finally recent advances have enabled the use of what are termed *hybrid–functionals*. These are functionals such as B3LYP [56, 57] where traditional functionals are mixed with a certain proportion of Hartree–Fock like exact–exchange. It still remains however that exchange–correlation functionals are an open problem.

[3] Usually Ceperley & Alder's QMC calculations are used [55].

2.2.4 Periodic boundary conditions and plane–waves

Perfect crystal systems are obviously of great interest. Determining the bulk properties of such systems relies on accurate modeling. While extrapolations using finite clusters of atoms can be used to model these systems it is generally advantageous to consider them in terms of periodic boundary condiditons (PBC).

PBCs as the name implies aim to model these extended systems by enforcing their inherent periodicity. A given primitive cell is surrounded by an infinite array of identical images allowing the infinite system to be modeled directly.

2.2.4.1 k-point sampling

In order to avoid having to explicitly calculate the interactions of an infinite number of electrons directly an appeal is made to Bloch's Theorem. This states that for a periodic system the energy eigenstates can be written as,

$$\phi_k(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_k(\boldsymbol{r}) \tag{2.38}$$

where $u_k(\boldsymbol{r})$ is some function periodic with the primitive cell. The vector $\mathbf{k}$ defines the *crystal momentum* of the electron in the given state. It is unique only to within translations by a reciprocal lattice vector hence $\mathbf{k} = \mathbf{k} + \mathbf{b}_n$ where $\mathbf{b}_n$ is a multiple of the reciprocal lattice vector. The structure of the eigenstates in k–space essentially encodes the information about arbitrary single particle states in the infinite system. Therefore any system wide property such as total energy or dipole moment must be intergrated across k–space in the first Brillouin zone[4].

In essence what Bloch's theorem has done is to replace an infinite sum over electrons with an integral over k–space.

While in practice as this would require an infinite number of evaluations in k–space we settle for summing over a number of pre–defined 'k–points'. This is possible due to the smooth nature of most functions of interest in k–space. Many schemes for generating such points have been devised but perhaps the most commonly used is that created by Monkhorst and Pack [58] which uses the following expression,

$$\boldsymbol{k}_{n_1,n_2,n_3} = \sum_{i}^{3} \sum_{n_i}^{N_i} \frac{2n_i - N_i - 1}{2N_i} \boldsymbol{G}_i \tag{2.39}$$

where $\boldsymbol{G}_i$ is the primitive reciprocal lattice vectors, and N_i is the maximum number of points in a given direction.

This is very useful as not only does it avoid certain points of high–symmetry that are undesirable for insulator calculations, but also it allows systematic improvement via the N_i values. This is very important as with any scheme where a full integration is approximated by a summation you must test the convergence of the results with respect to the approximation. Examples of this can be seen in later sections.

[4] The periodicity of $\mathbf{k}$ ensures that any point in k–space can be mapped to a point in the first Brillouin zone.

2.2.4.2 Plane waves

Going back to Equation 2.38 it is clear that in order to find solutions of this form a suitable basis–set expansion must be found for the periodic functions $u_k(\mathbf{r})$.

Gaussian type orbitals have already been briefly discussed in Section 2.1 and it is possible to use these as the basis–set expansion as for instance the CRYSTAL [59] code does. This approach has drawbacks as convergence with respect to the basis–set can be hard to achieve. Diffuse functions nominally centred on one atom in the primitive cell end up overlapping with others causing problems of linear dependencies within the basis. A technique for overcoming this is described in Section 4.5.4 but it is far from straight forward.

Another solution is to expand in terms of plane–waves in the following way,

$$u_k(\mathbf{r}) = \sum_n c_{n,\mathbf{k}} e^{i\mathbf{b}_n \cdot \mathbf{r}}. \tag{2.40}$$

Here $\mathbf{b}_n$ is the n^{th} reciprocal lattice vector as defined previously. Substituting into Equation 2.38 gives,

$$\phi_k(\boldsymbol{r}) = \sum_n c_{n,\mathbf{k}} e^{i(\mathbf{b}_n + \boldsymbol{k}) \cdot \boldsymbol{r}}. \tag{2.41}$$

Obviously this infinite summation must be truncated in order to provide tractable calculations. It turns out that using an energy cut–off defined as,

$$E_{\text{cut}} = \frac{1}{2} \left| \mathbf{b}_{\text{cut}} + \boldsymbol{k} \right|^2 \tag{2.42}$$

a systematic truncation of the system can be found which gives good results. Here any plane–wave with a wavevector larger than b_{cut} is omitted. Convergence can be tested by gradually increasing E_{cut} and therefore b_{cut} until the change in the desired observable is small enough.

2.2.5 Pseudopotentials

When using Gaussian basis–sets pseudopotential (pp) techniques are useful. The basic idea is that the nucleus and generally some of the core electrons are replaced by a smooth potential that recreates the scattering effect that the nucleus and core has on the valence electrons. This means the core electrons do not need to be calculated thus reducing the cost of the calculation. This potential, when correctly constructed, also causes the valence wavefunctions which would vary quite rapidly within the core region to vary quite slowly and smoothly, thus reducing the complexity of basis–set required to adequately describe them.

When using plane–waves however these pseudopotentials become almost essential. The fast varying core regions of wavefunctions that are a nuisance when using Gaussians are a disaster with plane–waves.

This section describes the usual norm–conserving type pseudopotentials used in this work, and a slightly different approach to the problem, that of the projector augmented wave method (PAW) also adopted in this work. Common to both is the idea of a non–local potential. This means that the potential felt by a given valence electron is determined partially by that electron's own angular momentum. This allows the potential to be separated into local and non–local parts.

$$V(\mathbf{r}) = V_{\mathrm{loc}}(\mathbf{r}) + \delta V_l(\mathbf{r}) \tag{2.43}$$

The PWscf code that is used in this work adopts what is known as Kleinman–Bylander projectors [60] when dealing with non–local pseudopotentials. Here Equation 2.43 is recast as,

$$\hat{V} = V_{\mathrm{loc}}(\mathbf{r}) + \sum_{lm} \frac{|\psi_{lm}\delta V_l\rangle \langle \psi_{lm}\delta V_l|}{\langle \psi_{lm}|\delta V_l|\psi_{lm}\rangle}. \tag{2.44}$$

Here some atomic reference pseudo–wavefunctions ψ_{lm} are calculated once for each pseudopotential and then used in the creation of the projection operators. This is a very efficient way of implementing non–local pseudopotentials from a computational point of view, but it is an approximation that is dependent on the suitability of the atomic reference states used.

2.2.5.1 Norm–conserving

Pseudopotentials can be constructed using a variety of methods from fitting to experimental data to the *ab initio* approach that is adopted in this work. The advent of norm–conserving pseudopotentials (NCPPs) marked perhaps the start of an era where reliable pseudopotentials that can be used in a variety of situations and systems (a property known as transferability) are available for most light elements. The oft repeated four requirements of a good pseudopotential after Hamann *et al.* [61] are repeated here:

1. Real and pseudo valence eigenvalues agree for a chosen "prototype" atomic configuration.
2. Real and pseudo atomic wavefunctions agree beyond a chosen "core radius" r_{c}.
3. The integrals from 0 to r of the real and pseudo charge densities agree for $r > r_{\mathrm{c}}$ for each valence state (norm conservation).

4. The logarithmic derivatives of the real and pseudo wave function and their *first energy derivatives* agree for $r > r_c$.

Point four is required in order to make the pseudopotential transferable, as it means that any change in the eigenstates of the true system compared to the atomic one for which the pp was constructed will be correct to first order [54]. Points one two and three were shown by Hamann *et al.* to be sufficient to ensure this condition was upheld. This is what leads to the great success of NCPPs.

The stages in modern production of NCPPs are generally done in the following steps:

1. A core–valence partition is assumed i.e. the electrons that will be pseudized are decided upon.
2. A set of appropriate atomic configurations are chosen.
3. The corresponding all–electron (AE) wavefunctions are calculated using the desired xc functional. This is usually done on a radial grid which must be sufficiently fine.
4. The potential (with a given radial cut–off) is generated using one of the many methods available. This is usually done by fitting some analytical function to the AE wavefunction and then inverting the Schrödinger equation.
5. The pp should then be thoroughly tested to asses its transferability.

Obviously this process is often painstaking and is regarded as something of a mysterious art, requiring many different choices for r_c , the electronic reference configuration, fitting functions etc. to be explored before a good pp is found.

2.2.5.2 Projector augmented waves

The drawbacks of NCPPs are that often quite hard[5] pps will be required in order to obtain good transferability. The 'ultrasoft' pseudopotentials of Vanderbilt [62] recast the problem so that the norm–conserving condition could be dropped allowing for accurate pps that did not require such high energy cut–offs. This method has been generalized into the Projector Augmented Wave (PAW) method by Blöchl [63]. It is a slightly different approach to the problem of the fast-varying core of the wavefunctions. Here instead of using a pseudopotential and in effect discarding this part of the true wavefunction one defines a transformation $\boldsymbol{T}$, between the true one particle wavefunction $|\Psi\rangle$, and a smoothed pseudo-wavefunction $|\Psi^{\mathrm{PS}}\rangle$. This transformation is defined as,

$$\boldsymbol{T} = 1 + \sum_N \boldsymbol{T}_N \tag{2.45}$$

where $\boldsymbol{T}_N$ is some atom centered contribution that is non-zero inside a certain atomic cut-off radius and N is the number of atoms.

The atom centered contribution can be thought of as,

$$\sum_N \boldsymbol{T}_N = \sum_i (|\phi_i\rangle - \left|\phi_i^{\mathrm{PS}}\right\rangle) \langle p_i| \tag{2.46}$$

[5] A *hard* pp is usually one with a very small r_c that usually requires a correspondingly large plane–wave cut–off before energy convergence is obtained.

where $|\phi_i\rangle$ is the all-electron partial–wave which is obtained by performing reference calculations on the isolated atom, $\left|\phi_i^{\mathrm{PS}}\right\rangle$ is the smoothed pseudo–partial–wave for the system which is equal to $|\phi_i\rangle$ outside the atomic cut–off and a projector function defined as $\langle p_i|\phi_j^{\mathrm{PS}}\rangle = \delta_{ij}$.

Hence one can minimize the energy with respect to $|\Psi^{\mathrm{PS}}\rangle$, expanded in plane–waves, just as in a pseudopotential calculation but can then regain the accuracy afforded by the all-electron wavefunction by this transformation.

It is also normal for such calculations to treat a number of defined core–electrons within the frozen–core approximation. This is where these electrons are fixed to positions calculated with a reference calculation usually done on the isolated atom. This reduced the computational workload since fewer electrons have to be considered explicitly.

2.3 Post Hartree–Fock

One thing can be said about both DFT and HF methods. They are both quantitatively incorrect. DFT is incomplete due to the fact the exact xc functional is not known. Systematic improvement in this area is difficult as a result of this. HF theory however is incomplete due to the fact it does not take into account the correlated motion that arises due to the Coulomb interaction between individual electrons. This can be seen as a controlled approximation and it is because of this that a number of so–called '*post Hartree–Fock*' methods have been developed that use HF as a starting point and gradually approach the full many–body solution (still of course within the BO approximation).

2.3.1 Configuration interaction: Moving into Slater determinant space

Perhaps the most obvious way of correcting for the deficiencies of Hartree–Fock theory is what is termed configuration interaction (CI) theory. As has been demonstrated in Section 2.1.5 as well as producing the Hartree–Fock groundstate wavefunction the SCF procedure also produces all of the other eigenstates of the Fock operator through the various excitations of the groundstate with respect to the virtual single particle orbitals. As the Fock operator is Hermitian this set of states should completely span the Hilbert space of the true many–body wavefunction, in the given one–electron basis. It follows that the true many–body groundstate $|\Phi_0\rangle$ can be expanded in this set of Slater determinants,

$$|\Phi_0\rangle = \sum_{n}^{M} c_n |\Psi_n\rangle \tag{2.47}$$

$$= c_{\mathrm{HF}} |\Psi_{\mathrm{HF}}\rangle + \sum_{a>i} c_i^a |\Psi_i^a\rangle + \sum_{a>i,b>j} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle + \sum_{a>i,b>j,c>k} c_{ijk}^{abc} |\Psi_{ijk}^{abc}\rangle + \ldots + \sum_{a>i,b>j,\ldots,m>N} c_{ij\ldots N}^{ab\ldots m} |\Psi_{ij\ldots N}^{ab\ldots m}\rangle \,. \tag{2.48}$$

Here M is the number of Slater determinants that can be constructed using the given Hartree–Fock orbitals with N electrons. By minimizing the energy with respect to the expansion coefficients the exact groundstate within the given one–electron basis will be obtained. This approach of using the full expansion is termed full CI (FCI).

As Equation 2.18 shows the number of determinants in the full expansion increases factorially with the number of one–electron basis functions and electrons. This makes FCI calculations viable for only very small molecules in limited basis–sets.[6] An alternative is to truncate the expansion after a certain order of excitation. As Brillouin's theorem states that singly excited determinants cannot mix directly with the Hartree–Fock groundstate the lowest non-vanishing order of theory is CI doubles (CID) where only the double excitations are included in the expansion. Single excitations can mix via double excitations leading to CI singles, doubles (CISD).

[6] Recent work by the Alavi group has led to a stochastic implementation of FCI termed initiator quantum Monte Carlo FCI (i-QMCFCI) that should extend the applicability of this method [64, 65, 66]. However it only reduces the prefactor, not the actual scaling itself.

2.3.1.1 Size–consistency

Although such truncations can provide an effective way to approach the FCI limit they do have one major drawback, they lack size–consistency. So if fragments of a larger system are calculated then compared to the energy of the full system (say for cohesive energies) at a given level of excitation an error will be made. This is because the full system and the fragment will not be calculated at the same accuracy. This can be easily seen by imagining a system consisting of two helium atoms at an infinite distance. As these would be non-interacting at this point the energy of this system should be just twice the energy of a single helium atom. Calculating and isolated helium atom using CISD includes a double excitation, so when doubling the energy this can be seen as having a contribution where both helium atoms are doubly excited. However in the original system this would be termed a quadruple excitation and hence be discarded, making an error.

2.3.2 Coupled cluster methods

In order to recover size–consistency the coupled cluster (CC) *ansatz* is generally used. Here one defines a general excitation operator as the sum of all of the excitation operators at each specific order,

$$\hat{T} = \sum_i \hat{T}_i. \tag{2.49}$$

Where the individual excitation operators are defined as,

$$\hat{T}_1 = \sum_{a,i} c_i^a a_i a_a^\dagger \tag{2.50}$$

$$\hat{T}_2 = \frac{1}{4} \sum_{a,b,i,j} c_{ij}^{ab} a_i a_j a_a^\dagger a_b^\dagger \tag{2.51}$$

$$\vdots$$

The exact groundstate wavefunction is then expanded as,

$$|\Phi_0\rangle = e^{\hat{T}} |\Psi_{\mathrm{HF}}\rangle \tag{2.52}$$

$$= \left(1 + \hat{T}_1 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{\hat{T}_1^2}{2} + \frac{\hat{T}_2^2}{2} + O(\hat{T}_3)\right) |\Psi_{\mathrm{HF}}\rangle\,. \tag{2.53}$$

Now at this stage if the energy of the expansion was minimized with respect to the coefficients the FCI energy should be recovered as the expansion should contain all possible excitations. However if the general excitation operator was truncated so that it only contained the single and double excitation operators,

$$\hat{T} = \hat{T}_1 + \hat{T}_2, \tag{2.54}$$

the CISD energy would not be recovered. It is easy to see from Equation 2.53 that the wavefunction expansion would not only contain all of the single and double excitations (as in CISD) but also certain sets of higher excitations generated by products of excitation operators. It is the inclusion of such higher excitations that render the theory size–consistent and avoid the problem of the non-interacting helium atoms in the previous section. This size–consistency makes coupled cluster methods very popular when

it comes to accurate calculations on small systems. Indeed CCSD(T), where perturbation theory is used to estimate the contribution of the triple excitations is often referred to as the *gold–standard* in quantum chemistry.

2.3.2.1 Scaling

This size–consistency is not without its computational cost however. CCSD formally scales as $O(N^6)$ as does CISD. However CCSD has a larger prefactor due to the inclusion of higher excitations. This means that in terms of CPU time truncated CI methods converge faster than CC methods. Indeed a paper by Kàllay [67] indicates that while only CC(6) gives the same total energy convergence as CI(8) the CC calculation takes three times as long. Due to this poor scaling there is need for a quicker size consistent method for use where the high accuracy is not required. This comes in the form of second order Møller–Plesset theory.

2.3.3 Møller–Plesset perturbation theory

The correlation energy missing from HF theory is generally a small fraction of the total energy[7] of a given system. Another reasonable starting point would be therefore to use perturbation theory in order to recover the correlation effects. This was the approach of Møller and Plesset in what has become known as Møller–Plesset (MP) perturbation theory.

As with usual perturbation theory the full Hamiltonian is decomposed into an exactly solvable part ($\hat{H}_0$) and a perturbation ($\hat{H}'$) as follows,

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}'. \tag{2.55}$$

Here λ is a arbitrary parameter that allows the strength of the perturbation to be adjusted. As is usual the eigenvalues and eigenstates of the full Hamiltonian, $\hat{H}$ are expanded in a Maclaurin series in λ,

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \tag{2.56}$$

$$|n\rangle = |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots . \tag{2.57}$$

Here $E_n^{(0)}$ and $|n^{(0)}\rangle$ are the energy eigenvalues and and eigenstates of $\hat{H}_0$. Substituting Equations 2.55, 2.56 and 2.57 into the Schrödinger equation yields and infinite set of coupled equations. Solving in terms of powers of λ and imposing orthonormality constraints on the eigenstate corrections the energy corrections at each order can be written down as follows,

$$E_n^{(0)} = \langle \Psi_n^{(0)} | \hat{H}_0 | \Psi_n^{(0)} \rangle \tag{2.58}$$

$$E_n^{(1)} = \langle \Psi_n^{(0)} | \hat{H}' | \Psi_n^{(0)} \rangle \tag{2.59}$$

$$E_n^{(2)} = \langle \Psi_n^{(0)} | \hat{H}' | \Psi_n^{(1)} \rangle \tag{2.60}$$

$$\vdots$$

$$E_n^{(m)} = \langle \Psi_n^{(0)} | \hat{H}' | \Psi_n^{(m)} \rangle \tag{2.61}$$

[7] It can be a large fraction of any energy differences hence is still an important effect for many physical properties.

Modern day MP theory takes the Fock operator of Equation 2.7 as the unperturbed Hamiltonian $\hat{H}_0$. Obviously the perturbed Hamiltonian is taken from the many–body Schrödinger equation, hence the perturbation must be,

$$\hat{H}' = \hat{H} - \hat{f} \tag{2.62}$$

$$= \sum_{i<j}^{N} \frac{1}{r_{ij}} - \sum_{j}^{N} \left(\hat{\mathcal{J}}_j - \hat{\mathcal{K}}_j \right). \tag{2.63}$$

Calculating the first two corrections for the ground state is a trivial task. Substituting in gives,

$$E_0^{(0)} = \sum_{i}^{N} \epsilon_i \tag{2.64}$$

$$E_0^{(1)} = E_{\mathrm{HF}} - \sum_{i}^{N} \epsilon_i. \tag{2.65}$$

Here ϵ_i is the i^{th}Hartree–Fock eigenvalue and E_{HF} is the Hartree–Fock groundstate energy. So in this formulation MP1 (first order Møller–Plesset perturbation theory) recovers the Hartree–Fock groundstate. Hence the first improvement to the HF energy must come with the second order correction, or MP2 theory. In order to calculate $E_0^{(2)}$ the first order correction to the groundstate wavefunction, $|\Psi_0^{(1)}\rangle$ must be obtained. This can be expanded in the eigenstates of the Fock operator,

$$|\Psi_0^{(1)}\rangle = \sum_{i}^{\infty} c_j \, |j^{(0)}\rangle \, . \tag{2.66}$$

The coefficients c_j are given by,

$$c_j = \frac{\langle \Psi_j^{(0)} | \hat{H}' | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_j^{(0)}}. \tag{2.67}$$

Substituting this result into Equation 2.60 gives an expression for the second order correction,

$$\sum_{j} \frac{\left| \langle \Psi_j^{(0)} | \hat{H}' | \Psi_0^{(0)} \rangle \right|^2}{E_0^{(0)} - E_j^{(0)}}. \tag{2.68}$$

This sum over all excitations can be simplified further by the application of Brillouin's theorem which allows all of the singly excited determinants to be omitted. Also the Slater–Condon rules state that all determinants that are triply excited and above have a vanishing contribution. This is due to the perturbation operator being a sum of one and two–electron operators. Hence Equation 2.68 can be re–expressed in terms of double excitations,

$$\frac{1}{4} \sum_{i,j,a,b} \frac{\left| \langle \Psi_{ij}^{ab} | \hat{H}' | \Psi_0 \rangle \right|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \tag{2.69}$$

Recalling Equation 2.17 it is clear that in this double excitation two electrons from ground state orbitals i and j are promoted into the virtual orbitals a and b. The use of only double excitations allows the second order correction to be calculated with a formal scaling of $O(N^5)$.

Higher order of corrections can be calculated and there has been some work done using MP3 and MP4 theory. Unfortunately however the increase in accuracy compared to cost associated with these

calculations makes MP2 by far the more commonly used method. In fact studies on the higher orders of MP theory (up to MP39) indicate [68] that MP theory can show alarming oscillatory convergence and in some cases divergence.

2.3.3.1 Success or luck?

It is against this background it is always interesting to note the astounding success of MP2 theory. Much like DFT which is a fairly radical approximation, here the treatment of correlation as a small perturbation to only second–order, produces in some cases very accurate results. In Chapter 5 we will see that MP2 theory reproduces cluster adsorption energies to within ~ 3 meV of CCSD(T) for small basis–sets. It is perhaps worth noting the view of Case and Manby [69] who suggest that this is down to a remarkable cancellation of errors. They suggest that on the one hand MP2 theory overestimates the correlation between a pair of electrons which should produce a error of about 15% in the correlation energy. However this is cancelled by the inclusion of the eigenvalues of the Hartree–Fock virtual orbitals which contain the spurious self–interaction, noticed by Kelly in 1963 [70], which leads to an underestimation of the correlation effects by a similar amount.

2.3.3.2 Density–fitting

Density–fitting (DF) or the resolution of the identity method has become the main way in which the speed of MP2 calculations is increased [71]. An optimized Gaussian auxiliary basis is used to speed up the calculation of the four index two electron integrals that appear in Equation 2.69. Taking the usual Coulomb integral,

$$\int \chi_i^*(\mathbf{x}_1)\chi_j^*(\mathbf{x}_1)r_{ij}^{-1}\chi_i(\mathbf{x}_2)\chi_j(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 \tag{2.70}$$

$\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_1)$ can be expanded in an auxiliary basis $\tilde{\chi}_m$ giving,

$$\sum_m \left[\int \chi_i^*(\mathbf{x}_1)\chi_j^*(\mathbf{x}_1)\tilde{\chi}_m(\mathbf{x}_1)d\mathbf{x}_1\right] \int \tilde{\chi}_m^*(\mathbf{x}_1)r_{ij}^{-1}\chi_i(\mathbf{x}_2)\chi_j(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2. \tag{2.71}$$

The coefficients of this expansion can be calculated very quickly and the resulting three index integral is also far more computationally favorable than the original four index one. The same ideas can also be applied to Hartree–Fock theory [72] producing DF-HF theory. Obviously the auxiliary basis used is not complete so an error is made in this approximation, however with an auxiliary basis optimized for use with a given one–electron basis the errors are generally less than 2 meV and indeed in this work they are estimated to be of the order of 30 μeV. The advantages of the method are huge. For both MP2 and HF calculations their density fitted counterparts are usually around an order of magnitude faster increasing the range of applicability by a large amount.

2.3.4 Application to extended systems

The paucity of any efficient implementations of these methods within periodic boundary conditions leads to serious issues in applying them to extended systems.

The hierarchical method of Manby *et al.* [73] allows the correlation energy of perfect crystalline bulk and surfaces to be obtained and the work of Nolan *et al.* [36] has produced high quality data-sets for both LiH and LiF. A full discussion of the method and the difficulty in obtaining the required Hartree–Fock data will be presented in Section 4.4.

A different approach, which allows calculations of the adsorption energy with post–HF methods is that of the incremental scheme of Paulus and Stoll [40]. We use this method extensively in Chapter 5 and a full description of it can be found in Section 5.3.4.1.

Chapter 3

Quantum Monte Carlo

The term Quantum Monte-Carlo (QMC) is used in the literature to cover a broad number of semi–overlapping techniques from variational and diffusion Monte–Carlo to path integral Monte–Carlo and auxiliary field Monte–Carlo. This chapter aims to outline the basics of Monte–Carlo integration (perhaps best seen as the starting point for all of these methods) before going into detail on the two methods used in this work, namely variational Monte–Carlo (VMC) and diffusion Monte–Carlo (DMC). These methods have been the subject of several recent reviews [74, 75, 76] which are referenced in this work.

All of the QMC calculations performed in this work use the CASINO code [75].

3.1 Monte–Carlo integration

Lying at the heart of the various techniques that are usually grouped under the heading of Quantum Monte–Carlo (QMC) is the idea of Monte–Carlo integration.

The problem that Monte–Carlo integration seeks to solve is that of evaluating multi–dimensional definite integrals. Usual quadrature methods are generally extensions of the one dimensional case. Here one takes an evenly spaced set of points at which to evaluate the integrand, and then applies Simpsons's rule or similar. This gives an error in the estimate that scales as N^{-4}, where N is the number of points. For multi-dimensional integrals one can appeal to Fubini's theorem which allows the decomposition of a d–dimensional integral into d iterated, one–dimensional integrals on which we can apply Simpson's rule. The error however now scales spectacularly badly with respect to the number of dimensions, going as $N^{-4/d}$. If we were dealing with a wavefunction with say 100 electrons we would now have a error that scales as $N^{-1/75}$ requiring far too many function evaluations for any serious modeling attempt.

Monte-Carlo integration however allows us to have an error that scales as $N^{-1/2}$, regardless of dimensionality. At a stroke it allows us to evaluate this multi–dimensional integrals in a reasonable amount of computer time.

Let's consider an integral of the form,

$$I = \int_V f(\mathbf{x}) d\mathbf{x}. \tag{3.1}$$

The simplest realization of Monte–Carlo integration is where one takes N completely random points

$[\mathbf{x}_1, \ldots, \mathbf{x}_N]$ within the volume V and then sums over them as follows,

$$\tilde{I} = \frac{1}{N} \sum_i^N f(\mathbf{x}_i). \tag{3.2}$$

This gives us an estimate $\tilde{I}$, for the value of the integrand with an estimate of the error given by,

$$\tilde{\Delta I} = \frac{1\sigma}{\sqrt{N}} \tag{3.3}$$

where σ is the standard deviation of $f(\mathbf{x})$ estimated by,

$$\sigma^2 = \frac{1}{N-1} \sum_i^N \left(f(\mathbf{x}_i) - \tilde{I} \right)^2. \tag{3.4}$$

3.1.1 Importance sampling

The above method assumes no knowledge about $f(\mathbf{x})$ other than our ability to evaluate it at an arbitrary point. However if we know the approximate shape of $f(\mathbf{x})$ we can improve the efficiency of our integration dramatically, using importance sampling.

Let us take the following change of variables,

$$I = \int_V f(\mathbf{x}) d\mathbf{x} = \int_V \frac{f(\mathbf{x})}{p(\mathbf{x})} dP(\mathbf{x}) \tag{3.5}$$

where we have contrived $p(\mathbf{x})$ to satisfy the conditions of a probability density and also satisfy,

$$p(\mathbf{x}) = \frac{\partial^d}{\partial x_1 \ldots \partial x_d} P(\mathbf{x}). \tag{3.6}$$

Drawing our random points from the probability distribution $P(\mathbf{x})$ our estimate of the integral is now given by,

$$\tilde{I} = \frac{1}{N} \sum_i^N \frac{f(\mathbf{x}_i)}{p(\mathbf{x}_i)}. \tag{3.7}$$

It is easy to see that the better $p(\mathbf{x})$ matches $f(\mathbf{x})$ the lower the statistical error. In fact if $f(\mathbf{x}) = Ip(\mathbf{x})$ the error is zero with only one point, however this does require us already knowing the answer.

3.1.2 Metropolis accept/reject

Sampling large probability distributions without knowing their normalization constants, is far from a straightforward problem to solve. In 1953 Metropolis *et al.* [77] proposed a scheme that is now known as the Metropolis accept/reject algorithm, in order to overcome this.

The idea is simple, the set of points $\mathbf{x}_i$ are generated by a random walk controlled in the following way. At each step a move from $\mathbf{x}$ to $\mathbf{x}'$ is suggested by a transition probability $T(\mathbf{x}'; \mathbf{x})$. The move is accepted with probability,

$$p = \min \left(1, \frac{P(\mathbf{x}')T(\mathbf{x}'; \mathbf{x})}{P(\mathbf{x})T(\mathbf{x}; \mathbf{x}')} \right). \tag{3.8}$$

This keeps the walker near areas where $P(\mathbf{x})$ is large. It can be shown that after an initial period of equilibration this produces a set of points distributed as $P(\mathbf{x})$ which is ideal for importance sampled Monte-Carlo integration.

It is an interesting historical aside to note that the work on Monte–Carlo integration and the accept/reject algorithm were conducted on the MANIAC machine at Los Alamos for the US hydrogen bomb program. The machine was capable of 10,000 instructions per second. Part of the work in this thesis was done on the Oak Ridge Laboratory's Jaguar machine for which the nominal instructions per second measurement is of the order of 10^{15}.

3.2 Variational Monte–Carlo

The basic idea behind VMC is to combine the variational principle with importance sampled Monte-Carlo integration.

The well known variational principle states that for an arbitary trial wavefunction Ψ and an arbitrary Hamiltonian $\hat{H}$ the energy given by applying the Hamiltonian to the trial function provides an upper bound to the true groundstate energy.

$$E_{\mathrm{v}} = \frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} \geq E_0 \tag{3.9}$$

It is also the case that when $E_{\mathrm{v}} = E_0$ the trial wavefunction must be the groundstate.

If we recall Equation 3.5 it is simple to see how we might rearrange Equation 3.9 in order to apply importance sampled Monte–Carlo techniques.

$$E_{\mathrm{v}} = \frac{\int |\Psi|^2 [\Psi^{-1} \hat{H} \Psi] d\tau}{\int |\Psi|^2 d\tau} \tag{3.10}$$

This gives us a function to evaluate that we will term the *local energy*, $E_{\mathrm{l}} = \hat{H}\Psi/\Psi$ and an importance function from which to sample, $P(\mathbf{x}) = |\Psi(\mathbf{x})|^2 / \int |\Psi|^2 d\tau$.

Now by using the Metropolis accept/reject algorithm we are able to obtain an estimate of E_{v} to arbitrary precision for an arbitrary trial wavefunction.

3.2.1 Trial wavefunctions

None of that is much good however, if we have no way of proposing, or indeed improving trial wavefunctions. We also note that in order for the importance sampling to be efficient our trial wavefunction needs to be a reasonable approximation to the actual ground state.

With this in mind it seems reasonable that the first element in our trial wavefunction should be the single particle orbitals obtained by DFT or Hartree–Fock theory. These mean field approaches give a reasonably good first approximation to the actual wavefunction. So our first attempt would be a trial function of the form,

$$\Psi_T = \mathbf{D}_\uparrow \mathbf{D}_\downarrow \tag{3.11}$$

where $\mathbf{D}_\uparrow, \mathbf{D}_\downarrow$ are the spin-up and spin-down Slater determinants from single particle theory.

We find that this only gives us an energy estimate variational to the Hartree-Fock groundstate. This is due to the fact that a single Slater determinant is unable to describe any electron correlation beyond the Pauli principle.

3.2.1.1 Jastrow factor

In order to account for electron correlation we multiply the Slater determinant by a collection of terms known as a Jastrow factor. This gives us a trial wavefunction of the form,

$$\Psi_T = \mathbf{D}_\uparrow \mathbf{D}_\downarrow e^J. \tag{3.12}$$

The form J most commonly used, and the form used in this work, is what is known as a three–body Jastrow factor [75],

$$J = \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^{N} \chi_I(r_{Ii}) + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} u(r_{ij}) + \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} f_I(r_{iI}, r_{jI}, r_{ij}) \tag{3.13}$$

where χ, u and f are the one, two and three-body terms respectively, N the number of electrons, N_I the number of nuclei and r the interparticle distance. All three terms have cut-off radii beyond which they are zero in order to reduce the computational work needed in evaluating them. The actual form of each of the terms is subject to a lot of variation in the literature but in this work we use those suggested by Drummond *et al.* in 2004 [78]. These are polynomial expansions up to adjustable expansion orders, which can be tuned so as to obtain the best form for a given application.

The physical significance of the two body term is to reduce the magnitude of the wavefunction when the interparticle distance is small, correcting the pair correlation function. However this has a knock on effect by altering the electron density. This is unfortunate given that the electron density produced by single particle theory is fairly good and this degrades it. Thus the one-body term is needed to compensate for this and restore the correct charge density. The three body term is often important introducing interactions neglected by the other terms. [74, 75]

3.2.1.2 Kato cusp conditions

Something that also needs to be understood is the cusp conditions. These are the conditions imposed on the wavefunction in the event of electron–electron coalescence or electron–ion coalescence. In both cases if the wavefunction does not have the correct shape at these points the local energy will diverge. This causes problems in both VMC and DMC. The electron–electron cusp is usually imposed in a satisfactory manner via the Jastrow factor.

Kato's electron–nucleus cusp condition[1] is given by,

$$\left.\frac{\partial \langle \Psi(r) \rangle}{\partial r}\right|_{r=0} = Z\Psi(r=0). \tag{3.14}$$

Here $\langle\rangle$ indicates a spherical average, r is the electron–nuclear separation and Z the nuclear charge.

All–electron Slater determinants are usually constructed from Gaussian basis functions which do not have the correct shape at the point charge of the nucleus to satisfy this condition. While this can be corrected by the Jastrow factor this approach is far from ideal. More usually the s–functions are replaced

[1] The picture is actually more complex since recent work by Tóth [79] has suggested more fundamental first and second–order boundary conditions at the nucleus, with full angular dependence. In this work however Kato's original conditions are deemed sufficient.

within a pre–defined radial cut–off by the correct empirically observed cusp, the full scheme is described in Ref [80].

During the course of this research a more general scheme was devised by Drummond [81] where a function is added to the wavefunction in order to recover the form of the cusp. This generalized correction can be applied regardless of the basis the wavefunction is expanded in.

3.2.2 Optimization

Armed with an efficient way of calculating the energy and a flexible trial–wavefunction we now need to improve the wavefunction and therefore the estimation of the groundstate energy.

While the single particle orbitals supplied in the Slater determinant are generally left untouched the Jastrow factor we have introduced contains adjustable parameters for each of its terms. A simple approach would be to vary these parameters with the aim of reducing the ground state energy directly. This *energy–minimization* technique has had a rough time since its inception. Numerically it has always suffered from being quite unstable, prone to producing wildly unphysical low energies. It is also very sensitive to over–parameterization of the Jastrow factor. Developments by Umrigar *et al.* [82, 83] have improved matters and generally energy minimization gives the lowest VMC energy, however you still need a good initial guess at the Jastrow factor for it to truly work.

This is generally why the initial optimization is done using the *variance minimization* scheme. We can see from Section 3.1 that when the variance of the local energy is zero the trial function must be exact, therefore giving the lowest energy. It seems reasonable to attempt to minimize the variance in order to produce a better trial function. It can be shown however, that a wavefunction with a lower variance does not necessarily give a lower energy [75]. Though variance minimization still remains a useful technique in optimizing the trial wavefunction.

3.2.3 VMC algorithm

A basic algorithm for performing a VMC calculation runs as follows:

1. N configurations each of M electrons are randomly generated.
2. For each configuration n_i:
   (a) For each electron position $\mathbf{r}_j$ in configuration n_i:
      i. A move is proposed $\mathbf{r}_i \rightarrow \mathbf{r}'_i$.
      ii. The move is accepted with probability $\min\left(1, \left|\frac{\Psi(\mathbf{r}_i)}{\Psi(\mathbf{r}'_i)}\right|^2\right)$.
   (b) The new value of E_{l} (or any other required observable) for the configuration n_i is accumulated.

This process is repeated until the variance of the accumulated E_{l} or other observables is below an acceptable level. In practice a period of equilibration is needed initially so a few hundred iterations are run without step 2(b). During this time the timestep is usually optimized as well so as to ensure the acceptance rate is $\sim$50%. As Frenkel and Smit [84] point out there seems to be little physical or mathematical basis for this oft quoted rule of thumb in Monte–Carlo simulations, however in the case of

VMC this target seems to be as good as any with little to be gained from a more expensive analysis of the optimum timestep for each system.

A further modification known as "two–level sampling" is also made in order to improve computational efficiency. Here step 2(a)ii is split. An initial accept/reject is based on $\min\left(1, |D(\mathbf{r}_i)/D(\mathbf{r}'_i)|^2\right)$ is made and only if that is passed a secondary accept reject based on $\min\left(1, |\exp[J(\mathbf{r}_i)]/\exp[J(\mathbf{r}'_i)]|^2\right)$ is performed. This reduces the need for the relatively expensive step of evaluating the Jastrow factor for moves that are not going to be accepted anyway.

3.3 Diffusion Monte–Carlo

Diffusion Monte Carlo is part of a class of QMC methods that are known as projector methods. They all centre around the imaginary time Schrödinger equation (ITSE),

$$-\frac{\partial}{\partial\tau}\left|\Phi(\mathbf{r},\tau)\right\rangle = (\hat{H} - E_{\mathrm{r}})\left|\Phi(\mathbf{r},\tau)\right\rangle . \tag{3.15}$$

Here τ is imaginary time and E_{r} is a constant that will be explained later.

They all exploit the fact that the ITSE wavefunction can be expressed as,

$$\left|\Phi(\mathbf{r},\tau)\right\rangle = e^{-(\hat{H}-E_{\mathrm{r}})\tau}\left|\phi(\mathbf{r})\right\rangle \tag{3.16}$$

$$= \sum_n e^{-(E_n-E_{\mathrm{r}})\tau}\left|\phi_n(\mathbf{r})\right\rangle\left\langle\phi_n(\mathbf{r})|\phi(\mathbf{r})\right\rangle . \tag{3.17}$$

Where $\sum_n \left|\phi_n(\mathbf{r})\right\rangle\left\langle\phi_n(\mathbf{r})\right|$ is an expansion in eigenfunctions of $\hat{H}$. One can see that in the limit of $\tau \to \infty$ Equation 3.17 projects out $\left|\phi_0(\mathbf{r})\right\rangle$, the ground state of $\hat{H}$ providing of course our initial ϕ_{T} was not orthogonal to it. The offset E_{r} can be adjusted so that the multiplicative factor applied to $\left|\phi_0(\mathbf{r})\right\rangle$ is constant with respect to imaginary time and does not also decay exponentially.

3.3.1 Drift, diffusion and branching

In DMC the imaginary–time evolution of Equation 3.17 is performed via Brownian motion in configuration space.

The Hamiltonian for a general system can be regarded as the sum of a kinetic term $\hat{T}$ and a potential term $\hat{V}$.

$$\hat{H} = \hat{T} + \hat{V} \tag{3.18}$$

From inspection of Equation 3.17 it is clear that in the case of a system of N electrons and M nuclei within the Born–Oppenheimer approximation these terms become,

$$\hat{T} = -\sum_i^N \frac{1}{2}\nabla_i^2 \tag{3.19}$$

$$\hat{V} = -\sum_i^N \sum_k^M \frac{Z_\mu}{r_{i\mu}} + \sum_{i<j}^N \frac{1}{r_{ij}} + V_{\mathrm{nn}} \tag{3.20}$$

Neglecting $\hat{V}$ and setting E_r to zero Equation 3.15 is just equivalent to the master equation for a diffusion process given by,

$$\frac{\partial}{\partial t}\phi(\mathbf{r},t) = D\nabla^2\phi(\mathbf{r},t) \tag{3.21}$$

where D is the diffusion coefficient and ϕ is the particle density. Including the $\hat{V}$ term and relaxing the condition on E_r in this equation would be equivalent to a birth/death or branching process in the particle evolution, causing the population to fluctuate with time. Unfortunately in this case the fact that $\hat{V}$ includes Coulomb interactions would cause the population to diverge at an electron–nucleus coalescence. This obviously causes problems when trying to numerically model the system in a stable fashion.

The solution to this is bound up in the idea of importance sampling discussed in Section 3.1.1. Replacing $\Psi(\mathbf{r},\tau)$ by $f(\mathbf{r},\tau) = \Phi_{\mathrm{T}}(\mathbf{r})\Psi(\mathbf{r},\tau)$ where Φ_{T} is some trial–wavefunction still satisfies the requirements of a solution to Equation 3.15. Substituting in, the following is obtained,

$$-\frac{\partial}{\partial\tau}f(\mathbf{r},\tau) = \hat{T}f(\mathbf{r},\tau) - \nabla\cdot\left(f(\mathbf{r},\tau)\nabla\ln\Psi_{\mathrm{T}}\right) + \left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - E_{\mathrm{r}}\right)f(\mathbf{r},\tau). \tag{3.22}$$

It is immediately obvious that this solution describes again a diffusion process (given by the first term) with a fluctuating particle density (given by the third term). The $\hat{H}\left|\Psi_{\mathrm{T}}\right\rangle / \left|\Psi_{\mathrm{T}}\right\rangle$ component of the third term is what was described in Section 3.2 as the *local energy* E_{l}, of the trial wavefunction. Away from the nodal surface this quantity should not diverge and so the population should now be free of the instabilities such divergences can cause. Looking at the Smoluchowski equation for a drift–diffusion process,

$$\frac{\partial}{\partial t}\phi(\mathbf{r},t) = D\nabla^2\phi(\mathbf{r},t) - \zeta\nabla\cdot\left(\phi(\mathbf{r},t)\nabla U(\mathbf{r})\right), \tag{3.23}$$

where ζ represents some frictional coefficient and $U(\mathbf{r})$ some potential acting on the system, it is clear the second term in Equation 3.22 represents a drift motion in configuration space.

Therefore using importance sampling from some initial trial wavefunction the true ground state of the system can be projected out through imaginary time evolution which as has been shown can be performed using a drift–diffusion–branching process in configuration space.

3.3.2 Short timestep propagator

Obviously in order to perform the simulation the correct propagator for taking $f(\mathbf{r},\tau) \rightarrow f(\mathbf{r}',\tau+\delta\tau)$ in configuration space is required. This can easily be found by integrating Equation 3.22 with the appropriate Green's function,

$$f(\mathbf{r}',\tau+\delta\tau) = \int G(\mathbf{r}',\mathbf{r};\delta\tau)f(\mathbf{r},\tau)d\mathbf{r}. \tag{3.24}$$

The Green's function $G(\mathbf{r}',\mathbf{r};\delta\tau)$ can of course be found by the fact it must satisfy Equation 3.22 leading to the following solution,

$$G(\mathbf{r}',\mathbf{r};\delta\tau) = \left\langle\mathbf{r}\right| e^{-\left[\hat{T}'+\left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}-E_{\mathrm{r}}\right)\right]\delta\tau}\left|\mathbf{r}'\right\rangle, \tag{3.25}$$

where $\hat{T}'$ is defined as,

$$\hat{T}'f(\mathbf{r},\tau) = \hat{T}f(\mathbf{r},\tau) - \nabla\cdot\left(f(\mathbf{r},\tau)\nabla\ln\Psi_{\mathrm{T}}\right). \tag{3.26}$$

Using the Trotter–Suzuki formula the following approximation can be obtained,

$$G(\mathbf{r}',\mathbf{r};\delta\tau) = \langle\mathbf{r}|\, e^{-\left[\hat{T}'+\left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}-E_{\mathrm{r}}\right)\right]\delta\tau}\,|\mathbf{r}'\rangle\,, \tag{3.27}$$

$$\approx \langle\mathbf{r}|\, e^{\left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}-E_{\mathrm{r}}\right)\delta\tau/2} e^{-\hat{T}'\delta\tau} e^{\left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}-E_{\mathrm{r}}\right)\delta\tau/2}\,|\mathbf{r}'\rangle\,, \tag{3.28}$$

$$\approx (2\pi\delta\tau)^{-3N/2} e^{-(\mathbf{r}-\mathbf{r}'-\delta\tau\nabla\ln\Psi_{\mathrm{T}}(\mathbf{r}))^2/2\delta\tau} \times e^{-\delta\tau(E_{\mathrm{l}}(\mathbf{r})+E_{\mathrm{l}}(\mathbf{r}')-2E_{\mathrm{r}})/2}. \tag{3.29}$$

This approximation tends to the exact Green's function in the limit of $\delta\tau \to 0$ so represents a controlled approximation. Usually the first step of any actual DMC investigation is to determine what timestep to use in order to reduce this error to an acceptable level.

This approximation on its own does not obey the *detailed balance* condition whereby,

$$G(\mathbf{r}',\mathbf{r};\delta\tau)\Psi_{\mathrm{T}}(\mathbf{r}')^2 = G(\mathbf{r},\mathbf{r}';\delta\tau)\Psi_{\mathrm{T}}(\mathbf{r})^2. \tag{3.30}$$

By not satisfying this condition Equation 3.29 provides very poor sampling in regions near potential divergences and where the trial function goes to zero. Therefore detailed balance is re–introduced with an accept/reject step with the following acceptance probability,

$$p = \min\left[1, \frac{G(\mathbf{r}',\mathbf{r};\delta\tau)\Psi_{\mathrm{T}}(\mathbf{r}')^2}{G(\mathbf{r},\mathbf{r}';\delta\tau)\Psi_{\mathrm{T}}(\mathbf{r})^2}\right]. \tag{3.31}$$

3.3.3 The Fermion sign problem

Within the framework of quantum mechanics notions of spin and the Pauli exclusion principle can be shown to arise naturally from relativistic considerations using the Dirac equation. In non–relativistic quantum mechanics based on the Schrödinger equation spin considerations are very much added in as an after thought. Within VMC the use of Slater determinants which, as has already been discussed, enforce the fermion anti-symmetry by design leads automatically to a fermionic solution[2].

In pure DMC however no such symmetry is enforced. Even though the trial function is fermionic DMC merely projects out the lowest non–orthogonal energy eigenstate. This is almost always a bosonic solution. Many ideas have been tried to overcome this problem. One such method is to use walkers of different signs but for all but the smallest systems this approach always ends in sign decoherence and an exponentially decreasing signal–to–noise ratio.

The most successful solution has been the adoption of the *fixed node approximation*. Here the nodal surface, the hypersurface in configuration space where the wavefunction is zero, is enforced from the trial wavefunction. Configuration space is divided up into regions bounded by the nodal surface. Within each of these regions the simulation proceeds as usual but any attempt by a walker to cross the nodal surface is rejected. Indeed for the exact Green's function the drift term should always stop such a move being proposed due to the divergence of the local energy at the trial nodal surface. For the approximate Green's function however such node crossings are occasionally proposed as a result of the short timestep approximation.

[2]The Jastrow factor being positive definite cannot alter the symmetry of the wavefunction.

Obviously this approximation introduces an error into the final energy, only with an exact nodal surface will the exact groundstate energy be recovered. Reynolds *et al.* however showed [85] that the energy obtained is variational with respect to the ground state energy and as such errors made in the energy are expected to be second order to the errors in the nodal surface.

This uncontrolled error although usually small is a source of great frustration to the community. Many schemes have been devised in order to improve the quality of the nodal surface from backflow like schemes where an analogy with classical fluids is made to allow the co-ordinates of a given electron to be partially warped by the co-ordinates of other electrons, thereby moving the nodal surface, to the so called "self–healing" method of Reboredo *et al.* [86]. Here the kinks in the wavefunction that occur next to the fixed–nodes are used as a cost function so as to provide guidance in shifting the nodal surface. It must be noted however that as Troyer and Wiese have outlined [87] finding the exact nodal surface is nondeterministic polynomial (NP) hard putting it in the same class of problems as the infamous "travelling salesman" problem, for which it is conjectured no deterministic algorithm exists that can solve it in polynomial time.

3.3.4 The locality approximation

A further approximation must be made however if one is to use non–local pseudopotentials in the calculation. These assume a potential of the form,

$$\hat{V} = \hat{V}_{\mathrm{loc}}(\mathbf{r}) + \hat{V}_{\mathrm{NL}}(\mathbf{r}) \tag{3.32}$$

where the potential is split into a part local in the angular momentum channels $\hat{V}_{\mathrm{loc}}$ and a part non-local $\hat{V}_{\mathrm{NL}}$. The non–local part of the potential causes the ITSE to become,

$$\frac{\partial f}{\partial t} = \frac{1}{2}\nabla^2 f - \nabla\cdot(f\nabla \ln \Psi) - \left[\frac{(H-E_T)\Psi}{\Psi}\right] f - \left(\frac{V_{\mathrm{NL}}\Phi}{\Phi} - \frac{V_{\mathrm{NL}}\Psi}{\Psi}\right). \tag{3.33}$$

The last term here would represent a non-local branching, which is a problem for a local random walk. It is in fact analogous to the Fermi sign problem that leads us to introduce the fixed nodal surface in DMC. This non-local term is not positive definite and, if left in will cause sign changes in the walkers leading to a rapidly reducing signal–to–noise ratio.

The approximation made is to discard this term recovering the local random walk of all-electron DMC [88]. Obviously in doing this an error is made and it has been shown that this error is second order in the difference between the trial wavefunction and the exact wavefunction. This error means the energy is no longer variational with respect to the ground state.

This significantly changes the importance of the trial wavefunction in DMC. In all-electron DMC the quality of the trial wavefunction, assuming the nodal surface is good enough, merely affects the computational time required to gain a given accuracy. This is not the case with the locality approximation, here the quality of the trial wavefunction directly affects the accuracy of the answer.

Another problem introduced with the locality approximation is that of a spurious attractive potential. This, introduced due to the negative divergence of the local part of the potential at the nodal surface, can cause walkers to get trapped near the nodes. The result of this behavior is an uncontrollable population explosion, destroying the simulation run.

In 2006 Casula proposed a scheme where a second displacement was performed each step determined by a transition matrix [89], constructed from the non-local part of the potential in such a way that it is always positive. This has the advantage of removing the population control problems and restoring the variational property of the energy estimate. The disadvantages however are a much larger timestep error and a breaking of the detailed balance condition. Further work in 2010 also showed the original scheme to be size-inconsistent [90].

3.3.5 The DMC algorithm

With all of the basic framework in place all there is left is to outline the process of actually performing a DMC calculation. A basic algorithm for performing a DMC calculation runs as follows:

1. N configurations each of M electrons are generated usually using VMC.

2. For each configuration n_i:

   (a) For each electron position $\mathbf{r}_j$ in configuration n_i:

      i. A move is proposed $\mathbf{r}'_i = \mathbf{r}_i + \delta\tau\nabla \ln \Psi(\mathbf{r}_i) + \zeta$. Here ζ is a random number drawn from a Gaussian distribution of mean zero and variance $\delta\tau$.

      ii. Reject the move and continue the loop if the electron has attempted to cross a node.

      iii. Accept the move with probability according to Equation 3.31

   (b) $\text{integer}\left[e^{-\delta\tau(E_\text{l}(\mathbf{r})+E_\text{l}(\mathbf{r}')-2E_\text{R})/2} + \eta\right]$ copies of the configuration are then made, where η is a random number of uniform distribution, $0 \leq \eta \leq 1$.

   (c) The new value of E_l (or any other required observable) for the configuration n_i is accumulated.

3. E_R is then set to reflect the change in the configurations. It is also used to control the population in relation to the target of N walkers. Several schemes are possible for this.

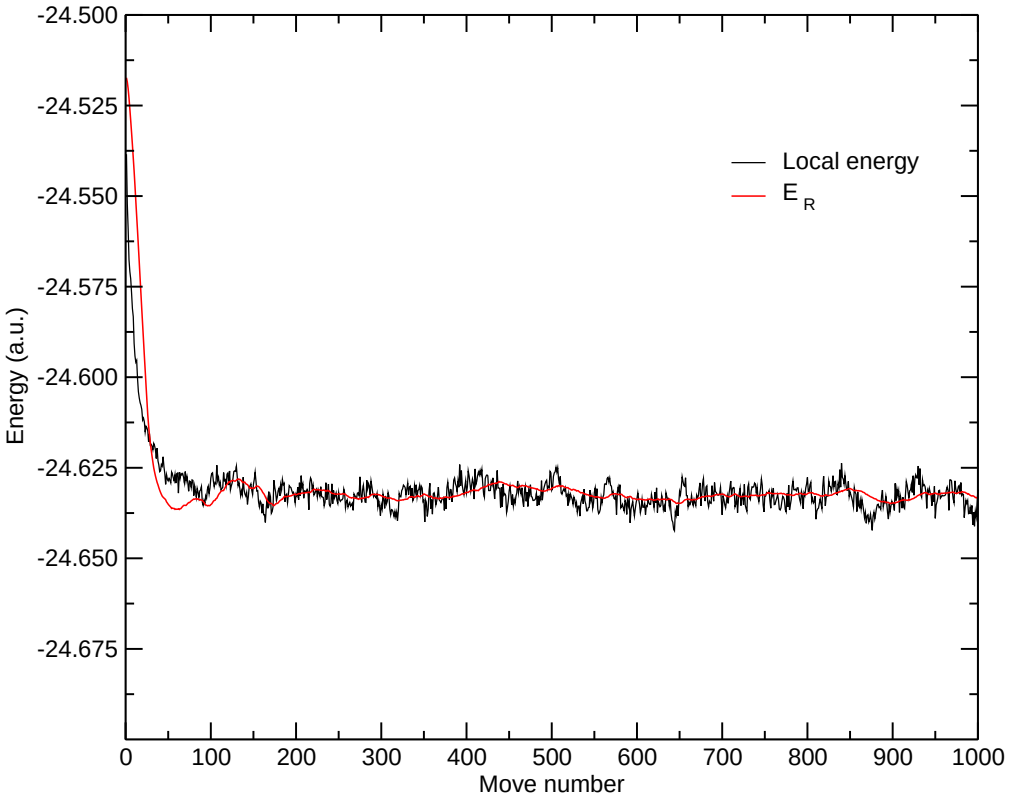
As with VMC a period of equilibration is needed initially. This is in order to project out the groundstate and dampen any excited states present in the trial function. Figure 3.1 shows a sample run where the first 200–300 steps show the characteristic equilibration behaviour. After that the local energy and E_R settle down. Usually during the equilibration phase E_R is updated more often than in the later statistics accumulation phase where the updates are usually only every few hundred steps.

3.4 Finite size effects

Just as the enforcing of periodic boundary conditions in single–particle theory introduces hidden complexity the same is true in many–body methods. The use of periodic boundary conditions is quite advanced within QMC methods and proves a very useful method of obtaining data on extended systems. The hidden complexities it introduces as a price for this can be broadly divided into two categories.

Figure 3.1: Example DMC run. This graph shows the initial part of a DMC calculation on crystalline LiF using a 4×4×4 supercell

.



3.4.1 Single–body finite–size effects.

The so–called *single–body finite–size effects* are analogous to the finite–size effects in single particle theory. This is essentially due to incomplete sampling of the Brillouin zone. Recall from Section 2.2.4.1 that in applying Bloch's theorem to an infinite crystal a summation should be made over all Bloch states or k–vectors. In single particle theory this is done by k–point sampling which approximates this as a discrete sum. In QMC theory the orbitals used are usually only calculated at a single k–point. This can introduce a significant bias. Enlarging the supercell being simulated to N×N×N primitive cells is equivalent to using an N×N×N k–point mesh. Hence the bias can be removed in this fashion. A less expensive way is by performing a correction calculation using single particle theory. Perhaps the most common way of doing this is via the following formula [91]:

$$E_{\infty}^{\mathrm{QMC}} = E_{N}^{\mathrm{QMC}} + a(E_{\infty}^{\mathrm{LDA}} - E_{N}^{\mathrm{LDA}}) \,, \tag{3.34}$$

where E_N^{QMC} is the QMC energy using a N×N×N supercell, E_N^{LDA} is the LDA DFT energy of the primitive cell using a N×N×N k–point mesh, E_∞^{LDA} is the DFT energy of the cell with perfect k-point sampling, E_∞^{QMC} is the corrected DMC energy, and a is some fitting parameter. The parameter a is often

taken to be unity and should only affect the rate of convergence as $\lim_{N\to\infty}(E_\infty^{\mathrm{LDA}} - E_N^{\mathrm{LDA}}) = 0$.

3.4.2 Many–body finite–size effects

The so–called *many–body* finite–size effects have no real analogy in single–particle theory except perhaps for the case of non–periodic systems being treated in a periodic fashion. In many–body theories the individual electrons interact with an infinite number of their counterparts via the Coulomb interaction. The sum over these periodic Coulomb interactions is only conditionally convergent, indeed naïve summation of these interactions does not necessarily lead to a Hamiltonian with the correct periodicity [76].

The Ewald interaction is usually adopted in this summation which can be variously thought of as imposing the correct periodicity on the Hamiltonian or imposing a boundary condition on the surface of the infinite crystal of no net surface polarization [74]. While it is convergent and unique at short ranges however the Ewald scheme deviates from the Coulomb potential of $1/r$ with terms involving higher powers of r. As would be expected these terms are inversely proportional to the volume of the supercell so in the limit of an infinite–size cell the correct behavior is recovered.

This deviation however is a manifestation of the spurious self–correlation that has been introduced when imposing PBCs. Simply put, in this system contributions arise from having two classes of electron interaction. When an electron is interacting with another electron they are both free to adjust their inter–particle distance (to some degree) to minimize the overall energy, this is the same interaction that occurs in real crystals. An electron in the PBC system however can also interact with its own periodic image. In this case the inter–particle distance is fixed by the PBCs. Indeed the electron can be thought of as being perfectly correlated with its image. In effect this spurious perfect correlation lowers the energy of the system. To leading order, with full 3D PBCs, this has been shown to be of the form [74],

$$E(N) = E_\infty + \frac{b}{N}. \tag{3.35}$$

Here $E(N)$ is the energy per fundamental unity using an N particle supercell, E_∞ is the true energy per fundamental unit and b is a system dependent parameter. In the case of 2D systems Drummond *et al.* [92] find the following form,

$$E(N) = E_\infty + \frac{b}{N^{\frac{5}{4}}}. \tag{3.36}$$

It is therefore possible to again remove this effect through extrapolation though this is costly. Hence this work makes use of two other methods for reducing this effect.

3.4.2.1 The Kwee–Zhang–Krakauer correction

The first is the correction of Kwee, Zhang and Krakauer (KZK) which can be applied to bulk calculations [93]. This corrects for both single–particle and many–body errors using the following formula,

$$E_\infty = E_N^{\mathrm{DMC}} + E_\infty^{\mathrm{LDA}} - E_N^{\mathrm{KZK}}\ . \tag{3.37}$$

Here, $E_{\mathrm{cell}}^{\mathrm{KZK}}$ is a DFT-like energy of the cell (no k-point sampling), which uses a volume dependent functional designed to mimic the sum of single-particle and many-body errors, while E_∞^{LDA} is the DFT energy of the cell with perfect k-point sampling using the LDA. The functional is constructed by parametrization of a series of calculations on the HEG. It seems to be an effective method [94, 95] but Drummond

et al. [92] point out that by being based on the LDA ultimately it is limited to cases where the LDA is already a reasonable approximation.

3.4.2.2 The model periodic Coulomb interaction

The Model Periodic Coulomb (MPC) interaction of Williamson *et al.* [96] offers an alternative to the Ewald interaction. It is formulated by separating the interacting electron term into the Hartree interaction, familiar from Section 2.2.2 and the interaction between the exchange–correlation hole density and the electron density.

The exchange–correlation hole density can be thought of as the change in the electron density in the vicinity of a given point in space caused by an electron being present at that point. The presence of the electron reduces the electron density around it due to the exchange and correlation interactions. This change can be thought of as a negative electron density (or a positive *hole* density) called the exchange–correlation hole density.

The Ewald interaction results in a correct treatment of the Hartree term with respect to the true extended system so this must be retained in any new interaction. It is in the interaction with the exchange–correlation hole that the Ewald interaction fails to reproduce the $1/r$ behavior.

The MPC interaction is constructed so that it retains the Ewald character with the Hartree term but has the correct r dependence for the exchange–correlation term. In this way it removes a large amount of the finite–size effects. However in formulating it in terms of the exact charge density it requires knowledge of this quantity *a priori*. But that quantity is, indirectly, what is being calculated in the QMC simulations. An approximation is therefore introduced where the charge density from the trial function is used instead. This creates an error that is second order in the error in the assumed charge density [96]. For most systems studied so far however this error is small.

Drummond *et al.* [92] also point out that when used in the walker propagation in DMC the MPC interaction, while representing the interaction with the exchange–correlation hole better than the Ewald interaction, actually produces a worse exchange–correlation hole. Therefore the correct approach is to use the Ewald interaction in the propagation of a DMC run but use the MPC interaction when calculating the local energy for the statistics accumulation. When ever the MPC correction is referred to in this work this implementation is used.

3.5 Scaling

As has already been outlined the main advantage of these Monte–Carlo methods is their scaling, both in respect to number of particles and number of processors.

3.5.1 Particles

Both VMC and DMC can be shown to scale as N^3, where N is the number of particles. This is because the overall time taken to update a given walker is given by the following relation,

$$T \propto \alpha N^2 + \beta N^3 + \gamma N^3. \tag{3.38}$$

Here the term in α is due to evaluating the electron-electron, electron-ion interactions, the term in

β updating the determinant value and γ evaluating the Slater orbitals expanded in plan–waves for each electron in a given configuration. The different "linear scaling" methods of Williamson *et al.* [97] and Alfè and Gillan [98] can reduce this scaling to $\sim N^2$, by making the scaling of the orbital evaluation $O(N)$. Because β tends to be a very small prefactor the overall scaling then tends to be dominated by the $O(N^2)$ terms.

Willamson *et al.* achieve this by re–expanding the orbitals in truncated maximally localized Wannier (MLW) functions. Here the quality of the expansion and as a result the accuracy of the final QMC energy is determined by the magnitude of the cut–off beyond which the MLWs are held to be zero.

Alfè and Gillan use what would seem to be a more elegant solution. Here the plane–waves are re–expanded in terms of B–spline functions (blips). These blips were initially envisaged for linear–scaling DFT applications [99] and consist of a real space grid of spacing a made up of grid centered cubic spline functions. These are defined as zero outside of a radius $2a$ centred on their respective grid points. This ensures that for any point in real space only 64 blips contribute to the wavefunction, regardless of the system size. The spacing is determined by,

$$a = \pi / n k'_{\max} \tag{3.39}$$

Here $k_{\max}$ is the length of the largest plane–wave vector in the original basis–set. In practice n is generally unity or two which results in aliasing errors introduced by the re–expansion begin small enough to be neglected. It can be shown however that in the limit of $n \to \infty$ the initial wavefunction is recovered.

3.5.1.1 DMC and the exponential scaling problem

It must be said however that this discussion of scaling has only taken into account the CPU time required to perform one move in configuration space. This does not necessarily specify the computational scaling with respect to system size *for a given statistical error*. In the case of DMC this turns out to be a crucial distinction.

Nemec [100] has shown that the branching/death process in DMC introduces an exponential scaling with respect to system size for a given statistical error. This is due to the branching effect causing the amount of statistical correlation in the configuration population to vary with time. Nemec proposes an effective configuration population $N_{\rm pop}^{\rm eff}$ which is the number of walkers at a given time that are needed to represent the energy to within the given variance if there were no serial correlation. He finds that for a given population of walkers its corresponding $N_{\rm pop}^{\rm eff}$ decays exponentially with respect to system size showing the exponential increase in the level of statistical correlation within the population.

It is clear that the prefactor involved is also strongly system dependant. Nemec introduces a measure of statistical efficiency and posits that when this exceeds two this exponential scaling should become apparent. For He atoms for instance this happens for system sizes above 2700 atoms but for C_2H_4 this occurs with only 38 molecules. He also notes that this is proportional to the variance of the trial wavefunction, another reason for careful optimization.

3.5.2 Processors

At the time most of the calculations in this work were performed the most powerful computer in the world (based in the Linpack score) was the Jaguarpf system[3] at the Oak Ridge National Laboratory, in the United States. This consists of 18,680 compute nodes each housing 12 cores. It is clear that one of the most important characteristics of any modern electronic structure method is its ability to scale over many cores.

In the case of VMC this scaling is what is usually referred to as *embarrassingly parallel*. When run in a parallel environment each core can host a separate VMC random walk with no communication. At the end of the simulation all of these separate runs can be gathered together and their results averaged as if it were one longer run. The only possible cause for concern in this prediction of it being an infinitely scalable algorithm is that it requires a good enough random number generator to ensure there is no correlation in the random number sequences generated on different cores. Otherwise a bias would be introduced into the final energy.

The DMC algorithm is however more complicated. Interprocess communication is required for population control i.e. setting E_{R}. This is generally quite low however. The main amount of communication comes from the need to redistribute the configurations between the processors. This arises because the algorithm that is generally used has a walker population that can fluctuate. Thus some processors end up having more walkers than others leading to poor load balancing. In order to rectify this after a certain number of moves walkers are shifted between nodes in order to balance out the walker distribution. This in practice tends to lead to an upper limit on the number of nodes one can use before this inter-process communication leads to inefficient runs. Recent work [101] seems to indicate that even with runs on 100,000 cores this effect can be negated.

[3] Unsurprisingly mere months before this thesis was completed Jaguarpf was unseated and now a GPU based machine, Tianhe-1A is the most powerful, hopefully until the ink is at least dry on this page.

Chapter 4

A Study of Ionic Surface Formation Energies

The aim of this thesis is to show the possibility of obtaining important properties relating to surface energetics to beyond chemical accuracy. This chapter is concerned with the calculation of the surface formation energy, introduced in Section 1.2, for a variety of ionic solids.

The first section presents the calculation of σ using DFT. As already outlined current scattered evidence strongly suggests that in order even to obtain qualitatively accurate values it is necessary to go beyond DFT. The DFT work presented here confirms this, showing a large dependence on the xc functional adopted. Presented alongside these data are DFT calculations of the bulk properties of the systems, a useful cross–check showing again the expected strong exchange–correlation functional dependence. Surface relaxation effects are also calculated which show a far weaker exchange–correlation functional dependence.

In an attempt to go beyond DFT and obtain high accuracy results a careful study of the LiH (001) surface formation energy using QMC methods is then presented. This includes calculations using the more traditional pseudopotential approximation within QMC and, as far as we are aware, a completely novel all–electron study, the first of its kind for the calculation of surface formation energies outside jellium within DMC. This shows the importance of core–effects and pseudopotential errors in the case of LiH (001) changing the result by $0.07(1)\,\mathrm{J\,m^{-2}}$.

The group has a long standing collaboration with the Manby group at the University of Bristol and they were able to calculate the correlation contribution to the surface energy using high level quantum chemistry and the hierarchical method. A detailed set of all–electron Hartree–Fock calculations on both the bulk and the surface have been performed to obtain the Hartree–Fock contribution. Special attention had to be paid to the difficult task of obtaining basis–set convergence for the crystalline system. The comparison with quantum chemistry and DMC yields spectacular agreement for both the bulk and surface allowing the absolute performance of the DFT functionals to be judged.

Finally so as to explore the applicability of the method pseudopotential calculations of LiF (001) using DMC are presented. In this case there are two competing experimental predictions, the DMC results favor the method of Burns and Webb outlined in Section 1.2.2. As before comparison with quantum chemistry is obtained with the correlation calculations from the Manby group begin combined with accurate Hartree–Fock calculations. For this it was necessary to optimize three sets of basis functions in

order to allow the Hartree–Fock complete basis–set limit to be approached systematically. The comparison again shows excellent agreement between quantum chemistry and DMC reinforcing the accuracy of both.

4.1 DFT Calculations

The main aim of the DFT work is to discover whether there is any exchange–correlation functional dependence when calculating the surface formation energies of some simple ionic systems. The systems chosen are the (001) planes of LiH, LiF, NaF and MgO. This is due to a balance of complexity, and interest. In terms of complexity LiH, LiF, NaF and MgO represent a useful progression in terms of computational effort allowing any computational problems to be discovered early on with small systems. It also allows more interesting materials such as MgO to be studied which provides a gateway into the metal–oxides.

It is also clear that in order to gain a full understanding of the strengths and weaknesses of the various xc functionals it is necessary to look at how they perform in the treatment of the bulk systems, this also provides useful convergence testing that can be taken into the surface calculations.

The overall schema is that exhaustive bulk calculations first be performed for each functional. The the surface formation energies can be calculated using the equilibrium lattice parameters specific to the particular functional/system combination. This removes any bias that may occur through a functional being in error on the bulk but being relatively accurate in its treatment of the surface.

4.1.1 Exchange–correlation functionals

The functionals used in this study are the usual LDA along with the PBE, PW91, RPBE and Wu–Cohen GGAs. The inclusion of the LDA is not just for mere historical interest. As outlined in Section 1.2.4.2 scattered evidence in the literature points toward the LDA providing the most accurate estimation of σ within the DFT method.

Of the GGA based functionals PBE is perhaps the most widely used with its fore–runner the PW91 functional being a close second. Again the scattered evidence already outlined suggests that this functional consistently under predicts σ by $\sim$30%. Therefore in order to verify this trend it is important that these functionals be included in the test set.

With PBE expected to be in error it is also worthwhile to include functionals that attempt to offer corrections to it. The first of these is the RPBE. This comprises of an alteration to the PBE functional by Hammer *et al.* [37]. This alteration improves both the molecular atomization energies predicted by the functional, and more importantly, surface chemisorption energies. While this might suggest a better description of the surface energy than other functionals this is not at all clear hence the interest in including it in the test set.

4.1.1.1 The Wu–Cohen functional

The final functional is the recent Wu and Cohen modification [102] of the PBE functional. This modification was made with bulk systems in mind. Wu and Cohen observe that while PBE and its predecessor PW91 impose a sharp cut–off of the exchange–correlation hole in real space the exchange–correlation

hole in solids actually has a diffuse tail.

So they firstly adopt the PBE *ansatz* for the exchange energy,

$$E_{\mathrm{X}} = \int n\epsilon_X(n)F_x(s)dr^3. \tag{4.1}$$

where E_{X} is the energy due to exchange, n is the electron-density, $\epsilon_X(n)$ is the parameterized energy of the free electron gas,s is a reduced gradient of n ($s = |\nabla n|/[2(3\pi^2)^{1/3}n^{4/3}]$) and $F_X(s)$ is termed the exchange enhancement factor.

In PBE F_{X} is given by,

$$F_{\mathrm{X}} = 1 + \kappa - \kappa(1 + x/\kappa) \tag{4.2}$$

where $x = \mu s$ with μ and κ numerical constants used to make the function satisfy certain conditions. By modifying x Wu and Cohen were able to impose known gradient expansions for small s while recovering the successful PBE F_{X} for large s. This modification is given by,

$$x = \frac{10}{81}s^2 + (\mu - \frac{10}{81})s^2 \exp^{-s^2} + \ln(1 + cs^4) \tag{4.3}$$

with c another numerical constant.

With this Wu and Cohen were able to significantly improve on the accuracy of the predicted lattice parameters and bulk moduli for a number of crystal systems. Also their predicted jellium surface energies were significantly better than either the LDA or the PBE functionals. This makes the WC functional a good candidate for producing accurate surface energies.

Implementation

Unlike the other functionals in the test set WC is not implemented in the standard VASP code that was used for the DFT study. As explained WC only differs from PBE in the exchange enhancement factor. Therefore in order to implement WC within VASP all that was done was to modify the existing PBE routines with this new exchange enhancement factor. This was tested both analytically, by confirming the forms of the various derivatives used in the code were indeed correct and then numerically using some test systems.

At the time no other PAW code was available that had implemented WC. Hence results supplied by Cazorla which had been obtained using Wien2k FLAPW implementation were used as a test set. This consisted of equilibrium volumes of both Niobium and Technetium crystals obtained by fitting to a Birch–Murnaghan curve [103]. Repeating these calculations in VASP produced the same results indicating the code was working correctly. This implementation was then later used in work on the phase-diagram of Molybdenum [104].

4.1.2 Bulk Calculations

As described in Chapter 2 periodic DFT calculations are dependent on several parameters. Apart from the xc functional, these consist of:

- The one–electron basis–set used in the expansion of the Kohn–Sham orbitals.

- The number and location in Fourier space of the k-points that are summed over in order to include all of the effects of the infinite crystal.
- Also within the PAW formalism there is the choice of PAW data set to be used for the core regions of each atom.

The code adopted for the calculations is the Vienna *ab initio* Simulations Package (VASP) [105, 106, 107, 108] includes with it PAW datasets for all of the elements studied in this work, constructed with the LDA, the PBE GGA and the PW91 GGA. These datasets, constructed by Kresse and Joubert [63, 109] have been widely used within the community and as such can be said to provide a transferable approximation to all-electron calculations, reproducing dimer bond lengths to within a few thousandths of an Ångström in some cases [109]. Table 4.1 shows the valence states for each dataset.

Table 4.1: Electron states defined as valence in the PAW datasets.

Element	Valence states
H	1s
Li	1s2s
O	2s2p
F	2s2p
Na	2p3s
Mg	2p3s

4.1.2.1 Convergence

In terms of basis–set convergence the VASP PAW datasets prove to be very convenient. As discussed in Section 2.2.4.2 the basis–set is systematically improved by increasing the energy cut–off. This is the highest allowed energy for a given plane–wave in the one–electron basis, and by increasing this the number of plane–waves in the calculation are also increased providing a better quality basis. The data sets each have optimal plane–wave cut–offs above which the total energy should be well converged. This was found to be the case and all of the initial testing for each system is done using this supplied values (Table 4.2 lists these) and then the final calculations of specific properties are also performed using a 30% increase to demonstrate the level of convergence obtained. Figure 4.2 shows the minimal difference between the recommended cut–off and the 30% increase.

For the k–point sampling $N \times N \times N$ Γ–centred Monkhorst–Pack grids were used. These are straightforward to converge as the dimension of the grid, N can be increased which should lead to convergence of the total energy of the primitive cell. Figure 4.1 shows this behavior for LiH. The overall target is total energies converged to better than 0.2 meV, Table 4.2 shows the final grids used and their convergence level for each system.

It is also worth noting that VASP also allows the grid–spacing used in its fast Fourier transform (FFT) of the wavefunction to be defined by the user. If these grids are not fine enough errors can occur.

Table 4.2: DFT energy cut–offs and k–point grids used with their convergence criteria.

System	Energy cut–off / eV	k–point grid	k–point convergence / meV
LiH	270.99	7×7×7	0.2
LiF	400.00	6×6×6	0.01
NaF	400.00	7×7×7	0.01
MgO	400.00	8×8×8	0.02

It can be shown that the FFT mesh must contain all of the points corresponding to plane–waves up to twice the plane–wave cut–off to avoid these errors [110]. All of the calculations here satisfy that requirement.

4.1.2.2 Production calculations

Satisfied that convergence is obtained with respect to the necessary parameters, the production runs consist of total energy calculations for nine different lattice parameters in the vicinity of the experimental lattice parameter. The spacing between lattice parameters was 0.1 Å which provides a good description of the energy–volume, $E(V)$ curve. A third order Birch–Murnaghan equation of state [111] as shown in Equation 4.4 was then fitted to the data in order to provide an estimate of the equilibrium lattice parameter and the bulk modulus of the material.

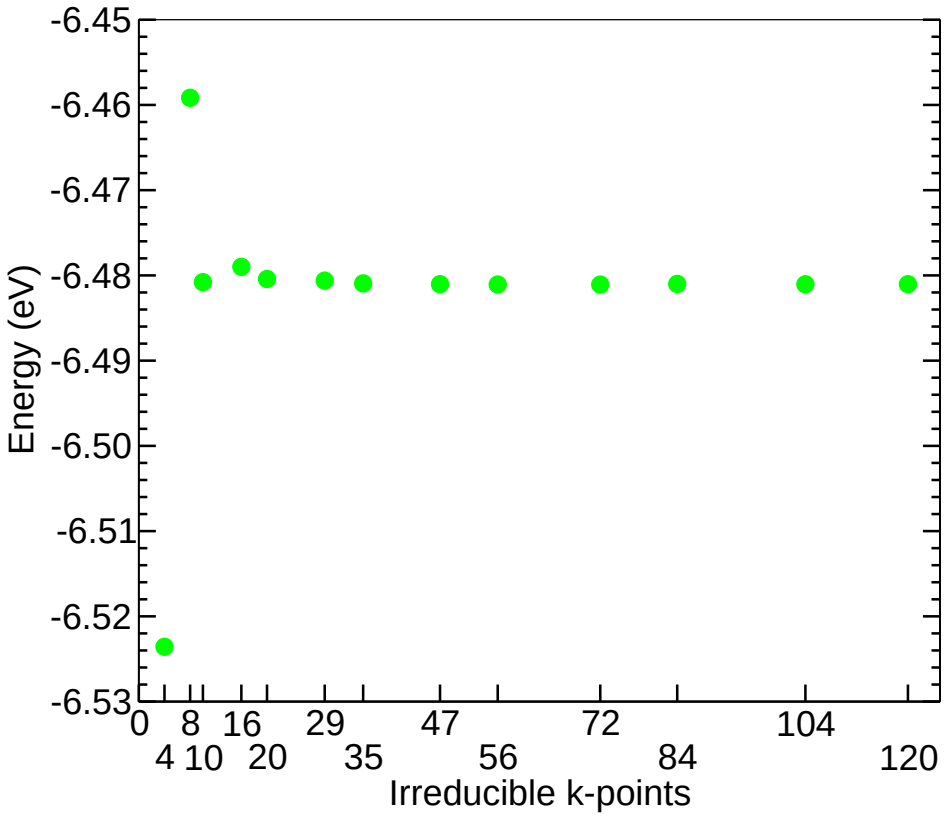
$$E(V) = E_0 - \frac{9V_0B_0}{16}\left((4 - B_0')\frac{V_0^3}{V^2} - (14 - 3B_0')\frac{V_0^{7/3}}{V^{4/3}} + (16 - 3B_0')\frac{V_0^{5/3}}{V^{2/3}}\right) \tag{4.4}$$

Here E_0, V_0 and B_0 represent the equilibrium energy, volume and bulk modulus, respectively. B_0' is the first derivative of the bulk modulus.

Table 4.3 shows the equilibrium lattice parameter and bulk modulus that results form these calculations. It is clear that there is a definite trend with respect to the functionals. It generally follows that the pattern obtained is LDA–WC–PW91–PBE–RPBE, going from over–binding to under–binding, and the same trend goes when applied to the bulk moduli, from over–estimates to under–estimates. The range of predicted lattice parameters with the different functionals cover from between 2–4% of the experimental values. For the bulk modulus the figure is closer to 20%. While these bulk parameters are not catastrophically bad they do show a noticeable exchange–correlation functional dependence for the relatively simple environment of the perfect bulk crystal. They do not give confidence for obtaining chemical accuracy for the surface case.

It must be pointed out that as such none of these calculations contain zero-point energy corrections or take into account thermal effects. Therefore their comparison with experiment can only be very limited. We cannot see from these values alone which functional is the most accurate, only that at least some of the functionals must be in error.

Figure 4.1: Total energy vs. k–points for bulk LiH.



4.1.2.3 Zero–point corrections

In order to be able to compare with experiment it is necessary to calculate the correction due to the lattice vibrations.

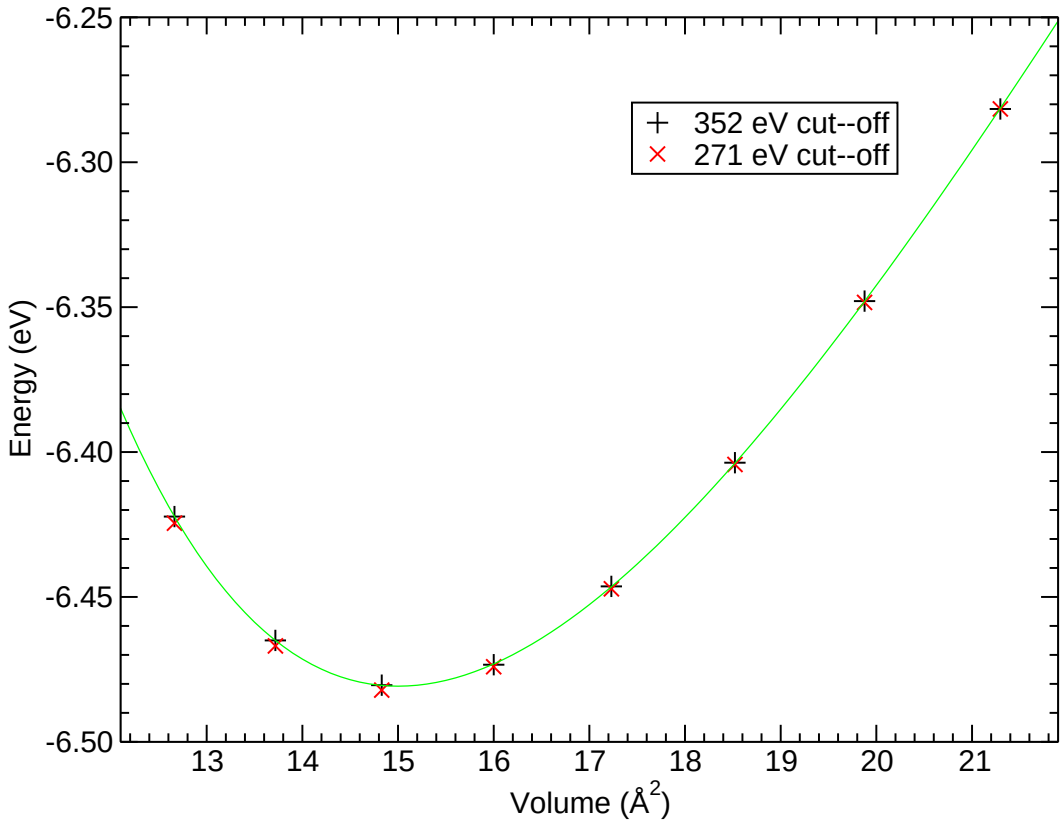
In the preceding DFT calculations the nuclei have been held fixed in space to simplify the calculations. Obviously in the real system not only are the nuclei free to move but they move in a quantum manner. In calculating the leading order correction however it is sufficient to treat them as classical particles.

In this lattice of classical nuclei the nuclei can be thought of as vibrating about a mean position, the allowed lattice vibrations are termed *phonons*. The lowest energy phonon state corresponds to the crystal at zero temperature and it is this correction that is termed the *zero–point energy*.

The method adopted in this work for calculating phonons utilizes the density functional perturbation theory (DFTPT) of Baroni *et al.* [116]. Here the forces producing the phonons are calculated from the first–order variation of the electron density.

This is implemented in the QUANTUM ESPRESSO package, and uses the ground–state wavefunction generated by the plane–wave PWscf code. The PBE GGA is adopted for these calculations and for ionic solids this is estimated to produce phonon frequencies accurate to 5% [117].

Figure 4.2: Two $E(V)$ plots for LDA LiH showing the effect of the higher energy cut–off. A Birch–Murnaghan fit to the data at 352 eV is also shown.



LiH

For the ground–state calculations the Li.pbe-n-van and H.pbe-van_ak pseudopotentials of the PWscf distribution were used. An energy cut–off of 1300 eV and a 4×4×4 k–point grid were enough to converge the calculated zero–point energies to 0.1 meV. The zero point energy for each volume of the E(V) curve was calculated and then added to the existing DFT fixed ion calculations. The Birch–Murnaghan curve was then fitted to these corrected values to obtain the corrected bulk parameters.

LiF

Here the F.pbe-n-van pseudopotential is used again from the PWscf distribution. This time an energy cut–off of 822 eV and a k–point grid of 4×4×4 is sufficient to obtain sub meV convergence of the zero–point energy.

Table 4.4 shows the results of these calculations along with cohesive energies calculated by taking the difference between the E_0 term in Equation 4.4 with the appropriate DFT energies of the isolated atomic species in a periodic supercell of sides 20 Å. Comparing with the 0K experimental reference data it is now clear that no one functional replicates the experimental results with any consistency.

Table 4.3: DFT predicted equilibrium lattice constant and bulk modulus for the ionic systems using different xc functionals. The experimental data quoted is measured at 293K and therefore cannot be directly compared with the raw DFT data.

	LiH		LiF		NaF		MgO	
	a_0 / Å	B_0 / GPa	a_0 / Å	B_0 / GPa	a_0 / Å	B_0 / GPa	a_0 / Å	B_0 / GPa
LDA	3.9510	40.5	3.9169	86.6	4.5101	61.3	4.1690	172.3
WC	3.9834	37.0	4.0329	69.6	4.6615	46.2	4.2356	155.9
PW91	4.0215	36.4	4.0978	67.0	4.7014	44.7	4.2543	151.3
PBE	4.0232	36.5	4.0932	67.6	4.7062	44.9	4.2585	150.4
RPBE	4.0302	34.8	4.1445	60.8	4.8181	37.4	4.2954	142.2
Expt	4.0834	22.8–34.7[a]	4.0262	67.1[b]	4.6234	46.4[c]	4.2117	162.5[d]

[a] Ref [112]

[b] Ref [113]

[c] Ref [114]

[d] Ref [115]

4.1.3 Surface Calculations

Briefly recalling from Section 1.2.3 there is a choice as to how to calculate the surface formation energy using a periodic code. For the reasons of accuracy previously outlined we adopt the slab method of Fiorentini and Methfessel. As this involves merely calculating the total energies of a number of slabs of increasing thickness and fitting to Equation 1.3 (reproduced here for convenience) no further bulk calculations are required.

$$E_{\mathrm{slab}}(N) = 2A\sigma + NE_{\mathrm{bulk}}. \tag{1.3}$$

For the slab calculations we are able to assume the same k-point grid as in the bulk (apart form the fact that due to the slab nature there is only one point in the z-direction). Hence for LiH a k-point grid of $7 \times 7 \times 1$ is used. The system is however still periodic in the z–direction, this means that to avoid the slab interacting with its periodic image in the z–direction we must increase the vacuum gap at the top of the super-cell until the total energy converges. The convergence demanded is better than 2 meV per formula unit (f.u.) and for LiH this results in a vacuum gap of $2a_0$, where a_0 is the equilibrium lattice parameter. The resulting values of σ are shown in Table 4.5.

In the cases shown it was necessary to discard the first 2–3 layers in order to achieve convergence. Also while the calculations went up to slabs of 15 layers thick it was found little extra was gained by using slabs containing more than 7 layers. All of this is demonstrated in Figure 4.3. The first three calculations show finite size errors that skew the result (and significantly worsen the goodness of fit). After this point the values plateau until layers 1-8 have been discarded, at which point noise is introduced by extrapolating over a large distance back to zero using only quite thick layers.

Table 4.4: Zero–point corrected DFT bulk parameters including cohesive energy. The zero–point corrections in all cases are calculated using PBE and the linear response method as outlined in the text. The experimental values are all at 0K unless otherwise stated.

	LiH			LiF		
	a_0 / Å	B_0 / GPa	E_{coh} / eV	a_0 / Å	B_0 / GPa	E_{coh} / eV
LDA	3.9993	35.6	−5.1231	3.9522	80.5	−9.7285
WC	4.0720	32.0	−4.4390	4.0630	67.2	−9.2579
PW91	4.1098	31.2	−4.5296	4.1431	57.4	−8.8538
PBE	4.1108	31.5	−4.4600	4.1379	58.2	−8.7536
RPBE	4.1217	30.0	−4.3395	4.1962	51.1	−8.4523
Expt	4.062	33–38	−4.778,−4.759[a]	4.005[b]	67.1[c]	−8.800(14)[d]

[a] Ref [36].

[b] Ref [118]. Constant between 40-60K.

[c] Ref [119]. Measured at 4.2K.

[d] Calculated from thermochemical cycles. See Appendix B for details.

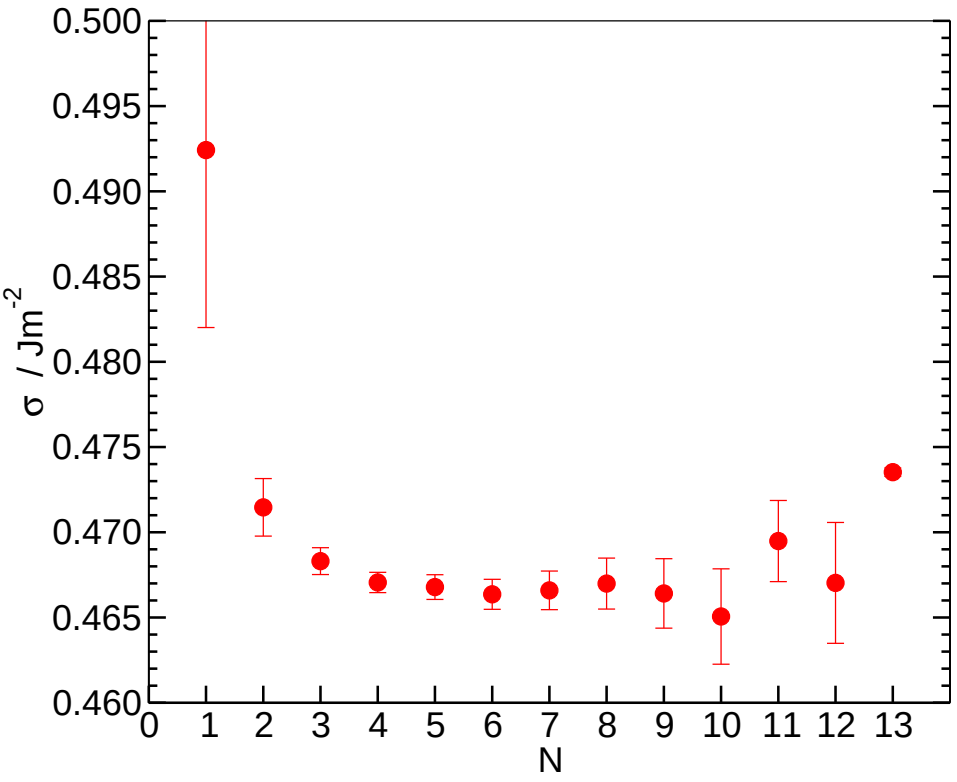
As before the calculations are initially done at the plane–wave cut–off suggested by the pseudopotentials then again with a cut–off 30% higher. This leads to surface energies which are converged to within 0.002 J m^{-2}. Table 4.5 again shows a definite trend LDA–WC–PW91–PBE–RPBE from high to low estimations of the surface formation energy. This time there exists little experimental evidence with which to compare. Only LiF and MgO are represented experimentally. Figures of 0.34 J m^{-2}from Gilman [16] and 0.48(5) from Burns and Webb [23, 24] are given for LiF and for MgO the range is 1.04-1.2 J m^{-2} [26] none of which being good enough to discriminate between the DFT values. There is no information as to which functional is the most accurate.

Even more it is not possible to know if the functionals are consistently out by a given amount across all of these systems. What can be seen from this however is that all of the functionals tend to agree on the impact of surface relaxation upon the surface formation energy. This effect is obviously quite low for these systems, justifying the decision to use only the bulk terminated structures in the QMC work.

Table 4.5: DFT surface energies for each system as calculated with the different xc approximations. Results for both bulk–terminated and DFT relaxed geometries are presented. The units used are J m^{-2}.

System		LDA	PW91	PBE	RPBE	Wu-Cohen
LiH	Relaxed	0.4586	0.3342	0.3273	0.2598	0.4024
	Bulk–term.	0.4663	0.3435	0.3365	0.2723	0.4105
LiF	Relaxed	0.4954	0.3235	0.3129	0.2320	0.3494
	Bulk–term.	0.5123	0.3415	0.3314	0.2534	0.3687
NaF	Relaxed	0.4012	0.2705	0.2719	0.2238	0.2707
	Bulk–term.	0.4016	0.2723	0.2733	0.2254	0.2721
MgO	Relaxed	1.1662	0.9051	0.8920	0.7592	1.0100
	Bulk–term.	1.1807	0.9233	0.9107	0.7807	1.0276

Figure 4.3: LiH σ calculated for N - 15 layers with LDA. Error–bars shown are that calculated from the goodness of fit.



4.2 LiH with DMC and the pseudopotential approach

The lack of any definitive answers within the DFT framework means it is clear that a different method has to be used in order to obtain an accurate estimate for σ. As discussed in Section 3, DMC offers a way to obtain an estimate to the ground state energy of a given system. Periodic boundary conditions such as bulk crystal and slabs can be incorporated. Most importantly by not relying upon the uncontrolled approximation of a exchange–correlation functional DMC should be able to provide a much more accurate estimate than DFT.

In order to test this assumption LiH, the simplest system from the DFT study, was adopted for the first part of the DMC work. In the bulk primitive cell LiH is a problem of only four electrons, by adopting pseudopotentials this can be reduced to a problem of two electrons. This was the approach taken in the first instance.

As with the DFT work in order to help gauge the success or otherwise of an estimate of σ it is useful to have the corresponding estimates for the bulk parameters hence we first present a DMC bulk study followed by the σ calculations.

4.2.1 Pseudopotential choice

Since the pseudopotential method is being used the first point to consider is the choice of pseudopotentials.

The Trail and Needs pseudopotentials [120, 121] offer a set of well tested norm–conserving pseudopotentials generated using the Dirac–Fock formalism. The 1s state of the lithium atom is pseudized reducing the lithium atom to a one electron problem. The divergent Coulomb potential of the hydrogen atom is smoothed allowing the trial–wavefunctions to be generated using a manageable plane–wave basis.

By calculating the bulk properties of the system first the pseudopotential transferability can be tested in the regime of the many–body method.

4.2.1.1 Trial Wavefunction

The trial wavefunction is of course generated using the same pseudopotentials and is the usual Slater–Jastrow kind as discussed in Section 3.2.1. The Slater part is generated using the plane–wave DFT code PWscf code found in the QUANTUM ESPRESSO package [122]. This allows the use of the LDA functional within the plane–wave formalism.

The rational behind this was three–fold: Firstly previous studies have tended to indicate that the LDA approximation within DFT produces the best nodal surface and overall trial–wavefunction for DMC calculations [123, 124]. Secondly the use of a plane–wave basis allows a high level of basis–set convergence that is unaffected by the use of PBCs. This is an important consideration as even though the final DMC energy should be independent of the quality of trial wavefunction the adoption of the locality approximation, as outlined in Section 3.3.4, introduces an error proportional to the square of the error in the trial function. Finally as discussed in Section 3.5 plane–waves can be easily converted into blip functions which offer good computational efficiency for the DMC calculations.

Table 4.6: VMC optimized energies for bulk LiH using a 2×2×2 supercell with different Jastrow factors.

Expansion order				VMC Energy / E_h	VMC Variance / $E_\mathrm{h}{}^2$
u	χ	f	p		
2	2	-	-	−7.142(1)	0.104
2	2	2	-	−7.142(1)	0.104
2	2	3	-	−7.147(2)	0.145
2	2	3	4	−7.148(1)	0.080
2	2	3	6	−7.147(1)	0.082
2	2	4	-	−7.149(1)	0.090
4	2	-	-	−7.142(1)	0.101
4	2	3	-	−7.147(1)	0.090
4	2	4	-	−7.145(1)	0.098
4	4	-	-	−7.138(1)	0.106
5	2	-	-	−7.142(1)	0.101

The plane–wave cut-off used for all of the pseudopotential LiH calculations was 300 Ryd. This provided well converged DFT total energies suggesting well–converged wavefunctions. The form of the Jastrow factor was optimized individually for both the bulk and the surface cases.

4.2.2 Bulk properties

Calculating the bulk properties of a crystalline system using DMC is very similar to calculating bulk properties with DFT. The overall schema consists of: testing total energy convergence with respect to the variable parameters of the method, calculation of the total energy of several different primitive cell volumes and fitting this $E(V)$ data to a Birch–Murnaghan curve.

The convergence parameters of DMC can be broken down into the following areas:

4.2.2.1 Jastrow factor

The form of the Jastrow factor was optimized using a 2×2×2 supercell. Variance minimization was used within VMC using 20,000 electronic configurations. Table 4.6 shows the optimized VMC total energies for the different expansion orders[1] of the polynomial expansions of the various Jastrow factor terms. These tests show no advantage in going beyond a Jastrow factor with expansion orders of four, two and three for the u, χ and f terms respectively. A small gain can be made with respect to the variation of the local energy in including a p term[2] but this was not worth the increase in computational cost. For each separate supercell size and primitive cell volume the individual parameters of the Jastrow factor were re–optimized using the variance minimization scheme. However the expansion orders remained fixed.

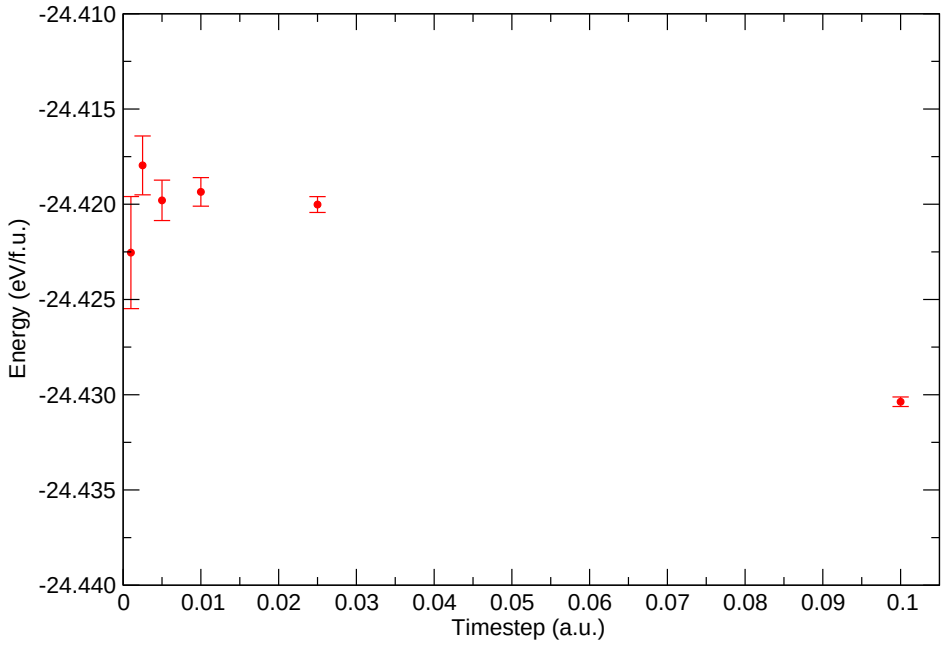
[1] These determine the order to which the polynomial expansion of each term is taken, and therefore the number of adjustable parameters in each term.

[2] This is a little used Jastrow factor term that involves a plane–wave expansion in the electon–electron separation.

4.2.2.2 Timestep

The results of the timestep tests performed on a 2×2×2 super–cell are shown in Figure 4.4 and as can be seen a timestep of 0.025 a.u. is sufficient to bring the timestep error to within the error bars on the points themselves, suggesting a timestep error of > 3 meV per formula unit (f.u.). It is also expected that the timestep error will be subject to a certain level of cancellation if energy differences are taken, in this case between primitive cells of differing volumes.

Figure 4.4: Energy vs. Timestep for pseudopotential bulk LiH with a 2×2×2 super-cell.



This figure also shows how for a given run length the statistical precision decreases with decreasing timestep. This is because for a given system setup with a given number of walkers a certain fixed length of imaginary time must be covered to obtain a given level of precision. Thus decreasing the length of the timestep while keeping the wall–time fixed reduces the amount of imaginary time simulated, reducing the statistical precision.

4.2.2.3 Finite–size effects

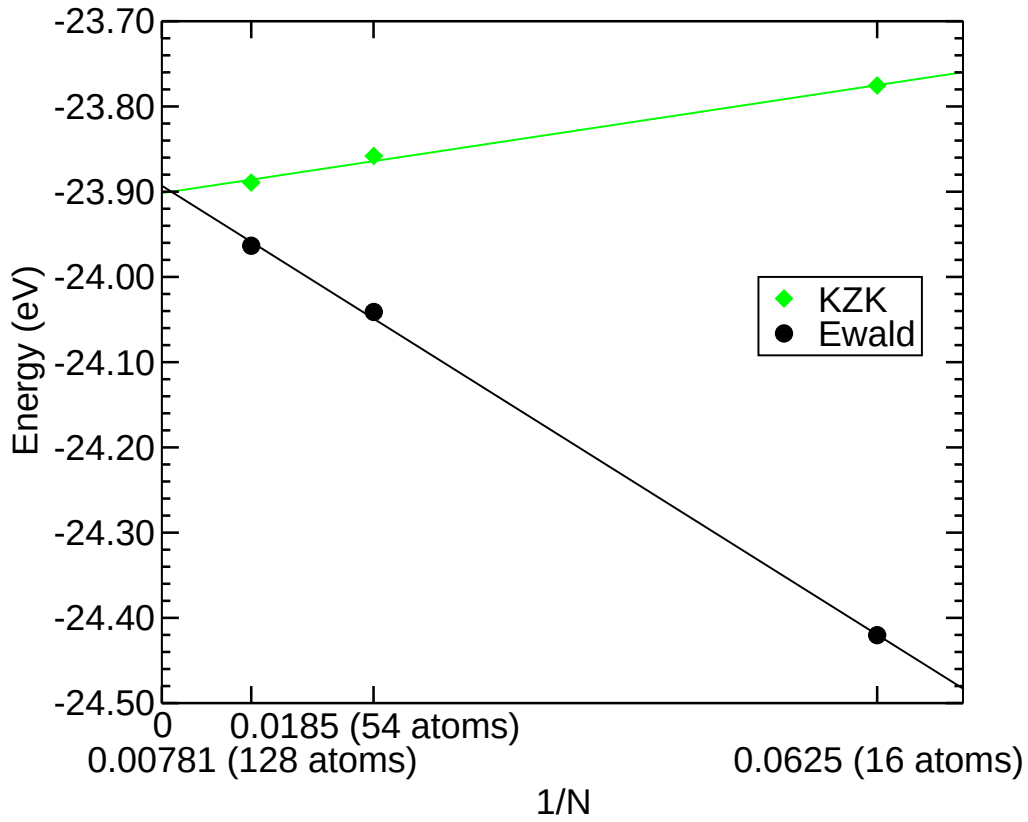
The finite–size effects detailed in Section 3.4 in this case are mitigated by the use of the KZK correction and extrapolation. The results of the finite–size tests are shown in Figure 4.5. With 128 atoms (a 4×4×4 supercell) the errors are converged to within 10 meV per f.u., when using the KZK correction. Indeed later we find when calculating the bulk properties for cells of 54 and 128 atoms the equilibrium lattice parameter does not change.

4.2.2.4 Results

With the convergence parameters now established the system properties can be calculated.

Figure 4.6 shows the $E(V)$ curves calculated with DMC using the two different supercell sizes. The $E(V)$ calculations were performed using both a 3×3×3 and 4×4×4 supercell both corrected for zero–

Figure 4.5: Energy vs. Supercell size: LiH Bulk. Shown are the original DMC total energies computed using the Ewald method and the values including the KZK correction. Both data sets have been fitted to Equation 3.35 to provide an infinite cell extrapolation. The statistical errors on each point are smaller than the symbols.

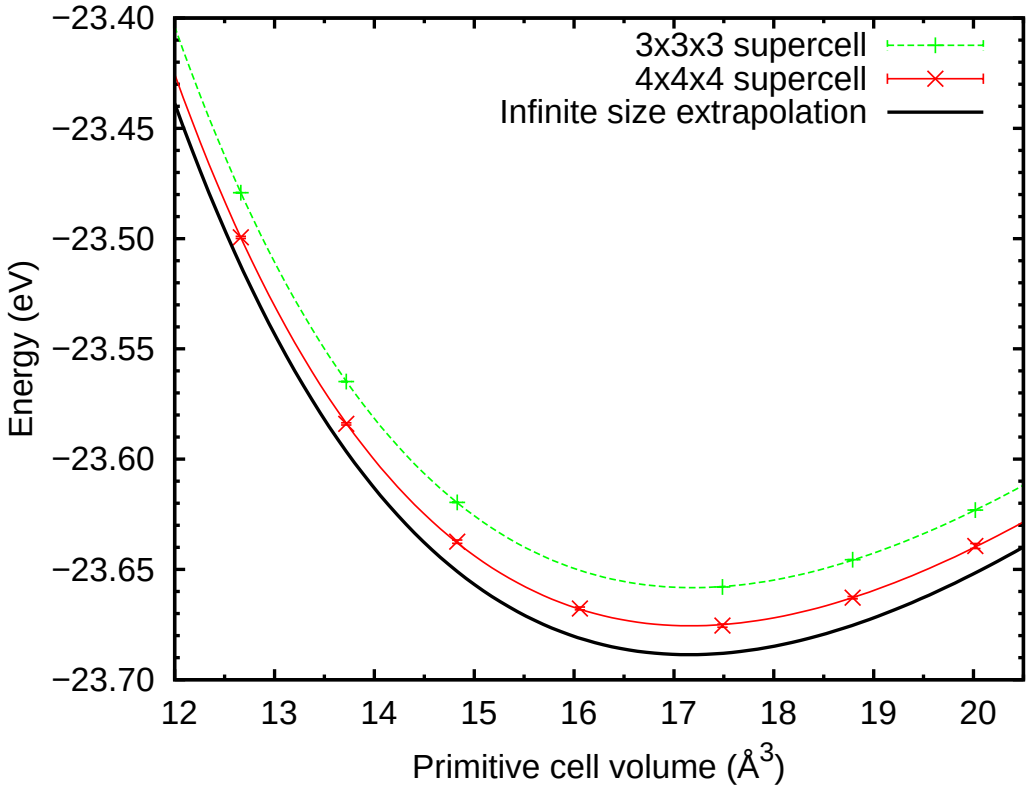


point energy contributions using the corrections determined in Section 4.1.2.3 and finite–size effects using the KZK corrections evaluated using PWscf. This allows the results to be compared directly with low temperature experimental work.

Table 4.7 shows the results of Birch–Murnaghan fits to the $E(V)$ data, with the error bars determined by Gaussian sampling of the DMC data. As can be seen the lattice parameter and bulk modulus are converged with respect to system size by the $3\times3\times3$ supercell. Comparing with experimental data the calculated lattice parameter is correct to within 0.8% and the bulk modulus to within the magnitude of experimental uncertainty. By extrapolating each point on the $E(V)$ curve to infinite–size using Equation 3.35 shows that the extrapolated lattice constant and bulk modulus are essentially unchanged. This can also be seen in Figure 4.6 where the only real change between the different $E(V)$ curves is a constant energy offset.

The cohesive energy was also determined requiring appropriate atomic energies for lithium and hydrogen. As both the hydrogen atom and the pseudo lithium atom are one electron problems Hartree–Fock theory provides the exact solution, and DMC calculations should just recover this. Thus by using

Figure 4.6: Total energy as a function of primitive cell volume for the pp-DMC calculations including the extrapolation to infinite system size. The lines indicate a Birch–Murnaghan fit to the data. Where error bars are not shown they are smaller than the points themselves.



the Hartree–Fock energies calculated by Trail and Needs in the generation of the pseudopotentials the DMC cohesive energies of the different supercell sizes were calculated. These results are also shown in Table 4.7.

This shows that while the lattice parameter and bulk modulus are both well converged by the $3\times3\times3$ supercell the cohesive energy is not. This is as expected as the shape of the $E(V)$ curve, which determines the lattice parameter and the bulk modulus, is only dependent on differences in energy between very similar supercells. However the cohesive energy is the difference between the total energy of a given supercell in DMC and two Hartree–Fock calculations on isolated atoms. In the former situation a large cancellation of errors like those caused by finite–size effects can be expected, in the latter situation this is not the case. By performing an extra total energy calculation with $5\times5\times5$ supercell a three–point extrapolation using Equation 3.35 is performed. This results in a cohesive energy within 12 meV of one of the experimental values quoted by Nolan *et al.* [36].

4.2.3 Surface

As with the bulk calculations the process of calculating σ with DMC is very similar to the way in which σ is calculated with DFT. The total energy of a number of slabs of different thickness are calculated, which

Table 4.7: Calculated bulk properties of LiH with both pseudopotential and all–electron DMC. Zero point corrections from PBE DFT are included throughout to allow comparison with the quoted experimental values. The cohesive energy is calculated at 4.084Å in all cases.

	a_0 / Å	B_0 / GPa	E_{coh} / eV
DMC 3×3×3[a]	4.0965(2)	30.5(1)	−4.6967(1)
DMC 4×4×4[a]	4.096(2)	31.1(8)	−4.7249(1)
DMC Extrap.	4.093(2)	31(1)	−4.7466(3)
Experiment	4.061(1)[b]	33-38[c]	−4.778,−4.759[d]

[a] Including the KZK correction

[b] Neutron diffraction data at 83 K from Ref [125]

[c] Scaled to 0 K from 300 K as outlined in Ref [36]

[d] 0 K experimental estimate using two separate thermochemical cycles, derived in Ref [36]

are then fitted to Equation 1.3. Unlike the plane–wave DFT calculations which require 3D periodicity to be enforced the DMC calculations can be performed in true 2D PBCs. This is due to the use of a blip basis–set expansion for the Slater determinant. As a result there is no dependence on vacuum gap unlike in the DFT case and all that must be ensured is that the DFT calculations generating the Slater determinants are converged with respect to the vacuum gap.

A choice of Jastrow parametrization must also be made along with the DMC timestep used and the size of, in this case the surface, supercell required for the finite–size effects to be eliminated.

4.2.3.1 Jastrow factor

A number of different Jastrow parameters were tested again using VMC variance minimization with a two–layer 2×2 surface supercell (8 atoms per layer). Table 4.8 shows that in the case of the surface a minimal Jastrow factor proved to be the best, with the addition of an f term tending to over–parameterize

Table 4.8: VMC optimized energies for a LiH 2×2 surface supercell with different Jastrow factors.

Expansion order			VMC Energy / E_{h}	VMC Variance / ${E_{\mathrm{h}}}^2$
u	χ	f		
2	2	-	−6.831(3)	0.172
4	2	-	−6.831(3)	0.168
4	4	-	−6.831(3)	0.169
4	2	2	−6.791(3)	0.199
4	2	3	−6.785(3)	0.210

the wavefunction and increase its energy[3]. Hence a Jastrow factor with just four parameters for the u and χ terms (corresponding to a expansion order of two for each) was used.

4.2.3.2 Timestep

Timestep tests were performed using a two–layer 2×2 surface supercell with timesteps down to 0.001 a.u. which indicated that a timestep of 0.01 a.u. brought the timestep bias to within ~ 5 meV per f.u., this was the timestep adopted for the rest of the calculations.

4.2.3.3 Finite–size effects

As mentioned in Section 3.4.2.1 KZK corrections depend on the volume of the DFT super-cell. As the DFT super–cells in this case contain a large vacuum gap in order to simulate slabs at infinite separation the KZK corrections could be made arbitrarily small by varying this parameter, which has no effect on the actual finite–size errors. So instead the usual one-body correction of Equation 3.34 was used with the LDA functional. Calculations for 3– and 4–layer slabs were performed with both 3×3 and 4×4 surface unit cells (18 and 32 ions per layer in the repeating supercell respectively). By fitting these data to Equation 1.3 estimates for the surface energy using the two different surface cells and hence the finite–size errors can be obtained. Comparing directly $\sigma_{3,4}^{3\times3}$ and $\sigma_{3,4}^{4\times4}$ shows they differ by only 0.006(5) J m^{-2} indicating the finite–size errors are well converged for 4×4 unit cells.

4.2.3.4 Results: Surface energy

Calculations of the total slab energy $E_{\mathrm{slab}}(N)$ for numbers of ionic layers from 3–6 using a 4×4 surface unit cell (18 ions per layer in the repeating supercell) were then performed. As with the DFT calculations estimates of σ are produced from subsets of these data to gauge convergence with respect to slab thickness. Table 4.9 shows this convergence of σ with respect to the slabs used when fitting to Equation 1.3.

Table 4.9: pp-DMC surface formation energy calculated using slabs of different thicknesses. Calculations performed on 4×4 surface unit cells.

Slabs used	σ / J m^{-2}
3,4,5,6	0.369(2)
4,5,6	0.373(3)
5,6	0.379(6)

The resulting best value for σ from the pseudopotential calculations is 0.373 J m^{-2}.

4.2.4 Discussion

What can be said for the validity of these estimates? Despite QMC's generally good record and solid theoretical basis it must be remembered that it is still an estimate. While we can be fairly confident that the errors incurred due to finite timestep and finite simulation cell are negligible because of the extensive

[3]Obviously this is an artifact of the variance minimization process. A perfect minimization would just set the redundant parameters to zero.

testing the same cannot be said of the two approximations that remain, that of the pseudopotential and fixed nodal surface.

The fixed–node error is very difficult to measure with most nodal improvement methods prohibitively expensive even for this relatively modest system. Besides improving the nodal surface does not necessarily indicate the size of the error as the exact surface will still not be attained.

Instead let us turn to the pseudopotential error. In this case the effect is threefold. Firstly there is the error in the pseudopotential itself. The pseudopotentials are designed to reproduce the all–electron Hartree–Fock valence states outside of a given radial cut–off, how well they reproduce the scattering for a fully correlated system can be variable. Bound up in this is also the issue of the locality approximation where an error proportional to the square in the error of the trial wavefunction is made to accommodate the pseudopotentials. Finally the is the issue of the core electrons on the lithium atom which are simply not present in the simulation. Hence changes in there correlation between systems is omitted entirely.

In order to address these issues we felt it was clear that an all–electron DMC study of LiH must be undertaken.

4.3 All-electron LiH

Calculations of extended systems within DMC without the use of pseudopotentials is a far from straightforward or standard task. We are only aware of one other study that has performed such all–electron bulk calculations [126] and no examples of surface calculations in the literature.

4.3.1 Cusp–correction

Usually ae–DMC calculations are performed using Gaussian basis–sets to represent the Slater orbitals. This is due to the shape of a wavefunction near to the nucleus being far more akin to that of a Gaussian than a plane–wave. Hence a much larger number of plane–waves must be used to describe it. However using Gaussians in PBCs have the problems described in Section 4.4.4 in reaching basis–set completeness. Also the computational efficiencies of blips are lost, with possibly thousands of Gaussians being non–zero at a given point in space and thus requiring evaluation at each move. It is for these reasons these ae–DMC calculations make use of plane–waves in the initial DFT generation of the Slater determinant. This is then expanded in blip functions as before.

As outlined in Section 3.2.1.2 the Kato cusp condition must be imposed at the electron–nuclear coalescence. In order to do this when using the blip–basis the method and implementation of Drummond [81] is used. This adds a small overhead to the calculation as the additive functions of this *general purpose cusp correction* must be generated at the beginning of the calculation and must be evaluated along with the blips during the random walk should any particle venture into the range of the correction.

4.3.2 Bulk properties

With this technical solution in place it was possible to proceed with the bulk calculations in much the same way as in the pseudopotential case, with thorough tests of the trial wavefunction construction, timestep dependency and finite–size effects.

4.3.2.1 Trial–wavefunction

Determining the size of the plane–wave basis required to generate a good trial wavefunction proved problematic. Table 4.10 shows the convergence of the VMC energy and variance of a $2\times2\times2$ supercell with respect to energy cut–off of the trial–wavefunction and fineness of the blip expansion. It is obvious that a blip–grid twice as fine as the natural grid spacing instantly reduces the variance by over half and is therefore a must in order to increase the efficiency of the DMC runs. Table 4.10 seems to indicate that an energy cut–off of $\sim 14 \times 10^3$ eV should be sufficient for the bulk calculations. Unfortunately it turns out that an energy cut–off of 6.8×10^4 eV is required in order to produce stable DMC runs. It appears that below this level the trial wavefunctions are not good enough to produce stable populations of DMC walkers. A consequence of this is the decision to represent the blips using only single precision, this reduces the wavefunction file to a manageable size. As before the LDA was used in the generation of the wavefunctions.

Table 4.10: Plane wave convergence of the all–electron LiH Slater determinant. Energy and variance of the local energy evaluated using VMC. n corresponds to that of Equation 3.39 which determines the blip grid spacing.

Energy cut–off / eV	Energy / E_h		Variance / ${E_\mathrm{h}}^2$	
	$n=1$	$n=2$	$n=1$	$n=2$
5,442	−8.056(8)	−8.05(1)	84(7)	31(1)
6,803	−8.083(8)	−8.09(1)	76(2)	21.3(9)
8,163	−8.080(8)	−8.07(1)	70(1)	20.7(9)
13,606	−8.063(8)	−8.10(1)	63(2)	18(2)
27,211	−8.095(8)	-	46(1)	-
40,817	−8.074(7)	−8.078(9)	43(4)	16(1)
54,423	−8.091(7)	-	37(1)	-

Table 4.11: VMC optimized energies for all–electron LiH using a 3×3×3 supercell with different Jastrow factors. The Slater determinant used is based on an energy cut–off of 2.7×10^4 eV with single blip spacing.

Expansion order			VMC Energy / E_h	VMC Variance / ${E_\mathrm{h}}^2$
u	χ	f		
4	2	3	−8.155(3)	88(1)
4	4	3	−8.154(3)	86(1)
8	4	3	−8.151(3)	89(1)
8	8	3	−8.153(3)	91(1)

Jastrow factor

It was assumed that the optimal all–electron Jastrow factor would have to contain at least as many terms as that used with the pseudopotential calculations. This Jastrow factor was re–optimized using the usual variance minimization along with other Jastrow factors with more parameters. The results of these tests are shown in Table 4.11 and indicate that the form used in the pseudopotential calculations is sufficient. As before the parameters were then re–optimized for every individual supercell used.

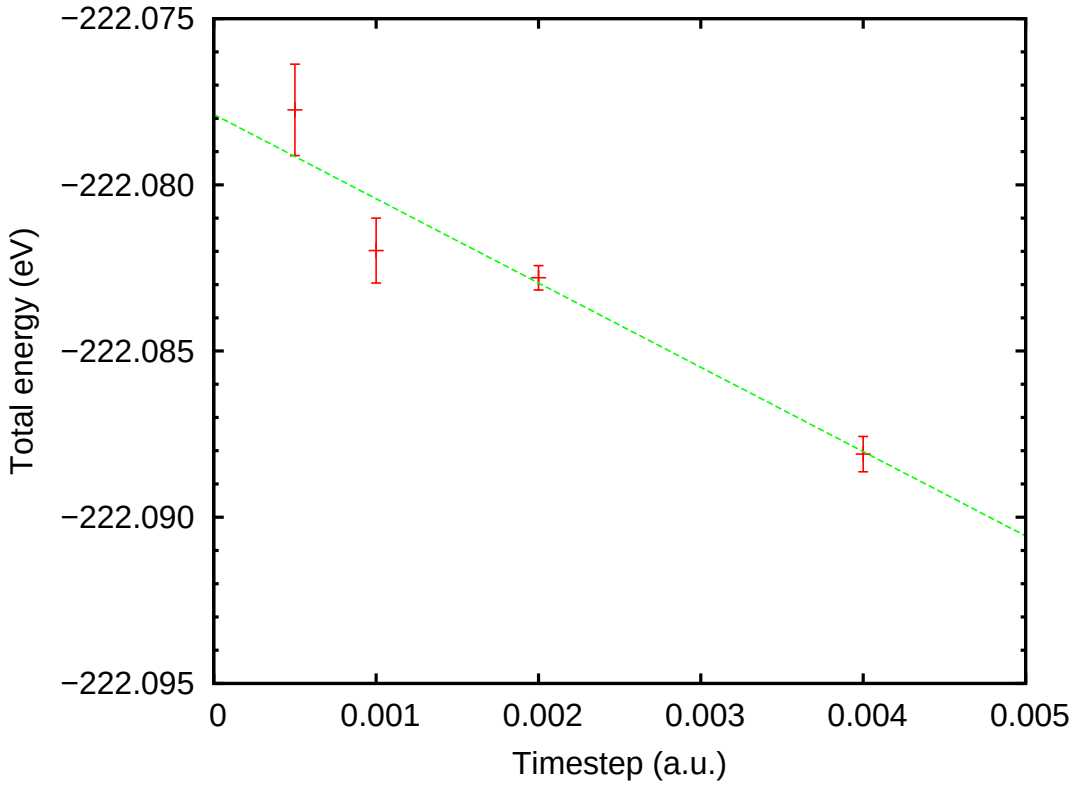
4.3.2.2 Timestep

Due to the fact that the length scales involved in the all–electron wavefunctions near the atomic cores are smaller than that of the wavefunctions in the pseudopotential case the timestep error is expected to be larger. Figure 4.7 shows the convergence of the energy with respect to timestep for the $3 \times 3 \times 3$ supercell.

With the lack of the locality approximation in the all–electron case it is possible to extrapolate to

zero timestep using this data. The figure also shows a linear fit to the results, which is clearly adequate, as expected from earlier work [127]. A timestep of 0.004 a.u. gives an error of only 10 meV/f.u., which more than suffices to give accurate results for the equilibrium a_0 and B_0.

Figure 4.7: Total energy vs. timestep. This shows the convergence of the ae-DMC calculations with respect to timestep and is for a 3×3×3 supercell of bulk LiH. A linear extrapolation to zero timestep is included.



4.3.2.3 Atomic energies

As before in order to calculate the cohesive energy it is necessary to know the individual atomic energies. Since the hydrogen atom is being treated with the full Coulomb potential it is sufficient to use the well known analytic ground-state energy of the hydrogen atom in the BO approximation.

For the lithium atom in order to maintain some hope of a cancellation of errors DMC calculations on the atom are used. These are done with LDA trial wavefunctions generated using the ATOMIC code, part of the QUANTUM ESPRESSO distribution [122]. These wavefunctions are generated by solving the Kohn–Sham equations on a radial grid. The grid used in this case is the same as that used for the Trail and Needs lithium pseudopotential.

The reason the wavefunctions are generated in this way as opposed to the more usual method of solving the KS equations using plane–waves for an atom in a large supercell was mainly one of computational efficiency. In order to avoid spurious interactions between the atom and its periodic

images the box has to be very large (of the order of 20 Å), with the large plane–wave cut–off required this makes the calculation very resource intensive and consequently produces a very large wavefunction file that the CASINO code has to then store in memory for the entire run. The radial grid calculation on the other hand can be performed in seconds and produces a very compact wavefunction file.

The same Jastrow factor as used in the bulk is used for the atomic calculation, the parameters being re–optimized with the variance minimization scheme.

Finally when calculating the cohesive energy at different timesteps a consistent timestep is used for both the bulk and the atomic DMC calculation, again in order to maximize the possible error cancellation.

4.3.2.4 Finite size effects

The pseudopotential calculations of Section 4.2.1 indicated that while the total energy is not fully converged with a 3×3×3 supercell the energy differences are. Also the work on LiF presented in Section 4.5 indicates that the MPC correction scheme performs at least as well as the KZK scheme. With this in mind the $E(V)$ curve for ae-LiH is calculated using a 3×3×3 supercell with the MPC correction. The reason for the choice of the MPC correction is due to the difficulty in converging the DFT total energies with respect to plane–wave cut–off, something which is necessary in order to accurately calculate the KZK corrections.

For the cohesive energy the MPC corrected energy is calculated with both a 3×3×3 and 4×4×4 supercell with a lattice parameter of 4.084 Å. This is then extrapolated to infinite–size using Equation 3.35. The offset between this and the MPC corrected 3×3×3 supercell value is only ~ 30 meV, which offers further justification for the use of the 3×3×3 supercell in generating the $E(V)$ curve.

4.3.2.5 Results

Table 4.12 shows the bulk properties obtained from the usual Birch–Murnaghan fit to the DMC generated $E(V)$ data. The procedure in calculating the ae–DMC cohesive energy uses six separate cohesive energy calculations performed with timesteps of 0.001, 0.002 and 0.004 a.u., on both the $3 \times 3 \times 3$ and $4 \times 4 \times 4$ supercells. At each timestep, the usual two point extrapolation to infinite supercell size was made, and then a final linear extrapolation to zero timestep. Table 4.13 shows the intermediate extrapolations along with the DMC Li atomic energies at each timestep. The fully extrapolated cohesive energy is also presented in Table 4.12.

The extrapolated Li atomic value is ~ 4 meV higher than the highly accurate non–relativistic results obtained by Puchalski *et al.* by minimizing the energy of wavefunctions expanded in a basis of Hylleraas functions [128, 129].

The agreement of the lattice parameter and the bulk modulus with experiment is very good indicating that the initial error in the pp–DMC calculations is down to the pseudopotential approximation rather than any fixed–node considerations. The all–electron DMC also offers some improvement over the pp–DMC cohesive energy. This brings the cohesive energy into line with the experimental predictions. Overall it is plain that in the bulk case the all–electron treatment offers noticeable improvement over the already quite accurate pseudopotential treatment.

Table 4.12: Calculated bulk properties of LiH with both pseudopotential and all–electron DMC. The cohesive energy is calculated at 4.084 Å in all cases.

	a_0 / Å	B_0 / GPa	E_{coh} / eV
Pseudopotential DMC	4.093(2)	31(1)	−4.7466(3)
All–electron DMC	4.061(1)	31.8(4)	−4.758(1)[a]
Experiment	4.061(1)[b]	33-38[c]	−4.778,−4.759[d]

[a] This value is extrapolated to infinite–size, zero–timestep using six separate calculations with 3×3×3 and 4×4×4 supercells and timesteps of 0.004, 0.002 and 0.001 a.u. Further details are presented in Table 4.13 .

[b] Neutron diffraction data at 83 K from Ref [125]

[c] Scaled to 0 K from 300 K as outlined in Ref [36]

[d] 0 K experimental estimate using two separate thermochemical cycles, derived in Ref [36]

4.3.3 Surface calculations

The familiar protocol of Section 4.2.3 was again adopted in order to calculate the surface formation energy. This consisted of calculating the total energy of a series of slabs and then fitting to Equation 1.3. In order to aid comparison with the previous work all the calculations were done with the lattice parameter $a_0 = 4.084$ Å.

4.3.3.1 Trial wavefunction

The vacuum gap required when generating the trial–wavefunctions added additional constraints on the choice of energy cut–off. For the same reasons as the isolated atom calculations were unfeasible the energy cut–off for the slabs was limited to 3.4×10^4 eV. This primarily made the DMC runs more susceptible to population control issues and resulted in a higher statistical error on the final values compared to the pseudopotential work.

Jastrow factor

Testing on various expansion orders for the terms in the Jastrow factor again found no benefit in changing from the established values of 4, 4 and 3 for the u, χ and f terms. Variance minimization of the parameters of such a Jastrow factor for a two layer 3×3 surface cell yielded a VMC energy of $-16.279(1)\ E_{\mathrm{h}}$ with a variance of $14.1(7)\ E_{\mathrm{h}}^2$.

By performing a subsequent optimization using the energy minimization method the VMC energy was reduced by 10 mE_{h} to $-16.289(1)\ E_{\mathrm{h}}$ with a variance of $13.17(7)\ E_{\mathrm{h}}^2$. Further tests using DMC indicated that the energy minimized trial wavefunctions seemed far less susceptible to population control issues than their variance minimized counterparts. This observation also seems to hold for LiF and would appear to be absent from the literature. As a result all of the Jastrow factors for the surface slabs were

Table 4.13: All–electron DMC cohesive energies of LiH calculated at three different timesteps and two supercell sizes. At each timestep the infinite size extrapolation calculated using Eq. 3.35 is shown. The result of a final linear extrapolation of these values to zero timestep is also shown. The ae–DMC Li atomic energies used at each stage are included for completeness, the exact Hartree–Fock atomic energy for H is used throughout. All energies are in eV.

Timestep / a.u.	Li atom / eV	E_{coh} / eV		
		3×3×3	4×4×4	Infinite size (extrap.)
0.004	−203.4881(4)	−4.7824(7)	−4.7728(56)	−4.7659(8)
0.002	−203.4875(9)	−4.7778(9)	−4.7675(22)	−4.7601(36)
0.001	−203.4850(8)	−4.7795(13)	−4.7684(11)	−4.7603(16)
0 (extrap.)	−203.4847(9)			−4.758(1)

first optimized with variance minimization before being re–optimized with energy minimization.

4.3.3.2 Finite size effects

The memory requirements of the blip functions also limited the size of the systems that could be studied. Due to the pseudopotential work indicating that the finite–size errors on the surface formation energy were negligible by the 3×3 surface supercell this supercell was adopted for the all–electron calculations.

Also only slabs of thickness 3-5 atomic layers were calculated, again with the pseudopotential calculations indicating the errors were negligible.

4.3.3.3 Results

Finally the timestep adopted was 0.004 a.u., the same as used in the all-electron bulk work. This is justified by the close agreement between the pp–LiH bulk and surface timestep behavior.

Table 4.14: Calculated surface formation energy of LiH with both pseudopotential DMC and all–electron DMC. The calculations are performed at 4.084 Å in both cases. DFT data from Section 4.1 is included.

	σ / J m^{-2}
Pseudopotential DMC	0.373(3)
All–electron DMC	0.44(1)
DFT LDA	0.466
DFT PBE	0.337
DFT RPBE	0.272

The final all-electron DMC result for the surface formation energy is given in Table 4.14. Again it is the case that the all–electron calculations offer a noticeable improvement over the pseudopotential approach. In this case the surface energy is increased by ~ 7 J m^{-2}.

4.3.3.4 Summary

At this point it is important to note what has been achieved. Estimates of σ have been obtained using highly converged DMC calculations, both with and without the pseudopotential approximations. From the benchmarks also obtained for the bulk properties which agree very well with experiment we would deduce that the all–electron estimate of $\sigma = 0.44(1)\ \mathrm{J\,m^{-2}}$ is highly accurate. This would make the LDA DFT estimate fairly close to the actual value as we might expect from the evidence found in the literature.

However in order to really be able to claim accuracy to $0.01\ \mathrm{J\,m^{-2}}$ it is not enough to offer such inductive reasoning, an independent cross check is desired.

4.4 LiH with quantum chemistry

With the case of the bulk properties independent verification of the DMC predictions could easily be obtained from the high quality experimental data. As we have seen however the experimental methods used for measuring σ can be questioned. In this particular case, LiH(001), no experimental data even exists, therefore our verification must come from another branch of theory.

Manby *et al.* [73] demonstrated in 2006 that the high–level quantum chemistry techniques of Section 2.3 can be applied to perfect crystalline systems. Nolan *et al.* [36] have since performed an in–depth study of LiH using this so–called hierarchical method and obtained very good agreement with experiment in the case of the bulk.

As we shall see the hierarchical method only provides estimates of the correlation contributions, these must be combined with accurate calculations of the Hartree–Fock part in order to provide any meaningful results.

The Manby group of the University of Bristol have already produced very accurate estimates of such correlation contributions to both the bulk and σ for LiH and LiF. What remained an issue however was the Hartree–Fock contributions. Previous work on the LiH bulk suggested that this was not a trivial problem and so it proved.

What follows is a brief description of the so–called hierarchical method and how it can be applied to surfaces.

A detailed account of our own Hartree–Fock calculations for LiH is then presented.

Finally the correlation results of Nolan (part of the Manby group) are shown combined with our Hartree–Fock results to give an accurate estimate of σ which can be compared with DMC.

4.4.1 Hierarchical method

This method starts off with one fairly uncontested assumption. It is that after partitioning the total energy of a crystal in the usual quantum chemical manner into a Hartree–Fock part and correlation part,

$$E = E_{\mathrm{HF}} + E_{\mathrm{corr}}, \tag{4.5}$$

the correlation part, E_{corr} is fairly short ranged.

In the extreme case this would allow the infinite system correlation energy to be calculated just by using large clusters. In reality however these clusters are still too large to be treated with accurate coupled–cluster methods. What Manby *et al.* did was to suggest the ansatz that the correlation energy of a cluster can be broken down into separate elements.

To the first approximation they suggested that the correlation energy of a cluster of LiH is just N times the correlation energy of an isolated LiH molecule with an inter atomic distance equal to that of the bulk system.

$$E_{\mathrm{corr}}^{lmn} = \frac{lmn}{2} e_{\mathrm{mol}} + \delta E_{lmn}. \tag{4.6}$$

Here E_{corr}^{nlm} is the correlation energy of a cuboid dimensions l, m and n, e_{mol} is the correlation energy of an isolated, in this case LiH, molecule and δE_{lmn} is the so–called *correlation residual* or the difference between the molecular assumption and the real cluster energy.

Now if this assumption is reasonably good a large fraction of the correlation energy can be obtained with a relatively easy molecular calculation. Obviously though this is still not quite good enough so they turned their attention as to how to calculate the correlation residual.

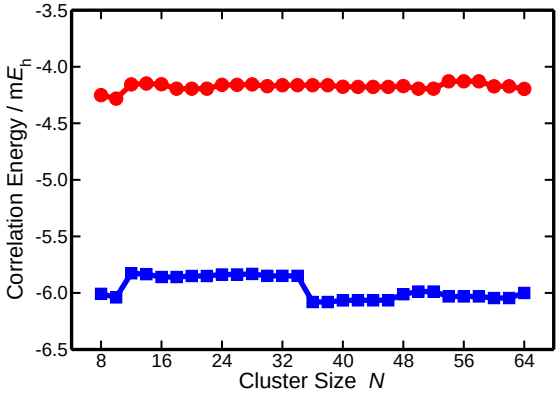
A further ansatz can be made as to the form of δE_{lmn},

$$\begin{aligned}\delta E_{lmn} =& 8E^{000} + 4[(l-2)+(m-2)+(n-2)]E^{001} \\ &+ 2[(m-2)(n-2)+(n-2)(l-2)+(l-2)(m-2)]E^{011} \\ &+ (l-2)(m-2)(n-2)E^{111} \ . \end{aligned} \tag{4.7}$$

This can be seen as breaking down the contributions to δE_{lmn} into that of atomic sites in a given cuboid structure. There are 8 corner sites represented by the E^{000} coefficient, $4[(l-2)+(m-2)+(n-2)]$ edge sites (that are not also corners) represented by the E^{001} coefficient, $2[(m-2)(n-2)+(n-2)(l-2)+(l-2)(m-2)]$ surface sites (that are not also edges or corners) represented by the E^{011} coefficient and finally $(l-2)(m-2)(n-2)$ bulk sites represented by E^{111}. It is straightforward to see that in the limit of the infinite system $n, l, m \to \infty$ the bulk term E^{111} dominates. By calculating the correlation energy of four different cluster sizes (the hierarchy of the hierarchical method) estimates of the coefficients can be determined by solving the resulting simultaneous equations.

It turns out that the convergence of these coefficients is very quick with respect to cluster size. Figure 4.8, reproduced from calculations by S.J. Nolan [81] demonstrates this for LiH.

Figure 4.8: Convergence of E^{111} and E^{011} with respect to cluster size courtesy of S.J. Nolan [81]. The circles represent the E^{011} coefficients with the squares representing the E^{111} coefficients.



4.4.2 Bulk comparison

This combined with periodic Hartree–Fock calculations using PWscf allowed Nolan *et al.* to calculate the same bulk properties for LiH. Table 4.15 shows the comparison to the DMC calculations.

The agreement between quantum chemistry and ae–DMC is astounding. For the lattice parameter and bulk modulus they both agree fully not only with each other but also experiment. With the cohesive energy they both agree to within the experimental range (19 meV), but beyond that it is hard to say which number is definitive. It can be noted that in the ae–DMC case, as has been outlined in Section 4.3.2.5, the ae-DMC result for the Li atom is within ~ 4 meV of what can be regarded as the definitive result.

Table 4.15: Calculated bulk properties of LiH with both pseudopotential and all–electron DMC and hierarchical quantum chemistry with and without core effects. Zero–point effects have been included in both the DMC calculations and the quantum chemistry calculations via the PBE-DFT phonon calculations outlined in Section 4.1.2.3. The cohesive energy is calculated at 4.084 Å in all cases.

	a_0 / Å	B_0 / GPa	E_{coh} / eV
DMC Extrap.	4.093(2)	31(1)	$-4.7466(3)$
All–electron DMC	4.061(1)	31.8(4)	$-4.758(1)$[a]
Quantum Chemistry (no core)[b]	4.099	31.9	-4.7087
Quantum Chemistry (with core)[b]	4.062	33.1	-4.7710
Experiment	4.061(1)[b]	33-38[c]	-4.778,-4.759[d]

[a] This value is extrapolated to infinite–size, zero–timestep as outlined in Section 4.3.2.5.

[b] Neutron diffraction data at 83 K from Ref [125]

[c] Scaled to 0 K from 300 K as outlined in Ref [36]

[d] 0 K experimental estimate using two separate thermochemical cycles, derived in Ref [36]

Any remaining error in E_{coh} would be expected to come from fixed–node considerations on the LiH bulk calculations. These, being variational, would only serve to reduce the cohesive energy still leaving the DMC estimate much closer to -4.759 eV than -4.778 eV.

What is interesting is the agreement between frozen core quantum chemistry and pp–DMC. For the lattice parameter and bulk modulus the agreement is again exact. This is not something you would necessarily expect to see, while both the pp approach and the frozen core approach neglect the correlation of the core electrons it is not correct to suggest the pp–DMC is performing frozen core calculations. Other errors might be expected to arise from the pps as the core is actually absent (replaced by a potential). The disagreement in the cohesive energy is likely caused by this.

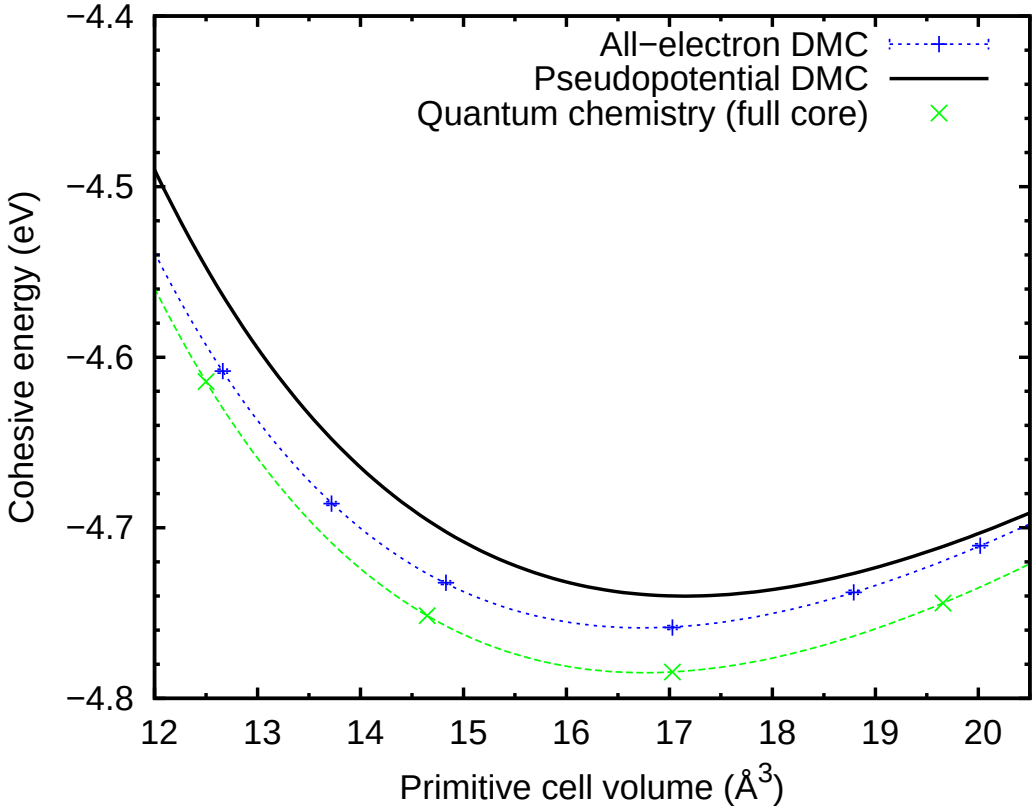
Figure 4.9 shows the $E(V)$ curves for the different methods. As can be seen the shape of the resultant ae–DMC Birch–Murnaghan fit agrees extremely well with that of the quantum chemistry calculations. In fact both curves only seem to differ by an additive constant a fact borne out by the agreement in Table 4.12.

The all–electron calculations in Figure 4.9 also show where the error in the pp–DMC calculations is to be found, with the two curves diverging from each other as the crystal is compressed. This is due to the lack of the lithium 1s electrons in the pp–DMC calculations.

4.4.3 Surface energies

Looking back at Equation 4.7 the limit of an infinitely large and infinitely thin slab could be taken, $n \to 0, l, m \to \infty$. This, as previous discussions on the extrapolation of Equation 1.3 have indicated,

Figure 4.9: Cohesive energy vs. primitive cell volume of ae-LiH. Here the DMC energy differences have been calculated using a 3×3×3 supercell and timestep of 0.004 a.u. and the point at 17.03 Å^3 calculated using the extrapolation procedure in the text.



should yield the surface formation energy, which indeed is given by [81],

$$\sigma = \sigma_{\mathrm{HF}} + \frac{4}{a^2}(E^{011} - E^{111}). \tag{4.8}$$

Where σ_{HF} is the Hartree–Fock surface formation energy, and a is the lattice parameter. Now the E^{011} coefficient is calculated as a consequence of the bulk calculations hence this correlation contribution to σ can easily be obtained from the Nolan *et al.* data. It seems that with the calculation of σ_{HF} there would be an accurate quantum chemistry based calculation of σ with which the DMC calculation could be compared.

4.4.4 Hartree–Fock surface energies

The original periodic Hartree–Fock calculations for the bulk used in Reference [36] were performed with PWscf using a plane–wave basis and pseudopotentials. In order to eliminate the pseudopotential error a rather long and arduous set of corrections had to be calculated by Hartree–Fock calculations on clusters with Gaussian basis–sets by Gillan *et al.* [130]. Repeating these calculations for surface energies was not an enviable task. Two alternatives were possible.

Firstly the CRYSTAL code is capable of performing all–electron Hartree–Fock calculations in PBCs.

The drawback is the use of Gaussian basis–sets. The large basis–sets required to approach the compete basis limit are computationally troublesome. Due to overlapping diffuse Gaussian centered on adjacent atomic sites linear dependencies are introduced into the basis–set making solving the Roothan–Hall equations very difficult and the whole SCF procedure numerically unstable.

The second approach is to use the PAW Hartree–Fock implementation in the VASP code. Here the basis–set is straightforward to converge being based on plane–waves, but the use of PAW data generated using DFT approximations introduces an error similar in type to the pseudopotential error of the PWscf calculations.

4.4.4.1 CRYSTAL calculations

Luckily the basis issue of the CRYSTAL calculations can be sidestepped. This is due to the work of Paier *et al.* [131]. They performed benchmark calculations on bulk LiH using both the CP2K and GAUSSIAN codes. While the methods they used were far too computationally demanding to be of use in this situation they, as a by product, produced a highly optimized bulk LiH Gaussian basis–set. Civalleri *et al.* [132] demonstrated that the CRYSTAL code could be altered to accommodate this basis and indeed reproduce the calculations of Paier *et al.* at a fraction of the computational cost. Thus the technical convergence parameters can be taken from Reference [132] and the basis from [131].

Since the basis was optimized for the bulk it relies on significant overlap from adjacent basis functions to fully describe a given atom. Hence for the surface it was found that a number of *ghost* layers needed to be present both above and below the surface. These ghost layers consist of the atomic basis functions of a usual layer without the associated electrons or nuclear charge. Table 4.16 shows the convergence of σ_{HF} with respect to the number of ghost layers. showing that with two–layers this effect is well converged.

Table 4.16: Convergence of σ_{HF} for LiH using CRYSTAL with respect to the number of layers of *ghost* atoms above and below the surface. Based on two-point extrapolations from slabs of 4-5 layers. All energies are quoted in $\mathrm{J\,m^{-2}}$.

Ghost layers	σ_{HF}
0	0.43835
1	0.19886
2	0.19849
3	0.19854

With this a sequence of slabs in true 2D PBCs were calculated using the CRYSTAL code. Table 4.17 shows the usual convergence with respect to slabs discarded yielding the best estimate for $\sigma_{\mathrm{HF}}^{\mathrm{CRYSTAL}}$ of $0.198\ \mathrm{J\,m^{-2}}$.

Table 4.17: Hartree-Fock approximation to surface formation energy for LiH, $a = 4.084$ Å, using CRYSTAL and VASP. All energies are quoted in J m^{-2}.

Slabs used	CRYSTAL	VASP
2-8	0.20005	0.19363
3-8	0.19883	0.19114
4-8	0.19825	0.19001
5-8	0.19819	0.18944
6-8	0.19836	0.18926
7-8	0.19703	0.18864

4.4.4.2 VASP calculations

In order to provide some corroboration for the surface energy calculations using the PAW implementation of Hartree–Fock theory in VASP were also performed.

These calculations follow much the same pattern as the VASP DFT calculations performed in Section 4.1. The only difference is the use of an initial DFT calculation to provide the starting wavefunctions for the Hartree–Fock SCF cycles. This is merely a computational convenience in order to speed up the convergence of the SCF method and should not effect the final result.

The PAW potentials used were the hardest PBE potentials available in the VASP library. While not ideal tests using the PBE datasets with hybrid functionals by Paier *et al.* [133, 134, 135] suggest the errors introduced should be small. The energy cut–off used was 700 eV again enough to ensure convergence with these particular potentials.

Table 4.18: k–point convergence of VASP Hartree–Fock energy for the LiH bulk.

k–point grid	Total energy / eV
3×3×3	-7.3722
4×4×4	-7.2308
5×5×5	-7.1990
6×6×6	-7.1825
7×7×7	-7.1741
8×8×8	-7.1696
9×9×9	-7.1666
10×10×10	-7.1649
11×11×11	-7.1637
12×12×12	-7.1628
13×13×13	-7.1622

Table 4.18 shows the convergence of the bulk system with respect to k–point grid. As expected of HF calculations this is slow, but it appears that by $12\times12\times12$ the total energy is converged to within a couple of meV, which is consistent with the k—point grid used in the CRYSTAL calculations. This is expected to be adequate when energy differences of the slab system are taken.

Table 4.17 shows the usual convergence of $\sigma_{\mathrm{HF}}^{\mathrm{VASP}}$ with respect to discarded slabs. Comparing this to $\sigma_{\mathrm{HF}}^{\mathrm{CRYSTAL}}$ a consistent discrepancy of $\sim$0.008 J m^{-2} is noticed. The convergence pattern of both methods however seem very similar. Given both methods have their sources of error the discrepancy of 0.008 J m^{-2} can be taken as the confidence interval for the final value which is quoted as $\sigma_{\mathrm{HF}} = 0.19$ J m^{-2}.

4.4.5 Comparing σ

With the HF estimate for σ it is now possible to compare the full hierarchical value of σ with the calculated DMC value. Table 4.19 shows this comparison.

Table 4.19: Calculated surface formation energy of LiH with both pseudopotential and all–electron DMC and hierarchical quantum chemistry with and without core effects. The calculations are performed at 4.084 Å in both cases.

method	σ / J m^{-2}
DMC pseudopotential	0.373(3)
DMC all–electron	0.44(1)
Quantum Chemistry (frozen core)	0.402
Quantum Chemistry (full core)	0.434

The ae–DMC value for σ agrees with the full–core quantum chemistry value to within the statistical error on the DMC value (0.01 J m^{-2}). This gives very compelling evidence that the true value (less any thermal and zero–point considerations) is indeed 0.44(1) J m^{-2}.

As we might expect the frozen–core quantum chemistry does not agree exactly with the pp–DMC, indeed the difference between ae–DMC and pp–DMC is somewhat larger than just the core–correlation.

We can now say with a large amount of confidence that for LiH(001) $\sigma = 0.44(1)$ J m^{-2} with LDA DFT again providing a relatively good estimate compared to other functionals. In this case we have indeed gone beyond both DFT and chemical accuracy.

4.5 LiF: The Gilman method tested

As outlined in Section 1.2.2 one of the popular experimental methods for measuring the surface formation energy of easily cleaved materials like LiF is the method of Gilman [16].

Section 1.2.2 also outlined the alternative method and measurements of Burns and Webb [23, 24] which call into question the accuracy of Gilman's results for LiF. The discrepancy is striking with Gillman quoting 0.34 J m^{-2}and Burns and Webb 0.48(5) J m^{-2}.

In the previous section we managed to obtain a definitive benchmark value of σ using both QMC and the hierarchical scheme for LiH(001). This allowed us to arbitrate between the various estimates coming from DFT.

In this section we will do the same for LiF this time not only arbitrating between the DFT predictions but also the experimental ones. In this case pp–DMC will be used due to the prohibitive cost of all–electron calculations on LiF.

As before we begin by presenting a DMC study of the bulk in order to check the accuracy of our pp–DMC assumptions. Then a pp–DMC study of σ is presented.

Finally accurate Hartree–Fock calculations on both the bulk and surface are performed using a specially constructed basis–set hierarchy. These are again combined with correlation data supplied by Nolan of the Manby group [136] allowing us to quote ΔCCSD(T) predictions for both the bulk and the surface.

4.5.1 Bulk properties

The bulk calculations followed the prototype of the pp–LiH work. The pseudopotentials of Trail and Needs were again used which in this case reduces the fluorine atom to a seven electron problem. This gives eight electrons per ion pair. From this it is obvious that these calculations are more expensive in terms of computer time than the two electron pp–LiH calculations.

4.5.1.1 Trial wavefunction

As before the trial wavefunction adopted was of a Slater–Jastrow type, with the Slater part obtained from plane–wave LDA DFT calculations, re–expanded in blip functions. The variance of the local energy for the trial wavefunctions proved to be much higher than that encountered in the pp–LiH calculations. There the variance was of the order of 0.1 E_{h}here it was $\sim 80\ E_{\mathrm{h}}$. It turns out that this posed a problem when performing the DMC calculations, seemingly causing population explosions of the walkers.

Several different Slater parts were tried, from using PBE and PBE0 to the use of different wavefunctions for the Kleinman–Bylander projectors. These changes made little difference to the overall population stability. Instead it was found that the best stability could be obtained by using the LDA functional but with a very high cut–off of 13,506 eV with a blip grid that used half the natural spacing. This combined with the use of energy minimization of the Jastrow parameters reduced the variance to $\sim 20\ E_{\mathrm{h}}$ and led to DMC runs which proceeded without incident.

Jastrow factor

Table 4.20 shows the results of the testing different forms of the Jastrow factor. The parameters have been generated with variance minimization within VMC. The Jastrow factor formed with expansion

Table 4.20: VMC optimized energies for a 3×3×3 supercell of bulk LiF with different Jastrow factors. The Slater determinant used is based on an energy cut–off of 4.08×10^3 eV with single blip spacing.

Expansion order			VMC Energy / E_h	VMC Variance / E_h^2
u	χ	f		
2	2	-	−24.468(3)	81(1)
2	4	-	−24.478(4)	90(10)
2	6	-	−24.486(4)	80(2)
4	2	-	−24.570(3)	79(3)
4	4	-	−24.559(3)	76(2)
4	6	-	−24.569(3)	74(2)
6	2	-	−24.569(3)	77(2)
6	4	3	−24.565(3)	75(1)
6	6	3	−24.558(3)	72(1)
4	2	3	−24.5751(9)	73.6(6)
4	4	3	−24.580(1)	73.7(6)
4	6	3	−24.5751(9)	73.8(5)
6	2	3	−24.579(1)	73.3(6)
6	4	3	−24.5760(9)	77(2)
6	6	3	−24.5759(9)	74.3(6)

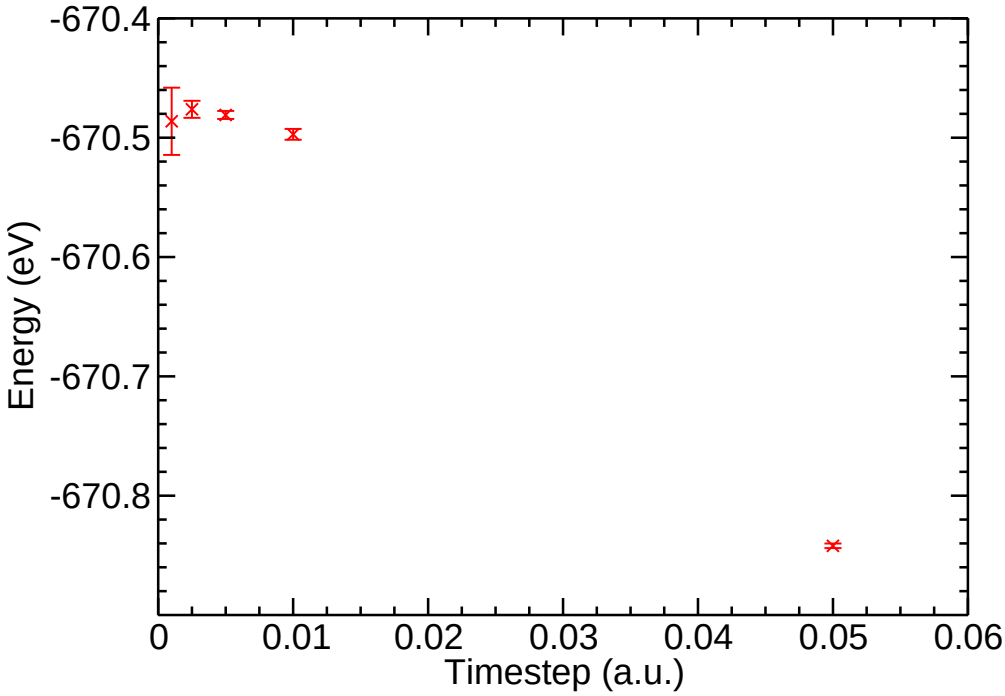
orders of 4, 4 and 3 for u, χ and f was selected as the optimal form for the remaining calculations. Further energy minimization of this form leads to a VMC energy of $-24.5973(9)$ E_h with a variance of 76.0(7) E_h^2. This reduction in energy but increase in variance compared to the variance minimization was replicated across all of the bulk LiF calculations.

Despite this the DMC runs were better behaved with the energy minimized wavefunction, experiencing far fewer population control issues. Hence for each individual supercell the Jastrow factor was first optimized with variance minimization which offered a good starting point for further energy minimization.

4.5.1.2 Timestep

As shown in Figure 4.10 a timestep of 0.005 a.u. was required in order to reduce the timestep error to less than 7 meV/f.u.. As before the locality approximation was used throughout.

Figure 4.10: DMC timestep tests on bulk LiF using a 3×3×3 supercell.



4.5.1.3 Finite–size effects

Due to the expense of the LiF calculations it was decided that any savings that could be made through the use of finite–size corrections should be taken. Hence results of extensive tests using both the MPC and the KZK methods of Section 3.4 are presented in Figure 4.11. This could be done relatively cheaply as due to the particular implementation of the MPC method used. As outlined in Section 3.4.2.2, this provides both the Ewald energy and the MPC energy with just a single calculation. As can be seen from Figure 4.11 both the MPC and the Ewald schemes extrapolate to within 19(19) meV of each other (-670.5089(19) and -670.5070(19) eV respectively).

They are both within $\sim$ 20 meV of the extrapolated KZK value, which seems to suffer due to over correcting the 3×3×3 supercell. As is clear from this a 3×3×3 supercell (54 atoms) using the MPC scheme is enough to reduce the finite–size errors to $\sim$ 30 meV. This is in line with the work on pp and ae–LiH in Sections 4.2.1 and 4.3.

4.5.1.4 Results

Total energy calculations on 3×3×3 supercells at varying lattice parameters were then performed. Again both the MPC and the KZK schemes were employed. The results are fitted to the usual third order Birch-Murnaghan equation of state, the results shown in Table 4.21.

It is clear that while the uncorrected Ewald scheme is not good enough for accurate bulk parameters, the MPC and KZK schemes provide excellent agreement with each other and experiment providing lattice parameters that agree with experiment to better than 0.3% and cohesive energies to within 27(15)

Figure 4.11: Finite size tests for bulk LiF performed with a timestep of 0.005 a.u. and both the KZK and MPC correction schemes. Fits of Equation 3.35 are shown along with the uncorrected Ewald values.

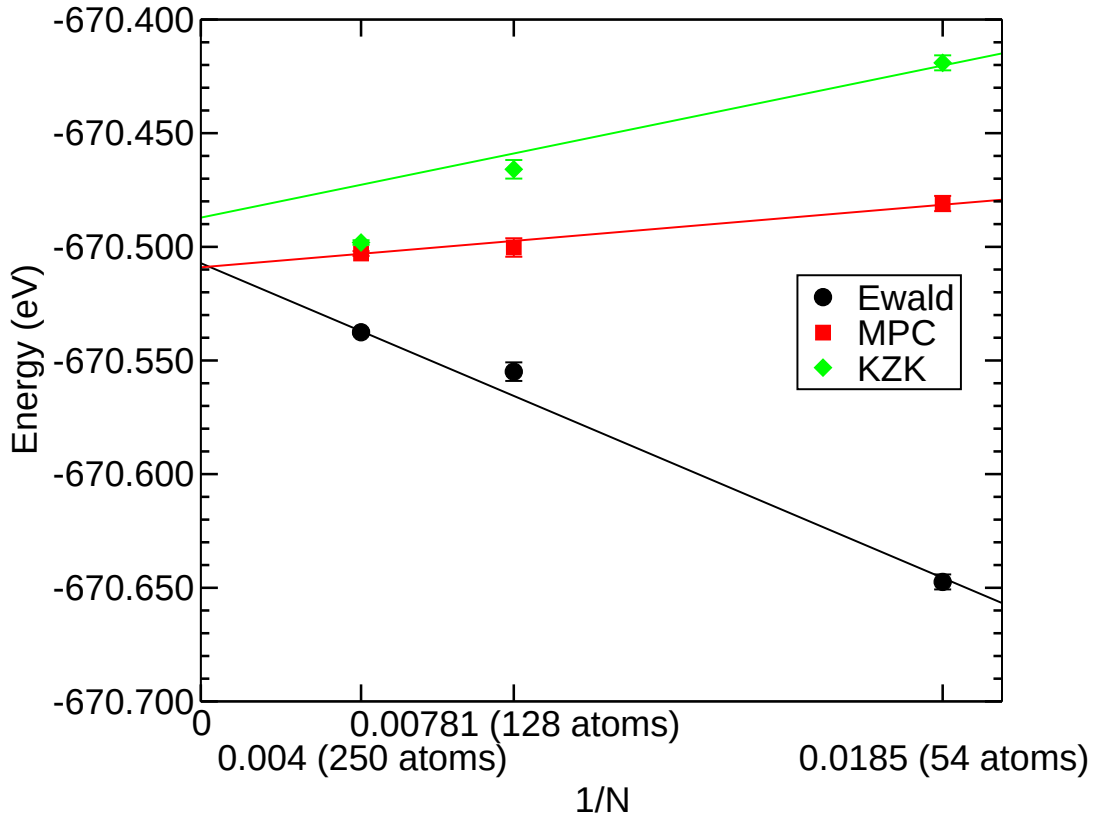


Table 4.21: Calculated bulk properties of LiF using uncorrected DMC, MPC corrected DMC and KZK corrected DMC. The cohesive energy is calculated at 4.004 Å in all cases.

	a_0 / Å	B_0 / GPa	E_{coh} / eV
Ewald DMC	3.984(6)	70(1)	−8.827(4)
MPC DMC	4.004(7)	64(2)	−8.827(4)
KZK DMC	4.015(6)	65(2)	−8.827(4)
Experiment [118, 119]	4.005	69.8	−8.800(14)

meV. The cohesive energies themselves are calculated using 3×3×3 and 4×4×4 supercells extrapolated to infinite–size using Equation 3.35.

The atomic energy for the pseudopotential lithium atom as discussed in Section 4.2.1 can be taken directly from the Trail and Needs Hartree–Fock calculations. The fluorine atom however had to be calculated using DMC being as it is a many–electron problem. The Hartree-Fock Slater determinant of Trail and Needs was used with an optimized Jastrow factor of the same construction as the bulk calculations. The resulting energy is -656.203(2) eV at a timestep of 0.005 a.u..

Table 4.22 shows the results of these calculations and as expected shows that all three methods

extrapolate to the same limiting value with the MPC interaction already being within ~ 6 meV of it with only a 3×3×3 supercell.

Table 4.22: DMC cohesive energy of LiF at 4.004 Å with the different correction schemes. All energies in eV.

Supercell	Ewald	KZK	MPC
3×3×3	−8.9880(40)	−8.7439(40)	−8.8305(31)
4×4×4	−8.8950(63)	−8.7920(63)	−8.8283(58)
Extrapolation	−8.827(4)	−8.827(4)	−8.827(4)

4.5.2 Surface formation energy

The surface calculations used the same pseudopotentials and trail wavefunction generation procedure as was used for the bulk. As has been described previously slabs of increasing thickness were used and the total energies fitted to Equation 1.3.

Timestep testing indicated that a timestep of 0.005 a.u. was sufficient to remove the timestep error as in the bulk.

4.5.3 Trial wavefunction

The population control problems of Section 4.5.1.1 were not observed in the slab geometries. This meant that an energy cut–off of 4,080 eV could be used, re–expanded with a blip grid that used half the natural spacing. Table 4.23 shows the tests of the form of the Jastrow factor. This indicated that expansion orders of 6, 6 and 2 for the u, χ and f terms respectively is the most optimal. In this case energy minimization offered no decrease in VMC energy for the trial functions so at each supercell the Jastrow parameters were optimized using variance minimization only.

4.5.3.1 Finite size

Calculations on 3, 4 and 5 layer slabs yielded the results of Table 4.24. This compares the Ewald and MPC surface formation energies using both the 3×3×3 and 4×4×4 supercell. The errors on these values are larger than their LiH counterparts due to the higher statistical error on the raw DMC data. This in turn is due to the much higher cost of the calculations. Within the errors there seems to be no discernible difference between the Ewald and MPC data.

The Ewald method was selected for the final calculations due to the somewhat odd behavior of the MPC method with the six layer slab. For the five layer case the difference between the MPC energy and the Ewald energy was 2.9 meV where as for the six layer case this was 1.6 eV. This suggests that in the six layer case the MPC approximation broke down in some way. Analysis of the Fourier components of the MPC interaction yielded no indication as to why this was the case, hence the Ewald method was adopted from this point on. The single particle LDA correction from Eq 3.34 amounted to less than a 0.003 J m^{-2} change in the surface energy so was neglected.

Table 4.23: VMC optimized energies for a two layer 3×3 LiF supercell with different Jastrow factors. The Slater determinant used is based on an energy cut–off of 4.08×10^3 eV re–expanded with a blip grid that used half the natural spacing.

Expansion order			VMC Energy / E_h	VMC Variance / E_h^2
u	χ	f		
2	2		−49.124(5)	27.7(5)
2	4		−49.107(4)	26.7(4)
2	6		−49.121(4)	27.1(4)
4	2		−49.139(4)	24.6(4)
4	4		−49.141(4)	23.8(3)
4	6		−49.134(4)	24.7(4)
6	2		−49.145(4)	24.4(4)
6	4		−49.142(5)	24.1(4)
6	6		−49.145(4)	23.9(3)
2	2	2	−49.203(3)	22.2(4)
2	4	2	−49.210(4)	22.1(4)
2	6	2	−49.203(3)	23.1(7)
4	2	2	−49.215(3)	23(1)
4	4	2	−49.199(4)	19.8(3)
4	6	2	−49.216(4)	20.9(6)
6	2	2	−49.206(4)	21.2(4)
6	4	2	−49.216(3)	20.2(3)
6	6	2	−49.225(4)	20.4(4)
4	2	3	−49.210(4)	20.7(5)
4	4	3	−49.212(3)	19.9(3)
4	6	3	−49.213(3)	20.8(5)
6	2	3	−49.226(4)	19.8(3)
6	4	3	−49.218(4)	19.3(3)
6	6	3	−49.207(4)	24(2)

4.5.3.2 Results

The best estimate for the DMC surface formation energy of LiF is therefore 0.48(3) J m^{-2} based on an Ewald calculation using a 3×3 supercell with slabs of 3-6 atomic layers.

The zero–point energy contribution is estimated to be −0.029 J m^{-2}. This gives a value of 0.45(3) to compare with the Burns and Webb experimental number of 0.48(5). This would seem to indicate the accuracy of Burns and Webb's method as opposed to Gilman's, however as has been explained in Section 4.4 it should be possible to obtain an independent estimate from quantum chemistry to verify

Table 4.24: DMC surface formation energy of LiF at 4.004 Å with the different correction schemes. All energies in $\mathrm{J\,m^{-2}}$.

Supercell	MPC	Ewald
3×3	0.44(3)	0.46(3)
4×4	0.43(6)	0.47(6)

this. Indeed the correlation part has been calculated by Nolan *et al.* [136] so for a comparison all that remains is the Hartree–Fock contribution. That shall be calculated presently.

4.5.4 Hartree–Fock

Unlike the calculations of Section 4.4 we are not fortunate enough to have already a fully optimized Gaussian basis–set suitable for the bulk. The specific optimization scheme of Paier *et al.* [131] is technically very challenging and computationally expensive, this is because it attempts to reach the basis–set limit in one bound so to speak, obtaining as it does a single basis–set producing calculations of micro–Hartree accuracy. The approach adopted in this work is more akin to that of Jensen [53] itself inspired by Dunning [137] where a hierarchy of basis–sets is produced each more accurate than the last allowing the basis–set limit to be reached via extrapolation.

4.5.4.1 Bulk

The basis was optimized with the bulk calculations with the assumption that as in Section 4.4.4 it could be used in the surface calculations merely by adding ghost layers.

As the starting point the lithium basis from the highly accurate work of Paier *et al.* [131] is assumed. Some of the more diffuse functions are omitted to reduce problems of linear dependency. This is justified on the basis the LiF is an ionic solid hence the basis on the lithium ions is really only attempting to describe the tightly bound Li^+ ion.

For the fluorine ion the polarisation consistent basis of Jensen [52, 138] (pc-n) was taken as the starting point. Here the default contractions were discarded and re–optimized using Hartree–Fock calculations performed using the Molpro code [139]. The full details of the basis–sets are given in Appendix A.

In this way a hierarchy is created by going up the pc-N chain on the fluorine atom while keeping the basis fixed on the lithium. The assumption is the lithium basis incompleteness can be compensated for by overlapping functions from the fluorine sites.

The bulk properties were calculated in the usual way by fitting the total energies of a number of different cell volumes, calculated using the CRYSTAL code to a third order Birch–Murnaghan curve. Convergence with respect to the integral tolerance parameters outlined in Reference [132] was checked and the k–point grid [4] 12×12×12.

[4] referred to as the shrinking factor in CRYSTAL

BSSE: Lattice energies vs. cohesive energies

The basis–set superposition error described in 2.1.4.3 manifests itself in such $E(V)$ curves as a negative shifting of the minimum. This is because when not at the basis–set limit the increase in Gaussian overlap caused by ions moving together results in a better basis–set. Hence the energy of larger volumes in underestimated compared to smaller volumes. In order to get an estimate of this the $E_{\mathrm{lattice}}(V)$ can be calculated, where E_{lattice} is the lattice energy, or the total energy of the cell less the energy of the individual ions at infinite separation. The separated ions can be calculated separately for each point on the curve, embedded in a crystal of ghost atoms, with the required lattice parameter.

Of course the quantity of interest so far has been the cohesive energy, and it would be expected that if the separated atoms are calculated surrounded by ghosts the same effect could be achieved. This is not the case in this work. While the basis on the lithium provided a good description of the ion in the bulk its description of the isolated atom was so bad as to be unusable even when embedded in the ghost lattice.

Hence the cohesive energy was obtained by first calculating the lattice energy. This is defined as the energy of the bulk in excess of the energy of the free ions (Li^+ and F^- in this case). For this the energy of the free ion was calculated with the ion surrounded by the appropriate ghost atoms using the CRYSTAL code. This lattice energy converged much faster with respect to basis set than the cohesive energy.

Then the first electron affinity of fluorine and the first ionization energy of lithium were calculated at the Hartree–Fock level using the aug-cc-pV6Z and aug-cc-pV5Z basis respectively. For this the Molpro code was used. By adding these results to the lattice energy the cohesive energy was recovered.

Results

Table 4.25 shows the convergence of these properties with respect to the fluorine basis. The BSSE corrected and uncorrected results bound the lattice parameter from above and below allowing us to claim an accuracy of 0.002 Å on the equilibrium lattice parameter. The cohesive energy also would appear to be within ~ 5 meV of the basis–set limit, which was estimated using the extrapolation formula of Equation 2.14.

Table 4.25: Calculated Hartree–Fock lattice parameters and cohesive energies of LiF using systematically improved basis functions on the fluorine atom. The cohesive energy is calculated at 4.005 Å in all cases.

	a_0 / Å	a_0 / Å (BSSE corrected)	E_{coh} / eV
pc-0	3.8481	4.0632	−7.05115
pc-1	3.9669	4.0563	−6.84276
pc-2	3.9989	4.0016	−6.80693
Extrap.	-	-	−6.79949

4.5.4.2 Surface

Using the new basis–set hierarchy the protocol of Section 4.4.4 is followed.

In summary the total energies of slabs with increasing thickness are calculated using CRYSTAL. These energies are fitted to Equation 1.3 with the first few thin slabs discarded until σ has converged. In this case slabs of 4-7 atomic layers were used, which led to convergence of better than 1×10^{-4} J m^{-2}. As before it was necessary to include two layers of 'ghost' atoms above and below each slab to ensure proper convergence with respect to basis–set, going from one layer of ghost atoms to two layers changed σ by ~ 0.007 J m^{-2}.

Table 4.26 shows the convergence of σ with respect to basis–set. The best value for σ, 0.332 J m^{-2} should be accurate to 0.001 J m^{-2} which is more than sufficient for the comparison.

Table 4.26: The convergence of the Hartree–Fock surface formation energy of LiF at 4.005 Å with respect to basis–set. All energies in J m^{-2}.

Basis	σ / J m^{-2}
pc-0	0.4068
pc-1	0.3364
pc-2	0.3317

4.5.5 Comparison

Table 4.27 lists the DMC results for the bulk and surface and the full quantum chemistry results obtained through the hierarchical calculations of Nolan and the Hartree–Fock calculations of the previous section.

As with the bulk calculations an estimate of the zero–point effects can be obtained from DFT. Using the same computational set–up as Section 4.1.2.3 the zero–point energies of slabs of atomic thickness 3–6 was calculated. As usual by fitting these data to Equation 1.3 and estimate of σ_{ZPE} was obtained of 0.029 J m^{-2}. This is included in the values of Table 4.27 allowing a full comparison to be made with experiment.

Table 4.27: Comparison of calculated and experimental properties of lithium fluoride.

Method	a_0 / Å	B_0 / GPa	E_{coh} / eV	σ / J m^{-2}
DMC	4.004(7)	64(2)	−8.827(4)	0.45(3)
Quantum Chemistry (frozen core)	4.049	71.6	−8.705	0.44
Quantum Chemistry	4.010	74.6	−8.857	0.50
Experiment	4.005	69.8	−8.800(14)	0.48(5), 0.34

The DMC calculations and the quantum chemistry calculations agree very closely with each other and experiment for the lattice parameter and bulk modulus as would be expected. In this case it appears that the error made by the pseudopotential approximation is negligible. The same is also true for the

cohesive energy where the DMC calculations out perform the quantum chemistry reproducing the experimental value to within 30(15) meV. It is thought that the error on the quantum chemistry is due to the difficulty in converging the correlation part with respect to basis–set, particularly in the case of the diffuse tails of the fluorine atom [140].

As was the case with LiH the quantum chemistry calculations indicate that for σ the contribution from the explicitly correlated cores on both the lithium and fluorine core are $\sim 0.06\ \mathrm{J\,m^{-2}}$. With the DMC and quantum chemistry agreeing to within 0.05(3) $\mathrm{J\,m^{-2}}$ it is possible to make a more definitive judgment on the disagreement within the experimental numbers.

It is quite clear that Burns and Webb did indeed produce the better estimate suggesting the Gilman number is in error by 41%. Indeed if the effect of surface relaxation (which is not included in either the DMC or quantum chemistry methods) is estimated from the DFT work in Table 4.5 the computational estimates would be reduced by $\sim 0.02\ \mathrm{J\,m^{-2}}$ giving even better agreement with the Burns and Webb estimate.

Chapter 5

The Adsorption of Water on Ionic Surfaces

Having demonstrated the need and possibility of going beyond DFT in order to achieve highly accurate surface energies in the previous chapter now we move on to the case of surface adsorption.

Section 1.3 outlined the importance and basic ideas behind the surface adsorption process. We also outlined why accurate surface adsorption energies are needed and the evidence that DFT can be sorely lacking. While lots of DFT work has been done on the computationally demanding case of molecules on heavy metal surfaces we have adopted the water on LiH(001) system due to the fact the relatively few electrons mean that more exhaustive studies using higher levels of theory can be performed.

The chapter is split into three separate sections with the first reporting the results of our DFT calculations of the water–LiH(001) adsorption energy. As before these have been carried out with a variety of different xc functionals and calculations of both the adsorption energy and the adsorption energy curve show a striking xc functional dependence highlighting the need to go beyond DFT.

In the second section we turn again to QMC methods to give a more accurate prediction of the adsorption energy and adsorption energy curve. Exhaustive timestep and finite–size testing is presented showing the very high level of convergence achieved.

Finally as outlined in Section 1.3.2 several calculations have shown it is possible to obtain accurate adsorption energies using quantum chemistry methods. We present calculations that use both the cluster methods of Li *et al.* and the incremental scheme of Paulus and Stoll [40] in order to produce an adsorption energy curve accurate to a few meV.

Convergence with respect to basis–set, level of theory and cluster size is obtained in the manner of Li *et al.*. A collaboration with the Paulus group of the Free University of Berlin has then allowed us to use the incremental scheme to calculate CCSD and CCSD(T) energies directly on otherwise unfeasibly large clusters.

In combination with the QMC data we show clearly that both QMC and quantum chemistry methods can indeed be used to produce adsorption energies to almost an order of magnitude better than chemical accuracy for extended systems, presenting a significant improvement upon DFT.

5.1 DFT

The DFT work falls naturally into three parts:

The first part is needed not just for all of the DFT work that follows but for all of the calculations using other levels of theory. This is determining a physical adsorbed water–LiH geometry which can serve as the test system. This has to be done in DFT due to the lack of any efficient means of ionic relaxation within QMC or the large scale quantum chemistry calculations. As the only real interest in the water on LiH system is for benchmarking the performance of individual levels of theory against each other the absolute accuracy of this geometry is not of much concern, just that it represents a reasonable water–LiH configuration. Hence the qualitative accuracy of GGA DFT should be good enough.

Secondly a series of DFT water–LiH interaction curves need to be plotted for several different functionals. These serve the same purpose as the DFT calculations of Section 4.1 namely to establish the level of exchange–correlation functional dependence for the interaction energy and provide a sequence of references with which the higher levels of theory can be compared.

Finally the issue of dispersion and dispersion corrected DFT is briefly explored using the semi–empirical correction scheme of Grimme *et al.* [141].

The VASP code was used throughout unless otherwise indicated.

5.1.1 Geometry calculation

As has been highlighted the absolute accuracy of the geometry was of secondary importance. Due to this and the very small surface relaxation effects discovered for the LiH(001) surface in Section 4.1.3 it was decided to only relax the water molecule, leaving the LiH ions in their bulk terminated positions. This has the advantage of keeping the geometry consistent when different sized supercells or fragments are used in the QMC and quantum chemistry calculations.

A starting point for the geometry was that of water on MgO [142]. Here water on MgO was studied with the PBE functional. The adsorbed water geometry found was the oxygen atom centered over the surface magnesium with the two O–H bonds pointing towards the two adjacent surface oxygen atoms slightly tilted down with respect to the plane of the surface.

Hence for the starting geometry the water molecule was placed in approximately this position (substituting lithium for magnesium and surface hydrogen for surface oxygen). A two–layer 4×4 supercell (32 atoms per layer) was used with a 32 Å vacuum gap[1]. The PAW potentials and cut–offs of Section 4.1 were also used along with Γ–point sampling.

As expected the water did relax into the water–MgO like configuration. Several different starting geometries (with the water further from the surface and slightly distorted) were tried but the final configuration appeared quite stable. Figure 5.1 shows the water geometry. As can be seen in the plan view there is a diagonal plane of symmetry that can be exploited in the calculations.

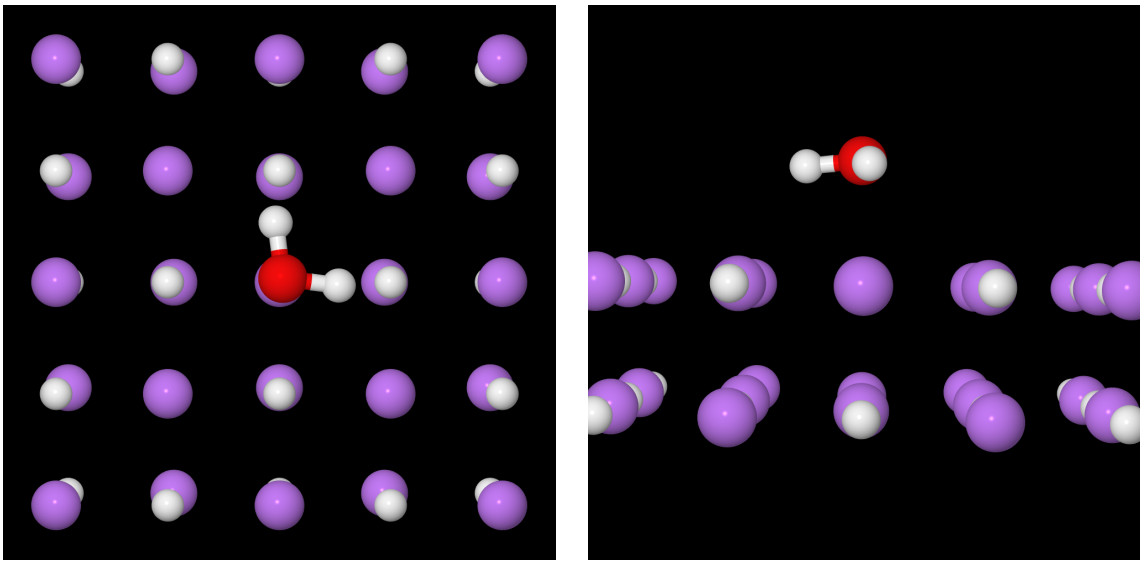
5.1.2 Adsorption energies

With the test geometry now established it was possible to go about calculating the adsorption energy curves with the different functionals.

Obviously in order to calculate this convergence tests had to be done with respect to the following parameters:

[1] Measured between the upper surface of the LiH slab to the lower surface directly above it.

Figure 5.1: The plan and elevation view of the adsorbed water–LiH geometry, demonstrated on a two–layer 5×5 fragment of LiH.



1. Surface supercell size. Since Γ–point sampling is being used if the surface supercell is too small then finite–size effects in the surface will be large. Also the interaction of the water molecule with its periodic image will be large. Both of these factors would lead to errors in the adsorption energy.

2. Thickness of the slab used. As seen in Section 4.1.3 if a slab is too thin it does not properly represent the case of a surface bounding an infinite solid.

3. Vacuum gap. Again familiar from Section 4.1.3 the plane–wave basis requires a spurious periodicity in the z–direction, if the vacuum gap is not large enough this spurious interaction is large enough to cause errors.

5.1.2.1 Defining adsorption energy

A lot of these effects can be mitigated through the judicious choice of how the adsorption energy is calculated. As described in Section 1.3.1 the adsorption energy is defined as the difference in energy between the interacting system (here the water on LiH) and its non–interacting fragments (water and the LiH surface). The most naïve calculation would be to take the energy of the fully converged interacting system less the energies of the fully converged fragments.

However much as with surface energies if all three systems can be calculated in a similar way cancellation of errors means that individual energies that are not fully converged can combine to give an adsorption energy that is converged.

One way of doing this is to calculate the adsorption energy as follows. First the energy of the interacting system is taken with a given surface supercell, number of layers and vacuum gap. Then the two fragments are calculated using exactly the same supercell, so the water is placed in the same

box as the full system and its energy evaluated. This is the repeated with the plain LiH surface. This means any spurious interactions present in the interacting system at a given set of parameters should be approximately present in the isolated fragments. Hence when the difference is taken the spurious interactions should cancel out. This is the method adopted for the DFT calculations.

5.1.2.2 Convergence

With that in mind Table 5.1 shows the convergence of the adsorption energy of water on a two–layer slab with a 20 Å vacuum gap calculated using the PBE functional with different surface supercell sizes.

Table 5.1: Convergence of the water–LiH adsorption energy with respect to surface supercell size calculated using DFT.

Surface	E_{ads}/ eV
1×1	0.16697
2×2	−0.60613
3×3	−0.20259
4×4	−0.21217
5×5	−0.21189
6×6	−0.21182

The convergence is relatively quick with a 4×4 supercell showing better than 0.5 meV convergence. The effects of the water–water interactions can be seen with the 1×1 and 2×2 cells which are significantly in error.

Adopting the 4×4 supercell the convergence with respect to slab thickness was then tested with the results listed in Table 5.2.

Table 5.2: Convergence of the water–LiH adsorption energy with respect to slab thickness calculated using DFT and a 4×4 surface supercell.

Slab thickness / Atomic layers	E_{ads}/ eV
2	−0.21217
3	−0.21202
4	−0.21201
5	−0.21206
6	−0.21214
7	−0.21213

This shows that by two layers $E_{\mathrm{ads}}$is already converged to better than 0.1 meV justifying the use of two layer slabs in the geometry generation and indicating two layer slabs are sufficient for the remaining calculations.

Similar tests on the vacuum gap indicated that with a gap of 12.3 Å the convergence achieved was of the same order as the other parameters.

This means that the adsorption energy at the equilibrium water–LiH distance, E_{ads} for the PBE functional is -0.211 eV

5.1.2.3 Other functionals

In order to test exchange–correlation functional dependence E_{ads} was then evaluated using the same DFT functionals as were used in Section 4.1. The motivation for using these functionals in this case was much the same, with them representing two of the most popular functionals along with others that have been optimized for solids or surface adsorption. These results are tabulated in Table 5.3.

Table 5.3: DFT functional dependence of E_{ads}. All calculations used the same PBE reference geometry. All calculations were performed with VASP apart from the BLYP estimates which used the PWscf code.

Functional	E_{ads}/ eV
LDA	-0.475
PBE	-0.211
RPBE	-0.090
WC	-0.273
BLYP	-0.006

The work done in Section 4.1 indicated that the convergence parameters did not depend on the functional. This is to be expected since the LDA and GGA approaches are both quite short ranged suggesting that the spurious self interactions should all converge with roughly the same rate. Hence the same parameters were used for all of the calculations. Obviously were hybrid–functionals or meta–GGAs to have been used it would be expected that extra convergence testing would have had to be done.

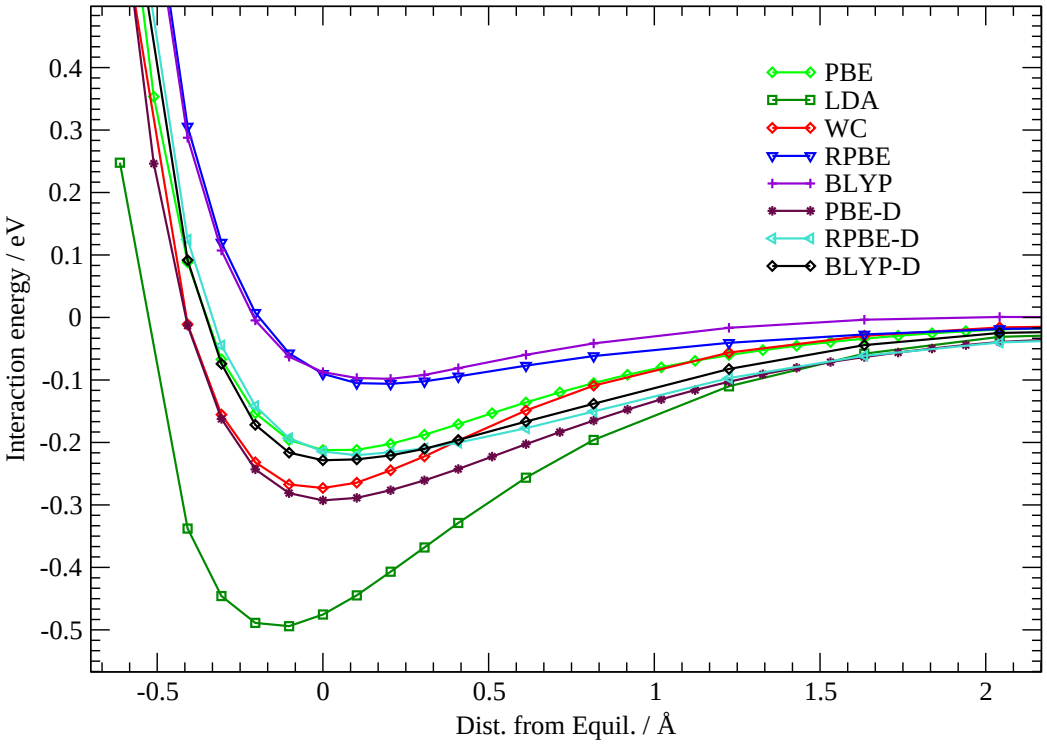
Table 5.3 also shows an estimate calculated using the BLYP functional [143, 56]. This was included due to its popularity for use in systems including water and to provide a GGA that was not based upon the PBE GGA. As BLYP is not implemented in VASP the PWscf code was used for these calculations. The convergence parameters were assumed to be the same as the PBE case but the norm–conserving pseudopotentials, Li.blyp-s-mt, H.blyp-vbc and O.blyp-mt from the PWscf distribution were used. These are all–electron in their treatment of lithium and hydrogen and reduce the oxygen atom to a six electron problem. An energy cutoff of 680 eV was enough to obtain convergence of E_{ads} to ~ 0.3 meV.

By shifting the water molecule in the z–direction and recalculating the adsorption energy at that point an adsorption energy curve could be produced. Figure 5.2 shows this for all of the functionals studied.

5.1.3 Dispersion correction

As can be seen from the adsorption energies of Table 5.3 and the adsorption energy curves of Figure 5.2 the agreement between the functionals is spectacularly bad. Not only do they not agree on the adsorption

Figure 5.2: The water–LiH adsorption curve calculated using different DFT functionals. The equilibrium point is defined from the PBE equilibrium geometry. –D indicates the dispersion correction of Grimme *et al.* has been used.



energy or even the equilibrium distance but the shape of the curves appears to be qualitatively different as well.

Part of the problem might be expected to be the lack of dispersion forces within LDA and GGA DFT. With these functionals being local in the electron density it is impossible for them to correctly describe the long range induced dipole interactions (normally referred to as London dispersion forces) which have a characteristic form of r^{-6}. If these forces are important in the adsorption of water on LiH then the errors in the DFT treatment might arise from this neglect.

Several schemes have been proposed for introducing dispersion into the DFT framework. [144, 145, 146, 141] Perhaps the most formally correct approach, from the point of view of an *ab initio* computational scientist is that of the van der Waals DFT (vdW–DFT) method [144, 145]. Here the semi–local parts of the GGA functional are used to explicitly introduce a r^{-6} dependence in order to correct the functional itself. Klimeš *et al.* [147] have demonstrated the power of this method when coupled with a suitably optimized functional but the lack of an accessible implementation led to it being discounted for this work.

A simpler if not fully *ab initio* approach is that of Grimme *et al.* [141] where a *post hoc* correction to the DFT total energy is made after the SCF cycle is converged. This correction is functional and atomic species specific and is derived from a parametrized set of TD–DFT and experimental observa-

tions. Grimme's own code was used to evaluate the corrections using large ten layer 20×20 fragments to ensure convergence. These calculations took minutes to run and the results are shown in Figure 5.2.

These corrections do indeed seem to increase the agreement of their respective functionals which are now just about within chemical accuracy with each other. The price paid for this is heavy however, these are no longer true *ab initio* calculations[2] and it is not clear whether the functionals are being corrected towards the true curve or just towards an different incorrect answer. It is again clear that higher levels of theory are needed to arbitrate in this problem.

[2]Though efforts have been made by Tkatchenko and Scheffler [148] to put this type of correction on a non–empirical footing.

5.2 H_2O on LiH with quantum Monte Carlo

As with the case of surface energies in Chapter 4 DFT is found to be lacking. The issue of exchange–correlation functional dependence has, as expected, reduced its predictive powers drastically. Obviously one solution to this problem is to resort to DMC to provide a form of benchmark that can be used to start building some sort of consensus of theoretical predictions.

The main idea behind calculating the adsorption energies remains the same, differences between the total energies of the adsorbed water system and the non–interacting system are sought. As before the PBE optimized geometry is adopted for the H_2O–LiH complex. With that in mind the usual convergence tests of the QMC energies are required.

5.2.1 Trial wavefunction

The usual Slater–Jastrow form of the trial wavefunction was assumed with the Slater determinant generated using plane–wave LDA–DFT calculations within PWscf. The wavefunctions were then re–expanded in blip functions. The familiar Trail and Needs Dirac–Fock pseudopotentials [120, 121] were used in order to reduce the computational effort of the DMC calculations.

With these a plane–wave cut–off of 3400 eV was found sufficient for the DFT calculations and the natural blip grid spacing produced a trial wavefunction with an acceptable variance.

5.2.1.1 Jastrow factor

Since the LiH calculations of Section 4.2.1 required the presence the u, χ and f terms in the Jastrow factor it was assumed that a Jastrow factor of this complexity would be needed in this system also. Table 5.4 shows the results of the variance optimization of different Jastrow factors of this form within VMC.

Table 5.4: VMC optimized total energies for an H_2O molecule on a two–layer 3×3 LiH surface supercell with different Jastrow factors. The Slater determinant used is based on an energy cut–off of 3.4×10^3 eV with single blip spacing.

Expansion order			VMC Energy / E_h	VMC Variance / ${E_\mathrm{h}}^2$
u	χ	f		
2	2	2	−32.65(1)	1.22(8)
4	2	3	−32.694(5)	1.03(2)
4	4	3	−32.710(4)	1.01(2)
4	6	3	−32.712(5)	1.04(2)
6	4	3	−32.708(5)	1.03(2)

On this basis the expansion orders of 4, 4 and 3 are chosen for the u, χ and f terms of the Jastrow factor. Energy minimization in this case does reduce the energy by a significant amount, to −32.740(5) E_h with a variance of 1.28(3) E_h^2. Again the increase of the variance when using energy minimization is noted.

Hence all of the Jastrow factors in the subsequent calculations were of this form and their parameters were re–optimized first using variance minimization then energy minimization for each separate calculation.

5.2.2 Timestep and the adsorption energy

As with any DMC calculation it was also necessary to determine what timestep would provide the required trade–off between accuracy and computational efficiency.

As mentioned in Section 5.1.2.1 it is possible to chose a particular way of defining the adsorption energy so as to maximize possible cancellation of errors. That section was concerned mainly with the finite–size errors that occur within single–particle theory. As the DMC calculations are also subject to the same finite–size errors (interaction between neighboring water molecules) and more (many–body finite–size errors in the surface) it might be expected that the same prescription is followed. This will be termed the 'fragment' approach, where the adsorption energy is calculated by the difference between an interacting system (H_2O on LiH) and two fragments (H_2O and LiH) each in identically sized and shaped supercells.

Another definition however allows the timestep error of the DMC calculations to also be the subject of the error cancellation. This will be termed the 'separated' definition where the difference is between the interacting system of H_2O on LiH and the system where the H_2O–LiH distance is suitably far apart. The DFT calculations indicate that a distance of 10 Å is more than enough, leaving for a significant margin for error.

Figure 5.3 shows the timestep dependence for both the separated and fragment approaches when calculated using the locality approximation. Also calculations with fragments using Casula's T–move approach [89] are shown.

Considering the locality approximation for just a moment it is clear that for very small timesteps the separated and the fragment methods produce identical adsorption energies. This is what would be expected. Also at 0.02 a.u. it is clear that the cancellation of errors with the separated scheme is superior to the fragment scheme showing convergence to the assumed zero timestep value while the fragment scheme is still in error by ~ 60 meV. Overall these tests indicate that by 0.02 a.u. the locality–separated scheme has plateaued and can be taken as a reasonable approximation to the zero timestep adsorption energy.

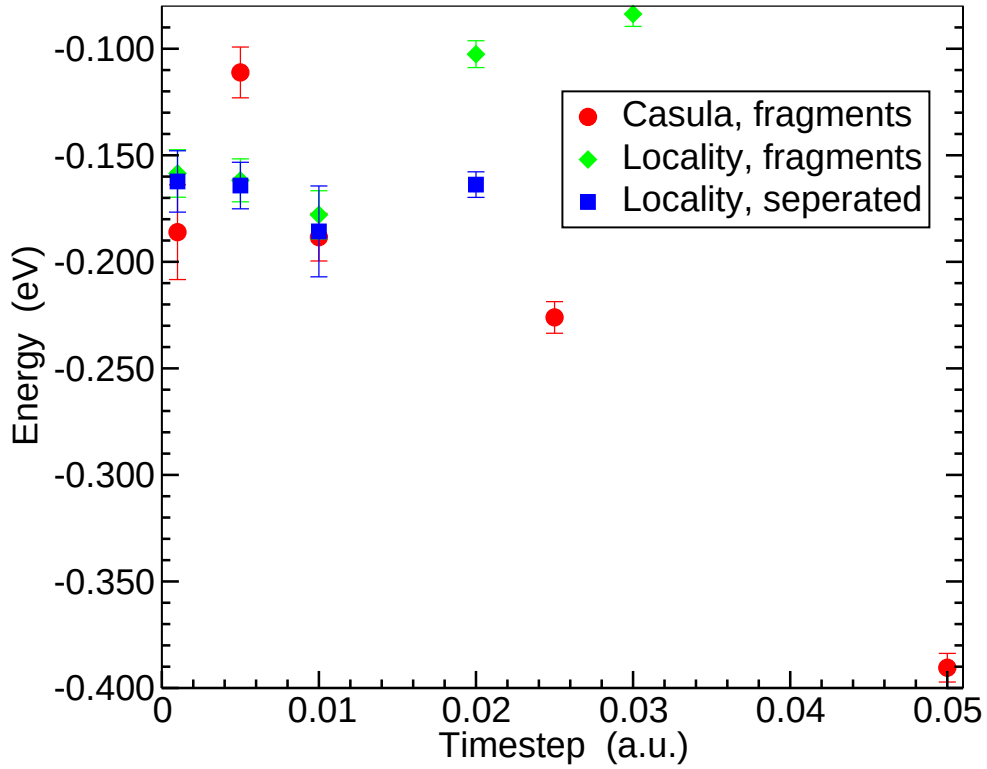
The T–move scheme of Casula does, as expected, show a much larger timestep dependence. More worryingly it also shows an erratic behavior around 0.005 a.u. that makes predicting convergence very difficult. This is in all likelihood a demonstration of the lack of size–consistency recently highlighted by Casula *et al.* [90]. For these reasons the timestep of 0.02 a.u. with the locality–separated scheme was adopted for the remaining DMC calculations.

5.2.3 Finite–size

The finite–size effects can first be divided into two categories, those that are dependent on slab thickness and those dependent on the area of the surface supercell.

Beginning with the slab thickness it is first useful to explore where these errors arise from. The

Figure 5.3: Plot of the DMC adsorption energy for the two–layer 3×3 supercell against timestep. Both the locality approximation and T-moves were examined. Also the adsorption energy was calculated using fragments and again using a 10 Å seperation. See text for details.



real world situation that the modeling is attempting to capture is that of a molecule on the surface of a large fragment of crystal. To all intents and purposes the crystal is assumed to extend to infinity across half of space. Hence the molecule is only interacting with the surface and the bulk beneath it. In the computational calculations the infinite extent of the crystal is approximated by the periodic boundary conditions in two dimensions and by having a set number of atomic layers making up the slab.

The error introduced by the latter approximation is that if the slab is too thin then the lower surface of the slab will affect the electronic structure of the bulk like material and the upper surface. This in turn will affect the interaction of the molecule with these parts. Also the lower surface itself could be interacting with the molecule, an interaction that is simply not present in the real world case. Obviously in the limit of an infinity thick slab the real–world case will be recovered, hence by increasing the thickness of the slab the adsorption energy should converge.

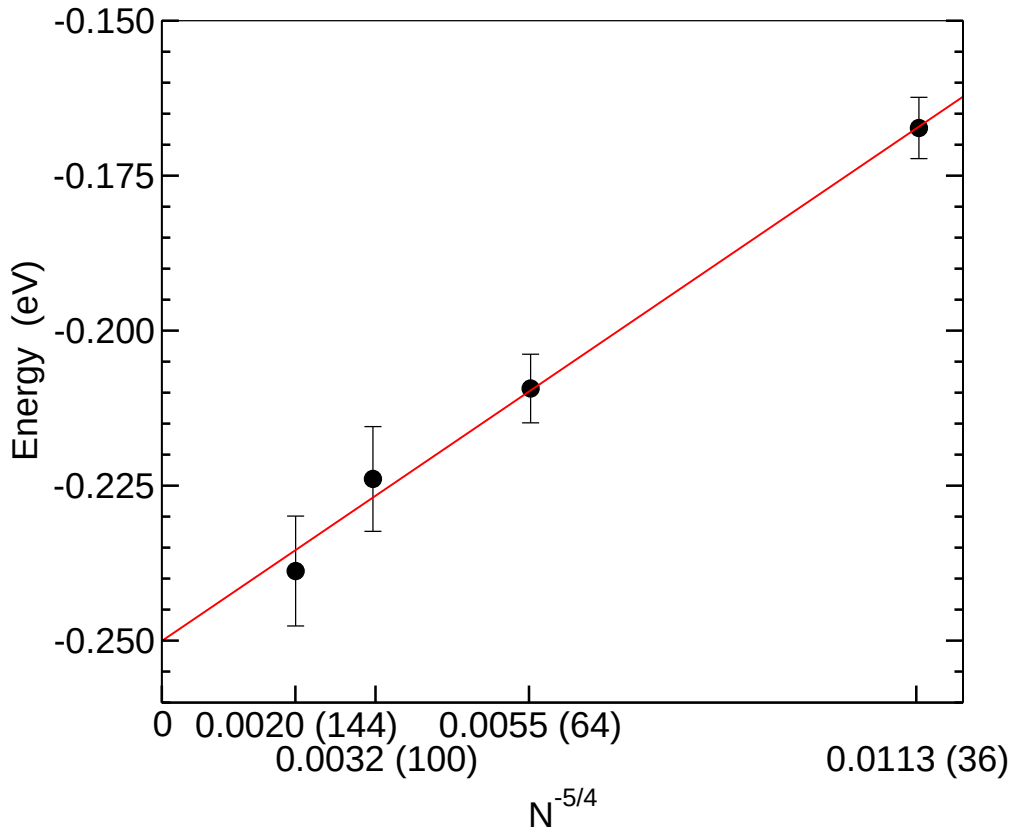
It is noted that the DFT calculations of Section 5.1.2.2 indicate that the adsorption energy is converged to within less than 0.1 meV of its limiting value with only two–layers in the slab. This finding is repeated in the Hartree–Fock calculations of Section 5.3.5 and in the correlated quantum chemistry estimates of Section 5.3.3.2 an error of $\sim$ 3 meV is suggested by limiting the slab to two–layers. As

the finite–size effects for all four cases (DFT, HF, post–HF and DMC) arise due to the same reason, spurious interactions with the lower surface, it is reasonable to assume the errors in DMC will be of the same magnitude. These errors are below the statistical error of 5 meV which is attained in the DMC calculations and hence two–layer slabs are adopted.

The finite–size errors associated with the surface supercell are not as easy however. DFT, HF and DMC, in PBCs, all suffer from a finite size error related to the water and its periodic image interacting through its dipole moment[3]. The DFT calculations show that this effect is fairly short ranged and the adsorption energy is converged to within a few tenths of an meV by the 4×4 supercell. However as has been discussed the DMC method also includes many–body finite–size errors which occur due to the self–correlation between electrons and their own periodic images. This is also present to some extent within HF theory and is why the Hartree–Fock adsorption energy converges slower than its DFT counterpart.

Figure 5.4 shows the result of calculating the adsorption energy at the PBE equilibrium geometry within DMC. Several two–layer surface supercells are used.

Figure 5.4: DMC adsorption energy at different surface supercell sizes for the two–layer slab. N is the number of atoms in the slab for the given supercell. The numbers in brackets are N. The line shows the fit to Equation 3.36.



[3]The post–HF calculations do not make use of PBCs and their only finite–size error is based on spurious interactions with the molecule and edges and corners of the LiH fragments. Much like the slab thickness finite–size error.

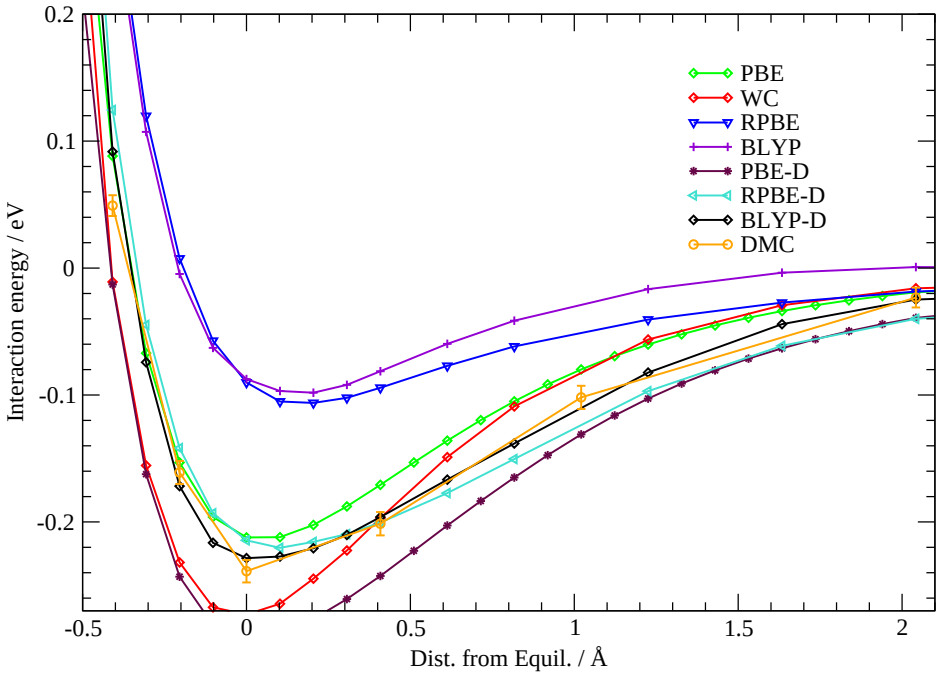
It is striking just how slowly the DMC energy converges with respect to surface supercell. Due to the rapid convergence of the single–particle finite–size effects and the dipole–dipole interaction within DFT it seems fair to assume that this convergence is due to the many–body finite–size effects. Subsequent calculations for H_2O on MgO [149] have shown DMC having a similar convergence to DFT, therefore it is likely that the effect of the water molecule on the LiH surface sufficiently alters its electronic structure in particular the exchange–correlation hole. This does not increase the absolute magnitude of the finite–size errors per se, but it does remove a lot of the cancellation of errors expected when the interacting and non–interacting systems are subtracted from one another.

With this in mind it seems sensible to fit the data to Equation 3.36 in an attempt to extrapolate to infinite system size. This yields an adsorption energy, E_{ads} of 250(7) meV which is quoted as the best estimate for the DMC method.

5.2.4 Adsorption curve

A four point extrapolation for every point on the adsorption curve would be very costly. It would also be of dubious value given the value using the 6×6 surface is 238(9) meV, in error by only 12(11) meV. For this reason the adsorption curve was generated solely using the 6×6 surface supercell. Figure 5.5 shows the DMC curve compared to some of the DFT functionals.

Figure 5.5: The water–LiH adsorption curve calculated using DMC. Selected data from the DFT calculations are also presented to compare with the DMC curve. The PBE equilibrium geometry is used throughout.



The DMC work seems to agree quite closely with the semi–empirical BLYP–D calculations. However as in the surface case these values are subject to both pseudopotential error and fixed–node error. It is clear that in order to obtain a convincing estimate our second appeal to theory beyond DFT must also be made.

5.3 H_2O on LiH with incremental coupled–cluster

Chapter 4 has demonstrated the amazing agreement that can be obtained between DMC and high–level quantum chemistry. The surface energy of both LiH and LiF seem to be described to a very high level of accuracy by both DMC and CCSD(T) methods. It seems reasonable to expect that the CCSD(T) estimate of the adsorption energy of a water molecule on the LiH(001) surface would also be very accurate. Therefore this would seem a good choice of method with which to build a consensus with DMC.

Unfortunately Chapter 4 also indicates the possible problems with this method. Periodic implementations of coupled–cluster theory are currently limited to MP2 with CCSD present in the experimental version of VASP [150, 151]. The hierarchical method was used in order to overcome this limitation and apply CCSD(T) and higher methods to the bulk and surface but it is not immediately clear how this method could be applied to adsorption problems.

As has already been outlined there are two related methods by which adsorption energies have already been calculated using quantum chemistry.

The first of these is a cluster approach. This is perhaps best represented by the work of Li *et al.* [47]. The adsorption energy is split into the Hartree–Fock and correlation contributions. An MP2 calculation is performed to obtain the correlation part of the adsorption energy ($E_{\mathrm{ads}}^{\mathrm{corr}}$) of the molecule on a crystal cluster large enough to accurately reproduce the infinite system. In this calculation the basis–set limit is reached (usually through extrapolation). Corrections are then calculated as the difference between the MP2 value and the CCSD(T) value (so–called δCCSD(T) corrections) using a smaller, more manageable basis thus improving upon the MP2 estimate and bringing it closer to a true CCSD(T) estimate.

The Hartree–Fock part ($E_{\mathrm{ads}}^{\mathrm{HF}}$) is also estimated by the adsorption of the molecule on a small cluster of solid. Using this approach Li *et al.* were able to calculate the adsorption energy of water on NaCl(001) to better than 45 meV.

The second method is the incremental scheme of Paulus and Stoll. Originally used to provide estimates for bulk systems Paulus and Rosciszewski [43] and Müller *et al.* [41, 42] have applied it to adsorption problems. Again the adsorption energy is partitioned into a correlation part and a Hartree–Fock part. The Hartree–Fock part is this time calculated in PBCs with a Gaussian basis–set. An estimate of $E_{\mathrm{ads}}^{\mathrm{corr}}$ using CCSD or CCSD(T) is then obtained in one bound using a single cluster by calculating the contributions of a number of representative orbital groups[4] within a suitable basis–set.

Both of these methods also make full use of point–charge embedding to reduce the size of the cluster(s) required.

5.3.1 Our calculation schema

Both methods appear capable of achieving adsorption energies to around chemical accuracy. However our aim is to be much better than chemical accuracy so we have therefore combined the two.

In the first instance we intend to present an estimate of the adsorption energy at the PBE equilibrium

[4] This process is fully explained in Section 5.3.4.1

geometry to the highest level of accuracy we can obtain.

Then we shall present the complete adsorption energy curve, necessarily calculated with a slightly reduced accuracy.

5.3.1.1 Correlation

First let us look at $E_{\mathrm{ads}}^{\mathrm{corr}}$. It is useful to first state what each of the two methods, the cluster approach and the incremental scheme tells us.

The calculations of Li *et al.* make an important point about convergence with respect to cluster size. They show us that the difference in $E_{\mathrm{ads}}^{\mathrm{corr}}$ calculated with two levels of theory (in their case MP2 and CCSD(T)) converges quicker with cluster size than the absolute value of $E_{\mathrm{ads}}^{\mathrm{corr}}$ itself. The same is also true of the difference between frozen–core estimates and fully correlated estimates, and indeed between estimates using one basis–set compared to another.

The incremental scheme allows much larger clusters, or fragments to be accessible to a given level of theory by only including the correlation effect between sets of orbitals when it is of a significant level.

Both methods highlight the importance of some sort of embedding in improving the convergence with respect to cluster size.

It is clear that if our initial estimate within the cluster approach can be made more accurate (through the use of a bigger basis–set, larger fragment or higher level of theory) the corrections would necessarily be smaller reducing the overall error. So by using the incremental scheme on a large fragment with a high level of theory we can indeed start off with a much better estimate upon which to apply the corrections of the cluster approach.

Hence our schema for calculating $E_{\mathrm{ads}}^{\mathrm{corr}}$ is as follows:

- Use the incremental scheme to calculate $E_{\mathrm{ads}}^{\mathrm{corr}}$ using a 6×6×2 fragment (72 atoms) at the CCSD/V[DT]Z level.
- Use the incremental scheme to calculate a δCCSD(T)/V[DT]Z[5] correction using a 4×4×2 (32 atom) fragment.
- Calculate a core correction using traditional MP2 calculations. This is the difference between a fully correlated MP2 calculation on a 3×3×2 fragment and one performed using the frozen–core approximation.
- Calculate a basis–set correction defined as E(MP2/V[Q5]Z)$-E$(MP2/V[DT]Z), using a 4×4×2 fragment.
- Estimate convergence with respect to level of theory by calculating δCCSDT and δCCSDT(Q) corrections using small fragments in a small basis.

Obviously we must justify our choice of basis and fragment sizes in the above scheme, the majority of this section shall be concerned with reporting the results of these tests.

[5] Obviously here we define δCCSD(T) as E(CCSD(T))$-E$(CCSD).

All of the correlation calculations were performed using the Molpro quantum chemistry package [139] unless otherwise stated. These calculations were also the result of a collaboration with the Paulus group of the Free University Berlin. Their support and advice on the technical implementation of these calculations, the embedding scheme and the basis–sets adopted is gratefully acknowledged.

5.3.1.2 Hartree–Fock

In this work we have chosen to calculate the Hartree–Fock component of the adsorption energy using the VASP code. The PBC scheme offers a much more robust convergence of $E_{\mathrm{ads}}^{\mathrm{HF}}$ than a cluster approach, which Li *et al.* have demonstrated is very sensitive to the embedding scheme used.

Also the PAW formalism allows a much more controlled method of reaching the complete basis limit than the Gaussian formalism of the CRYSTAL code, which as Section 4.4.4 has shown require layers of ghost atoms above the surface. These would cause linear dependency issues with the inclusion of the basis from the adsorbed molecule.

Section 4.4.4 also shows that for the surface energy of LiH(001) the Guassian approach and the PAW approach agree to better than 1%. As we will see the HF contribution is only $\sim$ 10 meV making errors on this scale negligible.

5.3.2 Embedding

For all of the fragment calculations we perform the fragments are embedded in the surface of a $20\times20\times20$ lattice of point charges. This lattice is the same as the crystal lattice and each point charge is either $+1$ a.u. if it is a lithium site or -1 a.u. if it is a hydrogen site. For those sites adjacent to the face of the fragment the absolute charge is reduced to $\frac{1}{2}$, for an edge $\frac{1}{4}$ and for a corner $\frac{1}{8}$ a.u..

For an ionic system such as LiH this should represent a good simulation of the infinite system. Our convergence tests presented in Table 5.6 show a reasonable convergence using this scheme.

5.3.3 $E_{\mathrm{ads}}^{\mathrm{corr}}$ convergence

Most of the convergence tests with respect to fragment and basis–set are performed using MP2 theory. Previous work [47, 43] has suggested that MP2 theory agrees quite well with the more accurate CCSD(T) treatment, and our own tests shown in Table 5.6 will show that this is the case with this system. Therefore it seems reasonable to assume that the MP2 convergence will be qualitatively the same as that of CCSD(T)

As has been outlined the density–fitting approximation significantly reduces the cost of both Hartree–Fock and MP2 calculations while only incurring a small loss of accuracy. Tests showed that density fitting introduced an error in the adsorption energies of $\sim$ 0.2 meV at the worst.

As a result all of the traditional MP2 and HF calculations in this work utilize the DF approximation. As we also used the cc-pVDZ basis–sets of Dunning *et al.* [137, 152, 153] the appropriate cc-pVDZ/JKFIT [72] auxiliary basis were adopted for the DF-HF calculations[6]. The cc-pVDZ/MP2FIT [154, 155] basis were adopted for the DF-MP2 calculations.

[6]The cc-pVDZ/MP2FIT basis was used for the lithium atoms as no JKFIT basis is available

5.3.3.1 Basis–set convergence

Due to the ionic nature of lithium hydride crystal the outermost angular momentum function on the lithium atom was removed to increase computational efficiency. The error introduced by this is negligible as shown in Table 5.5.

Table 5.5: Adsorption energy calculated at the MP2 level with an embedded 4×4×2 fragment for a variety of basis–sets. The reduced basis–set is cc-pVxZ with the outermost angular momentum function on the Li atom removed. Full refers to the standard cc-pVxZ basis and augmented the aug-cc-pVxZ basis. All energies are quoted in meV.

Basis	Reduced	Full	Augmented
VDZ	−90.42	−125.25	−156.87
VTZ	−166.78	−176.41	−189.58
VQZ	−191.95	−194.18	-
V5Z	−199.49	−199.89	-
V[DT]Z	−198.97	−197.96	−203.35
V[TQ]Z	−210.43	−207.16	-
V[Q5]Z	−207.35	−205.88	-

The notation V[xy]Z refers to the usual two point extrapolation of VxZ and VyZ with Equation 2.15.

We can also see from these data that our reduced basis using a V[Q5]Z extrapolation is enough to get to within a few meV of the basis–set limit. The largest fragment for which these calculations were feasible was 4×4×2 and so this was the fragment size we adopted for our basis–set correction.

5.3.3.2 MP2 calculations on fragments

Table 5.6 shows $E_{\mathrm{ads}}^{\mathrm{corr}}$ calculated with increasing large fragments using MP2 theory and our reduced VDZ basis. For those systems where it is feasible full CCSD and CCSD(T) estimates have also been calculated.

Several conclusions can be drawn from these data. Firstly it seems that by the time we get to the 6×6×2 fragment the MP2 correlation contribution is converged to within a few meV. Secondly the convergence of the CCSD method seems to be roughly the same as that of MP2 theory, suggesting we are justified in performing our tests with MP2 theory. Thirdly the difference between CCSD and CCSD(T) is converged to better than 0.1 meV by the 4×4×2 fragment.

Finally it indicates that two–layer fragments are enough to obtain convergence of ~ 3 meV.

5.3.3.3 Additivity

There is an implicit assumption when we use MP2 theory in order to provide basis–set corrections and core corrections. Namely this is that the additivity principle holds. This is the idea that the difference between calculations using say two different basis–sets within MP2 theory is exactly the same as the

Table 5.6: $E_{\mathrm{ads}}^{\mathrm{corr}}$ calculated using a VDZ basis with embedded fragments of varying sizes. All energies are quoted in meV.

Fragment	MP2	CCSD	CCSD(T)
3×3×2	−75.70	−58.04	−70.89
4×4×2	−90.37	−69.72	−85.82
5×5×2	−93.06	−73.66	−89.69
6×6×2	−96.11	-	-
7×7×2	−97.17	-	-
8×8×2	−98.61	-	-
4×4×1	−74.21	−45.20	−57.12
4×4×2	−90.37	−69.72	−85.82
4×4×3	−91.02	−72.00	−87.59
4×4×4	−92.19	−72.74	−89.42
4×4×5	−90.97	-	-

difference between the same calculations when performed with CCSD(T) theory. Obviously this is an approximation.

We have analysed the error this approximation introduces by performing a number of calculations on the 2×2×2 fragment. These are summarized in Table 5.7 and indicate the errors are generally ~ 1 meV. The one case where this error is larger, ~ 7 meV relates to the manner in which we calculate

Table 5.7: Additivity errors in the adsorption energy, calculated using the embedded 2×2×2 fragment. $E_{\mathrm{VxZ}}^{\mathrm{CC}}$ refers to CCSD(T)/VxZ. CVxZ refers to the standard cc-pCVxZ basis. All energies in meV.

	Energy
$E_{\mathrm{V[Q5]Z}}^{\mathrm{CC}} - (E_{\mathrm{V[DT]Z}}^{\mathrm{CC}} + E_{\mathrm{V[Q5]Z}}^{\mathrm{MP2}} - E_{\mathrm{V[DT]Z}}^{\mathrm{MP2}})$	-1.31
$E_{\mathrm{V[Q5]Z}}^{\mathrm{CC}} - (E_{\mathrm{VTZ}}^{\mathrm{CC}} + E_{\mathrm{V[Q5]Z}}^{\mathrm{MP2}} - E_{\mathrm{VTZ}}^{\mathrm{MP2}})$	-7.40
$E_{\mathrm{CVTZ}}^{\text{ae-CC}} - (E_{\mathrm{VTZ}}^{\mathrm{CC}} + E_{\mathrm{CVTZ}}^{\text{ae-MP2}} - E_{\mathrm{VTZ}}^{\mathrm{MP2}})$	-1.14

the basis–set correction. Here we see that when going from VTZ to V[Q5]Z the additivity error is much larger than if we go from V[DT]Z to V[Q5]Z. As a result of this we adopt the latter route for our basis–set corrections.

5.3.3.4 Further contributions

In order to test convergence with respect to level of theory, the MRCC code of Kálay [156, 67] was used to calculate δCCSDT and δCCSDT(Q) corrections using the VDZ basis and the 3×3×2 and 2×2×2 fragments respectively. While these fragments and their associated basis–sets are rather small these calculations should be enough to give us an estimate of the convergence. The final values for these

contributions are shown in Table 5.10 and suggest that higher levels of theory are not going to improve the adsorption estimate by more than a few meV at the very most.

Finally all of our calculations are performed using the frozen–core approximation and thus neglect the correlation of the lithium and oxygen 1s cores. An all–electron MP2/CVTZ[7] calculation on the embedded 3×3×2 fragment gives us an estimation of the core–valence effect and the core contribution. This is again presented in Table 5.10 with its magnitude ($\sim$ 14 meV) suggesting that it is a small enough consideration to be included as a correction.

5.3.4 Incremental calculations

Table 5.6 suggests that a CCSD calculation using the 6×6×2 fragment and a δCCSD(T) (defined as $E_{\delta\mathrm{CCSD(T)}} = E_{\mathrm{CCSD(T)}} - E_{\mathrm{CCSD}}$) calculation with the 4×4×2 fragment, should give us a very accurate estimation of the correlation part of the adsorption energy of the full periodic system.

However we have seen from the basis–set convergence that these calculations need to be performed using both the VDZ basis and the VTZ basis. Unfortunately full CCSD/VTZ and CCSD(T)/VTZ calculations on fragments of this size are simply not feasible. Instead we must turn to the incremental scheme.

5.3.4.1 Theory

The process of any post–HF method is to take the (usually delocalized) Hartree–Fock orbitals and introduce electron correlation, usually through mixing the occupied and virtual orbitals. In this way the scaling of such methods is often determined by the number of occupied and virtual orbitals.

The insight behind the incremental scheme is to recognize that in real–space electron correlation is a fairly nearsighted effect, the correlation effect is much stronger between electrons that are near by. With this in mind the incremental scheme firstly deals with localized orbitals, that is normal Hartree–Fock orbitals that have undergone some sort of localization process in real space (in this work the scheme of Foster and Boys [157, 158] is used). This allows us to break down the correlation energy in the following way,

$$E_{\mathrm{corr}} = \sum_{i}^{N} \epsilon_i + \sum_{i>j}^{N} \Delta\epsilon_{i,j} + \sum_{i>j>k}^{N} \Delta\epsilon_{i,j,k} + \ldots + \sum_{i>\ldots>p}^{N} \Delta\epsilon_{i,\ldots,p}. \tag{5.1}$$

In this equation there are N localized orbitals, ϵ_i is the correlation energy of the i^{th} orbital, that is to say the change in energy relative to the Hartree–Fock solution generated by correlating the two electrons in the orbital, holding all of the other orbitals frozen at the Hartree–Fock level. $\Delta\epsilon_{i,j}$ is the two–body increment term, that is the subsequent change in energy of correlating all of the electrons in orbitals i and j with each other,

$$\Delta\epsilon_{i,j} = \epsilon_{i,j} - \epsilon_i - \epsilon_j. \tag{5.2}$$

Higher terms such as the three–body increment term are defined recursively, for example,

$$\Delta\epsilon_{i,j,k} = \epsilon_{i,j,k} - \Delta\epsilon_{i,j} - \Delta\epsilon_{j,k} - \Delta\epsilon_{i,k} - \epsilon_i - \epsilon_j - \epsilon_k. \tag{5.3}$$

[7] Here and throughout this chapter CVxZ refers to the standard cc-pCVxZ basis, used when the frozen core approximation is removed.

These individual calculations offer much more favorable scaling as they are only performed with the fraction of electrons of the full system. So while the number of virtual orbitals involved remains the same the number of occupied orbitals is greatly reduced [159].

Obviously if all the increments had to be calculated this scaling would still not be of very much use, but it turns out generally that most of the increments are very close to zero. Usually only orbitals close to each other in real space produce increments that have any real contribution, and generally only one and two–body increments are needed also.

In the best cases this can reduce the scaling of a method such as CCSD(T) to that of traditional MP2 theory [159].

Adsorption

Previous work on adsorption problems has been done with the incremental scheme by Müller *et al.* [41, 42] but there it was not combined with any additive corrections and there is no other benchmark data (apart from DFT) with which to estimate accuracy. Nevertheless they indicated that the incremental scheme was able to recover the full fragment energies to within a few meV.

The increments can be defined in a slightly different way to aid understanding [42]. Here the correlation part of the adsorption energy is given by,

$$E_{\mathrm{corr}}^{\mathrm{ads}} = \eta_M + \sum_{i}^{N} \eta_i + \sum_{i>j}^{N} \eta_{i,j} + \sum_{i}^{N} \eta_{M,i} + \sum_{i>j>k}^{N} \eta_{i,j,k} + \sum_{i>j}^{N} \eta_{M,i,j} + \ldots . \tag{5.4}$$

η_i represents the difference between increment i for the adsorbed system and the isolated surface, $\eta_i = \epsilon_i^{\mathrm{adsorbed}} - \epsilon_i^{\mathrm{surface}}$. This is easily generalized to the higher increments.

The subscript M refers to the orbital group on the adsorbed molecule so η_M is the difference in correlation of the orbitals on the molecule when the frozen HF orbitals and nuclei of the surface is present compared to when it is not.

Again Müller *et al.* found that these increment sums converged very quickly with respect to the distance of the orbitals from the molecule. Also they found that the error introduced by ignoring three–body and higher contributions was very small.

5.3.4.2 Practice

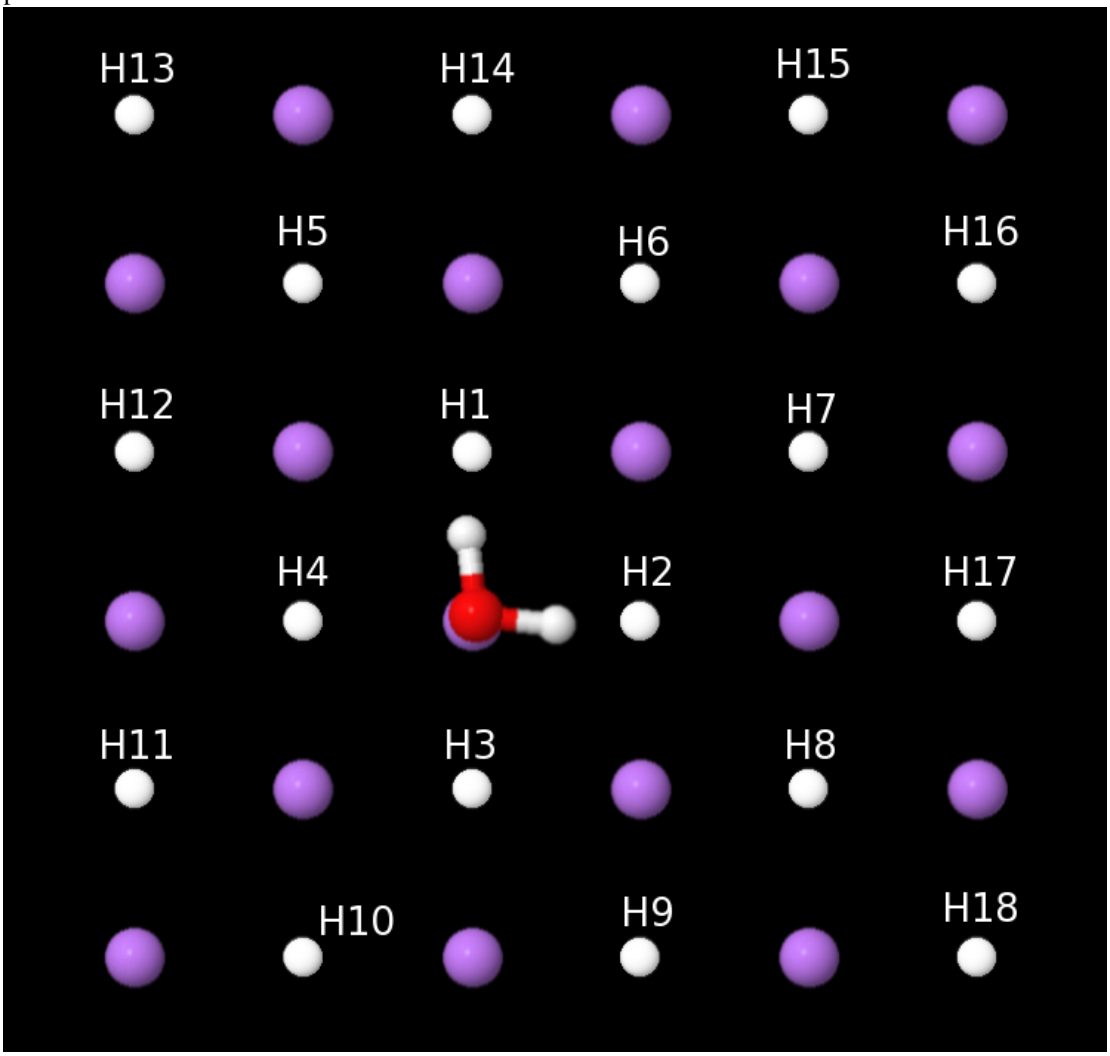
Since it was possible to calculate the full MP2/VDZ adsorption energies for all of the fragments we were interested in we used MP2 theory to indicate which were the most important increments to include in the calculation. Our method for doing this was similar to the energy screening approach of Friedrich *et al.* [160] and assumes that the most important increments when using MP2 theory are the same as the most important increments in coupled–cluster theory.

In order to maintain a consistent labeling of increments across fragment sizes the orbitals correlated in each increment were labeled according to the hydrogen atom lying at their orbital centres. Increments that correlated the orbitals on the water molecule were labeled H_2O.

The labels given to each hydrogen and hence each two electron orbital, are shown in Figure 5.6. With this system the increment H1-H2-H_2O for instance, would correspond to $\eta_{M,1,2}$, the correlation

part of the adsorption energy when just the orbitals on the H1 and H2 hydrogens and the water molecule are correlated.

Figure 5.6: Plan view of the 6×6×1 fragment. The labels correspond to the orbitals centered on the hydrogen atoms. Orbitals on the subsequent layers are labeled in the same manner with $100(n-1)$ added to the numeral of the orbital on the n^{th} layer. E.g. the hydrogen on layer 2, directly beneath the position of the water molecule is labelled H101.



Using the 4×4×2, 5×5×2 and 6×6×2 fragments our procedure was to first perform a Hartree-Fock calculation and then Foster–Boys localization in order to generate the localized orbitals. Then all of the one–body; H1 ... HN, one–body plus water; H1-H_2O ... HN-H_2O, two-body; H1-H2 ... HN-H$(N-1)$ and finally two-body plus water; H1-H2-H_2O ... HN-H$(N-1)$-H_2O increments were calculated at the MP2/VDZ level. By discarding all of the increments less than 0.27 meV the error in the total adsorption energy for each fragment was at worst 1.6 meV. The results of this screening are presented in Table 5.8.

Also as shown in Figure 5.6 the system included a plane of symmetry running across the diagonal of

the fragment. This was exploited so as increments that were identical by symmetry were only calculated once and then double counted.

The worst agreement came with the 4×4×2 fragment so this was used to test our assumption that the MP2 increments represented a good choice when moving on to coupled–cluster theory. The increments shown in Table 5.8 were recalculated for the 4×4×2 fragment at the CCSD(T)/VDZ level of theory. This gave an estimate of $E_{\mathrm{ads}}^{\mathrm{corr}}$ of -88.80 meV comparing to the full CCSD(T)/VDZ estimate for this fragment given in Table 5.6 of -85.82 meV. This suggests the incremental scheme is still only in error by a few meV in this worst case scenario suggesting our CCSD and CCSD(T) incremental estimates should be correct to within a few meV.

With this in mind two incremental CCSD calculations were then performed on the embedded 6×6×2 fragment using the modified VDZ and VTZ basis–sets. This gave us our starting estimate for $E_{\mathrm{ads}}^{\mathrm{corr}}$. The perturbative triples corrections were then calculated using the same increments on an embedded 4×4×2 fragment using the same basis, this gave us our δCCSD(T) correction. These two values are presented in Table 5.10.

5.3.5 Hartree–Fock contribution

As outlined we calculated the Hartree-Fock part of E_{ads} using VASP and as such these calculations proceeded in much the same way as the DFT calculations. Γ–point sampling was used with the standard PBE PAW potentials which had a recommended energy cut–off of 500 eV.

We found a vacuum gap of 16 Å was enough to reduce this effect to less than 0.03 meV. As before two layer slabs were found to converge E_{ads} to ~ 1.5 meV.

Table 5.9 shows the convergence of $E_{\mathrm{ads}}^{\mathrm{HF}}$ with respect to surface supercell size. This seems to exhibit some of the finite–size effects that were found in the QMC work. However using two-point extrapolation with Equation 3.36 we can get convergence of better than 1 meV leading us to quote -10.26 as our best estimate for $E_{\mathrm{ads}}^{\mathrm{HF}}$.

As stated near the start of this section the fact that this is such a small contribution to the adsorption energy suggests that even if there were a large error introduced by using the PAW formalism to calculate $E_{\mathrm{ads}}^{\mathrm{HF}}$ this would still result in a small error in the final adsorption energy.

5.3.6 Results: Adsorption energy

Our best estimate for the full adsorption energy including our incremental calculations, all of the subsequent corrections and our Hartree–Fock estimate is now presented in Table 5.10.

Focusing first on the level of convergence we can see that the basis set correction is less than 10 meV suggesting we must be very close to the basis–set limit. Also our corrections levels of theory beyond CCSD(T) show only about a 2.5 meV change again indicating a high level of convergence, though suggesting for more complex systems CCSDT(Q) and higher levels of theory may be needed. The large size of the δCCSD(T) contribution shows that the perturbative triples are definitely needed in these types of calculations in order to obtain chemical accuracy.

Comparing to our DMC number of 250(7) meV we can see that the agreement between quantum chemistry and DMC is once again astounding. The statistical error present in the DMC result means we

Table 5.8: Increments calculated at MP2/VDZ level for various fragments. All have energies greater than 0.27 meV. Dashes indicate increments that are physically absent in a given fragment. Numbers in italics are increments that are greater than 0.27 meV in other fragments but less than 0.27 meV in the given fragment. All energies are in meV and increments that are identical by symmetry have been omitted. The relevant full MP2 results from Table 5.6 have also been repeated for convenience.

Increment	6×6×2	5×5×2	4×4×2
H1	2.3402	2.5034	2.9660
H103	-0.3810	*0.0095*	-
H1–H2	2.0681	1.9864	2.2041
H1–H3	0.4626	0.4898	0.5170
H1–H4	1.7960	1.8504	1.9864
H3–H4	2.3674	2.4762	2.4490
H3–H9	-0.4082	-0.4082	-
H1–H101	1.0340	0.9796	1.0340
H1–H105	0.9252	0.8436	1.1157
H1–H106	0.6259	0.5442	0.6259
H1–H114	0.6531	0.6259	0.7347
H3–H102	-0.6259	-0.5442	-0.2721
H3–H103	-0.5714	-0.4082	-
H4–H12	-0.4082	-0.4082	-
H_2O	50.2866	53.2255	57.0895
H1–H_2O	-46.5315	-47.2934	-49.7696
H3–H_2O	-22.0684	-22.3950	-22.4222
H5–H_2O	-1.7415	-1.7143	-2.5307
H6–H_2O	-1.7960	-1.8232	-1.7687
H9–H_2O	-1.4422	-1.2517	-
H10–H_2O	-1.0340	-1.0068	-
H14–H_2O	-0.4354	-	-
H101–H_2O	-4.1906	-4.5715	-4.2178
H102–H_2O	-1.7143	-1.5510	-2.2858
H103–H_2O	-0.5170	-0.5442	-
H104–H_2O	-1.1973	-1.1701	-1.3061
H106–H_2O	-1.7415	-1.9048	-1.7415
H107–H_2O	-1.0068	-0.8980	-1.3606
Total	−96.63	−93.94	−88.80
Full calculation	−96.11	−93.06	−90.37

Table 5.9: Hartree-Fock adsorption energy calculated using VASP for a variety of two–layer surface super-cells. Two–point extrapolations using Eq. 3.36 are also given. All energies in meV.

Surface–cell	E_{ads}	Extrapolated
4×4	−7.35	
		−9.50
5×5	−8.38	
		−10.26
6×6	−9.17	
		−10.26
7×7	−9.58	

Table 5.10: Final adsorption energy. The CCSD and δCCSD(T) contributions were calculated using the incremental scheme. The δCCSDT and higher calculations were performed on embedded clusters using the MRCC module with Molpro. The basis and core correction calculations were performed on embedded clusters, with the basis correction making use of density fitting. The Hartree–Fock contribution was calculated using VASP.

Contribution	E_{ads} / meV	Notes
CCSD	−190.67	Incremental CCSD/V[DT]Z 6×6×2 fragment
δCCSD(T)[a]	−34.04	Incremental CCSD(T)/V[DT]Z 4×4×2 fragment
δCCSDT[b]	−0.52	CCSDT/VDZ 3×3×2 fragment
δCCSDT(Q)[c]	−2.07	CCSDT(Q)/VDZ 2×2×2 fragment
Basis correction[d]	−8.41	MP2/V[Q5]Z 4×4×2 fragment
Core correction[e]	−14.56	MP2/CVTZ 3×3×2 fragment
Hartree–Fock	−10.26	VASP calculation
E_{ads}	−245.97	(frozen–core)
E_{ads}	−260.53	

[a] Calculated as CCSD(T)/V[DT]Z−CCSD/V[DT]Z.

[b] Calculated as CCSDT/VDZ−CCSD(T)/VDZ

[c] Calculated as CCSDT(Q)/VDZ−CCSDT/VDZ

[d] Calculated as MP2/V[Q5]Z−MP2/V[DT]Z

[e] Calculated as ae-MP2/CVTZ−MP2/VTZ

are unable to discriminate between the frozen–core and full core calculations but all three agree to each other to well within chemical accuracy.

However in order to fully be convinced of the agreement the full adsorption energy curve is needed.

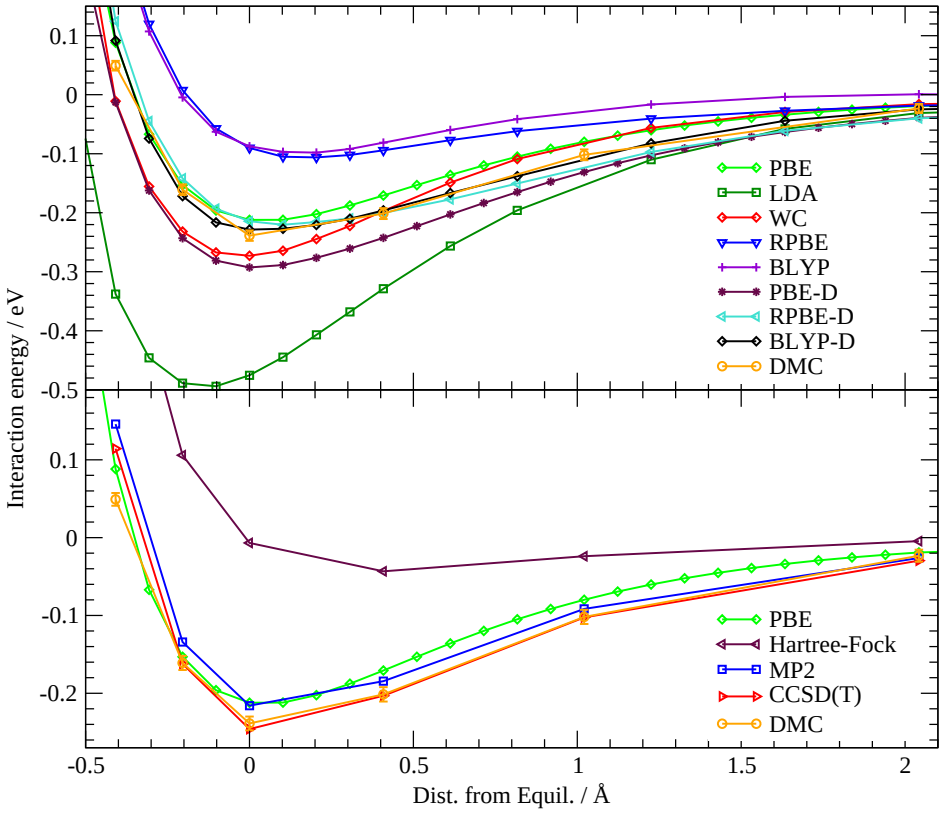
5.3.7 Adsorption curve

As explained at the start of this section the adsorption curve shown in Figure 5.7 was calculated in a less costly way. Instead of CCSD, MP2/V[D,T]Z on a 6×6×2 fragment was used as the reference energy

at each point. As before a δCCSD(T)/V[D,T]Z correction[8] was applied using a 4×4×2 fragment along with the usual basis correction. The reason for this is the extreme cost of the CCSD calculations. With the VTZ basis these calculations took around a month to complete on a large shared memory machine. In contrast the MP2 calculations took less than a day.

The expensive small higher order corrections (CCSDT etc.) and core–correction were omitted as they were also quite costly and only produced energy differences of the same magnitude as the statistical error on the DMC data.

Figure 5.7: The top graph shows the various DFT adsorption curves, based on the PBE equilibrium geometry with the water molecule displaced by a given amount in the z-direction, with the DMC data for a reference. The second graph shows the DMC and quantum chemistry curves.



5.3.8 Summary

The agreement between quantum chemistry and DMC shown in Figure 5.7 again shows that chemical accuracy has indeed been attained. More than this we can see that the quantum chemistry estimates and the DMC estimates agree to within the DMC error bars ($\sim$ 7 meV) for all but one of the points. This is almost an order of magnitude better agreement than chemical accuracy. Our extensive testing of the convergence for both the QMC and quantum chemistry calculations also allow us to say that this

[8] This time calculated as E(CCSD(T))$-E$(MP2).

is no mere lucky cancellation of errors. As we have seen the error estimate on the quantum chemistry adsorption energy would place it at no more than 10 meV in the worst case scenario.

The discrepancy at short distance is similar, but of an opposite sign, to that found in the DMC–quantum chemistry study of the water–benzene interaction [124]. In our case it is possibly down to the use of pseudopotentials in the DMC work. This discrepancy may be a symptom of the pseudopotential description breaking down at these short distances. However more work, going beyond the pseudopotential approximation, would be needed to produce a definitive answer in this case.

The graphs also highlights just why the quantum chemistry and QMC calculations were needed, with the large discrepancies between the different DFT functionals. In the end it would appear that WC, PBE and dispersion corrected PBE, RPBE and BLYP can just be said to be correct to within chemical accuracy. This must however be set within the context that the dispersion correction used is semi–empirical so these dispersion corrected calculations are not truly *ab initio* estimates. Also in the case of PBE the dispersion correction actually worsens the agreement with our highly accurate benchmark curves.

Chapter 6

Summary and Discussion

The remarkable diversity of electronic structure methods currently in use and development is a testament to the large amount of human effort that has gone into computational modelling over the last 60 years. It is also indicative of the fact that no one method appears to be ideal for every application. The ubiquity of DFT within the field would suggest that it at least comes close to being regarded as the best all–round theory, a suitable trade off between computational effort and accuracy.

The broad aim of this thesis has been to investigate two fairly simple situations where DFT produces at best ambiguous results due to a large dependence on minor variations within the exchange–correlation functionals. Moreover we have also shown that quantum Monte Carlo methods such as diffusion Monte Carlo can be used in these circumstances to produce otherwise unobtainable benchmark values against which the DFT work can be measured and information on the accuracy of the functionals obtained. Central to this has been the principle that if it is to be accurate it must be able to agree with wavefunction based quantum chemistry methods whose own accuracy can be ensured through rigorous hierarchies of levels of theory and basis–sets.

6.1 Surface energies and bulk crystal

In Chapter 4 more evidence was presented for the large exchange–correlation functional dependence of both the bulk properties and surface energies of simple ionic crystals. This is in line with the scattered evidence in the literature that prompted this investigation. Across the functionals tested variations of up to 50% in the surface formation energy have been observed for individual systems casting serious doubt on the predictive power of DFT not only for surface formation energies but on the numerous properties that depend upon them.

The detailed studies using DMC on LiH and LiF that have been presented have shown an exceptional level of agreement with experiment for the bulk properties, with lattice parameters agreeing to better than 0.2 % (within the statistical accuracy of the DMC calculations) and the cohesive energy to within chemical accuracy. This adds to the growing body of evidence of the quality of DMC for bulk crystals. Also in the case of LiF the MPC and KZK finite–size correction schemes were directly compared confirming that they both extrapolated to the same limiting value but indicating that in the case of the bulk the MPC correction significantly outperformed the KZK correction at a given supercell size.

Hartree–Fock calculations on both LiH and LiF also led to accurate quantum chemistry estimates being constructed from existing post Hartree–Fock data. The agreement between these numbers and DMC was no less impressive again with better than chemical accuracy achieved. As a side effect a tentative scheme for the construction of Gaussian basis–sets for use in accurate bulk calculations has been presented.

The quantum chemistry also allowed us to see in the case of LiH where the pseudopotential approach of DMC broke down, something which we were able to verify with our novel all–electron DMC calculations on LiH. Together the DMC and quantum chemistry data present benchmark values for σ which when compared to the DFT data provide more confirmation that the LDA seems to predict surface formation energies to within 10 %, with the PBE GGA underestimating σ by $\sim$ 30%. The calculations on LiF in paticular have provided new insight into the experimental methods that are often used to determine σ confirming the view of Burns and Webb that this method can be in error by at least 30 %.

6.2 Adsorption energy

Chapter 5 considered a specific system where accuracy of surface calculations is paramount, namely that of surface adsorption. Water on LiH (001) was taken as a relatively simple test system. A series of adsorption energy curves calculated using different DFT functionals were presented. These showed the same trends as the σ calculations, with LDA predicting the largest adsorption energy of 0.5 eV and RPBE the smallest with 0.01 eV. The dispersion correction of Grimme was also used suggesting an adsorption energy of 0.3–0.15 eV.

Diffusion Monte Carlo calculations showed a significant and unexpected many–body finite–size error. The error was substantially larger than those encountered in both DFT and HF calculations on the system. This is an important result as often convergence of DFT calculations with respect to finite–size can be taken as inductive evidence for the same sized supercell to be converged in DMC for adsorption problems. The DMC adsorption curve and adsorption energy of 0.250(7) eV agreed very well with the dispersion corrected BLYP calculations, and also agreed with the Wu–Cohen and PBE functionals to within chemical accuracy.

Accurate quantum chemistry calculations of the adsorption energy at the CCSD(T) level were also presented. These made use of the so–called incremental scheme of Stoll and Paulus and convergence tests indicated the final adsorption energy was converged to 10 meV, better than chemical accuracy with respect to fragment size, basis–set and level of theory. The predicted adsorption energy of 0.261 eV provided a level of agreement with the DMC calculations that is unprecedented for extended systems. This was also true of the adsorption curve which, apart form a discrepancy at very close distances, agreed with the DMC curve to within 10 meV for every point.

6.3 Further work

This study has shown that for fixed geometries diffusion Monte Carlo can be used to give surface energies and adsorption energies to better than chemical accuracy. These can be verified by a collaborative approach including wavefunction based quantum chemistry techniques. An obvious extension of this work

would be to consider more complex systems such as water on MgO which is a currently outstanding problem [142].

Perhaps one of the major drawbacks of the QMC and quantum chemistry methods used in this work is that neither, at the present time, are able to efficiently relax the geometries of these extended systems. In QMC this is part of the long standing problem of calculating accurate forces. Much progress has been made on this recently [161, 162, 163, 164] and other promising methods have been suggested circumventing the need for forces [165]. While it is hoped that these problems will eventually be solved the requirement for geometries assumed from other methods remains a reality for the moment.

With this in mind the role of studies such as this one in validating specific DFT functionals for specific properties must be considered. While this approach can be less than satisfying for an intellectual point of view, it does have the benefit of pragmatism. It is easy to see a situation where accurate bench–marks from more expensive methods could be used to indicate a specific functional that could be used in the majority of cases, with the calculations passing back to the expensive methods when very high accuracy is required, or in the borderline cases where the DFT predictions become ambiguous. In the cases of very high accuracy or where the supporting data is lacking it should perhaps become the norm to require the agreement of two such high accuracy methods, as outlined in this work.

Appendix A

Optimized LiF basis

There are well known problems that arise when trying to use diffuse atomic basis that have been optimized for a molecular setting in the close packed environs of a crystal. These include problems of the basis becoming linearly-dependant and making the SCF cycles numerically unstable. The usual approach has been to use severely truncated basis–sets, but this approach does not allow you to approach the basis–set limit with any confidence.

In this paper we have taken the approach of using a fixed basis–set that has previously been optimized for crystal calculations, on one atomic species, in our case lithium. We then attempt to reach the basis–set limit by using an established hierarchy of basis–sets on the other atomic species, fluorine. This approach relies on the diffuse tails of the fluorine centred basis also being able to describe the areas of the lithium sites that are poorly describe by its own basis functions. We then apply a well-established fitting formula, Equation 2.14 to obtain the basis–set limit.

In practice we were able to use Jensen's polarization consistent basis up to pc-2 almost without modification. The only change that was made was to use a more conservative set of basis contractions based on Hartree–Fock calculations of the F^- ion. These basis–sets are listed in Tables A.1, A.2 and A.3.

Table A.1: Details of the modified pc-0 basis adopted for flourine.

l	Exponent	Coefficient
s	1035.100000	0.019232
	156.380000	0.132989
s	35.196000	1.000000
s	9.392100	1.000000
s	0.753420	1.000000
p	10.599000	1.000000
p	2.173300	1.000000
p	0.448710	1.000000

The basis adopted for the lithium sites was the one used by Paier *et al.* in their study of crystalline

Table A.2: Details of the modified pc-1 basis adopted for flourine.

l	Exponent	Coefficient
s	2954.600000	0.005296
	444.620000	0.039541
	101.050000	0.176746
s	28.068000	1.000000
s	8.583700	1.000000
s	1.398600	1.000000
s	0.399690	1.000000
p	21.883000	0.047277
	4.776900	0.243154
p	1.282300	1.000000
p	0.311270	1.000000
d	1.100000	1.000000

LiH. We found it necessary to remove the most diffuse s and p functions in order to maintain the numerical stability of the SCF cycles. We would expect this reduction in the accuracy of the basis to be corrected for by the overlapping basis functions from the flourine sites. Table A.4 details our modified basis–set.

It must also be pointed out that this lithium basis proved spectacularly unsuited to calculations of just the lithium ion. Even when embedding the atom in the extended crystal basis this was the case. Much better performance was found when calculating the Li^+ ion, perhaps not surprising given the basis has been optimized for an ionic crystal. We therefore calculated all of our energies relative to isolated ions as opposed atoms (lattice energies as opposed cohesive energies etc.).

Table A.3: Details of the modified pc-2 basis adopted for flourine.

l	Exponent	Coefficient
s	19127.000000	0.000519
	2868.900000	0.004016
	653.040000	0.020574
	184.850000	0.080165
s	59.949000	1.000000
s	21.190000	1.000000
s	7.784600	1.000000
s	2.230500	1.000000
s	0.854090	1.000000
s	0.283360	1.000000
p	77.639000	1.000000
p	17.965000	1.000000
p	5.400500	1.000000
p	1.842300	1.000000
p	0.615320	1.000000
p	0.184890	1.000000
d	2.800000	1.000000
d	0.800000	1.000000
f	1.200000	1.000000

Table A.4: Details of the modified basis from Paier *et al.* adopted for lithium.

l	Exponent	Coefficient
s	70681.000000	0.00000544
	13594.000000	0.00003328
	3100.400000	0.00019175
s	1336.034100	1.00000000
s	444.299820	1.00000000
s	147.797020	1.00000000
s	49.209451	1.00000000
s	16.428957	1.00000000
s	5.529399	1.00000000
s	1.905282	1.00000000
s	0.700259	1.00000000
s	0.299587	1.00000000
p	28.500000	0.00036754
	6.640000	0.00322359
p	1.570911	1.00000000
p	0.748759	1.00000000
p	0.386141	1.00000000
d	0.779208	1.00000000
d	0.407899	1.00000000
f	0.737063	1.00000000

Appendix B

Cohesive energy of LiF

Figure B.1 shows the thermochemical cycle used to calculate the cohesive energy of crystalline LiF. Table B.1 shows the values of these contributions, sourced from Reference [166]. The E_{coh} estimate produced is $-849.1(13)$ kJ mol^{-1} or $-8.800(14)$ eV.

Figure B.1: Thermochemical cycle for calculating the cohesive energy of LiF

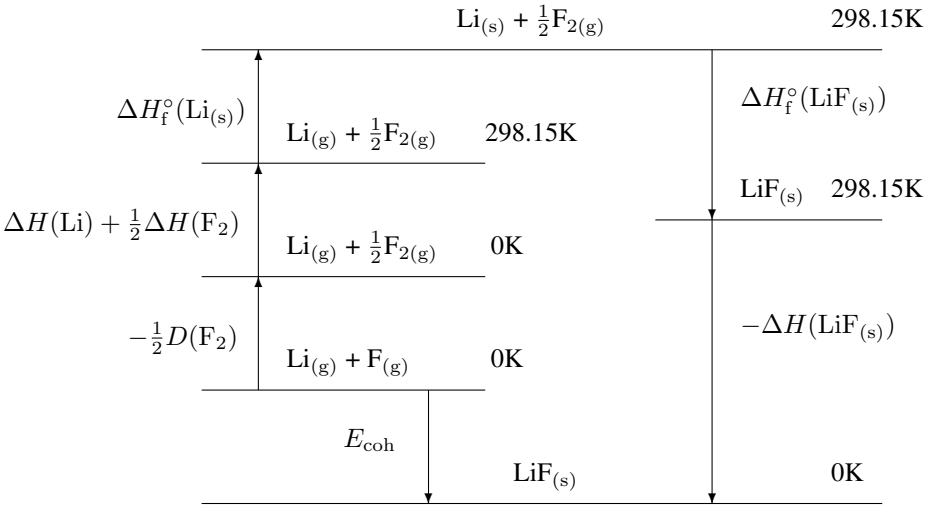


Table B.1: Contributions to the thermochemical cycle. All values in kJ mol^{-1}.

Contribution	Energy / kJ mol^{-1}
$-\frac{1}{2}D(\mathrm{F_2})$	$-77.28(30)$
$\Delta H(\mathrm{Li})$	-6.473
$\frac{1}{2}\Delta H(\mathrm{F_2})$	4.4125
$\Delta H_\mathrm{f}^\circ(\mathrm{LiF_{(s)}})$	$-159.3(10)$
$\Delta H_\mathrm{f}^\circ(\mathrm{LiF_{(s)}})$	$-616.931(800)$
$-\Delta H(\mathrm{LiF_{(s)}})$	6.473
E_coh	$-849.1(13)$

Bibliography

[1] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, **136**: (1964) B864.

[2] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, **140**: (1965) A1133.

[3] M. J. Gillan. The virtual matter laboratory. *Contemp. Phys.*, **38**: (1997) 115.

[4] E. R. M. Davidson, A. Alavi, and A. Michaelides. Dynamics of quantum tunneling: Effects on the rate and transition path of OH on Cu(110). *Physical Review B*, **81**: (2010) 153410.

[5] X. 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**: (2010) 066102.

[6] B. Walker and A. Michaelides. Direct assessment of quantum nuclear effects on hydrogen bond strength by constrained-centroid ab initio path integral molecular dynamics. *J. Chem. Phys.*, **133**: (2010) 174306.

[7] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, **21**: (1982) 467.

[8] S. Lloyd. Universal Quantum Simulators. *Science*, **273**: (1996) 1073.

[9] D. Alfè, M. J. Gillan, and G. D. Price. The melting curve of iron at the pressures of the Earth's core from ab initio calculations. *Nature*, **401**: (1999) 462.

[10] I. Lin, A. P. Seitsonen, M. D. Coutinho-Neto, I. Tavernelli, and U. Rothlisberger. Importance of van der Waals Interactions in Liquid Water. *J. Phys. Chem. B*, **113**: (2009) 1127.

[11] P. J. Feibelman, B. Hammer, J. K. Nørskov, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, and J. Dumesic. The CO/Pt(111) Puzzle. *J. Phys. Chem. B*, **105**: (2001) 4018.

[12] G. W. Crabtree, M. S. Dresselhaus, and M. V. Buchanan. The Hydrogen Economy. *Phys. Today*, **57**: (2004) 39.

[13] A. Züttel, A. Borgschulte, and L. Schlapbach. *Hydrogen as a future energy carrier* (Wiley-VCH, Weinheim, 2008).

[14] M. J. Pollitt. *Protein Coating of Nanoparticles*. Ph.D. thesis, University of London, London [2007].

[15] T. Miyazaki, D. R. Bowler, M. J. Gillan, and T. Ohno. The Energetics of Hut-Cluster Self-Assembly in Ge/Si(001) from Linear-Scaling DFT Calculations. *J. Phys. Soc. Jpn.*, **77**: (2008) 123706.

[16] J. J. Gilman. Direct Measurements of the Surface Energies of Crystals. *J. Appl. Phys.*, **31**: (1960) 2208.

[17] P. L. Gutshall and G. E. Gross. Cleavage Surface Energy of NaCl and MgO in Vacuum. *J. Appl. Phys.*, **36**: (1965) 2459.

[18] P. Kraatz and T. Zoltai. Cleavage surface energy of the (111) plane of strontium fluoride. *J. Appl. Phys.*, **45**: (1974) 4741.

[19] C. Bombis, A. Emundts, M. Nowicki, and H. P. Bonzel. Absolute surface free energies of Pb. *Surf. Sci.*, **511**: (2002) 83.

[20] H. Shi, R. I. Eglitis, and G. Borstel. *Ab initio* calculations of the CaF_2 electronic structure and F centers. *Phys. Rev. B*, **72**: (2005) 045109.

[21] J. G. Swadener, M. I. Baskes, and M. Nastasi. Stress-induced platelet formation in silicon: A molecular dynamics study. *Phys. Rev. B*, **72**: (2005) 201202.

[22] Y. V. Bhargava, Q. A. S. Nguyen, and T. M. Devine. Initiation of Organized Nanopore/Nanotube Arrays in Titanium Oxide. *J. Electrochemical Soc.*, **156**: (2009) E62.

[23] S. J. Burns and W. W. Webb. Fracture Surface Energies and Dislocation Processes during Dynamical Cleavage of LiF. I. Theory. *J. Appl. Phys.*, **41**: (1970) 2078.

[24] S. J. Burns and W. W. Webb. Fracture Surface Energies and Dislocation Processes during Dynamical Cleavage of LiF. II. Experiments. *J. Appl. Phys.*, **41**: (1970) 2086.

[25] V. Fiorentini and M. Methfessel. Extracting convergent surface energies from slab calculations. *J. Phys.: Condens. Matt.*, **8**: (1996) 6525.

[26] M. P. Tosi. Cohesion of Ionic Solids in the Born Model. In F. Seitz and D. Turnbull (editors), *Advances in Research and Applications*, volume 16 of *Solid State Physics*, 1 – 120 (Academic Press, 1964).

[27] P. H. Acioli and D. M. Ceperley. Diffusion Monte Carlo study of jellium surfaces: Electronic densities and pair correlation functions. *Phys. Rev. B*, **54**: (1996) 17199.

[28] Z. Yan, J. P. Perdew, S. Kurth, C. Fiolhais, and L. Almeida. Density–functional versus wave–function methods: Toward a benchmark for the jellium surface energy. *Phys. Rev. B*, **61**: (2000) 2595.

[29] L. M. Almeida, J. P. Perdew, and C. Fiolhais. Surface and curvature energies from jellium spheres: Density functional hierarchy and quantum Monte Carlo. *Phys. Rev. B*, **66**: (2002) 075115.

[30] B. Wood, N. D. M. Hine, W. M. C. Foulkes, and P. Garca-Gonzlez. Quantum Monte Carlo calculations of the surface energy of an electron gas. *Phys. Rev. B*, **76**: (2007) 035403.

[31] M. Methfessel, D. Hennig, and M. Scheffler. Trends of the surface relaxations, surface energies, and work functions of the 4d transition metals. *Phys. Rev. B*, **46**: (1992) 4816.

[32] J. L. D. Silva, C. Stampfl, and M. Scheffler. Converged properties of clean metal surfaces by all-electron first-principles calculations. *Surf. Sci.*, **600**: (2006) 703.

[33] D. Yu and M. Scheffler. First-principles study of low-index surfaces of lead. *Phys. Rev. B*, **70**: (2004) 155417.

[34] J. Goniakowski, J. M. Holender, L. N. Kantorovich, M. J. Gillan, and J. A. White. Influence of gradient corrections on the bulk and surface properties of TiO_2 and SnO_2. *Phys. Rev. B*, **53**: (1996) 957.

[35] D. Alfè and M. J. Gillan. The energetics of oxide surfaces by quantum Monte Carlo. *J. Phys.: Condens. Matt.*, **18**: (2006) L435.

[36] S. J. Nolan, M. J. Gillan, D. Alfè, N. L. Allan, and F. R. Manby. Calculation of properties of crystalline lithium hydride using correlated wave function theory. *Phys. Rev. B*, **80**: (2009) 165109.

[37] B. Hammer, L. B. Hansen, and J. K. Nørskov. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B*, **59**: (1999) 7413.

[38] A. Stroppa and G. Kresse. The shortcomings of semi–local and hybrid functionals: what we can learn from surface science studies. *New J. Phys.*, **10**: (2008) 063020.

[39] L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse. Accurate surface and adsorption energies from many-body perturbation theory. *Nat. Mater.*, **9**: (2010) 741.

[40] B. Paulus. The method of increments–a wavefunction-based ab initio correlation method for solids. *Phys. Rep.*, **428**: (2006) 1.

[41] C. Müller, B. Herschend, K. Hermansson, and B. Paulus. Application of the method of increments to the adsorption of CO on the CeO_2(110) surface. *J. Chem. Phys.*, **128**: (2008) 214701.

[42] C. Müller, K. Hermansson, and B. Paulus. Electron correlation contribution to the N_2O/ceria(111) interaction. *Chem. Phys.*, **362**: (2009) 91.

[43] B. Paulus and K. Rosciszewski. Application of the method of increments to the adsorption of H_2S on graphene. *Int. J. Quantum Chem.*, **109**: (2009) 3055.

[44] N. Lopez, F. Illas, N. Rosch, and G. Pacchioni. Adhesion energy of Cu atoms on the MgO(001) surface. *J. Chem. Phys.*, **110**: (1999) 4873.

[45] T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter. Periodic density functional embedding theory for complete active space self-consistent field and configuration interaction calculations: Ground and excited states. *J. Chem. Phys.*, **116**: (2002) 42.

[46] C. D. Valentin, G. Pacchioni, T. Bredow, D. Dominguez–Ariza, and F. Illas. Bonding of NO to NiO(100) and $Ni_xMg_{1x}O$(100) surfaces: A challenge for theory. *J. Chem. Phys.*, **117**: (2002) 2299.

[47] B. Li, A. Michaelides, and M. Scheffler. How strong is the bond between water and salt? *Surf. Sci.*, **602**: (2008) L135.

[48] C. Filippi, S. B. Healy, P. Kratzer, E. Pehlke, and M. Scheffler. Quantum Monte Carlo Calculations of H_2 Dissociation on Si(001). *Phys. Rev. Lett.*, **89**: (2002) 166102.

[49] M. Pozzo and D. Alfè. Hydrogen dissociation on Mg(0001) studied via quantum Monte Carlo calculations. *Phys. Rev. B*, **78**: (2008) 245313.

[50] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, Mineola, N.Y, 1996), 1st edition.

[51] J. C. Slater. Atomic Shielding Constants. *Phys. Rev.*, **36**: (1930) 57.

[52] F. Jensen. Polarization consistent basis sets: Principles. *J. Chem. Phys.*, **115**: (2001) 9113.

[53] F. Jensen. Estimating the Hartree–Fock limit from finite basis set calculations. *Theo. Chem. Acc.*, **113**: (2005) 267.

[54] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[55] D. M. Ceperley and B. J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.*, **45**: (1980) 566.

[56] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, **98**: (1993) 5648.

[57] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.*, **98**: (1994) 11623.

[58] H. J. Monkhorst and J. D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, **13**: (1976) 5188.

[59] R. Dovesi, V. R. Saunders, R. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. B. an d P D'Arco, and M. Llunell. *CRYSTAL09 User's Manual* (University of Torino, Torino, 2009).

[60] L. Kleinman and D. M. Bylander. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.*, **48**: (1982) 1425.

[61] D. R. Hamann, M. Schlüter, and C. Chiang. Norm–Conserving Pseudopotentials. *Phys. Rev. Lett.*, **43**: (1979) 1494.

[62] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, **41**: (1990) 7892.

[63] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, **50**: (1994) 17953.

[64] G. H. Booth, A. J. W. Thom, and A. Alavi. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.*, **131**: (2009) 054106.

[65] G. H. Booth and A. Alavi. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials. *J. Chem. Phys.*, **132**: (2010) 174104.

[66] D. Cleland, G. H. Booth, and A. Alavi. Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.*, **132**: (2010) 041103.

[67] M. Kállay and P. R. Surján. Higher excitations in coupled-cluster theory. *J. Chem. Phys.*, **115**: (2001) 2945.

[68] M. Leininger, W. Allen, H. S. III, and C. Sherrill. Is Møller–Plesset perturbation theory a convergent ab initio method? *J. Chem. Phys.*, **112**: (2000) 9213.

[69] D. Case and F. R. Manby. The Ornstein–Zernike equation in molecular electronic structure theory. *Mol. Phys.*, **108**: (2010) 307.

[70] H. P. Kelly. Correlation Effects in Atoms. *Phys. Rev.*, **131**: (1963) 684.

[71] M. Feyereisen, G. Fitzgerald, and A. Komornicki. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.*, **208**: (1993) 359.

[72] F. Weigend. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.*, **4**: (2002) 4285.

[73] F. R. Manby, D. Alfè, and M. J. Gillan. Extension of molecular electronic structure methods to the solid state: computation of the cohesive energy of lithium hydride. *Phys. Chem. Chem. Phys.*, **8**: (2006) 5178.

[74] W. M. C. Foulkes, L. Mitáš, R. J. Needs, and G. Rajagopal. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.*, **73**: (2001) 33.

[75] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos. Continuum variational and diffusion quantum Monte Carlo calculations. *J. Phys.: Condens. Matt.*, **22**: (2010) 023201.

[76] J. Kolorenc and L. Mitáš. Applications of quantum Monte Carlo methods in condensed systems. *arXiv:1010.4992*.

[77] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.*, **21**: (1953) 1087.

[78] N. D. Drummond, M. D. Towler, and R. J. Needs. Jastrow correlation factor for atoms, molecules, and solids. *Phys. Rev. B*, **70**: (2004) 235119.

[79] P. V. Tóth. Boundary conditions for many–electron systems. *arXiv:1010.2700*.

[80] A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs. Scheme for adding electronnucleus cusps to Gaussian orbitals. *J. Chem. Phys.*, **122**: (2005) 224322.

[81] S. J. Binnie, S. J. Nolan, N. D. Drummond, D. Alfè, N. L. Allan, F. R. Manby, and M. J. Gillan. Bulk and surface energetics of crystalline lithium hydride: Benchmarks from quantum Monte Carlo and quantum chemistry. *Phys. Rev. B*, **82**: (2010) 165431.

[82] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig. Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions. *Phys. Rev. Lett.*, **98**: (2007) 110201.

[83] J. Toulouse and C. J. Umrigar. Optimization of quantum Monte Carlo wave functions by energy minimization. *J. Chem. Phys.*, **126**: (2007) 084102.

[84] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 1996).

[85] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester. Fixed-node quantum Monte Carlo for molecules. *J. Chem. Phys.*, **77**: (1982) 5593.

[86] F. A. Reboredo, R. Q. Hood, and P. R. C. Kent. Self-healing diffusion quantum Monte Carlo algorithms: Direct reduction of the fermion sign error in electronic structure calculations. *Phys. Rev. B*, **79**: (2009) 195117.

[87] M. Troyer and U. Wiese. Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations. *Phys. Rev. Lett.*, **94**: (2005) 170201.

[88] L. Mitáš, E. L. Shirley, and D. M. Ceperley. Nonlocal pseudopotentials and diffusion Monte Carlo. *J. Chem. Phys.*, **95**: (1991) 3467.

[89] M. Casula. Beyond the locality approximation in the standard diffusion Monte Carlo method. *Phys. Rev. B*, **74**: (2006) 161102.

[90] M. Casula, S. Moroni, S. Sorella, and C. Filippi. Size–consistent variational approaches to non-local pseudopotentials: Standard and lattice regularized diffusion Monte Carlo methods revisited. *J. Chem. Phys.*, **132**: (2010) 154113.

[91] G. E. Engel, Y. Kwon, and R. M. Martin. Quasiparticle bands in a two-dimensional crystal found by GW and quantum Monte Carlo calculations. *Phys. Rev. B*, **51**: (1995) 13538.

[92] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes. Finite-size errors in continuum quantum Monte Carlo calculations. *Phys. Rev. B*, **78**: (2008) 125106.

[93] H. Kwee, S. Zhang, and H. Krakauer. Finite-Size Correction in Many-Body Electronic Structure Calculations. *Phys. Rev. Lett.*, **100**: (2008) 126404.

[94] E. Sola and D. Alfè. Melting of Iron under Earth's Core Conditions from Diffusion Monte Carlo Free Energy Calculations. *Phys. Rev. Lett.*, **103**: (2009) 078501.

[95] L. Spanu, S. Sorella, and G. Galli. Nature and Strength of Interlayer Binding in Graphite. *Phys. Rev. Lett.*, **103**: (2009) 196401.

[96] A. J. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang, and M. Chou. Elimination of Coulomb finite-size effects in quantum many-body simulations. *Phys. Rev. B*, **55**: (1997) R4851.

[97] A. J. Williamson, R. Q. Hood, and J. C. Grossman. Linear-Scaling Quantum Monte Carlo Calculations. *Phys. Rev. Lett.*, **87**: (2001) 246406.

[98] D. Alfè and M. J. Gillan. Efficient localized basis set for quantum Monte Carlo calculations on condensed matter. *Phys. Rev. B*, **70**: (2004) 161101.

[99] E. Hernández, M. J. Gillan, and C. M. Goringe. Basis functions for linear-scaling first-principles calculations. *Phys. Rev. B*, **55**: (1997) 13485.

[100] N. Nemec. Diffusion Monte Carlo: Exponential scaling of computational cost for large systems. *Phys. Rev. B*, **81**: (2010) 035119.

[101] M. J. Gillan, M. D. Towler, and D. Alfè. Petascale computing opens new vistas for quantum Monte Carlo. In Ψ_k *Highlight*, volume 103 (2011).

[102] Z. Wu and R. E. Cohen. More accurate generalized gradient approximation for solids. *Phys. Rev. B*, **73**: (2006) 235116.

[103] C. Cazorla [2008]. Private communication.

[104] C. Cazorla, S. Binnie, D. Alfè, and M. J. Gillan. Effect of the exchange-correlation energy and temperature on the generalized phase diagram of the 4d transition metals. *High Pressure Res.*, **28**: (2008) 449.

[105] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, **47**: (1993) 558.

[106] G. Kresse and J. Hafner. Ab initio molecular-dynamics simulation of the liquid-metalamorphous-semiconductor transition in germanium. *Phys. Rev. B*, **49**: (1994) 14251.

[107] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comp. Mat. Sci.*, **6**: (1996) 15.

[108] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, **54**: (1996) 11169.

[109] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, **59**: (1999) 1758.

[110] G. Kresse, M. Marsman, and J. Furthmüller. *VASP Users Guide* (Computational Physics, Faculty of Physics, Universität Wien, Sensengasse 8, A-1130 Wien, Austria).

[111] F. Birch. Finite Elastic Strain of Cubic Crystals. *Phys. Rev.*, **71**: (1947) 809.

[112] C. R. Fischer, T. A. Dellin, S. W. Harrison, R. D. Hatcher, and W. D. Wilson. Repulsive Interactions for LiH. *Phys. Rev. B*, **1**: (1970) 876.

[113] R. N. Euwema, G. G. Wepfer, G. T. Surratt, and D. L. Wilhite. Hartree-Fock calculations for crystalline Ne and LiF. *Phys. Rev. B*, **9**: (1974) 5249.

[114] R. L. Erikson, L. E. Eary, and C. J. Hostetler. Energetics, structure, and compressibility of NaF determined by the periodic Hartree–Fock method. *J. Chem. Phys.*, **99**: (1993) 336.

[115] C.-S. Zha, H.-k. Mao, and R. J. Hemley. Elasticity of MgO and a primary pressure scale to 55 GPa. *Proc. Natl. Acad. Sci.*, **97**: (2000) 13494.

[116] S. Baroni, P. Giannozzi, and A. Testa. Green's-function approach to linear response in solids. *Phys. Rev. Lett.*, **58**: (1987) 1861.

[117] G. Barrera, D. Colognesi, P. Mitchell, and A. Ramirez-Cuesta. LDA or GGA? A combined experimental inelastic neutron scattering and ab initio lattice dynamics study of alkali metal hydrides. *Chem. Phys.*, **317**: (2005) 119.

[118] M. E. Straumanis and J. S. Shah. Low temperature lattice parameters and expansion coefficients of AI2Au and LiF Gruneisen constants of LiF. *Z. Anorg. Allg. Chem.*, **391**: (1972) 79.

[119] C. V. Briscoe and C. F. Squire. Elastic Constants of LiF from 4.2K to 300K by Ultrasonic Methods. *Phys. Rev.*, **106**: (1957) 1175.

[120] J. R. Trail and R. J. Needs. Smooth relativistic Hartree–Fock pseudopotentials for H to Ba and Lu to Hg. *J. Chem. Phys.*, **122**: (2005) 174109.

[121] J. R. Trail and R. J. Needs. Norm-conserving Hartree–Fock pseudopotentials and their asymptotic behavior. *J. Chem. Phys.*, **122**: (2005) 014112.

[122] 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. Matt.*, **21**: (2009) 395502.

[123] P. R. C. Kent, R. Q. Hood, M. D. Towler, R. J. Needs, and G. Rajagopal. Quantum Monte Carlo calculations of the one-body density matrix and excitation energies of silicon. *Phys. Rev. B*, **57**: (1998) 15293.

[124] J. Ma, D. Alfè, A. Michaelides, and E. Wang. The water-benzene interaction: Insight from electronic structure theories. *J. Chem. Phys.*, **130**: (2009) 154303.

[125] J. P. Vidal and G. Vidal-Valat. Accurate Debye–Waller factors of ^{7}LiH and ^{7}LiD by neutron diffraction at three temperatures. *Acta Crystallogr. B*, **42**: (1986) 131.

[126] K. P. Esler, R. E. Cohen, B. Militzer, J. Kim, R. J. Needs, and M. D. Towler. Fundamental High–Pressure Calibration from All–Electron Quantum Monte Carlo Calculations. *Phys. Rev. Lett.*, **104**: (2010) 185702.

[127] S. M. Rothstein, N. Patil, and J. Vrbik. Time step error in diffusion Monte Carlo simulations: An empirical study. *J. Comp. Chem.*, **8**: (1987) 412.

[128] M. Puchalski and K. Pachucki. Ground-state wave function and energy of the lithium atom. *Phys. Rev. A*, **73**: (2006) 022503.

[129] M. Puchalski, D. Kęziera, and K. Pachucki. Ground state of Li and Be^{+} using explicitly correlated functions. *Phys. Rev. A*, **80**: (2009) 032521.

[130] M. J. Gillan, D. Alfè, S. de Gironcoli, and F. R. Manby. High-precision calculation of Hartree-Fock energy of crystals. *J. Comp. Chem.*, **29**: (2008) 2098.

[131] J. Paier, C. V. Diaconu, G. E. Scuseria, M. Guidon, J. VandeVondele, and J. Hutter. Accurate Hartree-Fock energy of extended systems using large Gaussian basis sets. *Phys. Rev. B*, **80**: (2009) 174114.

[132] B. Civalleri, R. Orlando, C. M. Zicovich-Wilson, C. Roetti, V. R. Saunders, C. Pisani, and R. Dovesi. Comment on "Accurate Hartree-Fock energy of extended systems using large Gaussian basis sets". *Phys. Rev. B*, **81**: (2010) 106101.

[133] J. Paier, R. Hirschl, M. Marsman, and G. Kresse. The Perdew-Burke-Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set. *J. Chem. Phys.*, **122**: (2005) 1.

[134] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. Gerber, and J. Angyn. Screened hybrid density functionals applied to solids. *J. Chem. Phys.*, **124**.

[135] J. Paier, M. Marsman, and G. Kresse. Why does the B3LYP hybrid functional fail for metals? *J. Chem. Phys.*, **127**.

[136] S. J. Nolan, S. J. Binnie, D. Alfè, N. L. Allan, F. R. Manby, and M. J. Gillan. To be published.

[137] T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.*, **90**: (1989) 1007.

[138] F. Jensen. Polarization consistent basis sets. II. Estimating the Kohn–Sham basis set limit. *J. Chem. Phys.*, **116**: (2002) 7372.

[139] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz and others. MOLPRO, version 2008.1, a package of ab initio programs [2008]. See http://www.molpro.net.

[140] S. J. Nolan [2010]. Private communication.

[141] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, **132**: (2010) 154104.

[142] D. Alfè and M. Gillan. Ab initio statistical mechanics of surface adsorption and desorption. I. H2 O on MgO (001) at low coverage. *J. Chem. Phys.*, **127**.

[143] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, **37**: (1988) 785.

[144] Y. Andersson, D. C. Langreth, and B. I. Lundqvist. van der Waals Interactions in Density-Functional Theory. *Phys. Rev. Lett.*, **76**: (1996) 102.

[145] D. C. Langreth, M. Dion, H. Rydberg, E. Schröder, P. Hyldgaard, and B. I. Lundqvist. Van der Waals density functional theory with applications. *Int. J. Quantum Chem.*, **101**: (2005) 599.

[146] P. L. Silvestrelli, K. Benyahia, S. Grubisic, F. Ancilotto, and F. Toigo. Van der Waals interactions at surfaces by density functional theory using Wannier functions. *J. Chem. Phys.*, **130**: (2009) 074702.

[147] J. Klimeš, D. R. Bowler, and A. Michaelides. Chemical accuracy for the van der Waals density functional. *J. Phys.: Condens. Matt.*, **22**: (2010) 022201.

[148] A. Tkatchenko and M. Scheffler. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.*, **102**: (2009) 073005.

[149] D. Alfè. To be published.

[150] M. Marsman, A. Grüneis, J. Paier, and G. Kresse. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.*, **130**: (2009) 184103.

[151] Grüneis, M. Marsman, and G. Kresse. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.*, **133**: (2010) 074107.

[152] R. A. Kendall, T. H. Dunning, and R. J. Harrison. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.*, **96**: (1992) 6796.

[153] D. E. Woon and T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. V. Core–valence basis sets for boron through neon. *J. Chem. Phys.*, **103**: (1995) 4572.

[154] F. Weigend, A. Köhn, and C. Hättig. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.*, **116**: (2002) 3175.

[155] C. Hättig. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.*, **7**: (2005) 59.

[156] M. Kállay. MRCC, a string-based quantum chemical program suite. See http://www.mrcc.hu.

[157] S. F. Boys. Construction of Some Molecular Orbitals to Be Approximately Invariant for Changes from One Molecule to Another. *Rev. Mod. Phys.*, **32**: (1960) 296.

[158] J. M. Foster and S. F. Boys. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.*, **32**: (1960) 300.

[159] C. Müller. *Physisorption of CO and N_2O on Ceria Surfaces*. Ph.D. thesis, Uppsala University [2009].

[160] J. Friedrich, M. Hanrath, and M. Dolg. Energy Screening for the Incremental Scheme: Application to Intermolecular Interactions. *J. Phys. Chem. A*, **111**: (2007) 9830.

[161] A. Badinski and R. J. Needs. Accurate forces in quantum Monte Carlo calculations with nonlocal pseudopotentials. *Phys. Rev. E*, **76**: (2007) 036707. PMID: 17930361.

[162] A. Badinski and R. J. Needs. Total forces in the diffusion Monte Carlo method with nonlocal pseudopotentials. *Phys. Rev. B*, **78**: (2008) 035134.

[163] A. Badinski, P. D. Haynes, and R. J. Needs. Nodal Pulay terms for accurate diffusion quantum Monte Carlo forces. *Phys. Rev. B*, **77**: (2008) 085111.

[164] A. Badinski, P. D. Haynes, J. R. Trail, and R. J. Needs. Methods for calculating forces within quantum Monte Carlo simulations. *J. Phys.: Condens. Matt.*, **22**: (2010) 074202.

[165] L. K. Wagner and J. C. Grossman. Quantum Monte Carlo Calculations for Minimum Energy Structures. *Phys. Rev. Lett.*, **104**: (2010) 210201.

[166] M. W. Chase (editor). *NIST-JANAF Thermochemical Tables*. Number 9 in J. Phys. Chem. reference data (American Chemical Society and American Institute of Physics for the National Institute of Standards and Technology, Washington, D.C., 1998), 4th edition.